Two studies in risk management: portfolio insurance under risk measure constraint and quadratic hedge for jump processes.

Carmine de Franco

UNIVERSITÉ PARIS 7 - Denis Diderot
UFR de MATHÉMATIQUES
École Doctorale de Sciences Mathématiques de Paris Centre

Thèse presentée pour obtenir le titre de

Docteur de l'Université Paris Diderot

Specialité: mathématiques appliquées

presentée par

Carmine DE FRANCO

Deux études en gestion de risque: assurance de portefeuille avec contrainte en risque et couverture quadratique dans les modèles à sauts

Two studies in risk management: portfolio insurance under risk measure constraint and quadratic hedge for jump processes.

Thèse dirigée par

Prof. Peter TANKOV

Soutenue publiquement le 29 Juin 2012 devant le jury composé de:

ACHDOU Yves, Professeur, Université Paris Diderot
BOUCHARD Bruno, Professeur, Université Paris Dauphine
LAMBERTON Damien, Professeur, Université de Marne-la-Vallée
PHAM Huyên, Professeur, Université Paris Diderot
SULEM Agnès, Directeur de recherche, INRIA
WARIN Xavier, Département R&D, EDF France

au vu des rapports de:

KALLSEN Jan, Professeur, Christian-Albrechts Universität zu Kiel
LAMBERTON Damien, Professeur, Université de Marne-la-Vallée

A mia madre Iride,
mio padre Francesco,
a Caterina e Vincenzo

Remerciements

Je tiens à remercier mon directeur de thèse Peter Tankov, qui m'a fortement soutenu pendant ces années de recherche et avec lequel j'ai passé de très beaux moments intenses et agréables. J'ai pu profiter de toute son expérience en mathématiques financières et des nombreuses discussions qu'on a eu le long de mon doctorat. Pour cela je lui serai éternellement reconnaissant.

Je tiens à exprimer ma gratitude à Jan Kallsen et Damien Lamberton pour avoir accepté de rapporter cette thèse, et à tous les membres du jury. Je remercie aussi l'Université Paris VII et le Laboratoire de Probabilités et Modèles Aléatoires qui ont été des environnements de travail très sérieux. En particulier, je veux exprimer toute ma reconnaissance à Prof. Yves Achdou pour l'aide précieuse qu'il m'a fourni, en premier lieu lorsque j'étais un étudiant Erasmus, puis comme professeur à Paris VII et enfin comme son assistant-chargé de travaux dirigés. Ensuite aux Prof. Laure Elie et Prof. Huyên Pham pour leurs cours de finance et optimisation stochastique que j'ai eu la chance de suivre pendant mon Master. Je remercie aussi le Centre de Mathématiques Appliquées de l'École Polytechnique et son groupe de travail auquel j'ai partecipé et à qui ma formation doit beaucoup. À Xavier Warin et Nadia Oudjane du département R&D de EDF va toute ma reconnaissance pour la précieuse collaboration à ce travail.

Si aujourd'hui j'ai pu compléter ma formation en mathématiques, je le dois aussi aux excellents cours que j'ai pu suivre à l' Università degli studi di Roma II -Tor Vergata, et une pensée particulière va aux Prof. Claudio D'Antoni, Roberto Longo, Mauro Nacinovich et Stefano Trapani. J'exprime ma gratitude la plus sincère aux Prof. Paolo Baldi et Prof. Domenico Marinucci pour avoir su motiver ma passion pour le probabilités. Affectueusement ma gratitude va aux Prof. Carmine Armentano et Prof. Francesco Romano de Mormanno.

Nombreux sont ceux qui, pendant ces années, ont toujours manifesté de l'intêret pour mon travail. En particulier, pour les très beaux moments passés ensemble et son support infatigable, toute ma reconnaissance va à Germain Griette. Également à Emmanuel et Jean-Baptiste Meffre, Amber Leesa French, Claire Jarry, Jean-Marc et Stéphane Boulais, Cyprien Larrouy, Marc Viader, Bruno Ludwig, Caroline Remy, Angeliki Kordoni, Sophie Albert, Eric Vignat, Sarah Orford, Sarah Louise Porter, Neil Hancox, John Gareth Wyn Williams, Roberto Arcone et Liam D. Gray. Pour leur amitié depuis toujours je remercie Luigi Bloise, Pierluigi Perrone, Antonio Oliva et Gianni Regina.

Au *Maestro* Jorge De Leon, selon lequel les Mathématiques et l'Opéra sont deux faces d'une même pièce, va toute mon admiration et ma gratitude pour l'expérience qu'il a partagé avec moi.

A mamma, papà, fratima e sorma, ca m'anu duvutu suppurtà pi tuttu stu tempu, tra iute, vinute e spirciate. "...N'dra vita unu adda sapí manià la pinna e la zappa..." dicia quaccunu ca di nummari angunu picchi ni capiscia. P'a zappa c'é sempri tempu e terri ierse certu non ti mancanu..a pinna po jessi ca mò m'ajjiu m'barata a la manià.

Paris , 29 Juin 2012 *Carmine DE FRANCO*

Resumé

Dans cette thèse, je me suis intéressé à deux aspects de la gestion de portefeuille : la maximisation de l'utilité d'un portefeuille financier lorsque on impose une contrainte sur l'exposition au risque, et la couverture quadratique en marché incomplet.

Part I. Dans la première partie, j'étudie un problème d'assurance de portefeuille du point de vue du manager d'un fond d'investissement, qui veut structurer un produit financier pour les investisseurs du fond avec une garantie sur la valeur du portefeuille à la maturité. Si, à la maturité, la valeur du portefeuille est au dessous d'un seuil fixé, l'investisseur sera remboursé à la hauteur de ce seuil par une troisième partie, qui joue le rôle d'assureur du fond (on peut imaginer que le fond appartient à une banque et que donc c'est la banque elle même qui joue le rôle d'assureur). En échange de cette assurance, la troisième partie impose une contrainte sur l'exposition au risque que le manager du fond peut tolérer, mesuré avec une mesure de risque monétaire convexe. Je donne la solution complète de ce problème de maximisation non convexe en marché complet et je prouve que le choix de la mesure de risque est un point crucial pour avoir existence d'un portefeuille optimal. J'applique donc mes résultats lorsque on utilise la mesure de risque entropique (pour laquelle le portefeuille optimal existe toujours), les mesures de risque spectrales (pour lesquelles le portefeuille optimal peut ne pas exister dans certains cas) et la G-divergence.
Mots-clés : Assurance de portefeuille ; maximisation d'utilité ; mesure de risque convexe ; VaR, CVaR et mesure de risque spectrale ; entropie et G-divergence.

Part II. Dans la deuxième partie, je m'intéresse au problème de couverture quadratique en marché incomplet. J'assume que le marché est décrit par un processus Markovien tridimensionnel avec sauts. La première variable d'état décrit l'actif financier, échangeable sur le marché, qui sert comme instrument de couverture ; la deuxième variable d'état modélise un actif financier que intervient dans la dynamique de l'instrument de couverture mais qui n'est pas échangeable sur le marché : il peut donc être vu comme un facteur de volatilité de l'instrument de couverture, ou comme un actif financier que l'on ne peut pas acheter (pour de raisons légales par exemple) ; la troisième et dernière variable d'état représente une source externe de risque qui affecte l'option européenne qu'on veut couvrir, et qui, elle aussi, n'est pas échangeable sur le marché. Pour résoudre le problème j'utilise l'approche de la programmation dynamique, qui me permet d'écrire l'équation de *Hamilton-Jacobi-Bellman* associée au problème de couverture quadratique, qui est non locale en non linéaire. Je prouve que la fonction valeur associée au problème de couverture quadra-

tique peut être caractérisée par un système de trois équations integro-différentielles aux dérivées partielles, dont l'une est semilinéaire et ne dépends pas du choix de l'option à couvrir, et les deux autres sont simplement linéaires , et que ce système a une unique solution régulière dans un espace de Hölder approprié, qui me permet donc de caractériser la stratégie de couverture optimale . Ce résultat est démontré lorsque le processus est non dégénéré (c'est à dire que la composante Brownienne est strictement elliptique) et lorsque le processus est à sauts purs. Je conclus avec une application de mes résultats dans le cadre du marché de l'électricité.
Mots-clés : Couverture quadratique ; modèle à sauts ; programmation dynamique ; équation de Hamilton-Jacobi-Bellman ; équations aux dérivées partielles integro-différentielles ; espaces de Hölder ; processus de Lévy ; marché de l'électricité.

Abstract

In this thesis I'm interested in two aspects of portfolio management: the portfolio insurance under a risk measure constraint and quadratic hedge in incomplete markets.

Part I. I study the problem of portfolio insurance from the point of view of a fund manager, who guarantees to the investor that the portfolio value at maturity will be above a fixed threshold. If, at maturity, the portfolio value is below the guaranteed level, a third party will refund the investor up to the guarantee. In exchange for this protection, for which the investor pays a given fee, the third party imposes a limit on the risk exposure of the fund manager, in the form of a convex monetary risk measure. The fund manager therefore tries to maximize the investor's utility function subject to the risk measure constraint. I give a full solution to this non-convex optimization problem in the complete market setting and show in particular that the choice of the risk measure is crucial for the optimal portfolio to exist. An interesting outcome is that the fund manager's maximization problem may not admit an optimal solution for all convex risk measures, which means that not all convex risk measures may be used to limit fund's exposure in this way. I provide conditions for the existence of the solution. Explicit results are provided for the entropic risk measure (for which the optimal portfolio always exists), for the class of spectral risk measures (for which the optimal portfolio may fail to exist in some cases) and for the G-divergence.
Key words: Portfolio Insurance; utility maximization; convex risk measure; VaR, CVaR and spectral risk measure; entropy and G-divergence.

Part II. In the second part I study the problem of quadratic hedge in incomplete markets. I work with a three-dimensional Markov jump process: the first compo-

nent is the state variable representing the hedging instrument traded in the market, the second component models a risk factor which "perturbs" the dynamics of the hedging instrument and is not traded in the market (as a volatility factor for example in stochastic volatility models); the third one is another source of risk which affects the option's payoff at maturity and is also not traded in the market. The problem can be seen then as a constrained quadratic hedge problem. I privilege here the dynamic programming approach which allows me to obtain the HJB equation related to the value function. This equation is semi linear and non local due the presence of jumps. The main result of this thesis is that this value function, as a function of the initial wealth, is a second order polynomial whose coefficients are characterized as the unique smooth solutions of a triplet of PIDEs, the first of which is semi linear and does not depend on the particular choice of option one wants to hedge, the other two being simply linear. This result is stated when the Markov process is assumed to be a non-generate jump-diffusion and when it is a pure jump process. I finally apply my theoretical results to an example of quadratic hedge in the context of electricity markets.
Key words: Quadratic Hedge; jump processes; dynamic programming; Hamilton-Jacobi-Bellman equations; partial integro-differential equations; Lévy processes; Hölder spaces; electricity markets.

Contents

1 Introduction (French version) 11

2 Introduction 25

I Portfolio Insurance under risk measure constraint 29

3 An overview on Risk measures 31

3.1 Practical needs of the risk measures: empirical evidence 31
3.2 Law invariant risk measures 32
3.2.1 Definition and main properties 32
3.2.2 Representation of convex risk measures 33
3.3 VaR, CVaR and spectral risk measures 35
3.4 G-divergence and entropy 37
3.5 Risk Measures on $\mathbb{L}^p$-spaces 38

4 Portfolio Insurance 39

4.1 The Problem 40
4.2 The decoupling and the solution 42
4.3 The fee 50
4.4 Explicit result: the Entropic risk measure 51
4.4.1 The result 51
4.4.2 Numerical example 53
4.5 Explicit result: the spectral risk measure 56
4.5.1 The result 56
4.5.2 Proof of Lemma 4.14 60
4.5.3 The special case: $CVaR_\beta$ 60
4.6 Explicit result: the G-divergence 61
4.7 Portfolio Insurance: a short review 63

II Quadratic Hedge 65

5 Quadratic hedge: introduction and main properties 67

5.1 Introduction 67
5.2 The Model 71
5.3 The pure investment problem: a priori estimate 74

5.3.1 Proof of Lemma 5.5 82
5.3.2 Proof of Lemma 5.6 83
5.3.3 Proof of Lemma 5.7 86
5.4 The structure of the quadratic hedge value function 89
5.5 The pure investment problem: verification 91
5.6 The quadratic hedge problem: verification 94
5.7 Viscosity solutions 96

6 Smooth solutions: the jump-diffusion case 101
6.1 Hölder regularity of the differential operators 101
6.2 Smoothness and characterization of the function a 106
6.2.1 The approximation sequence 106
6.2.2 Weak convergence and uniqueness 107
6.2.3 Characterization of the function a 111
6.2.4 Comments 115
6.3 Smoothness and characterization of the function v^f 118

7 Smooth solutions: the pure jump case 121
7.1 Infinite activity processes: the model 122
7.2 The pure investment problem: HJB characterization 125
7.2.1 Formal derivation of the PIDE 125
7.3 Operators regularity 126
7.4 Smoothness and characterization of the function a 133
7.4.1 The approximating sequence and its main properties 133
7.4.2 Characterization of the function a 138
7.5 The change of variable 140
7.6 Smoothness and characterization of the function v^f 143
7.7 Finite activity processes: the model 144
7.7.1 The pure investment problem in the finite variation case . . . 147
7.7.2 Proof of Lemma 7.23 151
7.7.3 The quadratic hedge problem in the finite variation case . . . 153

8 Quadratic hedge in electricity markets 155
8.1 Electricity market: a short survey 155
8.2 Future contracts 158
8.3 Numerical approximation of the functions a and b 162
8.3.1 Truncation for the function a 163
8.3.2 Numerical algorithm for the function a 168
8.3.3 Resolution methodology to calculate the function b 169
8.4 A numerical example 171
8.4.1 The martingale case 173

9 The PIDE truncation effect 177
9.1 Motivations 177
9.2 A first approximation 179
9.3 The semi linear PIDE on a bounded domain 182
9.3.1 Existence and Uniqueness in the viscosity sense 182
9.3.2 Existence and Uniqueness in the viscosity sense 184

9.4 Estimate on the truncated PIDE 191

III Appendix **197**

A Doléans-Dade exponential and other estimations **199**

B About a cubic ordinary differential equation **209**

C Hölder spaces **211**
C.1 Introduction 211
C.2 Parabolic Hölder spaces of type 1 212
C.3 Parabolic Hölder spaces of type 2 215

D Itô's formula for pure jump processes **219**

E Density of an α-stable Lévy process **223**

Notations **227**

Index **229**

Bibliography **231**

List of Figures

4.1 The Portfolio insurance structure 40
4.2 Optimal payoff profile for the fund manager 54
4.3 Value function for the fund manager 55
4.4 Payoff profile for the investor 55

8.1 Spikes in the electricity price: European comparison 156
8.2 Spikes in the electricity price in Ontario's market 157
8.3 Realized volatility in the electricity price: European comparison . . . 157
8.4 Future contract. France vs Germany 158
8.5 The function a for the NIG process 173
8.6 Optimal strategy for the pure investment problem 174
8.7 The function b for an at-the-money call option 174
8.8 The function b for an at-the-money put option 175

Chapter 1

Introduction (French version)

Dans cette thèse, composée de deux parties indépendantes, je me suis interessé à la gestion de portefeuille lorsque des contraintes sont imposées sur les stratégies d'investissement possibles.

Dans la première partie, on étudie un problème non standard de maximisation d'utilité de portefeuille. L'idée de fond de ce problème est la suivante: un manager d'un fond d'investissement garantit à ses investisseurs que la valeur du portefeuille à la maturité sera au dessus d'un seuil fixé z. Lorsque ce n'est pas le cas, les investisseurs seront remboursés à la hauteur de ce seuil par une troisième partie, dans le rôle d'assureur du fond. A ce stade, le payoff terminal pour l'investisseur sera $\max(V_T^*, z)$, où V_T^* est la valeur du portefeuille optimal à maturité. Le gestionnaire du fond choisira la stratégie qui maximise l'utilité du fond au-dessus de cette garantie z:

$$V_T^* := \arg \sup_{V_T,\, V_0 = v_0} \mathbb{E}\left[u((V_T - z)^+)\right]$$

On observe que le critère appliqué par le gestionnaire du fond est non-standard, car la fonction d'utilité s'applique seulement au gain réel de l'investisseur. On peut motiver ce choix sur un exemple très simple: prenons la fonction d'utilité exponentielle $u(x) = -\exp(-x)$, la garantie fixée à $z = 1$ et deux portefeuilles dont le profile à maturité est

$$V_T^1 := \begin{cases} 1.50 & \text{on } A \\ 0.80 & \text{on } A^c \end{cases} \qquad V_T^2 := \begin{cases} 1.40 & \text{on } A \\ 0.90 & \text{on } A^c \end{cases}$$

où $\mathbb{P}(A) = 1/2$. Un calcul élémentaire prouve que $\mathbb{E}[u(V_T^1)] < \mathbb{E}[u(V_T^2)]$, mais un investisseur qui est garanti à la hauteur de 1 choisira toujours le portefeuille V_T^1 plutôt que V_T^2, ce qui explique notre choix d'appliquer u seulement au gain effectif. La contrainte imposée par troisième partie est donnée à travers une mesure de risque monétaire convexe: le problème devient donc

$$V_T^* := \arg \sup_{V_T,\, V_0 = v_0} \mathbb{E}\left[u((V_T - z)^+)\right] \text{ tel que } \rho(-(V_T - z)^-) \leq \rho_0$$

où ρ_0 est le seuil de risque que l'assureur tolère.

Dans le Chapitre 3 on commence donc par une introduction rapide sur les mesures de risque, en partant de leur définition axiomatique pour ensuite étudier en détail des classes de mesures de risque très populaires qui seront utilisés par la suite dans nos exemples: la *Value at Risk* (VaR); la *Conditional Value at risk* ($CVaR$) et, de manière plus générale, les mesures de risque spectrales; la mesure de risque entropique et les mesures de risque communément appelées *G-divergence*.

La solution du problème est exposé dans le Chapitre 4, où on fait l'hypothèse que le marché est complet. La grande difficulté dans ce problème est due à sa nature non convexe. On suppose donc, sans perdre de généralité que $z = 0$ et on va introduire deux problèmes , cette fois-ci bien convexes, associés au problème de depart:

$$\begin{array}{ll} U(A,x^+): & \text{maximum } \mathbb{E}\left[u(Z^+)\right] \\ & \text{sous la contrainte } Z \in \mathcal{H}_1\left(A,x^+\right) \text{ou' } \mathcal{H}_1\left(A,x^+\right) := \\ & \left\{Z \in L^1\left(\xi\mathbb{P}\right) \mid \mathbb{E}\left[\xi Z\right] \leq x^+,\, Z = 0 \text{ on } A^c,\, Z \geq 0 \text{ on } A\right\} \\ & \\ \triangle\left(A\right): & \text{minimum } \mathbb{E}\left[\xi Y\right] \\ & \text{sous la contrainte } Y \in \mathcal{H}_2\left(A\right) \text{ou' } \mathcal{H}_2\left(A\right) := \\ & \left\{Y \in L^1\left(\xi\mathbb{P}\right) \mid \rho\left(Y\right) \leq \rho_0,\, Y = 0 \text{ on } A,\, Y \leq 0 \text{ on } A^c\right\} \end{array}$$

ou' ξ est la densité de la probabilité martingale et $A \in \mathcal{F}$ est un ensemble mesurable. Ces deux problèmes sont parametrisés par le couple $(x^+, A) \in \mathbb{R}^+ \times \mathcal{F}$. On va donc chercher la solution optimale de la forme $V_T^* := Z^* + Y^*$, avec Z^* et Y^* solutions optimales des ces deux problèmes, correspondant au couple optimal $((x^*)^+, A^*)$. En effet on prouve que si x_0 est le capital initial à disposition du gestionnaire du fond alors

Si pour tout $A \in \mathcal{F}$, $\triangle\left(A\right) > -\infty$ alors

$$\sup_{\rho(-(X)^-)\leq\rho_0,\,\mathbb{E}[\xi X]\leq x_0} \mathbb{E}[u(X^+)] = \sup_{A\in\mathcal{F}} U\left(A, x^+\left(A\right)\right)$$

ou' $x^+\left(A\right) = x_0 - \triangle\left(A\right)$. Si de plus $\sup_x u(x) = +\infty$ et $\inf_{A\in\mathcal{F}} \triangle\left(A\right) > -\infty$ alors

$$\sup_{\rho(-(X)^-)\leq\rho_0,\,\mathbb{E}[\xi X]\leq x_0} \mathbb{E}[u(X^+)] < +\infty$$

Ce résultat nous donne un algorithme pour résoudre le problème initial: pour un ensemble $A \in \mathcal{F}$, on calcule d'abord $\triangle\left(A\right)$, ensuite $U(A, x^+\left(A\right))$ et on maximise enfin sur tous les ensembles A. Si A^* est ce supremum, et $Z(A^*), Y(A^*)$ sont les solutions optimales des deux problèmes convexes associés, alors une solution optimale pour le gestionnaire de fond sera $V_T^* = Z(A^*)\mathbb{1}_{A^*} + Y(A^*)\mathbb{1}_{(A^*)^c}$. On remarque que le problème initial n'étant pas convexe, on ne peut pas conclure que cette solution est unique. De plus, si pour un ensemble A donné on peut toujours trouver le $Z(A)$ associé, la solution optimale $Y(A)$ peut ne pas exister. Dans ce cas, on n'a pas de solution optimale pour le problème initial mais on peut quand même parler de solutions ε-optimales.

La condition $\triangle(A) > -\infty$ pour tout A est fondamental pour obtenir une solution optimale finie: si en effet $\triangle(A) = -\infty$ et $\sup_x u(x) = +\infty$ alors on peut trouver une suite de portefeuilles admissible X_n tels que $\mathbb{E}[u(X_n^+)] \to +\infty$. Pour cela il est important de bien choisir la mesure de risque ρ et, comme on montrera dans le paragraphe (4.5.3), lorsque ξ est la densité de la probabilité martingale dans un modèle de Black-Scholes et $\rho = CVaR$ alors le problème n'a pas de solution car $\triangle(A) = -\infty$. En pratique, tester si $\triangle(A) > -\infty$ peut ne pas être facile: on donne donc une condition nécessaire pour que cela soit verifiée:

> Soit γ_{min} la fonction de penalité minimale associée à la mesure de risque ρ. Si
> $$\gamma_{min}(\xi\mathbb{P}) < +\infty$$
> alors
> $$\inf_{A\in\mathcal{F}} \triangle(A) > -\infty$$

Pour les mesures de risque les plus populaires, la fonction γ_{min} est suffisamment explicite pour pouvoir tester cette condition et déduire si le problème a une solution finie. Encore plus difficile peut paraître la maximisation de $U(A, x^+(A))$ lorsque A décrit l'ensemble des événements mesurables $\mathcal{F}$. Le résultat suivant prouve qu'on peut réduire cette maximisation à une sous-classe d'événements mesurables indexée par un paramètre réel:

> Si la loi de ξ n'a pas d'atome et $\triangle(A) > -\infty$ pour tout $A \in \mathcal{F}$ alors
> $$\sup_{\rho(-(X)^-)\leq\rho_0,\, \mathbb{E}[\xi X]\leq x_0} \mathbb{E}[u(X^+)] = \sup_{c\in[\underline{\xi},\overline{\xi}]} U(\{\xi \leq c\},\, x^+(c))$$
> où $\underline{\xi} := essinf\,\xi$, $\overline{\xi} := esssup\,\xi$ et $x^+(c) := x^+(\{\xi \leq c\})$.

L'existence d'un maximum pour la fonction $c \mapsto U(\{\xi \leq c\},\, x^+(c))$ est difficile à montrer en général pour toute mesure de risque. Cependant ce résultat nous permet de trouver la solution explicite de notre problème de départ pour tout une grande classe de mesures de risque, parmi lesquelles il y a certainement les plus connues et utilisées en pratique: lorsque ρ est la mesure de risque entropique (pour laquelle la solution optimale existe toujours, Section 4.4); lorsque ρ est une mesure de risque spectrale (pour laquelle la solution optimale peut ne pas exister, Section 4.5) et lorsque ρ est une mesure de risque de type G-divergence (Section 4.6). Pour conclure, on peut remarquer que l'algorithme de résolution issu du dernier résultat est facilement implémentable numèriquement: dans le Paragraphe 4.4.2, on a pu effectivement le tester pour la mesure de risque entropique, couplé avec la fonction d'utilitè exponentielle et un modèle de Black and Scholes pour obtenir le payoff optimal pour le gestionnaire de fond (Figure 4.2) et pour l'investisseur (Figure 4.4).

Dans la deuxième partie de cette thèse, je me suis interessé au problème de couverture quadratique avec contraintes sur les stratégies. Le problème ensoi est très classique dans la littérature et plusieurs méthodes ont été developpées pour le résoudre dans un cadre très général. Ce type de couverture est devenue très populaire pour les praticiens car elle est relativement facile à mettre en place lorsque il s'agit de couvrir une option en marché incomplet, dans lequel il est bien connu que la couverture parfaite est rarement possible. Dans sa formulation générale, le problème de la couverture quadratique est le suivant:

> Soit $H \in \mathbb{L}^2(\mathcal{F}_T, \mathbb{P})$ et S une semimartingale. Sous des conditions appropriées d'intégrabilité, on cherche à
>
> $$\text{minimiser } \mathbb{E}^{\mathbb{P}}\left[\left(x + \int_0^T \theta_t dS_t - H\right)^2\right]$$
>
> lorsque $x \in \mathbb{R}$ et θ décrit un ensemble de stratégies que l'on appellera "admissibles"

Si la solution de ce problème est connue, elle n'est néanmoins pas explicite pour tout type de semimartingale. A ma connaissance, une solution semi-explicite est disponible lorsque S est une martingale ou lorsque S a des propriétés particulières (par example lorsque S est aux accroissements indépendants). D'un point de vue pratique, il est donc important de pouvoir expliciter cette solution ou proposer des méthodes numériques qui peuvent l'approcher.

Le problème reste également intéressant lorsqu'on le modifie de la manière suivante:

> Supposons que S est une semimartingale multidimensionnelle et on cherche à
>
> $$\text{minimiser } \mathbb{E}^{\mathbb{P}}\left[\left(x + \int_0^T \theta_t dS_t - H\right)^2\right]$$
>
> lorsque $x \in \mathbb{R}$ et $\theta_i = 0$ pour tout $i > 1$.

Cette formulation est intéressante en pratique car il est possible que l'on n'ait pas le droit d'investir dans une certaine classe d'actifs financiers qui, de même, peuvent interagir avec la dynamique des actifs qui entrent dans notre portefeuille. Ou aussi lorsque certains actifs financiers ne sont pas échangés sur le marché ou encore ne peuvent pas être considérés comme des actifs financiers tout court (on pense, par exemple, aux modèles à volatilité stochastique, ou' le processus de volatilité ne peut pas être utilisé comme actif de couverture).

Le modèle auquel je m'intéresse est donc le suivant:

$$\begin{aligned}
dZ_r &:= \mu\left(r, U_r, Z_r\right) dr + \sigma(r, U_r, Z_r) dW_r^1 + \int_{\mathbb{R}} \gamma\left(r, U_{r-}, Z_{r-}, y\right) \bar{J}\left(dydr\right) \\
dU_r &:= \mu^U(r, U_r) dr + \sigma^U(r, U_r) dB_r + \int_{\mathbb{R}} \gamma^U\left(r, U_{r-}, y\right) \bar{N}\left(dydr\right) \\
dP_r &:= \mu^P(r, P_r) dr + \sigma^P(r, P_r) dW_r^2 + \int_{\mathbb{R}} \gamma^P(r, P_r, y) \bar{J}\left(dydr\right)
\end{aligned}$$

où W, B sont deux mouvements Brownien et J, N deux mesures aléatoires de Poisson. L'actif financier dans lequel on peut investir est donné par $S := \exp(Z)$ et le problème de couverture quadratique devient:

Pour une fonction $f : \mathbb{R}^3 \to \mathbb{R}$ on cherche à

$$\text{minimiser } \mathbb{E}^{\mathbb{P}}\left[\left(f(U_T, P_T, Z_T) - x - \int_0^T \theta_t d\exp(Z_t)\right)^2\right]$$

lorsque $x \in \mathbb{R}$ et θ est une stratégie admissible.

Cette formulation explique bien le rôle de U et P: on imagine que U soit un facteur de risque qui "perturbe" la dynamique de notre actif financier (comme un facteur de volatilité par exemple) et P est une autre source de risque qui influence la valeur de l'option à la maturité. Ce type de problème est typique dans le marché des commodités, en particulier du marché de l'électricité, duquel d'ailleurs je me suis inspiré: en effet, dans ce marché l'actif financier qui représente le prix spot de l'électricité ne peut pas être pensé comme un instrument de couverture, même s'il influence la dynamique des autres actifs financiers. De plus, on peut bien imaginer que les options sur livraison d'électricité peuvent dépendre d'un facteur externe de risque (comme par exemple la température). Dans ce contexte, on notera par U le prix spot de l'électricité et par P la température, qui donc ne feront pas partie de la classe d'instruments financiers pour construire le portefeuille de couverture.

Dans le Chapitre 5, on commence à étudier le problème et donner ses propriétés générales. Vu la nature Markovienne de notre modèle, on utilise les techniques de la programmation dynamique pour caractériser la stratégie optimale à l'aide de l'équation de Hamilton-Jacobi-Bellman. Par des arguments de projection orthogonale dans les espaces de Hilbert, on montre d'abord que

Si v^f denote la fonction valeur du problème:

$$v^f\,(t,u,p,z,x) := \inf_\theta \mathbb{E}\left[\left(f\,(U_T,P_T,Z_T) - x - \int_t^T \theta_{r-} d\exp(Z_r)\right)^2\right]$$

où $(U_T,P_T,Z_T) := (U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})$, alors

$$v^f(t,u,p,z,x) = a\,(t,u,z)\,x^2 + b(t,u,p,z)x + c(t,u,p,z)$$

où

$$a\,(t,u,z) = \inf_\theta \mathbb{E}\left[\left(1 + \int_t^T \theta_{r-} d\exp(Z_r^{t,u,z})\right)^2\right]$$

et b^f et c^f sont deux fonctions qui dépendent de f.

La fonction a ne dépend pas de l'option à couvrir f: elle est donc universelle dans ce problème. Elle correspond à la mesure martingale optimale, qui est un outil fondamental pour résoudre le problème de couverture quadratique avec des méthodes duales. Il est important pour la suite d'avoir des propriétés de régularité sur la fonction a. En effet on peut montrer que

Il existe une constante $C > 0$ telle que

$$e^{-C(T-t)} \leq a\,(t,u,z) \leq 1, \quad \text{pour tout } t,u,z.$$

De plus, il existe $T^* > 0$ et $K^a_{lip} \geq 0$ tels que si $T < T^*$ alors on a

$$|a(t,u,z') - a(t,u,z)| \leq K^a_{lip}|z - z'|, \quad \text{pour tout } t,u,z$$

Un théorème de vérification nous permet de caractériser les fonctions a,b,c et la stratégie optimale du problème de couverture quadratique, si les fonctions a,b et c sont les uniques solutions régulières d'un système de trois PIDEs. L'étude de la régularité des ces fonctions sera fait dans les Chapitres 6 et 7.

Dans le chapitre 6, on étudie le problème lorsque on impose une condition de stricte ellipticité sur la matrice de voltilité σ. On étudie d'abord les opérateurs différentiels associés au processus (Z,U,P) et, en utilisant des techniques de contraction dans des espaces de Hölder appropriés, on arrive à montrer notre résultat principal (Théorème 6.8):

Si $T < T^*$ alors les fonctions a, b et c sont les uniques solutions de

$$\begin{aligned} 0 &= -\frac{\partial a}{\partial t} + \mathcal{A}_t a - \mathcal{B}_t a - \inf_{|\pi|\le\bar{\Pi}} \left\{2\pi\mathcal{Q}a + \pi^2\mathcal{G}a\right\} \\ 0 &= -\frac{\partial b}{\partial t} + \mathcal{A}_t b - \mathcal{B}_t b - \pi^*\mathcal{Q}_t b \\ 0 &= -\frac{\partial c}{\partial t} + \mathcal{A}_t c - \mathcal{B}_t c + \frac{1}{4}\frac{(\mathcal{Q}_t b)^2}{\mathcal{G}_t a} \end{aligned}$$

avec $a(T,.) = 1$, $b(T,.) = -2f$ et $c(T,.) = f^2$ dans l'espace de Hölder $C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^3)$ pour un $\delta \in (0,1)$. Ici $\mathcal{A} - \mathcal{B}$ denotes l'opérateur integro-différentiel associé au processus (U,P,Z) et Q,G sont introduits dans la Définition 5.10. Le contrôle optimal est donné par

$$\theta^*(t,u,p,z,x) := e^{-z}\left(\pi^*(t,u,z)x - \frac{1}{2}\frac{\mathcal{Q}_t b}{\mathcal{G}_t a}(t,u,p,z)\right)$$

où

$$\pi^*(t,u,z) := \frac{\mathcal{Q}_t a(t,u,z)}{\mathcal{G}_t a(t,u,z)}$$

On retrouve ici une des raisons qui ont fait de la couverture quadratique un outil très efficace dans la gestion de portefeuille: en effet, pour trouver la stratégie optimale, on doit résoudre une équation semi-linéaire (pour la fonction a) une fois pour toutes et après on peut déterminer la stratégie optimale en résolvant une équation linéaire (pour la fonction b), qui est relativement facile au moins numériquement. Cette structure permet donc de déterminer la stratégie optimale pour plusieurs options à couvrir au même temps, qui est numériquement efficace. La régularité de la fonction valeur permet aussi d'implémenter des schémas numériques très fiables avec des bons contrôles sur l'erreur d'approximation. De la structure de la fonction valeur, on retrouve facilement le prix de couverture optimale, simplement en minimisant sur x:

$$x^*(f)(t,u,p,z) := -\frac{b(t,u,p,z)}{2a(t,u,z)}$$

On retrouve ici une autre caractéristique de la couverture quadratique, c'est à dire la linéarité de la strategie optimale et du prix optimal par rapport à l'option f. Cet aspect est très pratique lorsque on veut couvrir une option qui est une complexe combinaison linéaire d'options simples. Non négligeable est également le fait que la linéarité du prix optimal par rapport à f est une propriété importante qu'on peut observer sur le marché au moins pour les options liquides (les options vanille pour exemple).

Dans le Chapitre 7, je donne un équivalent du résultat précédent lorsque on travaille avec des processus à sauts purs. Ce cas est très intéressant dans le contexte des marchés de commodités car, comme cela a été observé dans plusieurs travaux empiriques, les mouvements des prix des actifs financiers sont dus essentiellement à l'activité des sauts. De plus, la présence de pics dans les courbes des prix et des queues de distribution très épaisses ne peuvent pas être expliqués par un comportement Gaussien des actifs au moins à une petite échelle de temps. Le modèle simplifié qu'on va étudier est donc

$$dZ_r := \mu\left(r, Z_r\right) dr + \int_{\mathbb{R}} \gamma\left(r, Z_{r-}, y\right) \bar{J}\left(dydr\right)$$

et

$$v^f\left(t, z, x\right) := \inf_{\theta} \mathbb{E}\left[\left(f\left(Z_T^{t,z}\right) - x - \int_t^T \theta_{r-} d\exp(Z_r^{t,z})\right)^2\right]$$

On ne pourra pas s'attendre à une regularité de la fonction valeur comme dans le cas précédent, qui, on le rappelle, était due à la présence du mouvement Brownien. Pour avoir la régularité nécessaire, on assume que les petits sauts de la mesure J se comportent comme dans le cas d'un processus α-stable avec $\alpha \in (1, 2)$:

$$\nu(dy) := g(y)|y|^{-(1+\alpha)}$$

avec g positive, bornée et avec une décroissance appropriée à l'infini, pour garantir l'intégrabilité de Z.

Le choix d'une mesure de Lévy de ce type est dû au fait qu'on pourra montrer comment l'opérateur integro-différentiel associé à Z peut être approché par l'opérateur integro-différentiel associé à un processus de Lévy α-stable. Pour ce type de processus on a des estimations sur leur densité de probabilité, ce qui nous permettra finalement de réutiliser les techniques de contractions appliquées dans le cadre précédent:

$$\begin{aligned}
\mathcal{B}_t\varphi(z) &:= \int \left(\varphi(t, z + \gamma(., y)) - \varphi(t, z) - \gamma(., y)\frac{\partial\varphi}{\partial z}(t, z)\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy) \\
&\rightsquigarrow \\
\mathcal{B}_t^{st}\varphi(z) &:= \int \left(\varphi(t, z + y) - \varphi(t, z) - y\frac{\partial\varphi}{\partial z}(t, z)\mathbb{1}_{\{|y|\leq 1\}}\right)\nu^{st}(dy) \\
\nu^{st}(dy) &:= \frac{g(0^+)}{|y|^{1+\alpha}}\mathbb{1}_{\{0<y\}} + \frac{g(0^-)}{|y|^{1+\alpha}}\mathbb{1}_{\{y<0\}}
\end{aligned}$$

Néanmoins, on remarque que dans ce contexte, on n'a pas besoin de travailler avec des fonctions valeur deux fois différentiables. Le résultat auquel on parvient est le suivant:

Supposons que $\frac{d}{dy}\gamma(t,z,0) = 1$ pour tout t, z. Si $0 < \delta < \alpha - 1$ et $T < T^*$ alors les fonctions a, b et c sont les uniques solutions de

$$\begin{aligned} 0 &= -\frac{\partial a}{\partial t} - \mu\frac{\partial a}{\partial z} - \mathcal{B}_t a - \inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}a + \pi^2\mathcal{G}a\right\} \\ 0 &= -\frac{\partial b}{\partial t} - \mu\frac{\partial b}{\partial z} - \mathcal{B}_t b - \pi^*\mathcal{Q}_t b \\ 0 &= -\frac{\partial c}{\partial t} - \mu\frac{\partial c}{\partial z} - \mathcal{B}_t c + \frac{1}{4}\frac{(\mathcal{Q}_t b)^2}{\mathcal{G}_t a} \end{aligned}$$

avec $a(T,.) = 1$, $b(T,.) = -2f$ et $c(T,.) = f^2$ dans l'espace de Hölder de type 2 $H^{\alpha+\delta}([0,T]\times\mathbb{R})$ et différentiables par rapport à t; le contrôle optimal est donné par

$$\theta^*(t,z,x) := e^{-z}\left(\pi^*(t,z)x - \frac{1}{2}\frac{\mathcal{Q}_t b}{\mathcal{G}_t a}(t,z)\right), \text{ où } \quad \pi^*(t,z) := \frac{\mathcal{Q}_t a(t,z)}{\mathcal{G}_t a(t,z)}$$

La structure du contrôle optimal et ses propriétés sont les mêmes que dans le cadre précédent. Par contre, l'hypothèse sur la régularité de γ au point zero peut paraître très contraignante: à titre d'exemple, la fonction $\gamma(t,z,y) := \hat{\gamma}(t,z)y$ la vérifie si et seulement si $\hat{\gamma}(t,z) := 1$ qui réduit énormément la classe de modèles qu'on peut étudier. Cependant, dans la Section 7.5 du même chapitre, on montrera que cette hypothèse peut être supprimée si on impose des conditions de bornitude sur les dérivées de la fonction γ par rapport à y en zéro, ces conditions étant verifiés dans la plupart des modèles qu'on retrouve en pratique. Dans la Section 7.7 qui conclut le chapitre, on traite le cas ou' le processus Z est à variation finie. Dans ce cas, on n'a plus besoin d'imposer de conditions sur la fonction γ: en effet, dans ce cas là, les équations de Hamilton-Jacobi-Bellman peuvent être dérivées sans supposer de régularité particulière si la fonction de dérive $\mu = 0$. On va donc faire un changement de variable $L_t = \phi(t, Z_t)$ pour que le nouveau processu L n'ait pas de drift. Si on réécrive le problème en termes de L et on note par v^L la fonction valeur du problème alors le résultat qu'on a est le suivant:

On a $v^L(t,l,x) = x^2 a^L(t,l) + x b^L(t,l) + c^L(t,l)$ et si $T < T^*$ alors les fonctions a^L, b^L et c^L sont les uniques solutions de

$$\begin{aligned} 0 &= -\frac{\partial a^L}{\partial t} - \mathcal{B}^L a^L - \inf_{\pi \in \mathbb{R}} \left\{ 2\pi \mathcal{Q}^L a^L + \pi^2 \mathcal{G}^L a^L \right\} \\ 0 &= -\frac{\partial b^L}{\partial t} - \mathcal{B}^L b^L - \pi^* \mathcal{Q}^L b^L \\ 0 &= -\frac{\partial c^L}{\partial t} - \mathcal{B}^L c^L + \frac{(\mathcal{Q}^L b^L)^2}{4\mathcal{G}^L a^L} \end{aligned}$$

avec $a(T,.) = 1$, $b(T,.) = -2f$ et $c(T,.) = f^2$ dans l'espace de Hölder de type 2 $H^1([0,T] \times \mathbb{R})$ et différentiables par rapport à t; le contrôle optimal est donné par

$$\theta^*(t,l,x) := e^{-\phi^{-1}(t,l)} \left(\pi^*(t,l)x - \frac{1}{2} \frac{\mathcal{Q}^L b(t,l)}{\mathcal{G} a^L(t,l)} \right), \text{ où}$$

$$\pi^*(t,l) := -\frac{\mathcal{Q}_t^L a^L(t,l)}{\mathcal{G}_t^L a^L(t,l)}$$

Dans le Chapitre 8 on applique les résultats obtenus sur le problème de la couverture quadratique à un problème pratique du conteste des marchés de l'électricité. Ce Chapitre est le fruit d'une intense collaboration avec Xavier Warin de l'équipe R&D de EDF France. Après avoir donné une brève description de ces marchés (Section 8.1), on introduit dans la Section 8.2 le "*future*", un produit financier très populaire qu'on va utiliser comme instrument de couverture:

Un contrat *future* de maturité T et durée de livraison d est un produit qui permet d'acheter de l'électricité à prix fixé qui sera livré à la date T pour une période d. Son prix à la date t est noté $F_{d,T,t}$. Le problème de couverture quadratique est

$$\textbf{minimiser} \quad \mathbb{E}\left[\left(\tilde{f}(F_{d,T,t}) - x - \int_t^T \theta_{u-} dF_{d,T,t} \right)^2 \right]$$

Un modèle classique proposé dans la littérature est d'assumer que le prix du *future* soit une déformation aléatoire de la courbe des prix à la date 0:

On introduit d'abord

$$L_s = \zeta s + \int_0^s \int_{\mathbb{R}} y\bar{J}(dyds) \text{ et } A_t := \int_0^t e^{cs} dL_s$$

pour $\zeta \in \mathbb{R}$, $c \geq 0$ et J mesure de Poisson. On prend ensuite $F_{d,T,t} := \exp(\Phi(A_t))$ ou'

$$\Phi(A) := \log \left(\frac{1}{d} \int_T^{T+d} \psi(0,s) \exp\left(e^{-cs}A\right) ds \right)$$

et $s \to \psi(0,s)$ est la curve des prix à terme à la date zéro.

On montre d'abord que la dynamique de ce produit financier satisfait les hypothèses du modèle décrit dans le Chapitre 7 et ensuite on réécrit le problème de la couverture quadratique de la manière suivante:

Le processus $Z_t := \log(F_{d,T,t})$ vérifie l'EDS:

$$dZ_t = \mu(t, Z_t)dt + \int \gamma(t, Z_{t-}, y)\bar{J}(dydt)$$

où les coefficients sont donnés par

$$\begin{aligned} \gamma(t,z,y) &:= \Phi(\Phi^{-1}(z) + ye^{ct}) - z \\ \mu(t,z) &:= \zeta e^{ct}\Phi'(\Phi^{-1}(z)) + \int_{|y|\leq 1} \left(\gamma(t,z,y) - ye^{ct}\Phi'(\Phi^{-1}(z))\right) \nu(dy) \end{aligned}$$

et le problème de couverture quadratique se transforme

$$v^f(t,z,x) = \inf_\theta \mathbb{E}\left[\left(f(Z_T^{t,z}) - x - \int_t^T \theta_{u-} d\exp(Z_s^{t,z}) \right)^2 \right]$$

On peut donc appliquer nos résultats pour caractériser la fonction valeur et déterminer la stratégie optimale. Cette modélisation permet notamment de couvrir les options dont le sous-jacent est un *future* avec durée de livraison différente: par exemple

Soit $p(x) = (G - x)^+$, $d' \neq d$ et

$$h(A) := \frac{1}{d'} \int_T^{T+d'} \psi(0,s) e^{g(s)A} ds$$

Il s'ensuit que $h \circ \Phi^{-1}(Z_t) = F_{d',T,t}$. En particulier, le problème de couverture quadratique pour $f = p \circ h \circ \Phi^{-1}(Z_t)$ devient

$$\textbf{minimiser}\, \mathbb{E}\left[\left((G - F_{d',T,t})^+ - x - \int_t^T \theta_{u-} dF_{d,T,u}\right)^2\right]$$

ce qui correspond à couvrir une option *put* dont le sous-jacent est $F_{d',T}$ avec un portefeuille composé de contrats *futures* avec durée de livraison d. Cet aspect est intéressant lorsque on veut couvrir des options dont le sous-jacent n'est pas échangé sur le marché (en effet, les *future* qui sont échangés sur le marché ont des durées de livraison standardisées, 1 mois, 3 mois, etc.).

Dans la Section 8.3, on propose un schéma numérique pour résoudre les PIDEs associées introduites dans le Chapitre 7. On conclut avec la Section 8.4 ou̓ on teste nos schémas lorsque le processus de Lévy L est un NIG (Normal Inverse Gaussian). Pour ce type de processus, qui est très populaire pour les praticiens, nos résultats ne peuvent pas s'appliquer directement car les petits sauts de ce processus se comportent comme les sauts d'un processus α-stable avec $\alpha = 1$. Cependant, les résultats numériques qu'on trouve semblent être très satisfaisants et suggèrent que la fonction valeur dans ce cas particulier aussi est régulière.
Les schémas numériques utilisés dans le Chapitre 8 on montré l'importance d'étudier des PIDE sur un domaine tronqué, de la forme $[0,T] \times [-\underline{Z}, \underline{Z}]$, avec une condition de Dirichlet artificielle au bord. Dans le Chapitre 9, qui conclut cette thèse, on s'intéresse donc à une PIDE de la forme

$$\begin{cases} 0 = -\dfrac{\partial a}{\partial t} + \eta a + \mathcal{A}_t a - \mathcal{B}_t a - \mathcal{H}[a] & (t,z) \in [0,T) \times (-\underline{Z},\, \underline{Z}) \\ a(T,z) = e^{\eta T} & z \in (-\underline{Z},\, \underline{Z}) \\ a(t,z) = e^{\eta t} q(t,z) & (t,z) \in [0,T] \times (-\underline{Z},\, \underline{Z})^c \end{cases}$$

où q est une fonction régulière.

Pour simplifier, on suppose que les coefficients du processus Z ne dépendent pas de U et que la fonction de volatilité σ est strictement positive. Analyser cette PIDE directement peut s'avérer très compliqué si la mesure de Lévy n'est pas finie. L'idée est donc de transformer le problème initial, qui correspond au choix des paramètres $\big(\mu, \sigma^2, \gamma, \nu(dy)\big)$, en remplaçant les petits sauts du processus Z par un mouvement Brownien. À ces nouveaux paramètres correspond un nouveau problème d'optimisation:

Pour $h > 0$ on définit $\gamma^h(t,z) = \int_{|y|\le h} \gamma^2(t,z,y)\nu(dy)$ et les nouveaux paramètres $\big(\mu, \sigma^2 + \gamma^h, \gamma, \nu(dy)\mathbb{1}_{\{h<|y|\}}\big)$. Soit a^h la fonction valeur du problème de couverture quadratique lorsque $f = 0$ et $x = 1$, correspondant à ces nouveaux paramètres. Alors

1. $\left\|a - a^h\right\|_{2-\delta,H} \to 0$ lorsque $h \to 0$
2. $\left\|\pi^* - (\pi^h)^*\right\|_{1-\delta,H} \to 0$ lorsque $h \to 0$

où a est la fonction valeur du même problème avec les paramètres initiaux et π^* est le contrôle optimal correspondant.

Ce résultat nous donne une première approximation pour la fonction a. De plus, comme la nouvelle mesure de Lévy est finie, on déduit que les opérateurs non locales, associés au processus Z avec les nouveaux paramètres, sont d'ordre zéro. On va donc tronquer la PIDE qui caractérise la fonction a^h plutôt que celle de a et on prouve que

Si la condition de Dirichlet est suffisamment régulière alors la PIDE

$$\begin{cases} 0 = -\dfrac{\partial a^{tr}}{\partial t} + \eta a^{tr} + \mathcal{A}_t^h a^{tr} - \mathcal{B}_t^h a^{tr} - \mathcal{H}^h[a^{tr}] & (t,z) \in [0,T) \times (-\underline{Z}, \underline{Z}) \\ a^{tr}(T,z) = e^{\eta T} & z \in (-\underline{Z}, \underline{Z}) \\ a^{tr}(t,z) = e^{\eta t} q(t,z) & (t,z) \in [0,T] \times (-\underline{Z}, \underline{Z})^c \end{cases}$$

a une unique solution $a^{tr} \in C^{1+\kappa/2, 2+\kappa}([0,T] \times [-\underline{Z}, \underline{Z}])$, où $\kappa \in (0,1)$, et les opérateurs $\mathcal{A}^h$, $\mathcal{B}^h$ et $\mathcal{H}^h$ sont les opérateurs différentiels usuels, correspondant aux nouveaux paramètres.

Cela nous permet d'évaluer l'erreur entre la fonction a^{tr} et a^h:

Soit

$$\beta^{t,z} := T \wedge \inf\left\{s > t;\ \left|Z_s^{t,z}\right| \geq \underline{Z}\right\}$$

le premier instant de sortie du processus Z (correspondant aux nouveaux paramètres). Pour tout $(t,z) \in [0,T] \times [-\underline{Z}, \underline{Z}]$ on a

$$\left|a^h(t,z) - a^{tr}(t,z)\right| \leq M_1 \left\|a^h - q\right\|_{\infty,[0,T]\times[-\underline{Z},\underline{Z}]^c} \mathbb{P}\left(\beta^{t,z} < T\right)^{1/2}$$

où M_1 est une constante positive qui ne dépend pas de t, z, a^h, q ou $\underline{Z}$. De plus, il existe une constante $M_2 > 0$ tel que

$$\mathbb{P}\left(\beta^{t,z} < T\right) \leq \frac{M_2}{\underline{Z}^2}(1 + z^2)$$

On a donc une estimation de l'erreur entre la fonction a et a^{tr}, qui est dû, à la fois, à la troncature des petits sauts du processus Z et à la troncature du domaine de la PIDE. À une constante près, cette erreur est majorée par la probabilité de sortie du domaine du processus Z, qui décroit à zéro lorsque $\underline{Z} \to +\infty$.

En appendice, on trouvera des résultats techniques qui ont été utilisés au cours de cette thése: sur l'exponentielle stochastique d'une semi-martingale (Appendice A); sur une équation différentielle cubique (Appendice B); sur les espaces de Hölder (Appendice C); sur la formule de Ito pour les processus à sauts (Appendice D); sur la densité d'un processus d'Itô α-stable (Appendice E).

Chapter 2

Introduction

The main object of this thesis is to propose, investigate and solve some problems on portfolio management theory. The work is composed of two parts. In the first one we propose a new problem concerning the utility maximization theory, where the usual convex structure of the problem is removed (by a modification of the maximization criterion) and a new type of constraint is imposed on the admissible strategies, inspired by portfolio insurance problems. In the second one we solve the quadratic hedge problem for a class of discontinuous Markovian models which turns out to be well adapted in the context of commodities markets, which partially inspired this work. The mathematical tools and the methodologies used in the two parts are completely different. In the first one we privilege the so called "dual" formulation which is more adapted in the context of utility maximization theory when the market is complete, whereas in the second one, where no assumptions are made on the completeness on the market, we exploit the Markovian structure of the model in order to implement the well known dynamic programming principle and the relative Hamilton-Jacobi-Bellman equations.

The thesis starts with a brief but sufficiently complete introduction on risk measures (Chapter 3), which have become an important tool in finance. After a short discussion of their properties (Sections 3.1–3.2) we recall some of the most popular risk measures: the *Value at Risk* VaR, the *Conditional Value at Risk* CVaR and more generally the spectral risk measures (Section 3.3); the entropic risk measure and the G-divergence (Section 3.4). These special risk measures present many nice properties and are analytically tractable, so that they will be used to deduce explicit results in our non-standard utility maximization problem. The problem is presented in Chapter 4 and in Section 4.2 we develop our methodology to solve it and propose its solution. An important issue of this chapter is to show how the problem may fail to have a finite solution if the risk measure does not fill a non-degeneracy assumption. For this we provide a criterion, easy to check, which guarantees the existence of a finite solution and an algorithm to explicitly compute the optimal solution. We then start to test our results on practical examples: in Section 4.4 we use the entropic risk measure and we provide a simple numerical experiment; Section 4.5 is devoted to the study of the maximization problem when one uses a general spectral risk measure and we provide a criterion under which the problem has a finite solution. The special case of the CVaR is treated in Paragraph 4.5.3, whereas the G-divergence case is studied in Section 4.6. Section 4.7, which concludes the chapter, is devoted to a comparison with other types of portfolio insurance which have been studied in the literature.

The second part of the thesis is devoted to the quadratic hedge problem for discontinuous Markovian models. The problem in all its generality is presented in Chapter 5, starting by a short survey on what is already done in the literature (Section 5.1) and what is new in our work. As already pointed out in many previous works, a fundamental step to solve the quadratic hedge problem is the so called *pure investment problem*: basically it is the quadratic hedge problem when one wants to hedge the option with payoff $f = 0$. Both the quadratic hedge and the pure investment problem are introduced in Section 5.2 together with a general class of Markovian models used in the thesis. The model consists of a three-dimensional process (Z, U, P), where $\exp(Z)$ is the hedging instrument traded in the market, U is a risk factor in the dynamics of Z which cannot be used as a hedging instrument (as a volatility factor for example) whereas P is another factor of risk which influences the option one wants to hedge and is also not traded in the market. In Section 5.3 we recall many properties of the value function corresponding to the pure investment problem, and we use them to prove that it is uniformly bounded from below by a strictly positive constant, and Lipschitz continuous, whereas Section 5.4 shows many general properties on the structure of the quadratic hedge problem. In Sections 5.5, we first introduce the integro-differential operators related to the Markovian discontinuous model, and then we characterize the value function of the pure investment problems as the solution of a semi linear PIDE, provided that this PIDE has a unique smooth solution. This is done with a verification argument, and it also give us the optimal strategy for the pure investment problem. We repeat this procedure in Section 5.6 for the value function of the quadratic hedge problem. The existence and uniqueness of the solution of these PIDEs are studied in Chapters 6–7. We finally give a survey on the viscosity solution theory and see how it can be used in our context (Section 5.7).

In Chapter 6 we assume that the Markovian model used in the quadratic hedge problem is a non degenerate jump-diffusion, which is done by assuming strict ellipticity on the matrix of the Brownian component. Section 6.1 is devoted to the study of the integro-differential operators associated to the jump-diffusion model: in particular we focus on their behavior when one considers them as operators in an appropriate Hölder space. We obtain some fundamental results on their continuity in this space. In Section 6.2, we expose the methodology we use to prove that the HJB equation corresponding to the value function of the pure investment problem has a unique solution in a Hölder space of smooth functions. The proof is a mixture of contraction techniques in Banach spaces (classical tool for specialists in differential equations) and probabilistic techniques. Other methods to solve this problem are discussed in Paragraph 6.2.4, especially the ones making use of Backward SDEs or Sobolev spaces. Once one knows the value function of the pure investment problem, it is straightforward to characterize the value function of the quadratic hedge problem. In Section 6.3 we prove our main result concerning the quadratic hedge problem for jump-diffusion models: its value function can be characterized as the solution of a triplet of Partial integro-differential equations, the first of which is semi linear and it corresponds to the value function of the pure investment problem; the other two are linear, so relatively easy to solve (at least numerically).

The results of Chapter 6 are derived by assuming, in particular, that the matrix of the Brownian component is strictly elliptic. This assumption seems to be very restrictive if one wants to apply our results, for example, to the quadratic hedge problem in commodities markets. In these markets, it is popular to model the stock price as a purely discontinuous process, which basically corresponds to assume that Brownian component is equal to zero. Motivated by many discussions with practitioners in commodities markets, in Chapter 7 we assume that the stock price process is driven by a Poisson random measure. We start Chapter 7 by introducing a pure jump model for the stock price used for the quadratic hedge problem, for which we assume some properties on its Lévy measure (Section 7.1). In particular we assume that the small jumps of the process "look like" the jumps of an α-stable process, i.e. the Lévy measure has a density w.r.t. the Lebesgue measure, which is assumed to be a weighted deformation of the density of an α-stable Lévy process with $\alpha \in (1, 2)$. This is done since many properties are known for these processes, in particular on their density, and this will allow us to prove the smoothness of the value function. We proceed then as in Chapter 6 by studying the integro-differential operators in a new functional space that we call Hölder space of type 2. This is done since in the pure jump case we cannot expect the same regularity for the value function as before. The fundamental result in this case is that we can replace the principal term of the HJB solved by the value function with the integro-differential operator associated to an α-stable Lévy process, for which we know many properties (Paragraph 7.3). We prove (Paragraphs 7.4.1–7.4.2) that the value function of the pure investment problem can be characterized as the unique smooth solution in an appropriate Hölder space of type 2 of a semi linear PIDE. We can finally characterize the value function of the quadratic hedge problem in the pure jump case (Section 7.6), and, as in Chapter 6, we find that it solves a triplet of semi linear PIDEs. Section 7.7, which concludes the chapter, is devoted to the study of the problem when the stock price is modeled by a finite variation pure jump process (which includes the case $\alpha \in (0, 1)$ excluded before): in this relatively simple case we also find that the value function is characterized by a triplet of PIDEs which have a unique solution in the space of Lipschitz continuous functions.

We can finally apply the results provided in Chapter 7 on a practical problem from the portfolio management in electricity markets. Chapter 8 summarizes an intense and fruitful collaboration with Xavier Warin of R&D department of EDF France. We first discuss why financial instruments in electricity markets are generally modeled by pure jump processes (Section 8.1) and then we present the *future contract*, a popular hedging instrument in these markets (Section 8.2). Section 8.3 is devoted to the numerical methodology used to solve the PIDEs related to the value function of the quadratic hedge problem. We conclude the Chapter by using these schemes when the future contract is modeled as a random deformation of the forward curve, the randomness coming from a NIG process, which corresponds to the case $\alpha = 1$ in Chapter 7. Although for this case we cannot directly apply our result and then it should be considered as a degenerate case in some sense, the numerical results that we obtain are encouraging. In particular we obtain a numerical approximation for the value function of the pure investment problem and its optimal control and the profiles for the value functions of an at-the-money call and put options written on the future contract.

The numerical schemes used in Chapter 8 showed that it is important to study the PIDEs on a bounded domain and to quantify the truncation error. We do this analysis in Chapter 9 where, to simplify, we assume that the process Z does not depend on U. In a general framework it is not easy to study these PIDEs on a bounded domain, unless one assumes that the intensity measure of the process Z is finite. Since the method can be readapted for all the PIDEs, we just study the PIDE characterizing the value function of the pure investment problem a. We provide a first approximation of this value function by cutting the small jumps of the process Z and replacing them with a Brownian motion. This is equivalent to consider the model with new parameters, where, in particular, one has a finite intensity measure. This new model leads to a new value function of the pure investment problem and in Section 9.2 we are able to give an estimate on the error between the value function a and the new value function, and prove that we can make this error as small as we want, provided that the level at which we cut the jumps is small enough. Once we have approximated this value function, we concentrate on the PIDE that characterizes this new value function. We first prove that the truncated version of this PIDE also has a unique smooth solution (Section 9.3) and finally give an estimate on the error between the new value function and the unique solution of the truncated PIDE (Section 9.4).

We conclude the thesis with several appendices in which one can find many interesting technical results that we used throughout the thesis: on stochastic exponentials for semimartingales (Appendix A); on a cubic differential equation (Appendix B); on Hölder spaces (Appendix C); on Itô's formula for pure jump processes (Appendix D) and on the density of α-stable Lévy processes (Appendix E).

Part I

Portfolio Insurance under risk measure constraint

Chapter 3

An overview on Risk measures

In this chapter we will recall the axiomatic definition of risk measures and their main properties (Section 3.2). We then describe some of the most popular risk measures used in finance: the Value-at-Risk (VaR), the Conditional Value-at-Risk (CVaR) and, more generally, the spectral risk measures (Section 3.3); the Entropic risk measure and the G-divergence (Section 3.4). Since we are more interested in the use of them in risk management, we restrict ourselves to a brief survey on their axiomatic definition and their main properties. Excellent works on the subject can be found in our references.

Contents

3.1 Practical needs of the risk measures: empirical evidence 31
3.2 Law invariant risk measures 32
3.2.1 Definition and main properties 32
3.2.2 Representation of convex risk measures 33
3.3 VaR, CVaR and spectral risk measures 35
3.4 G-divergence and entropy 37
3.5 Risk Measures on $\mathbb{L}^p$-spaces 38

3.1 Practical needs of the risk measures: empirical evidence

The last decade of 1980 has seen the increasing interest for risk measures. Explicit references to them can be found in many reports by the Basel Committee of Banking Supervision (BCBS) and the well known Basel accords (Basel Committee, 1996, 2004). The main objective was to find a way to measure the exposure to risks for investors, banks, and, more generally, for financial institutions. They become a fundamental tool in risk management for banks and insurance companies since they use them to compute, for example, minimal capital requirements:

> [...]A significant innovation of the revised Framework is the greater use of assessments of risk provided by banks internal systems as inputs to capital calculations. In taking this step, the Committee is also putting forward a

> detailed set of minimum requirements designed to ensure the integrity of these internal risk assessments. [...]Basel Committee (2004)

The above citation suggests a way to define these risk measures: roughly speaking a risk measure can be thought as the extra capital one needs to add to her portfolio in order to have a new portfolio with zero risk. The question now is the following: what is a portfolio with zero risk? This procedure is not yet satisfactory but gives us an important property for any reasonable definition of risk measure: in order to define it, one only needs to be able to identify the portfolios with zero risk. In particular two portfolios with the same risk should keep the same risk exposure if one adds the same amount of capital to both of them. Remark however that deciding which portfolios have zero risk is a subjective choice. Following these ideas Artzner et al. (1999) first gave a precise definition of what should be a reasonable definition of a risk measure.
We want to give here a simple example of what should **not** be a good way to measure the risk: assume that there are two portfolios, say P_1 and P_2, at time $t = 0$, such that at time $t = 1$ they have the following distribution:

$$P_1 = \begin{cases} 1 & \text{on the set } A \\ -1 & \text{on the set } A^c \end{cases} \qquad P_2 = \begin{cases} 100 & \text{on the set } A \\ -100 & \text{on the set } A^c \end{cases}$$

where A is a set of possible scenarios with $\mathbb{P}(A) = 1/2$. If we agree to measure the portfolio's risk with the probability of being negative then $risk(P_1) = risk(P_2)$, whereas $risk(P_1 + 1) > risk(P_2 + 1)$. It follows that adding the same amount to both the portfolios changes their risk in a different way. This violate the property seen before that any reasonable risk measure should have. Remark that any investor would agree that the portfolio P_2 is more risky than P_1, so that this way of measure the portfolio's risk is not reasonable.

3.2 Law invariant risk measures

3.2.1 Definition and main properties

We now present the construction of risk measures for bounded random variables in the static case, by following the ideas of Föllmer and Schied (2004). For the dynamic definition of risk measures we refer to Frittelli and Gianin (2004); Bion-Nadal (2008, 2009) and references therein. The use of quadratic BSDEs in the dynamic risk measures theory can be found in Barrieu and El Karoui (2004, 2008) and their related bibliography.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{L}^\infty := \mathbb{L}^\infty(\Omega, \mathbb{P})$ the Banach space of (essentially) bounded random variable $X : \Omega \to \mathbb{R}$. The multidimensional case has to be carefully treated because there are some non trivial technical difficulties; however the main results that we will present can be extended in the multidimensional case (Jouini et al., 2004; Ekeland et al., 2009; Ekeland and Schachermayer, 2011).

Definition 3.1. *A law invariant risk measure ρ on $\mathbb{L}^\infty(\Omega)$ is a functional ρ : $\mathbb{L}^\infty(\Omega) \to \mathbb{R}$ verifying the following properties:*

i). For any $X \le Y$ $\mathbb{P}$-a.s. $\rho(X) \ge \rho(Y)$ (Monotonicity) *ii). For any $m \in \mathbb{R}$ $\rho(X+m) = \rho(X) - m$ (Cash Invariance)*

We say that ρ is normalized if $\rho(0) = 0$. ρ is said to be law invariant if

$$\rho(X) = \rho(Y) \text{ whenever } X \stackrel{\mathcal{L}}{=} Y$$

From now on, except when mentioned, we consider all risk measures to be law invariant.

Definition 3.2. *Let ρ be a risk measure on $\mathbb{L}^\infty(\Omega)$.*

i). We say that ρ is convex if for any $\lambda \in [0,1]$ and any $X, Y \in L^\infty(\Omega)$ one has $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$.

ii). We say that a convex risk measure ρ is coherent if for any $m \geq 0$ and any $X \in \mathbb{L}^\infty(\Omega)$ one has $\rho(mX) = m\rho(X)$

The financial meaning of Definition 3.1 is clear. More interesting are the conditions given in Definition 3.2: condition *i*) essentially says that if the risk is measured with a convex risk measure ρ then diversification decreases the risk, whereas condition *ii*) says that proportional portfolios have proportional risks.

The cash invariance and monotonicity property also gives that any risk measure is Lipschitz continuous : $|\rho(X) - \rho(Y)| \leq \|X - Y\|_\infty$. An important object related to risk measures is the so called *acceptance set*:

$$\mathcal{A}_\rho := \{X \in \mathbb{L}^\infty \mid \rho(X) \leq 0\} \tag{3.1}$$

This set has some interesting properties, in particular the fact that any risk measure can be recovered from its acceptance set. We list here some properties of $\mathcal{A}_\rho$:

Lemma 3.3. *Let $X, Y \in \mathbb{L}^\infty$ and $\mathcal{A}_\rho$ as in (3.1). Then:*

i). If $X \in \mathcal{A}_\rho$ and $Y \geq X$ then $Y \in \mathcal{A}_\rho$ and $\inf\{x \in \mathbb{R} \mid x \in \mathcal{A}_\rho\} > -\infty$

ii). ρ admits the representation :$\rho(X) = \inf\{x \in \mathbb{R} \mid x + X \in \mathcal{A}_\rho\}$

iii). If ρ is a convex risk measure then $\mathcal{A}_\rho$ is a convex subset of $\mathbb{L}^\infty$

iv). If ρ is a coherent risk measure then $\mathcal{A}_\rho$ is a convex cone in $\mathbb{L}^\infty$

Conversely, a risk measure can be defined from a suitable acceptance set: let $\mathcal{A} \subseteq \mathbb{L}^\infty$ be a set of bounded random variables which verifies the property *i*) of Lemma 3.3. Then the functional ρ defined in Lemma 3.3 *iii*) is a risk measure, which is convex if $\mathcal{A}$ is a cone, and coherent if $\mathcal{A}$ is a convex cone.

3.2.2 Representation of convex risk measures

In this paragraph we will recall some well known results on the representation of convex risk measures. We keep following Föllmer and Schied (2004) and we refer to them for the proofs. The general result is the following:

Theorem 3.4. *For any convex risk measure ρ (not necessarily law invariant) there exists a functional $\gamma_{min} : \mathcal{M}_{1,f} \to [0,1]$, where $\mathcal{M}_{1,f}$ is the set of normalized finite additive measures on $\mathcal{F}$, such that*

$$\rho(X) : \sup_{\mathcal{Q}\in\mathcal{M}_{1,f}} \left(\mathbb{E}^{\mathcal{Q}}[-X] - \gamma_{min}(\mathcal{Q})\right), \; X \in \mathbb{L}^{\infty}$$

The functional γ_{min} is called the minimal penalty function and it is related to ρ by

$$\gamma_{min}(\mathcal{Q}) := \sup_{X\in\mathcal{A}_{\rho}} \mathbb{E}^{\mathcal{Q}}[-X] \tag{3.2}$$

If ρ is coherent then γ_{min} takes values in $\{0, +\infty\}$.

The above Theorem shows that any risk measure can be characterized by a functional on $\mathcal{M}_{1,f}$. However it is preferable to have a representation for which the minimal penalty takes finite values only on *true* probabilities, which have to be $\sigma-$additive. This is possible if the risk measure has a regular behavior in the sense of the Definition below:

Definition 3.5. *We say that a convex risk measure is continuous from below if for any $X_n \nearrow X$ we have $\rho(X_n) \searrow \rho(X)$, $X_n, X \in \mathbb{L}^{\infty}$. We say that it is continuous from above if for any $X_n \searrow X$ we have $\rho(X_n) \nearrow \rho(X)$*

Theorem 3.6. *Let ρ be a convex risk measure (not necessarily law invariant) continuous from below. Then*

$$\rho(X) : \sup_{\mathcal{Q}\in\mathcal{M}_{1}} \left(\mathbb{E}^{\mathcal{Q}}[-X] - \gamma_{min}(\mathcal{Q})\right), \; X \in \mathbb{L}^{\infty}$$

where $\mathcal{M}_1$ is the set of probabilities on $(\Omega, \mathcal{F})$. In this case ρ is also continuous from above and satisfies the Fatou's property

$$X_n \to X \;\mathbb{P}\text{-a.s. then } \rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

Theorems 3.4–3.6 hold true for any convex risk measure not necessarily law invariant: when it is the case, the minimal penalty functional takes values in a particular subset of $\mathcal{M}_1$:

Theorem 3.7. *Let ρ be a law invariant convex risk measure. Then*

$$\rho(X) : \sup_{\mathbb{Q}\ll\mathbb{P}} \left(\mathbb{E}^{\mathbb{Q}}[-X] - \gamma_{min}(\mathcal{Q})\right), \; X \in \mathbb{L}^{\infty}$$

if and only if ρ is continuous from above, or equivalently, if and only ρ has the Fatou's property

$$X_n \to X \;\mathbb{P}\text{-a.s. then } \rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

In the next sections we will present some of the most popular risk measures and their main properties.

3.3 VaR, CVaR and spectral risk measures

Let $\lambda \in (0,1)$ and $X \in \mathbb{L}^{\infty}$. A λ-quantile of X is a real number in $\left[q_{\lambda}^{-}(X), q_{\lambda}^{+}(X)\right]$ where

$$q_{\lambda}^{-}(X) := \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) \geq \lambda\}$$
$$q_{\lambda}^{+}(X) := \sup\{x \in \mathbb{R} | \mathbb{P}(X < x) \geq \lambda\}$$

Definition 3.8. *The Value-at-Risk of X at level λ is defined as*

$$VaR_{\lambda}(X) := -q_{\lambda}^{+}(X) = \inf\{m \in \mathbb{R} \,|\mathbb{P}(X + m < 0) \leq \lambda\} \tag{3.3}$$

Equivalently we also have that $VaR_{\lambda}(X) = -F_X^{-1}(\lambda)$ where F_X^{-1} is a generalized inverse distribution function of X. Since the generalized inverse distribution function has at most a countable number of discontinuities, this definition does not depend on the particular choice of this function (right-continuous or left-continuous). We shall always use the definition

$$F_X^{-1}(\lambda) := \inf\{x : F(x) \geq \lambda\} \tag{3.4}$$

with the convention $\inf \emptyset = +\infty$.

It is not difficult to prove that the VaR_{λ} is a law invariant risk measure. Many examples have shown that the VaR is not a convex risk measure. This feature has some important financial consequences: in risk management diversification in the portfolio selection should decrease its risk, but if we measure this risk with the VaR then this is not always the case.

Example 3.9. *Let P_1 and P_2 the two portfolios in Section 3.1 and assume that P_1 is independent from P_2. It is easy to verify that*

$$VaR_{\lambda}(P_1) = \begin{cases} 1 & \text{if } \lambda < 0.5 \\ -1 & \text{if } \lambda \geq 0.5 \end{cases} \qquad VaR_{\lambda}(P_2) = \begin{cases} 100 & \text{if } \lambda < 0.5 \\ -100 & \text{if } \lambda \geq 0.5 \end{cases}$$

and

$$VaR_{\lambda}(\frac{1}{2}P_1 + \frac{1}{2}P_2) = \begin{cases} 50.5 & \text{if } \; 0 < \lambda < 0.25 \\ 45.5 & \text{if } \; 0.25 \leq \lambda < 0.5 \\ -45.5 & \text{if } \; 0.5 \leq \lambda < 0.75 \\ -50.5 & \text{if } \; 0.75 \leq \lambda < 1 \end{cases}$$

If now we take $0.5 \leq \lambda < 0.75$ *we obtain*

$$-45.5 = VaR_{\lambda}(\frac{1}{2}P_1 + \frac{1}{2}P_2) > \frac{1}{2}VaR_{\lambda}(P_1) + \frac{1}{2}VaR_{\lambda}(P_2) = -50.5$$

which shows that diversification does not decrease the risk.

In spite of the above example, the VaR is a popular risk measure which is widely used by practitioners since it has a simple financial interpretation. Our first example of risk measure which is, at the same time, convex and simple to use, is the so called Conditional Value-at-Risk:

Definition 3.10. *The Conditional Value-at-Risk of X at level λ is defined as*

$$CVaR_\lambda(X) := \frac{1}{\lambda}\int_0^\lambda VaR_u(X)\,du = -\frac{1}{\lambda}\int_0^\lambda F_X^{-1}(u)du \tag{3.5}$$

The following Proposition lists some properties of the $CVaR$:

Proposition 3.11. *Let $\lambda \in (0,1)$. The $CVaR_\lambda$ is a coherent risk measure and it admits the representation*

$$CVaR_\lambda(X) = \sup_{\mathbb{Q}\in\mathcal{H}_\lambda} \mathbb{E}^{\mathbb{Q}}[-X], \qquad \mathcal{H}_\lambda := \left\{\mathbb{Q} \ll \mathbb{P} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \le \frac{1}{\lambda}\,\mathbb{P}-a.s.\right\}$$

The supremum in the above representation is achieved by the probability

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}} := \frac{1}{\lambda}\left(\mathbf{1}_{\{X<q\}} + k\mathbf{1}_{\{X=q\}}\right)$$

where q is any λ-quantile in $\left[q_\lambda^-(X), q_\lambda^+(X)\right]$ and

$$k = \begin{cases} 0 & \text{if } \mathbb{P}(X=q)=0 \\ \frac{\lambda-\mathbb{P}(X<q)}{\mathbb{P}(X=q)} & \text{otherwise} \end{cases}$$

From the above Proposition we can determine the minimal penalty function for the $CVaR_\lambda$:

$$\gamma_{min}(\mathbb{Q}) := \begin{cases} 0 & \text{if } \frac{d\mathbb{Q}}{d\mathbb{P}} \le \frac{1}{\lambda}, \quad \mathbb{P}\text{-a.s} \\ +\infty & \text{otherwise} \end{cases} \tag{3.6}$$

It can also be proved that if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless then the $CVaR$ is the smallest law invariant convex risk measure to be continuous from below that dominates the VaR (Föllmer and Schied, 2004).

Proposition 3.12. *Let $\lambda \in (0,1)$, $X \in \mathbb{L}^\infty$ and q be a λ-quantile for X in $\left[q_\lambda^-(X), q_\lambda^+(X)\right]$. Then*

$$CVaR_\lambda(X) = \frac{1}{\lambda}\inf_{s\in\mathbb{R}}\left(\mathbb{E}\left[(s-X)^+\right] - \lambda s\right) = \frac{1}{\lambda}\mathbb{E}\left[(q-X)^+\right] - q$$

The representation given in Proposition 3.12 is much more simple to handle then the one given in Proposition 3.11, especially in risk management problems, since the maximization can be carried out over the real line instead of a set of probabilities. The $CVaR$ is a special case of the so called *spectral risk measure*:

Definition 3.13. *Let μ be a probability measure on $(0,1)$. The related spectral risk measure is defined as*

$$\rho_\mu(X) := \int_0^1 CVaR_u(X)\,\mu(du) \tag{3.7}$$

It is straightforward to see that ρ_μ is a coherent risk measure continuous from above, since the $CVaR$ is. In particular, if $\lambda \in (0,1)$ and $\mu(du) = \delta_\lambda(du)$ then $\rho_\mu = CVaR_\lambda$. Furthermore, from the definition of $CVaR$ we also can write

$$\rho_\mu(X) := \int_0^1 \tilde{\mu}(u) VaR_u(X)\,du \qquad \text{where} \qquad \tilde{\mu}(u) := \int_u^1 \frac{\mu(dx)}{x} \tag{3.8}$$

The function $\tilde{\mu}$ is right-continuous, non increasing and normalized: $\int_0^1 \tilde{\mu}(u)du = 1$.

Lemma 3.14. *Let* $(\rho_i)_i$ *be a family of convex risk measures such that* $\sup_i \rho_i(0) < +\infty$. *Then*

$$\rho(X) := \sup_i \rho_i(X)$$

is a convex risk measure.

Using Lemma 3.14 we can define a wide class of risk measures: for any subset of probability measure $\mathcal{M}$ on $(0,1)$ the

$$\rho_{\mathcal{M}}(X) := \sup_{\mu \in \mathcal{M}} \int_0^1 CVaR_u(X)\mu(du) \tag{3.9}$$

is a coherent risk measure continuous from above. It can be proved that if $(\Omega,\ \mathcal{F},\ \mathbb{P})$ is atomless then any coherent risk measure continuous from above can be represented by a subset of probability measures on $(0,1)$ as in (3.9) (Föllmer and Schied, 2004).

3.4 G-divergence and entropy

In this section we will introduce another class of risk measures which are particularly simple to handle in risk management problems.
Let $G : \mathbb{R}^+ \to \mathbb{R} \cup \{+\infty\}$ be a convex, increasing and non constant function, with $G(1) < \infty$ and $G(x)/x \to +\infty$ when $x \to +\infty$. The $G-$divergence of any absolutely continuous probability $\mathbb{Q} \ll \mathbb{P}$ is defined as

$$I_G(\mathbb{Q} \mid \mathbb{P}) := \mathbb{E}\left[G\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] \tag{3.10}$$

By using the G-divergence as penalty function in Theorem 3.6 we can build a new risk measure (Csiszar, 1967):

$$\rho_G(X) := \sup_{\mathbb{Q} \ll \mathbb{P}} \left(\mathbb{E}^{\mathbb{Q}}[-X] - I_G(\mathbb{Q} \mid \mathbb{P})\right) \tag{3.11}$$

The fact that $G(x)/x \to +\infty$ when $x \to +\infty$ and de la Vallée-Poussin's criterion (See, for example, Doob (1994), Chapter VI, §17) show that the supremum in the above definition is achieved by some probability measure $\mathbb{Q}^*$. Furthermore, since G is convex, a Lagrangian-type argument allows us to rewrite the above risk measure as

$$\rho_G(X) := \inf_{y \in \mathbb{R}} \left(\mathbb{E}^{\mathbb{P}}[G^*(y - X)] - y\right) \tag{3.12}$$

where $G^*(x) = \sup_{y>0}(yx - G(y))$. For a detailed proof, see for example, Csiszar (1967); Liese and Vajda (1987); Föllmer and Schied (2004). For example, if $\lambda \in (0,1)$ and

$$G(y) = \begin{cases} 0 & \text{if } 0 \le y \le \frac{1}{\lambda} \\ +\infty & \text{otherwise} \end{cases} \quad \text{then} \quad G^*(y) = \begin{cases} \frac{y}{\lambda} & \text{if } y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

then

$$\rho(X) = \frac{1}{\lambda} \inf_{y \in \mathbb{R}} \left(\mathbb{E}\left[(y - X)^+\right] - \lambda y\right)$$

which is nothing but the $CVaR_\lambda$ as stated in Proposition 3.12.

Another special case is given by $G(x) := \beta x \log(x)$ for $\beta > 0$: $I_G(\mathbb{Q} \mid \mathbb{P})$ is the well known entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$. The related entropic risk measure has the form

$$\rho_\beta(X) := \beta \log \mathbb{E}\left[e^{-\frac{X}{\beta}}\right] \tag{3.13}$$

Remark 3.15. *A heuristic Taylor expansion yields*

$$\rho_\beta \approx -\mathbb{E}[X] + \frac{1}{2\beta}\mathbb{E}\left[X^2\right]$$

Small β implies high risk aversion.

3.5 Risk Measures on $\mathbb{L}^p$-spaces

In risk management problems generally one has to deal with unbounded random variables. The domain of definition of ρ may be taken equal, for example, to some L^p space (Kaina and Rüschendorf, 2009) or a more general Orlicz space (Section 5.4 in Biagini and Frittelli (2009)). A general theory for risk measures on such spaces is available and a generalization of the representation given in Theorem 3.4 is also available. We do not go into details since it is not the scope of this thesis, however it is not difficult to extend the risk measures introduced in Sections 3.3–3.4 to $\mathbb{L}^1(\mathbb{P})$. This extension is straightforward for the VaR, the $CVaR$ and, more generally, for all spectral risk measures.

For the entropic risk measure we first remark that $\mathbb{E}[\exp(-X/\beta)]$ is always well defined and it may take the value $+\infty$. Furthermore Jensen's inequality yields $\rho_\beta(X) \geq -\mathbb{E}[X]$[1], which allows us to extend the entropic risk measure to $\mathbb{L}^1(\mathbb{P})$. Remark that now it takes values in $(-\infty, +\infty]$. A slight difference appears in the dual representation: if $X \in \mathbb{L}^1(\mathbb{P})$ then we need to write

$$\rho_\beta(X) := \sup_{\mathbb{Q} \ll \mathbb{P}, \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) \in \mathbb{L}^1(\mathbb{Q})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - E^{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right) \tag{3.14}$$

to avoid ambiguities.
For the risk measures issued from the G-divergence, we can remark that

$$\limsup_{x \to 0^+} G(x) = 0 \quad \Rightarrow \quad G^* \geq 0$$

so the right-hand side of (3.12) is well defined and we can use it as the definition of a wide class of risk measures on $\mathbb{L}^1(\mathbb{P})$. This condition is a quite standard assumption on the function G.

[1] This is actually true for all law invariant, normalized and convex risk measures which also are continuous from above if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless

Chapter 4

Portfolio Insurance

We study the problem of portfolio insurance from the point of view of a fund manager, who guarantees to the investor that the portfolio value at maturity will be above a fixed threshold. If, at maturity, the portfolio value is below the guaranteed level, a third party will refund the investor up to the guarantee. In exchange for this protection, the third party imposes a limit on the risk exposure of the fund manager, in the form of a convex monetary risk measure (Section 4.1). To enter in this portfolio insurance, the investor pays an initial fixed fee. The fund manager therefore tries to maximize the investor's utility function subject to the risk measure constraint. We give a full solution to this non-convex optimization problem in the complete market setting and show in particular that the choice of the risk measure is crucial for the optimal portfolio to exist (Section 4.2). An interesting outcome is that the fund manager's maximization problem may not admit an optimal solution for all convex risk measures, which means that not all convex risk measures may be used to limit fund's exposure in this way. We provide conditions for the existence of the solution and we also study the impact of the fee paid by the investor (Section 4.3). Explicit results are provided for the entropic risk measure (for which the optimal portfolio always exists, Section 4.4); for the class of spectral risk measures (for which the optimal portfolio may fail to exist in some cases, Section 4.5) and for the G-divergence (Section 4.6). Finally, in Section 4.7, we briefly recall some of the recent work that have been done in the Portfolio Insurance management and the connections to our work (De Franco and Tankov, 2011).

Contents

4.1 The Problem **40**
4.2 The decoupling and the solution **42**
4.3 The fee **50**
4.4 Explicit result: the Entropic risk measure **51**
4.4.1 The result 51
4.4.2 Numerical example 53
4.5 Explicit result: the spectral risk measure **56**
4.5.1 The result 56
4.5.2 Proof of Lemma 4.14 60
4.5.3 The special case: $CVaR_\beta$ 60
4.6 Explicit result: the G-divergence **61**

4.7 Portfolio Insurance: a short review 63

4.1 The Problem

We consider the problem of a fund manager who wants to structure a portfolio insurance product where the investors pay the initial value v_0 at time 0 and are guaranteed to receive at least the amount z at maturity T. We assume that if, at time T, the value of the fund's portfolio V_T is smaller than z, a third party pays to the investor the shortfall amount $z - V_T$. In practice, this guarantee is usually provided by the bank which owns the fund, subject to a fee f. The final payoff for the investor will be

$$\text{Payoff} = \max(V_T, z) \tag{4.1}$$

In exchange, the third party imposes a limit on the risk of shortfall $-(V_T - z)^-$, represented by a law-invariant convex risk measure ρ . Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered

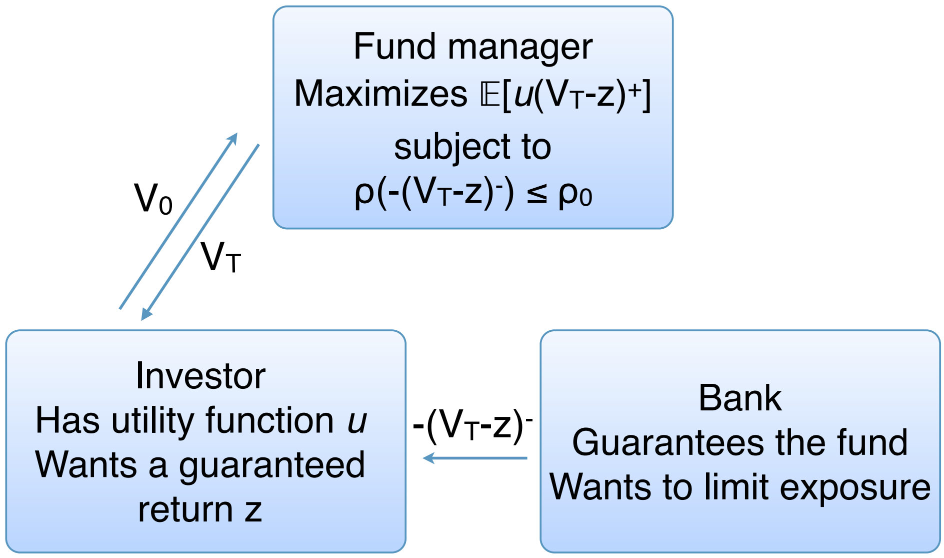

Figure 4.1: The structure of the portfolio insurance.

probability space. We consider an arbitrage-free complete financial market consisting of d risky assets with $(\mathcal{F}_t)$-adapted price processes $(S_t^i)_{0\leq t\leq T}^{i=1,\ldots,d}$ and the risk-free asset with price process $S_t^0 \equiv 1$. We do not specify the dynamics of risky assets and the precise definition of admissible strategies because they are not relevant for what follows. See Karatzas and Shreve (1998) for the standard example of a market which satisfies our assumptions in the Brownian filtration. For an admissible trading strategy π, the investor's portfolio value is

$$V_T^\pi = v_0 + \int_0^T \pi_u dS_u$$

The unique martingale measure will be denoted by $\mathbb{Q}$, and we define $\xi := \frac{d\mathbb{Q}}{d\mathbb{P}}$. The market completeness implies that for any $\mathcal{F}_T$-measurable random variable X with $\mathbb{E}[\xi|X|] < \infty$ such that $\mathbb{E}[\xi X] = v_0$, there exists an admissible trading strategy π such that $V_T^\pi := v_0 + \int_0^T \pi_t dS_t = X$ a.s. Since the interest rate is zero, $z \leq v_0$ to avoid direct arbitrage for the investor.

Without loss of generality we will assume $z = 0$ in the rest of the chapter. The attitude of the investor towards gains above 0 is measured, in the spirit of the Von Neumann-Morgenstern expected utility theory, by a twice differentiable, strictly concave and strictly increasing function $u : [0, +\infty) \to \mathbb{R}$, satisfying the usual condition $\lim_{x\to+\infty} u'(x) = 0$. We suppose $u(0) = 0$ and we denote $u^\star(y) = \sup_{x\geq 0}(u(x) - xy)$ the convex conjugate of u and $I(y) := (u')^{-1}(y)$ if $y < \lim_{x\downarrow 0} u'(x)$ and $I(y) = 0$ otherwise. Moreover, we assume that the following integrability condition holds: $\mathbb{E}[u^\star(\lambda\xi)] < \infty$ for all $\lambda > 0$. Remark that the investor payoff is given by $\max(V_T, 0)$ and that the utility function u takes value on the positive real line: in other worlds we are assuming that the utility of the portfolio is given by $u(V_T^+)$, as if the investor was indifferent to the portfolio's value below the guarantee $z = 0$.

The risks are measured using a convex law-invariant risk measure (not necessarily normalized) and continuous from above $\rho : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ where the domain of definition $\mathcal{X}$ is a subset of $\mathbb{L}^1(\xi\mathbb{P})$ (Chapter 3, Section 3.5). To simplify notation later on, we additionally define $\rho(X) = +\infty$ if $X \leq 0$ and $X \notin \mathcal{X}$.
The fund manager therefore faces the following problem:

$$\text{maximize} \quad E[u((V_T^\pi)^+)] \tag{4.2}$$

$$\text{subject to} \quad \rho(-(V_T^\pi)^-) \leq \rho_0 \quad \text{and} \quad V_0 := x_0 := v_0 - f. \tag{4.3}$$

where $\rho_0 > \rho(0)$ represents the risk tolerance allowed by the third party and f is a fee that the investor pays to enter in this portfolio insurance. This is a nonstandard maximization problem, because the objective function is not concave, therefore it cannot be solved using standard Lagrangian methods. We now assume that the fee $f = 0$, and we will discuss in Section 4.3 the case $f \neq 0$.

Using the market completeness, the optimization problem (4.2)–(4.3) can be reformulated as the problem to find, if it exists, a $X^* \in H$ such that

$$\mathbb{E}\left[u\left((X^*)^+\right)\right] = \sup_{X\in H} \mathbb{E}\left[u\left(X^+\right)\right] \tag{4.4}$$

where

$$H := \left\{X \in \mathbb{L}^1(\xi\mathbb{P}) \,\middle|\, \mathbb{E}[\xi X] \leq x_0, \rho\left(-X^-\right) \leq \rho_0\right\} \tag{4.5}$$

To simplify the notation, let us define $U(X) := \mathbb{E}[u(X^+)]$. Table 4.1 summarizes the assumption made above.

Market	Complete. Martingale measure: $\mathbb{Q}$ $\xi := d\mathbb{Q}/d\mathbb{P}$
Utility function	u' vanishing at infinity, $u(0) = 0$. We denote $I(y) := (u')^{-1}(y)$ if $y < u'(0^+)$, 0 otherwise
Integrability conditions for admissible pay-offs	$\mathbb{E}[u^\star(\lambda\xi)] < \infty$ for all $\lambda > 0$ where $u^\star(y) = \sup_{x\geq 0}(u(x) - xy)$.
Risk measure	ρ: law invariant convex risk measure continuous from above not necessarily normalized
Initial wealth, guarantee, risk tolerance and fee	x_0, $z = 0$, $\rho(0) < \rho_0$ and $f = 0$

Table 4.1: Portfolio Insurance problem: Assumptions.

4.2 The decoupling and the solution

In this Section we will decouple problem (4.2)–(4.3) into two convex problems for which Lagrangian methods are available. Let us start by remarking that for any $X \in H$ we have $\mathbb{E}[u(X^+)] = \mathbb{E}[u(X\mathbb{1}_A)]$, where $A := \{X \geq 0\}$. A financial interpretation of this equality is that only $X\mathbb{1}_A$ remains important for the investor. This remark suggests the following decoupling: let $(A, x^+) \in \mathcal{F} \times \mathbb{R}^+$ and consider

$$\begin{aligned} &\mathcal{P}_1 : \text{maximize } U(Z) \\ &\qquad \text{subject to } Z \in \mathcal{H}_1\left(A, x^+\right) \text{where} \qquad (4.6)\\ &\quad \mathcal{H}_1\left(A, x^+\right) := \left\{Z \in \mathbb{L}^1(\xi\mathbb{P}) \mid \mathbb{E}[\xi Z] \leq x^+,\ \ Z = 0 \text{ on } A^c,\ Z \geq 0 \text{ on } A\right\} \\ &\mathcal{P}_2 : \text{minimize } \mathbb{E}[\xi Y] \\ &\qquad \text{subject to } Y \in \mathcal{H}_2(A) \text{where} \qquad (4.7)\\ &\quad \mathcal{H}_2(A) := \left\{Y \in \mathbb{L}^1(\xi\mathbb{P}) \mid \rho(Y) \leq \rho_0,\ Y = 0 \text{ on } A,\ Y \leq 0 \text{ on } A^c\right\} \end{aligned}$$

Problem $\mathcal{P}_2$ is a minimization of a linear function over a convex set and, as we will see later, Problem $\mathcal{P}_1$ can be viewed as a concave maximization problem under a linear constraint.

Definition 4.1. *For all $A \in \mathcal{F}$ we define:*

$$U\left(A, x^+\right) := \sup_{Z\in\mathcal{H}_1(A,x^+)} U(Z) \qquad (4.8)$$

$$\triangle(A) := \inf_{Y\in\mathcal{H}_2(A)} \mathbb{E}[\xi Y] \qquad (4.9)$$

$$x^+(A) := x_0 - \triangle(A) \qquad (4.10)$$

We will often refer to $\triangle(A)$ as the value function of problem $\mathcal{P}_2$, to $U(A, x^+)$ as the value function of problem $\mathcal{P}_1$ and to $x^+(A)$ as the extra capital.

We first study Problems $\mathcal{P}_1$ and $\mathcal{P}_2$ and then we clarify the relationship between these problems and problem (4.4).

Remark 4.2. *Before going on let us investigate the behavior of $\mathcal{P}_1$ and $\mathcal{P}_2$ on trivial sets. If $\mathbb{P}(A) = 0$ then $0 \in \mathcal{H}_2(A)$ and then $\triangle(A) \leq 0$ which means that $x^+(A) \geq x_0 \geq 0$. Therefore, $0 \in \mathcal{H}_1(A, x^+(A))$ and $U(A, x^+(A)) = u(0)$.*

Lemma 4.3. *Suppose $\mathbb{P}(A) > 0$. The unique maximizer of problem $\mathcal{P}_1$ is given by*

$$Z\left(A, x^+\right) = I\left(\lambda\left(A, x^+\right)\xi\right)\mathbb{1}_A$$

where $\lambda(A, x^+)$ is the unique solution of

$$\mathbb{E}\left[\xi I\left(\lambda\left(A, x^+\right)\xi\right)\mathbb{1}_A\right] = x^+. \tag{4.11}$$

The value function $U(A, x^+)$ is strictly increasing and continuous in x^+, and for every $c > 0$ there exists $C < \infty$ such that

$$U(A, x^+) \leq C + cx^+ \tag{4.12}$$

for all $A \in \mathcal{F}$ and all $x^+ \geq 0$.

Proof.
Introduce the new probability space $(A, \mathcal{F}_A := \{B \cap A, B \in \mathcal{F}\}, \mathbb{P}(\cdot \mid A))$ and let $\mathbb{E}_A$ denote the expectation under the conditional probability $\mathbb{P}(\cdot \mid A)$. The maximizer of $\mathcal{P}_1$, if it exists, is given by

$$Z\left(A, x^+\right) = W\left(A, x^+\right)\mathbb{1}_A$$

where $W(A, x^+)$ is the maximizer of the following problem on the new probability space:

$$\sup_{W \geq 0} \mathbb{E}_A\left[u\left(W\right)\right] \qquad \text{subject to} \qquad \mathbb{E}_A\left[\xi W\right] = \frac{x^+}{\mathbb{P}(A)}$$

This is a classical problem of maximizing a concave function under a linear constraint which can be solved by Lagrangian methods (see e.g., Karatzas and Shreve (1998)). Remark first that $u^\star$ is continuously differentiable and the mapping $\lambda \mapsto \mathbb{E}[u^\star(\lambda\xi)]$ is convex and finite for all λ, so then it is almost everywhere differentiable. Moreover, from the definition of u^*, we have

$$(u^*)'(\lambda\xi) = -\lambda\xi I(\lambda\xi) = u^*(\lambda\xi) - u(I(\lambda\xi))$$

so that $\xi I(\lambda\xi) \in \mathbb{L}^1(\mathbb{P})$. The dominated convergence applies and we deduce that $\lambda \mapsto \mathbb{E}[u^\star(\lambda\xi)]$ is differentiable everywhere. In particular $\mathbb{E}[\xi(u^\star)'(\lambda\xi)] = -\mathbb{E}\left[\xi I\left(\lambda\xi\right)\right] < +\infty$ for all $\lambda > 0$. Therefore, the solution to the above optimization problem is

$$W\left(A, x^+\right) = I\left(\lambda\left(A, x^+\right)\xi\right)$$

where $\lambda(A, x^+)$ is the unique solution of $\mathbb{E}_A\left[\xi I\left(\lambda\xi\right)\right] = x^+/\mathbb{P}(A)$.

To show that $x^+ \mapsto U(A, x^+)$ is strictly increasing, let $x_1^+ < x_2^+$. Then the random variable

$$X = I\left(\lambda\left(A, x_1^+\right)\xi\right)\mathbb{1}_A + \frac{x_2^+ - x_1^+}{\mathbb{E}\left[\xi \mathbb{1}_A\right]}\mathbb{1}_A$$

belongs to $\mathcal{H}_1\left(A, x_2^+\right)$, which proves that $U\left(A, x_1^+\right) < U\left(A, x_2^+\right)$.

The continuity of

$$x^+ \mapsto U\left(A, x^+\right)$$

follows from inequality

$$u(I(\lambda\xi)) \leq u^\star(c\xi) + c\xi I(\lambda\xi) \tag{4.13}$$

which holds true for any $c > 0$, and the continuity of $x^+ \mapsto \lambda(A, x^+)$, which is straightforward since the function $\lambda \mapsto \mathbb{E}[\xi I(\lambda\xi)\mathbb{1}_A]$ is strictly decreasing and continuous. The upper bound on U is also a consequence of (4.13), after taking expectations, where $C = \mathbb{E}\left[u^*(c\xi)\mathbb{1}_A\right]$.

□

We can now clarify the relationship between Problems (4.4) and $\mathcal{P}_1$–$\mathcal{P}_2$.

Theorem 4.4. *Assume that*

$$\text{for all } A \in \mathcal{F},\ \triangle(A) > -\infty \tag{4.14}$$

Then,

$$\sup_{X \in H} U(X) = \sup_{A \in \mathcal{F}} U\left(A, x^+(A)\right) \tag{4.15}$$

If, in addition, $sup_x u(x) = \infty$ *and*

$$\inf_{A \in \mathcal{F}} \triangle(A) > -\infty \tag{4.16}$$

then both sides of (4.15) *are finite.*

Proof.
We start with the inequality "$\leq$". Let $X^n \in H$ such that $U(X^n) \uparrow \sup_{X \in H} U(X)$. Define $A_n := \{X^n \geq 0\}$ and $x_n := \mathbb{E}\left[\xi X^n \mathbf{1}_{A_n}\right]$. We have then

$$U(X^n) = U(X^n \mathbf{1}_{A^n}) \leq U(A^n, x_n)$$

because $X^n \mathbf{1}_{A^n} \in \mathcal{H}_1(A_n, x_n)$ and $U(A^n, x_n)$ is the supremum over $\mathcal{H}_1(A_n, x_n)$. The random variable $Y^n := X^n - X^n \mathbb{1}_{A_n}$ belongs to $\mathcal{H}_2(A_n)$ and verifies

$$x_0 - x_n = \mathbb{E}\left[\xi Y^n\right] \geq \inf_{Y \in \mathcal{H}_2(A_n)} \mathbb{E}\left[\xi Y\right] = \triangle(A_n) = x_0 - x^+(A_n)$$

It follows $x_n \leq x^+(A_n)$ and since $U(A, x^+)$ is nondecreasing in x^+ we deduce

$$U(X^n) = U(X^n \mathbf{1}_{A^n}) \leq U(A^n, x_n) \leq U(A^n, x_+(A_n)) \leq \sup_{A \in \mathcal{F}} U(A, x_+(A))$$

Let us prove the inequality "$\geq$". Select $A_n \in \mathcal{F}$ such that

$$U\left(A_n, x_+\left(A_n\right)\right) \uparrow \sup_{A \in \mathcal{F}} U\left(A, x_+\left(A\right)\right) := m, \text{ when } n \to +\infty$$

By the assumption of the Theorem, $x^+(A_n) < \infty$ for all n. If for fixed $\varepsilon > 0$ we could find, for all n, $X_n \in H$ such that

$$U(X_n) \geq U\left(A_n, x_+\left(A_n\right)\right) - \varepsilon \tag{4.17}$$

then we are done. If $\mathbb{P}(A_n) > 0$, by Lemma 4.3 there exists an explicit maximizer of Problem $\mathcal{P}_1$, denoted by $Z(A_n, x^+)$, and recall that $U(A_n, x^+) = U(Z(A_n, x^+))$ is continuous in x^+. Therefore, we can find $Y_n \in \mathcal{H}_2(A_n)$ with $\mathbb{E}[\xi Y_n]$ sufficiently close to $\triangle(A_n)$ so that $U(A_n, x_0 - \mathbb{E}[\xi Y_n]) \geq U(A_n, x^+(A_n)) - \varepsilon$. Then $X_n := Z(A_n, x_0 - \mathbb{E}[\xi Y_n]) + Y_n$ satisfies (4.17). If $\mathbb{P}(A_n) = 0$ then, as we saw in Remark 4.2, taking $0 \in H$ and $X_n = 0$ satisfies $U(X_n) = u(0) = U\left(A_n, x_+\left(A_n\right)\right)$.

Finally, the fact that $m < \infty$ under Assumption (4.16) follows directly from the estimate (4.12).

$\square$

Clearly, (4.14) depends on the particular choice of ρ and ξ. In particular, a choice under which $\triangle\left(A\right) = -\infty$ for some A is not appropriate in this kind of portfolio insurance. As we will see later on an example, the use of the CVaR_λ in the Black and Scholes model yields $\triangle\left(A\right) = -\infty$, whereas the same risk measure coupled with a bounded ξ satisfies (4.16). A simple example clarifies why assumption (4.14) is fundamental in this kind of portfolio insurance:

Example 4.5. *Assume that* $\sup_x u(x) = +\infty$ *anc fix* $A \in \mathcal{F}$ *with* $\mathbb{P}\left(A\right) > 0$ *and* $\triangle\left(A\right) = -\infty$*. It is then possible to find, for each* $n \in \mathbb{N}$ *a random variable* $Y^n \in \mathcal{H}_2\left(A\right)$ *such that* $\mathbb{E}\left[\xi Y^n\right] \leq -n$*. Define now*

$$X^n = \frac{x_0 + n}{\mathbb{E}\left[\xi \mathbb{1}_A\right]} \mathbb{1}_A + Y^n$$

It is clear that $X^n \in H$ *for all* n *and* $U\left(X^n\right) \to \sup_x u\left(x\right) = +\infty$*, which means that Problem* (4.4) *does not admit a maximizer.*

Nevertheless, in practice, it may be difficult to check whenever Assumptions (4.14)–(4.16) hold true: the following proposition, which is simpler to verify, guarantees them but it is not necessary.

Proposition 4.6. *Assume that*

$$\gamma_{min}\left(\xi \mathbb{P}\right) < +\infty, \tag{4.18}$$

where γ_{min} *is the minimal penalty function of* ρ *defined in* (3.2)*. Then the condition* (4.16) *holds true.*

Proof.
By definition of γ_{min} in (3.2) and the acceptance set related to ρ in (3.1) we have,

$$\begin{aligned}\gamma_{min}(\xi\mathbb{P}) &= \sup_{Y\in\mathcal{A}_\rho}\mathbb{E}[-\xi Y]\\ &= -\rho_0 + \sup\{\mathbb{E}[-\xi Y] \mid Y+\rho_0\in\mathcal{A}_\rho\}\\ &\geq -\rho_0 + \sup\{\mathbb{E}[-\xi Y] \mid Y+\rho_0\in\mathcal{A}_\rho,\, Y\leq 0\}\\ &\geq -\rho_0 + \sup\{\mathbb{E}[-\xi Y] \mid Y+\rho_0\in\mathcal{A}_\rho,\; Y\leq 0,\, Y=0 \text{ on } A\}\\ &= -\rho_0 - \triangle(A)\end{aligned}$$

from which the result follows.

□

Theorem 4.4 gives us a condition under which the value function of problem (4.4) is finite and a way to compute it:

Algorithm 4.7.

1. *fix* $A\in\mathcal{F}$
2. *solve* $\mathcal{P}_2(A)$ *and find* $\triangle(A)$
3. *solve* $\mathcal{P}_1(A)$ *with parameter* $(A, x^+(A))$
4. *maximize the value function of problem* $\mathcal{P}_1$ $U(A, x^+(A))$ *over* $A\in\mathcal{F}$

The next result establishes a relationship between the maximizers of problem (4.4) and $\mathcal{P}_1$–$\mathcal{P}_2$.

Theorem 4.8. *Let* (4.14) *hold true.*
If X^* *achieves the maximum in Problem* (4.4) *and* $A^* := \{X^*\geq 0\}$ *then*

- A^* *achieves the maximum in the right-hand side of* (4.15)
- $Y^* := X^* - X^*\mathbb{1}_{A^*}\in\mathcal{H}_2(A^*)$ *achieves the minimum in* $\mathcal{P}_2$.

Conversely, let $A^*\in\mathcal{F}$, $\mathbb{P}(A^*)>0$ *and* $Y^*\in\mathcal{H}_2(A^*)$ *such that*

$$U\left(A^*, x^+(A^*)\right) = \sup_{A\in\mathcal{F}} U\left(A, x^+(A)\right)$$
$$\mathbb{E}[\xi Y^*] = \triangle(A^*) = \inf_{Y\in\mathcal{H}_2(A^*)}\mathbb{E}[\xi Y]$$

Then a solution of problem (4.4) *is given by*

$$X^* := I(\lambda^*\xi)\,\mathbb{1}_{A^*} + Y^* \tag{4.19}$$

where $\lambda^* = \lambda(A^*, x^+(A^*))$ *verifies* (4.11). *In this case, the payoff for the investor will be*

$$\text{Payoff} = I(\lambda^*\xi)\,\mathbb{1}_{A^*} \tag{4.20}$$

Proof.
Let $X^* \in H$ be an optimal solution for (4.4), $A^* = \{X^* \geq 0\}$ and $Y^* = X^* \mathbb{1}_{\{X^*<0\}}$. It is clear that $Y^* \in \mathcal{H}_2(A^*)$. It is also clear that $\mathbb{P}(A) > 0$, since otherwise $\mathbb{E}[\xi X^*] < x_0$ which of course is not optimal. Theorem 4.4 and the fact that $U(A, x^+)$ is increasing in x^+ (Lemma 4.3) give:

$$\begin{aligned}\sup_{A\in\mathcal{F}} U\left(A, x^+(A)\right) = \sup_{X\in H} U(X) = U(X^*) = U\left(X^* \mathbf{1}_{A^*}\right)\\ = U\left(A^*, x_0 - \mathbb{E}[\xi Y^*]\right) \leq U\left(A^*, x^+(A^*)\right)\end{aligned}$$

which means that A^* achieves the supremum in (4.15). Since $U(A, x^+)$ is strictly increasing in x^+, we shall have $x^+(A^*) = x_0 - \mathbb{E}[\xi Y^*]$, otherwise we would obtain a strict inequality in the second line of the above estimate, which, of course, yields a contradiction. It follows then that Y^* achieves the minimum in $\mathcal{P}_2$.

Conversely, assume that A^* is a maximizer of (4.15) and Y^* is a minimizer of $\mathcal{P}_2$. We can then solve Problem $\mathcal{P}_1$ with parameters $(A^*, x_0 - \triangle(A^*))$ and we know, by Lemma 4.3, that its solution is given by $\left[I(\lambda^*\xi)^+\right]^+ \mathbb{1}_{A^*}$. Let then

$$X^* := I(\lambda^*\xi)\, \mathbb{1}_{A^*} + Y^*$$

We have $\rho(-(X^*)^-) = \rho(Y^*) \leq \rho_0$ and $\mathbb{E}[\xi X^*] \leq x_0$, i.e. $X^* \in H$. Using Theorem 4.4, we conclude our proof:

$$U(X^*) = U\left(X^* \mathbf{1}_{A^*}\right) = U\left(A^*, x^+(A^*)\right) = \sup_{A\in\mathcal{F}} U\left(A, x^+(A)\right) = \sup_{X\in H} U(X).$$

□

Remark 4.9. *Algorithm 4.7 and Theorem 4.4 give us a way to find an optimal solution for problem* (4.4) *if we are able to find a maximizer in* (4.15) *and the minimizer in* $\mathcal{P}_2$.

But what happens in the case when the maximizer in (4.15) *or the minimizer in* $\mathcal{P}_2$ *do not exist? In this case, under Assumption 4.14, following the steps of the proof of Theorem 4.4, one can show that for all* $\varepsilon > 0$ *there exist* $A^\varepsilon \in \mathcal{F}$, $\lambda^\varepsilon \in \mathbb{R}$ *and* $Y^\varepsilon \in \mathcal{H}_2(A^\varepsilon)$ *such that*

$$X^\varepsilon := \left[I(\lambda^\varepsilon\xi)\right] \mathbb{1}_{A^\varepsilon} + Y^\varepsilon \tag{4.21}$$

verifies $U(X^\varepsilon) + \varepsilon > \sup_{X\in H} U(X)$, *i.e.* X^ε *is a sequence of* ε*-optimal solutions.*

The main difficulty to apply Theorems 4.4–4.8 is to find a maximizer A^*. Generally, maximization of a set-valued function over $\mathcal{F}$ is not simple. Our aim now is to show that this latter maximization may be carried out over a subset of $\mathcal{F}$, parameterized by a real number. A similar approach was used in Jin and Yu Zhou (2008), where they faced the same difficulty.

Theorem 4.4 tells us that under Assumption (4.14) we have

$$\sup_{X\in H} U(X) = \sup_{A\in\mathcal{F}} U\left(A, x^+(A)\right) = \sup_{A\in\mathcal{F}} \sup_{X\in\mathcal{H}_1(A, x^+(A))} U(X)$$

In order to focus our attention on the set dependence, we will introduce the following notation:

$$v\left(A\right) := \sup_{X\in\mathcal{H}_1(A,x^+(A))} U\left(X\right) \tag{4.22}$$

Let us also define $\underline{\xi} := essinf\, \xi$ and $\overline{\xi} := esssup\, \xi$.

Theorem 4.10. *Suppose that the law of* ξ *has no atom and that Assumption* (4.14) *holds true. Let* $A \in \mathcal{F}$ *and* $c \in \left[\underline{\xi}, \overline{\xi}\right]$ *such that* $\mathbb{P}\left(\xi \leq c\right) = \mathbb{P}\left(A\right)$. *Then*

$$v\left(A\right) \leq v(c), \qquad \textit{where} \qquad v(c) := v\left(\{\xi \leq c\}\right) \tag{4.23}$$

which means that

$$\sup_{X\in H} U\left(X\right) = \sup_{A\in\mathcal{F}} v\left(A\right) = \sup_{c\in\left[\underline{\xi},\overline{\xi}\right]} v(c). \tag{4.24}$$

Proof.
We will use the methods developed in Jin and Yu Zhou (2008) (see the proof of Theorem 5.1 therein). There are however some important differences in our proof which are due to the presence of a risk measures in our context.

The theorem will be proved in two steps: in Step 1 we will prove that for every $A \in \mathcal{F}$, there exists $c \geq 0$ such that $\triangle\left(A\right) \geq \triangle\left(c\right) := \triangle\left(\{\xi \leq c\}\right)$ so that $x_+\left(c\right) := x_0 - \triangle\left(\{\xi \leq c\}\right) \geq x_+\left(A\right)$, and in Step 2 we will find, for every $X \in \mathcal{H}_1\left(A, x^+\right)$ some $\hat{X} \in \mathcal{H}_1\left(\{\xi \leq c\}, x^+\left(c\right)\right)$ such that $U(\hat{X}) \geq U\left(X\right)$. We conclude then that $v\left(c\right) \geq v\left(A\right)$

If $\mathbb{P}(A) = 1$ then the result trivially holds true, whereas if $\mathbb{P}(A) = 0$ we can use Remark 4.2 and again the result holds true. Assume then $0 < \mathbb{P}\left(A\right) < 1$ and define $\alpha = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. Let us fix $c \in \left[\underline{\xi}, \overline{\xi}\right]$ so that

$$\mathbb{P}\left(\xi \leq c\right) = 1 - \alpha$$

This is possible since ξ has no atom. Consider the following sets:

$$A_1 = \{\xi \leq c\} \cap A \qquad A_2 = \{\xi > c\} \cap A \tag{4.25}$$

$$B_1 = \{\xi \leq c\} \cap A^c \qquad B_2 = \{\xi > c\} \cap A^c \tag{4.26}$$

Since $\mathbb{P}(A_1)+\mathbb{P}(A_2) = \mathbb{P}(A_1)+\mathbb{P}(B_1) = 1-\alpha$ it follows $\mathbb{P}(A_2) = \mathbb{P}(B_1)$. If $P\left(A_2\right) = 0$ then $A = \{\xi \leq c\}$ and the result trivially holds true. We can suppose $\mathbb{P}(A_2) > 0$.
Step 1. Let $Y \in \mathcal{H}_2(A)$. Our aim is to construct $\hat{Y} \in \mathcal{H}_2(\{\xi \leq c\})$ with $\mathbb{E}[\xi\hat{Y}] = \mathbb{E}[\xi Y]$ and $\rho(Y) \geq \rho(\hat{Y})$. This will imply that $\triangle\left(A\right) \geq \triangle(c)$ since we can decrease $\hat{Y}$. Introduce the following notation:

1. $f_1\left(t\right) := \mathbb{P}\left(Y \leq t \,|\, B_1\right)$
2. $g_1\left(t\right) := \mathbb{P}\left(\xi \leq t \,|\, A_2\right)$
3. $Z_1 = g_1\left(\xi\right)$, that is, $\mathcal{L}\left(Z_1 \mid A_2\right) = \mathcal{U}([0,1])$, because ξ has no atom.
4. $W_1 = f_1^{-1}\left(Z_1\right)$, that is, the law of W_1 on A_2 is the same as the law of Y on B_1.

Let

$$k_1 := \begin{cases} 1 & \text{if } W_1 = 0 \text{ on } A_2 \\ \frac{\mathbb{E}[\xi Y \mathbb{1}_{B_1}]}{\mathbb{E}[\xi W_1 \mathbb{1}_{A_2}]} & \text{otherwise} \end{cases}$$

Observe that since $\xi \leq c$ on B_1, and $\xi > c$ on A_2, we have that $k \leq 1$. Now define

$$\hat{Y} = Y \mathbb{1}_{B_2} + k W_1 \mathbb{1}_{A_2}.$$

By definition, $\hat{Y} = 0$ on $\{\xi \leq c\}$ and $\hat{Y} \leq 0$ on $\{\xi \leq c\}$. In addition, since $k_1 \leq 1$, we easily get that $\mathbb{P}(-\hat{Y} > t) \leq \mathbb{P}(-Y > t)$ for every $t > 0$:

$$\begin{aligned} \mathbb{P}(-\hat{Y} > t) =& \mathbb{P}(B_2)\mathbb{P}(-\hat{Y} > t \mid B_2) + \mathbb{P}(A_2)\mathbb{P}(-\hat{Y} > t \mid A_2) \\ =& \mathbb{P}(B_2)\mathbb{P}(-Y > t \mid B_2) + \mathbb{P}(B_1)\mathbb{P}(-kW_1 > t \mid A_2) \\ \geq& \mathbb{P}(B_2)\mathbb{P}(-Y > t \mid B_2) + \mathbb{P}(B_1)\mathbb{P}(-W_1 > t \mid A_2) \text{ since } k_1 \leq 1 \\ =& \mathbb{P}(B_2)\mathbb{P}(-Y > t \mid B_2) + \mathbb{P}(B_1)\mathbb{P}(-Y > t \mid B_1) \\ =& \mathbb{P}(-Y > t) \end{aligned}$$

Let F and $\hat{F}$ be the distribution functions of, respectively, $-Y$ and $-\hat{Y}$, and F^{-1} and $\hat{F}^{-1}$ their generalized inverses (defined in (3.4)). From the above inequality, they satisfy $\hat{F}^{-1}(u) \leq F^{-1}(u)$ for all $u \in [0,1]$. Let U be a random variable with uniform distribution on $[0,1]$. Since ρ is law invariant, we obtain that $\rho(\hat{Y}) = \rho(-\hat{F}^{-1}(U)) \leq \rho(-\hat{F}^{-1}(U)) = \rho(Y) \leq \rho_0$ and therefore $\hat{Y} \in \mathcal{H}_2(\{\xi \leq c\})$. On the other hand, $\mathbb{E}[\xi\hat{Y}] = \mathbb{E}[\xi Y]$ (this is due to our choice of the constant k). Since the choice of Y was arbitrary, this means that $\triangle(A) \geq \triangle(c)$.

Step 2. Let X be feasible for $\mathcal{P}_1$ with parameter $(A, x_+(A))$, and define

1. $f_2(t) := \mathbb{P}(X \leq t \mid A_2)$
2. $g_2(t) := \mathbb{P}(\xi \leq t \mid B_1)$
3. $Z_2 = g_2(\xi)$
4. $W_2 = f_2^{-1}(Z_2)$, that is, the law of W_2 on B_1 is the same as the law of X on A_2.

Let

$$k_2 := \begin{cases} 1 & \text{if } W_2 = 0 \text{ on } B_1 \\ \frac{\mathbb{E}[\xi X \mathbb{1}_{A_2}]}{\mathbb{E}[\xi W_2 \mathbb{1}_{B_1}]} & \text{otherwise} \end{cases}$$

Note that now, $k_2 \geq 1$. We define a new random variable $\hat{X}$ by

$$\hat{X} := X\mathbf{1}_{A_1} + k_2 W_2 \mathbf{1}_{B_1} + \frac{x_+(c) - x^+(A)}{\mathbb{E}\left[\xi \mathbb{1}_{\{\xi \leq c\}}\right]} \mathbb{1}_{\{\xi \leq c\}}$$

Since $\mathbb{E}\left[\xi\hat{X}\right] = x^+(c)$ we deduce $\hat{X} \in \mathcal{H}_1(\{\xi \leq c\}, x^+(c))$. Moreover, since $k_2 \geq 1$, similar computations as before yield $\mathbb{P}\left(\hat{X} > t\right) \geq \mathbb{P}(X > t)$. By definition

$$U(X) = \mathbb{E}\left[u\left(X^+\right)\right] = \int_0^{+\infty} \mathbb{P}\left(X^+ > u^{-1}(t)\right) dt$$

and since $u^{-1}(t)$ is positive,

$$\left\{X^+ > u^{-1}(t)\right\} = \left\{X > u^{-1}(t)\right\}$$

we conclude that $U(X) \leq U\left(\hat{X}\right)$.

□

Theorem 4.10 allows us to simplify Algorithm 4.7:

Algorithm 4.11.

1. *fix* $c \in \left[\underline{\xi}, \overline{\xi}\right]$ *and consider* $A = \{\xi \leq c\}$
2. *solve* $\mathcal{P}_2$ *with parameter* $(\{\xi \leq c\})$ *and find* $\triangle(c) := \triangle(\{\xi \leq c\})$
3. *solve* $\mathcal{P}_1$ *with parameters* $(\{\xi \leq c\}, x^+(c))$, $x^+(c) := x_0 - \triangle(c)$
4. *find* c^*, *if it exists, that maximizes* $c \to v(c)$

The question of the existence of c^* which maximizes $c \to v(c)$, and the related question of the existence of the optimal pay-off for the fund manager is difficult to answer for general risk measures. A complete answer to this question will be given in Section 4.4 in the case of the entropic risk measure (see Theorem 4.12) and in Section 4.5 for spectral risk measures (Theorem 4.16).

4.3 The fee

The question of the role and the amount of the fee f, which the investor pays to the bank to enter the portfolio insurance scheme, is closely related to the more general issue of the economic rationality of the three-party structure bank—fund—investor. From the point of view of the bank, providing the guarantee is equivalent to providing a put option written on the optimal contingent claim with pay-off $(-X^*)^+$ to the investor (we still suppose $z = 0$), in exchange of the initial fee f. This transfer of risk from the investor towards the bank makes sense because the bank, as a large financial institution, can accommodate greater losses than the other two parties. It is less risk averse than the fund or the investor, and may even be risk seeking. While the fee leads to an immediate return for the bank, the tail risk associated to the put option may be reduced by diversification, or it may simply be kept on the balance sheet as an unhedgeable risk. These and similar considerations can induce a bank to provide the portfolio guarantee for a fee which is less than the replication price of the put option. On the other hand, the fee cannot be greater than the replication price, since in that case it would be optimal for the fund manager to replicate the guarantee himself, which is not what is observed in reality.

This leads us to a new formulation where the fee is assumed to be a percentage of the no-arbitrage price of the put option:

$$\text{maximize} \quad \mathbb{E}[u((V_T)^+)] \tag{4.27}$$

over v_T subject to

$$\rho(-(V_T)^-) \leq \rho_0 \text{ and } V_0 = x_0 - p\mathbb{E}^{\mathbb{Q}}\left[(-V_T)^+\right]. \tag{4.28}$$

for $p \in [0, 1]$. The case $p = 1$ corresponds to the situation when the fee is equal to the price of the put option: since the bank can completely hedge away her risk, we deduce that the value function does not depend on ρ_0, and therefore on the amount of risk which is being transferred to the bank, so one can take from the beginning $\rho_0 = 0$ (no risk transfer). In this case, the problem can be transformed into

$$\textbf{maximize}\,\mathbb{E}\left[u(V_T)\right] \text{ over } V_T \geq 0 \text{ and } V_0 = x_0$$

If the bank asks a fee which completely covers the price of the option she sells, then the fund manager has no reason to take any risk.

From now on let $0 \leq p < 1$. Assume that problem (4.2) has a solution, given by (4.19). Since we know that the optimal solution is of the form

$$V(c) := X(c)\mathbf{1}_{\xi \leq c} + Y(c)\mathbf{1}_{\xi > c}$$

for $c \in \left[\underline{\xi}, \overline{\xi}\right]$, the price of a put option written on $V(c)$ is $-\triangle(c)$. Hence the problem (4.27)–(4.28) can be transformed into

$$\begin{aligned} &\text{maximize} \quad \mathbb{E}[u(I(\lambda_c\xi))\mathbb{1}_{\xi \leq c}] && (4.29)\\ &\text{over } c \in \left[\underline{\xi}, \overline{\xi}\right] \text{ where } \lambda_c \text{ satisfies} \\ &\mathbb{E}^{\mathbb{Q}}\left[I(\lambda_c\xi)\mathbb{1}_{\xi \leq c}\right] = x_0 - (1-p)\triangle(c) && (4.30) \end{aligned}$$

If there exists a c^* which maximizes the above expression and $\mathbb{P}(\xi > c^*) > 0$ then the solution of this problem is given by

$$V_T^* := I(\lambda(c^*)\xi)\mathbf{1}_{\xi \leq c^*} + Y(c^*)\mathbf{1}_{\xi > c^*}$$

where $Y(c^*)$ is the solution of problem $\mathcal{P}_2$ with parameter $\{\xi \leq c^*\}$. The effect of the fee on the optimal portfolio will be further analyzed in the specific example of the entropic risk measure in Section 4.4.

4.4 Explicit result: the Entropic risk measure

4.4.1 The result

In this section we show how Theorems 4.8 and 4.10 can be used to solve problem (4.27)–(4.28) when one uses the entropic risk measure defined in (3.13):

$$\rho_\beta\left(X\right) := \beta \log \mathbb{E}\left[\exp\left(-\frac{1}{\beta}X\right)\right]$$

where $\beta > 0$. From (3.14):

$$\rho_\beta\left(X\right) = \sup_{\mathbb{Q} \ll \mathbb{P},\, \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) \in L^1(\mathbb{Q})} \left(\mathbb{E}^{\mathbb{Q}}\left[-X\right] - \beta \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right).$$

In particular, $\gamma_{min}\left(\xi\mathbb{P}\right) = \beta\mathbb{E}\left[\xi \log\left(\xi\right)\right]$.

Theorem 4.12. *Let ρ_β to be the entropic risk measure and assume that the state price density ξ has no atom and satisfies $\xi \log \xi \in \mathbb{L}^1(P)$. Assume also that the fee is a fraction $p \in [0,1]$ of the price of the put option corresponding to the guarantee, as in Section 4.3. Then the optimal claim for the fund manager is given by*

$$V^* := I(\lambda(c^*)\xi)\mathbb{1}_{\{\xi \le c^*\}} - \beta\left[\log\left(\frac{\beta}{\eta(c^*)}\xi\right)\right]^+ \mathbb{1}_{\{\xi > c^*\}}$$

where

- $\lambda(c)$ *is the unique solution of* $\mathbb{E}\left[\xi I(\lambda(c)\xi)\mathbb{1}_{\{\xi \le c\}}\right] = x_0 - (1-p)\triangle(c)$
- $\alpha(c) := \mathbb{P}(\xi > c)$
- $\triangle(c) = -\beta\mathbb{E}\left[\xi \log\left(\frac{\beta\xi}{\eta(c)} \vee 1\right)\right]$
- $\eta(c)$ *is the unique solution of:* $\mathbb{E}\left[\left(\frac{\beta\xi}{\eta(c)} \vee 1\right)\mathbb{1}_{\xi > c}\right] = e^{\frac{\rho_0}{\beta}} + \alpha(c) - 1.$
- c^* *attains the supremum of* $c \to \mathbb{E}\left[u(I(\lambda(c)\xi))\mathbb{1}_{\{\xi \le c\}}\right]$

Proof.
Remark first that the condition $\xi \log \xi \in \mathbb{L}^1(P)$ implies that Assumptions (4.16) holds true (Proposition 4.6). The proof is just a simple application of Theorems 4.8, 4.10, Lemma 4.3 and Lagrangian methods.

We first need to compute the map $c \to \triangle(c)$. Fix c and consider the problem:

$$\textbf{minimize}\,\mathbb{E}[\xi Y] \text{ over } \rho(Y) \le \rho_0,\ Y = 0 \text{ on } A \text{ and } Y \le 0 \text{ on } A^c$$

where $A = \{\xi \le c\}$. Working on the new space $\left(A^c, \hat{\mathcal{F}} := \{B \cap A^c, B \in \mathcal{F}\}, \hat{\mathbb{P}} := \mathbb{P}(\cdot|A^c)\right)$, we can transform this minimization into

$$\textbf{minimize} \quad \alpha(c)\hat{\mathbb{E}}[\xi W] \text{ over } \hat{\mathbb{E}}\left[\exp\left(-\frac{W}{\beta}\right)\right] \le \delta(c),\ W \le 0$$

$$\delta(c) = \frac{e^{\frac{\rho_0}{\beta}} + \alpha(c) - 1}{\alpha(c)}$$

where $\hat{\mathbb{E}}$ is the expectation under the new probability $\hat{\mathbb{P}}$ and W is a random variable on the new space. Using Lagrangian methods we can find the unique optimal solution:

$$W^*(c) := -\beta\left[\log\left(\frac{\beta}{\eta(c)}\xi\right)\right]^+$$

where $\eta(c)$ is the unique solution of:

$$\mathbb{E}\left[\left(\frac{\beta\xi}{\eta(c)} \vee 1\right)\mathbb{1}_{\xi > c}\right] = e^{\frac{\rho_0}{\beta}} + \alpha(c) - 1$$

so then

$$Y^*(c) := W^*(c)\mathbb{1}_{\{\xi > c\}}.$$

A simple calculation then gives:

$$\triangle(c) = -\beta\mathbb{E}\left[\xi\log\left(\frac{\beta\xi}{\eta(c)}\vee 1\right)\right]$$

If now we set $x_+(c,p) := x_0 - (1-p)\triangle(c)$, by Lemma 4.3, Problem $\mathcal{P}_1$ with parameters $(\{\xi \le c\}, x^+(c,p))$ can be easily solved and its unique solution is given by

$$X(c,p) = I(\lambda(c,p)\xi)\mathbb{1}_{\{\xi\le c\}}$$

where, by (4.11),

$$\mathbb{E}\left[\xi I(\lambda(c,p)\xi)\mathbb{1}_{\{\xi\le c\}}\right] = x^+(c,p).$$

By using Theorem 4.10 we find that the optimal c^* is the maximizer of the function

$$c \to \mathbb{E}\left[u(I(\lambda(c)\xi))\mathbb{1}_{\{\xi\le c\}}\right].$$

□

4.4.2 Numerical example

We will apply Theorem 4.12 in a simple case. Let the market be composed of one risky asset, S, which follows the Black and Scholes dynamics:

$$dS_t = S_t(bdt + \sigma dW_t)\ S_0 > 0$$

Suppose $\mu = b/\sigma > 0$. The unique equivalent martingale measure is given by $\mathbb{Q} = \xi\mathbb{P}$ where

$$\xi = \exp(-\mu W_T - \mu^2T/2) = \left[S_T\exp\left(T\left(\sigma^2 - b\right)/2\right)/S_0\right]^{-\frac{b}{\sigma^2}}$$

We will use the exponential utility function $u(x) = 1 - e^{-\delta x}$. For this example we take $b = 0.15$, $\sigma = 0.4$, $\mu = 0.375$, $T = 1$, $S_0 = 5$, $v_0 = 3.5$, $\rho_0 = 1.5$, $\beta = 1$, and $\delta = 0.6$.

The optimal pay-off is a spread of two options on the log contract $\log(S_T)$: one option is sold to match the desired risk tolerance and the second one is bought to obtain the gain profile desired by the investor.

$$X^* := \left[\frac{b}{\delta\sigma^2}\log(S_T) + K_1\right]^+\mathbf{1}_{\{S_T\ge s^*\}} - \beta\left[K_2 - \frac{b}{\sigma^2}\log(S_T)\right]^+\mathbf{1}_{\{S_T<s^*\}}$$

where

$$\begin{aligned}
s^* &= S_0\exp\left(T\left(b-\sigma^2\right)/2\right)(c^*)^{-\frac{-\sigma^2}{b}}\\
K_1 &= \frac{1}{\delta}\left(\frac{b\left(\sigma^2-b\right)}{2\sigma^2}T - \frac{b}{\sigma^2}\log(S_0) - \log\left(\frac{\lambda(c^*)}{\delta}\right)\right)\\
K_2 &= \frac{b}{\sigma^2}\log(S_0) - \frac{b\left(\sigma^2-b\right)}{2\sigma^2}T + \log\left(\frac{\beta}{\eta(c^*)}\right)
\end{aligned}$$

p	s^*	K_1	K_2	c^*	$\triangle(c^*)$	$\mathbb{P}(S_T \geq s^*)$	$v(c^*)$
0	2.340	1.458	5.863	2.028	-0.234	0.981	0.892
0.25	2.158	1.306	6.307	2.188	-0.164	0.989	0.890
0.5	1.491	1.139	8.798	3.095	-0.0202	0.9993	0.887
1	0.491	1.111				1	0.886

Table 4.2: Numerical results for different values of p.

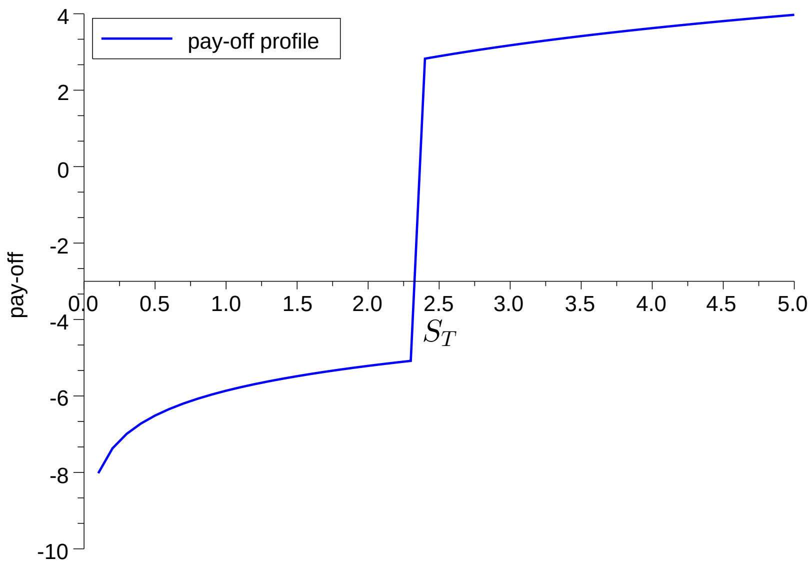

Figure 4.2: Optimal pay-off of the fund manager as function of the stock price value S_T for $p = 0$.

In the case $p = 1$ we know that the value function does not depend on ρ_0 (Section 4.3), and the problem becomes

$$\textbf{maximize} \quad \mathbb{E}[1 - e^{\delta X^+}] \text{ over } X \geq 0 \text{ an } \mathbb{E}^{\mathbb{Q}}[X] = x_0$$

whose solution is given by

$$X^* = \left[\frac{b}{\delta\sigma^2}\log(S_T) + K_1\right]^+ \mathbb{1}_{\{S_T \geq s^*\}}$$

The numerical values of various quantities of interest for different values of p are given in Table 4.2. The optimal pay-off of the fund manager as function of S_T is shown in Figure 4.2. Figure 4.3 shows the value function as function of c. Figure 4.4 shows the gain for the investor compared to the situation where no risk is allowed.

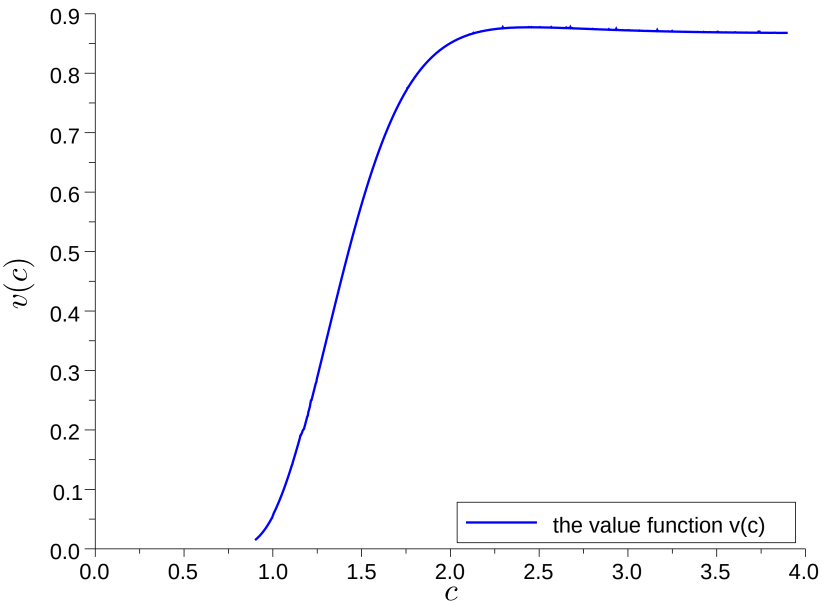

Figure 4.3: Value function of Problem $\mathcal{P}_1$ as function of c for $p = 0$.

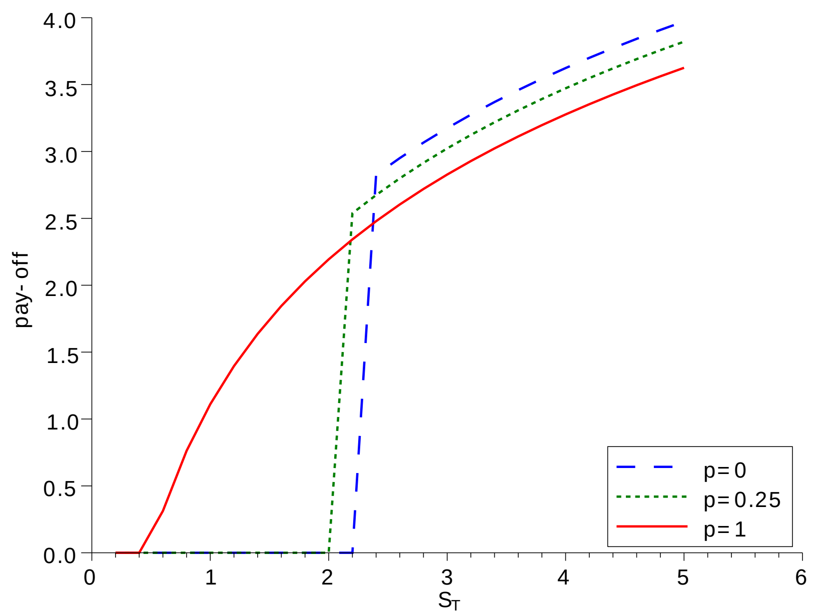

Figure 4.4: The pay-off profile for the investor for different values of p.

4.5 Explicit result: the spectral risk measure

4.5.1 The result

In this section we solve the portfolio optimization problem (4.27)–(4.28) when the risk constraint is given by a spectral risk measure. For a given probability on $[0,1]$ the related spectral risk measure is given by (3.7)–(3.8) :

$$\rho_\mu(X) := \int_0^1 CVaR_\beta(X)\,\mu(d\beta) \qquad \text{or equivalently}$$
$$\rho_\mu(X) := \int_0^1 \mu(u) VaR_u(X)\,du \qquad \text{where} \qquad \mu(u) := \int_u^1 \frac{\mu(dx)}{x}$$

Following Algorithms 4.7–4.11 we first need to compute the mappings $A \to \triangle(A)$ and $c \to \triangle(c)$:

Lemma 4.13. *For $A \in \mathcal{F}$ with $\mathbb{P}(A) < 1$, let $\hat{F}_\xi$ be the conditional distribution of ξ on A^c and define $\alpha_A := \mathbb{P}(A^c)$. $\triangle(A) > -\infty$ if and only if*

$$\lim_{x\to 0^+} \frac{\hat{F}_\xi^{-1}(1-x)}{\mu(x)} < +\infty \tag{4.31}$$

In this case

$$\triangle(A) = -\rho_0 \max_{x\in[0,1]} r(x) \tag{4.32}$$
$$r(x) := \frac{\alpha_A}{\int_0^{\alpha_A x} \mu(u)\,du} \int_0^x \hat{F}_\xi^{-1}(1-u)\,du \tag{4.33}$$

Proof.
In order to compute $\triangle(A)$ we reformulate Problem $\mathcal{P}_2$ in terms of the conditional distribution function of $Y \in \mathcal{H}_2(A)$ on A^c. Introduce a new probability $\hat{\mathbb{P}}$ via $\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \frac{\mathbf{1}_{A^c}}{\alpha_A}$. Let $\hat{F}_Y$ be the distribution function of Y under this probability and $\hat{F}_Y^{-1}$ its generalized inverse . Using this new probability we can rewrite the ingredients of our problem as

$$\mathbb{E}[\xi Y] = \alpha_A \hat{\mathbb{E}}[\xi Y]$$

and

$$CVaR_\beta(Y) = -\frac{1}{\beta}\int_0^\beta F_Y^{-1}(u) = -\frac{1}{\beta}\int_0^{\beta\wedge\alpha_A} \hat{F}_Y^{-1}(u/\alpha_A)du$$
$$= -\frac{\alpha_A}{\beta}\int_0^{\frac{\beta}{\alpha_A}\wedge 1} \hat{F}_Y^{-1}(u)du.$$

Fubini's theorem gives

$$\begin{aligned}\rho_\mu(Y) &= -\alpha_A \int_0^1\int_0^1 \mathbb{1}_{\{0\le\alpha_A u\le\beta\wedge\alpha_A\}} \hat{F}_Y^{-1}(u)\frac{\mu(d\beta)}{\beta}du \\ &= -\alpha_A \int_0^1\int_{\alpha_A u}^1 \frac{\mu(d\beta)}{\beta}\hat{F}_Y^{-1}(u)du \\ &= -\alpha_A \int_0^1 \mu(\alpha_A u)\hat{F}_Y^{-1}(u)du\end{aligned}$$

To express $\hat{\mathbb{E}}[\xi Y]$, we make use of the following Lemma:

Lemma 4.14. *Let F_1 and F_2 be distribution functions on $[0, \infty)$. Then*

$$\sup_{X \sim F_1,\ Y \sim F_2} \mathbb{E}[XY] = \int_0^1 F_1^{-1}(u) F_2^{-1}(u) du.$$

The proof of this result is given in Paragraph 4.5.2.

We can transform problem $\mathcal{P}_2$ as

$$\triangle(A) = \alpha_A \inf \int_0^1 \hat{F}_\xi^{-1}(u) \hat{F}_Y^{-1}(1-u) du \tag{4.34}$$

$$\text{subject to} \quad -\alpha_A \int_0^1 \mu(\alpha_A u) \hat{F}_Y^{-1}(u) du \leq \rho_0, \tag{4.35}$$

where the inf is taken over all generalized inverse distribution functions $\hat{F}_Y^{-1}$ of non-positive random variables. Such a function can always be written as

$$\hat{F}_Y^{-1}(u) := -\int_u^1 \zeta(du), \tag{4.36}$$

where ζ is a positive measure on $[0, 1]$. By using Fubini's theorem we can rewrite problem (4.34)–(4.35) in terms of this measure:

$$\triangle(A) = -\alpha_A \sup \left(\int_0^1 \zeta(ds) \int_0^s \hat{F}_\xi^{-1}(1-u) du \right)$$

$$\text{subject to} \quad \alpha_A \left(\int_0^1 \zeta(ds) \int_0^s \mu(\alpha_A u) du \right) \leq \rho_0.$$

The solution of this problem can easily be shown to be a point mass: $\zeta = h\delta_x$ where $h \geq 0$ and $x \in [0, 1]$ can be found from

$$\triangle(A) = -\alpha_A \sup \left(h \int_0^x \hat{F}_\xi^{-1}(1-u) du \right) \tag{4.37}$$

$$\text{subject to} \quad \alpha_A h \int_0^x \mu(\alpha_A u) du = \rho_0, \tag{4.38}$$

The constraint (4.38) gives us

$$h = h(x) = \frac{\rho_0}{\alpha_A \int_0^x \mu(\alpha_A s) ds}$$

and using definition (4.33) we get

$$\triangle(A) = -\alpha_A \sup_{x \in [0,1]} \left(\frac{\rho_0}{\alpha_A \int_0^x \mu(\alpha_A s) ds} \int_0^x \hat{F}_\xi^{-1}(1-u) du \right) = -\rho_0 \max_{x \in [0,1]} r(x)$$

The function r is differentiable on $(0, 1]$ and may only have a singularity at $x = 0$; using l'Hôpital's rule, we get

$$r(0^+) = \lim_{x \to 0} \frac{\hat{F}_\xi^{-1}(1-x)}{\mu(x)}$$

So $\triangle(A) > -\infty$ if and only if r is bounded on $[0, 1]$, which is true if and only if $r(0^+) < +\infty$.

□

Corollary 4.15. *The function* $\triangle(c)$ *is given by*

$$\triangle(c) = -\rho_0 \max_{0 \le z \le \alpha(c)} R(z), \quad R(z) := \frac{\mathbb{E}\left[\xi \mathbb{1}_{\{1-F_\xi(\xi)<z\}}\right]}{\int_0^z \mu(u)\, du}$$

Assume that the limit

$$\lim_{x \to 0^+} \frac{F_\xi^{-1}(1-x)}{\mu(x)} \tag{4.39}$$

exists. Then

$$\lim_{c \uparrow \bar{\xi}} \triangle(c) = -\rho_0 \lim_{x \to 0^+} \frac{F_\xi^{-1}(1-x)}{\mu(x)}.$$

Proof.
In order to make the dependence on c explicit, we introduce the notation

$$\triangle(c) := -\rho_0 \max_{x \in [0,1]} R(x,c)$$

where

$$R(x,c) := \frac{\alpha(c) \int_0^x \hat{F}_\xi^{-1}(1-u)\, du}{\int_0^{\alpha(c)x} \mu(u)\, du}$$

which holds true from Lemma 4.13. Noting that $\hat{F}_\xi^{-1}(1-u) = F_\xi^{-1}(1-\alpha(c)u) \ge c$ and making a change of variable,

$$\begin{aligned} R(x,c) &= \frac{\mathbb{E}\left[\xi \mathbb{1}_{\{c<\xi\}} \mathbb{1}_{\{\hat{F}_\xi^{-1}(1-x)<\xi\}}\right]}{\int_0^{\alpha(c)x} \mu(u)\, du} = \frac{\mathbb{E}\left[\xi \mathbb{1}_{\{F_\xi^{-1}(1-\alpha(c)x)<\xi\}}\right]}{\int_0^{\alpha(c)x} \mu(u)\, du} \\ &= \frac{\mathbb{E}\left[\xi \mathbb{1}_{\{1-F_\xi(\xi)<\alpha(c)x\}}\right]}{\int_0^{\alpha(c)x} \mu(u)\, du} \end{aligned}$$

The function $\triangle(c)$ can then be rewritten as

$$\triangle(c) = -\rho_0 \max_{0 \le z \le \alpha(c)} R(z), \quad R(z) := \frac{\mathbb{E}\left[\xi \mathbb{1}_{\{1-F_\xi(\xi)<z\}}\right]}{\int_0^z \mu(u)\, du}$$

□

The following theorem, which is the main result of this section, characterizes the solution of the problem (4.4) when the risk constraint is given by a spectral risk measure via an one-dimensional optimization problem.

Theorem 4.16. *Assume that there exists* c^* *with* $\mathbb{P}[\xi > c^*] > 0$ *such that* $v(c^*) = \max_{\underline{\xi} \le c \le \bar{\xi}} v(c)$ *with*

$$v(c) = \mathbb{E}[u(I(\lambda(c)\xi))\mathbb{1}_{\xi \le c}],$$

where $\lambda(c)$ *is the solution of*

$$\mathbb{E}[\xi I(\lambda(c)\xi)\mathbb{1}_{\xi \le c}] = x_0 + \frac{\rho_0 \mathbb{E}[\xi \mathbb{1}_{\xi > c}]}{\int_0^{\mathbb{P}[\xi > c]} \mu(u)du}.$$

Then the solution to the problem (4.4) *is given by*

$$X^* = I(\lambda(c^*)\xi)\mathbb{1}_{\xi \le c^*} - \frac{\rho_0}{\int_0^{\mathbb{P}[\xi > c^*]} \mu(u)du} \mathbb{1}_{\xi > c^*}.$$

Proof.
From Theorem 4.10 we need to maximize the function $c \to v(c)$ over $c \in [\underline{\xi}, \bar{\xi}]$. Assume that $v(c)$ achieves its maximum at the point c^* such that $\triangle(c^*) = -\rho R(z)$ with $z < \alpha(c)$ and let $c' = \alpha^{-1}(z)$. Then $\triangle(c)$ is constant on the interval $[c, c']$, which means that $x^+(c) = x^+(c')$,

$$\mathcal{H}_1\left(\{\xi \le c\}, x^+(c)\right) \subset \mathcal{H}_1\left(\{\xi \le c'\}, x^+(c')\right)$$

and therefore $v(c) \le v(c')$. This argument shows that the solution of the optimization problem appearing in the right-hand side of (4.24) does not change if we replace the expression for $\triangle(c)$ given by Corollary 4.15 by

$$-\rho_0 R(\alpha(c)) = -\frac{\rho_0 \mathbb{E}[\xi \mathbf{1}_{\xi > c}]}{\int_0^{\mathbb{P}[\xi > c]} \mu(u)du}.$$

Applying Lemma 4.3 we then find

$$v(c) = \mathbb{E}[u(I(\lambda(c)\xi))\mathbb{1}_{\xi \le c}],$$

where

$$\mathbb{E}[\xi I(\lambda(c)\xi)\mathbb{1}_{\xi \le c}] = x_0 + \frac{\rho_0 \mathbb{E}[\xi \mathbb{1}_{\xi > c}]}{\int_0^{\mathbb{P}[\xi > c]} \mu(u)du}.$$

If there exists a c^* with $\mathbb{P}(\xi > c^*) > 0$ which maximizes the value function $c \to v(c)$ then the optimal contingent claim is given by

$$X^* = I(\lambda(c^*)\xi)\mathbb{1}_{\xi \le c^*} - \frac{\rho_0}{\int_0^{\mathbb{P}[\xi > c^*]} \mu(u)du} \mathbb{1}_{\xi > c^*}.$$

where

$$-\frac{\rho_0}{\int_0^{\mathbb{P}[\xi > c^*]} \mu(u)du} \mathbb{1}_{\xi > c^*}.$$

is the optimal solution of Problem $\mathcal{P}_2$ corresponding to $\{\xi \le c^*\}$, which can be deduced from the proof of Lemma 4.13.

□

Remark 4.17. *If* $\sup_{\underline{\xi}\leq c\leq\overline{\xi}} v(c)$ *is attained only by* $c^* = \overline{\xi}$ *and*

$$\lim_{c\to\overline{\xi}} \frac{\mathbb{E}[\xi\mathbb{1}_{\xi>c}]}{\int_0^{\mathbb{P}[\xi>c]}\mu(u)du} < \infty,$$

(the latter condition holds, in particular, if $\overline{\xi} < \infty$*), then* $\inf_{A\in\mathcal{F}}\Delta(A) > -\infty$ *but this infimum is not achieved: the extra gain from allowing a risk tolerance is bounded, but the optimal claim does not exist. Intuitively, claims which are "almost optimal" will lead to a very large loss occurring with a very small probability.*

If

$$\limsup_{c\to\overline{\xi}} \frac{\mathbb{E}[\xi\mathbb{1}_{\xi>c}]}{\int_0^{\mathbb{P}[\xi>c]}\mu(u)du} = \infty$$

then $\inf_{A\in\mathcal{F}}\triangle(A) = -\infty$*: the extra gain from allowing a risk tolerance is unbounded.*

4.5.2 Proof of Lemma 4.14

Proof.
It will be easier to work with survival functions $\bar{F}_1(x) = 1 - F_1(x)$ and $\bar{F}_2(x) = 1 - F_2(x)$ rather than distribution functions. Let $\bar{F}(x,y) = \mathbb{P}(X > x, Y > y)$ denote the 2-dimensional survival function of (X,Y). By Fubini's theorem and elementary bounds on distribution functions,

$$\mathbb{E}[XY] = \int_0^\infty\int_0^\infty \bar{F}(x,y)dxdy \leq \int_0^\infty\int_0^\infty \min(\bar{F}_1(x),\bar{F}_2(y))dxdy,$$

which means that the maximum of $\mathbb{E}[XY]$ is attained when the survival function of X and Y is equal to $\min(\bar{F}_1(x),\bar{F}_2(y))$. But it is straightforward to check that the survival function of the couple $(F_1^{-1}(U), F_2^{-1}(U))$, where U is uniform on $[0,1]$, has exactly this form.

□

4.5.3 The special case: $CVaR_\beta$

Definition 3.13 gives that the $CVaR_\beta$ is a special case of spectral risk measure, when one takes $\mu\,(du) = \delta_\beta\,(du)$, which yields $\mu_\beta\,(x) := \frac{1}{\beta}\mathbb{1}_{\{\beta>x\}}$. The condition (4.31) appearing in Lemma 4.13 becomes

$$\lim_{x\to0^+} \beta\hat{F}_\xi^{-1}(1-x) = \beta\overline{\xi}$$

Corollary 4.15, Lemma 4.3 and Theorems 4.10 and 4.16 enable us to give the solution of Problem (4.4):

- If $\bar{\xi} := \operatorname{essup}\xi < \infty$, then the value function of problem (4.4) is:

$$\sup_{X\in H} U\,(X) = \sup_{c\in[\underline{\xi},\overline{\xi}]} \mathbb{E}\left[u\,(I\,(\lambda\,(c)\,\xi))\,\mathbb{1}_{\{\xi\leq c\}}\right] \tag{4.40}$$

where $\lambda(c)$ is the unique solution of

$$\mathbb{E}\left[\xi I\left(\lambda\left(c\right)\xi\right)\mathbb{1}_{\{\xi\leq c\}}\right]=x_0+\rho_0\frac{\mathbb{E}[\xi\mathbb{1}_{\{\xi>c\}}]}{1\wedge\frac{\alpha(c)}{\beta}}$$

- If $\bar{\xi}=+\infty$ then there exists $A\in\mathcal{F}$ with $\triangle(A)=-\infty$.

The maximum in (4.40) is always attained for some $c^*\in[\underline{\xi},\,\overline{\xi}]$ because the value function is continuous and $[\underline{\xi},\overline{\xi}]$ is compact. If $c^*<\overline{\xi}$ then Theorem 4.16 applies and then we have a optimal solution for Problem (4.4). If the maximum is attained at $c^*=\overline{\xi}$, then, as in Remark 4.17, the optimal claim does not exist.

Remark 4.18. *From* (3.6)*, the minimal penalty function for the $CVAR_\beta$ is given by:*

$$\gamma_{min}\left(\mathbb{Q}\right):=\begin{cases}0 & \textit{if } \frac{d\mathbb{Q}}{d\mathbb{P}}\leq\frac{1}{\beta}, \quad \mathbb{P}\text{-}a.s\\ +\infty & \textit{otherwise}\end{cases}$$

If ξ is bounded but $\mathbb{P}\left(\xi>\frac{1}{\beta}\right)>0$ then $\gamma_{min}\left(\xi\mathbb{P}\right)=+\infty$ and we have an example of a situation where Assumption (4.16) *holds true but the stronger assumption* (4.18) *does not.*

4.6 Explicit result: the G-divergence

The goal of this section is to solve problem (4.4) when ρ is related to some G-divergence. Let then G be a convex, increasing and non constant function, with $G(0)=0$, $G(1)<+\infty$ and $G(x)/x\to+\infty$ when $x\to+\infty$. The risk measure related to G was introduced in Section 3.4:

$$\rho_G(X):=\sup_{Q\ll\mathbb{P},\,I_G(\mathbb{Q}|\,\mathbb{P})<+\infty}\left(\mathbb{E}^{\mathbb{Q}}\left[-X\right]-I_G(\mathbb{Q}\,|\,\mathbb{P})\right)$$

or equivalently, by (3.12):

$$\rho_G(X):=\inf_{t\in\mathbb{R}}\left(\mathbb{E}\left[G^*(t-X)\right]-t\right)$$

where $G^*(u):=\sup_{u>0}(ut-G(u))$. In order to solve problem (4.4) let us first compute the map $A\to\triangle(A)$. For this we introduce, for $t\in\mathbb{R}$

$$\triangle(A,t):=\inf_{Y\in\mathcal{H}_2(A,t)}\mathbb{E}\left[\xi Y\right]\ \text{where} \tag{4.41}$$

$$\mathcal{H}_2(A,t):=\left\{Y\in\mathbb{L}^1(\xi\mathbb{P})\,|\ \mathbb{E}\left[G^*(t+Y)\right]\leq\rho_0+t,\,Y=0\text{ on }A,\,Y\leq0\text{ on }A^c\right\}$$

Lemma 4.19. *Let $G:\mathbb{R}\to\mathbb{R}^+\cup\{+\infty\}$ to be a convex, increasing function with $G(0)=0$, $G(x)/x\to+\infty$ when $x+\infty$ and assume that there exists some $\varepsilon>0$ such that $G(1+\varepsilon)<+\infty$. Then*

$$\triangle(A):=\inf_{t\in\mathbb{R}}\triangle(A,t)$$

Proof.
From the definition of $\triangle(A)$ in (4.9) and $\mathcal{H}_2(A)$ in (4.7) we have

$$\mathcal{H}_2(A,t) \subseteq \mathcal{H}_2(A), \text{ for all } t \in \mathbb{R}$$

which means that $\triangle(A,t) \geq \triangle(A)$, so then $\inf_t \triangle(A,t) \geq \triangle(A)$. The equality holds true if we will show that for any $Y \in \mathcal{H}_2(A)$ there exists a $t \in \mathbb{R}$ such that $Y \in H_2(A,t)$. For sake of clarity, let us introduce $\psi(t) := \mathbb{E}[G^*(t+Y)] - t$. It is straightforward to prove that ψ is convex. Furthermore for some $\eta \in (0,1)$ we have

$$\begin{aligned}
\lim_{t\to+\infty} \frac{\psi(t)}{t} &\geq \mathbb{E}\left[\liminf_{t\to+\infty} \frac{G^*(t+Y)}{t}\right] - 1 \\
&\geq \mathbb{E}\left[\liminf_{t\to+\infty} \left((1+\varepsilon)\left(1+\frac{Y}{t}\right) - \frac{G(1+\varepsilon)}{t}\right]\right) - 1 \\
&\geq \varepsilon
\end{aligned}$$

$$\begin{aligned}
\lim_{t\to-\infty} \frac{\psi(t)}{t} &\leq \mathbb{E}\left[\limsup_{t\to-\infty} \frac{G^*(t+Y)}{t}\right] - 1 \\
&\leq \mathbb{E}\left[\limsup_{t\to-\infty} \left((1-\eta)\left(1+\frac{Y}{t}\right) - \frac{G(1-\eta)}{t}\right]\right) - 1 \\
&\leq -\eta
\end{aligned}$$

which proves that ψ is a coercive function: for any $Y \in \mathcal{H}_2(A)$ there exists a $t_Y \in \mathbb{R}$ such that

$$\inf_{t\in\mathbb{R}} (\mathbb{E}[G^*(t+Y)] - t) = \mathbb{E}[G^*(t_Y+Y)] - t_Y$$

so then $Y \in \mathcal{H}_2(A,t_Y)$ which concludes our proof.

$\square$

Remark how both the $CVaR$ and the Entropy satisfy the assumptions of Lemma 4.19.

To compute $\triangle(A,t)$ we can use Lagrangian arguments: on the new space

$$\left(A,\, \mathcal{F}_A := \{A \cap B |\, B \in \mathcal{F}\},\, \hat{\mathbb{P}} := \mathbb{P}(|A)\right)$$

we can transform the problem into

$$\begin{aligned}
\textbf{minimize} \qquad & \alpha_A \hat{\mathbb{E}}[\xi Y] \text{ over } Y \leq 0 \text{ and} \\
& \hat{\mathbb{E}}[G^*(t+Y)] \leq \frac{\rho_0 + t - G^*(t)(1-\alpha_A)}{\alpha_A}
\end{aligned} \tag{4.42}$$

where $\alpha_A := \mathbb{P}(A)$ and $\hat{\mathbb{E}}$ is the expectation under $\hat{\mathbb{P}}$. Once we know $\triangle(A,t)$ for all t we first apply Lemma 4.19 and then Algorithm 4.11 gives us a way to solve problem (4.4).

4.7 Portfolio Insurance: a short review

Portfolio insurance is a widely popular concept in financial industry, and there exists an extensive literature on this topic. When the guarantee constraint is imposed in an almost sure way, a common strategy is the option based portfolio insurance, which uses put options written on the underlying risky asset as protection. The optimality of OBPI for European and American capital guarantee is studied in El Karoui et al. (2005). The difficulty of finding a sufficiently long-dated option for use in OBPI has lead to the appearance of strategies which involve only the underlying risky asset, of which the most popular is the Constant Proportion Portfolio Insurance (CPPI), (Black and Perold, 1992), where the exposure to the risky asset is proportional to the difference between the value of the fund and the discounted value of the guaranteed payment. If the price path of the underlying risky asset admits jumps, the CPPI strategy no longer ensures that the fund value will be a.s. above the guaranteed level at maturity, unless the portfolio is completely deleveraged (Cont and Tankov, 2009), which usually imposes too strong a restriction on the potential gains. The current market practice is therefore to require that the portfolio stays above the guaranteed level with a sufficiently high probability, or, for example, that it remains above the guarantee for a certain set of stress scenarios, chosen from historical data coming from highly volatile periods. A more flexible approach, which can take into account not only the probability of loss but also the sizes of potential losses, is to impose a constraint on a risk measure of the shortfall. This has led to the development of literature on portfolio insurance and, more generally, portfolio optimization under probabilistic / risk measure constraints.

Emmer et al. (2001) study one-period portfolio optimization under Capital-at-Risk constraint (the Capital-at-Risk is defined as the difference between the mean value of the portfolio and its VaR). Still in the one-period setting, Rockafellar and Uryasev (2000) provide an algorithm for minimizing the CVaR of a portfolio under a return constraint. Basak and Shapiro (2001) solve the utility optimization problem under the VaR constraint and Boyle and Tian (2007) discuss continuous-time portfolio optimization under the constraint to outperform a given benchmark with a certain confidence level. Like us, these authors also face some issues related to the non-convexity of the optimization problem, although the non-convexity appears for a different reason (non-convexity of the constraint itself). Another stream of literature (Föllmer and Leukert, 1999; Bouchard et al., 2009) considers hedging problems when the hedging constraint is imposed with a certain confidence level rather than almost surely. The viscosity solution approach of Bouchard et al. (2009) was extended in (Bouchard et al., 2010) to stochastic control problems under target constraint (that is, for example, under the constraint to outperform a benchmark with a certain probability) but it does not seem possible to treat risk measure constraints in this setting. He and Zhou (2010) have recently introduced a general methodology for solving law-invariant portfolio optimization problems by reformulating them in terms of the quantile function of the terminal value of the portfolio. While such a reformulation is in principle possible for our problem by using the dual representation results for law-invariant convex risk measures (see Föllmer and Schied (2004) and Jouini et al. (2006)), the resulting problem is still non-linear and non-convex so such a transformation does not necessarily simplify the treatment. Gundel and Weber (2007) solve the problem of maximizing the robust utility of a portfolio under a constraint on the expected shortfall, which includes, in particular, all coherent risk measures. Rogers (2009) discusses utility optimization when a portfolio constraint in the form of a coherent risk measure is present, and goes on to study optimal contracting problems in this context. The main difference/novelty of our work from these two studies is that in our approach, the utility function is only applied to positive gains while the risk measure is only applied to negative shortfall. This brings us much closer to the reality of portfolio insurance and at the same time allows to obtain explicit solutions.

Part II

Quadratic Hedge

Chapter 5

Quadratic hedge: introduction and main properties

The Chapter is organized as follows: we first introduce the Quadratic Hedge problem in its most general formulation and give a review of the literature on the subject. In Section 5.2 we introduce the general model, the value function associated to the Quadratic Hedge problem, we define the so called Pure investment problem and we study the structure of these value functions (Sections 5.3 and 5.4). Next, with a verification argument, we characterize the value functions of the pure investment and the quadratic hedge problems as the unique solution of a semi linear partial integro-differential equations (Sections 5.5 and 5.6). We finally do a short digression on the theory of viscosity solutions and how it can be used in this context (Section 5.7).

Contents

5.1 Introduction 67
5.2 The Model 71
5.3 The pure investment problem: a priori estimate 74
5.3.1 Proof of Lemma 5.5 82
5.3.2 Proof of Lemma 5.6 83
5.3.3 Proof of Lemma 5.7 86
5.4 The structure of the quadratic hedge value function . . 89
5.5 The pure investment problem: verification 91
5.6 The quadratic hedge problem: verification 94
5.7 Viscosity solutions 96

5.1 Introduction

In incomplete markets the perfect hedge does not exist in general, and most of the markets are incomplete. In these cases, pricing and hedging an option is a hard task since one cannot totally hedge away the risk. Once we accept that a residual risk may affect our hedging strategy, an important issue, especially from a practitioner's point of view, is to quantify and control this residual risk. A common way to measure this residual risk is to compute the expected squared distance between the option one wants to hedge and the portfolio. The quadratic hedge problem is to find the optimal portfolio which minimizes this residual risk. In its most general form the quadratic hedge problem can be formulated as follows: consider a random variable $H \in \mathbb{L}^2(\mathcal{F}_T, \mathbb{P})$ and a set of admissible strategies, which has to be carefully specified,

$$\theta \in \mathcal{X} \subseteq \mathbb{L}(S) \tag{5.1}$$

where S is a semimartingale modeling the stock price. Up to appropriate integrability conditions, the quadratic hedge problem becomes

$$\text{minimize } \mathbb{E}^{\mathbb{P}}\left[\left(x + \int_0^T \theta_{t-} dS_t - H\right)^2\right] \text{ over } x \in \mathbb{R} \text{ and } \theta \in \mathcal{X} \tag{5.2}$$

If (x^*, θ^*) achieves this infimum, we call θ^* the optimal mean-variance hedging strategy and x^* its price. When one defines

$$\mathbb{A} := \left\{x + \int_0^T \theta_t dS_t \mid x \in \mathbb{R},\, \theta \in \mathcal{X}\right\}$$

then problem (5.2) can be viewed as the $\mathbb{L}^2$-projection of H on the space $\mathbb{A}$: one tries to minimize the $\mathbb{L}^2$-distance between the contingent claim H and a set of all admissible portfolios in $\mathbb{A}$.

The quadratic hedge problem is a particular case of the so called *utility-based* pricing and hedging problem: for an utility function $\mathcal{U}$ one tries to solve

$$\text{maximize } \mathbb{E}^{\mathbb{P}}\left[U\left(x + \int_0^T \theta_t dS_t - H\right)\right] \text{ over } x \in \mathbb{R} \text{ and } \theta \in \mathcal{X}$$

For a complete overview on the *utility-based indifference price* problem we refer to El Karoui and Rouge (2000); Schweizer (2001); Delbaen et al. (2002). Assume now that (x_H, θ_H) achieves the maximum in the above problem and consider the following map $\mathcal{P}$ from $\mathbb{L}^2$ to $\mathbb{R}$:

$$\mathcal{P} : \mathbb{L}^2(\mathcal{F}_T, \mathbb{P}) \ni H \to x_H \in \mathbb{R}$$

Generally this map is not linear, unless one takes $U(x) = -x^2$, which corresponds to the quadratic hedge problem. Although one cannot speak of pricing rule (the utility function has to be increasing and this is not the case), the fact that the above mapping $\mathcal{P}$ is linear has several advantages: from a practical point of view, when one wants to price an entire portfolio, let us say $H = \sum_i H_i$, according to (5.2), she can first compute the prices corresponding to the single positions H_i, and then add them up to obtain the portfolio's price.

Several methods have been proposed to solve the quadratic hedge problem, depending on the features of the semimartingale S or on the set of admissible strategies $\mathcal{X}$. An elegant solution is provided when S is a martingale under the historical probability $\mathbb{P}$, by using the *Galtchouk-Kunita-Watanabe* decomposition (Kunita and Watanabe, 1967; Galtchouk, 1976): since $H \in \mathbb{L}^2(\mathbb{P}, \mathcal{F}_T)$, one can find a predictable process $\theta^H \in \mathbb{L}(S)$ such that

$$H = \mathbb{E}^P[H] + \int_0^T \theta^H_{u-} dS_u + N^H_T \text{ a.s.}$$

where $N_t^H := \mathbb{E}^P \left[N_T^H \mid \mathcal{F}_T \right]$ is a square integrable martingale such that $N_t^H \int_0^t \theta_{u-} dS_u$ is still a martingale under $\mathbb{P}$, i.e. $(N_t^H)_t$ is strongly orthogonal to the set $\mathbb{A}$. By using this decomposition it is straightforward to deduce that the problem (5.2) is solved by $\theta^* = \theta^H$ and $x^* = \mathbb{E}^P[H]$. Even if the Galtchouk-Kunita-Watanabe decomposition gives the optimal hedging strategy and the minimal price in problem (5.2), some important questions naturally arise when one wants to compute it in practice: firstly, does θ^H belong to $\mathcal{X}$? and if it is the case, can it be thought as a trading strategy, i.e. is it caglad? In many important cases, the answer is positive, and it is also possible to compute the hedging strategy semi-explicitly, in particular when S is Lévy process (Cont and Tankov, 2004) or, more generally, if S is a general Markov jump martingale (Cont, Tankov, and Voltchkova, 2007).

When S fails to be a martingale under $\mathbb{P}$, problem (5.2) becomes much more difficult because the *Galtchouk-Kunita-Watanabe* decomposition is no longer available. To solve the quadratic hedge problem many authors have exploited the particular features of the model S which allow to find the solution of problem (5.2), but it is no longer explicit enough (which is an important question since this procedure is commonly used in practice). For example, when S is a continuous Itô process, the solution can be found by making an appropriate change of probability in problem (5.2), and then using the Galtchouk-Kunita-Watanabe decomposition to obtain the optimal hedging strategy. For a complete review on this procedure we refer to Laurent and Pham (1999) or Pham (2000).

The idea of looking for a suitable change of probability that makes S a martingale turns out to be the key tool to solve the quadratic hedge problem in a general setting: it can be proved that this suitable martingale measure can be obtained by solving problem (5.2) for $H = 0$ and $x = 1$:

$$\text{minimize } \mathbb{E}^{\mathbb{P}} \left[\left(1 + \int_0^T \theta_t dS_t \right)^2 \right], \text{ over } \theta \in \mathcal{X} \tag{5.3}$$

In the literature this problem is known as the *pure investment problem*. In a general setting (i.e. S discontinuous semimartingale) Černỳ and Kallsen (2007) give many interesting properties of problem (5.3) and provide its solution. They derive then the optimal strategy of problem (5.2) and, in particular, they find that it is given by:

$$\theta_t^* = \hat{\alpha}_{t-} X_{t-}^{\theta^*} + \hat{\beta}_{t,-}, \quad X_t^{\theta^*} = x + \int_0^t \theta_{r-}^* dS_r$$

for some semi-explicit processes $\hat{\alpha}, \hat{\beta}$, where $\hat{\alpha}$ does not depend on the particular option H and it is related to the optimal solution of problem (5.3), whereas $\hat{\beta}$ linearly depends on H. This procedure completely characterizes the optimal hedging strategy in problem (5.2) but, nevertheless, this solution is no longer explicit for all types of semimartingale S, unless the stock price S has a particular structure: with stationary and independent increment (Hubalek, Kallsen, and Krawczyk, 2006) or affine stochastic volatility models (Černỳ and Kallsen, 2008; Kallsen and Vierthauer, 2009).

Our contribution is to give a systematic way to solve problem (5.2) when S is a general Markov jump process but not necessarily a $\mathbb{P}$-martingale by using stochastic optimization tools (value function and Hamilton-Jacobi-Bellman equation). All over the work, we tried to move according to the following objectives:

- Considering a sufficiently general model which allows us to apply our results to some practical problems (which partially motivated our work)
- Solving problem (5.2) and characterizing the optimal hedging strategy in the more explicit way
- Proving regularity, in a sense which has to be specified, of the optimal hedging strategy

As we will explain in Chapter 8, portfolio management in electricity markets motivated our work. We think it is helpful to briefly anticipate the main lines of that Chapter in order to understand our model on S.
The electricity spot price is generally modeled by a Lévy-driven process whereas a typical hedging instrument available in this market (the semimartingale S) is the future contract, which turns out to be Markov jump process and it does not posses the martingale property under $\mathbb{P}$. Moreover the spot price process may affect the dynamics of the future contract price, and since the electricity cannot be stored, one has to consider it as a non hedgeable source of risk like, for example, a volatility factor. We concentrate on European options written on S which may also depend on the spot price at maturity and on some other non hedgeable source of risk like, for example, the temperature.

The price of our hedging instrument is denoted by $S = \exp(Z)$, U denotes the spot price process and P the temperature, where (Z, U, P) is a $\mathbb{R}^3$-valued Markov jump process. The quadratic hedge problem for practitioners of the electricity markets can be formulated as follows

$$\begin{aligned} &\text{minimize } \mathbb{E}\left[\left(H(U_T, P_T, e^{Z_T}) - x - \int_0^T \theta_{r-} de^{Z_r}\right)^2\right] \\ &\text{over } x \in \mathbb{R},\ \theta \in \mathcal{X} \end{aligned} \tag{5.4}$$

Problem (5.4) can be also viewed as a constrained quadratic hedge problem: if $\tilde{S} := (S, U, P)$ then the problem above can be rewritten as

$$\text{minimize } \mathbb{E}\left[\left(H(\tilde{S}_T) - x - \int_0^T \theta'_{r-} d\tilde{S}_r\right)^2\right] \text{over } x \in \mathbb{R} \text{ and } \theta \in \mathbb{L}(\tilde{S}),\ \theta^2 = \theta^3 = 0$$

In order to solve problem (5.4) with the classical instruments of the stochastic optimization, we will assume that the dynamics of the state variable are of Markov type with appropriate assumptions on their coefficients. This will allow us to write and solve the partial integro-differential equation (PIDE) associated to the above problem in a particular space of smooth functions. A particular attention will be devoted to the case $H = 0$ since, as we have seen, this is a fundamental tool to solve the quadratic hedge problem.
Still inspired by practical problems, we will consider essentially three different cases: when Z is a jump-diffusion process (Chapter 6) and when it is a pure jump process with infinite/finite variation (Chapter 7). The pure jump case is quite interesting

since, in commodities markets, upward and downward movements of the stock price process S (and then Z) are essentially due to jumps. To conclude, we want to point out that our goal is proving that the value function of problem (5.4) is smooth (so that the optimal strategy also is): this is important from a numerical point of view and also because it will give us a better understanding of the optimal strategy behavior (their derivatives), which is undoubtedly an important task in risk management.

5.2 The Model

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space. On this space we introduce two independent Brownian motions W and B, taking values respectively in $\mathbb{R}^2$ and $\mathbb{R}$, and two independent Poisson random measures J and N on $\mathbb{R}\setminus\{0\}$, both independent from W and B. We assume that $\mathcal{F}_t$ is the natural filtration of J, N, W and B and that $\mathcal{F}_0$ is augmented with the null sets. We also assume $\mathcal{F} = \mathcal{F}_T$ where $T > 0$ is given. Furthermore $\left[W^1, W^2\right]_t = \lambda t$ for some $\lambda \in (-1, 1)$. The positive Lévy measure on $\mathbb{R}\setminus\{0\}$ related to J is denoted by $\nu(dy)$ and it satisfies the standard integrability condition $\int_{\mathbb{R}} \left(1 \wedge |y|^2\right) \nu(dy) < \infty$. The same holds true for the Lévy measure $\nu_n(dy)$ associated to N. We denote

$$\tilde{J}(dydt) = J(dydt) - \nu(dy)dt \qquad \tilde{N}(dydt) = N(dydt) - \nu_n(dy)dt$$

the compensated Poisson measures on $\mathbb{R}$ and

$$\bar{J}(dydt) := J(dydt) - dt \times \nu(dy)\mathbb{1}_{\{|y|\le 1\}}, \quad \bar{N}(dydt) := N(dydt) - dt \times \nu_n(dy)\mathbb{1}_{\{|y|\le 1\}}$$

On this probability space we introduce the family of $\mathbb{R}^3$-valued Markov jump processes (Z, U, P) as follows:

$$\begin{aligned}
dZ_r^{t,u,z} &:= \mu\left(r, U_r^{t,u}, Z_r^{t,u,z}\right) dr + \sigma(r, U_r^{t,u}, Z_r^{t,u,z}) dW_r^1 + \int_{\mathbb{R}} \gamma\left(r, U_{r-}^{t,u}, Z_{r-}^{t,u,z}, y\right) \bar{J}(dydr) \\
dU_r^{t,u} &:= \mu^U(r, U_r^{t,u}) dr + \sigma^U(r, U_r^{t,u}) dB_r + \int_{\mathbb{R}} \gamma^U\left(r, U_{r-}^{t,u}, y\right) \bar{N}(dydr) \\
dP_r^{t,p} &:= \mu^P(r, P_r^{t,p}) dr + \sigma^P(r, P_r^{t,p}) dW_r^2 + \int_{\mathbb{R}} \gamma^P(r, P_r^{t,p}, y) \tilde{J}(dydr)
\end{aligned} \tag{5.5}$$

with initial conditions $Z_t^{t,u,z} = z$, $U_t^{t,u} = u$ and $P_t^{t,p} = p$, for $t \in [0, T)$ and $z, u, p \in \mathbb{R}$. The stock price process S is given by $S = \exp(Z)$.

We make the following assumptions:

Assumption 5.1.
*[C]- **The coefficients-1**.*

i). There exists $\overline{\mu} \ge 0$ *such that* $\max\left(\|\mu\|_\infty, \left\|\mu^U\right\|_\infty, \left\|\mu^P\right\|_\infty\right) \le \overline{\mu}$.

ii). The volatility functions $\sigma, \sigma^U, \sigma^P$ *take values in* $[\sigma_{min}, \sigma_{max}]$, *for some* $0 \le \sigma_{min} \le \sigma_{max}$.

iii). For all $t \in [0, T]$ *and* $u, y \in \mathbb{R}$ *the functions* $z \to \mu(t, u, z)$, $z \to \sigma(t, u, z)$ *and* $z \to \gamma(t, u, z, y)$ *belong to* $C^1(\mathbb{R})$.

iv). There exist two constants $K^c_{lip} \geq 0$, $K^d_{lip} \geq 0$ and a positive locally bounded function $\rho : \mathbb{R} \to \mathbb{R}^+$ such that for all $y \in \mathbb{R}$ we have

$$\begin{aligned}
\left|\mu(t,u,z) - \mu(t,u',z')\right| &\leq K^c_{lip}\left(|z-z'| + |u-u'|\right)\\
\left|\sigma(t,u,z) - \sigma(t,u',z')\right| &\leq K^c_{lip}\left(|z-z'| + |u-u'|\right)\\
\left|\gamma(t,u,z,y) - \gamma(t,u',z',y)\right| &\leq K^d_{lip}\rho(y)\left(|z-z'| + |u-u'|\right)
\end{aligned}$$

for all $t \in [0,T]$, $z, z', u, u' \in \mathbb{R}$.

v). Property iv) holds true for μ^U, σ^U and γ^U (resp. μ^P, σ^P and γ^P) for some positive constant $K^U \geq 0$ and some positive locally bounded function ρ^U (resp. some $K^P \geq 0$ and some positive locally bounded function ρ^P)

[I1]- ***Integrability conditions-1****. The functions*

$$\begin{aligned}
\tau(y) &:= \max\left(\sup_{t,u,z}\left(|\gamma(t,u,z,y)|, \left|e^{\gamma(t,u,z,y)} - 1\right|\right), \rho(y)\right)\\
\tau^U(y) &:= \max\left(\sup_{t,u}|\gamma^U(t,u,y)|, \rho^U(y)\right)\\
\tau^P(y) &:= \max\left(\sup_{t,p}|\gamma^P(t,p,y)|, \rho^P(y)\right)
\end{aligned}$$

verify $\tau, \tau^P \in \mathbb{L}^2(\mathbb{R}, \nu(dy))$ and $\tau^U \in \mathbb{L}^2(\mathbb{R}, \nu_n(dy))$.

[I2]- ***Integrability conditions-2****. The function τ verifies $\tau \in \mathbb{L}^4(\{|y[\geq 1\}, \nu(dy))$.*

We define $K_{max} := \max(K^c_{lip}, K^d_{lip})$,

$$\tilde{\mu} := \mu + \frac{1}{2}\sigma^2 + \int_{|y|<1}(e^\gamma - 1 - \gamma)\nu(dy) \ \text{ and } \ \|\tilde{\mu}\| := \sup_{t,u,z}|\tilde{\mu}(t,u,z)| \tag{5.6}$$

and

$$|\Gamma| := \int_{\mathbb{R}} \Gamma(y)\nu(dy), \qquad \text{where} \qquad \Gamma(y) := \inf_{t,u,z}\left(e^{\gamma(t,u,z,y)} - 1\right)^2 \tag{5.7}$$

In the rest of the chapter we denote $\|\tau\|_{1,\nu} := \int_{|y|\geq 1}\tau(y)\nu(dy)$ whereas $\|\tau\|^2_{2,\nu} := \int_{\mathbb{R}}\tau^2(y)\nu(dy)$. The same convention holds for τ^P and τ^U (with respect to the Lévy measure $\nu_n(dy)$).

It is well known that there exists a unique semimartingale (U, Z, P) which solves the SDE (5.5) (see for example Jacod and Shiryaev (2003) or Protter (2004)).

Let us introduce the set of admissible strategies θ in problem (5.4): as already pointed out in Černỳ and Kallsen (2007), this set has to be carefully chosen: if it is too wide we may violate the principle of no arbitrage; if it is too small we may not be able to find the optimal strategies. We follow Černý and Kallsen (2007) to define a good set of admissible strategies: first let us introduce the sets of simple caglad strategies:

$$\mathcal{D} := \left\{ \theta := Y_0 + \sum_i Y_i \mathbf{1}_{]\varsigma_i, \varsigma_{i+1}]},\ Y_i \in \mathbb{L}^\infty(\mathcal{F}_{\varsigma_i}) \text{ and } \varsigma_i \le \varsigma_{i+1} \text{ are stopping times} \right\}$$

$$\mathcal{D}_t := \left\{ \theta \mathbb{1}_{(t,T]} \mid \theta \in \mathcal{D} \right\} \tag{5.8}$$

$\bar{\mathcal{D}}_t$ denotes the $\mathbb{L}^2(\mathbb{P})$-closure of $\mathcal{D}_t$, and for $t, u, z, x \in [0, T) \times \mathbb{R}^3$, $\theta \in \bar{\mathcal{D}}_t$ we define the wealth process as

$$dX_r^{t,u,z,x,\theta} := \theta_{r-} dS_r^{t,u,z}, \quad X_t^{t,u,z,x,\theta} := x \tag{5.9}$$

Here θ represents the number of shares in the portfolio at time t. We say that a control θ is admissible if it is caglad and $X_r^{t,u,z,x} \in \mathbb{L}^2(\mathbb{P})$ for any r, t, u, z, x: the set of admissible strategies is then defined as

$$\mathcal{X}(t, u, p, z) := \left\{ \theta \in \bar{\mathcal{D}}_t \mid x + \int_t^s \theta_{r-} dS_r^{t,u,z} \in \mathbb{L}^2(\mathbb{P}), \text{ for all } t \le s \le T \right\} \tag{5.10}$$

Consider a European option of the form $f(U_T, P_T, Z_T)$ where f is, for the moment, a measurable function with $f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}) \in \mathbb{L}^2(\mathbb{P})$ for all $(t, u, p, z) \in [0, T) \times \mathbb{R}^3$ (according to problem (5.4) we have $f(., z) = H(., e^z)$).
The quadratic hedging problem can be formulated as follows:

$$\textbf{QH}: \qquad \text{minimize } \mathbb{E}^{\mathbb{P}} \left[\left(f\left(U_T^{0,u}, P_T^{0,p}, Z_T^{0,u,z} \right) - X_T^{0,u,z,x,\theta} \right)^2 \right]$$
$$\text{over } \theta \in \mathcal{X}(0, u, z, x),\ x \in \mathbb{R}$$

The value function of **QH** is given by

$$v^f(t, u, p, z, x) := \inf_{\theta \in \mathcal{X}(t,u,p,z)} \mathbb{E}^{\mathbb{P}} \left[\left(f\left(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z} \right) - X_T^{t,u,z,x,\theta} \right)^2 \right] \tag{5.11}$$
$$v^f(T, u, p, z, x) = (f(u, p, z) - x)^2$$

Deeply related to the solution of problem (5.11) is the so called *Pure investment problem*, which essentially is Problem (5.11) when $f = 0$:

$$\begin{aligned} v^0(t, u, z, x) :=& \inf_{\theta \in \mathcal{X}(t,u,z)} \mathbb{E}^{\mathbb{P}} \left[\left(x + \int_t^T \theta_{r-} dS_r^{t,u,z} \right)^2 \right] \\ =& x^2 \inf_{\theta \in \mathcal{X}(t,u,z)} \mathbb{E}^{\mathbb{P}} \left[\left(1 + \int_t^T \theta_{r-} dS_r^{t,u,z} \right)^2 \right] \\ =& x^2 a(t, u, z) \end{aligned} \tag{5.12}$$

where

$$a(t, u, z) = \inf_{\theta \in \mathcal{X}(t,u,z)} \mathbb{E} \left[\left(1 + \int_t^T \theta_{r-} dS_r^{t,u,z} \right)^2 \right] \tag{5.13}$$

because the set $\mathcal{X}(t, u, z)$ is a cone.

5.3 The pure investment problem: a priori estimate

Problem **QH** when $f = 0$ is known in the literature as the *the pure investment problem*. Several properties have been shown for this problem by using its dual formulation. We recall here the most important ones. For a complete review in a more general setting we recommend Černỳ and Kallsen (2007).

Lemma 5.2. *Let a be the function defined in* (5.13) *and* θ^* *the optimal strategy which achieves the infimum. Then*

i). $a(t,u,z) = \mathbb{E}^{\mathbb{P}}\left[\left(1+\int_t^T \theta^*_{r-} dS_r^{t,u,z}\right)^2\right] = \mathbb{E}^{\mathbb{P}}\left[1+\int_t^T \theta^*_{r-} dS_r^{t,u,z}\right]$ *and it is strictly positive.*

ii). Define the set

$$\mathcal{M}_{t,u,z} := \left\{ \text{signed } \mathbb{Q} \ll \mathbb{P} \;\middle|\; \begin{array}{l} \mathbb{Q}(\Omega) = 1 \text{ and } \left(Z_r^Q S_r^{t,u,z}\right)_{r\geq t} \text{ is a martingale} \\ \text{where } Z_r^{\mathbb{Q}} := \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle| \mathcal{F}_r\right] \end{array} \right\}$$

$\mathcal{M}_{t,u,z}$ *is called the set of all absolutely continuous signed* $\sigma-$*martingale measures. Then*

$$d\mathbb{Q}^* := \frac{1}{a(t,u,z)}\left(1+\int_t^T \theta^*_{r-} dS_r^{t,u,z}\right) d\mathbb{P}$$

belongs to $\mathcal{M}_{t,u,z}$.

iii). The function a verifies

$$1 = \inf_{\mathbb{Q}\in\mathcal{M}_{t,u,z}} \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right] a\left(t,u,z\right) \tag{5.14}$$

and the infimum above is achieved by $\mathbb{Q}^*$.

In the literature $\mathbb{Q}^*$ *is called* Variance-optimal signed martingale measure .

Proof.

i). It can be proved that there exists a unique strategy θ^* which achieves the infimum in (5.13) (Černỳ and Kallsen, 2007). It follows that for any $\eta \in \mathcal{X}(t,u,z)$ and $\epsilon \neq 0$

$$a(t,u,z) \leq \mathbb{E}^{\mathbb{P}}\left[\left(1+\int_t^T (\theta^*_{r-} + \epsilon\eta_{r-}) dS_r^{t,u,z}\right)^2\right]$$

which implies

$$0 \leq \epsilon^2 \mathbb{E}^{\mathbb{P}}\left[\left(\int_t^T \eta_{r-} dS_r^{t,u,z}\right)^2\right] + 2\epsilon\mathbb{E}^{\mathbb{P}}\left[\left(1+\int_t^T \theta^*_{r-} dS_r^{t,u,z}\right)\int_t^T \eta_{r-} dS_r^{t,u,z}\right]$$

Dividing by ϵ and taking the limit $\epsilon \to 0$ we obtain

$$\mathbb{E}^{\mathbb{P}}\left[\int_t^T \theta^*_{r-} dS_r^{t,u,z} \int_t^T \eta_{r-} dS_r^{t,u,z}\right] = -\mathbb{E}^{\mathbb{P}}\left[\int_t^T \eta_{r-} dS_r^{t,u,z}\right] \tag{5.15}$$

If we take $\eta = \theta^*$ then $\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^T \theta^*_{r-} dS^{t,u,z}_r\right)^2\right] = -\mathbb{E}^{\mathbb{P}}\left[\int_t^T \theta^*_{r-} dS^{t,u,z}_r\right]$, which, in particular, gives

$$a(t,u,z) = \mathbb{E}^{\mathbb{P}}\left[\left(1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right)^2\right] = \mathbb{E}^{\mathbb{P}}\left[1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right]$$

ii). By using Lemma 3.1 in Černý and Kallsen (2007), we also have $a(t,u,z) > 0$, so then the following signed measure $\mathbb{Q}^*$:

$$\left.\frac{d\mathbb{Q}^*}{d\mathbb{P}}\right|_{\mathcal{F}_s} := Y^*_s, \quad Y^*_s := \frac{1}{a(t,u,z)}\mathbb{E}\left[\left(1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right)\middle| \mathcal{F}_s\right]$$

is well defined. It follows that $\mathbb{Q}^* \ll \mathbb{P}$ in the sense of signed measures (Jordan-Hahn decomposition Theorem) and trivially $\mathbb{Q}^*(\Omega) = 1$. $\mathbb{Q}^*$ is an absolutely continuous $\sigma-$martingale measure if and only if Y^*S is a $\mathbb{P}$-martingale. This is equivalent to prove that for any stopping time ϱ taking values in $(t,T]$ one has $\mathbb{E}^{\mathbb{P}}\left[Y^*_\varrho S_\varrho\right] = Y_t S_t$:

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[Y_\varrho S_\varrho\right] =& \frac{1}{a(t,u,z)}\mathbb{E}^{\mathbb{P}}\left[\left(1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right) S_\varrho\right] \\ =& \frac{1}{a(t,u,z)}\left(\mathbb{E}^{\mathbb{P}}\left[S_\varrho\right] + \mathbb{E}^{\mathbb{P}}\left[S_\varrho \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right]\right)\end{aligned}$$

By taking $\eta := \mathbb{1}_{(t,\varrho]}$ in (5.15) we obtain

$$\mathbb{E}^{\mathbb{P}}\left[\int_t^T \theta^*_{r-} dS^{t,u,z}_r (S_\varrho - S_t)\right] = -\mathbb{E}^{\mathbb{P}}\left[S_\varrho\right] + S_t$$

so then

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[Y_\varrho S_\varrho\right] =& \frac{1}{a(t,u,z)}\left(\mathbb{E}^{\mathbb{P}}\left[S_\varrho\right] + \mathbb{E}^{\mathbb{P}}\left[\int_t^T \theta^*_{r-} dS^{t,u,z}_r S_\varrho\right]\right) \\ =& \frac{S_t}{a(t,u,z)}\mathbb{E}^{\mathbb{P}}\left[1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right] = Y_t S_t\end{aligned}$$

iii). Take now any other absolutely continuous σ-martingale measure $\mathbb{Q}$ with $d\mathbb{Q}/d\mathbb{P} \in \mathbb{L}^2(\mathbb{P})$. Then

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right] - \mathbb{E}^{\mathbb{P}}\left[\left(\frac{d\mathbb{Q}^*}{d\mathbb{P}}\right)^2\right] \geq 2\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}^*}{d\mathbb{P}}\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - \frac{d\mathbb{Q}^*}{d\mathbb{P}}\right)\right] \\ =& \frac{2}{a(t,u,z)}\left(\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\int_t^T \theta^*_{r-} dS^{t,u,z}_r\right] - \mathbb{E}^{\mathbb{P}}\left[\frac{dQ^*}{d\mathbb{P}}\int_t^T \theta^*_{r-} dS^{t,u,z}_r\right]\right) \\ =& 0\end{aligned}$$

since both $\mathbb{Q}$ and $\mathbb{Q}^*$ are martingale measures. It follows that

$$\inf_{Q \in \mathcal{M}_{t,u,z}} \mathbb{E}^P\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right] = \mathbb{E}^P\left[\left(\frac{d\mathbb{Q}^*}{d\mathbb{P}}\right)^2\right] = \frac{1}{a(t,u,z)}$$

by using *i*).

□

The characterization given in Lemma 5.2-*iii*) can be used to deduce an upper and lower bound for function a:

Lemma 5.3. *Let Assumptions 5.1-*[**C**, **I1**] *hold true and assume that one of the conditions below also holds true*

i). $0 < \sigma_{min}$

ii). $0 < |\Gamma|$, *where* $|\Gamma|$ *is defined in 5.7*

Then the function a verifies

$$e^{-C(T-t)} \leq a(t,u,z) \leq 1 \tag{5.16}$$

where $C := 2\left(\|\tilde{\mu}\|^2 + \|\tau\|^2_{1,\nu}\right)\left(\max\left(\sigma^2_{min}, |\Gamma|\right)\right)^{-1}$

Proof.
Trivially $a \leq 1$ whereas from (5.14) we deduce

$$a(t,u,z) = \sup_{\mathbb{Q}\in\mathcal{M}_{t,u,z}} \frac{1}{\mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right]}$$

If $\sigma_{min} > 0$ then Itô's formula yields

$$dS_t = S_{t-}(\tilde{\mu}dt + \sigma dW^1_t + \int_{\mathbb{R}} (e^{\gamma} - 1)\bar{J}(dydt)$$

Girsanov's Theorem for Markov jump process (Jacod and Shiryaev, 2003) allows us to select $d\mathbb{Q}/d\mathbb{P} := \xi_T$ where $d\xi_t := -\xi_t\alpha_t dW^1_t$ and

$$\alpha := \frac{1}{\sigma}\left(\mu + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}}\left(e^{\gamma} - 1 - \gamma\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)\right)$$

so that

$$\mathbb{E}\left[\xi^2_T\right] = 1 + \int_t^T \mathbb{E}\left[\xi^2_r\alpha^2\left(r, U^{t,u}_r, Z^{t,u,z}_r\right)\right]dr \leq 1 + 2\frac{\|\tilde{\mu}\|^2 + \|\tau\|^2_{1,\nu}}{\sigma^2_{min}}\int_t^T \mathbb{E}\left[\xi^2_r\right]dr$$

If instead $|\Gamma| > 0$ then we can select $d\mathbb{Q}'/d\mathbb{P} := \eta_T$ where $d\eta_t = -\eta_t\int_{\mathbb{R}}\beta_t(y)\tilde{J}(dydt)$ and

$$\beta_t(y) = \frac{(e^{\gamma} - 1)}{\int_{\mathbb{R}}(e^{\gamma} - 1)^2\nu(dy)}\left(\mu + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}}\left(e^{\gamma} - 1 - \gamma\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)\right)$$

and again

$$\mathbb{E}\left[\eta^2_T\right] = 1 + \int_t^T \mathbb{E}\left[\eta^2_r\beta^2\left(r, U^{t,u}_r, Z^{t,u,z}_r\right)\right]dr \leq 1 + 2\frac{\|\tilde{\mu}\|^2 + \|\tau\|^2_{1,\nu}}{|\Gamma|}\int_t^T \mathbb{E}\left[\eta^2_r\right]dr$$

Gronwall's inequality, in both cases, gives $a(t,u,z) \geq e^{C(t-T)}$.

□

In order to solve Problem (5.11) by the Hamilton-Jacobi-Bellman approach we also need some a priori regularity on the function a.

Theorem 5.4. *Let Assumptions 5.1 hold true. The map $(t,u,z) \to a(t,u,z)$ is measurable. Furthermore let us assume that one of the conditions below also holds true*

i). $0 < \sigma_{min}$

ii). $0 < |\Gamma|$*, where* $|\Gamma|$ *is defined in 5.7*

Then there exists some $T^* > 0$ and $K^a_{lip} \geq 0$ depending on T^* such that for $T < T^*$ one has

$$|a(t,u,z') - a(t,u,z)| \leq K^a_{lip}|z - z'|$$

for all $t \in [0,T]$ and $u, z, z' \in \mathbb{R}$. T^* depends on $\overline{\mu}$, τ and τ^U, σ_{min}, σ_{max} and K_{max} and $T^* \to +\infty$ when $K_{max} \to 0$ and the other constants remain fixed.

Proof.
We start by considering the problem (5.12) when the minimization is only carried over piecewise constant simple strategies. This corresponds to discretize the value function v_0 on a partition of $[0,T]$. Let then $n \in \mathbb{N}^*$ and $0 = t_0 < t_1 < \cdots < t_{2^n} = T$ where $t_i = iT2^{-n}$. The set of admissible strategies at time t_k is given by

$$\mathcal{D}^n_k := \left\{ \theta : (t_k, T] \times \Omega \to \mathbb{R},\, \theta_r = \sum_{j=k}^{2^n-1} \tilde{\theta}_j \mathbb{1}_{r\in]t_j,t_{j+1}]} \mid \tilde{\theta}_j \in \mathbb{L}^\infty(\mathcal{F}_{t_j}) \right\}$$

The discretized wealth process is defined as:

$$\triangle X^{t_k,u,z,x,\theta}_{t_i} = \theta_{t_{i-1}} \left(S^{t_k,u,z}_{t_i} - S^{t_k,u,z}_{t_{i-1}} \right), \quad X_{t_k} = x,\, k < i \leq 2^n \tag{5.17}$$

We can write the corresponding value function for all $k < 2^n$:

$$v^n_k(u,z,x) := \inf_{\theta\in\mathcal{D}^n_k} \mathbb{E}^{\mathbb{P}} \left[\left(x + \sum_{j=k}^{2^n-1} \theta_j \left(S_{t^{t_k,u,z}_{j+1}} - S^{t_k,u,z}_{t_j} \right) \right)^2 \right] = x^2 a^n_k(u,z) \tag{5.18}$$

$$a^n_k(u,z) := \inf_{(\pi_j),\, \pi_j\in\mathbb{L}^2(\mathcal{F}_{t_j})} \mathbb{E}^{\mathbb{P}} \left[\left(1 + \sum_{j=k}^{2^n-1} \pi_j \left(e^{Z_{t^{t_k,u,z}_{j+1}} - Z^{t_k,u,z}_{t_j}} - 1 \right) \right)^2 \right] \tag{5.19}$$

At time $t_n = T$ we have $v^n_{2^n}(u,z,x) = x^2$ and $a^n_{2^n}(u,z) = 1$. If $\pi^*_k(u,z)$ is the optimal strategy in (5.19) then the optimal strategy in (5.18) is given by

$$\theta^*_k(u,z,x) := e^{-z}\pi^*_k(u,z)x \tag{5.20}$$

Remark also that

$$a(t_k,u,z) \leq a^n_k(u,z) \tag{5.21}$$

since $\mathcal{D}_k^n$ in included in $\mathcal{X}(t,u,p,z)$ defined in (5.10). To simplify the notations we introduce

$$C_e := \max(C_{e,2}, C_{e,3}, C_{e,4}) \qquad C_{dz} := \max(C_{dz,2}, C_{dz,3}) \tag{5.22}$$

where $C_{E,i}$ and $C_{DZ,i}$ are defined in Corollaries A.3–A.4 in Appendix A. The scheme of the proof is the following

Step 1: We prove that the functions a_k defined above are continuous and continuously differentiable w.r.t. z with bounded derivative, for all $n \in \mathbb{N}$ and $0 \leq k \leq 2^n$, and we give a relation between the Lipschitz constants of a_k and a_{k+1}.

Step 2: We prove that there exists a T^* such that if $T < T^*$ then the functions a_k are uniformly Lipschitz w.r.t.z.

Step 3: We consider the linear interpolation of the functions a_k and we prove that this interpolation converges pointwise to the function a defined in (5.13). We conclude.

Step 1: The functions a_k are all continuous and continuously differentiable w.r.t. z. The prove is done by recurrence. Fix $k \leq 2^n$ and let a_k the function defined in (5.19) where, to simplify the notation, we omit the superscript n. Assume then

Recurrence hypothesis: for all $l = k+1, \ldots, 2^n$

i). $a_l \in C^0(\mathbb{R}^2)$ and $a_l(u,.) \in C^1(\mathbb{R})$, for all $u \in \mathbb{R}$.

ii). There exist a family of positive constants $L_l \geq 0$ such that $|\partial_z a_l(u,z)| \leq L_l$ for all $u, z \in \mathbb{R}$

Remark that the above assumptions are trivially verified by a_{2^n} with $L_{2^n} = 0$. These regularity assumptions on the functions a_k allow us to prove a dynamic programming principle:

Lemma 5.5. *Let Assumptions 5.1 stand in force and assume that the recurrence hypothesis holds true and that for all $l = k+1, \ldots, 2^n - 1$*

$$a_l(u,z) = \inf_{\pi \in \mathbb{R}} \mathbb{E}^{\mathbb{P}} \left[\left(1 + \pi \left(e^{Z_{t_{l+1}}^{t_l,u,z} - z} - 1 \right) \right)^2 a_{l+1}\left(U_{t_{l+1}}^{t_l,u}, Z_{t_{l+1}}^{t_l,u,z} \right) \right] \tag{5.23}$$

then a_k also verifies the above expression.

The Proof of this Lemma is postponed in paragraph 5.3.1.

Lemma 5.5 tells us that the function a_k verifies (5.23) provided that the functions a_l, $l = k+1, \ldots, 2^n - 1$ verify the recurrence hypothesis and (5.23). By differentiating w.r.t. π we find the optimal control

$$\pi_{t_k}^*(u,z) := - \frac{\mathbb{E}^{\mathbb{P}} \left[\left(e^{Z_{t_{k+1}}^{t_k,u,z} - z} - 1 \right) a_{k+1}\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} \right) \right]}{\mathbb{E}^{\mathbb{P}} \left[\left(e^{Z_{t_{k+1}}^{t_k,u,z} - z} - 1 \right)^2 a_{k+1}\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} \right) \right]} \tag{5.24}$$

Since a_{k+1} is assumed to be Lipschitz we have

$$\left|\pi^*_{t_k}(u,z)\right| \leq \frac{\left|\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}},z\right)\right]\right| + L_{k+1}\mathbb{E}\left[\left|e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right|\left|Z^{t_k,u,z}_{t_{k+1}}-z\right|\right]}{\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}},Z^{t_k,u,z}_{t_{k+1}}\right)\right]}$$

$$\leq \frac{e^{CT}}{C_{e,5}T2^{-n}\vartheta_{2^{-n}}}\left[\left|\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}},z\right)\right]\right| + L_{k+1}C_{e,2}T2^{-n}\vartheta_{2^{-n}}\right]$$

where we used Lemma A.1, Corollaries A.3–A.4 given in Appendix A and estimation (5.21) together with the bounds on a stated in Lemma 5.3. Consider the process

$$V_s := \mathbb{E}\left[a_{k+1}(U^{t_k,u}_{t_{k+1}},z) \mid U^{t_k,u}_s\right],\ t_k \leq s \leq t_{k+1}$$

It is a bounded martingale with respect to the filtration generated by the Brownian motion B and the Poisson random measure N. The martingale representation property yields

$$V_s = \mathbb{E}\left[a_{k+1}(U^{t_k,u}_{t_{k+1}},z)\right] + \int_{t_k}^s \alpha_r dB_r + \int_{t_k}^s\int_{\mathbb{R}} \beta_{r-}(y)\tilde{N}(dydr) \tag{5.25}$$

for some predictable processes α,β. By using the Itô's formula and the independence of the Brownian motions and the Poisson random measures leading the processes Z and U we obtain

$$\left|\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}},z\right)\right]\right| = \left|\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)V_{t_{k+1}}\right]\right|$$

$$= \left|\mathbb{E}\left[\int_{t_k}^{t_{k+1}} e^{Z_r-z}V_r\left(\tilde{\mu}_r + \int_{|y|\geq 1}(e^\gamma-1)\nu(dy)\right)dr\right]\right| \leq C_{e,3}T2^{-n}\vartheta_{2^{-n}}$$

According to (5.22) we obtain an important estimation on the optimal π^*:

$$\left\|\pi^*_{t_k}\right\|_\infty \leq e^{CT}\frac{C_e}{C_{e,5}}\left(1+L_{k+1}\right)\vartheta_{2^{-n}} \tag{5.26}$$

and from (5.23) we can also write

$$a_k(u,z) = \mathbb{E}\left[a_{k+1}\left(U^{t_k,u}_{t_{k+1}},Z^{t_k,u,z}_{t_{k+1}}\right)\right] - \frac{\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}},Z^{t_k,u,z}_{t_{k+1}}\right)\right]^2}{\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}},Z^{t_k,u,z}_{t_{k+1}}\right)\right]}$$

Lemma 5.6. *Let Assumptions 5.1 stand in force and assume that the recurrence hypothesis hold true and that for all $l = k+1,\dots,2^n-1$ a_l verifies (5.23). Then a_k also verifies the recurrence hypothesis and*

$$L_k := L_{k+1} + \left(C_e\frac{e^{CT}}{C_{e,5}}\right)^2\left[\Lambda(K_{max}) + \Psi(K_{max})L_{k+1}(L^2_{k+1}+1)\right]T2^{-n}\vartheta_{2^{-n}} \tag{5.27}$$

where

$$\Lambda(K_{max}) := 2C_{dz}^{1/2}\left(6C_e^{1/2} + C_{dz}^{1/2}\right)$$

$$\Psi(K_{max}) := \frac{3}{2}\left(\Lambda(K_{max}) + C_{dz} + 2(C_e + K_{max}(\sigma_{max} + \|\tau\|_{2,\nu}^2)) + C_e\left(\frac{3+K_{max}}{2}\right)\right)$$

Furthermore

$$\frac{\Lambda(K_{max})}{\sqrt{K_{max}}} \to M \text{ for some } M \geq 0 \text{ when } K_{max} \to 0 \tag{5.28}$$

The Proof of this Lemma is postponed in paragraph 5.3.2.

We have already remarked that a_{2^n} trivially verifies the recurrence hypothesis: Lemma 5.5 tells us that a_{2^n-1} verifies the dynamic programming (5.23) and consequently we can use Lemma 5.6 to prove that a_{2^n-1} also verifies the recurrence hypothesis. By repeating the above argument we conclude that all the functions a_k finally verify the recurrence hypothesis.

Step 2: The functions a_k are uniformly Lipschitz w.r.t. z. We now prove that the constants L_k are uniformly bounded. Let us start by considering the following ODE:

$$-L'(t) = \left(C_E\frac{e^{CT}}{C_{e,5}}\right)^2\left[\Lambda(K_{max}) + \Psi(K_{max})L(t)(L(t)^2+1)\right],\ 0 \leq t < T,\ L(T) = 0$$

Lemma B.1 in Appendix B gives a way to compute L, which is a positive and non increasing function. It follows that for all $k \leq 2^n$,

$$\begin{aligned} L(t_k) =& L(t_{k+1}) + \int_{t_k}^{t_{k+1}} L'(r)dr \\ \geq& L(t_{k+1}) + T2^{-n}\left(C_E\frac{e^{CT}}{C_{e,5}}\right)^2\left[\Lambda(K_{max}) + \Psi(K_{max})L(t_{k+1})(L(t_{k+1})^2+1)\right] \end{aligned}$$

From (5.27) we deduce $L_k \leq L(t_k)$: our aim is then to prove that the function L is bounded. Again Lemma B.1 gives that the function L only depends on $y^* \in \mathbb{R}$ defined as follows:

$$(y^*)^3 + y^* + \frac{\Lambda(K_{max})}{\Psi(K_{max})} = 0$$

Remark that this y^* does not depend on T, n of k. In particular $\sup_{t \leq T} L(t) = L(0) < +\infty$ if and only if $T < T(y^*)$, and we have the explicit form of this $T(y^*)$ (see (B.2)):

$$T(y^*) = \frac{M_1(y^*)}{\Psi(K_{max})} = M_2(y^*)e^{-2CT} := T(y^*, T)$$

for some positive constants $M_1(y^*), M_2(y^*)$. It follows that if T^* is the unique solution of $T^* = T(y^*, T^*)$ and $T < T^*$, then $L(0) < +\infty$. We conclude that $\sup_{n\in\mathbb{N}, k\leq 2^n} L_k \leq L(0)$ and then

$$\sup_{n,k}\left|a_k^n(u,z) - a_k^n(u,z')\right| \leq L(0)|z - z'|$$

for all u, z, z', provided that $T < T^*$. Still from Lemma B.1 we have $T^* \to +\infty$ when $\Lambda(K_{max}) \to 0$, and this is the case when $K_{max} \to 0$ and the other constants appearing in the definition of $\Lambda(K_{max})$ remain fixed.

Step 3: The linear combination of a_k converges pointwise to a. From now on we will always highlight the dependence of a_k^n on n. Let us define the function $a^n(t, u, z)$ on $[0, T] \times \mathbb{R}^2$ as follows

$$\begin{aligned} &a^n(t, u, z) = 2^n \left(a_{k+1}^n(u, z) - a_k^n(u, z)\right)(t - t_k) + a_k^n(u, z) \\ &\text{where } k \text{ verifies } : \quad t_k \le t < t_{k+1} \end{aligned} \tag{5.29}$$

From the properties of a_k^n we have that the interpolation function a^n is continuous, continuously differentiable in the variable z and under the condition $T < T^*$ it is straightforward to see that

$$\left|a^n(t, u, z) - a^n\left(t, u, z'\right)\right| \le K_{lip}^a \left|z - z'\right|, \quad \text{for all } t \in [0, T], \, u, z, z' \in \mathbb{R} \tag{5.30}$$

where $K_{lip}^a = 3L(0)$.

Lemma 5.7. *Suppose Assumptions 5.1-*$[\mathbf{C}, \mathbf{I_1}]$ *hold true. Fix* $(t, u, z) \in [0, T] \times \mathbb{R}^2$, $\varepsilon > 0$ *and* $\bar{\theta} \in \mathcal{D}_t$ *so that*

$$\mathbb{E}\left[\left(1 + \int_t^T \bar{\theta}_r dS_r^{t,u,z}\right)^2\right] \le a(t, u, z) + \varepsilon$$

There exist $M > 0$ *and* $\bar{n} \in \mathbb{N}$ *only depending on* $\bar{\theta}$ *such that for all* $n \ge \bar{n}$ *we can find* $\theta^{\varepsilon,n} \in \mathcal{D}_k^n$ *with*

$$\left|\mathbb{E}\left[\left(1 + \int_t^T \bar{\theta}_r dS_r^{t,u,z}\right)^2\right] - \mathbb{E}\left[\left(1 + \int_{t_k}^T \theta_r^{\varepsilon,n} dS_r^{t_k,u,z}\right)^2\right]\right| \le M\varepsilon \tag{5.31}$$

where $0 \le k \le 2^n$ *verifies* $t_k \le t < t_{k+1}$. *The same result still holds true if we consider* t_{k+1} *instead of* t_k.

The Proof of this Lemma is postponed in paragraph 5.3.3.

The above result tells us that for all $\varepsilon > 0$ there exists $\bar{n}$ such that for all $n \ge \bar{n}$ we can select two controls $\theta^{\varepsilon,n,1} \in \mathcal{D}_k^n$ and $\theta^{\varepsilon,n,2} \in \mathcal{D}_{k+1}^n$ verifying

$$\begin{aligned} a_k^n(u, z) &\le \mathbb{E}\left[\left(1 + \int_{t_k}^T \theta_{r-}^{\varepsilon,n,1} dS_r^{t_k,u,z}\right)^2\right] \le a(t, u, z) + (M+1)\varepsilon \\ a_{k+1}^n(u, z) &\le \mathbb{E}\left[\left(1 + \int_{t_{k+1}}^T \theta_{r-}^{\varepsilon,n,2} dS_r^{t_{k+1},u,z}\right)^2\right] \le a(t, u, z) + (M+1)\varepsilon \end{aligned}$$

The above estimations and (5.21) yield

$$|a_{k+1}^n(u, z) - a_k^n(u, z)| \le (M+1)\varepsilon \tag{5.32}$$

so that, for fixed (t, u, z), $\varepsilon > 0$ and $n \geq \bar{n}$ as above we obtain:

$$\begin{aligned}|a^n(t,u,z) - a(t,u,z)| \leq & 2^n(t-t_k)\left|a^n_{k+1}(u,z) - a^n_k(u,z)\right| + |a^n_k(u,z) - a(t,u,z)| \\ \leq & (M+1)(T+1)\varepsilon\end{aligned}$$

Arbitrary $\varepsilon > 0$ allows us to deduce that $a^n \to a$ pointwise: in particular, if $T < T^*$ then the function a is Lipschitz continuous in the variable z and at least measurable in the variable t, u (since it is the limit of continuous functions).

□

5.3.1 Proof of Lemma 5.5

Proof.
If (5.23) holds true, then, by differentiating w.r.t. π we obtain

$$a_l(u,z) = \tag{5.33}$$
$$\mathbb{E}^{\mathbb{P}}\left[a_{l+1}\left(U^{t_l,u}_{t_{l+1}}, Z^{t_l,u,z}_{t_{l+1}}\right)\right] - \frac{\mathbb{E}^{\mathbb{P}}\left[\left(e^{Z^{t_l,u,z}_{t_{l+1}}-z} - 1\right) a_{l+1}\left(U^{t_l,u}_{t_{l+1}}, Z^{t_l,u,z}_{t_{l+1}}\right)\right]}{\mathbb{E}^{\mathbb{P}}\left[\left(e^{Z^{t_l,u,z}_{t_{l+1}}-z} - 1\right)^2 a_{l+1}\left(U^{t_l,u}_{t_{l+1}}, Z^{t_l,u,z}_{t_{l+1}}\right)\right]}$$

for all $l > k$. From the definition of a_k in (5.19) we can write

$$\begin{aligned}a_k(u,z) = & \inf_{(\pi_k,\dots,\pi_{2^n-1})} \mathbb{E}\left[\left(1 + \sum_{j=k}^{2^n-2} \pi_j \left(e^{Z^{t_j,u,z}_{t_{j+1}} - Z^{t_j,u,z}_{t_j}} - 1\right)\right)^2 + \right. \\ & 2\pi_{2^n-1}\left(1 + \sum_{j=k}^{2^n-1} \pi_j \left(e^{Z^{t_j,u,z}_{t_{j+1}} - Z^{t_j,u,z}_{t_j}} - 1\right)\right) \mathbb{E}^{\mathcal{F}_{t_{2^n-1}}}\left[\left(e^{Z^{t_k,u,z}_{t_{2^n}} - Z^{t_k,u,z}_{t_{2^n-1}}} - 1\right)\right] \\ & \left. + \pi^2_{2^n-1}\mathbb{E}^{\mathcal{F}_{t_{2^n-1}}}\left[\left(e^{Z^{t_k,u,z}_{t_{2^n}} - Z^{t_k,u,z}_{t_{2^n-1}}} - 1\right)^2\right]\right]\end{aligned}$$

where $\mathbb{E}^{\mathcal{F}_{t_{2^n-1}}}[\dots]$ stands for the conditional expectation w.r.t. $\mathcal{F}_{t_{2^n-1}}$. We can now minimize the above expression w.r.t. π_{2^n-1}: by using the Markov property of (U, Z) and (5.33) we obtain

$$a_k(u,z) = \inf_{(\pi_k,\dots,\pi_{2^n-2})} \mathbb{E}\left[\left(1 + \sum_{j=k}^{2^n-2} \pi_j \left(e^{Z^{t_j,u,z}_{t_{j+1}} - Z^{t_j,u,z}_{t_j}} - 1\right)\right)^2 a_{2^n-1}\left(U^{t_k,u}_{t_{2^n-1}}, Z^{t_k,u,z}_{t_{2^n-1}}\right)\right]$$

Remark that the right hand side of the above expression is well defined since a_{2^n-1} is supposed to verify the recurrence hypotheses. We can repeat the procedure for $\pi_{t_{2^n-2}}, \pi_{t_{2^n-3}}, \dots, \pi_{t_{k+1}}$ and finally we obtain

$$a_k(u,z) = \inf_{\pi_k \in \mathbb{R}} \mathbb{E}\left[\left(1 + \pi_k\left(e^{Z^{t_k,u,z}_{t_{k+1}} - z} - 1\right)\right)^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]$$

which concludes our proof.

□

5.3.2 Proof of Lemma 5.6

Proof.
We know that

$$a_k(u,z) = \mathbb{E}\left[a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right] - \frac{\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]^2}{\mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]}$$

provided that a_l, $l > k$ verifies the assumptions of Lemma 5.5. It is straightforward to see that a_k is continuous: for this one can use the fact that a_{k+1} is continuous and the estimations on the exponential of Z given in Corollary A.3, Appendix A. We now prove that a_k is also differentiable w.r.t. z and give an estimation of this derivative. To simplify, let us call

$$\begin{aligned}
\delta_k(u,z) :=& \mathbb{E}\left[a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]\\
\alpha_k(u,z) :=& \mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]\\
\beta_k(u,z) :=& \mathbb{E}\left[\left(e^{Z^{t_k,u,z}_{t_{k+1}}-z}-1\right)^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]
\end{aligned}$$

so that $a_k = \delta_k - (\alpha_k)^2/\beta_k$. Remark that

$$\begin{aligned}
\alpha_k(u,z) =& e^{-z}\mathbb{E}\left[e^{Z^{t_k,u,z}_{t_{k+1}}} a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right] - \delta_k(u,z)\\
\beta_k(u,z) =& e^{-2z}\mathbb{E}\left[(e^{Z^{t_k,u,z}_{t_{k+1}}})^2 a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right] - \delta_k(u,z) - 2\alpha_k
\end{aligned}$$

Proving that δ_k is continuously differentiable is trivial and we obtain

$$\partial_z\delta_k(u,z) = \mathbb{E}\left[\partial_z a_{k+1}(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}})DZ^{t_k,u,z}_{t_{k+1}}\right]$$

where DZ is the derivative of the flow Z with respect to z (see Appendix A):

$$DZ^{t,u,z}_s = 1 + \int_t^s DZ^{t,u,z}_{r-}\left(\frac{\partial\mu_r}{\partial z}dr + \frac{\partial\sigma_r}{\partial z}dW^1_r + \int_{\mathbb{R}}\frac{\partial\gamma_r(y)}{\partial z}\bar{J}\,(dydr)\right)$$

where $\partial\mu_r/\partial z := \partial\mu_r/\partial z(r, U^{t,u}_r, Z^{t,u,z}_r)$. We can prove that α_k and β_k are differentiable with the same type of computations, so we just detail them for β_k. For this, let us assume that α_k is differentiable, so we only need to prove that

$$\tilde{\beta}_k(u,z) := \mathbb{E}\left[e^{2Z^{t_k,u,z}_{t_{k+1}}} a_{k+1}\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)\right]$$

is differentiable w.r.t. z. We start with

$$\frac{\tilde{\beta}(u,z+\varepsilon) - \tilde{\beta}(u,z)}{\varepsilon} = \mathbb{E}\left[D_\sharp Z^{\varepsilon,t_k,u,z}_{t_{k+1}}\int_0^1 \varphi\left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}} + r\varepsilon D_\sharp Z^{\varepsilon,t_k,u,z}_{t_{k+1}}\right)dr\right]$$

where $\varphi(u,z) = e^{2z}(2a_{k+1}(u,z) + \partial_z a_{k+1}(u,z))$ and $D_\sharp Z_s^{\varepsilon,t_k,u,z} := (Z_s^{t_k,u,z+\varepsilon} - Z_s^{t_k,u,z})/\varepsilon$. From Lemma A.1 and Corollary A.4 we get

$$Z_{t_{k+1}}^{t_k,u,z+\varepsilon} \xrightarrow{\mathbb{P}} Z_{t_{k+1}}^{t_k,u,z} \qquad \text{and} \qquad D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z} \xrightarrow{\mathbb{P}} DZ_{t_{k+1}}^{t_k,u,z} \qquad \text{when } \varepsilon \to 0$$

Corollaries A.3–A.4 and the recurrence hypothesis on a_{k+1} allow us to apply dominated convergence:

$$D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z} \int_0^1 \varphi\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} + r\varepsilon D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z}\right) dr \xrightarrow{\mathbb{P}} DZ_{t_{k+1}}^{\varepsilon,t_k,u,z} \varphi\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z}\right)$$

If we prove that the family $\varepsilon \to D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z} \int_0^1 \varphi\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} + r\varepsilon D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z}\right) dr$ is uniformly integrable we are done: first we have

$$\begin{aligned}
&\left| D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z} \int_0^1 \varphi\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} + r\varepsilon D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z}\right) dr \right| \\
\leq& \frac{L_{k+1}+2}{\varepsilon} \left| Z_{t_{k+1}}^{t_k,u,z+\varepsilon} - Z_{t_{k+1}}^{t_k,u,z} \right| e^{2Z_{t_{k+1}}^{t_k,u,z}} \int_0^1 e^{2r\left(Z_{t_{k+1}}^{t_k,u,z+\varepsilon} - Z_{t_{k+1}}^{t_k,u,z}\right)} dr \\
\leq& \frac{L_{k+1}+2}{\varepsilon} \left| e^{2Z_{t_{k+1}}^{t_k,u,z+\varepsilon}} - e^{2Z_{t_{k+1}}^{t_k,u,z}} \right|
\end{aligned}$$

Take now the test function $g(x) = x^{1+\upsilon}$ for some $\upsilon > 0$. The Cauchy-Schwarz inequality yields

$$\begin{aligned}
&\mathbb{E}\left[g\left(\left| D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z} \int_0^1 \varphi\left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z} + r\varepsilon D_\sharp Z_{t_{k+1}}^{\varepsilon,t_k,u,z}\right) dr \right| \right) \right] \leq \\
&\left(\frac{L_{k+1}+2}{2\varepsilon} \right)^{1+\upsilon} \mathbb{E}\left[\left| e^{Z_{t_{k+1}}^{u,z+\varepsilon}} - e^{Z_{t_{k+1}}^{u,z}} \right|^2 \right]^{\frac{1+\upsilon}{2}} \mathbb{E}\left[\left| e^{Z_{t_{k+1}}^{u,z+\varepsilon}} + e^{Z_{t_{k+1}}^{u,z}} \right|^{2\frac{1+\upsilon}{1-\upsilon}} \right]^{\frac{1-\upsilon}{2}}
\end{aligned}$$

Recall that the $\exp(Z)$ admits a fourth moment: if we select $0 < \upsilon < 1/3$ then $\mathbb{E}\left[\left| \exp(Z_{t_{k+1}}^{u,z+\varepsilon}) + \exp(Z_{t_{k+1}}^{u,z}) \right|^{2\frac{1+\upsilon}{1-\upsilon}} \right]^{\frac{1-\upsilon}{2}}$ is uniformly bounded in ε, provided that $\varepsilon \leq \varepsilon_0$, for any ε_0. We use again Corollary A.3 to deduce that

$$\mathbb{E}\left[g\left(\left| D_\sharp Z_{t_{k+1}}^{\varepsilon,z} \int_0^1 \varphi(Z_{t_{k+1}}^z + r(Z_{t_{k+1}}^{z+\varepsilon} - Z_{t_{k+1}}^z))dr \right| \right) \right] \leq M$$

for some positive M which depends on u, z but not on ε: with the de La Vallée-Poussin criterion (See, for example, Doob (1994), Chapter VI, 17) we finally prove that the family $\varepsilon \to D_\sharp Z_{t_{k+1}}^\varepsilon \int_0^1 \varphi(U_{t_{k+1}}, Z_{t_{k+1}} + r\varepsilon D_\sharp Z_{t_{k+1}}^\varepsilon)dr$ is uniformly integrable for $\varepsilon \leq \varepsilon_0$. Dominated convergence applies and we can pass to the limit $\varepsilon \to 0$ and get

$$\begin{aligned}
&\frac{\tilde{\beta}(u,z+\varepsilon) - \tilde{\beta}(u,z)}{\varepsilon} \xrightarrow{\varepsilon\to 0} \\
&\mathbb{E}\left[DZ_{t_{k+1}}^{t_k,u,z} e^{2Z_{t_{k+1}}^{t_k,u,z}} \left(2a_{k+1}(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z}) + \partial_z a_{k+1}(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z}) \right) \right]
\end{aligned}$$

The continuity of $\partial_z \tilde{\beta}$ can be proved by using the continuity of a_{k+1} together with its derivative and the estimations on e^Z and DZ given in Corollaries A.3–A.4, and this is enough to prove that a_k is continuously differentiable w.r.t. z: a_k verifies part $i)$ of the recurrence hypothesis.

We now give an estimate on the derivative of a_k w.r.t. z. From now on, in order to lighten the notations, we omit the superscript (t_k, u, z) so that $(U, Z) := \left(U^{t_k,u}_{t_{k+1}}, Z^{t_k,u,z}_{t_{k+1}}\right)$, $\exp(Z) := \exp(Z^{t_k,u,z}_{t_{k+1}})$ and $DZ := DZ^{t_k,u,z}_{t_{k+1}}$. By using the derivative of flow DZ we obtain

$$\begin{aligned}\partial_z a_k(u,z) =& \mathbb{E}\left[\partial_z a_{k+1}(U,Z)DZ\right] + 2\pi^*_{t_k}(u,z)\mathbb{E}\left[e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right] \\ &+2\pi^*_{t_k}(u,z)\mathbb{E}\left[\left(e^{Z-z}-1\right)\partial_z a_{k+1}(U,Z)DZ\right] \\ &+2(\pi^*_{t_k}(u,z))^2\mathbb{E}\left[\left(e^{Z-z}-1\right)e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right] \\ &+(\pi^*_{t_k}(u,z))^2\mathbb{E}\left[\left(e^{Z-z}-1\right)^2\partial_z a_{k+1}(U,Z)DZ\right]\end{aligned}$$

where $\pi^*_{t_k}$ is defined in (5.24). We can rearrange the terms above to obtain

$$\partial_z a_k(u,z) = \mathbb{E}\left[\left(1+\pi^*_{t_k}(u,z)\left(e^{Z-z}-1\right)\right)^2\partial_z a_{k+1}(U,Z)DZ\right] \tag{5.34}$$

$$+2\pi^*_{t_k}(u,z)\mathbb{E}\left[e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right] \tag{5.35}$$

$$+2(\pi^*_{t_k}(u,z))^2\mathbb{E}\left[\left(e^{Z-z}-1\right)e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right] \tag{5.36}$$

and we know that $\|\pi^* t_k\|_\infty \le e^{CT}C_e\left(1+L_{k+1}\right)/C_{e,5}$.
Let us start with (5.36): by recalling that $0 < a_{k+1} \le 1$ and using Corollaries A.3–A.4 we have

$$\begin{aligned}&\left|2(\pi^*_{t_k}(u,z))^2\mathbb{E}\left[\left(e^{Z-z}-1\right)e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right]\right| \\ \le& 2\left\|\pi^*_{t_k}\right\|^2_\infty\left(\mathbb{E}\left[\left|e^{Z-z}-1\right|^2|DZ-1|\right]+\mathbb{E}\left[\left|e^{Z-z}-1\right||DZ-1|\right]\right) \\ \le& 4\left\|\pi^*_{t_k}\right\|^2_\infty C_e^{1/2}C_{dz}^{1/2}T2^{-n}\vartheta_{2^{-n}}\end{aligned}$$

For (5.35) we have

$$\begin{aligned}&\left|2\pi^*_{t_k}(u,z)\mathbb{E}\left[e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right]\right| \\ \le& 2\left\|\pi^*_{t_k}\right\|_\infty\left(\mathbb{E}\left[\left|e^{Z-z}-1\right||DZ-1|\right]+\left|\mathbb{E}\left[(DZ-1)\,a_{k+1}(U,Z)\right]\right|\right) \\ \le& 2\left\|\pi^*_{t_k}\right\|_\infty\left(C_e^{1/2}C_{dz}^{1/2}T2^{-n}\vartheta_{2^{-n}}+\left|\mathbb{E}\left[(DZ-1)a_{k+1}(U,z)\right]\right|+L_{k+1}C_e^{1/2}C_{dz}^{1/2}T2^{-n}\vartheta_{2^{-n}}\right)\end{aligned}$$

The term $|\mathbb{E}\left[(DZ-1)a_{k+1}(U,z)\right]|$ can be estimated by using the martingale representation in (5.25) and we obtain

$$|\mathbb{E}\left[(DZ-1)a_{k+1}(U,z)\right]| \le C_{dz,3}T2^{-n}\vartheta_{2^{-n}} \le C_{dz}T2^{-n}\vartheta_{2^{-n}}$$

hence

$$\begin{aligned}&\left|2\pi^*_{t_k}(u,z)\mathbb{E}\left[e^{Z-z}(DZ-1)a_{k+1}(U,Z)\right]\right| \\ \le& 2\left\|\pi^* t_k\right\|_\infty C_{dz}^{1/2}\left(C_e^{1/2}+C_{dz}^{1/2}+L_{k+1}C_e^{1/2}\right)T2^{-n}\vartheta_{2^{-n}}\end{aligned}$$

For (5.34) we use Lemma A.5 in Appendix A:

$$\begin{aligned}&\left|\mathbb{E}\left[\left(1+\pi^*_{t_k}(u,z)\left(e^{Z-z}-1\right)\right)^2\partial_z a_{k+1}(U,Z)DZ\right]\right|\\ \leq & L_{k+1}\mathbb{E}\left[\left(1+\pi^*_{t_k}\left(e^{Z-z}-1\right)\right)^2|DZ|\right]\leq L_{k+1}\left(1+C_{dz,e}(\pi^*_{t_k})T2^{-n}\vartheta_{2^{-n}}\right)\end{aligned}$$

where

$$C_{dz,e}(\pi^*_k)=C_{dz}+2|\pi^*_{t_k}|(C_e+K_{max}(\sigma_{max}+\|\tau\|^2_{2,\nu}))+|\pi^*_{t_k}|^2C_e\left(\frac{3+K_{max}}{2}\right)$$

From the estimations for (5.34)–(5.35)–(5.36) and the bound on $|\pi^*_{t_k}|$ we will obtain, for some constants Λ and c_i to be determined, a polynomial expression in L_{k+1}:

$$\begin{aligned}&|\partial_z a_k(u,z)|\\ \leq & L_{k+1}+\max\left(\left(C_e\frac{e^{CT}}{C_{e,5}}\right)^2,\ 1\right)\left(\Lambda+L_{k+1}\left(c_1+c_2L_{k+1}+c_3L^2_{k+1}\right)\right)T2^{-n}\vartheta_{2^{-n}}\\ = & L_{k+1}+\left(C_e\frac{e^{CT}}{C_{e,5}}\right)^2\left(\Lambda+\frac{3}{2}\max(c_1,c_2,c_3)L_{k+1}\left(L^2_{k+1}+1\right)\right)T2^{-n}\vartheta_{2^{-n}}\end{aligned}$$

since $C_e\geq C_{e,2}\geq C_{e,5}$. Precise computations yield

$$\begin{aligned}&\Lambda:=2C^{1/2}_{dz}\left(6C^{1/2}_e+C^{1/2}_{dz}\right)\\ &\frac{3}{2}\max(c_1,c_2,c_3):=\frac{3}{2}\left(\Lambda+C_{dz}+2(C_e+K_{max}(\sigma_{max}+\|\tau\|^2_{2,\nu}))+C_e\left(\frac{3+K_{max}}{2}\right)\right)\end{aligned}$$

Furthermore from the definition of C_{dz} in (5.22) we have $C_{dz}\to 0$ when $K_{max}\to 0$ (see Corollary A.4). We focus on this by writing $\Lambda=\Lambda(K_{max})$ so then

$$\frac{\Lambda(K_{max})}{\sqrt{K_{max}}}\to M \text{ for some } M\geq 0 \text{ when } K_{max}\to 0$$

If we set $\Psi(K_{max})=:\frac{3}{2}\max(c_1,c_2,c_3)$ then the following estimate holds true:

$$\|\partial_z a_k\|_\infty:=L_k=L_{k+1}+\left(C_e\frac{e^{CT}}{C_{e,5}}\right)^2\left[\Lambda(K_{max})+\Psi(K_{max})L_{k+1}(L^2_{k+1}+1)\right]T2^{-n}\vartheta_{2^{-n}}$$

The above estimation proves that a_k also verifies the part *ii)* of the recurrence hypothesis.

□

5.3.3 Proof of Lemma 5.7

Proof.
From definition of $\mathcal{D}_t$ is (5.8), we have

$$\bar{\theta}_r:=Y_{m-1}\mathbb{1}_{\{r\in]t,\varsigma_m]\}}+\sum_{i=m}^{N}Y_i\mathbb{1}_{\{r\in]\varsigma_i,\varsigma_{i+1}]\}}$$

for some $N \in \mathbb{N}$, $Y_i \in \mathbb{L}^\infty(\mathcal{F}_{\varsigma_i}, \mathbb{P})$ whereas $(\varsigma_i)_i$ are stopping times. Our idea here is build a strategy which follows $\bar{\theta}$ along the time: for $n \in \mathbb{N}$ let

$$\theta^{\varepsilon,n}_{t_j} := \bar{\theta}_{t_j} \text{ for all } k \leq j \leq 2^n$$

Obviously there are windows of the time grid on which the two strategies do not coincide. By using this strategy we can write

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(1+\int_t^T \bar{\theta}_{r-} dS_r^{t,u,z}\right)^2\right] - \mathbb{E}\left[\left(1+\int_{t_k}^T \theta^{\varepsilon,n}_{r-} dS_r^{t_k,u,z}\right)^2\right]\right| \\
\leq &\left|\mathbb{E}\left[\left(1+\int_{t_k}^T \bar{\theta}_{r-} dS_r^{t_k,u,z}\right)^2\right] - \mathbb{E}\left[\left(1+\int_{t_k}^T \theta^{\varepsilon,n}_{r-} dS_r^{t_k,u,z}\right)^2\right]\right| \\
+ &\left|\mathbb{E}\left[\left(1+\int_{t_k}^T \bar{\theta}_{r-} dS_r^{t_k,u,z}\right)^2\right] - \mathbb{E}\left[\left(1+\int_t^T \bar{\theta}_{r-} dS_r^{t,u,z}\right)^2\right]\right|
\end{aligned}$$

The strategy $\bar{\theta}$ is bounded: the function $s \to \mathbb{E}^{s,u,z}\left[\left(1+\int_s^T \bar{\theta}_{r-} dS_r\right)^2\right]$ is continuous: there exist $\bar{n}$ such that for all $n \geq \bar{n}$ one has

$$\left|\mathbb{E}\left[\left(1+\int_{t_k}^T \bar{\theta}_{r-} dS_r^{t_k,u,z}\right)^2\right] - \mathbb{E}\left[\left(1+\int_t^T \bar{\theta}_{r-} dS_r^{t,u,z}\right)^2\right]\right| \leq \varepsilon$$

since $|t - t_k| \leq 2^{-n}$. We can concentrate then on the first term:

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(1+\int_{t_k}^T \bar{\theta}_{r-} dS_r^{t_k,u,z}\right)^2\right] - \mathbb{E}\left[\left(1+\int_{t_k}^T \theta^{\varepsilon,n}_{r-} dS_r^{t_k,u,z}\right)^2\right]\right| \\
\leq &\mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right] + 2\mathbb{E}\left[\left|1+\int_{t_k}^T \bar{\theta}_{r-} dS_r^{t_k,u,z}\right|\left|\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right|\right] \\
\leq &\mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right] + 2\left(a(t_k,u,z)+\varepsilon\right)^{\frac{1}{2}} \mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right]^{\frac{1}{2}} \\
\leq &\mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right] + 2\left(1+\varepsilon\right)^{\frac{1}{2}} \mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right]^{\frac{1}{2}}
\end{aligned}$$

We conclude our proof if we show that for some $\bar{n}$ and for all $n \geq \bar{n}$ we have

$$\mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right] \leq \varepsilon^2 \tag{5.37}$$

The Doob-Meyer decomposition and Assumptions 5.1-$[\mathbf{C}, \mathbf{I_1}]$ yield

$$\mathbb{E}\left[\left(\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right) dS_r^{t_k,u,z}\right)^2\right] \leq M\mathbb{E}\left[\int_{t_k}^T \left(\bar{\theta}_{r-} - \theta^{\varepsilon,n}_{r-}\right)^2 S_{r-}^2 dr^{t_k,u,z}\right]$$

for some $M > 0$. Since $\bar{\theta}_r - \theta_r^{\varepsilon,n} \to 0$ and

$$\left|\left(\bar{\theta}_r - \theta_r^{\varepsilon,n}\right)^2 S_r^2\right| \leq M' \sup_{r\in[0,T]} S_r^2 \in \mathbb{L}^1(\mathbb{P})$$

for some $M' > 0$ depending on the bounds of $\bar{\theta}$, we can apply dominated convergence to deduce that

$$\mathbb{E}\left[\left(\int_{t_k}^{T} \left(\bar{\theta}_{r-} - \theta_{r-}^{\varepsilon,n}\right) dS_r^{t_k,u,z}\right)^2\right] \to 0 \text{ when } n \to +\infty$$

which proves (5.37).

□

The Proof of Theorem 5.4 actually gives us some important elements on the structure of the pure investment problem: we found a sequence of bounded controls (as stated in (5.24)):

$$\pi_k^n(u,z) := -\frac{\mathbb{E}\left[\left(e^{Z_{t_{k+1}}^{t_k,u,z}-z} - 1\right) a_{k+1}^n \left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z}\right)\right]}{\mathbb{E}\left[\left(e^{Z_{t_{k+1}}^{t_k,u,z}-z} - 1\right)^2 a_{k+1}^n \left(U_{t_{k+1}}^{t_k,u}, Z_{t_{k+1}}^{t_k,u,z}\right)\right]}$$

such that

$$a_k^n(u,z) := \mathbb{E}\left[\left(1 + \sum_{j=k}^{2^n-1} \pi_j^n \left(U_{t_j}^{t_k,u}, Z_{t_j}^{t_k,u,z}\right)\left(e^{Z_{t_{j+1}}^{t_k,u,z} - Z_{t_j}^{t_k,u,z}} - 1\right)\right)^2\right]$$

and

$$\sup_{n\in\mathbb{N}} \sup_{k\leq 2^n} \|\pi_k^n\|_\infty \leq \frac{e^{CT}}{\sigma_{min}^2 \vee |\Gamma|} C_e \left(1 + K_{lip}^a\right) := \bar{\Pi} \tag{5.38}$$

from the bound (5.26) and the fact that $L_{k+1} \leq L(0) \leq K_{lip}^a$ if $T < T^*$. Consider now the function a^n introduced in (5.29): according to the fact that $a^n \to a$ pointwise and (5.32) we also deduce that for any $t \in [0,T]$

$$a(t,u,z) = \lim_{n\to\infty} a_k^n(u,z) \qquad \text{where } t_k \leq t < t_{k+1}$$

In particular this implies that $(\pi^n)_n$ is a minimizing sequence in Problem (5.13), or equivalently, by using (5.20), the sequence $\theta_r^n := \theta^n(r, U_{r-}, Z_{r-}, X_{r-}^{\theta^n})$, where

$$\theta^n(r,u,z,x) := e^{-z} x \sum_{k=0}^{2^n-1} \pi_k^n(u,z) \mathbb{1}_{\{r\in]t_k,t_{k+1}]\}} \tag{5.39}$$

is a minimizing sequence in problem (5.12). Remark that since π^n is bounded the strategy θ^n is admissible.

5.4 The structure of the quadratic hedge value function

In this Section we will discuss the structure of the value function introduced in (5.11). We define $F^{t,u,p,z} := f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})$ and

$$\mathcal{S} := \left\{ \int_t^T \theta_{r-} dS_r \,|\theta \in \mathcal{X}(t,u,p,z) \right\} \tag{5.40}$$

In what follows $\mathrm{Pr}^{\mathcal{S}}$ denotes the $\mathbb{L}^2$-projection into the space of stochastic integrals $\mathcal{S}$. Remark that the projection is well defined since $\mathcal{S}$ is a convex closed subset in the Hilbert space $\mathbb{L}^2(\mathbb{P})$. With these notations, we can see the quadratic hedge problem (5.11) as the projection of the random variable $F - x$ on $\mathcal{S}$:

$$\begin{aligned} v^f(t,u,p,z,x) =& \mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(F^{t,u,p,z} - x)\right)^2\right] = \mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(F^{t,u,p,z}) - x\mathrm{Pr}^{\mathcal{S}}(1)\right)^2\right] \\ =& \mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(1)\right)^2\right] x^2 - 2\mathbb{E}^{\mathbb{P}}\left[\mathrm{Pr}^{\mathcal{S}}(F^{t,u,p,z})\mathrm{Pr}^{\mathcal{S}}(1)\right] x + \mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(F^{t,u,p,z})\right)^2\right] \end{aligned}$$

From the definition of $\mathrm{Pr}^{\mathcal{S}}$ we first obtain

$$\mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(1)\right)^2\right] := \inf_{\theta \in \mathcal{X}(t,u,z)} \mathbb{E}^{\mathbb{P}}\left[\left(1 + \int_t^T \theta_{r-} dS_r^{t,u,z}\right)^2\right] = a\,(t,u,z)$$

which does not depend on the particular structure of the function f. In conclusion, if we define

$$\begin{aligned} b^f(t,u,p,z) :=& -2\mathbb{E}^{\mathbb{P}}\left[\mathrm{Pr}^{\mathcal{S}}(f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}))\mathrm{Pr}^{\mathcal{S}}(1)\right] \text{ and} \\ c^f(t,u,p,z) :=& \mathbb{E}^{\mathbb{P}}\left[\left(\mathrm{Pr}^{\mathcal{S}}(f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}))\right)^2\right] \end{aligned}$$

then the value function v^f admits the following quadratic decomposition:

$$v^f(t,u,p,z,x) = a\,(t,u,z)\,x^2 + b^f(t,u,p,z)x + c^f(t,u,p,z) \tag{5.41}$$

This quadratic structure for the value function is well known in the literature (see for example Jeanblanc et al. (2011)). From Lemma 5.3 we have $a > 0$ so then it is straightforward to obtain the optimal price in (5.11):

$$x^*(f) := \arg\min_{x\in\mathbb{R}} v^f(t,u,p,z,x) = -\frac{b^f(t,u,p,z)}{2a(t,u,z)} \tag{5.42}$$

which is a linear function of the payoff f since b^f is. The following Lemma proves the stability of the optimal price $x^*(f)$ and the optimal hedging strategy under small perturbation of the function f:

Lemma 5.8. *Let f_1, f_2 be two measurable functions with $f_i(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}) \in \mathbb{L}^2(\mathbb{P})$ for all t,u,p,z, $i = 1,2$. Then for any $t < T$ and $(u,p,z) \in \mathbb{R}^3$*

$$|x^*(f_1)(t,u,p,z) - x^*(f_2)(t,u,p,z)| \le a(t,u,z)^{-1/2}\left\|(f_1 - f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2$$

$$\left|\left(v^{f_1} - v^{f_2}\right)(t,u,p,z,x)\right| \le$$

$$2\left(x + \left\|(f_1+f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2\right)\left\|(f_1 - f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2$$

Fix now (t,u,p,z,x) *and let* f_n *such that* $\left\|(f_n - f)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2 \to 0,\ n \to \infty$. *If* θ^n *is the optimal control in* (5.11) *when one uses* f_n *then, for all* $\varepsilon > 0$, *there exists some* $N > 0$ *such that for any* $n \geq N$ *one has*

$$\left| v^f(t,u,p,z,x) - \mathbb{E}^{\mathbb{P}}\left[\left(f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}) - x - \int_t^T \theta^n_{r-} dS_r^{t,u,z} \right)^2 \right] \right| \leq \varepsilon$$

***Proof*.**
Let $\triangle f := (f_1 - f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})$. From the definition of b^f and $x^*(f)$ we obtain

$$\begin{aligned}|x^*(f_1)(t,u,p,z) - x^*(f_2)(t,u,p,z)| \leq & a^{-1}(t,u,z)\mathbb{E}\left[\mathrm{Pr}^{\mathcal{S}}(1)^2\right]^{1/2} \mathbb{E}\left[\mathrm{Pr}^{\mathcal{S}}\left(\triangle f\right)^2\right]^{1/2} \\ \leq & a(t,u,z)^{-1/2} \left\|\triangle f\right\|_2\end{aligned}$$

and from (5.41) we obtain

$$\begin{aligned}&\left| v^{f_1}(t,u,p,z,x) - v^{f_2}(t,u,p,z,x) \right| \\ \leq & 2\left(a^{1/2}(t,u,z)|x| + \left\|(f_1 + f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2 \right) \mathbb{E}\left[\mathrm{Pr}^{\mathcal{S}}\left(\triangle f)\right)^2\right]^{1/2} \\ \leq & 2\left(x + \left\|(f_1 + f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2 \right) \left\|\triangle f\right\|_2\end{aligned}$$

We conclude the proof by remarking that

$$\begin{aligned}0 \leq & \mathbb{E}\left[\left(f\left(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}\right) - x - \int_t^T \theta^n_{r-} dS_r^{t,u,z} \right)^2 \right] - v^f(t,u,p,z,x) \\ = & \mathbb{E}\left[\left(((f - f_n) + f_n)\left(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}\right) - x - \int_t^T \theta^n_r dS_r^{t,u,z} \right)^2 \right] - v^f(t,u,p,z,x) \\ \leq & \mathbb{E}\left[(f - f_n)^2\right] + 2\mathbb{E}\left[(f - f_n)^2\right]^{1/2} v^{f_n}(t,u,p,z,x)^{1/2} \\ \leq & M(1 + |x|) \left\|f_n - f\right\|_\infty + |v^{f_n} - v^f|\end{aligned}$$

and by using the above estimation on $|v^{f_n} - v^f|$ we deduce that, for some positive constant M which depends on t, u, p, z and x we find

$$\begin{aligned}&\left| v^f(t,u,p,z,x) - \mathbb{E}^{\mathbb{P}}\left[\left(f(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}) - x - \int_t^T \theta^n_{r-} dS_r^{t,u,z} \right)^2 \right] \right| \\ \leq & M \left\|(f_n - f)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2\end{aligned}$$

which concludes our proof.

□

Remark 5.9. *Lemma 5.8 can be improved since, as stated in Lemma 5.3, we have* $e^{-CT} \leq a(t,u,z)$ *for some positive constant* C, *so then*

$$|x^*(f_1)(t,u,p,z) - x^*(f_2)(t,u,p,z)| \leq e^{CT/2} \left\|(f_1 - f_2)(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z})\right\|_2$$

The message coming from the above Lemma is really interesting: one can "replace" a potentially non-smooth payoff function f with some smooth functions f_n, by controlling the error on the value function and the optimal quadratic hedge price by $\|f - f_n\|_2$.

5.5 The pure investment problem: verification

When $f = 0$ in problem (5.11) we have seen that the value function is given by $v^0(t,u,z,x) = a(t,u,z)x^2$. We can characterize the function a as the solution of a semi linear partial integro-differential equation (PIDE). For this, let us introduce the differential operators associated to the process (U,P,Z):

Definition 5.10. *Let $\varphi : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ be a smooth function. We denote*

$$\begin{aligned}
\mathcal{A}_t\varphi :=& -\left(\mu\frac{\partial\varphi}{\partial z} + \mu^U\frac{\partial\varphi}{\partial u} + \mu^P\frac{\partial\varphi}{\partial p}\right) \\
& -\frac{1}{2}\left(\sigma^2\frac{\partial^2\varphi}{\partial z^2} + (\sigma^P)^2\frac{\partial^2\varphi}{\partial p^2} + (\sigma^U)^2\frac{\partial^2\varphi}{\partial u^2} - 2\lambda\sigma\sigma^P\frac{\partial^2\varphi}{\partial p\partial z}\right) \\
\mathcal{B}_t\varphi :=& \int_{\mathbb{R}}\left(\varphi(t,u,p+\gamma^P,z+\gamma) - \varphi(t,u,p,z) - \left(\gamma^P\frac{\partial\varphi}{\partial p} + \gamma\frac{\partial\varphi}{\partial z}\right)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy) \\
& + \int_{\mathbb{R}}\left(\varphi(t,u+\gamma^U,p,z) - \varphi(t,u,p,z) - \gamma\frac{\partial\varphi}{\partial u}\mathbb{1}_{\{|y|\le 1\}}\right)\nu_n(dy)
\end{aligned}$$

$$\begin{aligned}
\mathcal{Q}_t\varphi :=& \tilde{\mu}\varphi + \sigma^2\frac{\partial\varphi}{\partial z} + \lambda\sigma\sigma^P\frac{\partial\varphi}{\partial p} \\
& + \int_{\mathbb{R}}(e^\gamma - 1)\left(\varphi(t,u,p+\gamma^P,z+\gamma) - \varphi(t,u,p,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy) \\
\mathcal{G}_t\varphi :=& \sigma^2\varphi + \int_{\mathbb{R}}(e^\gamma - 1)^2\varphi(t,u,p,z+\gamma)\nu(dy)
\end{aligned}$$

where μ stands for $\mu(t,u,z)$ and so on.

We introduce the functional spaces with which we will work throughout this section:

□ $C^{l/2,l}([0,T]\times\mathbb{R}^3)$, the Hölder space of type 1, defined in Paragraph C.2, where $l \in [0\ ,\ 3)\setminus\{1,2\}$.
□ $H^l([0,T]\times\mathbb{R}^3)$, the Hölder space of type 2, defined in Appendix C, paragraph C.3, where $l \in [0,3)$.

Recall that $C^{l/2,l}([0,T]\times\mathbb{R}^3) \subset H^l([0,T]\times\mathbb{R}^3)$. With the following theorem we characterize the function a as the solution of a semi linear PIDE, provided that it has a unique smooth solution. This procedure is also known as the verification, which in general is the "easiest" part in a stochastic optimization problem.

Theorem 5.11. *Let Assumptions 5.1 hold true and $T < T^*$ as stated in Theorem* (5.4)*. Let also*

$$\mathcal{H}[\varphi] := \inf_{|\pi|\le\bar{\Pi}}\left\{2\pi\mathcal{Q}\varphi + \pi^2\mathcal{G}\varphi\right\} \tag{5.43}$$

where

$$\bar{\Pi} := \frac{e^{CT}}{\sigma^2_{min}\vee|\Gamma|}C_e\left(1 + K^a_{lip}\right) \tag{5.44}$$

and C_e is given in (5.22)*.*

Case $\sigma_{min} > 0$

Assume that the PIDE

$$0 = -\frac{\partial \varphi}{\partial t} + \mathcal{A}_t\varphi - \mathcal{B}_t\varphi - \mathcal{H}[\varphi] \quad \varphi(T,u,z) = 1 \tag{5.45}$$

has a unique solution $\varphi \in C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$ *for some* $\kappa \in (0,1)$*, which also is strictly positive.*

Case $\sigma_{max=0}$

Suppose that the intensity measure $\nu(dy)$ *(respectively* $\nu_n(dy)$*) has a density w.r.t. the Lebesgue measure:* $\nu(dy) = g(y)|y|^{-(1+\alpha)}dy$ *(respectively* $\nu_n(dy) = g_n(y)|y|^{-(1+\alpha)}dy$*), where* $\alpha \in (1,2)$ *and* g *(respectively* g_n*) is a bounded, measurable and positive function. Assume that the PIDE* (5.45) *has a unique solution* $\varphi \in H^{\alpha+\kappa}([0,T]\times\mathbb{R}^2)$*, for some* $\kappa \in (0,1)$*, which also is strictly positive and continuously differentiable w.r.t.* t*.*

Then $\varphi = a$ *defined in* (5.13) *and the optimal strategy in problem* (5.12) *is given by*

$$\theta^*_t = e^{-Z_{t-}}\pi^*(t,U_{t-},Z_{t-})\,X^{\theta^*}_{t-}, \qquad X^{\theta^*}_t := x + \int_0^t \theta^*_{r-}dS_r \tag{5.46}$$

where

$$\pi^*(t,u,z) := -\frac{\mathcal{Q}_t a(t,u,z)}{\mathcal{G}_t a(t,u,z)} \tag{5.47}$$

is the minimizer in the operator $\mathcal{H}$*.*

Proof.
We start with the case $\sigma_{min} > 0$. Let then $\varphi \in C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$ be the unique solution of (5.45) and define $w(t,z,x) := x^2\varphi(t,u,z)$. Take now the minimizing sequence θ^n of the problem (5.12) introduced in (5.39). If X^n is the wealth process stated in (5.9) corresponding to θ^n then, from Itô's formula, we obtain

$$\begin{aligned}
&\mathbb{E}\left[w(t+h,U_{t+h},Z_{t+h},X^n_{t+h})\right] = w(t,u,z,x) + \mathbb{E}\left[\int_t^{t+h}(X^n_{s-})^2\frac{\partial\varphi}{\partial t}(s,U_sZ_s)ds\right] + \\
&\mathbb{E}\left[\int_t^{t+h}(X^n_{s-})^2\left(-\mathcal{A}_s\varphi + \mathcal{B}\varphi\right)(s,U_{s-},Z_{s-})ds\right] + \\
&\mathbb{E}\left[\int_t^{t+h}(\theta_{s-}e^{Z_{s-}})\left(2\tilde{\mu}X^n_{s-}\varphi + \sigma^2\theta_{s-}e^{Z_{s-}}\varphi + 2\sigma^2X^n_{s-}\frac{\partial\varphi}{\partial z}\right)(s,U_s,Z_s)ds\right] + \\
&\mathbb{E}\left[\int_t^{t+h}(\theta_{s-}e^{Z_{s-}})^2\int_{\mathbb{R}}(e^\gamma-1)^2\varphi(s,U_s,Z_{s-}+\gamma)\nu(dy)ds\right] + \\
&\mathbb{E}\left[\int_t^{t+h}2\theta_{s-}e^{Z_{s-}}X^n_{s-}\int_{\mathbb{R}}(e^\gamma-1)\left(\varphi(s,U_s,Z_{s-}+\gamma)-\varphi(s,U_s,Z_{s-})\mathbb{1}_{\{|y|<1\}}\right)\nu(dy)ds\right]
\end{aligned}$$

where (U,Z,X^n) stands for $(U^{t,u},Z^{t,u,z},X^{t,u,z,n})$. Since $\varphi \in C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$ (which implies that φ and its derivatives are bounded) we can omit the martingale part after taking expectation. Fix now (s,ω): if $X^n_{s-}(\omega)=0$ then the expression in the expectation above becomes

$$\sigma^2(\theta_s e^{Z_s})^2\varphi(s,U_s,Z_s)+(\theta_{s-}e^{Z_{s-}})^2\int_{\mathbb{R}}(e^\gamma-1)^2\,\varphi(s,U_s,Z_{s-}+\gamma)\nu(dy)\geq 0$$

since φ is positive. If $X^n_{s-}(\omega)\neq 0$ we can introduce $\pi^n:=\theta^n_s e^{Z_{s-}}/X_{s-}$ and, after simple calculations, we deduce that the expression in the expectation turns out to be non negative (since φ verifies the PIDE (5.45) and $|\pi^n|\leq\bar{\Pi}$ as proved in (5.38)). We deduce

$$\mathbb{E}\left[w(t+h,U_{t+h},Z_{t+h},X^n_{t+h})\right]\geq w(t,u,z,x)$$

The continuity of w gives $w(t+h,U_{t+h},Z_{t+h},X^n_{t+h})\to\left(x+\int_t^T\theta^n_{r-}dS_r\right)^2$ when $h\to T-t$. Since φ is bounded we also have

$$\left|w(t+h,U_{t+h},Z_{t+h},X^n_{t+h})\right|\leq\|\varphi\|_\infty\sup_{s\in[t,T]}|X^n_s|^2\in\mathbb{L}^1(\mathbb{P})$$

This is true since X^n is a stochastic exponential (from the definition of θ^n in (5.39)) and π^n is bounded:

$$dX^n_r=\pi^n_{r-}X^n_{r-}e^{-Z_{r-}}d\exp(Z_r)$$

so that we can apply Lemma 3.1 in Pham (1998) to prove that $\sup_{s\in[t,T]}|X^n_s|^2\in\mathbb{L}^1(\mathbb{P})$. Dominated convergence yields $w(t,u,z,x)\leq\mathbb{E}\left[\left(x+\int_t^T\theta^n_{r-}dS_u\right)^2\right]$. Since θ^n is a minimizing sequence in problem (5.12), by taking the limit $n\to\infty$ we finally obtain $w(t,u,z,x)\leq v_0(t,u,z,x)$. To prove the equality, we use the strategy θ^* in (5.46), which is admissible since π^* is bounded, and then it belongs to $\mathcal{X}(t,u,p,z)$. Remark that

$$\|\pi^*\|_\infty\leq\frac{e^{CT}}{\sigma^2_{min}\vee|\Gamma|}\left(\|\tilde{\mu}\|+K^a_{lip}\sigma^2_{max}+\|\tau\|_{1,\nu}+K^a_{lip}\|\tau\|^2_{2,\nu}\right)\leq\bar{\Pi}$$

With a similar argument as before we can prove that

$$\mathbb{E}\left[w(t+h,U_{t+h},Z_{t+h},X^{\theta^*}_{t+h})\right]=w(t,u,z,x)$$

and then, by letting $h\to T-t$, we deduce

$$w(t,u,z,x)=\mathbb{E}\left[\left(x+\int_t^T\bar{\theta}_{r-}dS_u\right)^2\right]\geq v_0(t,u,z,x)$$

which implies $w(t,u,z,x)=v_0(t,u,z,x)$. We conclude $\varphi=a$ and θ^* is the optimal policy for the stochastic control problem in (5.12).
If $\sigma_{max}=0$ then (U,Z) is a pure jump process. In this case, one can use the Itô's formula for pure jump process given in Appendix D and repeat the same argument.

□

Theorem 5.11 characterizes the function a as the solution of PIDE (5.45), provided that we can prove that it has a unique smooth solution which also is strictly positive. In Chapter 6 we will prove that this PIDE has a unique smooth solution when $\sigma_{min} > 0$, whereas Chapter 7 is devoted to the analysis of this PIDE when $\sigma_{max} = 0$.

Remark 5.12. *The function a does not depend on the variable p, so the operators in Definition 5.10 appearing in PIDE* (5.45) *can be simplified.*

Remark 5.13. *If S is a martingale under the historical probability $\mathbb{P}$ then from Lemma 5.2 we obtain*

$$a(t,u,z) = \mathbb{E}^{\mathbb{P}}\left[1 + \int_t^T \theta^*_{r-} dS^{t,u,z}_r\right] = 1$$

This fact can be also seen on the HJB equation (5.45)*: by applying Itô's formula to S we find*

$$\mu + \frac{1}{2}\sigma^2 + \int \left(e^\gamma - 1 - \gamma \mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy) = 0$$

since it is a martingale. It follows that the non linear operator $\mathcal{H}$ in (5.43) *will only depend on ∂a, the first derivative of a, and then it is straightforward to deduce the unique solution of PIDE 5.45 is given by $a = 1$.*

5.6 The quadratic hedge problem: verification

As in Section 5.5, our aim is to characterize the function v given in (5.11) as the unique solution of a PIDE. From (5.41) we know that v_f has the following structure

$$v^f(t,u,p,z,x) = a(t,u,z)x^2 + b(t,u,p,z)x + c(t,u,p,z)$$

We already know that a verifies the PIDE (5.45) if it is smooth enough to apply Itô's formula.

Theorem 5.14. *Let Assumptions 5.1 hold true and $T < T^*$ as stated in Theorem* (5.4)*. Suppose that f is continuous and that the PIDE* (5.45) *has a unique smooth solution which also is strictly positive.*

Case $\sigma_{min} > 0$
Assume that the PIDEs

$$0 = -\frac{\partial b}{\partial t} + \mathcal{A}_t b - \mathcal{B}_t b - \pi^* \mathcal{Q}_t b, \qquad b(T,.) = -2f \qquad (5.48)$$

$$0 = -\frac{\partial c}{\partial t} + \mathcal{A}_t c - \mathcal{B}_t c + \frac{1}{4}\frac{(\mathcal{Q}_t b)^2}{\mathcal{G}_t a}, \qquad c(T,.) = f^2 \qquad (5.49)$$

have a smooth solutions $b, c \in C^{1+\kappa/2+1, 2+\kappa}([0,T]\times\mathbb{R}^3)$ for some $\kappa \in (0,1)$, where π^* is given in (5.47).

Case $\sigma_{max=0}$

Suppose that the intensity measure $\nu(dy)$ (respectively $\nu_n(dy)$) has a density w.r.t. the Lebesgue measure: $\nu(dy) = g(y)|y|^{-(1+\alpha)}dy$ (respectively $\nu_n(dy) = g_n(y)|y|^{-(1+\alpha)}dy$), where $\alpha \in (1,2)$ and g (respectively g_n) is a bounded, measurable and positive function. Assume that the PIDEs (5.48)–(5.49) have a unique solutions $b, c \in H^{\alpha+\kappa}([0,T] \times \mathbb{R}^3)$, for some $\kappa \in (0,1)$, which also are continuously differentiable w.r.t. t.

Then the value function of the problem (5.11) is

$$v^f(t,u,p,z,x) = a(t,u,z)x^2 + b(t,u,p,z)x + c(t,u,p,z)$$

where a is the unique smooth solution of (5.45), whereas b, c are, respectively, the unique smooth solution of PIDES (5.48)–(5.49). Furthermore the optimal strategy in problem (5.11) is given by

$$\begin{aligned} \theta_t^* :=& e^{-Z_{t-}}\left(\pi^*(t,U_{t-},Z_{t-})X_{t-}^{\theta^*} - \frac{1}{2}\frac{\mathcal{Q}_t b}{\mathcal{G}_t a}(t,U_{t-},P_{t-},Z_{t-})\right) \\ X_t^{\theta^*} :=& x + \int_0^t \theta_{r-}^* dS_r \end{aligned} \tag{5.50}$$

Proof.
As in the proof of Theorem 5.11, let us consider the case $\sigma_{min} > 0$. We start with $w_f(t,u,p,z,x) := a(t,u,z)x^2 + b(t,u,p,z)x + c(t,u,p,z)$, where a is solution of (5.45), whereas b and c are, respectively, the solutions of (5.48) and 5.49. Since $a > 0$ then

$$\begin{aligned} 0 &= x^2\left[\frac{\partial a}{\partial t} - \mathcal{A}_t a + \mathcal{B}_t a\right] + x\left[\frac{\partial b}{\partial t} - \mathcal{A}_t b + \mathcal{B}_t b\right] + \left[\frac{\partial c}{\partial t} - \mathcal{A}_t c + \mathcal{B}_t c\right] \\ &+ \inf_{\theta\in\mathbb{R}}\left[\theta e^z\left(2\mathcal{Q}_t a x + \mathcal{Q}_t b\right) + \theta^2 e^{2z}\mathcal{G}_t a\right] \end{aligned}$$

Let now $\theta \in \mathcal{X}(t,u,p,z)$ and apply Itô's formula to $w_f(t+h, U_{t+h}, P_{t+h}, Z_{t+h}, X_{t+h}^\theta)$. We skip the computations, since they are similar to the ones we did in the proof of Theorem 5.11: by using the continuity of f we obtain

$$w_f(t,u,p,z,x) \le \mathbb{E}\left[\left(f\left(U_T^{t,u}, P_T^{t,p}, Z_T^{t,u,z}\right) - X_T^{t,u,z,x,\theta}\right)^2\right]$$

From the arbitrariness of θ we deduce $w_f(t,u,p,z,x) \le v_f(t,u,p,z,x)$. The equality is obtained by using θ^* in (5.50). When $\sigma_{max} = 0$ one can use Itô's formula for pure jump processes stated in Appendix D to complete the proof.

□

The above result proves that one can characterize the value function of problem (5.11) by solving a triplet of PIDEs, provided that they have a unique smooth solution. This system has a "triangular" structure: the first one, which is semi linear, only depends on the function a; the second one is linear in b whenever we know a and the third one is again linear in c when we know a and b. The optimal strategy of problem (5.11) has an affine structure

$$\theta_t^* := e^{-Z_{t-}} \left(\pi^*(t, U_{t-}, Z_{t-}) X_{t-}^{\theta^*} - \frac{1}{2} \frac{\mathcal{Q}_t b}{\mathcal{G}_t a}(t, U_{t-}, P_{t-}, Z_{t-}) \right)$$

where the multiplier π^* does not depend on the particular form of the payoff profile f, so it is universal. This simplifies the implementation of the above strategy to solve problem (5.11): firstly one computes the function a, which does not depend on f, by solving a semi linear PIDE and then one only has to solve a linear PIDE for the function b. This also allows to compute the hedge ratios for different options at the same time once one has computed the function a.

In general it is not possible to find explicit solution for PIDE (6.1), so one has to employ numerical schemes. However the triangular structure for the functions a, b and c largely simplifies the problem: first one computes numerically the function a and then uses it to compute the function b, which is relatively simple since it solves a linear PIDE. Remark as well that in order to compute the optimal strategy and the optimal quadratic hedge price (defined in (5.42)) one does not need the function c.

5.7 Viscosity solutions

In Sections 5.5 and 5.6 we characterized the value functions a and v^f in terms of solution of PIDEs; however, at this point, we do not know whenever the functions a, b and c are smooth or not, and in general semi linear partial differential equations do not have smooth solutions. To give an idea of this difficulty, we would like to recall a really simple example taken from Cannarsa and Sinestrari (2004), which explains how even relatively simple non linear differentiable equations may fail to have smooth solutions. Let $u : [-1, 1] \to \mathbb{R}$ a smooth function verifying $u(-1) = u(1) = 0$ and $(u')^2 = 1$ in $(-1, 1)$. If such a u exists then one can find $x_0 \in (-1, 1)$ with $u'(x_0) = 0$, which contradicts the fact that $(u'(x_0))^2 = 1$. So there is no solution to the above problem. However one can easily check that the function $u(x) = |x| - 1$ verifies the above equation everywhere except at $x = 0$. For many applications it may be sufficient to know that the problem above has a solution, provided that one gives a precise sense of what is a non-smooth solution of a differential problem, or, in other word, provided that one relaxes the notion of classical solution. Other examples of this type, issued in particular from stochastic optimization problems (Pham, 2007), prove that the classical notion of solution of a differential equation was too restrictive and not appropriate for a wide class of interesting differential problems.

The general theory of viscosity solutions is a well adapted context in which one can give a precise sense for a non-smooth solution of a differential equation. Especially in the case of stochastic control problems and related differential equations, the notion of viscosity solutions allows to characterize in a unique way the value function.

Nowadays the theory of viscosity solutions has been highly developed and many references can be found in the literature. To our knowledge the notion of viscosity solution in the diffusion case has been introduced by Lions (1983). A complete review of this theory has been done in Crandall, Ishii, and Lions (1992) to sum up ten years of intense research on the viscosity solutions theory applied to second order semi linear partial differential equations. These ideas were afterwards adapted to semi linear partial integro-differential equations, as for example PIDE (5.45) (see Barles, Buckdahn, and Pardoux (1997); Pham (1998); Jakobsen and Karlsen (2005); Barles and Imbert (2008) and references therein). Let us start with the basic definition of viscosity solution for a second order parabolic PIDEs:

Definition 5.15. *Let $A \subseteq \mathbb{R}^n$ a possibly unbounded domain and $T < \infty$. Consider the PIDE*

$$\frac{\partial v}{\partial t} + F(t, w, v(w), Dv, D^2 v, v(.)) = 0,\ (t, w) \in [0, T) \times A$$
$$v = \psi \text{ on the parabolic boundary } [0, T] \times \partial A \cup \{T\} \times A$$

where $F : [0, T) \times A \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \times C^2_p([0, T) \times A)$ is a given functional, $\psi : A \to \mathbb{R}$. and $C^2_p([0, T) \times A)$ denotes the space of twice continuously differentiable functions with polynomial growth at infinity with power p. A locally bounded map v is a viscosity sub-solution (resp. super-solution) of the above PIDE if for any $w \in A$ and any $\varphi \in C^2_p([0, T) \times A)$ such that w is a local maximizer of $v^* - \varphi$ (resp. minimizer of $v_* - \varphi$) one has

$$\frac{\partial \varphi}{\partial t} + F(t, w, v^*(w), D\varphi(w), D^2\varphi(w)v, \varphi(.)) \leq 0, \quad \textit{(sub-solution)}, \quad v^*(T, .) = \psi^*(T, .)$$
$$\frac{\partial \varphi}{\partial t} + F(t, w, v_*(w), D\varphi(w), D^2\varphi(w)v, \varphi(.)) \geq 0, \quad \textit{(super-solution)}, \quad v_*(T, .) = \psi_*(T, .)$$

where v^* (resp v_*) is the upper (resp lower) semi-continuous envelope of v . A map v is a viscosity solution if it is at the same time a super-solution and sub-solution of the above PIDE.

Remark 5.16. *We omitted to list some necessary assumptions on the functional F. One can find them, for example, in Jakobsen and Karlsen (2006) or Barles and Imbert (2008) when the non local component is a Lévy operator as $\mathcal{B}$ given in definition 5.10.*

It is clear at this point how the notion of viscosity solution relaxes the classical definition of smooth solution for PIDEs: according to the above definition, one only asks that the solution has to be locally bounded. Equivalent definitions of viscosity solution can be stated in terms of the so called *sub* and *superjet* (see for this Crandall, Ishii, and Lions (1992) where there is no non-local component, or Pham (1998) when a non local component is allowed). Proving the existence of a viscosity solution is a relatively simple task under mild conditions on the functional F. The main difficulties arise when one wants to prove its uniqueness: in many cases this is done by proving a so called "comparison theorem" or "maximum principle", which allows us to compare a sub-solution and a super-solution on the entire domain. In its general form this is stated as follows:

> if v is a sub-solution and v' is a super-solution with $v^* \leq v'_*$ on the parabolic boundary of $[0, T] \times A$ then
>
> $$v^* \leq v'_* \text{ on } [0, T] \times \bar{A}$$

This comparison result depends on the form of the functional F. In the literature a key tool to prove the above comparison theorem is given by the Jensen-Ishii Lemma (Jensen, 1988; Ishii, 1989) and its extension to more general cases (see for example Jakobsen and Karlsen (2006); Pham (1998); Barles and Imbert (2008)). We point out that, although general comparison principles are stated for many types of functionals F, the uniqueness problem remains an open problem in many situations, even when the non-local component is not allowed. We can say that the viscosity approach shifts the difficulty of proving the existence of a solution (which was the hard task in classical setting) to its uniqueness.

The viscosity solutions theory turns out to be well adapted to stochastic optimization problems when one can prove a so called *dynamic programming principle*. In the rest of the section we will present it for the problem (5.12). Since what follows is not fundamental for our work, we will do it in the simple case where Z does not depend on U. The principle can be stated as follows:

Dynamic programming principle: for any $h > 0$ and $(t, z, x) \in [0, T) \times \mathbb{R}^3$

$$v^0(t, z, x) = \inf_{\theta \in \mathcal{X}_{t+h}(t,z,x)} \mathbb{E}\left[v^0\left(t + h, Z^{t,u,z}_{t+h}, X^{t,u,z,x,\theta}_{t+h}\right)\right] \tag{5.51}$$

where $\mathcal{X}_{t+h}(t, z, x)$ is the set of admissible controls on the time window $[t, t+h]$. To prove the above principle one needs some a priori regularity on the value function v^0 (or equivalently on the function a). For example, if we could prove that the function a is continuous then (5.51) holds true by using Proposition 3.2 in Pham (1998). We do not insist on this but we just focus on the fact that when the admissible strategies are not bounded, as in our case, proving the dynamic programming principle is a delicate task, in particular when the state variable process can jump.

Fix (t, z, x) and let ψ be a smooth function with appropriate polynomial growth such that

$$0 = (v^0 - \psi)(t, z, x) = \sup_{[0,T]\times\mathbb{R}^2} (v^0 - \psi)(t', z', x')$$

It follows

$$0 \le \inf_{\theta \in \mathcal{X}_{t+h}(t,z,x)} \mathbb{E}\left[\psi\left(t + h, Z^{t,z}_{t+h}, X^{t,z,x,\theta}_{t+h}\right) - \psi(t, z, x)\right]$$

We can now apply Itô's formula and, by letting $h \to 0^+$, we obtain

$$0 \ge -\partial_t\psi - \mu\partial_z\psi - \frac{1}{2}\sigma^2\partial_z^2\psi - \Bigg(\theta e^z\tilde{\mu}\partial_x\psi + \frac{1}{2}\sigma^2(\theta e^z)^2\partial_x^2\psi + \sigma^2\theta e^z\partial^2_{zx}\psi +$$

$$\int \left(\psi(t, z + \gamma, x + \theta e^z(e^\gamma - 1)) - \psi(t, z, x) - (\gamma\partial_z\psi + \theta e^z(e^\gamma - 1))\partial_x\psi\right) \mathbb{1}_{\{|y|\le 1\}}\nu(dy)\Bigg)$$

for any $\theta \in \mathbb{R}$. From the fact that $v^0 = x^2 a(t, z)$ we deduce that

$$-\frac{\partial\varphi}{\partial t}(t, z) + \mathcal{A}\varphi(t, z) - \mathcal{B}\varphi(t, z) - \inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}\varphi(t, z) + \pi^2\mathcal{G}\varphi(t, z)\right\} \le 0$$

for any smooth φ verifying $0 = (a - \varphi)(t, z) = \sup(a - \varphi)(t', z')$. Remark that since $\varphi \ge a$ then

$$\inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}\varphi(t, z) + \pi^2\mathcal{G}\varphi(t, z)\right\} > -\infty$$

It follows that the function a is a viscosity sub-solution of

$$-\frac{\partial a}{\partial t}+\mathcal{A}a-\mathcal{B}a-\inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}a+\pi^2\mathcal{G}a\right\}=0$$

For the super-solution one has to take care of the fact that $\inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}\varphi+\pi^2\mathcal{G}\varphi\right\}$ can take the value $-\infty$: in fact the test function φ is bounded from above by a and then we do not know whenever $\mathcal{G}\varphi(t,z)$ is strictly positive or not. We do not detail it but it can be proved that a is a super-solution and then a viscosity solution of

$$\max\left(-\frac{\partial a}{\partial t}+\mathcal{A}a-\mathcal{B}a-\inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}a+\pi^2\mathcal{G}a\right\},\inf_{\pi\in\mathbb{R}}\left\{2\pi\mathcal{Q}a+\pi^2\mathcal{G}a\right\}\right)=0 \quad (5.52)$$

Remark that at this point we never used any strict elliptic condition on the volatility coefficient of Z. This is one of the reasons to use the viscosity approach when the coefficient σ of the processes in (5.5) may be degenerate. We do not discuss here the uniqueness of the (viscosity) solution of (5.52) since we will not use this PIDE in the sequel.

We conclude this Section with the following remark: the viscosity solution approach is a powerful tool to characterize the value function of a stochastic optimization problem, provided that one can prove the dynamic programming principle (for the existence) and a comparison principle (for the uniqueness). However finding the optimal control remains a open problem: from (5.47) we deduce that this optimal control depends on the derivative of the value function which is not defined if the value function is not differentiable.

Chapter 6

Smooth solutions: the jump-diffusion case

The Chapter is organized as follows: we start by giving some regularity properties of the differential operators introduced in Definition 5.10 (Section 6.1). To prove the regularity of the value function a, we introduce an iterative sequence (Paragraph 6.2.1), for which we first give some fundamental a priori properties (Paragraph 6.2.2), and then we prove its convergence to the function a in an appropriate Hölder space (Paragraph 6.2.3). We also compare our methods with two other ones, which make use of, respectively, BSDEs theory and Sobolev spaces, that may be used in some cases to prove that the function a is smooth. We conclude the Chapter with Section 6.3 where we prove that the value function of the quadratic hedge problem is smooth.

Contents

6.1 Hölder regularity of the differential operators 101
6.2 Smoothness and characterization of the function a . . . 106
6.2.1 The approximation sequence 106
6.2.2 Weak convergence and uniqueness 107
6.2.3 Characterization of the function a 111
6.2.4 Comments 115
6.3 Smoothness and characterization of the function v^f . . . 118

6.1 Hölder regularity of the differential operators

This chapter is devoted to the analysis of the PIDE (5.45) when $\sigma_{min} > 0$. We will always assume that Assumptions 5.1 hold true together with the following:

Assumption 6.1.
*[C]-**The coefficients**. There exists some $m \geq 0$ such that for all $t, t' \in [0,T]$ and $u, z, y \in \mathbb{R}$ we have*

$$\big|\mu(t,u,z) - \mu(t',u,z)\big| + \big|\sigma(t',u,z) - \sigma(t',u,z)\big| \leq m|t-t'|$$
$$\big|\gamma(t,u,z,y) - \gamma(t',u,z,y)\big| \leq m\rho(y)|t-t'|$$

and the same holds true for the coefficients of U (respectively P) where one uses positive constants m^U and ρ^U (resp. $0 \leq m^P$ and ρ^P).
[I]-Integrability condition *There exists some $\delta \in (0,1)$ such that*

$$\int_{|y|\leq 1} \left(\tau^{2-\delta}(y) + (\tau^P(y))^{2-\delta}\right) \nu(dy) + \int_{|y|\leq 1} (\tau^U(y))^{2-\delta} \nu_n(dy) < +\infty$$

[ND]-No degeneracy*. There volatility functions are uniformly bounded from below by some positive constant: $0 < \sigma_{min} \leq \sigma_{max}$.*

Recall that, according to Theorem 5.11, we need to prove that the following PIDE

$$0 = -\frac{\partial a}{\partial t} + \mathcal{A}a - \mathcal{B}a - \mathcal{H}[a], \qquad a(T,u,z) = 1 \tag{6.1}$$

has a unique smooth solution which also is strictly positive. We also recall the functional spaces used throughout this chapter:

□ $C^{l/2,l}([0,T] \times \mathbb{R}^3)$, the Hölder space of type 1, defined in Paragraph C.2, where $l \in [0\ ,\ 3) \setminus \{1,2\}$.

□ $H^l([0,T] \times \mathbb{R}^3)$, the Hölder space of type 2, defined in Appendix C, paragraph C.3, where $l \in [0,3)$.

Remark that $C^{l/2,l}([0,T] \times \mathbb{R}^3) \subset H^l([0,T] \times \mathbb{R}^3)$. In the rest of the chapter $\| \|_{l/2,l}$ denotes then the Hölder norm relatively to the Hölder space of type 1 (see definition (C.3)), whereas $\| \|_{l,H}$ is relative to the Hölder space of type 2 (see definition C.7).

In this section we will study the operators defined in Definition 5.10 and we prove that they are Lipschitz continuous in their appropriate Hölder space. This regularity is needed is order to prove that PIDE (6.1) has a unique smooth solution.

Lemma 6.2. *Suppose that Assumptions 5.1-*[**C1**, **I1**] *and Assumptions 6.1-***C** *hold true and fix $\kappa \in (0,1)$. Then*

$$\mathcal{B}, \mathcal{Q}, \mathcal{G} : C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^3) \to C^{\kappa/2,\kappa}([0,T] \times \mathbb{R}^3)$$

Moreover there exist a positive constant $M > 0$ and two functions $\varrho, \varsigma : (0,1) \to \mathbb{R}^+$ such that for all $\epsilon \in (0,1)$, $r \in (0,1)$

$$\begin{aligned} \|\mathcal{Q}\varphi\|_{\kappa/2,\kappa} + \|\mathcal{G}\varphi\|_{\kappa/2,\kappa} &\leq M \left(\epsilon^{1-\kappa} \|\varphi\|_{\kappa/2+1,\kappa+2} + \epsilon^{-(1+\kappa)} \|\varphi\|_\infty\right) \\ \|\mathcal{B}\varphi\|_{\kappa/2,\kappa} &\leq M \left(\left(\varrho(r) + \epsilon^{1-\kappa}\varsigma(r)\right) \|\varphi\|_{\kappa/2+1,\kappa+2} + \epsilon^{-(1+\kappa)}\varsigma(r) \|\varphi\|_\infty\right) \end{aligned}$$

for all $\varphi \in C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^3)$.

Let now $\delta \in (0,1)$ given in Assumption 6.1-[**I**]*, which is now supposed to hold true: there exists a positive constant $M > 0$ such that for all $\epsilon, r \in (0,1)$ and for all*

$\varphi \in H^{2-\delta}([0,T] \times \mathbb{R}^3)$ *one has*

$$\|\mathcal{Q}\varphi\|_\infty + \|\mathcal{G}\varphi\|_\infty \le M \|\varphi\|_{1,H}$$
$$\|\mathcal{B}\varphi\|_\infty \le M \left(\left(\varrho(r) + \varsigma(r)\epsilon^{1-\delta} \right) \|\varphi\|_{2-\delta,H} + \varsigma(r)\epsilon^{-1} \|\varphi\|_{\infty,H} \right)$$

The constant M does not depend on φ, ϵ or r. Furthermore $\varrho(r) \to 0$ when $r \to 0$ whereas $\varsigma(r) \to +\infty$ if τ or τ^U are not integrable around zero, respectively, w.r.t. $\nu(dy)$ and $\nu_n(dy)$.

Proof.
From now on $M > 0$ denotes a positive constant which only depend on the market parameters given in Assumptions 5.1 but not on φ. It may also change from line to line.

For the operator $\mathcal{Q}$ we can use the definition of $\tilde{\mu}$ given in Assumptions 5.1 to rewrite it in the following form

$$\mathcal{Q}\varphi := \left(\mu + \frac{1}{2}\sigma^2 + \int \left(e^\gamma - 1 - \gamma \mathbb{1}_{\{|y|\le 1\}} \right) \nu(dy) \right) \varphi + \sigma^2 \frac{\partial \varphi}{\partial z} + \lambda \sigma \sigma^P \frac{\partial \varphi}{\partial p}$$
$$+ \int_0^1 d\theta \int_{|y|\le 1} (e^\gamma - 1) \left(\gamma \frac{\partial \varphi}{\partial z}(t,u,p+\theta\gamma^P, z+\theta\gamma) + \gamma^P \frac{\partial \varphi}{\partial p}(t,u,p+\theta\gamma^P, z+\theta\gamma) \right) \nu(dy)$$

Hence it is straightforward to deduce

$$\|\mathcal{Q}\varphi\|_{\kappa/2,\kappa} \le M \left(\|\varphi\|_{\kappa/2,\kappa} + \|D_x\varphi\|_{\kappa/2,\kappa} \right) \le M \|\varphi\|_{(\kappa+1)/2,\kappa+1}$$

and then we use Proposition C.2 to conclude

$$\|\mathcal{Q}\varphi\|_{\kappa/2,\kappa} \le M \left(\epsilon^{1-\kappa} \|\varphi\|_{\kappa/2+1,\kappa+2} + \epsilon^{-(1+\kappa)} \|\varphi\|_\infty \right)$$

For $\mathcal{G}$ we obtain

$$\|\mathcal{G}\varphi\|_{\kappa/2,\kappa} \le M \left(\epsilon^{1-\kappa} \|\varphi\|_{\kappa/2+1,\kappa+2} + \epsilon^{-(1+\kappa)} \|\varphi\|_\infty \right)$$

From the above definition of $\mathcal{Q}$ it is straightforward to see that $\|\mathcal{Q}\varphi\|_\infty \le M \|\varphi\|_{1,H}$ and also $\|\mathcal{G}\varphi\|_\infty \le M \|\varphi\|_\infty \le M \|\varphi\|_{1,H}$

We can start our analysis of $\mathcal{B}$. We first write $\mathcal{B}_t\varphi := \mathcal{I}_t\varphi + \mathcal{J}_t\varphi$ where

$$\mathcal{I}_t\varphi := \int_{\mathbb{R}} \left(\varphi(t,u,p+\gamma^P, z+\gamma) - \varphi(t,u,p,z) - \left(\gamma^P \frac{\partial \varphi}{\partial p} + \gamma \frac{\partial \varphi}{\partial z} \right) \mathbb{1}_{\{|y|\le 1\}} \right) \nu(dy)$$
$$\mathcal{J}_t\varphi := \int_{\mathbb{R}} \left(\varphi(t,u+\gamma^U,p,z) - \varphi(t,u,p,z) - \gamma^U \frac{\partial \varphi}{\partial u} \mathbb{1}_{\{|y|\le 1\}} \right) \nu_n(dy)$$

If we prove the result for $\mathcal{I}$ and $\mathcal{J}$ then it also holds true for $\mathcal{B}$ by applying triangular inequality. Since similar computations can be done for $\mathcal{I}$ and $\mathcal{J}$, we give details only for $\mathcal{J}_t$. Define

$$\varrho(r) := \int_{|y|\le r} \tau^U(y)^{2-\delta} \nu_n(dy) + \int_{|y|\le r} \tau^U(y)^2 \nu_n(dy), \quad \text{and} \quad \varsigma(r) := \int_{r<|y|} \tau^U(y) \nu_n(dy)$$

The integrability conditions stated in Assumptions 5.1-[**I1**] yield that $\varrho(r) \to 0$ when $r \to 0$ and $\varsigma(r) \to +\infty$ if τ^U is not integrable around zero.

We start with the estimations in the Hölder space of type 1. For any $r \in (0,1)$ we have

$$\begin{aligned}\mathcal{J}_t\varphi = &\int_0^1 d\theta' \int_0^{\theta'} d\theta \int_{|y|\leq r} (\gamma^U)^2 \frac{\partial^2\varphi}{\partial u^2}(t, u+\theta\gamma^U, p, z)\nu_n(dy) \\ &+ \int_0^1 d\theta \int_{r<|y|} \gamma^U \left(\frac{\partial\varphi}{\partial u}(t, u+\theta\gamma^U, p, z) - \frac{\partial\varphi}{\partial u}\mathbb{1}_{\{|y|\leq 1\}}\right)\nu_n(dy)\end{aligned}$$

It follows that for some $M > 0$

$$\|\mathcal{J}\varphi\|_\infty \leq M\left(\varrho(r)\left\|D_x^2\varphi\right\|_\infty + \varsigma(r)\left\|D_x\varphi\right\|_\infty\right)$$

Similarly we can prove that

$$\begin{aligned}\langle\mathcal{J}\varphi\rangle_{x,Q_T}^{(\kappa)} &\leq M\left(\varrho(r)\left(\left\|D_x^2\varphi\right\|_\infty + \langle D_x^2\varphi\rangle_{x,Q_T}^{(\kappa)}\right) + \varsigma(r)\left(\|D_x\varphi\|_\infty + \langle D_x\varphi\rangle_{x,Q_T}^{(\kappa)}\right)\right) \\ \langle\mathcal{J}\varphi\rangle_{t,Q_T}^{(\kappa/2)} &\leq M\left(\varrho(r)\left(\left\|D_x^2\varphi\right\|_\infty + \langle D_x^2\varphi\rangle_{t,Q_T}^{(\kappa/2)}\right) + \varsigma(r)\left(\|D_x\varphi\|_\infty + \langle D_t\varphi\rangle_{t,Q_T}^{(\kappa/2)}\right)\right)\end{aligned}$$

By adding up the above estimations we obtain

$$\begin{aligned}\|\mathcal{J}\varphi\|_{\kappa/2,\kappa} \leq &M\left(\varrho(r)\|\varphi\|_{\kappa/2+1,\kappa+2} + \varsigma(r)\|\varphi\|_{(\kappa+1)/2,\kappa+1}\right) \\ \leq &M\left(\left(\varrho(r) + \epsilon^{1-\kappa}\varsigma(r)\right)\|\varphi\|_{\kappa/2+1,\kappa+2} + \epsilon^{-(1+\kappa)}\varsigma(r)\|\varphi\|_\infty\right)\end{aligned}$$

by applying Proposition C.2. A similar result can be obtained for $\mathcal{I}\varphi$ with, of course, some different functions ϱ and ς involving the functions τ and τ^P.

We can prove the estimations in the Hölder space of type 2 by slightly modifying the previous argument:

$$\begin{aligned}\mathcal{J}_t\varphi = &\int_0^1 d\theta \int_{|y|\leq r} \gamma^U \left(\frac{\partial\varphi}{\partial u}(t, u+\theta\gamma^U, p, z) - \frac{\partial\varphi}{\partial u}\right)\nu_n(dy) \\ &+ \int_0^1 d\theta \int_{r<|y|<1} \gamma^U \left(\frac{\partial\varphi}{\partial u}(t, u+\theta\gamma^U, p, z) - \frac{\partial\varphi}{\partial u}\mathbb{1}_{\{|y|\leq 1\}}\right)\nu_n(dy)\end{aligned}$$

It follows

$$\|\mathcal{J}\varphi\|_\infty \leq M\left(\varrho(r)\langle D_x\varphi\rangle_{x,Q_T}^{(1-\delta)} + \varsigma(r)\|D_x\varphi\|_\infty\right)$$

We conclude by applying Proposition C.3 so then

$$\|\mathcal{J}\varphi\|_\infty \leq M\left(\left(\varrho(r) + \varsigma(r)\epsilon^{1-\delta}\right)\|\varphi\|_{2-\delta,H} + \varsigma(r)\epsilon^{-1}\|\varphi\|_\infty\right)$$

□

For the non linear operator $\mathcal{H}$ we have the following

Lemma 6.3. *Suppose that Assumptions 5.1-*[**C1**, **I1**] *hold true and fix* $\beta > 0$. *Then*

$$\mathcal{H} : H^{1+\beta}([0,T] \times \mathbb{R}^3) \to H^{\beta}([0,T] \times \mathbb{R}^3)$$

and there exists a positive constant $M > 0$ *such that*

$$\begin{aligned}&\|\mathcal{H}[\varphi]\|_{\beta,H} \leq M \|\varphi\|_{1+\beta,H}\\&\|\mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi]\|_{\infty} \leq M \|\psi\|_{1,H}\end{aligned}$$

for all $\varphi, \psi \in H^{1+\beta}([0,T] \times \mathbb{R}^3)$.

*Suppose in addition that Assumptions 6.1-***C** *also holds true and fix* $\kappa \in (0,1)$. *Then*

$$\mathcal{H} : C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^3) \to C^{\kappa/2,\kappa}([0,T] \times \mathbb{R}^3)$$

and there exists a positive constant $M > 0$ *such that*

$$\|\mathcal{H}[\varphi]\|_{\kappa/2,\kappa} \leq M \|\varphi\|_{(\kappa+1)/2,\kappa+1}$$

for all $\varphi \in C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^3)$.

Proof.
If we define $H(q,g) := \inf_{|\pi| \leq \bar{\Pi}}\{2\pi q + \pi^2 g\}$ then $\mathcal{H}[\varphi] = H(\mathcal{Q}\varphi, \mathcal{G}\varphi)$. From Lemma 6.2 we obtain

$$\|\mathcal{H}[\varphi]\|_{\infty} \leq M\left(\|\mathcal{Q}\varphi\|_{\infty} + \|\mathcal{G}\varphi\|_{\infty}\right) \leq M \|\varphi\|_{1+\beta,H}$$

Let now $w = (t,u,p,z) \in [0,T] \times \mathbb{R}^3$ and assume now $0 \leq \mathcal{H}[\varphi](w) - \mathcal{H}[\varphi](w')$: it follows

$$\mathcal{H}[\varphi](w) - \mathcal{H}[\varphi](w') \leq 2\pi^*\left(\mathcal{Q}\varphi(w) - \mathcal{Q}\varphi(w')\right) + (\pi^*)^2\left(\mathcal{G}\varphi(w) - \mathcal{G}\varphi(w')\right)$$

where $\pi^* \in [-\bar{\Pi}, \bar{\Pi}]$ is the minimizer for $\mathcal{H}[\varphi](w')$. A similar estimation can be stated if $\mathcal{H}[\varphi](w) - \mathcal{H}[\varphi](w') \leq 0$. We deduce then

$$\langle\mathcal{H}[\varphi]\rangle^{(\beta)}_{x,Q_T} \leq M\left(\langle\mathcal{Q}\varphi\rangle^{(\beta)}_{x,Q_T} + \langle\mathcal{G}\varphi\rangle^{(\beta)}_{x,Q_T}\right) \leq M\left(\|\varphi\|_{1+\beta,H} + \|\varphi\|_{\beta,H}\right)$$

by using the definition of $\mathcal{Q}$ and $\mathcal{G}$. Together with the estimation for $\|\mathcal{H}[\varphi]\|_{\infty}$ we obtain $\|\mathcal{H}[\varphi]\|_{\beta,H} \leq \|\varphi\|_{1+\beta,H}$.

For the second estimation we use the concavity of H to obtain

$$\mathcal{H}[\psi] \leq \mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi] \leq \sup_{|\pi| \leq \bar{\Pi}}\left\{2\pi\mathcal{Q}\psi + \pi^2\mathcal{G}\psi\right\}$$

so then

$$\|\mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi]\|_{\infty} \leq M\left(\|\mathcal{Q}\psi\|_{\infty} + \|\mathcal{G}\psi\|_{\infty}\right) \leq M \|\psi\|_{1,H}$$

again from Lemma 6.2.

For the estimation in the Hölder space of type 1 $C^{\kappa/2,\kappa}([0,T]\times\mathbb{R}^3)$ we have

$$\|\mathcal{H}[\varphi]\|_\infty + \langle\mathcal{H}[\varphi]\rangle^{(\kappa)}_{x,Q_T} = \|\mathcal{H}[\varphi]\|_{\kappa,H} \le M\,\|\varphi\|_{1+\kappa,H} \le M\,\|\varphi\|_{(1+\kappa/2),1+\kappa}$$

whereas with the same type of computations we obtain

$$\langle\mathcal{H}[\varphi]\rangle^{(\kappa/2)}_{t,Q_T} \le M\left(\langle\varphi\rangle^{(\kappa/2)}_{t,Q_T} + \langle D\varphi\rangle^{(\kappa/2)}_{t,Q_T}\right) \le M\,\|\varphi\|_{(1+\kappa/2),1+\kappa}$$

which concludes our proof.

□

6.2 Smoothness and characterization of the function a

6.2.1 The approximation sequence

In this section we will prove that PIDE (6.1) has a unique, smooth and strictly positive solution. First of all let us transform the PIDE as follows: for $\eta > 0$

$$-\frac{\partial\varphi}{\partial t} + \mathcal{A}\varphi - \mathcal{B}\varphi - \mathcal{H}[\varphi] + \eta\varphi = 0, \quad \varphi(T,u,z) = e^{\eta T} \tag{6.2}$$

In particular, if φ is a solution of the above PIDE then $e^{-\eta t}\varphi(t,u,z)$ is a solution of (6.1).

Let now $\kappa \in (0,1)$, which we will determine later on, and fix $\varphi_0 \in C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$. Consider the sequence $(\varphi^n)_{n\in\mathbb{N}}$ defined recursively by

$$\begin{cases} \varphi^0 = \varphi_0 \\ -\dfrac{\partial}{\partial t}\varphi^{n+1} + \mathcal{A}\varphi^{n+1} + \eta\varphi^{n+1} = \mathcal{B}\varphi^n + \mathcal{H}[\varphi^n] \\ \varphi^{n+1}(T,u,z) = e^{\eta T} \end{cases} \tag{6.3}$$

This sequence is well defined in the Hölder space $C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$: by recurrence, if $\varphi^n \in C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$ then by Lemmas 6.2 and 6.3 we have $\mathcal{B}_t\varphi^n + \mathcal{H}_t[\varphi^n] \in C^{\kappa/2,\kappa}([0,T]\times\mathbb{R}^2)$ and we can apply Theorem 5.1 in Ladyzenskaja et al. (1967), a classical result on second order parabolic PDE, to deduce that φ^{n+1} is well defined in the Hölder space $C^{\kappa/2+1,\kappa+2}([0,T]\times\mathbb{R}^2)$ and

$$\left\|\varphi^{n+1}\right\|_{\kappa/2+1,\kappa+2} \le M\left(e^{\eta T} + \|\mathcal{B}\varphi^n + \mathcal{H}[\varphi^n]\|_{\kappa/2,\kappa}\right) \tag{6.4}$$

Furthermore we can write

$$\begin{aligned} \varphi^{n+1}(t,w) &= e^{\eta t} \\ &+ \int_t^T e^{-\eta(s-t)} \int_{\mathbb{R}^2} \Phi(T-t,w,T-s,\xi)\,(\mathcal{B}+\mathcal{H})\,\varphi^n(s,\xi)ds d\xi \end{aligned} \tag{6.5}$$

where $w = (u,z)$ and Φ is the fundamental solution of the linear parabolic PDE (Friedman, 1964; Ladyzenskaja et al., 1967). Estimations on the derivative of Φ are also available:

Lemma 6.4. *There exist some positive constant m_1, m_2 such that the following estimations hold true:*

i). For $2r + s \leq 2$, $s < t$

$$\left|D_t^i D_w^j \Phi(t, w, s, \xi)\right| \leq m_1 (t-s)^{-\frac{2+2i+j}{2}} \exp\left(-m_2 \frac{|w-\xi|^2}{t-s}\right)$$

ii). For $2i + j = 2$

$$\begin{aligned}&\left|D_t^i D_w^j \Phi(t, w, s, \xi) - D_t^i D_w^j \Phi(t, w', s, \xi)\right| \\ &\leq m_1 \left(|w-w'|^{\iota}(t-s)^{-\frac{4+\iota}{2}} + |w-w'|^{\iota'}(t-s)^{-\frac{4+\alpha-\iota}{2}}\right) \exp\left(-m_2 \frac{|w-\xi|^2}{t-s}\right)\end{aligned}$$

for any $\alpha \in (0,1)$, $\iota \in [0,1]$ and $\iota' \in [0, \alpha]$.

iii). For $2i + j = 1, 2$

$$\begin{aligned}&\left|D_t^i D_w^j \Phi(t, w, s, \xi) - D_t^i D_w^j \Phi(t', w, s, \xi)\right| \\ &\leq m_1 |t'-s|^{-\frac{4+2i+j}{2}} |t-t'|^{\frac{2-2i-j+\alpha}{2}} \exp\left(-m_2 \frac{|w-\xi|^2}{t-s}\right)\end{aligned}$$

If $s < t' < t$

A detailed proof of this Lemma can be found in Ladyzenskaja et al. (1967), Ch. IV, §13.

6.2.2 Weak convergence and uniqueness

Our aim now is to prove that the sequence defined in (6.3) converges in a (bigger) Hölder space of type 2. The method we will develop can also be used to prove that PIDE (6.1) has at most one solution. The main result of this paragraph is the following

Proposition 6.5. *Let Assumptions 5.1–6.1 hold true and $\delta \in (0,1)$ given in Assumptions 6.1-*[**I**]*. There exists a $\eta^* > 0$ such that for any $\eta > \eta^*$ the sequence $(\varphi^n)_n$ defined in* (6.3) *verifies*

$$\left\|\varphi^{n+1} - \varphi^n\right\|_{2-\delta,H} \leq (1+\eta) \left\|\varphi^1 - \varphi^0\right\|_{2-\delta,H} \beta^n$$

for some $\beta \in (0,1)$ which does not depend on η, φ^0 or φ^1. In particular $\varphi^n \to \varphi^*$ in $H^{2-\delta}([0,T] \times \mathbb{R}^2)$, the Hölder space of type 2, for some $\varphi^* \in H^{2-\delta}([0,T] \times \mathbb{R}^2)$.

Furthermore for any $\upsilon \in (0,1)$ there exists some positive constant M_υ which depends on υ, η and the other parameters given in Assumptions 5.1–6.1 such that

$$\left|\varphi^*(t, u, z) - \varphi^*(t', u, z)\right| \leq M_\upsilon |t-t'|^{\upsilon}, \quad \textit{for any } t, t', u, z$$

and

$$\left|D\varphi^*(t, u, z) - D\varphi^*(t', u, z)\right| \leq M_\varepsilon |t-t'|^{\upsilon/2}, \quad \textit{for any } t, t', u, z$$

where $D\varphi^* = (\partial_u \varphi*, \partial_z \varphi^*)$ is the spatial gradient of φ^*.

Proof.
Let us start by remarking that $\triangle^{n+1} := \varphi^{n+1} - \varphi^n$ verifies

$$-\frac{\partial}{\partial t}\triangle^{n+1} + \mathcal{A}_t\triangle^{n+1} + \eta\triangle^{n+1} = \mathcal{B}_t\triangle^n + \mathcal{H}_t[\varphi^n] - \mathcal{H}_t[\varphi^{n-1}]$$
$$\triangle^{n+1} = 0$$

For sake of compactness let us call $r(s,w) := (\mathcal{B}\triangle^n + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}])(s,w)$, $w := (u,z) \in \mathbb{R}^2$: according to (6.5) we can write

$$\triangle^{n+1}(t,w) = \int_t^T e^{-\eta(s-t)} \int_{\mathbb{R}^2} \Phi(T-t,w,T-s,\xi) r(s,\xi) d\xi ds$$

Let us also recall the definition of the $\| \, \|_{2-\delta,H}$ norm:

$$\|\varphi\|_{2-\delta,H} = \|\varphi\|_\infty + \|D_w\varphi\|_\infty + \langle D_w\varphi\rangle^{(1-\delta)}_{w,Q_T}$$

We easily obtain $\left\|\triangle^{n+1}\right\|_\infty \leq M\eta^{-1}\|r\|_\infty$ and $\left\|D_w\triangle^{n+1}\right\|_\infty \leq M\|r\|_\infty$ by using Lemma 6.4-*i*). The last thing we need to estimate is

$$\begin{aligned}
&\left|D_w\triangle^{n+1}(t,w) - D_w\triangle^{n+1}(t,w')\right| \\
=&\int_t^T e^{-\eta(s-t)} \int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,T-s,\xi) - D_w\Phi(T-t,w',s(t,\xi)\right| r(s,\xi) d\xi ds
\end{aligned}$$

From Lemma 6.4-*i*) we have

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,s-t,\xi) - D_w\Phi(T-t,w',s-t,\xi)\right| r(s+t,\xi) d\xi \qquad (6.6)\\
\leq& \|r\|_\infty \int_{\mathbb{R}^2} |D_w\Phi(T-t,w,T-s,\xi)| + \left|D_w\Phi(T-t,w',T-s,\xi)\right| d\xi \\
\leq& M\|r\|_\infty (s-t)^{-1/2}
\end{aligned}$$

and also

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,T-s,\xi) - D_w\Phi(T-t,w',T-s,\xi)\right| r(s,\xi) d\xi \qquad (6.7)\\
\leq& \|r\|_\infty |w-w'| \int_{\mathbb{R}^2}\int_0^1 d\theta \left|D^2_w\Phi(T-t,w'+\theta(w-w'),T-s,\xi)\right| d\xi d\theta \\
\leq& M\|r\|_\infty |w-w'|(s-t)^{-1}
\end{aligned}$$

Using the above estimation we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,s,\xi) - D_w\Phi(T-t,w',T-s,\xi)\right| r(s,\xi) d\xi \\
\leq& M\|r\|_\infty |w-w'|^{1-\delta}(s-t)^{-\frac{2-\delta}{2}}
\end{aligned}$$

so that finally

$$\begin{aligned}
&\left|D_w\triangle^{n+1}(t,w) - D_w\triangle^{n+1}(t,w')\right| \\
=&\int_t^T e^{-\eta(s-t)} \int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,T-s,\xi) - D_w\Phi(T-t,w',T-s,\xi)\right| r(s,\xi) d\xi ds \\
\leq& M\|r\|_\infty |w-w'|^{1-\delta} \int_t^T (s-t)^{-(2-\delta)/2} ds \leq M\|r\|_\infty |w-w'|^{1-\delta}
\end{aligned}$$

or, equivalently $\langle D_w \triangle^{n+1} \rangle^{(1-\delta)}_{w,Q_T} \leq M \|r\|_\infty$. We can now add up the previous estimations to deduce

$$\left\|\triangle^{n+1}\right\|_{2-\delta,H} \leq M \left\|\mathcal{B}\triangle^n + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty$$
$$\left\|\triangle^{n+1}\right\|_\infty \leq M\eta^{-1} \left\|\mathcal{B}\triangle^n + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty$$

Lemma 6.3 gives

$$\|\mathcal{B}\triangle^n\|_\infty \leq M\left(\left(\varrho(r) + \varsigma(r)\epsilon^{1-\delta}\right)\|\triangle^n\|_{2-\delta,H} + \varsigma(r)\epsilon^{-1}\|\triangle^n\|_\infty\right)$$

whereas from Lemma 6.3 and Proposition C.3 we obtain

$$\left\|\mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty \leq M\|\triangle^n\|_{1,H} \leq M\left(\epsilon^{1-\delta}\|\triangle^n\|_{2-\delta,H} + \epsilon^{-1}\|\triangle^n\|_\infty\right)$$

so then

$$\left\|\triangle^{n+1}\right\|_{2-\delta,H} \leq M\left(\left(\varrho(r) + \epsilon^{1-\delta}\varsigma(r)\right)\|\triangle^n\|_{2-\delta,H} + \epsilon^{-1}\varsigma(r)\|\triangle^n\|_\infty\right)$$
$$\left\|\triangle^{n+1}\right\|_\infty \leq M\eta^{-1}\left(\left(\varrho(r) + \epsilon^{1-\delta}\varsigma(r)\right)\|\triangle^n\|_{2-\delta,H} + \epsilon^{-1}\varsigma(r)\|\triangle^n\|_\infty\right)$$

Select ϵ^* and r^* small enough such that $2M\left(\varrho(r^*) + (\epsilon^*)^{1-\delta}\varsigma(r^*)\right) := \beta < 1$. Remark that we can do it since $\varrho(r) \to 0$ and M does not depend on r or ϵ. Select also

$$\eta^* > 2M\frac{(\epsilon^*)^{-1}\varsigma(r^*)}{\beta}$$

If we define an equivalent norm on $H^{2-\delta}([0,T]\times\mathbb{R}^2)$ as follows

$$\|\ \|_{2-\delta,\eta,H} := \|\ \|_{2-\delta,H} + \eta\|\ \|_\infty$$

then for $\eta \geq \eta^*$ then we obtain $\left\|\triangle^{n+1}\right\|_{2-\delta,\eta,H} \leq \beta\|\triangle^n\|_{2-\delta,\eta,H}$ which implies

$$\left\|\triangle^{n+1}\right\|_{2-\delta,\eta,H} \leq \beta^n\left\|\varphi^1 - \varphi^0\right\|_{2-\delta,\eta,H}$$

or equivalently

$$\left\|\varphi^{n+1} - \varphi^n\right\|_{2-\delta,H} \leq (1+\eta)\left\|\varphi^1 - \varphi^0\right\|_{2-\delta,\eta,H}\beta^n$$

which in particular proves that $(\varphi^n)_n$ is a Cauchy sequence in $H^{2-\delta}([0,T]\times\mathbb{R}^2)$ and then converges to some φ^*. Remark that β does not depend on η, whereas the function φ^1 does.

Let us prove the regularity of φ^* w.r.t. t. To simplify, assume that $\eta = 0$: from (6.5) we get

$$\left|D_w\varphi^{n+1}(t,w) - D_w\varphi^{n+1}(t',w)\right|$$
$$\leq \|\mathcal{B}\varphi^n + \mathcal{H}\varphi^n\|_\infty \int_t^T \int_{\mathbb{R}^2} \left|D_w\Phi(T-t,w,T-s,\xi) - D_w\Phi(T-t',w,T-s,\xi)\right| d\xi ds$$
$$+ \|\mathcal{B}\varphi^n + \mathcal{H}\varphi^n\|_\infty \int_{t'}^t \int_{\mathbb{R}^2} \left|D_w\Phi(T-t',w,T-s,\xi)\right| d\xi ds$$

for $t' < t$. The operators $\mathcal{B}$, $\mathcal{H}$ are continuous and since $\varphi^n \to \varphi^*$, we have

$$\|\mathcal{B}\varphi^n + \mathcal{H}\varphi^n\|_\infty \leq M\left(1 + \|\varphi^*\|_{2-\delta,H}\right)$$

Furthermore

$$\int_{t'}^t \int_{\mathbb{R}^2} \left|D_w\Phi(T-t', w, T-s, \xi)\right| d\xi ds \leq M \int_{t'}^t (s-t')^{-1/2} ds \leq M|t-t'|^{1/2}$$

again from Lemma 6.4-*i*). We deduce then

$$\begin{aligned}&\left|D_w\varphi^{n+1}(t,w) - D_w\varphi^{n+1}(t',w)\right| \\ \leq &M\left(|t-t'|^{1/2} + \int_t^T \int_{\mathbb{R}^2} \left|D_w\Phi(T-t, w, T-s, \xi) - D_w\Phi(T-t', w, T-s, \xi)\right| d\xi ds\right)\end{aligned}$$

The above integral can be estimated as in (6.6)–(6.7): we finally obtain

$$\left|D_w\varphi^{n+1}(t,w) - D_w\varphi^{n+1}(t',w)\right| \leq M\left(|t-t'|^{1/2} + |t-t'|^{\upsilon/2}\right)$$

for any $0 < \upsilon < 1$. It follows then that $D_w\varphi^n$ is Hölder continuous w.r.t. t so that we let $n \to \infty$ we deduce

$$\left|D_w\varphi^*(t,w) - D_w\varphi^*(t',w)\right| \leq M|t-t'|^{\upsilon/2}$$

We the same argument we can prove

$$\left|\varphi^{n+1}(t,w) - \varphi^{n+1}(t,w)\right| \leq M|t-t'|^{\upsilon}$$

and by taking the limit $n \to \infty$ we obtain the same property for φ^*.

□

Remark 6.6. *Proposition 6.5 gives us a fundamental property of the sequences defined in* (6.3) *when* $\eta \geq \eta^*$. *However this is not restrictive: if* φ^n *is a sequence corresponding to some* $\eta < \eta^*$ *we can always transform it into* $\tilde{\varphi}^n := \exp((\eta^* - \eta)t)\varphi^n$ *and then deduce all the properties for* $\tilde{\varphi}^n$. *With an abuse on language we will then say that the above Lemma holds for every* $\eta > 0$, *where, of course, one has to modify the constant which may change with* η.

Corollary 6.7. *Let Assumptions 5.1–6.1 hold true. The PIDE* (6.1) *has at most one solution in the Hölder space* $C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^2)$.

Proof.
Proving that the PIDE (6.1) has a unique solution is equivalent to prove that PIDE (6.2) has a unique solution. Suppose then that φ^i, $i = 1, 2$ are two solution of PIDE (6.2). Consider then the sequences $\varphi^{n,i}$ where $\varphi^{0,i} = \varphi^i$ for $i = 1, 2$. By construction it is clear that $\varphi^{n,i} = \varphi^i$ for all n. Let $\Delta^n := \varphi^{n,1} - \varphi^{n,2}$: with the same type of computation given in the proof of Proposition 6.5, it is possible to prove that

$$\left\|\Delta^{n+1}\right\|_{2-\delta,H} \leq (1+\eta)\beta^{n+1}\left\|\varphi^1 - \varphi^2\right\|_{2-\delta,H}$$

for some η big enough and $\beta \in (0,1)$. In particular $\Delta^n \to 0$ in $H^{2-\delta}([0,T] \times \mathbb{R}^2)$, and since $\Delta^n = \varphi^1 - \varphi^2$ we conclude that $\varphi^1 = \varphi^2$.

□

We can let $n \to \infty$ in (6.5) to deduce, for $\eta = 0$,

$$\varphi^*(t,u,z) = 1 + \mathbb{E}\left[\int_t^T (\mathcal{B}\varphi^* + \mathcal{H}[\varphi^*])(s, \tilde{U}_s^{t,u}, \tilde{Z}_s^{t,u,z})ds\right] \tag{6.8}$$

where

$$\begin{cases} d\tilde{Z}_r^{t,u,z} & := & \mu(r, \tilde{U}_r^{t,u}, \tilde{Z}_r^{t,u,z})dr + \sigma(r, \tilde{U}_r^{t,u}, \tilde{Z}_r^{t,u,z})dW_r^1 \\ d\tilde{U}_r^{t,u} & := & \mu^U(r, \tilde{U}_r^{t,u})dr + \sigma^U(r, \tilde{U}_r^{t,u})dB_r \end{cases} \tag{6.9}$$

since the operators $\mathcal{B}$ and $\mathcal{G}$ are continuous in $H^{2-\delta}([0,T]\times\mathbb{R}^2)$. We also have the optimal $\hat{\pi}$ related to φ^* in the right hand side of the above equality:

$$\hat{\pi}(t,u,z) := -\bar{\Pi} \vee -\frac{\mathcal{Q}\varphi^*(t,u,z)}{\mathcal{G}\varphi^*(t,u,z)} \wedge \bar{\Pi} \tag{6.10}$$

The regularity of φ^* shows that $\hat{\pi}$ is well defined and trivially bounded. Furthermore, from the definition of $\mathcal{Q}$ and $\mathcal{G}$, we have that $\hat{\pi}$ essentially depends on φ^* and its derivative w.r.t. z. In particular it is straightforward to deduce that $\hat{\pi}$ is Hölder continuous: more precisely, for any t, u, u', z, z'

$$\left|\hat{\pi}(t,u,z) - \hat{\pi}(t,u',z')\right| \le M\left(|u-u'|^{1-\delta} + |z-z'|^{1-\delta}\right)$$

since $\varphi^* \in H^{2-\delta}([0,T]\times\mathbb{R}^2)$. Furthermore, by using the regularity condition w.r.t. t given in Proposition 6.5 we also have

$$\left|\hat{\pi}(t,u,z) - \hat{\pi}(t',u,z)\right| \le M|t-t'|^{\upsilon/2}$$

for any t, t', u, z and any $\upsilon \in (0,1)$. The above Hölder conditions implies that

$$\hat{\pi} \in C^{(1-\delta)/2, 1-\delta}([0,T]\times\mathbb{R}^2)$$

6.2.3 Characterization of the function a

We now have all the elements to prove that the PIDE (6.1) has a unique smooth solution:

Theorem 6.8. *Let Assumptions 5.1–6.1 hold true. The PIDE* (6.1) *has a unique and strictly positive solution* $a \in C^{(1-\delta)/2+1, 2+(1-\delta)}([0,T]\times\mathbb{R}^2)$*, where* $\delta \in (0,1)$ *is given in Assumptions 6.1-*[**I**]*. Moreover*

$$\|\varphi^n - a\|_{2-\delta,H} \le M\beta^n, \quad n \to \infty$$

for some $M > 0$ *and* $\beta \in (0,1)$*.*

***Proof*.**
Let φ^* be the limit of the sequence φ^n when $\eta = 0$, as stated in Proposition 6.5, and $\hat{\pi}$ given in (6.10) which, as we know, belongs to $C^{(1-\delta)/2,1-\delta}([0,T]\times\mathbb{R}^2)$. For sake of clarity we summarize here the scheme of the proof:

Step 1. We prove that φ^* is the unique viscosity solution of

$$-\frac{\partial \varphi^*}{\partial t} + \mathcal{A}\varphi^* - \mathcal{B}\varphi^* - 2\hat{\pi}\mathcal{Q}\varphi^* - \hat{\pi}^2\mathcal{G}\varphi^* = 0, \quad \varphi^*(T,.) = 1 \tag{6.11}$$

Step 2. We prove that PIDE (6.11) has a unique smooth solution

Step 3. We deduce that $\varphi^* \in C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^2)$, it is strictly positive and, from Theorem 5.11, we conclude that $\varphi^* = a$.

Step 1. From (6.8) and the Markov property of the process $(\tilde{U}, \tilde{Z})$ given in (6.9) we have

$$\varphi^*(t,u,z) = \mathbb{E}\left[\int_t^{t+h} (\mathcal{B}+\mathcal{H})\,\varphi^*\left(s, \tilde{U}_s^{t,u}, \tilde{Z}_s^{t,u,z}\right) ds + \varphi^*(t+h, \tilde{U}_{t+h}^{t,u}, \tilde{Z}_{t+h}^{t,u,z}\right)$$

In Proposition 6.5 we proved that φ^* Hölder continuous w.r.t t so that the right hand side of the above equality is well defined. Remark that the above equality is nothing but the dynamic programming principle. Let now $(t,u,z) \in [0,T] \times \mathbb{R}^2$ and take $\Psi_1, \Psi_2 \in C^{1,2}([0,T] \times \mathbb{R}^2)$ such that

$$\begin{aligned} 0 =& \varphi^*(t,u,z) - \Psi_1(t,u,z) = \max_{t',u',z'} (\varphi^* - \Psi_1)(t',u',z') \\ 0 =& \varphi^*(t,u,z) - \Psi_2(t,u,z) = \min_{t',u',z',} (\varphi^* - \Psi_2)(t',u',z') \end{aligned}$$

It follows then

$$\begin{aligned} \Psi_1(t,u,z) \leq& \mathbb{E}\left[\int_t^{t+h} (\mathcal{B}+\mathcal{H})\,\varphi^*\left(s, \tilde{U}_s^{t,u}, \tilde{Z}_s^{t,u,z}\right) ds + \Psi_1\left(t+h, \tilde{U}_{t+h}^{t,u}, \tilde{Z}_{t+h}^{t,u,z}\right)\right] \\ \Psi_2(t,u,z,x) \geq& \mathbb{E}\left[\int_t^{t+h} (\mathcal{B}+\mathcal{H})\,\varphi^*\left(s, \tilde{U}_s^{t,u}, \tilde{Z}_s^{t,u,z}\right) ds + \Psi_2\left(t+h, \tilde{U}_{t+h}^{t,u}, \tilde{Z}_{t+h}^{t,u,z}\right)\right] \end{aligned}$$

We can now apply Itô's formula to obtain

$$-\frac{\partial \Psi_1}{\partial t} + \mathcal{A}\Psi_1 - \mathcal{B}\varphi^* - \mathcal{H}\varphi^* \leq 0 \quad \text{and} \quad -\frac{\partial \Psi_2}{\partial t} + \mathcal{A}\Psi_2 - \mathcal{B}\varphi^* - \mathcal{H}\varphi^* \geq 0$$

and by definition of $\hat{\pi}$

$$\begin{aligned} &-\frac{\partial \Psi_1}{\partial t} + \mathcal{A}\Psi_1 - \mathcal{B}\varphi^* - 2\hat{\pi}\mathcal{Q}\varphi^* - \hat{\pi}^2\mathcal{G}\varphi^* \leq 0 \\ &-\frac{\partial \Psi_2}{\partial t} + \mathcal{A}\Psi_2 - \mathcal{B}\varphi^* - 2\hat{\pi}\mathcal{Q}\varphi^* - \hat{\pi}^2\mathcal{G}\varphi^* \geq 0 \end{aligned}$$

According to Definition 5.15 we deduce that φ^* is a viscosity solution of (6.11). We do not detail it here but one can prove that φ^* is the unique viscosity solution of the above PIDE. We refer to Barles et al. (1997) or Pham (1998) for uniqueness results, which are stated in a more general context.

Step 2. Remark that the PIDE (6.11) is linear, so we can hope to prove that its unique viscosity solution is a classical solution. For this, let $\eta > 0$ and consider the map Ξ_η as follows: for $\psi \in C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$, $\Xi_\eta(\psi)$ denotes the unique solution of

$$-\frac{\partial \Xi_\eta(\psi)}{\partial t} + \mathcal{A}\Xi_\eta(\psi) + \eta\Xi_\eta(\psi) = \mathcal{B}\psi + 2\hat{\pi}\mathcal{Q}\psi + \hat{\pi}^2\mathcal{G}\psi \tag{6.12}$$
$$\Xi_\eta(\psi)(T,.) = e^{\eta T}$$

Since $\hat{\pi} \in C^{(1-\delta)/2,1-\delta}([0,T]\times\mathbb{R}^2)$ and $\psi \in C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$, we can apply Lemma 6.2 to deduce that

$$\mathcal{B}\psi + 2\hat{\pi}\mathcal{Q}\tilde{\psi} + \hat{\pi}^2\mathcal{G}\psi \in C^{(1-\delta)/2,1-\delta}([0,T]\times\mathbb{R}^2)$$

Theorem 5.1 in Ladyzenskaja et al. (1967) (as in paragraph 6.2.1) proves that Ξ_η is well defined and maps $C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$ into itself. Moreover there exists some $M > 0$ not depending on η or ψ such that

$$\|\Xi_\eta(\psi)\|_{(1-\delta)/2+1,2+(1-\delta)} \le M\left(e^{\eta T} + \left\|\mathcal{B}\psi + 2\hat{\pi}\mathcal{Q}\psi + \hat{\pi}^2\mathcal{G}\psi\right\|_{(1-\delta)/2,(1-\delta)}\right)$$

which proves that Ξ_η is well defined.
If we prove that Ξ_η is a contraction in $C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$ and ψ^* denotes its unique fixed point, then $e^{-\eta t}\psi^*(t,u,z) \in C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$. Moreover it satisfies the PIDE (6.11): by the uniqueness of the viscosity solution, proved in Step 1, we deduce $\varphi^*(t,u,z) = e^{-\eta t}\psi^*(t,u,z)$.

To prove that Ξ_η is a contraction we use the method introduced in Chapter III of Bensoussan and Lions (1984). Fix $\psi_1,\psi_2 \in C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$: by using Lemma 6.2 and the regularity of $\hat{\pi}$, we have

$$\begin{aligned}
&\|\Xi_\eta(\psi_1) - \Xi_\eta(\psi_1)\|_{(1-\delta)/2+1,2+(1-\delta)} \\
\le& M\left(\left\|\mathcal{B}(\psi_1-\psi_2) + 2\hat{\pi}\mathcal{Q}(\psi_1-\psi_2) + \hat{\pi}^2\mathcal{G}(\psi_1-\psi_2)\right\|_{(1-\delta)/2,(1-\delta)}\right) \\
\le& M\left(\left(\varrho(r) + \varsigma(r)\epsilon^\delta\right)\|\psi_1-\psi_1\|_{(1-\delta)/2+1,2+(1-\delta)} + \epsilon^{-(2-\delta)}\varsigma(r)\|\psi_1-\psi_1\|_\infty\right)
\end{aligned} \tag{6.13}$$

for some positive M which does not depend on η or ψ_1,ψ_2. Moreover, the Feynman-Kac formula gives

$$\Xi_\eta(\psi_1) - \Xi_\eta(\psi_1) = \mathbb{E}\left[\int_t^T e^{-\eta(s-t)}\left(\mathcal{B} + 2\hat{\pi}\mathcal{Q} + \hat{\pi}^2\mathcal{G}\right)(\psi_1-\psi_2)(s,\tilde{U}_s,\tilde{Z}_s)\right]$$

where the process $(\tilde{U},\tilde{Z})$ is given in (6.9). It follows

$$\begin{aligned}
&\|\Xi_\eta(\psi_1) - \Xi_\eta(\psi_1)\|_\infty \\
\le& M\eta^{-1}\left(\left\|\mathcal{B}(\psi_1-\psi_2) + 2\hat{\pi}\mathcal{Q}(\psi_1-\psi_2) + \hat{\pi}^2\mathcal{G}(\psi_1-\psi_2)\right\|_\infty\right) \\
\le& M\eta^{-1}\left(\left(\varrho(r) + \varsigma(r)\epsilon^\delta\right)\|\psi_1-\psi_1\|_{(1-\delta)/2+1,2+(1-\delta)} + \epsilon^{-(2-\delta)}\varsigma(r)\|\psi_1-\psi_1\|_\infty\right)
\end{aligned} \tag{6.14}$$

Take now r^* and ϵ^* small enough to verify

$$2M\left(\varrho(r^*)+\varsigma(r^*)(\epsilon^*)^{\delta}\right)<\omega$$

for some $\omega\in(0,1)$. Remark that we can do it since $\varrho(r)\to 0$ when $r\to 0$ and $\delta>0$. Select also

$$\eta>2M\frac{\varsigma(r^*)(\epsilon^*)^{-(2-\delta)}}{\omega}$$

If we introduce the norm

$$\|\,\|_{\eta,\delta}:=\|\,\|_{1+(1-\delta)/2,\,2+(1-\delta)}+\eta\,\|\,\|_{\infty}$$

on $C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$, which is equivalent to $\|\,\|_{1+(1-\delta)/2,\,2+(1-\delta)}$, then from (6.13) and (6.14) we obtain

$$\|\Xi_\eta(\psi_1)-\Xi_\eta(\psi_1)\|_{\eta,\delta}\leq\omega\,\|\psi_1-\psi_1\|_{\eta,\delta}$$

which proves that Ξ_η is a contraction. As we already said, this proves that φ^* belongs to $C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$.

Step 3. The process

$$d\hat{X}_s^{t,u,z,x}:=\hat{\pi}_{s-}\hat{X}_{s-}^{t,u,z,x}e^{-Z_{s-}^{t,u,z}}de^{Z_s^{t,u,z}},\;\hat{X}_t^{t,u,z,x}=x$$

is well defined since $\hat{\pi}$ is bounded. If follows that the function

$$w(t,u,z,x):=\mathbb{E}\left[\left(\hat{X}_T^{t,u,z,x}\right)^2\right] \tag{6.15}$$

is also well defined, continuous and $w(t,u,z,x)=x^2\tilde{\varphi}(t,u,z)$ for some $\tilde{\varphi}$. The Markov property of the process (U,Z,X) gives

$$w(t,u,z,x)=\mathbb{E}\left[w\left(t+h,U_{t+h}^{t,u},Z_{t+h}^{t,u,z},\hat{X}_{t+h}^{t,u,z,x}\right)\right]$$

and, as before, it is not complicated to prove that $\tilde{\varphi}$ is a viscosity solution of

$$-\frac{\partial\tilde{\varphi}}{\partial t}+\mathcal{A}\tilde{\varphi}-\mathcal{B}\tilde{\varphi}-2\hat{\pi}\mathcal{Q}\tilde{\varphi}-\hat{\pi}^2\mathcal{G}\tilde{\varphi}=0,\quad\tilde{\varphi}(T,.)=1$$

Again the uniqueness of the viscosity solution yields $\varphi^*=\tilde{\varphi}$. In particular

$$x^2\varphi^*(t,u,z)=x^2\tilde{\varphi}(t,u,z)=\mathbb{E}\left[\left(\hat{X}_T^{t,u,z,x}\right)^2\right]\geq v^0(t,u,z,x)=x^2a(t,u,z)$$

where v^0 is the value function defined in (5.12), since $\hat{X}$ is an admissible portfolio. From the above estimation and Lemma 5.3 we deduce $e^{-CT}<a(t,u,z)\leq\varphi^*(t,u,z)$ for all t,u,z.

To summarize, we proved that $\varphi^*\in C^{1+(1-\delta)/2,\,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$, it is the unique solution the PIDE (6.1) and it is strictly positive: we can then apply Theorem 5.11 to deduce $a=\varphi^*$ and characterize the optimal strategy of problem (5.12). Finally we use Proposition 6.5 to obtain that $\varphi^n\to a$ in $H^{2-\delta}([0,T]\times\mathbb{R}^2)$, which concludes our proof.

□

6.2.4 Comments

Another method to prove that PIDE (6.1) has a solution is provided by the theory of backward stochastic differential equations with jumps (BSDEs). For a complete review on different aspects of BSDEs see for example El Karoui et al. (1997); Rong (1997); Royer (2006); Crépey and Matoussi (2008). We do not go deeper in details but we think it is interesting to see how this method works. We just make a (short) digression which essentially uses the ideas developed in Barles et al. (1997). They start from a semi linear PIDE

$$-\partial_t \varphi + \mathcal{L}\varphi - f(t, x, D\varphi, \tilde{\mathcal{B}}\varphi) = 0, \quad \varphi(T, .) = g$$

where $\mathcal{L}$ is the Dynkin operator associated to some (eventually discontinuous) process X:

$$dX_s = b_X(X_s)ds + a_X(X_s)dW_s + \int_E \beta_X(X_s, e)\tilde{P}(de, ds)$$

for some Brownian motion W and some Poisson measure P with Lévy measure $\lambda(de)$. $\tilde{\mathcal{B}}$ is a first-order non local operator:

$$\tilde{\mathcal{B}}\varphi = \int_T (\varphi(t, x + \beta_X(x, e)) - \varphi(t, x))\tilde{\beta}(x, e)\lambda(de)$$

They prove then that the (unique) solution φ of this non linear PIDE can be related to the unique solution of a BSDE with jumps: $\varphi(t, x) := Y_t^{t,x}$ where (Y, Σ, α) is the unique solution of

$$-dY_s^{t,x} = f(t, x, Y_s^{t,x}, \Sigma_s^{t,x}, \alpha_s^{t,x}) - \Sigma_s^{t,x} dW_s - \int_E \alpha_s^{t,x}(e)\tilde{P}(de, ds)$$

and where the "driver" f has the particular form

$$f(t, x, Y, \Sigma, \alpha) = f(t, x, Y, \Sigma, \int_E \alpha(e)\tilde{\gamma}(e)\lambda(de))$$

for some $\tilde{\gamma}$. Under classical assumptions for BSDEs they prove existence and regularity for the solution φ. In our context the process X is (U, Z) whereas the driver f will be the non linear operator $\mathcal{H}$:

$$\begin{aligned} f(t, (u, z), Y, \Sigma, \alpha) :=& \inf_\pi f(\pi, (u, z), Y, S, \alpha) \\ f(\pi, x, Y, \Sigma, \alpha) :=& Y(2\pi\tilde{\mu}(t, u, z) + \sigma^2\pi^2 + \pi^2 \int (e^\gamma - 1)^2 \nu(dy)) \\ &+ \Sigma(2\pi\sigma(t, u, z)) + \int_{\mathbb{R}} \alpha(y) \left(2\pi(e^\gamma - 1)\mathbb{1}_{\{|y| \le 1\}} + \pi^2(e^\gamma - 1)^2\right) \nu(dy) \end{aligned}$$

Unfortunately they do not fulfill all the assumptions given in Barles et al. (1997): in particular the dynamics of the process X is not stationary so we cannot directly apply their result. Also the comparison theorem cannot be directly applied since the driver is not increasing in the α argument: this is due to the fact that we do not control the sign of $\pi(t, u, z)$.

Another possibility would be to use the theory of weak solutions, i.e. to work in (weighted) Sobolev spaces instead of Hölder spaces. A complete review on the use of Sobolev spaces in the resolution of PIDEs can be found in Bensoussan and Lions (1984), Ch III, §1-§3. For example they prove that a wide class of semi linear PIDEs has a unique solution in some appropriate weighted Sobolev space over the parabolic domain $[0,T]\times\mathbb{R}^2$. This method presents at the same time a great advantage and a serious disadvantage compared to our method. Let us start by presenting its advantage: in Bensoussan and Lions (1984), the existence is obtained by proving that $\Psi\to\Xi_\eta(\Psi)$ is a contraction in the weighted Sobolev space $\mathcal{W}^{2,1,p,\lambda}(([0,T]\times\mathbb{R}^n)$ for η big enough, where $\Xi_\eta(\Psi)$ is the (unique) solution of

$$-\partial_t\Xi_\eta(\Psi)+\mathcal{A}\Xi_\eta(\Psi)+\eta\Xi_\eta(\Psi)=(\mathcal{B}+\mathcal{H})\Psi$$

Let us assume, to simplify, that $p=\infty$ (Bensoussan and Lions (1984), Ch III, §3, Theorem 3.3), so we look at $\mathbb{L}^\infty$-norms. To prove the contraction they *only* need an estimation on $\|\mathcal{B}(\Psi_1-\Psi_2)\|_\infty$ and $\|\mathcal{H}[\Psi_1]-\mathcal{H}[\Psi_2]\|_\infty$ to ensure that the sequence $\Psi_{n+1}=\Xi_\eta(\Psi_n)$ converges in this Sobolev space.

In our case things are more complicated: firstly the space

$$\left\{\varphi\in C^2([0,T]\times\mathbb{R}^n),\ \sum_{i=0}^{2}\sum_{(j)}\left\|D^j\varphi\right\|_\infty<\infty\right\}$$

is not a Banach space: we cannot prove that the map Ξ_η is a contraction in this space. In other words, having a control on $\left\|D^2\Xi_\eta(\Psi_1-\Psi_2)\right\|_\infty$ is not enough to ensure that the sequence $\Psi_{n+1}=\Xi_\eta(\Psi_n)$ converges in a classical sense. For this, we need to have a control, for example, on $\langle D^2\Xi_\eta(\Psi_1-\Psi_2)\rangle_{Q_T}^{(\beta)}$ for some $\beta\in(0,1)$. As we have already done several times, we can use (6.5) and find

$$\begin{aligned}&D^2\Xi_\eta(\Psi_1-\Psi_2)(t,w)-D^2\Xi_\eta(\Psi_1-\Psi_2)(t,w')\\&=\int_t^T e^{-\eta(s-t)}\int_{\mathbb{R}^2}(D^2\Phi(T-t,w,T-s,\xi)-D^2\Phi(T-t,w',T-s,\xi))r(s,\xi)dsd\xi\end{aligned}$$

where $r=(\mathcal{B}+\mathcal{H})(\Psi_1-\Psi_2)$. If we use the estimation given in Lemma 6.4 then

$$\begin{aligned}&\left|D^2\Xi_\eta(\Psi_1-\Psi_2)(t,w)-D^2\Xi_\eta(\Psi_1-\Psi_2)(t,w')\right|\\&\le M\|r\|_\infty\int_t^T\int_{\mathbb{R}^2}\left|D^2\Phi(T-t,w,T-s,\xi)-D^2\Phi(T-t,w',T-s,\xi)\right|dsd\xi\\&\sim\int_t^T(s-t)^{-1-\beta/2}ds\end{aligned}$$

which of course is not finite. To make this term finite we need to exploit the regularity of r in its arguments: contrary to Bensoussan and Lions (1984), in our case some Hölder regularity on r is needed (Lemmas 6.2–6.3). The message coming for this short digression is that the use of Hölder spaces is more constraining if compared to the use of Sobolev spaces.

On the other side the main difficulty when one uses Sobolev spaces is to find good embeddings into some space of real and, possibly, smooth functions. In my knowledge embedding results are stated in bounded domains, which also verify the so called *cone condition* (Adams and Fournier, 2009). Also it is not immediate to find embeddings in the parabolic case: in Bensoussan and Lions (1984) for example the Sobolev space is defined as

$$\mathcal{W}^{2,1,p,\lambda}([0,T]\times\mathbb{R}^n) := \left\{ u \in \mathbb{L}^p\left([0,T] \to \mathcal{W}^{2,p,\lambda}(\mathbb{R}^n)\right) \middle| \frac{\partial u}{\partial t} \in \mathbb{L}^p\left([0,T] \to \mathbb{L}^{p,\lambda}(\mathbb{R}^n)\right) \right\}$$

for which direct embeddings are not available in my knowledge. We want to remark, however, that the ideas developed by these authors are quite universal and can be applied in different context, as we just did.

Before we conclude, we want to point out that, however, our model does not fulfill the assumptions of Theorem 3.3 (Bensoussan and Lions, 1984) concerning the regularity of the non local operator $\mathcal{B}$: for these authors it has to be defined, in its general form, as follows:

$$\mathcal{B}f = \int \left(f(x+y) - f(x) - \langle y, Df\rangle \mathbb{1}_{\{|y|\le 1\}}\right) M(x,dy)$$

where $M(x,dy) = c_0(x,y)m(dy)$ is an unbounded measure, for some measure $m(dy)$ and $0 \le c_0 \le 1$. In our case this is not always true: for this we should be able to invert the jump function γ and rewrite $\mathcal{B}$, for example, as

$$\mathcal{B}f = \int \left(f(x+y) - f(x) - \langle y, Df\rangle \mathbb{1}_{\{|y|\le 1\}}\right) \nu(\gamma^{-1}(x,y))D\gamma^{-1}(x,y)dy$$

This can be done if, for example, the Lévy measure has a density with respect to the Lebesgue measure, which is not always the case.

We conclude the section with a technical result concerning the sequence defined in (6.3), which will be used in Chapter 9. For sake of simplicity we prove it under the assumption that the coefficients of the process Z do not depend on the process U, but it can be easily extended to the general case.

Lemma 6.9. *Let Assumptions 5.1–6.1 hold true and* $\varphi^n_\eta \in C^{1+\frac{(1-\delta)}{2},2+(1-\delta)}([0,T]\times\mathbb{R})$ *be the sequence given in* (6.3)*. There exist a positive constant* $M \ge 0$ *depending on* η *such that*

$$\sup_{n\in\mathbb{N}} \left\|\varphi^{n+1}\right\|_{2,H} \le M$$

Proof.
Before we start the proof let us remark that $\|\varphi^n - a\|_{2-\delta,H} \to 0$, $n \to \infty$, which in particular proves that

$$\sup_{n\in\mathbb{N}} \left\|\varphi^{n+1}\right\|_{2-\delta,H} \le M$$

for some positive M. The proof will be completed if we can prove that $\sup_{n\in\mathbb{N}} \left\|\partial^2_z\varphi^{n+1}\right\|_\infty < \infty$. Consider φ^n_0, the sequence corresponding to $\eta = 0$. To simplify the notations we will omit the subscript. For any $\lambda > 0$ we obtain

$$\varphi^{n+1}(t,z) = \varphi^{n+1}(t+\lambda,z) - \int_t^{t+\lambda} \left(\mathcal{A}\varphi^{n+1} - \mathcal{B}\varphi^n - \mathcal{H}[\varphi^n]\right)(s,z)ds$$

or equivalently

$$\int_t^{t+\lambda} \mathcal{A}\varphi^{n+1}(s,z)ds = \varphi^{n+1}(t,z) - \varphi^{n+1}(t+\lambda,z) + \int_t^{t+\lambda} \left(\mathcal{B}\varphi^n + \mathcal{H}[\varphi^n]\right)(s,z)ds$$

From the definition of $\mathcal{A}$ we get

$$-\frac{1}{2}\int_t^{t+\lambda} \sigma^2 \frac{\partial^2 \varphi^{n+1}}{\partial^2 z}(s,z)ds = \Lambda^n(\lambda) \tag{6.16}$$

where

$$\begin{aligned}
\Lambda^n(\lambda) \quad =& \varphi^{n+1}(t,z) - \varphi^{n+1}(t+\lambda,z) + \int_t^{t+\lambda} \mu\frac{\partial \varphi^{n+1}}{\partial z}(s,z)ds \\
&+ \int_t^{t+\lambda} \left(\mathcal{B}\varphi^n + \mathcal{H}[\varphi^n]\right)(s,z)ds \\
\overset{n\to\infty}{\to} \quad & a(t,z) - a(t+\lambda,z) + \int_t^{t+\lambda} \mu\frac{\partial a}{\partial z}(s,z)ds + \int_t^{t+\lambda} \left(\mathcal{B}a + \mathcal{H}[a]\right)(s,z)ds
\end{aligned}$$

so then $\sup_{\lambda>0}\sup_n \|\Lambda^n(\lambda)/\lambda\|_\infty \leq m\,\|a\|_{1+(1-\delta)/2,2+(1-\delta)}$ for some positive constant m. If $\sup_{n\in\mathbb{N}} \|\varphi^n\|_{2,H} = \infty$ then for any positive $R > 0$ we could find some $n \in \mathbb{N}$ and $(t,z) \in [0,T] \in \mathbb{R}$ such that, for example, $\partial_z^2\varphi^{n+1}(t,z) > R$ (the same argument stands in force is the second derivative is negative). In particular, the continuity of $\partial_z^2\varphi^{n+1}$ proves that for some small λ, we will have

$$0 \leq \sigma^2_{min} R \leq \frac{1}{\lambda}\int_t^{t+\lambda} \sigma^2 \frac{\partial^2 \varphi^{n+1}}{\partial^2 z}(s,z)ds \leq -2\frac{\Lambda^n(\lambda)}{\lambda}$$

Since the above inequality trivially contradicts the fact that $\Lambda^n(\lambda)/\lambda$ is uniformly bounded, we can conclude that $\sup_{n\in\mathbb{N}} \|\varphi^n\|_{2,H} \leq M$, for some positive M. Obviously this constant depends on η when one consider the sequence in (6.3) for some $\eta > 0$.

□

6.3 Smoothness and characterization of the function v^f

According to Theorem 5.14, we now need to prove that the PIDEs (5.48)–(5.49) have a unique smooth solution.

Theorem 6.10. *Let Assumptions 5.1–6.1 hold true and $\delta \in (0,1)$ given in Assumptions 6.1-[**I**]. Assume also that $f \in H_e^{2+(1-\delta)}(\mathbb{R}^3)$. The value function v^f defined in* (5.11) *admits the decomposition given in* (5.41)*:*

$$\begin{aligned}
v^f(t,u,p,z,x) =& a(t,u,z)x^2 + b(t,u,p,z)x + c(t,u,p,z) \\
v^f(T,u,p,z,x) =& (f(u,p,z) - x)^2
\end{aligned}$$

where $a \in C^{(1-\delta)/2+1,2+(1-\delta)}([0,T]\times\mathbb{R}^2)$ is the unique solution of 6.1, so it does not depend on f, and

$$b, c \in C^{(1-\delta)/2+1,2+(1-\delta)}([0,T]\times\mathbb{R}^3)$$

are the unique solutions of the following linear parabolic PIDEs

$$0 = -\frac{\partial b}{\partial t} + \mathcal{A}_t b - \mathcal{B}_t b - \pi^* \mathcal{Q}_t b \qquad b(T,.) = -2f \tag{6.17}$$

$$0 = -\frac{\partial c}{\partial t} + \mathcal{A}_t c - \mathcal{B}_t c + \frac{1}{4}\frac{(\mathcal{Q}_t b)^2}{\mathcal{G}_t a} \qquad c(T,.) = f^2 \tag{6.18}$$

where π^* *is defined in* (5.47).

Remark 6.11. *The result also holds true if* $f \in H_e^{2+\kappa}(\mathbb{R}^3)$, *for some other* $0 < \kappa$. *In this case one would have*

$$b, c \in C^{\kappa'/2+1,\kappa'+2}([0,T] \times \mathbb{R}^3)$$

where $\kappa' = \min(1-\delta, \kappa)$.

Proof.
We first prove that PIDEs (6.17)-(6.18) have a unique solution and then we use Theorem 5.14 to conclude.
We already know that the unique solution of PIDE in (6.1) is the function a in (5.13) and it belongs to the Hölder space $C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^2)$ (Theorem 6.8). It follows that $\pi^* \in C^{(1-\delta)/2,(1-\delta)}([0,T] \times \mathbb{R}^2)$ (π^* is Lipschitz continuous in the space variable (u,z)). We can then rewrite the PIDE (6.17) into:

$$0 = -\frac{\partial b}{\partial t} + \mathcal{A}b - \mathcal{B}b - \pi^* \mathcal{Q}b + \eta b, \; b(T,.) = -2fe^{\eta T} \tag{6.19}$$

and use the contraction principle as in the proof of Theorem 6.8: let $\Xi_\eta(\psi)$ be the unique smooth solution of

$$-\frac{\partial}{\partial t}\Xi_\eta(\psi) + \mathcal{A}\Xi_\eta(\psi) + \eta\Xi_\eta(\psi) = \mathcal{B}\psi + \pi^*\mathcal{Q}\psi, \text{ on } [0,T) \times \mathbb{R}^3$$
$$\Xi_\eta(\psi)(T,.) = -2fe^{\eta T}$$

where $\psi \in C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^3)$. We skip the details, which are the same as in the proof of Theorem 6.8, but it is possible to select η big enough such that Ξ_η is a contraction in the Hölder space $C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^3)$: its fixed point ψ^* is then the unique solution of (6.19), or, equivalently, $e^{-\eta t}\psi^*$ is the unique solution of (6.17). For the PIDE (6.18) we can proceed in the same way to deduce that it has a unique smooth solution in $C^{(1-\delta)/2+1,2+(1-\delta)}([0,T] \times \mathbb{R}^3)$.

□

Theorem 6.10 holds true for smooth payoff functions. When $f \in H_e^\beta(\mathbb{R}^3)$ for some $\beta \in (0,1]$, it is possible to find a sequence $f_n \in C_b^\infty(\mathbb{R}^3)$, the space of infinitely differentiable functions with bounded derivatives, such that:

i). $\|f_n\|_{\beta,e} \leq 2\|f\|_{\beta,e}$ and $\|f\|_{\beta,e} \leq \liminf_n \|f_n\|_{\beta,e}$

ii). $\|f_n - f\|_{\beta',e} \to 0$, $n \to \infty$ for any $\beta' < \beta$

See for example Mikulevicius and Pragarauskas (2009) for a complete proof. We can then replace f with f_n and use the argument exposed in Chapter 5, Section 5.4 to control the error.

Chapter 7

Smooth solutions: the pure jump case

The Chapter is organized as follows: in Section 7.1 we modify the model proposed in (5.5) *by taking* $\sigma = 0$*, and assume that* Z *is an infinite variation jump process. We then recall the quadratic hedge problem and the pure investment problem, by deriving the PIDE verified by the value function* a *(Paragraph 7.2). We study the integro-differential operators associated to the model, by proving their continuity in the appropriate Hölder space of type 2, (Paragraph 7.3). We finally introduce a special sequence of smooth functions and we prove that they converge to the value function* a *in an appropriate functional space, which allows us to characterize the value function* a *and the pure investment optimal strategy, under a particular assumption on the jump size function appearing in the dynamic of* Z *(Paragraph 7.4.2). We then explain how to relax that particular assumption, and show that it is not too restrictive as it may seem (Section 7.5). We finally study the quadratic hedge problem by characterizing its value function and the optimal strategy (Section 7.6). We conclude the Chapter by studying the case when* Z *is a finite variation pure jump process (Section 7.7).*

Contents

7.1 Infinite activity processes: the model 122

7.2 The pure investment problem: HJB characterization . . 125

7.2.1 Formal derivation of the PIDE 125

7.3 Operators regularity 126

7.4 Smoothness and characterization of the function a . . . 133

7.4.1 The approximating sequence and its main properties . . . 133

7.4.2 Characterization of the function a 138

7.5 The change of variable 140

7.6 Smoothness and characterization of the function v^f . . . 143

7.7 Finite activity processes: the model 144

7.7.1 The pure investment problem in the finite variation case . 147

7.7.2 Proof of Lemma 7.23 151

7.7.3 The quadratic hedge problem in the finite variation case . 153

7.1 Infinite activity processes: the model

In Chapter 5 we introduced a $\mathbb{R}^3$-valued Markov process and the quadratic hedge problem was stated in terms of the value functions a and v^f. We also gave some a priori estimates and properties, in particular for the function a, and we finally deduced, with a verification argument, what type of PIDEs these functions have to satisfy. Chapter 6 was devoted to a complete analytic study of these PIDEs in the case of jump-diffusion processes, i.e. $\sigma_{min} > 0$. By using viscosity solution theory and contraction principles in Hölder spaces, we ended up with a complete characterization of the quadratic hedge problem in terms of unique smooth solution of a system of PIDEs. The fundamental Assumption $\sigma_{min} > 0$ has been used in order to apply a classical result of existence and uniqueness of smooth solutions for linear PDEs. Nevertheless, it is possible, under appropriate assumptions on the jump activity of the Poisson random measure, to repeat those arguments, even under the assumption $\sigma = 0$. This is the goal of this Chapter.

We propose a simplified model which allows us to focus on the main features of the quadratic hedge problem in the pure jump case: for this we denote then

$$dZ_r^{t,z} := \mu\left(r, Z_r^{t,u,z}\right) dr + \int_{\mathbb{R}} \gamma\left(r, Z_{r-}^{t,z}, y\right) \bar{J}\left(dydr\right), \; Z_t^{t,z} = z \tag{7.1}$$

for $t \in [0, T)$ and $z \in \mathbb{R}$ where, as usual, the stock price process S is defined to be $S = \exp(Z)$. In the rest of Chapter we will always assume that Assumptions 5.1 (with, of course, $\sigma_{min} = \sigma_{max} = 0$) hold true together with:

Assumption 7.1.

*[L]-**The Lévy measure**. The Lévy measure $\nu(dy)$ verifies*

$$\nu(dy) = \nu(y)dy \qquad \text{where} \qquad \nu(y) := g(y)|y|^{-(1+\alpha)}$$

for some $\alpha \in (1, 2)$, where g is a measurable function verifying $0 < m_g \leq g(y) \leq M_g$, $\forall y \in \mathbb{R}$, for some positive constants m_g, M_g. We also assume that

$$\lim_{y \to 0^-} g(y) = g(0^-) \quad \text{and} \quad \lim_{y \to 0^+} g(y) = g(0^+) \quad \text{with } g(0^+), g(0^-) > 0.$$

*[I]-**Integrability condition**. The function τ defined in Assumptions 5.1 verifies, for some $y_0 \in (0, 1)$ and some $m > 0$*

$$\sup_{0 < |y| \leq y_0} \frac{\tau(y)}{|y|} \leq m$$

*[ND]- **No degeneracy**. The function Γ in (5.7) verifies*

$$|\Gamma| := \int_{\mathbb{R}} \Gamma(y)\nu(dy) > 0$$

It is well known that there exists a unique semimartingale Z which solves the SDE defined above. Let us comment on these Assumptions, especially in view of Assumptions 5.1. The main difference here is that the process Z is only driven by a Poisson random measure, whose intensity measure has a precise structure. This is done since, as we will see, the non local linear operator arising from this process can be approximated, in a special sense, by the integro-differential operator associated to an α-stable Lévy process, and it is well known that this process has an infinitely differentiable density. We will use this density, and estimations on its derivatives, to prove that, in this case, non local linear PDEs do have a unique smooth solution. This essentially is the equivalent of the results we took from Ladyzenskaja et al. (1967).

The other main difference with the model in (5.5) is the fact that here we do not consider the processes U and P. Adding the process U in this model could be possible if one assumes some more regularity on the γ function, but it would have increased the technical complexity and decreased the clarity of our discussion. More interesting, instead, is the case of the process P: in the model (5.5), the non-degenerate volatility matrix $[\sigma;\sigma^U;\sigma^P]$ was the key property to deduce that linear parabolic PIDEs have smooth solutions. In this context, this role is played by the jump part: if we add the process P as in (5.5)

$$dP_r^{t,p} := \mu^P\left(r, P_r^{t,u,p}\right) dr + \int_{\mathbb{R}} \gamma^P\left(r, P_{r-}^{t,p}, y\right) \bar{J}\left(dydr\right),\ P_t^{t,p} = p$$

then the jump matrix becomes

$$\begin{pmatrix} \gamma(t,z,y) & 0 \\ \gamma^P(t,p,y) & 0 \end{pmatrix}$$

which is clearly degenerate: there will be no hope to prove regularity for the value function v^f in this case. On the other side, one could take

$$dP_r^{t,p} := \mu^P\left(r, P_r^{t,u,p}\right) dr + \int_{\mathbb{R}} \gamma^P\left(r, P_{r-}^{t,p}, y\right) \bar{N}\left(dydr\right)$$

where N is Poisson random measure independent from J, with intensity $\nu_n(dy) = g_n(y)|y|^{-(1+\alpha)}$: in this case the jump matrix will be

$$\begin{pmatrix} \gamma(t,z,y) & 0 \\ 0 & \gamma^P(t,p,y) \end{pmatrix}$$

which is non degenerate under appropriate assumptions on the function γ^P. The independence of J and N implies that one can easily repeat the argument we will expose in this chapter when considering the quadratic hedge problem in (5.11) for the couple (P,Z). Another choice would be to consider

$$dP_r^{t,p} := \mu^P\left(r, P_r^{t,u,p}\right) dr + \int_{\mathbb{R}} \gamma^{P,J}\left(r, P_{r-}^{t,p}, y\right) \bar{J}\left(dydr\right) + \int_{\mathbb{R}} \gamma^{P,N}\left(r, P_{r-}^{t,p}, y\right) \bar{N}\left(dydr\right)$$

This case can also be treated under some more restrictive assumptions on the functions $\gamma^{P,J}$ and $\gamma^{P,N}$. This shows why we prefer to consider only the process Z and privilege the clarity of our exposition.

Remark 7.2. *The no degeneracy condition in this context is given on the jump function γ. Remark that Assumptions 7.1-[**ND**] is equivalent to the following:*

There exists a Borel set with $0 < \nu(B) < +\infty$ such that

$$\inf_{t,z,y\in B} |\gamma(t,z,y)| \geq \epsilon > 0 \text{ for some } \epsilon > 0.$$

If this condition is true then trivially

$$\int \Gamma(y)\nu(dy) \geq \int_B \Gamma(y)\nu(dy) \geq \left(e^{-\epsilon} - 1\right)^2 \nu(B) > 0$$

If this condition was not true then for any Borel set B with $0 < \nu(B) < +\infty$ one has $\inf_{t,z,y\in B} |\gamma(t,z,y)| = 0$ and then one can find (t_n, z_n, y_n) such that $\gamma(t_n, z_n, y_n) \to 0$, $n \to \infty$. In particular

$$\inf_{y\in B} \Gamma(y) = \inf_{t,z,y\in B} \left(e^{\gamma(t,z,y)} - 1\right)^2 \leq \lim_{n\to\infty} \left(e^{\gamma(t_n,z_n,y_n)} - 1\right)^2 = 0$$

which implies $\int \Gamma(y)\nu(dy) = 0$.

It is clear then why we called it "no degeneracy" condition as in Assumptions 6.1 (where, we recall, we assumed $\sigma_{min} > 0$).

The quadratic hedge problem in this context becomes

$$\mathbf{QH}: \qquad \text{minimize } \mathbb{E}^{\mathbb{P}}\left[\left(f\left(Z_T^{0,z}\right) - X_T^{0,z,x,\theta}\right)^2\right]$$
$$\text{over } \theta \in \mathcal{X}(0,z,x)$$

where X is given in (5.9) and the set of admissible strategies is given in (5.10). The dynamic version of it is defined as:

$$v^f(t,z,x) := \inf_{\theta\in\mathcal{X}(t,z,x)} \mathbb{E}^{\mathbb{P}}\left[\left(f\left(Z_T^{t,z}\right) - X_T^{t,z,x,\theta}\right)^2\right] \tag{7.2}$$
$$v^f(T,z,x) = (f(z) - x)^2$$

From (5.41) we know that

$$v^f(t,z,x) = a(t,z)x^2 + b(t,z)x + c(t,z) \tag{7.3}$$

and when $f = 0$:

$$v^0(t,z,x) := x^2 \inf_{\theta\in\mathcal{X}(t,z,1)} \mathbb{E}\left[\left(1 + \int_t^T \theta_{r-} dS_r^{t,z}\right)^2\right] = x^2 a(t,z) \tag{7.4}$$

$$a(t,z) := \inf_{\theta\in\mathcal{X}(t,z,1)} \mathbb{E}\left[\left(1 + \int_t^T \theta_{r-} dS_r^{t,z}\right)^2\right] \tag{7.5}$$

We already know many properties of the *pure investment problem* above (Lemma 5.2). Furthermore Assumptions 5.1– 7.1 allow us to use Lemma 5.3 to obtain

$$e^{-C(T-t)} \leq a(t,z) \leq 1 \qquad \text{where} \qquad C := \frac{2(\|\tilde{\mu}\|^2 + \|\tau\|_{1,\nu}^2)}{|\Gamma|} \tag{7.6}$$

whereas Theorem 5.4 gives us the existence of $T^* > 0$ and $K^a_{lip} \geq 0$ such that if $T < T^*$ then

$$|a(t, z') - a(t, z)| \leq K^a_{lip}|z - z'| \tag{7.7}$$

with $T^* \to +\infty$ when $K_{max} \to 0$. Remark that these results hold true without making any specific assumption on the form of the Lévy measure $\nu(dy)$.

7.2 The pure investment problem: HJB characterization

7.2.1 Formal derivation of the PIDE

From now on we will work with the Hölder spaces of type 2: $H^{\alpha+\delta}([0,T] \times \mathbb{R})$ for some $\delta \in [0,1]$ (which of course is not the one of Assumptions 6.1-[**I**]) to be determined. The Hölder space of type 2 is defined in Appendix C, paragraph C.3. Remark that in Chapter 6 we proved that a belongs to some Hölder space of type 1: $C^{\kappa/2+1,\kappa+2}([0,T] \times \mathbb{R}^2)$ (here $\kappa = 1 - \delta$, δ now given in Assumptions 6.1-[**I**]), so in particular, twice continuously differentiable w.r.t. the space variable, and once w.r.t. the time variable.

In this case, we can restrict ourselves to the Hölder space of type 2 since, in the pure jump case, Itô's formula can be used if $a \in C^{1,1}(\mathbb{R})$ with Hölder condition on $\partial_z a$ (see Theorem D.1, Appendix D), so the natural choice is to look at solution in a less constraining functional space, i.e. $a \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$ and $a(.,z) \in C^1([0,T])$ for all $z \in \mathbb{R}$.

For sake of clarity, we recall the integro-differential operators associated to the process Z in the pure jump case:

Definition 7.3. *Let $\delta \in (0,1]$. For a real valued function $\varphi \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$ we denote*

$$\mathcal{A}\varphi(t,z) := \mu(t,z)\frac{\partial \varphi}{\partial z}(t,z)$$

$$\mathcal{B}\varphi(t,z) := \int_{\mathbb{R}} \left(\varphi(t, z + \gamma(t,z,y)) - \varphi(t,z) - \gamma(t,z,y)\frac{\partial \varphi}{\partial z}\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)$$

$$\mathcal{Q}\varphi(t,z) := \tilde{\mu}(t,z)\varphi(t,z) + \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1\right)\left(\varphi(t, z + \gamma(t,z,y)) - \varphi(t,z)\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)$$

$$\mathcal{G}\varphi(t,z) := \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1\right)^2 \varphi(t, z + \gamma(t,z,y))\nu(dy)$$

Remark that $\mathcal{B}$ is well defined in the Hölder space $H^{\alpha+\delta}([0,T] \times \mathbb{R})$. According to Theorem 5.11, we need to prove that the PIDE

$$0 = -\frac{\partial a}{\partial t} - \mathcal{A}a - \mathcal{B}a - \mathcal{H}[a], \qquad a(T,u,z) = 1 \tag{7.8}$$

has a unique solution in $H^{\alpha+\delta}([0,T] \times \mathbb{R})$, which also is strictly positive and continuously differentiable w.r.t. t, where

$$\mathcal{H}[a] := \inf_{|\pi|\leq \bar{\Pi}} \left\{2\pi\mathcal{Q}a + \pi^2\mathcal{G}a\right\} \tag{7.9}$$

and $\bar{\Pi}$ is defined in 5.44.

7.3 Operators regularity

In order to prove that the PIDE (7.8) has a unique smooth solution, we need to study the regularity of the operators introduced in Definition 7.3.

Lemma 7.4. *Suppose that Assumptions 5.1 and Assumptions 7.1-*[**I**] *hold true. Then there exists a positive constant $M > 0$ such that for all $\epsilon \in (0,1)$*

$$\|\mathcal{A}\varphi\|_\infty + \|\mathcal{G}\varphi\|_\infty + \|\mathcal{Q}\varphi\|_\infty \leq M\left(\epsilon^\delta \|\varphi\|_{1+\delta,H} + \epsilon^{-1}\|\varphi\|_\infty\right)$$

for all $\varphi \in H^{1+\delta}([0,T]\times\mathbb{R})$, and

$$\|\mathcal{A}\varphi\|_{\delta,H} + \|\mathcal{G}\varphi\|_{\delta,H} + \|\mathcal{Q}\varphi\|_{\delta,H} \leq M\left(\epsilon^{\alpha-\delta} \|\varphi\|_{\alpha+\delta,H} + \epsilon^{-(1+\delta)}\|\varphi\|_\infty\right)$$

for all $\varphi \in H^{\alpha+\delta}([0,T]\times\mathbb{R})$. The constant M does not depend on ϵ or φ.

Proof.
Definition 7.3 implies that $\|\mathcal{A}\varphi\|_\infty \leq M\|\varphi\|_{1,H}$ and $\|\mathcal{G}\varphi\|_\infty \leq M\|\varphi\|_\infty$ whereas $\|\mathcal{Q}\varphi\|_\infty \leq M\|\varphi\|_{1,H}$. It follows

$$\|\mathcal{A}\varphi\|_\infty + \|\mathcal{G}\varphi\|_\infty + \|\mathcal{Q}\varphi\|_\infty \leq \|\varphi\|_{1,H} \leq M\left(\epsilon^\delta \|\varphi\|_{1+\delta,H} + \epsilon^{-1}\|\varphi\|_\infty\right)$$

by using Proposition C.3.

Still from Definition 7.3 we obtain $\|\mathcal{G}\varphi\|_{\delta,H} \leq M\|\varphi\|_{\delta,H}$, $\|\mathcal{A}\varphi\|_{\delta,H} \leq M\|\varphi\|_{1+\delta,H}$ and $\|\mathcal{Q}\varphi\|_{\delta,H} \leq M\|\varphi\|_{1+\delta,H}$. It follows

$$\|\mathcal{A}\varphi\|_{\delta,H} + \|\mathcal{G}\varphi\|_{\delta,H} + \|\mathcal{Q}\varphi\|_{\delta,H} \leq M\|\varphi\|_{1+\delta,H} \leq M\left(\epsilon^{\alpha-\delta} \|\varphi\|_{\alpha+\delta,H} + \epsilon^{-(1+\delta)}\|\varphi\|_\infty\right)$$

□

For the operator $\mathcal{H}$ we have

Lemma 7.5. *Suppose that Assumptions 5.1 and Assumptions 7.1-*[**I**] *hold true. Then*

$$\mathcal{H} : H^{\alpha+\delta}([0,T]\times\mathbb{R}) \to H^\delta([0,T]\times\mathbb{R})$$

There exists a positive constant $M > 0$ such that for all $\epsilon \in (0,1)$

$$\|\mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi]\|_\infty \leq M\|\psi\|_{1,H} \leq M\left(\epsilon^\delta \|\psi\|_{1+\delta,H} + \epsilon^{-1}\|\psi\|_\infty\right)$$

for all $\varphi, \psi \in H^{1+\delta}([0,T]\times\mathbb{R})$. The constant M does not depend on ϵ or φ, ψ.

Proof.
Proving that $\mathcal{H}[\varphi] \in H^\delta([0,T]\times\mathbb{R})$ is straightforward.

As in Lemma 6.3 we use the concavity of the function H to get

$$\mathcal{H}[\psi] \leq \mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi] \leq \sup_{|\pi|\leq\bar{\Pi}} \left(2\pi\mathcal{Q}\psi + \pi^2\mathcal{G}\psi\right)$$

for which we deduce

$$\|\mathcal{H}[\varphi+\psi] - \mathcal{H}[\varphi]\|_\infty \leq (M(\|\mathcal{Q}\psi\|_\infty + \|\mathcal{G}\psi\|_\infty) \leq M\left(\epsilon^\delta \|\psi\|_{1+\delta,H} + \epsilon^{-1}\|\psi\|_\infty\right)$$

by Lemma 7.4.

□

Our next goal it to prove that it is possible to replace the operator $\mathcal{B}$ with the integro-differential operator associated to an α-stable Lévy process, if one assumes some more regularity of the jump function γ around $y = 0$. We start with the following assumption:

Assumption 7.6. *There exist two positive constants m_1, m_2 such that*

H1 *For any t, z the mapping $y \to \gamma(t,z,y)$ is twice continuously differentiable around zero and*

$$0 < m_1 \leq \inf_{t,z,|y|\leq y_0} |\gamma_y(t,z,y)| \text{and} \qquad \sup_{t,z,|y|\leq y_0} |\gamma_{yy}(t,z,y)| \leq m_2$$

for some $y_0 > 0$. It is not restrictive to assume that y_0 is the same as in Assumptions 7.1. In particular γ is invertible in $(-y_0, y_0)$: we call $\gamma^{-1}(t,z,y)$ its inverse.

H2 *For all $t, z \in [0,T) \times \mathbb{R}$ $\gamma_y(t,z,0) = 1$*

H3 *The function γ_y is Lipschitz continuous in the variable z:*

$$\sup_{t,z,|y|\leq y_0} |\gamma_y(t,z+h,y) - \gamma_y(t,z,y)| \leq m_2|h|$$

H4 *For all $y, y' \in (-y_0, 0) \cup (0, y_0)$ with $yy' > 0$*

$$|g(y) - g(y')| \leq m_2|y - y'|$$

i.e. the function g is Lipschitz continuous away from zero.

Let us comment on the fact that, among the above Assumptions, [**H2**] may seem to be very restrictive and many models do not verify it. However in Section 7.5 we will show how to avoid it.

Lemma 7.7. . *Let Assumptions 5.1–7.1 hold true together with Assumptions 7.6. For any $(t,z) \in [0,T] \times \mathbb{R}$ let*

$$\tilde{\nu}(t,z,y) := \frac{\nu(\gamma^{-1}(t,z,y))}{\gamma_y(t,\gamma^{-1}(t,z,y),y)} \tag{7.10}$$

where $0 < |y| \leq y_0$. There exists then a positive $M > 0$ such that for any t, z

$$|\tilde{\nu}(t,z+h,y) - \tilde{\nu}(t,z,y)| \leq M|h|\left(|\tilde{\nu}(t,z,y)| + |\tilde{\nu}(t,z+h,y)|\right) \tag{7.11}$$

$$|\tilde{\nu}(t,z,y) - \nu(y)| \leq M\left(|\nu(y)| + |\tilde{\nu}(t,z,y)|\right)|y| \tag{7.12}$$

Proof.
Let us start by remarking that the density $\tilde{\nu}$ is well defined on $\{|y| \leq y_0\}$ since the function γ is invertible. Let M be a positive constant which may change from line to line and, in order to simplify the exposition, we omit the dependence in the variable t. For (7.11) we can write

$$\left|\frac{\nu(\gamma^{-1}(z+h,y))}{\gamma_y(\gamma^{-1}(z+h,y))} - \frac{\nu(\gamma^{-1}(z,y))}{\gamma_y(\gamma^{-1}(z,y))}\right| \leq$$
$$|\tilde{\nu}(z,y)|\left|1 - \frac{\gamma_y(\gamma^{-1}(z,y))}{\gamma_y(\gamma^{-1}(z+h,y))}\right| + |\tilde{\nu}(z+h,y)|\left|1 - \frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\left|\frac{\gamma^{-1}(z+h,y)}{\gamma^{-1}(z,y)}\right|^{1+\alpha}\right|$$

The function γ_y is bounded from below (Assumptions 7.6-[**H1**]) and Lipschitz continuous w.r.t. z (Assumptions 7.6-[**H3**]). It follows then

$$\begin{aligned}
&\left|1 - \frac{\gamma_y(\ z,\gamma^{-1}(z,y))}{\gamma_y(z+h,\gamma^{-1}(z+h,y))}\right| \\
\leq &\frac{1}{m_1}\left(\left|\gamma_y(z+h,\gamma^{-1}(z+h,y)) - \gamma_y(z+h,\gamma^{-1}(z,y))\right| + \left|\gamma_y(z+h,\gamma^{-1}(z,y)) - \gamma_y(z,\gamma^{-1}(z,y))\right|\right) \\
\leq &\frac{1}{m_1}\left(\sup_{t,z,|y|\leq y_0}|\gamma_{yy}(t,z,y)|\,\gamma^{-1}(z+h,y) - \gamma^{-1}(z,y)\,|| + m_2|h|\right) \leq M|h|
\end{aligned}$$

since

$$\left|\frac{\partial\gamma^{-1}}{\partial z}(t,z,y)\right| = \left|\frac{\partial\gamma}{\partial z}(t,z,\gamma^{-1}(t,z,y))\left(\frac{\partial\gamma}{\partial y}(t,z,\gamma^{-1}(t,z,y))\right)^{-1}\right| \leq M \qquad (7.13)$$

For the other term we first write

$$\begin{aligned}
&\left|1 - \frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\left|\frac{\gamma^{-1}(z+h,y)}{\gamma^{-1}(z,y)}\right|^{1+\alpha}\right| \\
\leq &\left|1 - \frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\right| + \left|\frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\right|\left|1 - \left|\frac{\gamma^{-1}(z+h,y)}{\gamma^{-1}(z,y)}\right|^{1+\alpha}\right|
\end{aligned}$$

From (7.13) we have $\left|\gamma^{-1}(z+h,y) - \gamma^{-1}(z,y)\right| \leq K^{-1}|h|$ and since g is bounded from above and below we deduce

$$\left|1 - \frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\left|\frac{\gamma^{-1}(z+h,y)}{\gamma^{-1}(z,y)}\right|^{1+\alpha}\right| \leq \left|g(\gamma^{-1}(z+h,y)) - g(\gamma^{-1}(z,y))\right| + M|h|$$

Remark that $|\gamma^{-1}(z,y)| > 0$: if this is not true then for some $\hat{y} \in (0,y)$ (or $(y,0)$ if $y < 0$) one would have $\partial_y^{-1}(t,z,\hat{y}) = 0$ which contradicts the Assumption 7.6-[**H1**]. From the Lipschitz condition on g away from zero (Assumption 7.6-[**H4**]) we obtain

$$\left|1 - \frac{g(\gamma^{-1}(z,y))}{g(\gamma^{-1}(z+h,y))}\left|\frac{\gamma^{-1}(z+h,y)}{\gamma^{-1}(z,y)}\right|^{1+\alpha}\right| \leq M|h| \qquad (7.14)$$

so then

$$\left|\frac{\nu(\gamma^{-1}(z+h,y))}{\gamma_y(\gamma^{-1}(z+h,y))} - \frac{\nu(\gamma^{-1}(z,y))}{\gamma_y(\gamma^{-1}(z,y))}\right| \leq M|h| \left(|\tilde{\nu}(z,y)| + |\tilde{\nu}(z+h,y)|\right)$$

For (7.12) we write

$$\begin{aligned}&\left|\frac{\nu(\gamma^{-1}(z,y))}{\gamma_y(\gamma^{-1}(z,y))} - \nu(y)\right| \\ \leq &|\nu(y)| \left|1 - \frac{1}{\gamma_y(z,\gamma^{-1}(z,y))}\right| + |\tilde{\nu}(z,y)| \left|1 - \frac{g(y)}{g(\gamma^{-1}(z,y))}\left|\frac{\gamma^{-1}(z,y)}{y}\right|^{1+\alpha}\right|\end{aligned}$$

Assumptions 7.6-**H2** guarantees that $\gamma_y(z,0) = 1$ for all z: it follows then

$$\begin{aligned}\left|1 - \frac{1}{\gamma_y(z,\gamma^{-1}(z,y))}\right| \leq &\frac{1}{m_1}\left|\gamma_y(z,\gamma^{-1}(z,y)) - \gamma_y(z,0)\right| \\ \leq &\frac{1}{m_1}\sup_{t,z,|y|\leq y_0}|\gamma_{yy}(t,z,y)|\left|\gamma^{-1}(z,y)\right| \\ = &\frac{1}{m_1}\sup_{t,z,|y|\leq y_0}|\gamma_{yy}(t,z,y)|\left|\gamma^{-1}(z,y) - \gamma^{-1}(z,0)\right| \\ \leq &\frac{|y|}{m_1}\sup_{t,z,|y|\leq y_0}|\gamma_{yy}(t,z,y)|\frac{1}{\inf_{t,z,|y|\leq y_0}|\gamma_y(z,\gamma^{-1}(z,y))|} \leq M|y|\end{aligned}$$

since $\gamma^{-1}(z,0) = 0$. For the second term we have

$$\begin{aligned}\left|1 - \frac{g(y)}{g(\gamma^{-1}(z,y))}\left|\frac{\gamma^{-1}(z,y)}{y}\right|^{1+\alpha}\right| \leq &\left|1 - \frac{g(y)}{g(\gamma^{-1}(z,y))}\right| \\ + \left|\frac{g(y)}{g(\gamma^{-1}(z,y))}\right|\left|1 - \left|\frac{\gamma^{-1}(z,y)}{y}\right|^{1+\alpha}\right|&\end{aligned}$$

We can expand $\gamma^{-1}(z,y)$ around $y = 0$:

$$\gamma^{-1}(z,y) = \frac{1}{\gamma_y(z,0))}y - \frac{y^2}{2}\int_0^1 \frac{\gamma_{yy}(z,\theta y)}{(\gamma_y(z,\theta y))^2}d\theta$$

ans since $\gamma_y(z,0) = 1$ we deduce $\left|\gamma^{-1}(z,y) - y\right| \leq M|y|^2$. The same argument in (7.14) allows us to deduce

$$\left|1 - \frac{g(y)}{g(\gamma^{-1}(z,y))}\right| + \left|1 - \left|\frac{\gamma^{-1}(z,y)}{y}\right|^{1+\alpha}\right| \leq M|y| + o(y)$$

from which we conclude

$$\left|\frac{\nu(\gamma^{-1}(z,y))}{\gamma_y(\gamma^{-1}(z,y))} - \nu(y)\right| \leq M\left(|\nu(y)| + |\tilde{\nu}(z,y)|\right)\left(|y| + o(y)\right)$$

□

For $(t,z) \in [0,T) \times \mathbb{R}$ we introduce the following:

$$\bar{\mathcal{B}}_t\varphi(z) := \int_{\mathbb{R}} \left(\varphi(t,z+y) - \varphi(t,z) - y\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy) \tag{7.15}$$

$$\mathcal{B}^{st}_t\varphi(z) := \int_{\mathbb{R}} \left(\varphi(t,z+y) - \varphi(t,z) - y\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu^{st}(y)dy \tag{7.16}$$

where

$$\nu^{st}(y) := \frac{g(0^+)}{|y|^{1+\alpha}}\mathbb{1}_{\{0<y\}} + \frac{g(0^-)}{|y|^{1+\alpha}}\mathbb{1}_{\{y<0\}} \tag{7.17}$$

Both are integro-differential operators associated to Lévy processes, the second one, in particular, to some α-stable Lévy process. Our goal now is to prove that $\mathcal{B} - \mathcal{B}^{st}$ has a nice behavior in the Hölder space $H^{\alpha+\delta}([0,T]\times\mathbb{R})$, and this will allow us to replace the operator $\mathcal{B}$ in the PIDE (7.8) with $\mathcal{B}^{st}$:

$$-\frac{\partial a}{\partial t} - \mathcal{B}^{st}a = \mathcal{A}a + (\mathcal{B} - \mathcal{B}^{st})a + \mathcal{H}[a], \quad a(T,z) = 1$$

Proposition 7.8. *Suppose that Assumptions 5.1–7.1 hold true together with Assumptions 7.6 and let $\delta \in (0,1]$. There exists a positive constant $M > 0$ such that for any $\epsilon \in (0,1)$ and any $r \in (0,y_0)$ one has*

$$\left\|(\mathcal{B} - \bar{\mathcal{B}})\varphi\right\|_\infty \le M\left((r^{2-\alpha} + \epsilon^\delta |r|^{1-\alpha})\left\|\varphi\right\|_{1+\delta,H} + \epsilon^{-1}|r|^{1-\alpha}\left\|\varphi\right\|_\infty\right)$$

and

$$\left\|(\mathcal{B} - \bar{\mathcal{B}})\varphi\right\|_{\delta,H} \le M\left(\left(r^{2-\alpha} + r^\delta + \epsilon^{\min(\alpha-\delta,\delta)}r^{1-\alpha}\right)\left\|\varphi\right\|_{\alpha+\delta,H} + \epsilon^{-(1+\delta)}r^{1-\alpha}\left\|\varphi\right\|_\infty\right)$$

for any $\varphi \in H^{\alpha+\delta}([0,T]\times\mathbb{R})$. The constant M does not depend on φ, ϵ or r.

Proof.
Let $r \in (0,y_0)$ and split the operator $\mathcal{B} = \int_{|y|\le r}\cdots + \int_{|y|>r}$. In the first integral one has $|y| \le r \le y_0$ so we can invert the function γ since $\gamma_y(z,y) \neq 0$: the change the variable allows us to rewrite $\mathcal{B}$ as follows:

$$\begin{aligned}\mathcal{B}\varphi(t,z) := &\int_{|y|\le r}\left(\varphi(t,z+y) - \varphi(t,z) - y\frac{\partial\varphi}{\partial z}(t,z)\right)\tilde{\nu}(t,z,y)dy\\ &+ \int_{|y|>r}\left(\varphi(t,z+\gamma(t,z,y)) - \varphi(t,z) - \gamma(t,z,y)\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)\end{aligned}$$

where $\tilde{\nu}$ is given in (7.10). We obtain then $(\mathcal{B}_t - \bar{\mathcal{B}}_t)\varphi := \mathcal{F}^1_t\varphi + \mathcal{F}^2_t\varphi$ where

$$\mathcal{F}^1_t\varphi(z) := \int_{|y|\le r}\left(\varphi(t,z+y) - \varphi(t,z) - y\frac{\partial\varphi}{\partial z}(t,z)\right)(\tilde{\nu}(t,z,y) - \nu(y))\,dy$$

$$\mathcal{F}^2_t\varphi(z) := \int_{|y|>r}\left(\varphi(t,z+\gamma) - \varphi(t,z+y) - (\gamma - y)\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)$$

It follows

$$\left\|\mathcal{F}^1\varphi\right\|_\infty \leq M\left\|\varphi\right\|_{1,H} \sup_{t\leq T, z\in\mathbb{R}} \int_{|y|\leq r} |y|\left|\tilde{\nu}(t,z,y)-\nu(y)\right| dy$$

$$\left\|\mathcal{F}^2\varphi\right\|_\infty \leq M\left\|\varphi\right\|_{1,H}\left(1+\int_{r<|y|\leq 1}(\tau(y)+|y|)\nu(dy)\right) \leq M\left\|\varphi\right\|_{1,H} r^{1-\alpha}$$

By using (7.12) we get

$$\left\|\mathcal{F}^1\varphi\right\|_\infty \leq M\left\|\varphi\right\|_{1,H} \sup_{t\leq T, z\in\mathbb{R}} \int_{|y|\leq r} y^2\left(\left|\tilde{\nu}(t,z,y)\right|+\left|\nu(y)\right|\right) dy \leq M\left\|\varphi\right\|_{1,H} r^{2-\alpha}$$

and then

$$\left\|(\mathcal{B}-\bar{\mathcal{B}})\varphi\right\|_\infty \leq M\left(r^{2-\alpha}\left\|\varphi\right\|_{1+\delta,H}+|r|^{1-\alpha}\left\|\varphi\right\|_{1,H}\right)$$

We finally use Proposition C.3 to obtain

$$\left\|(\mathcal{B}-\bar{\mathcal{B}})\varphi\right\|_\infty \leq M\left((r^{2-\alpha}+\epsilon^\delta|r|^{1-\alpha})\left\|\varphi\right\|_{1+\delta,H}+\epsilon^{-1}|r|^{1-\alpha}\left\|\varphi\right\|_\infty\right)$$

and this proves the first inequality.

For the second inequality, we can use the above estimation on the $\mathbb{L}^\infty$−norm:

$$\begin{aligned}\left\|(\mathcal{B}-\bar{\mathcal{B}})\varphi\right\|_\infty \leq & M\left((r^{2-\alpha}+\epsilon^\delta|r|^{1-\alpha})\left\|\varphi\right\|_{1+\delta,H}+\epsilon^{-1}|r|^{1-\alpha}\left\|\varphi\right\|_\infty\right)\\ \leq & M\left((r^{2-\alpha}+\epsilon^\delta|r|^{1-\alpha})\left\|\varphi\right\|_{\alpha+\delta,H}+\epsilon^{-1}|r|^{1-\alpha}\left\|\varphi\right\|_\infty\right)\end{aligned} \tag{7.18}$$

We now need to estimate $\langle(\mathcal{B}-\bar{\mathcal{B}})\varphi\rangle^{(\delta)}_{z,Q_T}$. For the finite variation part it is straightforward to deduce

$$\left|\mathcal{F}^2\varphi(z+h)-\mathcal{F}^2\varphi(z)\right| \leq M|h|^\delta r^{1-\alpha}\left\|\varphi\right\|_{1+\delta,H}$$

Again from Proposition C.3 we obtain

$$\langle\mathcal{F}^2\varphi\rangle^{(\delta)}_{z,Q_T} \leq M\left(\epsilon^{\alpha-\delta}r^{1-\alpha}\left\|\varphi\right\|_{\alpha+\delta,H}+\epsilon^{-(1+\delta)}r^{1-\alpha}\left\|\varphi\right\|_\infty\right) \tag{7.19}$$

For the infinite variation part, since

$$\mathcal{F}^1_t\varphi(z) := \int_{|y|\leq r} y\int_0^1\left(\frac{\partial\varphi}{\partial z}(t,z+\theta y)-\frac{\partial\varphi}{\partial z}(t,z)\right)d\theta\left(\tilde{\nu}(t,z,y)-\nu(y)\right)dy$$

we have

$$\begin{aligned}\left|\mathcal{F}^1\varphi(t,z+h)-\mathcal{F}^1\varphi(t,z)\right| \leq & \int_0^1 d\theta\int_{|y|\leq r}|y|\left|\tilde{\nu}(t,z,y)-\nu(y)\right|\\ & \left|\frac{\partial\varphi}{\partial z}(t,z+h+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+h)-\frac{\partial\varphi}{\partial z}(t,z)\right| dy\\ +\int_{|y|\leq r}|y|\int_0^1 & \left|\frac{\partial\varphi}{\partial z}(t,z+h+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+h)\right| d\theta\left|\tilde{\nu}(t,z+h,y)-\tilde{\nu}(t,z,y)\right| d\theta dy\end{aligned}$$

Since

$$\left|\frac{\partial\varphi}{\partial z}(t,z+h+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+h)-\frac{\partial\varphi}{\partial z}(t,z)\right|$$
$$\leq\begin{cases}\|\varphi\|_{\alpha+\delta}\,|h|^{\alpha+\delta-1}\\ \\ \|\varphi\|_{\alpha+\delta}\,|\theta|^{\alpha+\delta-1}|y|^{\alpha+\delta-1}\end{cases}$$

we deduce

$$\left|\frac{\partial\varphi}{\partial z}(t,z+h+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+\theta y)-\frac{\partial\varphi}{\partial z}(t,z+h)-\frac{\partial\varphi}{\partial z}(t,z)\right|\leq\|\varphi\|_{\alpha+\delta}\,|h|^{\delta}|y|^{\alpha-1}$$

and then

$$\begin{aligned}\left|\mathcal{F}^1\varphi(z+h)-\mathcal{F}^1\varphi(z)\right|\leq&M\,\|\varphi\|_{\alpha+\delta}\,|h|^{\delta}\int_{|y|\leq r}|y|^{\alpha}\,|\tilde{\nu}(t,z,y)-\nu(y)|\,dy\\&+M\,\|\varphi\|_{\alpha+\delta}\int_{|y|\leq r}|y|^{\alpha+\delta}\,|\tilde{\nu}(t,z+h,y)-\tilde{\nu}(t,z,y)|\,dy\end{aligned}$$

Lemma 7.7 yields

$$\langle\mathcal{F}_t^1\varphi\rangle^{(1)}_{z,Q_T}\leq Mr^{\delta}\,\|\varphi\|_{\alpha+\delta,H}$$

Together with (7.18) and (7.19) we obtain

$$\left\|(\mathcal{B}-\bar{\mathcal{B}})\varphi\right\|_{\delta,H}\leq M\left(\left(r^{2-\alpha}+r^{\delta}+\epsilon^{\min(\alpha-\delta,\delta)}r^{1-\alpha}\right)\|\varphi\|_{\alpha+\delta,H}+\epsilon^{-(1+\delta)}r^{1-\alpha}\,\|\varphi\|_{\infty}\right)$$

□

Corollary 7.9. *Suppose that Assumptions 5.1–7.1 hold true together with Assumptions 7.6. There exists a positive constant $M>0$ such that for any $\epsilon\in(0,1)$ and any $r\in(0,y_0)$ one has*

$$\left\|(\mathcal{B}-\mathcal{B}^{st})\varphi\right\|_{\infty}\leq M\left((r^{2-\alpha}+\epsilon^{\delta}|r|^{1-\alpha})\,\|\varphi\|_{1+\delta,H}+\epsilon^{-1}|r|^{1-\alpha}\,\|\varphi\|_{\infty}\right)$$

and

$$\left\|(\mathcal{B}-\mathcal{B}^{st})\varphi\right\|_{\delta,H}\leq M\left(\left(r^{2-\alpha}+r^{\delta}+\epsilon^{\min(\alpha-\delta,\delta)}r^{1-\alpha}\right)\|\varphi\|_{\alpha+\delta,H}+\epsilon^{-(1+\delta)}r^{1-\alpha}\,\|\varphi\|_{\infty}\right)$$

for any $\varphi\in H^{\alpha+\delta}([0,T]\times\mathbb{R})$, where $\mathcal{B}^{st}$ is given in (7.16). The constant M does not depend on φ,ϵ or r.

***Proof*.**
We can easily estimate the difference $\bar{\mathcal{B}}-\mathcal{B}^{st}$:

$$\begin{aligned}\left\|(\bar{\mathcal{B}}-\mathcal{B}^{st})\varphi\right\|_{\infty}\leq&\left\|\int_{0^+}^{+\infty}\left(\varphi(t,z+y)-\varphi(t,z)-y\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\leq1\}}\right)\frac{g(y)-g(0^+)}{|y|^{1+\alpha}}dy\right\|_{\infty}\\&+\left\|\int_{-\infty}^{0^-}\left(\varphi(t,z+y)-\varphi(t,z)-y\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\leq1\}}\right)\frac{g(y)-g(0^-)}{|y|^{1+\alpha}}dy\right\|_{\infty}\end{aligned}$$

Remark that

$$|g(y) - g(0^+)| + |g(y) - g(0^-)| \leq M|y| \quad \text{if } 0 < |y| \leq y_0$$

from the definition of $g(0^+), g(0^-)$ given in Assumptions 7.1 and the Lipschitz continuity of g away from 0 (Assumptions 7.6-[**H4**]). We can use the same arguments of the proof of Proposition 7.8 to deduce

$$\left\|(\bar{\mathcal{B}} - \mathcal{B}^{st})\varphi\right\|_\infty \leq M\left((r^{2-\alpha} + \epsilon^\delta |r|^{1-\alpha}) \|\varphi\|_{1+\delta,H} + \epsilon^{-1}|r|^{1-\alpha} \|\varphi\|_\infty\right)$$

Since $\left\|(\mathcal{B} - \mathcal{B}^{st})\varphi\right\|_\infty \leq \left\|(\mathcal{B} - \bar{\mathcal{B}})\varphi\right\|_\infty + \left\|(\bar{\mathcal{B}} - \mathcal{B}^{st})\varphi\right\|_\infty$, the result follows from the above estimation and Proposition 7.8. In a similar way we can prove

$$\left\|(\mathcal{B} - \mathcal{B}^{st})\varphi\right\|_{\delta,H} \leq M\left(\left(r^{2-\alpha} + r^\delta + \epsilon^{\min(\alpha-\delta,\delta)} r^{1-\alpha}\right) \|\varphi\|_{\alpha+\delta,H} + \epsilon^{-(1+\delta)} r^{1-\alpha} \|\varphi\|_\infty\right)$$

□

7.4 Smoothness and characterization of the function a

7.4.1 The approximating sequence and its main properties

As in Section 6.2, we will build a sequence in the Hölder space $H^{\alpha+\delta}([0,T] \times \mathbb{R})$ which converges to the unique solution of (7.8). We first recall a basic result on linear PIDE with constant coefficients:

Theorem 7.10. *Let $\psi \in H^\lambda([0,T] \times \mathbb{R})$ for some $\lambda \in (0,1]$, $f \in H_e^{\alpha+\lambda}(\mathbb{R})$ and consider the following PIDE*

$$-\partial_t \varphi - \mathcal{B}_t^{st} \varphi = \psi, \quad \varphi(T,.) = f$$

where $\mathcal{B}^{st}$ is given in (7.16). *There exists a unique solution of the above PIDE in the Hölder space $H^{\alpha+\lambda}([0,T] \times \mathbb{R})$ which also is differentiable w.r.t. t. Furthermore*

$$\|\varphi\|_{\alpha+\lambda,H} \leq M\left(\|\psi\|_{\lambda,H} + \|f\|_{\alpha+\lambda,e}\right)$$

for some constant $M > 0$ which does not depend on φ and

$$\|\varphi(t,.) - \varphi(s,.)\|_{\frac{\alpha}{2}+\lambda,H} \leq M(t-s)^{1/2}\left(\|\psi\|_{\lambda,H} + \|f\|_{\alpha+\lambda,e}\right)$$

for all $0 \leq s \leq t \leq T$.

This result is stated in Mikulevicius and Pragarauskas (2009) (Lemma 7 and 17) or Mikulevicius and Pragarauskas (2011) (Lemma 8) when $f = 0$, since the operator $\mathcal{B}^{st}$ fulfills the assumptions of these Lemmas. When $f \neq 0$ the result can be adapted by means of the Feynman-Kac formula and the density of an α-stable process (See Appendix E). This is the equivalent of Theorem 5.1 in Ladyzenskaja et al. (1967), in the case of pure jump processes. By using the estimates on this density it is possible to relax the assumption $f \in H_e^{\alpha+\lambda}(\mathbb{R})$ and consider $f \in H_e^0(\mathbb{R})$.

For $\eta > 0$ let us consider the PIDE

$$-\frac{\partial \varphi}{\partial t} - \mathcal{B}^{st}\varphi + \eta\varphi = \mathcal{H}[\varphi] + (\mathcal{B} - \mathcal{B}^{st})\varphi + \mathcal{A}\varphi,\ \varphi(T,z) = e^{\eta T} \tag{7.20}$$

If φ is the unique solution of the above PIDE then $a(t,z) := e^{-\eta t}\varphi(t,z)$ is the unique solution of (7.8). As in paragraph 6.2.1 we fix $\varphi_0 \in H^{\alpha+\delta}([0,T]\times\mathbb{R})$ and consider the sequence

$$\begin{cases} \varphi^0 = \varphi_0 \\ -\frac{\partial}{\partial t}\varphi^{n+1} - \mathcal{B}^{st}\varphi^{n+1} + \eta\varphi^{n+1} = \ (\mathcal{B} - \mathcal{B}^{st})\varphi^n + \mathcal{H}[\varphi^n] + \mathcal{A}\varphi^n \\ \varphi^{n+1}(T,z) = e^{\eta T} \end{cases} \tag{7.21}$$

This sequence is well defined in the Hölder space $H^{\alpha+\delta}([0,T]\times\mathbb{R})$: by recurrence, if $\varphi^n \in H^{\alpha+\delta}([0,T]\times\mathbb{R})$ then by Lemmas 7.4–7.5 and Corollary 7.9 we have $r^n := (\mathcal{B} - \mathcal{B}^{st})\varphi^n + \mathcal{H}[\varphi^n] + \mathcal{A}\varphi^n \in H^{\delta}([0,T]\times\mathbb{R})$. Theorem 7.10 gives then that φ^{n+1} belongs to $H^{\alpha+\delta}([0,T]\times\mathbb{R})$ and verifies

$$\left\|\varphi^{n+1}\right\|_{\alpha+\delta,H} \leq M\left(e^{\eta T} + \left\|\left((\mathcal{B} - \mathcal{B}^{st}) + \mathcal{H} + \mathcal{A}\right)\varphi^n\right\|_{\delta,H}\right) \tag{7.22}$$

for some $M > 0$ which does not depend on η. Remark also that φ^n are all differentiable w.r.t. t. The Feynman-Kac formula also gives a probabilistic interpretation of the above linear PIDE

$$\varphi^{n+1}(t,z) = e^{\eta t} + \mathbb{E}\left[\int_t^T e^{-\eta(s-t)}\left((\mathcal{B} - \mathcal{B}^{st}) + \mathcal{H} + \mathcal{A}\right)\varphi^n(s, \tilde{Z}_s^{t,z})ds\right] \tag{7.23}$$

where

$$\tilde{Z}_s^{t,z} := z + \int_t^s \int_{\mathbb{R}} y\bar{J}^{\alpha}(dydr)$$

is the Lévy process associated to the integro-differential operator $\mathcal{B}^{st}$, i.e. J^{α} is a Poisson random measure whose intensity measure is given in (7.17). Although the functions φ^n may fail to be twice differentiable, the Feynman-Kac formula holds true: this is a direct consequence of the Itô's formula for pure jump processes (see Corollary D.2, Appendix D).
At this point it is clear why we decided to replace the operator $\mathcal{B}$ with the Lévy gradient $\mathcal{B}^{st}$: such Lévy processes have an infinitely differentiable density and a priori estimations on this density are available (see Appendix E).

We will now give the equivalent of Proposition 6.5 and Corollary 6.7 in the pure jump case.

Proposition 7.11. *Let Assumptions 5.1 –7.1 hold true together with Assumptions 7.6. If* $0 < \delta < \alpha - 1$ *then there exists a* $\eta^* > 0$ *such that for any* $\eta > \eta^*$ *the sequence* $(\varphi^n)_n$ *defined in* (7.21) *verifies*

$$\left\|\varphi^{n+1} - \varphi^n\right\|_{1+\delta,H} \leq (1+\eta)\left\|\varphi^1 - \varphi^0\right\|_{1+\delta,H}\beta^n$$

for some $\beta \in (0,1)$ which does not depend on η, φ^0 or φ^1. In particular $\varphi^n \to \varphi^* \in H^{1+\delta}([0,T]\times\mathbb{R})$. Furthermore

$$\sup_{n\in\mathbb{N}} \|\varphi^n\|_{1+\delta,H} \le c(\eta)$$

where $c(\eta)$ is a positive constant depending on η and φ^0. For any $\upsilon \in (0,1)$ there exists some positive constant $M_\upsilon > 0$ such that

$$\sup_{z\in\mathbb{R}} |\varphi^*(t,z) - \varphi^*(s,z)| \le M_\upsilon |t-s|^{\upsilon/2}$$

Proof.
The proof is really similar to the one we gave for Proposition 6.5, even more simple since the definition of Hölder norm of type 2 is less constraining: if $\triangle^{n+1} := \varphi^{n+1} - \varphi^n$ then

$$-\frac{\partial}{\partial t}\triangle^{n+1} - \mathcal{B}^{st}\triangle^{n+1} + \eta\triangle^{n+1} = \big((\mathcal{B}-\mathcal{B}^{st}) + \mathcal{A}\big)\triangle^{n-1} + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]$$
$$\triangle^{n+1}(T,z) = 0$$

Let now $r(t,z) := \Big(\big((\mathcal{B}-\mathcal{B}^{st}) + \mathcal{A}\big)\triangle^{n-1} + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\Big)(t,z)$ so that

$$\triangle^{n+1}(t,z) = \int_0^{T-t} e^{-\eta s}\int_{\mathbb{R}} r(s+t,\xi) m_s(\xi - z) d\xi ds$$

where m_s is the probability density of $\tilde{Z}_s^{0,z}$, for which estimations are given in Lemma E.1. Using this Lemma it is straightforward to deduce $\left\|\triangle^{n+1}\right\|_\infty \le M\eta^{-1}\|r\|_\infty$ and

$$\left\|D_z\triangle^{n+1}\right\|_\infty \le M\|r\|_\infty \int_0^{T-t} s^{-1/\alpha} ds$$

and the integral in the right-hand side is finite since $\alpha > 1$. We only need to estimate $\langle D_z\triangle^{n+1}\rangle^{(\delta)}_{z,Q_T}$:

$$\langle D_z\triangle^{n+1}\rangle^{(\delta)}_{z,Q_T} \le M\|r\|_\infty \sup_{t,z,0<|h|\le 1}\int_0^{T-t}\int_{\mathbb{R}} |D_z m_s(\xi - z - h) - D_z m_s(\xi - z)|\, d\xi ds$$

From Lemma E.1 we obtain

$$\int_{\mathbb{R}} |D_z m_s(\xi - z - h) - D_z m_s(\xi - z)|\, d\xi \le M s^{-\frac{1}{\alpha}}$$
$$\int_{\mathbb{R}} |D_z m_s(\xi - z - h) - D_z m_s(\xi - z)|\, d\xi \le M|h|\int_0^1 d\theta \int_{\mathbb{R}} |D_z^2 m_s(\xi - z - \theta h) d\xi \le |h| s^{-\frac{2}{\alpha}}$$

so then

$$\int_{\mathbb{R}} |D_z m_s(\xi - z - h) - D_z m_s(\xi - z)|\, d\xi \le M|h|^\delta s^{-\frac{2\delta}{\alpha}} s^{-\frac{1-\delta}{\alpha}} = M|h|^\delta s^{-\frac{1+\delta}{\alpha}}$$

We can use this estimation to obtain

$$\langle D_z\triangle^{n+1}\rangle^{(\delta)}_{z,Q_T} \le M\|r\|_\infty \int_0^{T-t} s^{-\frac{1+\delta}{\alpha}} ds$$

where the integral in the right hand side is finite since $\delta < \alpha - 1$: we can conclude $\left\|\triangle^{n+1}\right\|_{1+\delta,H} \leq M \|r\|_\infty$, or, equivalently

$$\begin{aligned}\left\|\triangle^{n+1}\right\|_{1+\delta,H} &\leq M \left\|\left((\mathcal{B}-\mathcal{B}^{st})+\mathcal{A}\right)\triangle^n + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty \\ \left\|\triangle^{n+1}\right\|_\infty &\leq M\eta^{-1} \left\|\left((\mathcal{B}-\mathcal{B}^{st})+\mathcal{A}\right)\triangle^n + \mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty\end{aligned}$$

From Lemmas 7.4-7.5 and Corollary 7.9 we obtain

$$\begin{aligned}\|\mathcal{A}\triangle\varphi^n\|_\infty &\leq M\left(\epsilon^\delta \|\triangle^n\|_{1+\delta,H} + \epsilon^{-1}\|\triangle^n\|_\infty\right) \\ \left\|\mathcal{H}[\varphi^n] - \mathcal{H}[\varphi^{n-1}]\right\|_\infty &\leq M\left(\epsilon^\delta \|\triangle^n\|_{1+\delta,H} + \epsilon^{-1}\|\triangle^n\|_\infty\right) \\ \left\|(\mathcal{B}-\mathcal{B}^{st})\triangle\varphi^n\right\|_\infty &\leq M\left((r^{2-\alpha} + \epsilon^\delta |r|^{1-\alpha})\|\triangle^n\|_{1+\delta,H} + \epsilon^{-1}|r|^{1-\alpha}\|\triangle^n\|_\infty\right)\end{aligned}$$

which implies

$$\begin{aligned}\left\|\triangle^{n+1}\right\|_{2,H} &\leq M\left(\left(r^{2-\alpha} + \epsilon^\delta r^{1-\alpha}\right)\|\triangle^n\|_{1+\delta,H} + \epsilon^{-1}r^{1-\alpha}\|\triangle^n\|_\infty\right) \\ \left\|\triangle^{n+1}\right\|_\infty &\leq M\eta^{-1}\left(\left(r^{2-\alpha} + \epsilon^\delta r^{1-\alpha}\right)\|\triangle^n\|_{1+\delta,H} + \epsilon^{-1}r^{1-\alpha}\|\triangle^n\|_\infty\right)\end{aligned}$$

We can now repeat the same argument of the proof of Proposition 6.5 to deduce

$$\left\|\triangle^{n+1}\right\|_{1+\delta,H} \leq (1+\eta)\beta^n \left\|\varphi^1 - \varphi^0\right\|_{1+\delta,H}$$

for some $\beta \in (0,1)$ which does not depend on η. It follows that $(\varphi^n)_n$ is a Cauchy sequence in $H^{1+\delta}([0,T]\times\mathbb{R})$ and then converges to some $\varphi^* \in H^{1+\delta}([0,T]\times\mathbb{R})$. In particular

$$\sup_{n\in\mathbb{N}} \|\varphi^n\|_{1+\delta,H} \leq c(\eta)$$

for some constant which depends on η. Furthermore, from (7.21) with $\eta = 0$, we have for $t' < t$

$$\begin{aligned}\left|\varphi^{n+1}(t,z) - \varphi^{n+1}(s,z)\right| \leq &M\left\|\left(\mathcal{B}-\mathcal{B}^{st}+\mathcal{A}+\mathcal{H}\right)\varphi^n\right\|_\infty \int_t^T ds \int_R |m_{T-t}(\xi) - m_{T-t'}(\xi)|\, d\xi \\ &+ M\left\|\left(\mathcal{B}-\mathcal{B}^{st}+\mathcal{A}+\mathcal{H}\right)\varphi^n\right\|_\infty \int_{t'}^t ds \int_R |m_{T-t'}(\xi)|\, d\xi\end{aligned}$$

First remark that

$$\sup_n \left\|\left(\mathcal{B}-\mathcal{B}^{st}+\mathcal{A}+\mathcal{H}\right)\varphi^n\right\|_\infty \leq M \sup_n \|\varphi^n\|_{1+\delta,H} \leq M$$

It follows

$$\left|\varphi^{n+1}(t,z) - \varphi^{n+1}(s,z)\right| \leq M\left(\int_t^T ds \int_R |m_{T-t}(\xi) - m_{T-t'}(\xi)|\, d\xi + |t-t'|\right)$$

From Lemma E.1 we have

$$\int_R |m_{T-t}(\xi) - m_{T-t'}(\xi)|\, d\xi \leq \int_R \left(m_{T-t}(\xi) + m_{T-t'}(\xi)\right) d\xi \leq M$$

but also

$$\int_R |m_{T-t}(\xi) - m_{T-t'}(\xi)|\, d\xi \leq |t-t'| \int_0^1 \int_{\mathbb{R}} \left|\partial_t m_{T-t'+\theta(t-t')}(\xi)\right| d\xi d\theta \leq M(T-t)^{-2/\alpha}$$

so then, by using both these estimates, we obtain, for any $\upsilon \in (0,1)$

$$\int_R |m_{T-t}(\xi) - m_{T-t'}(\xi)|\, d\xi \leq M_\upsilon |t-t'|^{\upsilon/2}(T-t)^{-\frac{\upsilon}{\alpha}}$$

for some positive constant M_υ depending on υ. Finally

$$\sup_z \left|\varphi^{n+1}(t,z) - \varphi^{n+1}(s,z)\right| \leq M_\upsilon |t-t'|^{\upsilon/2}$$

It follows then that, when we pass to the limit $n \to \infty$, φ^* inherits the same property:

$$\sup_{z\in\mathbb{R}} |\varphi^*(t,z) - \varphi^*(s,z)| \leq M|t-s|^{\upsilon/2}$$

□

Remark 7.12. *By using Theorem 7.10 one would obtain, in particular*

$$\left\|\varphi^{n+1}(t,.) - \varphi^{n+1}(s,.)\right\|_\infty \leq M(t-s)^{1/2}\left(1 + \left\|\left(\mathcal{B} - \mathcal{B}^{st} + \mathcal{A} + \mathcal{H}\right)\varphi^n\right\|_{\lambda,H}\right)$$

for $\lambda \in (0,1]$. But we only know how to estimate

$$\left\|\left(\mathcal{B} - \mathcal{B}^{st} + \mathcal{A} + \mathcal{H}\right)\varphi^n\right\|_\infty$$

which corresponds to the case $\lambda = 0$. So we cannot directly apply the above estimate to deduce that $t \to \varphi^*(t,z)$ is 1/2-Hölder continuous, but just $\upsilon/2$-Hölder continuous for any $\upsilon \in (0,1)$. Nevertheless, we will improve this estimate and show that the map $t \to \varphi^*(t,z)$ is 1/2-Hölder continuous.

As explained in Remark 6.6 we can say, with an abuse of language, that the above Proposition holds true for all $\eta > 0$. For the uniqueness we have

Corollary 7.13. *Let Assumptions 5.1 –7.1 hold true together with Assumptions 7.6 and fix $\delta \in (0, \alpha - 1)$ as in Proposition 7.11. Then the PIDE* (7.8) *has at most one solution in the Hölder space $H^{\alpha+\delta}([0,T] \times \mathbb{R})$.*

Proof.
Proving that PIDE (7.8) has a unique solution is equivalent to prove that (7.20) has a unique solution. We can then follow the argument of Corollary 6.7: if φ^i, $i = 1, 2$ are two solutions of PIDE (7.20) and $\varphi^{n,i}$ is the sequence given in (7.21) where $\varphi^{0,i} = \varphi^i$ for $i = 1, 2$, then, by construction it is clear that $\varphi^{n,i} = \varphi^i$ for all n. If $\Delta^n := \varphi^{n,1} - \varphi^{n,2}$ then, as in the proof of Proposition 7.11, one can prove that

$$\left\|\Delta^{n+1}\right\|_{1+\delta,H} \leq (1+\eta)\beta^{n+1}\left\|\varphi^1 - \varphi^2\right\|_{1+\delta,H}$$

for η big enough and $\beta \in (0,1)$. In particular $\Delta^n \to 0$ in $\mathbb{L}^\infty(\mathbb{R}^2)$, and since $\Delta^n = \varphi^1 - \varphi^2$, we conclude that $\varphi^1 = \varphi^2$.

□

Proposition 7.11 stated that the sequence φ^n converges in the Hölder space $H^{1+\delta}([0,T]\times\mathbb{R})$ to some φ^*. Let $n\to\infty$ in (7.23) for $\eta=0$ to deduce

$$\varphi^*(t,z) = 1+\mathbb{E}\left[\int_t^T \left((\mathcal{B}-\mathcal{B}^{st})+\mathcal{H}+\mathcal{A}\right)\varphi^*(s,\tilde{Z}_s^{t,z})ds\right] \tag{7.24}$$

since the operators $\mathcal{B}-\mathcal{B}^{st}$, $\mathcal{A}$ and $\mathcal{G}$ are continuous in $H^{1+\delta}([0,T]\times\mathbb{R})$. We denote with $\hat{\pi}$ the optimal control related to φ^* in the right hand side of the above equality:

$$\hat{\pi}(t,u,z) := -\bar{\Pi}\vee-\frac{\mathcal{Q}\varphi^*(t,z)}{\mathcal{G}\varphi^*(t,z)}\wedge\bar{\Pi} \tag{7.25}$$

The regularity of φ^* shows that $\hat{\pi}$ is well defined and bounded. Furthermore, from the definition of $\mathcal{G}$ and $\mathcal{Q}$, we deduce

$$\left|\hat{\pi}(t,z)-\hat{\pi}(t,z')\right|\leq M|z-z'|^\delta$$

for any t,z,z', since $\varphi^*\in H^{1+\delta}([0,T]\times\mathbb{R})$. In other words, $\hat{\pi}\in H^\delta([0,T]\times\mathbb{R})$ and, by using the regularity condition w.r.t. t given in Proposition 7.11, we also have that $t\to\hat{\pi}(t,z)$ is Hölder continuous.

7.4.2 Characterization of the function a

We now are able to prove that the function a given in (7.5) is the unique smooth solution of PIDE (7.8).

Theorem 7.14. *Let Assumptions 5.1 –7.1 hold true together with Assumptions 7.6. The PIDE* (7.8) *has a unique, smooth and strictly positive solution* $a\in H^{\alpha+\delta}([0,T]\times\mathbb{R})$*, for* $\delta\in(0,\alpha-1)$*. The function* $t\to a(t,z)$ *is also continuously differentiable in* $(0,T)$ *and*

$$\sup_{z\in\mathbb{R}}\|a(t,z)-a(s,z)\|_\infty\leq M(t-s)^{1/2}$$

Moreover

$$\|\varphi^n-a\|_{1+\delta,H}\leq M\beta^n,\quad n\to\infty$$

for some $M>0$ *and* $\beta\in(0,1)$*.*

Proof.
The proof is really similar to the one of Theorem 6.8, so that we will skip all similar computations. As in Step 1 of the proof of Theorem 6.8, we first prove that φ^* is the unique viscosity solution of

$$-\frac{\partial\varphi^*}{\partial t}-\mathcal{A}\varphi^*-\mathcal{B}\varphi^*-2\hat{\pi}\mathcal{Q}\varphi^*-\hat{\pi}^2\mathcal{G}\varphi^*=0,\quad \varphi^*(T,.)=1 \tag{7.26}$$

where $\hat{\pi}$ is given in (7.25). We then prove that PIDE (7.26) admits a unique smooth solution: for this, let $\eta>0$ and consider the map Ξ_η defined as follows:

$$-\frac{\partial\Xi_\eta(\psi)}{\partial t}-\mathcal{B}^{st}\Xi_\eta(\psi)+\eta\Xi_\eta(\psi)=\left(\mathcal{B}-\mathcal{B}^{st}\right)\psi+\mathcal{A}\psi+2\hat{\pi}\mathcal{Q}\psi+\hat{\pi}^2\mathcal{G}\psi \tag{7.27}$$
$$\Xi_\eta(\psi)(T,.)=e^{\eta T}$$

for $\psi \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$. Since $\hat{\pi} \in H^{\delta}([0,T] \times \mathbb{R})$, we can apply Lemma 7.4 and Corollary 7.9 to deduce that

$$\left(\mathcal{B} - \mathcal{B}^{st}\right) \psi + \mathcal{A}\psi + 2\hat{\pi}\mathcal{Q}\psi + \hat{\pi}^2\mathcal{G}\psi \in H^{\delta}([0,T] \times \mathbb{R})$$

Theorem 7.10 ensures that $\Xi_\eta(\psi)$ is well defined in the Hölder space $H^{\alpha+\delta}([0,T]\times\mathbb{R})$, it is continuously differentiable w.r.t. t and, for some $M > 0$ not depending on η or ψ, one has

$$\begin{aligned} \|\Xi_\eta(\psi)\|_{\alpha+\delta,H} \leq & M \left(e^{\eta T} + \left\| \left(\mathcal{B} - \mathcal{B}^{st} + \mathcal{A} + 2\hat{\pi}\mathcal{Q} + \hat{\pi}^2\mathcal{G}\right) \psi \right\|_{\delta,H} \right) \\ \|\Xi_\eta(\psi)(t,.) - \Xi_\eta(\psi)(s,.)\|_{\frac{\alpha}{2}+\delta,H} \leq & M(t-s)^{1/2} \left(e^{\eta T} + \right. \\ & \left. \left\| \left(\mathcal{B} - \mathcal{B}^{st} + \mathcal{A} + 2\hat{\pi}\mathcal{Q} + \hat{\pi}^2\mathcal{G}\right) \psi \right\|_{\delta,H} \right) \end{aligned}$$

In particular $\Xi_\eta(\psi)$ is 1/2-Hölder w.r.t. t. By using the method developed in Step 2 of the proof of Theorem 6.8, together with Lemma 7.4 and Corollary 7.9, we prove that Ξ_η is a contraction for η big enough. Denote with ψ^* its unique fixed point. It follows in particular

$$\begin{aligned} & \sup_{z \in R} \|\psi^*(t,z) - \psi^*(s,z)\|_\infty = \|\Xi_\eta(\psi^*)(t,.) - \Xi_\eta(\psi^*)(s,.)\|_\infty \\ \leq & M(t-s)^{1/2} \left(e^{\eta T} + \left\| \left(\mathcal{B} - \mathcal{B}^{st}\right) \psi^* + \mathcal{A}\psi^* + 2\hat{\pi}\mathcal{Q}\psi^* + \hat{\pi}^2\mathcal{G}\psi^* \right\|_{\delta,H} \right) \\ \leq & M(t-s)^{1/2} \left(1 + \|\psi^*\|_{\alpha+\delta,H} \right) \end{aligned} \tag{7.28}$$

so then ψ^* is 1/2-Hölder continuous w.r.t. t. But

$$\begin{aligned} & -\frac{\partial \Xi_\eta(\psi^*)}{\partial t} - \mathcal{B}^{st}\Xi_\eta(\psi^*) + \eta\Xi_\eta(\psi^*) = \left(\mathcal{B} - \mathcal{B}^{st}\right) \psi^* + \mathcal{A}\psi^* + 2\hat{\pi}\mathcal{Q}\psi^* + \hat{\pi}^2\mathcal{G}\psi^* \\ & \Xi_\eta(\psi^*)(T,.) = e^{\eta T} \end{aligned}$$

so that $\Xi_\eta(\psi^*)$ is continuously differentiable w.r.t. t (Theorem 7.10), and then $\psi^* = \Xi_\eta(\psi^*)$ also is. By the uniqueness of the viscosity solution we deduce that $\varphi^*(t,z) = e^{-\eta t}\psi^*(t,z)$ and it belongs to $H^{\alpha+\delta}([0,T] \times \mathbb{R})$.

We complete the proof by following the Step 3 of the proof of Theorem 6.8: let

$$d\hat{X}_s^{t,z,x} := \hat{\pi}_{s-}\hat{X}_{s-}^{t,z,x} e^{-Z_{s-}^{t,z}} de^{Z_s^{t,z}}, \; \hat{X}_t^{t,z,x} = x$$

which is well defined since $\hat{\pi}$ is bounded. The function

$$w(t,z,x) := \mathbb{E}\left[\left(\hat{X}_T^{t,z,x} \right)^2 \right] \tag{7.29}$$

is continuous and $w(t,z,x) = x^2\tilde{\varphi}(t,z)$ for some $\tilde{\varphi}$. The continuity of w and the Markov property of Z allow us to write

$$w(t,z,x) = \mathbb{E}\left[w\left(t+h, Z_{t+h}^{t,z}, \hat{X}_{t+h}^{t,z,x} \right) \right]$$

which, in particular, proves that $\tilde{\varphi}$ is a viscosity solution of (7.26), and again from the uniqueness of the viscosity solution of this PIDE, we deduce $\tilde{\varphi} = \varphi^*$. In particular

$$x^2\varphi^*(t,z) = x^2\tilde{\varphi}(t,z) = \mathbb{E}\left[\left(\hat{X}_T^{t,z,x}\right)^2\right] \geq v^0(t,z,x) = x^2 a(t,z)$$

because $\hat{X}$ is an admissible portfolio and v^0 is defined in (7.4). From the above estimation and Lemma 5.3 we deduce $e^{-CT} < a(t,z) \leq \varphi^*(t,z)$ for all t,z.

To summarize we proved that $\varphi^* \in H^{\alpha+\delta}([0,T]\times\mathbb{R})$, it verifies the PIDE (7.8) and it also is strictly positive: we can then apply Theorem 5.11 to deduce $a = \varphi^*$. Finally we use Proposition 7.11 to obtain that $\varphi^n \to a$ in $H^{1+\delta}([0,T]\times\mathbb{R})$ and the estimate (7.28) to conclude our proof.

□

Remark 7.15. *The above result is the equivalent of Theorem 6.8 in the case of jump-diffusion. It is important to remark that the structure of the semi linear PIDE* (6.1) *is not the same as that PIDE* (7.8)*: in PIDE* (6.1) *the role of "regularizer" was played by the strictly elliptic matrix of second derivatives D^2a whereas in PIDE* (7.8) *this role is played by the non local Lévy operator $\mathcal{B}$. Nevertheless this does not change substantially the proof.*

The last thing we want to point out is that we work here with Hölder space of type 2, whereas we used Hölder spaces of type 1 in Theorem 6.8: as we explained at the beginning of paragraph 7.2.1, in the pure jump case we do not need to prove that a is twice differentiable w.r.t. z. It means that we do not need to assume any a priori regularity w.r.t. t for the driver $(\mathcal{B}-\mathcal{B}^{st})\psi + \mathcal{A}\psi + 2\hat{\pi}\mathcal{Q}\psi + \hat{\pi}^2\mathcal{G}\psi$ in the PDE (7.27)*. On the other hand, it was necessary to assume some Hölder regularity w.r.t. t for the driver $\mathcal{B}\psi + \mathcal{H}[\psi]$ in the PDE* (6.12) *to deduce the regularity of $\Xi_\eta(\psi)$. This explains why here we work in the functional space $H^{\alpha+\delta}([0,T]\times\mathbb{R})$, for which no regularity w.r.t. t is required.*

7.5 The change of variable

In Theorem 7.14 we proved that the function a can be characterized as the unique solution of a semi linear PIDE if one imposes, in particular, Assumptions 7.6. Remark that Assumption 7.6-[**H2**] could appear very restrictive: if for example the jump function is of the form $\gamma(t,z,y) = \hat{\gamma}(t,z)y$ then the only possible choice would be $\hat{\gamma}(t,z) = 1$ for all t,z.

Our goal here is to prove that it is possible to find a process $L_t = \phi(t,Z_t)$ for some smooth function ϕ such that

$$dL_s^{t,l} := \mu^L(s, L_s^{t,l})ds + \int \gamma^L(s, L_{s-}^{t,l}, y)\bar{J}(dyds) \tag{7.30}$$

with μ^L and γ^L verifying Assumptions 7.1–7.6, with especially $\partial_y\gamma^L(t,l,0) = 1$ for all t,l. If this is possible then one could rewrite problem (7.4) in terms of L instead of the state variable Z :

$$v^{0,L}(t,l,x) := \inf_{\theta \in \mathcal{X}(t,l)} \mathbb{E}^{\mathbb{P}} \left[\left(X_T^{t,l,x,\theta} \right)^2 \right], \quad v^{0,L}(T,l,x) := x^2$$

where

$$\begin{aligned} X_s^{t,l,x,\theta} &:= x + \int_t^s \theta_{r-} dS_r^{t,l} \\ \frac{dS_u^{t,l}}{S_{u-}^{t,l}} &:= \tilde{\mu}(u, \phi^{-1}(u, L_u^{t,l}))du + \int \left(e^{\gamma(u,\phi^{-1}(u,L_u^{t,l}),y)} - 1 \right) \bar{J}(dydu) \end{aligned}$$

It is clear that the new value function $v^{0,L}$ and v^0 in (7.4) are related: $v^0(t,z,x) = v^{0,L}(t,\phi(t,z),x)$. Similarly the function

$$a^L(t,l) := \inf_{\theta \in \mathcal{X}(t,l,1)} \mathbb{E} \left[\left(1 + \int_t^T \theta_{u-} dS_u^{t,l} \right)^2 \right] \tag{7.31}$$

is related to the function a defined in (7.5) by $a(t,z) = a^L(t,\phi(t,z))$. With the same argument of Section 5.5 one can derive the PIDE verified by a^L and find that it has the same structure of PIDE (7.8), where in particular the operator $\mathcal{B}$ is replaced with

$$\mathcal{B}^L\varphi(t,l) := \int_{\mathbb{R}} \left(\varphi(t, l + \gamma^L(t,l,y)) - \varphi(t,l) - \gamma^L(t,l,y)\frac{\partial \varphi}{\partial l}(t,l)\mathbb{1}_{\{|y|\le 1\}} \right) \nu(dy)$$

We can then apply Theorem 7.14 to a^L (since by construction γ^L verifies the Assumptions 7.6) and the prove the regularity of the function a simply by using the regularity of the function a^L. This explains why it is not restrictive to suppose $\partial_y \gamma(t,z,0) = 1$.

Let then ϕ be a real valued function defined on $[0,T] \times \mathbb{R}$ and $L_t := \phi(t, Z_t)$. Assume that for all t the function $z \to \phi(t,z)$ is invertible and that ϕ is smooth enough to apply Itô's formula. We obtain

$$\gamma^L(t,l,y) := \phi(t, \phi^{-1}(t,l) + \gamma(t, \phi^{-1}(t,l), y)) - l \tag{7.32}$$

$$\begin{aligned} \mu^L(t,l) := &\frac{\partial \phi}{\partial t}(t, \phi^{-1}(t,l)) + \mu(t, \phi^{-1}(t,l))\frac{\partial \phi}{\partial z}(t, \phi^{-1}(t,l)) \\ &+ \int_{|y|\le 1} \left(\gamma^L(t,l,y) - \gamma(t,\phi^{-1}(t,l),y)\frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l)) \right) \nu(dy) \end{aligned} \tag{7.33}$$

In particular one has

$$\gamma_y^L(t,l,0) = \frac{\partial \phi}{\partial z}(t, \phi^{-1}(t,l)))\gamma_y(t, \phi^{-1}(t,l), 0)$$

If we select for example

$$\phi(t,z) := \int_0^z \frac{ds}{\gamma_y(t,s,0)} \tag{7.34}$$

then trivially $\gamma_y^L(t,l,0) = 1$ for all t,l. The following Lemma shows that this choice guarantees that the coefficients μ^L and γ^L verify Assumptions 5.1 –7.1 and Assumptions 7.6.

Lemma 7.16. *Assume that there exist some positive constants* $0 < m_1, m_2$ *such that*

i). For all $t, z \in [0,T] \times \mathbb{R}$ *the mapping* $y \to \gamma(t,z,y)$ *is differentiable at* $y = 0$ *and*

$$0 < m_1 \le |\gamma_y(t,z,0)| \le m_2 \text{ for all } t,z \in [0,T] \times \mathbb{R}$$

ii). The function $(t,z) \to \gamma_y(t,z,0)$ *is differentiable and*

$$\left|\frac{d}{dt}\gamma_y(t,z,0)\right| + \left|\frac{d}{dz}\gamma_y(t,z,0)\right| \le m_2 \text{ for all } t,z \in [0,T] \times \mathbb{R}$$

iii). The function $z \to \frac{d}{dt}\gamma_y(t,z,0)$ *is Lipschitz continuous:*

$$\left|\frac{d}{dt}\gamma_y(t,z,0) - \frac{d}{dt}\gamma_y(t,z',0)\right| \le m_2|z-z'| \text{ for all } t \in [0,T],\ z,z' \in \mathbb{R}$$

Then the functions μ^L *and* γ^L *defined in* (7.32)–(7.33) *with the choice of* ϕ *given by* (7.34) *verify Assumptions 5.1–7.1 and Assumptions 7.6.*

Proof.
First we remark that we can assume $m_1 \ge 1$ otherwise we can normalize the process Z_t by m_1: $\tilde{Z}_t := Z_t/m_1$ so then the new jump function will verify $1 \le \tilde{\gamma}_y(t,\tilde{z},0)$ for all $t, \tilde{z}$. Also, under assumptions $i)$ and $ii)$ the function

$$\phi(t,z) := \int_0^z \frac{ds}{\gamma_y(t,s,0)}$$

is well defined, invertible and $\|\partial\phi/\partial z\|_\infty \le 1/m_1 \le 1$. From now on M denotes a positive constant which may change from line to line. Let start by studying γ^L:

$$\gamma^L(t,l,y) = \gamma(t,\phi^{-1}(t,l),y)\int_0^1 \frac{\partial\phi}{\partial z}(t,\phi^{-1}(t,l) + \theta\gamma(t,\phi^{-1}(t,l),y)d\theta$$

It follows

$$\sup_{t,l}\left|\gamma^L(t,l,y)\right| \le M\tau(y) \qquad \text{and} \qquad \sup_{t,l}\left|e^{\gamma^L(t,l,y)} - 1\right| \le M\tau(y)$$

and by using (7.32) and $ii)$ we obtain

$$\begin{aligned}\left|\frac{\partial\gamma^L}{\partial l}(t,l,y)\right| =& \left|-1 + \frac{\frac{\partial\phi}{\partial z}(t,\phi^{-1}(t,l)+\gamma(t,\phi^{-1}(t,l),y))}{\frac{\partial\phi}{\partial z}(t,\phi^{-1}(t,l))}\left(1 + \frac{\partial\gamma}{\partial l}(t,\phi^{-1}(t,l),y)\right)\right| \\ \le& M\left(\sup_{t,l}|\gamma(t,\phi^{-1}(t,l),y)|\left\|\frac{\partial^2\phi}{\partial z^2}\right\|_\infty + \rho(y)\right) \le M\tau(y)\end{aligned}$$

In conclusion γ^L verifies Assumptions 5.1-[$\mathbf{C},\mathbf{I1},\mathbf{I2}$] with $\tau^L(y) := M\tau(y)$. Assumptions 7.1-[$\mathbf{L},\mathbf{I}$] have not been modified, whereas for the no degeneracy property we can use Remark 7.2: since $|\gamma^L| \ge |\gamma|/m_2$ then γ^L also verifies Assumption 7.1-[$\mathbf{ND}$].

For the drift function μ^L, it is straightforward to deduce

$$(t,l) \to \frac{\partial \phi}{\partial t}(t,\phi^{-1}(t,l)) + \mu(t,\phi^{-1}(t,l))\frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l))$$

is bounded and Lipschitz continuous in the variable l, whereas

$$\int_{|y|\le 1}\left(\tilde{\gamma}^L(t,l,y) - \gamma(t,\phi^{-1}(t,l),y)\frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l))\right)\nu(dy) =$$
$$\int_{-1}^{1}\gamma(t,\phi^{-1}(t,l),y)\int_0^1\left(\frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l)+\theta\gamma(t,\phi^{-1}(t,l),y) - \frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l))\right)d\theta\nu(dy)$$

from which we deduce that

$$(t,l) \to \int_{|y|\le 1}\left(\gamma^L(t,l,y) - \gamma(t,\phi^{-1}(t,l),y)\frac{\partial \phi}{\partial z}(t,\phi^{-1}(t,l))\right)\nu(dy)$$

is also bounded and Lipschitz continuous in the variable l: the function μ^L verifies Assumption 7.1-[**C**].

For the Assumptions 7.6 we have that, by construction, **H2** is verified, whereas Assumptions 7.6-[**H1**, **H3**] hold true by using $i), ii), iii)$, the bounds on $\partial\phi/\partial z$ and the properties of γ.

□

The above Lemma proves that it is possible to replace Assumption 7.6-[**H2**] by assuming more regularity of γ at $y = 0$:

H2$_{\text{bis}}$ The function $(t,z) \to \gamma_y(t,z,0)$ is differentiable and

$$\left|\frac{d}{dt}\gamma_y(t,z,0)\right| + \left|\frac{d}{dz}\gamma_y(t,z,0)\right| \le m_2 \text{ for all } t,z \in [0,T]\times\mathbb{R}$$

H2$_{\text{ter}}$ The function $z \to \frac{d}{dt}\gamma_y(t,z,0)$ is Lipschitz continuous:

$$\left|\frac{d}{dt}\gamma_y(t,z,0) - \frac{d}{dt}\gamma_y(t,z',0)\right| \le m_2|z-z'| \text{ for all } t \in [0,T],\ z,z' \in \mathbb{R}$$

7.6 Smoothness and characterization of the function v^f

We conclude with the study of Problem (7.2) when $f \neq 0$. According to Theorem 5.14, we need to prove that the following PIDEs

$$0 = -\frac{\partial b}{\partial t} - \mathcal{A}b - \mathcal{B}b - \pi^*[a]\mathcal{Q}b, \qquad b(T,.) = -2f \tag{7.35}$$

$$0 = -\frac{\partial c}{\partial t} - \mathcal{A}c - \mathcal{B}c + \frac{1}{4}\frac{(\mathcal{Q}b)^2}{\mathcal{G}a}, \qquad c(T,.) = f^2 \tag{7.36}$$

have a unique smooth solution, where π^* is the minimizer of the operator $\mathcal{H}$ in (7.9).

Theorem 7.17. *Let Assumptions 5.1–7.1 hold true together with Assumptions 7.6. Fix* $\delta \in (0, \alpha - 1)$ *and let* $f \in H_e^{\alpha+\delta}(\mathbb{R})$. *The function* v^f *in* (7.2) *admits the decomposition* (5.41)*:*

$$v^f(t,z,x) = a(t,z)x^2 + b(t,z)x + c(t,z), \quad v^f(T,z,x) = (f(z) - x)^2$$

where $a \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$ *is the unique solution of* (7.8)*, so it does not depend on* f*, and*

$$b, c \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$$

are the unique solutions of the linear PIDEs (7.35)*–*(7.36). *The functions* $t \to a(t,.), b(t,.), c(t,.)$ *are continuously differentiable in* $(0,T)$.

Proof.
By applying Theorem (7.14) we have that the function a in (7.5) is the unique solution of the PIDE in (7.8) and it belongs to the Hölder space $H^{\alpha+\delta}([0,T] \times \mathbb{R})$. Furthermore, from the regularity of the function a, it is not difficult to verify that $\pi \in H^1([0,T] \times \mathbb{R})$. We can proceed as in the proof of Theorem 6.10: transform the PIDE (7.35) into

$$-\frac{\partial b}{\partial t} - \mathcal{B}^{st} b + \eta b = \pi^* \mathcal{Q} b + (\mathcal{B} - \mathcal{B}^{st}) b + \mathcal{A} b, \ bi(T,.) = -2fe^{\eta T} \qquad (7.37)$$

where $\mathcal{B}^{st}$ is given in (7.16). Consider the map Ξ_η defined as follows: for any $\psi \in H^{\alpha+\delta}([0,T] \times \mathbb{R})$, $\Xi_\eta(\psi)$ verifies

$$-\frac{\partial}{\partial t}\Xi_\eta(\psi) + \mathcal{B}^{st}\Xi_\eta(\psi) + \eta\Xi_\eta(\psi) = \pi^* \mathcal{Q}\psi + (\mathcal{B} - \mathcal{B}^{st})\psi + \mathcal{A}\psi, \ \Xi_\eta(\psi)(T,.) = -2fe^{\eta T}$$

This map is well defined by applying Theorem 7.10. By using Lemma 7.4 and Corollary 7.9 it is possible to select η big enough such that Ξ_η is a contraction in the Hölder space $H^{\alpha+\delta}([0,T] \times \mathbb{R})$: its unique fixed point ψ^* is then the unique solution of (7.37), or, equivalently, $e^{-\eta t}\psi^*$ is the unique solution of (7.35). For the PIDE (7.36) we can proceed in the same way to deduce that it has a unique smooth solution in $H^{\alpha+\delta}([0,T] \times \mathbb{R})$.

□

The Remark (6.11) holds true in this case: in particular one can allow $f \in H_e^{\alpha+\delta'}(\mathbb{R})$ for some $\delta' \in (0,1)$, and then obtain

$$b, c \in H^{\alpha+\min(\delta,\delta')}([0,T] \times \mathbb{R}).$$

7.7 Finite activity processes: the model

We conclude the chapter by studying the quadratic hedge problem when the process Z is driven by a Poisson random measure whose intensity measure $\nu(dy)$ is of finite variation, which, in particular, covers the cases $\nu(\mathbb{R}) < \infty$ or $\nu(dy) = g(y)|y|^{-1+\alpha}$ for $\alpha \in (0,1)$. As we will see later, the main difficulty here is to prove that the value function v^f is smooth. In general this is not true, especially for a finite measure. However, if the drift satisfies $\mu = 0$ then the argument given in Sections 5.5 and 5.6 can be rigorously justified even for non a smooth value function v^f. When Z is a finite variation process without drift, the Itô's formula, which is the key tool in the proof of Theorems 5.11–5.14, applies whenever v^f is once differentiable in t and Lipschitz continuous in z[1] (see for example Proposition 8.12 in Cont and Tankov (2004)). For sake of clarity, we recall the model and the Assumptions we will need:

$$dZ_r^{t,z} := \mu\left(r, Z_r^{t,u,z}\right) dr + \int_{\mathbb{R}} \gamma\left(r, Z_{r-}^{t,z}, y\right) J\left(dydr\right), \; Z_t^{t,z} = z \tag{7.38}$$

for $t \in [0, T)$ and $z \in \mathbb{R}$ and we assume:

Assumption 7.18.

[I]- ***Integrability conditions****. The function τ defined in Assumptions 5.1 verifies, for some $y_0 \in (0,1)$ and some $m > 0$*

$$\sup_{0<|y|\le y_0} \frac{\tau(y)}{|y|} \le m \qquad \textit{and} \qquad \tau \in \mathbb{L}^1(\{|y| \le y_0\}, \nu(dy))$$

[ND]- ***No degeneracy****. The function Γ in (5.7) verifies*

$$|\Gamma| := \int_{\mathbb{R}} \Gamma(y)\nu(dy) > 0$$

Remark that there is no need to truncate the jump measure J as in (7.1) since, in this case, the function τ is integrable around zero. Our first objective is to prove that, under an appropriate change of variable, it is possible to remove the drift function in (7.38):

Lemma 7.19. *Let Assumptions 5.1–7.18 hold true and define the function*

$$\phi(t,z) = F_T^{t,z} \quad \textit{where } F_u^{t,z} \textit{ verifies} \quad F_u^{t,z} := z + \int_t^u \mu(s, F_s^{t,z})ds$$

Then $z \to \phi(t,z)$ is invertible for all t and there exist two positive constants m, M such that

$$0 < m \le \frac{\partial \phi}{\partial z} \le M \;\; \textit{for all } t, z \in [0,T] \times \mathbb{R}.$$

The process $L_t = \phi(t, Z_t)$ verifies

$$dL_s^{t,l} = \int \gamma^L(s, L_{s-}^{t,l}, y) J(dydt), \quad L_T = Z_T \tag{7.39}$$

where $\gamma^L(t,l,y) = \phi(t, \phi^{-1}(t,l) + \gamma(t, \phi^{-1}(t,l), y)) - l$ verifies Assumptions 7.18.

[1] In this case, Itô's formula is nothing but the usual Lebesgue-Stieltjes change of variables formula.

Proof.
Since the drift function $\mu(t,z)$ is Lipschitz continuous and differentiable in the variable z, the deterministic function $\digamma$ is well defined. Trivially

$$
\begin{aligned}
D_t \digamma_u^{t,z} &= -\mu(t,z) + \int_t^u \partial_z \mu(s, \digamma_s^{t,z}) D_t \digamma_s^{t,z} ds \\
D_z \digamma_u^{t,z} &= 1 + \int_t^u \partial_z \mu(s, \digamma_s^{t,z}) D_z \digamma_s^{t,z} ds
\end{aligned}
$$

so that

$$D_t \digamma_u^{t,z} + \mu(t,z) D_z \digamma_u^{t,z} = 0, \qquad \text{for all } t, z, u$$

Furthermore we can solve the linear ODE defining $D_z \digamma_u^{t,z}$ to obtain

$$D_z \digamma_u^{t,z} = \exp\left(\int_t^u \partial_z \mu(s, \digamma_s^{t,z}) ds \right) > 0$$

and therefore

$$0 < m = \exp\left(-T \left\|\partial_z \mu\right\|_\infty\right) \le \frac{\partial \phi}{\partial z} \le M := \exp\left(T \left\|\partial_z \mu\right\|_\infty\right)$$

In particular for fixed t the map $z \to \digamma_T^{t,z}$ is invertible. By applying Itô's formula to $L_t := \phi(t, Z_t)$ we obtain

$$
\begin{aligned}
dL_t &= \left(D_t \digamma_T^{t,Z_t} + \mu D_z \digamma_T^{t,Z_t} \right) dt + \int \left(\phi(t, Z_{t-} + \gamma) - \phi(t, Z_{t-}) \right) J(dydt) \\
&= \int_{\mathbb{R}} \gamma^L(t, L_{t-}, y) J(dydt)
\end{aligned}
$$

where $\gamma^L(t,l,y) := \phi(t, \phi^{-1}(t,l) + \gamma(t, \phi^{-1}(t,l), y)) - l$. The upper and lower bound on $\partial\phi/\partial z$ can be used to deduce that the new jump function γ^L verifies Assumptions 7.18:

□

If we use L as the state variable instead of Z then the control problem 7.2 becomes

$$\mathbf{QH}: \qquad \text{minimize } \mathbb{E}^{\mathbb{P}} \left[\left(f\left(L_T^{0,l}\right) - X_T^{0,l,x,\theta} \right)^2 \right] \text{ over } \theta \in \mathcal{X}(0,l,x)$$

where

$$
\begin{aligned}
X_s^{t,l,x,\theta} &:= x + \int_t^s \theta_{r-} dS_u^{t,l} \\
\frac{dS_u^{t,l}}{S_{u-}^{t,l}} &:= \tilde{\mu}(u, \phi^{-1}(u, L_u^{t,l}))du + \int \left(e^{\gamma(u, \phi^{-1}(u, L_u^{t,l}), y)} - 1 \right) J(dydu)
\end{aligned}
$$

and its dynamic version:

$$v^{f,L}(t,l,x) := \inf_{\theta \in \mathcal{X}(t,l)} \mathbb{E}^{\mathbb{P}}\left[\left(f\left(L_T^{t,l}\right) - X_T^{t,l,x,\theta}\right)^2\right] \tag{7.40}$$
$$v^{f,L}(T,z,x) := (f(l) - x)^2$$

It is clear that the new value function $v^{f,L}$ and v^f in (7.2) are related: $v^f(t,z,x) = v^{f,L}(t,\phi(t,z),x)$, and, as we have already seen many times, $v^{f,L}$ admits the decomposition

$$v^{f,L}(t,l,x) := x^2 a^L(t,l) + x b^L(t,l) + c^L(t,l) \tag{7.41}$$

7.7.1 The pure investment problem in the finite variation case

As in Section 5.5 we first start with the case $f = 0$. In this case $v^{0,L}(t,l,x) = x^2 a^L(t,z)$, where

$$a^L(t,l) := \inf_{\theta \in \mathcal{X}(t,l,1)} \mathbb{E}\left[\left(1 + \int_t^T \theta_{u-} dS_u^{t,l}\right)^2\right] \tag{7.42}$$

which is related to the function a defined in (7.5) by $a(t,z) = a^L(t,\phi(t,z))$. We already know that the function a is bounded from above and below: $e^{-CT} \le a(t,z) \le 1$ and it is Lipschitz continuous w.r.t. z if $T < T^*$ as stated in Theorem 5.4. These properties also hold true for a^L by using the fact that $\partial\phi/\partial z$ is uniformly bounded from above and below (Lemma 7.19). For sake of clarity we redefine the differential operators given in Definition 7.3 since, in the finite variation case, they can be simplified:

Definition 7.20. *For a function* $\varphi \in H^1([0,T] \times \mathbb{R})$ *let*

$$\begin{aligned}
\mathcal{B}^L\varphi(t,l) &:= \int_{\mathbb{R}} \big(\varphi(t,l+\gamma^L(t,l,y)) - \varphi(t,l)\big)\,\nu(dy) \\
\mathcal{Q}^L\varphi(t,l) &:= \mu(t,\phi^{-1}(t,l))\varphi + \int_{\mathbb{R}} (e^{\gamma(t,\phi^{-1}(t,l),y)} - 1)\varphi(t,l+\gamma^L(t,l,y))\nu(dy) \\
\mathcal{G}^L\varphi(t,l) &:= \int_{\mathbb{R}} (e^{\gamma(t,\phi^{-1}(t,l),y)} - 1)^2\varphi(t,l+\gamma^L(t,l,y))\nu(dy) \\
\mathcal{H}^L[\varphi](t,l) &:= \inf_{|\pi| \le \bar{\Pi}^L} \left\{2\pi\mathcal{Q}_t^L\varphi(t,l) + \pi^2\mathcal{G}_t^L\varphi(t,l)\right\}
\end{aligned}$$

where

$$\bar{\Pi}^L := \frac{e^{CT}}{|\Gamma|} C_e\left(1 + mK_{lip}^a\right) \tag{7.43}$$

is obtained from (5.38) *and the fact that* $|\partial_l a^L| = |\partial_z a|/|\partial_l \phi^{-1}| \le mK_{lip}^a$.

The equivalent of Theorem 5.11, when the underlying process, is given by L is the following:

Proposition 7.21. *Let Assumptions 5.1–7.18 hold true and let $T < T^*$ as stated in Theorem* (5.4). *Assume that the PIDE*

$$0 = -\frac{\partial \varphi}{\partial t} - \mathcal{B}_t^L \varphi - \mathcal{H}_t^L[\varphi], \qquad \varphi(T, l) = 1 \tag{7.44}$$

has a unique solution $\varphi \in H^1([0,T] \times \mathbb{R})$, which also is strictly positive and for all $l \in \mathbb{R}$ the map $t \to \varphi(t,l)$ is continuously differentiable. Then $\varphi = a^L$ and the optimal strategy in problem (7.42) *is given by*

$$\theta_t^* = e^{-\phi^{-1}(t, L_{t-})} \pi^* \left(t, \phi^{-1}(t, L_{t-})\right) X_{t-}^{\theta^*}, \quad X_t^{\theta^*} := x + \int_0^t \theta_{r-}^* dS_r \tag{7.45}$$

where

$$\pi^*(t,l) := -\frac{\mathcal{Q}_t^L a^L(t,l)}{\mathcal{G}_t^L a^L(t,l)} \tag{7.46}$$

is the minimizer in the operator $\mathcal{H}^L$.

Proof.
The proof follows the ideas of Theorem 5.11. Remark, however that, as we already pointed out at the beginning, we only need that a belongs to $H^1([0,T] \times \mathbb{R})$ and is differentiable in t to apply Itô's formula.

□

The above Proposition gives us a way to characterize the function $v^{0,L}$ in (7.40) when $f = 0$, by proving that PIDE (7.44) has a unique solution which also has to be strictly positive. It also gives us the optimal policy for the pure investment problem. Our goal now is to prove that PIDE (7.44) has a unique solution in its appropriate Hölder space.

Theorem 7.22. *Let Assumptions 5.1–7.18 hold true. The PIDE* (7.44) *has a unique solution in the Hölder space $H^1([0,T] \times \mathbb{R})$, it is strictly positive and continuously differentiable w.r.t. t. Furthermore, the sequence φ^n defined by*

$$\varphi^{n+1}(t,l) := \mathbb{E}\left[\int_t^T \mathcal{H}_t^L[\varphi^n](s, L_s^{t,l}) ds\right] + 1, \, \varphi^0 \in H^1([0,T] \times \mathbb{R}) \tag{7.47}$$

verifies

$$\left\|\varphi^n - a^L\right\|_\infty \le M\beta^n, \quad n \to \infty$$

for some $M > 0$ and $\beta \in (0,1)$

Proof.
Transform the PIDE (7.44) into

$$0 = -\frac{\partial a}{\partial t} - \mathcal{B}_t^L a - \mathcal{H}_t^L[a] + \eta a, \quad a(T,l) = e^{\eta T} \tag{7.48}$$

for $\eta > 0$ and define the sequence

$$\varphi^{n+1}(t,l) := \mathbb{E}\left[\int_t^T e^{-\eta(s-t)}\mathcal{H}_t^L[\varphi^n](s,L_s^{t,l})ds\right] + e^{\eta t} \tag{7.49}$$

for some $\varphi^0 \in H^1([0,T]\times\mathbb{R})$. Remark that, by recurrence, one can prove that the sequence is well defined in $H^1([0,T]\times\mathbb{R})$. It is easy to prove that this sequence converges, at least in $\mathbb{L}^\infty(\mathbb{R})$: from the structure of $\mathcal{H}$ in Definition 7.20, we get

$$\begin{aligned}\left\|\varphi^{n+1}-\varphi^n\right\|_\infty \leq& \left\|\mathcal{H}^L[\varphi^n]-\mathcal{H}^L[\varphi^{n-1}]\right\|_\infty \int_t^T e^{-\eta(s-t)}ds\\ \leq& M\eta^{-1}\left\|\varphi^n-\varphi^{n-1}\right\|_\infty \leq \cdots \leq \left(\frac{M}{\eta}\right)^n \left\|\varphi^1-\varphi^0\right\|_\infty\\ \leq& M(\eta)\beta^n \end{aligned} \tag{7.50}$$

for some $M(\eta) > 0$ depending on η and some $\beta \in (0,1)$, by taking η big enough. Remark that the estimation concerning the operator $\mathcal{H}$ is different from the one we gave in Lemma 7.5 since the process L is of finite variation, but it can be obtained with the same type of computations. This proves that the sequence converges in $\mathbb{L}^\infty(\mathbb{R})$. Let us call φ^* this limit function. As in Corollary 7.13, this also proves that the PIDE (7.44) has at most one solution in $H^1([0,T]\times\mathbb{R})$ which also is differentiable w.r.t. t. Furthermore, from (7.49) we have

$$\begin{aligned}\left\|\varphi^{n+1}\right\|_{1,H} \leq& 1 + M\left\|\mathcal{H}^L[\varphi^n]\right\|_{1,H}\int_t^T e^{-\eta(s-t)}ds\\ \leq& 1 + M\eta^{-1}\left\|\varphi^n\right\|_{1,H} \leq \cdots \leq \left(\frac{M}{\eta}\right)^{n+1}\left\|\varphi^0\right\|_{1,H} + \sum_{i=0}^n\left(\frac{M}{\eta}\right)^i\end{aligned}$$

which implies that $\sup_n \|\varphi^n\|_{1,H} \leq c(\eta)$ for some constant $c(\eta)$, provided that η is big enough. As usual (see Remark 6.6) we can say that this sequence is bounded in $H^1([0,T]\times\mathbb{R})$ for any $\eta \geq 0$. In particular, for $\eta = 0$ and $0 \leq t' \leq t \leq T$ we have

$$\begin{aligned}\left|\varphi^{n+1}(t,l)-\varphi^{n+1}(t',l)\right| \leq& \mathbb{E}\left[\int_{t'}^t \left|\mathcal{H}[\varphi^n](s,L_s^{t,l})\right| ds\right]\\ &+\mathbb{E}\left[\int_t^T \left|\mathcal{H}[\varphi^n](s,L_s^{t,l})-\mathcal{H}[\varphi^n](s,L_s^{t',l})\right| ds\right]\\ \leq& \left\|\mathcal{H}[\varphi^n]\right\|_\infty |t-t'| + T\left\|\mathcal{H}[\varphi^n]\right\|_{1,H}|t-t'|^{1/2}\\ \leq& M|t-t'|^{1/2}\left\|\varphi^n\right\|_{1,H} \leq M|t-t'|^{1/2}\end{aligned}$$

i.e. the functions φ^n are all uniformly 1/2-Hölder w.r.t. t: the limit function φ^* inherits of this property.
By taking the limit $n \to \infty$ in (7.49) with $\eta = 0$, we have

$$\varphi^*(t,l) := 1 + \mathbb{E}\left[\int_t^T \mathcal{H}_t^L[\varphi^*](s,L_s^{t,l})ds\right] \tag{7.51}$$

Exactly as in the proof of Theorem 7.14 we can prove that φ^* is the unique viscosity solution of (7.44) and, if we define

$$w(t,l,x) := \mathbb{E}\left[\left(\hat{X}_T^{t,l,x}\right)^2\right]$$

where $d\hat{X}_s^{t,l,x} := \hat{\pi}_{s-}^L \hat{X}_{s-}^{t,l,x} \exp\left(-\phi^{-1}(s, L_{s-}^{t,l})\right) de^{\phi^{-1}(s,L_s^{t,l})}$ and

$$\hat{\pi}^L := -\bar{\Pi}^L \vee -\frac{\mathcal{Q}^L \varphi^*}{\mathcal{G}^L \varphi^*} \wedge \bar{\Pi}^L$$

then we can prove that $w(t,l,x) = x^2 \varphi^*(t,l,x)$, from which we deduce that $e^{-CT} \leq a^L(t,l) \leq \varphi^*(t,l)$. It follows then that $e^{-CT} \leq \varphi^n(t,l)$ for all $n \geq \bar{n}$ and $(t,l) \in [0,T] \times \mathbb{R}$, $\bar{n} \in \mathbb{N}$.

Our aim now is to prove that φ^* belongs to $H^1([0,T]\times\mathbb{R})$ and is differentiable w.r.t. t. For this we could repeat the argument of the proof of Theorem 7.14, but in this case, due to the simple form of the operator $\mathcal{H}^L$, we are able to prove it directly. For this we have the following technical result

Lemma 7.23. *Let Assumptions 5.1–7.18 hold true. Then*

$$\mathcal{H}^L : H^1([0,T]\times\mathbb{R}) \to H([0,T]\times\mathbb{R}^1)$$

and there exists some positive constant $M > 0$ such that

$$\left\|\mathcal{H}^L[\varphi+\psi] - \mathcal{H}^L[\varphi]\right\|_{1,H} \leq M\left(\|\psi\|_{1,H} + \|\psi\|_\infty \|\varphi\|_{1,H}\right)$$

for any $\psi, \varphi+\psi \geq m \geq 0$. The constant M does not depend on φ or ψ but on m.

We postpone the proof to paragraph 7.7.2.

From (7.49) and the previous Lemma we obtain

$$\begin{aligned}\langle \varphi^{n+1} - \varphi^n \rangle_{l,Q_T}^{(1)} &\leq M_1 \langle \mathcal{H}^L[\varphi^n] - \mathcal{H}^L[\varphi^{n-1}]\rangle_{l,Q_T}^{(1)} \int_t^T e^{-\eta(s-t)} ds \\ &\leq M_1 \eta^{-1}\left(\left\|\varphi^n - \varphi^{n-1}\right\|_{1,H} + \left\|\varphi^n - \varphi^{n-1}\right\|_\infty \min\left(\|\varphi^n\|_{1,H}, \left\|\varphi^{n-1}\right\|_{1,H}\right)\right)\end{aligned}$$

by using the fact that $\mathbb{E}\left[|L_s^{t,l_1} - L_s^{t,l_2}|\right] \leq M|l_1 - l_2|$. From to the fact that φ^n is uniformly bounded in $H^1([0,T]\times\mathbb{R})$ and (7.50) we get

$$\langle \varphi^{n+1} - \varphi^n \rangle_{l,Q_T}^{(1)} \leq M\eta^{-1}\left(\left\|\varphi^n - \varphi^{n-1}\right\|_{1,H} + c(\eta)M(\eta)\beta^{n-1}\right)$$

Together with (7.50) we deduce

$$\begin{aligned}\left\|\varphi^{n+1} - \varphi^n\right\|_{1,H} \leq & M\eta^{-1}\left(\left\|\varphi^n - \varphi^{n-1}\right\|_{1,H} + c(\eta)M(\eta)\beta^{n-1}\right) + M(\eta)\beta^n \\ \cdots \leq & \left(\frac{M}{\eta}\right)^n \left\|\varphi^1 - \varphi^0\right\|_{1,H} + \beta^n c(\eta)(M(\eta)+1)\sum_{i=0}^n \left(\frac{M}{\beta\eta}\right)^i \\ \leq & \tilde{M}(\eta)\left(\left(\frac{M}{\eta}\right)^n + \beta^n\right) \leq \tilde{M}(\eta)\tilde{\beta}^n\end{aligned}$$

for some positive constant $\tilde{M}(\eta)$ and $\tilde{\beta} \in (0,1)$ if η is big enough. We finally deduce that $\varphi^n \to \varphi^*$ in $H^1([0,T]\times\mathbb{R})$, since it is a Cauchy sequence. To prove that φ^* is differentiable w.r.t t one can use the Markov property of L and (7.51) to obtain

$$\varphi^*(t,l) = \mathbb{E}\left[\int_t^{t+h} \mathcal{H}^L[\varphi^*](s, L_s^{t,l})ds + \varphi^*(t+h, L_{t+h}^{t,l})\right]$$

Apply then Itô's formula to obtain

$$\begin{aligned}\frac{\varphi^*(t,l)-\varphi^*(t+h,l)}{h} =& \frac{1}{h}\mathbb{E}\left[\int_t^{t+h} \mathcal{H}^L[\varphi^*](s, L_s^{t,l})ds + \varphi^*(t+h, L_{t+h}^{t,l} - \varphi^*(t+h,l)\right]\\ =& \frac{1}{h}\mathbb{E}\left[\int_t^{t+h} \mathcal{H}^L[\varphi^*]ds + \int_t^{t+h} \mathcal{B}^L\varphi^*(t+h, L_s^{t,l})ds\right]\end{aligned}$$

Since the process L_s has right continuous paths and φ^* is 1/2-Hölder continuous w.r.t. t, we can take the limit $h \to 0$ and, by dominated convergence, we deduce that the function φ^* is continuously differentiable w.r.t. t. Furthermore, it verifies the PIDE (7.44).
To summarize, we have found a function $\varphi^* \in H^1([0,T]\times\mathbb{R})$, which also is continuously differentiable w.r.t. t and it is bounded from below by e^{-CT}, and satisfies PIDE (7.48). Furthermore this φ^* is the unique solution of this PIDE, as pointed out when we proved that $\varphi^n \to \varphi$ in $\mathbb{L}^\infty(\mathbb{R})$. We conclude by applying Proposition 7.21.

□

Remark 7.24. *Remark that the above theorem does not tell us anything new on the regularity of a^L w.r.t. l that we already did not know. In fact the theorem proves that the equation* (7.44) *has a unique solution which belongs to $H^1([0,T]\times\mathbb{R})$, i.e. it is bounded and Lipschitz continuous w.r.t. l, and it is also differentiable w.r.t. t. From Proposition 7.21 this unique solution has to be a^L, and we already knew that it is bounded and Lipschitz continuous. Nevertheless the theorem proves that equation* (7.44) *has a unique solution which also is strictly positive, which is necessary to apply Proposition 7.21. We cannot have any further regularity on a^L since there is no "regularizing" operator in* (7.44)*. However, the regularity that we obtain for a^L is enough to apply Itô's formula in the finite variation pure jump case.*

The above Theorem and Proposition 7.21 allow us to characterize the value function $v^{0,L}$ and the optimal strategy. Remark that we only have Lipschitz regularity for the function a^L, and then Lipschitz regularity for the function a. If one does not change the variable, then the presence of a drift term in the dynamic of Z will demand at least C^1 regularity for a, which is far from being always true.

7.7.2 Proof of Lemma 7.23

Proof.
Let us define $H(q,g) := \inf_{|\pi|\le\bar{\Pi}}\{2\pi q + \pi^2 g\}$ so then $\mathcal{H}[\varphi] = H(\mathcal{Q}^L\varphi, \mathcal{G}^L\varphi)$. It follows then

$$\mathcal{H}^L[\psi] \le \mathcal{H}^L[\varphi+\psi] - \mathcal{H}^L[\varphi] \le \sup_{|\pi|\le\bar{\Pi}}\left\{2\pi\mathcal{Q}^L\psi + \pi^2\mathcal{G}^L\psi\right\}$$

so then

$$\left\|\mathcal{H}^L[\varphi+\psi]-\mathcal{H}^L[\varphi]\right\|_\infty \leq M\left(\left\|\mathcal{Q}^L\psi\right\|_\infty+\left\|\mathcal{G}^L\psi\right\|_\infty\right)\leq M\left\|\psi\right\|_{1,H} \tag{7.52}$$

form Definition 7.20. We now need to estimate $\langle\mathcal{H}^L[\varphi+\psi]-\mathcal{H}^L[\varphi]\rangle^{(1)}_{l,Q_T}$, i.e.

$$\left|\mathcal{H}[\varphi+\psi](t,l)-\mathcal{H}[\psi](t,l)-\mathcal{H}[\varphi+\psi](t,l')+\mathcal{H}[\psi](t,l')\right|$$

Remark first that the function H is Lipschitz continuous, twice differentiable with bounded second derivative when g is bounded from below by some strictly positive $\bar{g}$:

$$\sup_{q\in\mathbb{R},\,0<\bar{g}\leq g}\left|D^2H(q,g)\right|\leq M$$

for some positive M. If now $g, g+\eta, g', g'+\eta'\geq\bar{g}$ then

$$\begin{aligned}
&H(q+h,g+\eta)-H(q,g)-H(q'+h',g'+\eta')+H(q',g')=\\
&H(q+h,g+\eta)-H(q+h',g+\eta)+H(q+h',g+\eta)-H(q+h',g+\eta')\\
+&H(q+h',g+\eta')-H(q,g)-H(q'+h',g'+\eta')+H(q',g')
\end{aligned}$$

In particular

$$\begin{aligned}
&|H(q+h,g+\eta)-H(q+h',g+\eta)+H(q+h',g+\eta)-H(q+h',g+\eta')|\\
\leq&M(|h-h'|+|\eta-\eta'|)
\end{aligned}$$

and

$$\begin{aligned}
&|H(q+h',g+\eta')-H(q,g)-H(q'+h',g'+\eta')+H(q',g')|\\
\leq&M(|h'|+|\eta'|)(|q-q'|+|g-g'|)
\end{aligned}$$

We now use the above estimation with $q=\mathcal{Q}^L\varphi(t,l)$, $h=\mathcal{Q}^L\psi(t,l)$, $g=\mathcal{G}^L\varphi(t,l)$ and $\eta=\mathcal{G}^L\psi(t,l)$ and the same for (t,l'). Since both $\mathcal{G}^L\varphi$ and $\mathcal{G}^L(\varphi+\psi)$ are bounded from below by $m|\Gamma|$, as stated in Assumptions 7.18-[**ND**], we deduce

$$\begin{aligned}
&\left|\mathcal{H}[\varphi+\psi](t,l)-\mathcal{H}[\psi](t,l)-\mathcal{H}[\varphi+\psi](t,l')+\mathcal{H}[\psi](t,l')\right|\leq\\
&M\left(\langle\mathcal{Q}^L\psi\rangle^{(1)}_{Q_T}l+\langle\mathcal{G}^L\psi\rangle^{(1)}_{l,Q_T}+\left(\left\|\mathcal{Q}^L\psi\right\|_\infty+\left\|\mathcal{G}^L\psi\right\|_\infty\right)\left(\langle\mathcal{Q}^L\varphi\rangle^{(1)}_{l,Q_T}+\langle\mathcal{G}^L\varphi\rangle^{(1)}_{l,Q_T}\right)\right)\\
\leq&M\left(\|\psi\|_{1,H}+\|\psi\|_\infty\|\varphi\|_{1,H}\right)|l-l'|
\end{aligned}$$

By taking the supremum over t we obtain

$$\langle\mathcal{H}^L[\varphi+\psi]-\mathcal{H}^L[\varphi]\rangle^{(1)}_{l,Q_T}\leq M\left(\|\psi\|_{1,H}+\|\psi\|_\infty\|\varphi\|_{1,H}\right)$$

The above estimations and (7.52) allow us to conclude our proof.

□

7.7.3 The quadratic hedge problem in the finite variation case

The general case $f \neq 0$ in problem (7.40) is very easy:

Theorem 7.25. *Let Assumptions 5.1–7.18 hold true and $T < T^*$ as in Theorem* (5.4)*. Fix then $f \in H^1_e(\mathbb{R})$. The function $v^{f,L}$ defined in* (7.40) *admits the representation $v^{f,L}(t,l,x) = x^2 a^L(t,l) + x b^L(t,l) + c^L(t,l)$ where a^L is the unique solution of PIDE* (7.44) *in the Hölder space $H^1([0,T] \times \mathbb{R})$ and it is continuously differentiable w.r.t. t, whereas $b, c \in H^1([0,T] \times \mathbb{R})$ are the unique solutions of the following linear parabolic PIDEs*

$$0 = -\frac{\partial b^L}{\partial t} - \mathcal{B}^L_t b^L - \pi^* \mathcal{Q}^L_t b^L, \qquad b^L(T,.) = -2f; \tag{7.53}$$

$$0 = -\frac{\partial c^L}{\partial t} - \mathcal{B}^L_t c^L + \frac{1}{4}\frac{(\mathcal{Q}^L_t b^L)^2}{\mathcal{G}^L_t a^L}, \qquad c^L(T,.) = f^2 \tag{7.54}$$

where π^* is defined in (7.46) *and the functions $t \to b(t,.), c(t,.)$ also are continuously differentiable in $(0,T)$. The optimal strategy in the control problem* (7.40) *is given by*

$$\theta^*(t,l,x) := e^{-\phi^{-1}(t,l)}\left(\pi^*(t,l)x - \frac{1}{2}\frac{\mathcal{Q}^L b(t,l)}{\mathcal{G} a^L(t,l)}\right) \tag{7.55}$$

Proof.
We know that $v^{f,L}(t,l,x) = x^2 a^L(t,l) + x b^L(t,l) + c^L(t,l)$. We first prove that the PIDEs (7.53)–(7.54) have a unique solution in $H^1([0,T] \times \mathbb{R})$ and then we conclude with a verification argument.

We know that a^L defined in (7.42) is the unique solution of semi linear PIDE (7.44), it belongs to the Hölder space $H^1([0,T] \times \mathbb{R})$ and it is differentiable in time (Theorem 7.22). Furthermore $\|\pi^*\|_\infty \leq \bar{\Pi}^L$. But since a^L is Lipschitz continuous in the variable l and bounded from above and below, we also have

$$\begin{aligned}
\left|\pi^*(t,l) - \pi^*(t,l')\right| \leq & \frac{\left|\mathcal{Q}^L a^L(t,l) - \mathcal{Q}^L a^L(t,l')\right|}{|\mathcal{G}^L a^L(t,l)|} \\
& + \left|\frac{\mathcal{Q}^L a^L(t,l')}{\mathcal{G}^L a^L(t,l')}\right| \frac{1}{|\mathcal{G}^L a^L(t,l)|} \left|\mathcal{G}^L a^L(t,l) - \mathcal{G}^L a^L(t,l')\right| \\
\leq & M|l - l'|
\end{aligned}$$

which implies that $\pi^* \in H^1([0,T] \times \mathbb{R})$. For $\eta > 0$ consider the map Ξ_η in $H^1([0,T] \times \mathbb{R})$ as follows:

$$\Xi_\eta(\psi)(t,l) := \mathbb{E}\left[\int_t^T e^{-\eta(s-t)} \left(\pi^* \mathcal{Q}\psi\right)(s, L^{t,l}_s) ds - 2e^{\eta t} f(L^{t,l}_T)\right] \tag{7.56}$$

The Lipschitz condition on π^*, μ and f prove that $\Xi_\eta(\psi) \in H^1([0,T] \times \mathbb{R})$ and that

$$\|\Xi_\eta(\psi_1) - \Xi_\eta(\psi_2)\|_{1,H} \leq \frac{M}{\eta} \|\psi_1 - \psi_2\|_{1,H}$$

for some positive M. It follows that, for η big enough, Ξ_η is a contraction in $H^1([0,T]\times\mathbb{R})$. If we call ψ^* its unique fixed point then it verifies

$$\psi^*(t,l) = \mathbb{E}\left[\int_t^T e^{-\eta(s-t)}\left(\pi^*\mathcal{Q}\psi^*\right)(s,L_s^{t,l})ds - 2e^{\eta t}f(L_T^{t,l})\right]$$

We can apply the Markov property of L to get

$$\psi^*(t,l) = \mathbb{E}\left[\int_t^{t+h} e^{-\eta(s-t)}\left(\pi^*\mathcal{Q}\psi^*\right)(s,L_s^{t,l})ds + \psi^*(L_{t+h}^{t,l})\right]$$

from which we deduce that ψ^* is differentiable w.r.t. t (as we did in Theorem 7.22). Then it is the unique solution of

$$0 = -\frac{\psi^*}{\partial t} - \mathcal{B}^L\psi^* - \pi^*\mathcal{Q}^L\psi^* + \eta\psi^*$$

or, equivalently, $b^L(t,l) := e^{-\eta t}\psi^*(t,l)$ is the unique solution of (7.53) in the Hölder space $H^1([0,T]\times\mathbb{R})$ and is also continuously differentiable w.r.t. t. The same method can be used to prove that PIDE (7.54) has a unique solution. A verification argument (as in the proof of Theorem 5.14) can be used here to conclude that the value function in (7.40) is given by $v^{f,L}(t,l,x) = x^2a^L(t,l) + xb^L(t,l) + c^L(t,l)$, and the optimal strategy of the problem is given by (7.55).

□

Chapter 8

Quadratic hedge in electricity markets

In this Chapter we apply the results obtained in Chapter 7 to the electricity market, which inspired us to consider pure jump models for the quadratic hedge problem. In Section 8.1 we briefly describe some features of these electricity markets. We then introduce the future *contracts, which play the role of hedging instrument, and we model them in order to satisfy Assumptions 5.1–7.1 and Assumptions 7.6 (Section 8.2). We derive HJB equations as stated in Theorem 5.14 and propose a numerical scheme to solve these PIDEs (Section 8.3). We finally test these schemes for the NIG process which is a degenerate model ($\alpha = 1$).*

Contents

8.1 Electricity market: a short survey 155
8.2 Future contracts 158
8.3 Numerical approximation of the functions a and b 162
8.3.1 Truncation for the function a 163
8.3.2 Numerical algorithm for the function a 168
8.3.3 Resolution methodology to calculate the function b 169
8.4 A numerical example 171
8.4.1 The martingale case 173

8.1 Electricity market: a short survey

In the last two decades energy markets have been liberalized by many governments, with the idea that a more open market should lead to a better distribution of supply and demand, stabilization of prices and more competition between the actors.[1]. For the electricity, the first example was given by Chile (early 1980s), followed by Argentina and some other countries in South America. This liberalization is rather advanced and, beyond standard operations (buy and sell electricity, secure

[1]In France, see for example the really recent "loi NOME",n.2010-1488 du 7 décembre 2010 portant nouvelle organisation du marché de l'électricité, http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000023174854&categorieLien=id

the supply), it is possible to make more sophisticated financial operations (buying insurance, short-term trading, exchange of financial derivatives).

However, due to the nature of the electricity, this particular market presents some problematic aspects. First of all, electricity is not exactly as any other commodity, since, it is not possible, at present, to store the excess production (although many studies have been done in this direction). This means that a "unit" of electricity has to be used when it is bought. With a macroeconomic language we can say that, structurally, the supply follows the demand, which generally is "instantaneous" whereas the supply is not. This mismatch implies the presence of spikes in the electricity price, as showed for example in figures 8.1–8.2: upward movements are followed by quick return to initial level. Consequently, a non-Gaussian behavior is observed in empirical estimations of electricity price time series, as pointed out in Geman and Roncoroni (2006) or Meyer-Brandis and Tankov (2008). This first

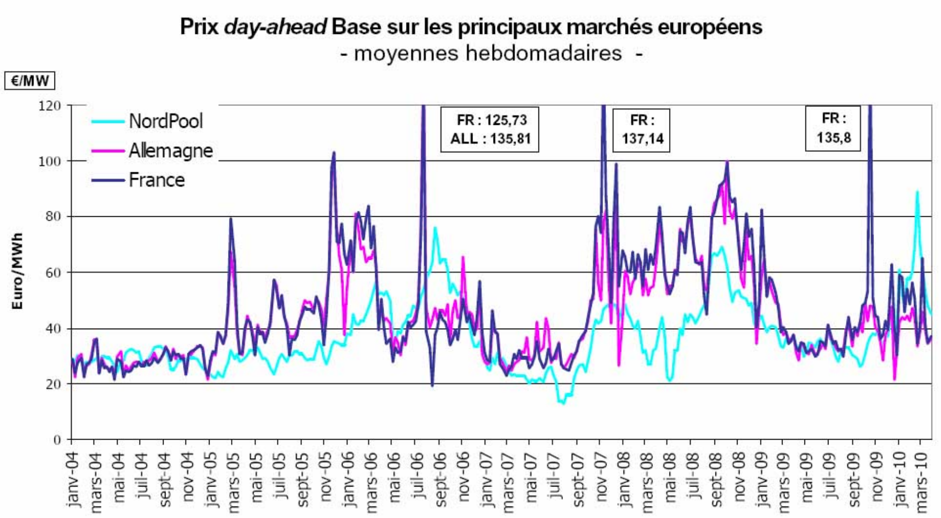

Figure 8.1: The spikes impact on the weekly averages of electricity prices. Comparison between France, Germany and the Nord Pool (Norway, Denmark, Sweden, Finland, and Estonia). Source: Commission de Regulation de l'énergie www.cre.fr

remark suggests that a reasonable model for the electricity price should present jumps in the path, together with a mean-reverting behavior. In practice, upward and downward movements of electricity price are essentially due to jumps. Seasonality also should be taken into account when modeling the electricity price. Another important feature of electricity market is the procedure of price formation: in contrast with liquid and deep markets (the euro-dollar exchange for example), where the price formation essentially reflects the supply and demand, the price formation procedure for the electricity is more complex. We do not enter into the details, but to understand this procedure one should consider many factors: the structurally biased supply and demand in the electricity context; the role of former monopolistic public companies (which, at the same time, are producers, retail actors and exercise their natural monopolistic function of distribution); the role of governments, which care of the prices of socially sensible goods. These consideration suggests that, when modeling, one should consider the electricity market as an illiquid and

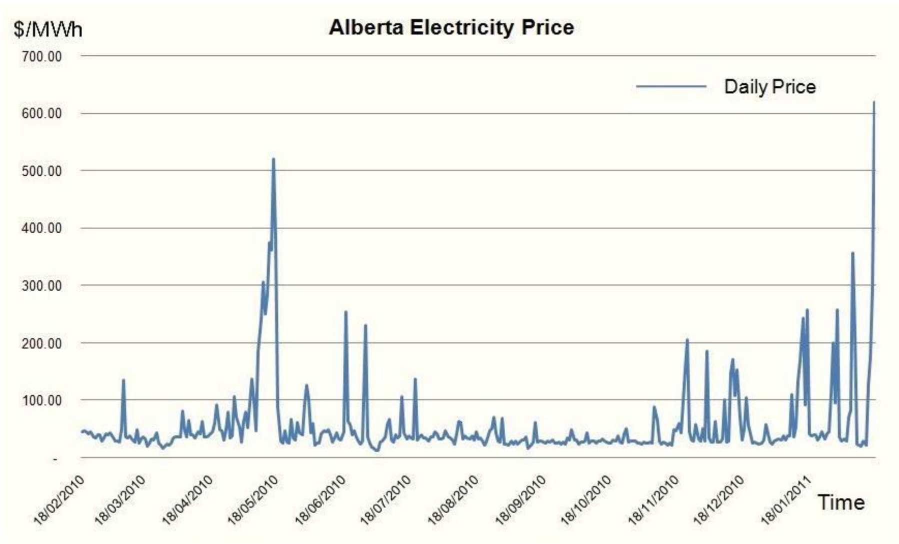

Figure 8.2: The observed price of electricity in the Ontario's market. Source: *http : //www.energyadvantage.com*

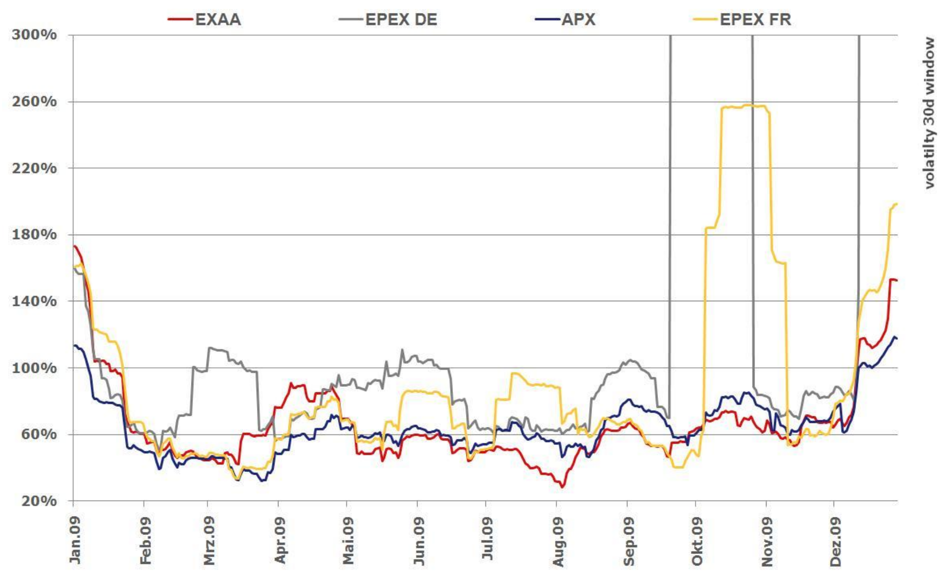

Figure 8.3: The realized volatility of the electricity price in different electricity markets: Energy Exchange Austria (EXAA), Short-term Trading Deutschland (EPEX DE), Amsterdam Power Exchange (APX), Short-term Trading France (EPEX FR). Source EXAA Abwicklungsstelle für Energieprodukte AG.

an incomplete one (which is partially due to the mismatch between the production and the consumption) and should not consider the electricity spot price process as a hedging instrument (since, as we said above, the electricity cannot be stored and therefore, in some sense, is not tradeable).

8.2 Future contracts

In this section we will describe the future contract, a popular hedging instrument which is traded in many electricity markets. For further details we refer to Clewlow and Strickland (2000). Buying a future contract with maturity T and duration d essentially means that at the maturity T one will be delivered a certain quantity of electricity up to time $T + d$. We denote the price of this contract at time t with $F_{d,T,t}$. To model this financial instrument we introduce L as follows

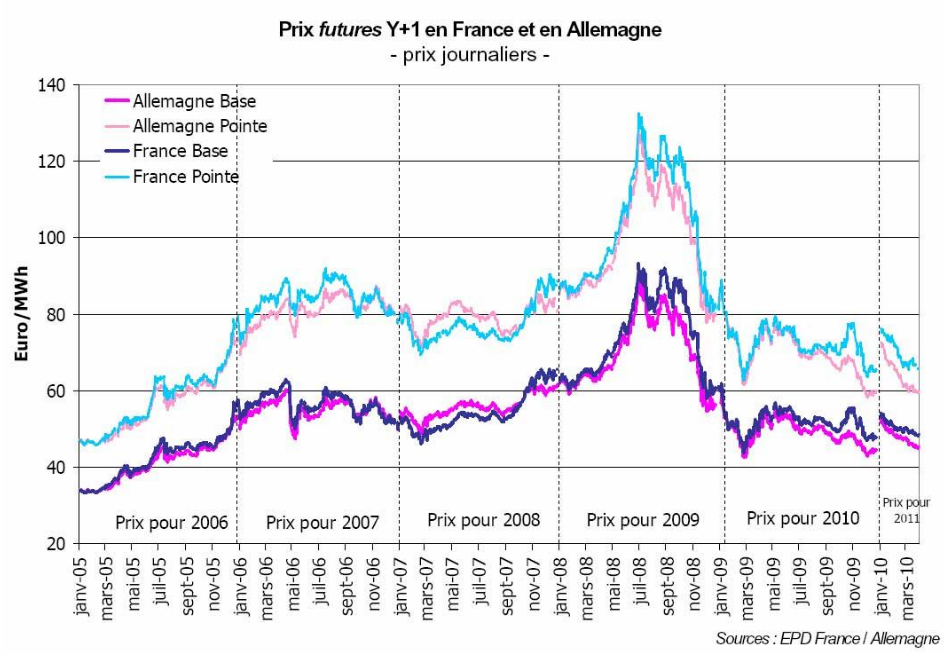

Figure 8.4: Price of futures with maturity 1 year, comparison France and Germany. Source EPD France/Germany.

$$L_s = \zeta s + \int_0^s \int_{\mathbb{R}} y\bar{J}(dyds) \tag{8.1}$$

where $\zeta \in \mathbb{R}$ and J is a Poisson random measure, whose intensity measure is denoted by $\nu(dy)$. Fix $c \in \mathbb{R}^+$, $l(s) = e^{-cs}$ and

$$A_t := \int_0^t e^{cs} dL_s \tag{8.2}$$

If $T \to \psi(0, T)$ denotes the price at time 0 of a future contract with maturity T and instantaneous delivery (which is supposed known), then we will model the price at time t of the same future contract as a random perturbation of the forward curve ψ: with the previous notations we have

$$F_{0,T,t} = \psi(0,T)e^{l(T)A_t}$$

We can imagine that the price at time t of a future contract with duration d is the average on the time period $[T, T+d]$ of the future contract prices with instantaneous delivery. It is then reasonable to model the price at time t of a future contract with delivery time T and duration $d > 0$ by

$$F_{d,T,t} = \frac{1}{d}\int_T^{T+d} F_{0,s,t}ds = \frac{1}{d}\int_T^{T+d} \psi(0,s)e^{l(s)A_t}ds$$

For some particular reason that would be clear in the sequel, we prefer the following notation:

$$F_{d,T,t} := \exp(\Phi(A_t)) \quad \text{where} \quad \Phi(A) := \log\left(\frac{1}{d}\int_T^{T+d} \psi(0,s)e^{l(s)A}ds\right) \tag{8.3}$$

The quadratic hedge problem, in this context, becomes

$$\textbf{minimize} \qquad \mathbb{E}\left[\left(\tilde{f}(F_{d,T,t}) - x - \int_t^T \theta_{u-}dF_{d,T,u}\right)^2\right] \text{over } \theta \text{ and } x \in \mathbb{R} \tag{8.4}$$

for a given map $\tilde{f}$. The process $F_{d,T,t}$ corresponds to S in the formulation (5.2).

Lemma 8.1. *The process* $Z_t := \log(F_{d,T,t})$ *verifies*

$$dZ_t = \mu(t, Z_t)dt + \int \gamma(t, Z_{t-}, y)\bar{J}(dydt)$$

where

$$\begin{aligned}\gamma(t,z,y) :=& \Phi(\Phi^{-1}(z) + ye^{ct}) - z\\ \mu(t,z) :=& \zeta e^{ct}\Phi'(\Phi^{-1}(z)) + \int_{|y|\le 1}\left(\gamma(t,z,y) - ye^{ct}\Phi'(\Phi^{-1}(z))\right)\nu(dy)\end{aligned}$$

Assume that the Lévy measure $\nu(dy)$ *is given by* $\nu(dy) = g(y)|y|^{-(1+\alpha)}$*, for some* $\alpha \in (1,2)$ *and a bounded, positive and measurable* g *such that the following condition hold true:*

i). there exists some positive $m \ge 0$ *such that for all* $y, y' \in (-y_0, 0) \cup (0, y_0)$ *with* $yy' > 0$*,* $|g(y) - g(y')| \le m|y - y'|$

ii). $\lim_{y\to 0^-} g(y) = g(0^-)$ *and* $\lim_{y\to 0^+} g(y) = g(0^+)$ *with* $g(0^+), g(0^-) > 0$

iii). $\int_{y\le -1} y^4\nu(dy) + \int_{1<y} e^{4y}\nu(dy) < +\infty$

then the functions μ and γ verify the Assumptions 5.1–7.1, where the function τ introduced in Assumption 5.1-[$\mathbf{I_1}$] *is given by*

$$\tau(y) = e^{cd} \max\left(|y|, |e^y - 1|\right)$$

Furthermore the function γ verifies the Assumptions 7.6-[$\mathbf{H1}, \mathbf{H3}, \mathbf{H4}$].

Proof.
Before we start, remark that the function $A \mapsto F_{d,T}(A)$ is strictly increasing, so invertible, and infinitely differentiable: in particular

$$\Phi'(A) = \frac{\int_T^{T+d} \psi(0,s) l(s) e^{l(s)A} ds}{\int_T^{T+d} \psi(0,s) e^{l(s)A} ds}$$

$$\Phi''(A) = \frac{\left(\int_T^{T+d} \psi(0,s) l^2(s) e^{l(s)A} ds\right)\left(\int_T^{T+d} \psi(0,s) e^{l(s)A} ds\right) - \left(\int_T^{T+d} \psi(0,s) l(s) e^{l(s)A} ds\right)^2}{\left(\int_T^{T+d} \psi(0,s) e^{l(s)A} ds\right)^2}$$

from which we deduce

$$e^{-c(T+d)} \leq \Phi'(A) \leq e^{-cT} \quad \text{and} \quad e^{-2c(T+d)} - e^{-2cT} \leq \Phi''(A) \leq e^{-2cT} - e^{-2c(T+d)}$$

From Itô's formula, we obtain

$$dZ_t = \left(\Phi'(A_t) e^{ct} \zeta + \int_{|y| \leq 1} \left(\Phi(A_{t-} + e^{ct} y) - \Phi(A_{t-}) - y e^{ct} \Phi'(A_{t-})\right) \nu(dy)\right) dt + \int_{\mathbb{R}} \left(\Phi(A_{t-} + e^{ct} y) - \Phi(A_{t-})\right) \bar{J}(dydt)$$

or equivalently

$$dZ_t = \mu(t, Z_t) dt + \int \gamma(t, Z_{t-}, y) \bar{J}(dydt)$$

We can now prove that μ and γ verify the Assumptions 5.1. We detail the computations only for the function γ, since similar computations can be done for μ. First we remark that $z \to \gamma(t, z, y)$ is differentiable and we can compute this derivative to obtain

$$\begin{aligned} \partial_z \gamma(t, z, y) &= -1 + (\Phi'(\Phi^{-1}(z)))^{-1} \Phi'(\Phi^{-1}(z) + y e^{ct}) \\ &= e^{ct} y (\Phi'(\Phi^{-1}(z)))^{-1} \int_0^1 \Phi''(\Phi^{-1}(z) + r e^{ct}) dr \end{aligned}$$

so that

$$\begin{aligned} |\partial_z \gamma(t, z, y)| &= \left| e^{ct} y (\Phi'(\Phi^{-1}(z)))^{-1} \int_0^1 \Phi''(\Phi^{-1}(z) + r e^{ct}) dr \right| \\ &\leq |y| e^{cT} (\inf_A |\Phi(A)|)^{-1} \left\|\Phi''\right\|_\infty \leq e^{cT} e^{-c(T+d)} \left\|\Phi''\right\|_\infty \\ &\leq |y| e^{cT} e^{c(T+d)} \left(e^{-2cT} - e^{-2c(T+d)}\right) \leq e^{cd} |y| \end{aligned}$$

From the bounds on the first and second derivative of Φ we obtain $\sup_{t,z} |\partial_z\gamma(t,z,y)| \leq e^{cd}|y|$, which gives us the function ρ introduced in Assumptions 5.1. From the bounds on the first and second derivative of Φ we obtain $\sup_{t,z} |\partial_z\gamma(t,z,y)| \leq e^{-cd}|y|$, which gives us the function ρ introduced in Assumptions 5.1. Also by the definition of Φ in (8.3) we have

$$\exp(e^{-c(T+d)}y) - 1 \leq e^{\gamma(t,z,y)} - 1 \leq e^y - 1$$

if $y > 0$ and the inverse inequality stands in force if $y < 0$, which yield $\sup_{t,z} |e^{\gamma(t,z,y)} - 1| \leq |e^y - 1|$. According to the definition of the function τ given in Assumptions 5.1 and the estimations above we deduce that

$$\tau(y) := \max\left(\sup_{t,u,z}\left(|\gamma(t,z,y)|, \left|e^{\gamma(t,z,y)} - 1\right|\right), \rho(y)\right) = e^{cd}\max\left(|y|, |e^y - 1|\right)$$

which verifies Assumptions 5.1-[**I1**, **I2**] and Assumptions 7.1-[**I**] from *iii*). For Assumption 7.1-[**ND**] we have, from the definition of γ

$$\left(e^{\gamma(t,z,y)} - 1\right)^2 \geq \left(\exp(e^{-c(T+d)}y) - 1\right)^2$$

so then, for some positive $M > 0$ we have

$$\begin{aligned}\Gamma(y) :=& \int_{\mathbb{R}} \inf_{t,z}\left(e^{\gamma(t,z,y)} - 1\right)^2 \nu(dy) \geq \int_{\mathbb{R}} \inf_{t,z}\left(\exp(e^{-c(T+d)}y) - 1\right)^2 \nu(dy)\\ \geq& M\int_{|y|\leq\epsilon} |y|^{1-\alpha} g(y)dy > 0\end{aligned}$$

since $g(0^+)$ and $g(0^-)$ are strictly positive, we can select ϵ small enough and obtain

$$\Gamma(y) \geq M\int_{|y|\leq\epsilon} |y|^{1-\alpha} dy > 0$$

For the Assumptions 7.6 we can differentiate γ w.r.t y to obtain $\partial_y\gamma(t,z,y) = e^{ct}\Phi'(\Phi^{-1}(z) + e^{ct}y)$ so then $e^{-c(T+d)} \leq |\partial_y\gamma(t,z,y)| \leq 1$, which proves that Assumption 7.6-[**H1**] holds true. For Assumption 7.6-[**H3**], one can differentiate $\partial_y\gamma$ w.r.t. z and give for it an upper bound to prove that $z \to \partial_y\gamma(t,z,y)$ is Lipschitz continuous, uniformly in t, y. Assumption 7.6-[**H4**] trivially holds true.

□

We can transform the problem (8.4) by using the process Z to obtain

$$v^f(t,z,x) = \inf_{\theta\in\mathcal{X}(t,z,x)} \mathbb{E}\left[\left(f(Z_T^{t,z}) - x - \int_t^T \theta_{u-} d\exp(Z_u^{t,z})\right)^2\right],\ x, z \in \mathbb{R} \quad (8.5)$$

where $\mathcal{X}(t,z,x)$ is defined in (5.10) and $f(z) = \tilde{f}(e^z)$.

In order to apply our results (Theorems 7.14 and 7.17) we need to verify all the Assumptions 7.6. It is easy to prove that the function γ does not verify Assumption 7.6-[**H2**]: however, as we have already seen, this can be avoided by using Lemma

7.16, whose assumptions are verified by γ. We could change the state variable Z into a new one, say Z', study the problem (8.5) in the new state variable Z' (see the discussion in Section 7.5, page 141) and finally obtain the characterization of the optimal Markovian strategy as a function of the state variable Z'. And by applying the inverse change of variable, we can express this optimal strategy in terms of Z. From a practical and numerical point of view, one can avoid to make this change of variable, and apply directly Theorem 5.14 to obtain the optimal strategy.

To conclude this section, we want to introduce a special class of options one can use in problem (8.5). First define, for some $G > 0$ the function $p(x) : (G-x)^+$, the usual put function, and

$$h(A) := \frac{1}{d'} \int_T^{T+d'} \psi(0,s) e^{l(s)A} ds$$

for some $d' \neq d$. From (8.3) it follows that $h \circ \Phi^{-1}(Z_t) = F_{d',T,t}$, and then, by defining $f := p \circ h \circ \Phi^{-1}$, we obtain $f(Z_t) = (G - F_{d',T,t})^+$, which is a put option written on a future contract with different duration d'. Using this particular option we can rewrite problem (8.5) as follows

$$\text{minimize over } \theta \quad \mathbb{E}\left[\left((G - F_{d',T,t})^+ - x - \int_t^T \theta_{u-} dF_{d,T,u}\right)^2\right]$$

The financial meaning of the above problem is particularly interesting: one tries to hedge (in the quadratic sense) a put option written on a future contract with duration $d' \neq d$ only by using, as hedging instrument, the future contract with duration d. This may be useful when, for example, one sells a future contract with a non-standardized duration in the OTC market and hedges its position only by using the instruments available in the market.

8.3 Numerical approximation of the functions a and b

In this section we will briefly discuss how one can solve the PIDEs obtained in Theorems 5.11–5.14 in the particular example presented in Section 8.2. This was done with the precious collaboration of Xavier Warin, EDF R&D (De Franco, Tankov, and Warin, 2012).

As we proved in Theorem 7.17, the optimal Markovian strategy in problem (8.5) is given by $\theta_t^* = \theta(t, Z_{t-}, X_{t-})$, where

$$\theta^*(t,z,x) := e^{-z}\left(\pi^*(t,z)x - \frac{1}{2}\frac{\mathcal{Q}_t b(t,z)}{\mathcal{G}_t a(t,z)}\right)$$

and

$$\pi^*(t,z) := -\frac{\mathcal{Q}_t a(t,z)}{\mathcal{G}_t a(t,z)} \; (:= \pi_t[a])$$

and the optimal price is given in (5.42) $x^* := -b(t,z)(2a(t,z))^{-1}$ To solve the problem then we only need to compute the function a and b, which in this case, verify

$$0 = -\frac{\partial a}{\partial t} - \mu\frac{\partial a}{\partial z} - \mathcal{B}_t a - \inf_{|\pi|\leq\bar{\Pi}} \left\{2\pi\mathcal{Q}_t a + \pi^2\mathcal{G}_t a\right\}, \qquad a(T,z) = 1$$

$$0 = -\frac{\partial b}{\partial t} - \mu\frac{\partial b}{\partial z} - \mathcal{B}_t b - \pi_t[a]\mathcal{Q}_t b, \qquad b(T,z) = -2f$$

We will present here a numerical method to approximate the functions a and b. For this, we prefer to rewrite the above PIDEs in a more comfortable way: with an abuse of notation, we first invert the time direction $\mu(t,z) \mapsto \mu(T-t,z)$, and we redefine our operators as follows:

$$\mu(t,z) \overset{redef}{:=} \mu(t,z) + \int_{|y|\geq 1} \gamma(t,z,y)\nu(dy)$$

$$\mathcal{B}\varphi(t,z) \overset{redef}{:=} \int_{\mathbb{R}} \left(\varphi(t,z+\gamma(t,z,y)) - \varphi(t,z,y) - \gamma(t,z,y)\frac{\partial\varphi}{\partial z}(t,z)\right)\nu(dy)$$

$$\mathcal{Q}\varphi(t,z) \overset{redef}{:=} \left[\mu(t,z) + \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1 - \gamma(t,z,y)\right)\nu(dy)\right]\varphi(t,z) + \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1\right)\left(\varphi(t,z+\gamma(t,z,y)) - \varphi(t,z)\right)\nu(dy)$$

$$\mathcal{G}\varphi(t,z) \overset{redef}{:=} \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1\right)^2 \varphi(t,z+\gamma(t,z,y))\nu(dy)$$

In particular, we will write $\pi[\varphi] \overset{redef}{:=} -\mathcal{Q}\varphi(t,z)(\mathcal{G}\varphi(t,z))^{-1}$, where $\mathcal{Q}$ and $\mathcal{G}$ are the redefined operators above. With these notations we obtain the new PIDEs verified by a and b

$$0 = -\frac{\partial a}{\partial t} + \mu\frac{\partial a}{\partial z} + \mathcal{B}a + \inf_{|\pi|\leq\bar{\Pi}} \left\{2\pi\mathcal{Q}a + \pi^2\mathcal{G}a\right\}, \qquad a(0,z) = 1 \tag{8.6}$$

$$0 = -\frac{\partial b}{\partial t} + \mu\frac{\partial b}{\partial z} + \mathcal{B}b + \pi[a]\mathcal{Q}b, \qquad b(0,z) = -2f \tag{8.7}$$

In order to solve the above PIDEs, we need to truncate the domain, i.e. we will numerically solve the above PIDEs in $[0,T] \times [-\underline{Z}, \underline{Z}]$. Due to the presence of the integro-differential operator, the boundary conditions must be imposed not only at the boundary $\partial[0,T] \times \{\underline{Z}, \underline{Z}\}$ but also outside this parabolic boundary, let us say on the region $[0,T] \times [-\hat{Z}, -\underline{Z}] \cup [\underline{Z}, \hat{Z}]$. Moreover, we also need to truncate at some $\hat{Y}$ the integrals appearing in the definition of the coefficients in (8.6)–(8.7). We finally assume the following condition:

for all $t \in [0,T]$

$$\underline{Z} + \sup_{z\in\mathbb{R}, y\in[-\hat{Y},\hat{Y}]} \gamma(t,z,y) \leq \hat{Z} \qquad \text{and} \qquad -\underline{Z} + \inf_{z\in\mathbb{R}, y\in[-\hat{Y},\hat{Y}]} \gamma(t,z,y) \geq -\hat{Z}$$

8.3.1 The algorithm for the function a: truncation and first approximation

The truncation procedure transforms PIDE (8.6) into the following:

- on $[0,T] \times [-\underline{Z},\, \underline{Z}]$ we solve

$$\frac{\partial a}{\partial t} + \sup_{\pi \in [-\bar{\Pi}, \bar{\Pi}]} \left\{ -\mu \frac{\partial a}{\partial z} - \mathcal{B}^{tr} a - 2\pi \mathcal{Q}^{tr} a - \pi^2 \mathcal{G}^{tr} a \right\} = 0, \quad a(0,z) = 1 \tag{8.8}$$

where Π is the constant given in (5.44) and the truncated operators are defined as follows:

$$\mathcal{B}^{tr} a(t,z) := \int_{-Y}^{Y} \left(a(t, z + \gamma(t,z,y)) - a(t,z,y) - \gamma(t,z,y) \frac{\partial a}{\partial z}(t,z) \right) \nu(dy)$$

$$\mu^{\mathcal{Q}}(t,z) := \mu(t,z) + \int_{-Y}^{Y} \left(e^{\gamma(t,z,y)} - 1 - \gamma(t,z,y) \right) \nu(dy)$$

$$\mathcal{Q}^{tr} a(t,z) := \mu^{\mathcal{Q}}(t,z) a(t,z) + \int_{-Y}^{Y} \left(e^{\gamma(t,z,y)} - 1 \right) \left(a(t, z + \gamma(t,z,y)) - a(t,z) \right) \nu(dy)$$

$$\mathcal{G}^{tr} a(t,z) := \int_{-Y}^{Y} \left(e^{\gamma(t,z,y)} - 1 \right)^2 a(t, z + \gamma(t,z,y) \nu(dy)$$

- on $[0,T] \times [-\overline{Z},\, -\underline{Z}) \cup (\underline{Z},\, \overline{Z}]$ we impose

$$a(t,z) := 1$$

Remark 8.2. *The effect of truncating the coefficients in the PIDE has been studied in Jakobsen and Karlsen (2005) and we refer to it for an error estimation.*

The choice of the boundary condition $a(t,z) = 1$ is motivated as follows: if $\exp(Z)$ *was a martingale, this is the value of the function a (see for example Remark 5.13). When* $\exp(Z)$ *fails to be a martingale, we should expect that the effect of the drift term should not be too strong. Alternatively, one can replace the process Z, outside the boundary, by a Lévy process, for which the solution can be computed explicitly (it will be a simple time-dependent function).*

We now adapt the methodology proposed by Forsyth et al. (2007) to solve the truncated PIDE (8.8) on a regular grid $z_j = j\Delta z$, for some $\Delta z > 0$ and $j \in (-N, N)$. We also define two integers $j_{-\hat{Z}}$ and $j_{\hat{Z}}$ such that $k_j \in (-\hat{Z},\, \hat{Z})$ if and only if $j_{-\hat{Z}} < j < j_{\hat{Z}}$. To avoid interpolation of the values of a when estimating the integral term, we use a space and time dependent grid to define the intervals of integration: the integration points y_i are chosen to verify $\gamma(t, z, y_i(t,z)) = i\Delta z$. (Remark that this is possible since the function γ is invertible in the variable y). For an integer $k \geq 1$ we split the region $[-\hat{Y},\, \hat{Y}]$ in three domains:

$$\begin{aligned} \hat{\Omega}_0(t,z) &= \left\{ y | y_{-k-\frac{1}{2}}(t,z) \leq y \leq y_{k+\frac{1}{2}}(t,z) \right\}, \\ \hat{\Omega}_1(t,z) &= \left\{ y | y_{k+\frac{1}{2}}(t,z) < y < 1 \text{ or } -1 < y < y_{-k-\frac{1}{2}}(t,z) \right\}, \\ \hat{\Omega}_2(t,z) &= \{ y | 1 \leq |y| \leq Y \}, \end{aligned} \tag{8.9}$$

The parameter k is used to fit the size of the domain $\Omega_0(t,z)$. Due to the very high infinite activity of the jump process near 0, we will numerically fix k equal to 2 or 3 whereas Forsyth et al. (2007) take $k = 1$. We can write then

$$\mathcal{B}^{tr} a(t,z) := \int_{\hat{\Omega}_0(t,z)} \cdots + \int_{\hat{\Omega}_1(t,z)} \cdots + \int_{\hat{\Omega}_2(t,z)} \cdots$$

For the first term we define

$$D(t,z) := \int_{\hat{\Omega}_0(t,z)} \gamma^2(t,z,y)\nu(dy)$$

so that

$$\begin{aligned}
&\int_{\hat{\Omega}_0(t,z)} \left(a(t, z+\gamma(t,z,y)) - a(t,z) - \gamma(t,z,y)\frac{\partial a}{\partial z}(t,z) \right) \nu(dy) \\
&= \frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2} + \int_{\hat{\Omega}_0(t,z)} O(\gamma^3)\nu(dy)
\end{aligned}$$

Since $|\gamma(t,z,y)| \leq M|y|$ around zero for some positive M (see Lemma 8.1), we deduce

$$\begin{aligned}
&\int_{\hat{\Omega}_0(z)} \left(a(t, z+\gamma(t,z,y)) - a(t,z) - \gamma(t,z,y)\frac{\partial a}{\partial z}(t,z) \right) \nu(dy) \\
&= \frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2} + \int_{\hat{\Omega}_0} O(|y|^3)\nu(dy) = \frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2} + O(|y_{k+1/2}(t,z) - y_{-(k+1/2)}|^{3-\alpha}) \\
&= \frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2} + O(\Delta z^{3-\alpha}) \qquad (8.10)
\end{aligned}$$

since for some $\xi \in \hat{\Omega}_0$ we have

$$\begin{aligned}
(2k+1)\Delta z =& |\gamma(t,z,y_{k+1/2}(t,z)) - \gamma(t,z,y_{-(k+1/2)}(t,z))| \\
=& |\partial_y \gamma(t,z,\xi)||y_{k+1/2}(t,z) - y_{-(k+1/2)}|
\end{aligned}$$

and $\partial_y \gamma$ is bounded from below.

In the region $\hat{\Omega}_2$, away from zero, we can subdivide the domain in disjoint intervals centered in y_i and expand the function γ around the integration points y_i, as it is done in Forsyth et al. (2007), to obtain

$$\begin{aligned}
&\int_{\hat{\Omega}_2(z)} \left(a(t, z+\gamma(t,z,y)) - a(t,z) - \gamma(t,z,y)\frac{\partial a}{\partial z}(t,z) \right) \nu(dy) \\
&= \sum_{y_i \in \hat{\Omega}_2} \omega(t,z,y_i) \left(a(t, z+\gamma(t,z,y_i)) - a(t,z) - \gamma(t,z,y_i)\frac{\partial a}{\partial z} \right) + O(\Delta z^2) \\
&= \sum_{y_i \in \hat{\Omega}_2} \omega(t,z,y_i) \left(a(t, z+i\Delta z) - a(t,z) - z\Delta z\frac{\partial a}{\partial z} \right) + O(\Delta z^2)
\end{aligned}$$

where

$$\omega(t,z,y_i) := \int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)} \nu(dy),\ y_i \in \hat{\Omega}_2$$

For the region $\hat{\Omega}_1$ we need to transform the Lévy measure since it has infinite activity close to zero. Following Forsyth et al. (2007), define $\tilde{\nu}(dy) := \nu(y)y^2 dy$ and then

$$y^{-2}\left(a(t,z+\gamma(t,z,y))-a(t,z)-\gamma(t,z,y)\frac{\partial a}{\partial z}(t,z)\right)$$
$$=y_i^{-2}\left(a(t,z+\gamma(t,z,y_i))-a(t,z)-\gamma(t,z,y_i)\frac{\partial a}{\partial z}(t,z)\right)+e(i,y)$$
$$e(i,y):=(y-y_i)\frac{d}{dy}\left(y^{-2}\left(a(t,z+\gamma(t,z,y))-a(t,z,y)-\gamma(t,z,y)\frac{\partial a}{\partial z}(t,z)\right)\right)|_{y=y_i}$$
$$+\frac{(y-y_i)^2}{2}\int_0^1 d\theta$$
$$\int_0^\theta \frac{d^2}{dy^2}\left(y^{-2}\left(a(t,z+\gamma(t,z,u))-a(t,z,u)-\gamma(t,z,u)\frac{\partial a}{\partial z}(t,z)\right)\right)|_{u=y_i+r(y-y_i)}\,dr$$

Then, as before,

$$\int_{\hat\Omega_1}\left(a(t,z+\gamma(t,z,y))-a(t,z)-\gamma(t,z,y)\frac{\partial a}{\partial z}(t,z)\right)\nu(dy)$$
$$=\sum_{i:\,y_i\in\hat\Omega_1}\omega(t,z,y_i)\left(a(t,z+\gamma(t,z,y_i))-a(t,z)-\gamma(t,z,y_i)\frac{\partial a}{\partial z}(t,z)\right)$$
$$+\sum_{i:\,y_i\in\hat\Omega_1}\int_{y_{i-1/2}}^{y_{i+1/2}} e(i,y)\tilde\nu(dy)$$

where

$$\omega(t,z,y_i):=y_i^{-2}\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)}|y|^{1-\alpha}g(y)dy,\ y_i\in\hat\Omega_1$$

We can use the bound on the function γ and its derivative w.r.t. y to control the error term (See appendix A in Forsyth et al. (2007))

$$\sum_{i:\,y_i\in\hat\Omega_1}\int_{y_{i-1/2}}^{y_{i+1/2}} e(i,y)\tilde\nu(dy)=O(\Delta z^{\min(2-\epsilon,3-\alpha)})$$

for any $\epsilon>0$. We finally add up all the above estimations to obtain

$$\mathcal{B}^{tr}a(t,z):=\frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2}$$
$$+\sum_{y_i,\,|i|>k}\omega(t,z,y_i)\left(a(t,z+i\Delta z)-a(t,z)-i\Delta z\frac{\partial a}{\partial z}(t,z)\right)+O\left(\Delta z^{\min(2-\epsilon,3-\alpha)}\right)$$

where

$$\omega(t,z,y_i):=\begin{cases} y_i^{-2}\displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)}|y|^{1-\alpha}g(y)dy & \text{if}\quad y_{i+1/2}(t,z),y_{i-1/2}(t,z)\in\hat\Omega_1\\ \displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)}|y|^{-(1+\alpha)}g(y)dy & \text{otherwise}\end{cases}\tag{8.11}$$

The above terms can be calculated with the trapezoidal rule with an integration over 5 points, if the function g is twice continuously differentiable with bounded second derivative away from zero: when $y_i \in \hat{\Omega}_1$ we obtain an error of order $O(\Delta z^{3-\alpha})$ whereas for $y_i \in \hat{\Omega}_2$ we obtain an error of order $O(\Delta z^3)$.

Similarly we will treat the operators $\mathcal{Q}^{tr}$ and $\mathcal{G}^{tr}$. We skip the details to obtain:

$$\begin{aligned}\mathcal{Q}^{tr} a(t,z) :=& \mu^{\mathcal{Q}}(t,z)a(t,z) + D(t,z)\frac{\partial a}{\partial z} \\ &+ \sum_{y_i,\,|i|>k} I^{\mathcal{Q}}(t,z,y_i)\left(a(t,z+i\Delta z) - a(t,z)\right) + O(\Delta z^{\min(2-\epsilon,3-\alpha)}) \\ \mathcal{G}^{tr} a(t,z) :=& D(t,z)a(t,z) + \sum_{y_i,\,|i|>k} I^{\mathcal{G}}(t,z,y_i)a(t,z+i\Delta z) + O(\Delta z^{2-\epsilon})\end{aligned}$$

where

$$I^{\mathcal{Q}}(t,z,y_i) := \begin{cases} y_i^{-2}\displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)} (e^{\gamma}-1)\,|y|^{1-\alpha} g(y)dy & \text{if} \quad y_{i+1/2}(t,z), y_{i-1/2}(t,z) \in \hat{\Omega}_1 \\ \displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)} (e^{\gamma}-1)\,|y|^{-(1+\alpha)} g(y)dy & \text{otherwise} \end{cases}$$

and

$$I^{\mathcal{G}}(t,z,y_i) := \begin{cases} y_i^{-2}\displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)} (e^{\gamma}-1)^2\,|y|^{1-\alpha} g(y)dy & \text{if} \quad y_{i+1/2}(t,z), y_{i-1/2}(t,z) \in \hat{\Omega}_1 \\ \displaystyle\int_{y_{i-1/2}(t,z)}^{y_{i+1/2}(t,z)} (e^{\gamma}-1)^2\,|y|^{-(1+\alpha)} g(y)dy & \text{otherwise} \end{cases}$$

Again a five point approximation can be used to estimate the above integrals.

The above computation allows us to rewrite PIDE (8.8) into the following:

$$\begin{aligned}\frac{\partial a}{\partial t}(t,z) - \frac{D(t,z)}{2}\frac{\partial^2 a}{\partial z^2} - \mu(t,z)\frac{\partial a}{\partial z} + \sup_{\pi\in[-\bar{\Pi},\bar{\Pi}]} \Big\{ \Big(\tilde{V}(t,z) \; -2\pi D(t,z)\Big)\frac{\partial a}{\partial z} \\ - \sum_{|i|\geq k} \tilde{W}(t,z,y_i,\pi)a(t,z+i\Delta z) + \tilde{R}(t,z,\pi)a(t,z)\Big\} \\ a(0,z) = 1 \qquad (8.12)\end{aligned}$$

where

$$\tilde{W}(t,z,y_i,\pi) := \omega(t,z,y_i) + 2\pi I^{\mathcal{Q}}(t,z,y_i) + \pi^2 I^{\mathcal{G}}(t,z,y_i)$$

$$\begin{aligned}\tilde{V}(t,z) &:= \sum_{|i|\geq k} i\Delta z\omega(t,z,y_i) \\ \tilde{R}(t,z,\pi) &:= \sum_{|i|>k}\left(\omega(t,z,y_i) + 2\pi I^{\mathcal{G}}(t,z,y_i)\right) - 2\pi\mu^{\mathcal{Q}}(t,z) - \pi^2 D(t,z)\end{aligned}$$

Remark 8.3. *The same type of computations can be used to approximate the coefficients appearing in the PIDEs* (8.6)–(8.7), *like, for example,*

$$\int_{|y|\geq 1} \gamma(t,z,y)\nu(dy) \qquad \textit{and} \qquad \int_{\mathbb{R}} \left(e^{\gamma(t,z,y)} - 1 - \gamma(t,z,y)\right)\nu(dy)$$

8.3.2 The algorithm for the function a: numerical scheme and convergence

Fix $\bar{n} \in \mathbb{N}$, $\Delta t = T/\bar{n}$ and consider $t_n = n\Delta t$. To solve PIDE (8.12) we use an implicit scheme for linear part, which basically corresponds to a classical diffusion, and an explicit one is used for the integral part: if a^n stands for $a(n\Delta t, .)$ then this will lead to

$$\begin{gathered}\frac{a^{n+1}-a^n}{\Delta t} - \frac{D(t^{n+1},.)}{2}\frac{\partial^2 a^{n+1}}{\partial z^2} - \\ \mu(t^{n+1},z)\frac{\partial a^{n+1}}{\partial z} + \sup_{\pi\in[-\Pi,\Pi]}\Bigg[(\tilde{V}(t^n,.) - 2\pi D(t^n,z))\frac{\partial a^n}{\partial z} + \\ \tilde{R}(t^n,.\pi)a^n - \sum_{|i|>k}\tilde{W}(t^n,.,y_i,\pi)a^n(z+i\Delta z)\Bigg] = 0 \end{gathered} \qquad (8.13)$$

For the implicit term a classical central difference scheme (order two) is used for the first order differential term coupled with forward/backward differencing when matrix coefficients due to diffusion are negative (see for example Forsyth et al. (2005)). The explicit first order differential term is treated to have monotony of the scheme. This transforms equations (8.13) into

$$\begin{aligned} &a_j^{n+1}(1+\Delta t(\alpha_j(t^{n+1})+\beta_j(t^{n+1})) - \Delta t\alpha_j(t^{n+1})a_{j-1}^{n+1} - \Delta t\beta_j(t^{n+1})a_{j+1}^{n+1} \\ &+ \sup_{\pi\in[-\Pi,\Pi]}\Bigg[a_j^n(-1+\Delta t(\tilde{R}(t^n,z_j,\pi^n) + \frac{|\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j)|}{\Delta z}) \\ &-a_{j-1}^n\Delta t\frac{(\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j))^+}{\Delta z} - a_{j+1}^n\Delta t\frac{(\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j))^-}{\Delta z} \\ &-\Delta t\sum_{|i|>k}\tilde{W}(t^n,z_j,y_i,\pi_j^n)a_{j+i}^n\Bigg] = 0 \end{aligned} \qquad (8.14)$$

where a_j^{n+1} stands for an approximation of $a^{n+1}(z_j)$ calculated at point z_j and α_j and β_j are positive weights (see for example Forsyth and Labahn (2007)) given by :

$$\begin{aligned} \alpha_{j,central}(t,) &= \frac{D(t,z_j)}{2\Delta z^2} - \frac{\mu(t,z_j)}{2\Delta z} \\ \beta_{j,central}(t) &= \frac{D(t,z_j)}{2\Delta z^2} + \frac{\mu(t,z_j)}{2\Delta z} \end{aligned}$$

if $\alpha_{j,central}(t)$ or $\beta_{j,central}(t,\pi)$ is negative, we use

$$\begin{aligned} \alpha_{j,forward/backward}(t) &= \frac{D(t,z_j)}{2\Delta z^2} + max(0, -\frac{\mu(t,z_j)}{\Delta z}) \\ \beta_{j,forward/backward}(t) &= \frac{D(t,z_j)}{2\Delta z^2} + max(0, \frac{\mu(t,z_j)}{\Delta z}) \end{aligned}$$

Denote with $\mathcal{M}$ the matrix

$$\begin{aligned}
\mathcal{M}_{j,j}(t^{n+1}) &= \alpha_j(t^{n+1}) + \beta_j(t^{n+1}),\\
\mathcal{M}_{j,j+1}(t^{n+1}) &= \beta_j(t^{n+1}),\\
\mathcal{M}_{j,j-1}(t^{n+1}) &= \alpha_j(t^{n+1}),
\end{aligned}$$

and with B the matrix

$$\begin{aligned}
B_{j,j}(t^n,\pi) &= \tilde{R}(t^n,\pi,z_j) + \frac{|\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j)|}{\Delta z},\\
B_{j,j-1}(t^n,\pi) &= -(\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j))^+,\\
B_{j,j+1}(t^n,\pi) &= -(\tilde{V}(t^n,z_j) - 2\pi D(t^n,z_j))^-,\\
B_{j,j+i}(t^n,\pi) &= -\tilde{W}(t^n,\pi,z_j,y_i),\ \text{if } |i| > k,\\
B_{j,j+i}(t^n,\pi) &= 0,\ \text{for } 1 < |i| \le k,
\end{aligned}$$

The system can be written in terms of these matrices:

$$(I + \Delta t M(t^{n+1}))a^{n+1} + \sup_{\pi\in(-\Pi,\Pi)} (-I + \Delta t B(t^n,\pi))a^n = 0 \tag{8.15}$$

Theorem 8.4. *Under the CFL condition*

$$\sup_{t,j}\left[\frac{\tilde{V}(t,z_j)}{\Delta z} + 2\Pi(|\mu(t,z_j)| + D(t,z_j)) + \sum_{|i|>k}(2\Pi I^{\mathcal{Q}}(t,z_j,y_i) + \omega(t,z_j,y_i)) + \Pi^2 D(t,z_j)) + \right]\Delta t < 1$$

the scheme 8.15 is consistent, monotone, L_∞ stable and converges to the viscosity solution of equation (8.8)*.*

The proof of this result can be found in De Franco, Tankov, and Warin (2012). We just remark that the integer $k \ge 1$ in (8.9) is chosen to improve the convergence of the scheme. This is needed since the jump activity of the process may be very high, and then the error we do in our approximations (in the Taylor expansions as, for example, in (8.10), or in the approximated integrals as (8.11)) can be very important. Taking $k \ge 1$ means that we extend the critical region $\hat{\Omega}_0(t,z)$ in (8.9) and this allows us to be more precise.

8.3.3 Resolution methodology to calculate the function b

We will use the same methodology developed in paragraphs 8.3.1–8.3.2 for the PIDE (8.7): we first rescale the function b to $\tilde{b}(t,z) := e^{-\eta t} b(t,z)$ and then we proceed by truncating the domain of definition of b to obtain the new PIDE:

- on $[-\underline{Z}, \underline{Z}]$

$$\begin{cases} \dfrac{\partial \tilde{b}}{\partial t} - \mu\dfrac{\partial \tilde{b}}{\partial z} - \mathcal{B}^{tr}\tilde{b} + \eta\tilde{b} - \pi^{tr}[a]\mathcal{Q}^{tr} b = 0 \\ \tilde{b}(0,z) = -2f(z) \end{cases} \tag{8.16}$$

  where $\pi^{tr}[a] := -\mathcal{Q}^{tr} a \left(\mathcal{G}^{tr} a\right)^{-1}$

- on $[-\overline{Z}, -\underline{Z}) \cup (\underline{Z}, \overline{Z}]$ we impose

$$\tilde{b}(t,z) = -2f(z)a(t,z)e^{-\eta t}$$

Remark 8.5. *The choice of the boundary condition outside the domain is justified as follows: the value $-b(t,z)/2a(t,z)$ can be interpreted as the cost of hedging the pay-off f, that is, the wealth at time t which leads to the minimal hedging error at maturity, as stated in (5.42). In the regions far from the money (and under the assumption of zero interest rate), the cost of hedging can be approximated by the option's pay-off, whence the boundary condition for b.*

Using the same discretization as before we get the following equation to solve on $[-\underline{Z}, \underline{Z}]$:

$$\begin{aligned} &\frac{\partial \tilde{b}}{\partial t} - \frac{D(t,z)}{2}\frac{\partial^2 \tilde{b}}{\partial z^2} + \frac{\partial \tilde{b}}{\partial z}\Big((\tilde{V}(t,z) - \mu(t,z) - \pi^{tr}[a]D(t,z)\Big) \\ &+\tilde{b}\left(\eta + \hat{R}(t,z,\pi^{tr}[a])\right) - \sum_{|i|\geq k} \hat{W}(t,z,y_i,\pi^{tr}[a])\tilde{b}(t,z+i\Delta z) \\ &\tilde{b}(0,z) = -2f(z) \end{aligned} \tag{8.17}$$

where

$$\hat{W}(t,z,y_i,\pi^{tr}[a]) := \omega(t,z,y_i) + \pi^{tr}[a]I^{\mathcal{Q}}(t,z,y_i)$$

$$\hat{R}(t,z,\pi^{tr}[a]) := \sum_{|i|>k}\left(\omega(t,z,y_i) + \pi^{tr}[a]I^{\mathcal{Q}}(t,z,y_i)\right) - \pi^{tr}[a]\mu^{\mathcal{Q}}(t,z)$$

and $\tilde{V}$ and $D(t,z)$ are the functions introduced in paragraph 8.3.1. We propose, as in paragraph 8.3.2, the following time discretization scheme :

$$\begin{aligned} &\frac{\tilde{b}^{n+1} - \tilde{b}^n}{\Delta t} - \frac{D(t^{n+1},.)}{2}\frac{\partial^2 \tilde{b}^{n+1}}{\partial z^2} + \\ &(\tilde{V}(t^{n+1},.) - \pi[a]^{n+1}D(t^{n+1},z) - \mu(t^{n+1},z))\frac{\partial \tilde{b}^{n+1}}{\partial z} + (\hat{R}(t^{n+1},.,\pi[a]^{n+1}) + \eta)\tilde{b}^{n+1} - \\ &\sum_{|i|>k}(\hat{W}(t^{n+1},.,y_i,\pi[a]^{n+1}))^+ b^n(z+i\Delta z) + \\ &\sum_{|i|>k}(\hat{W}(t^{n+1},.,y_i,\pi[a]^{n+1}))^- b^n(z+i\Delta z) = 0 \end{aligned}$$

which becomes

$$\begin{aligned} &\tilde{b}_j^{n+1}(1 + \Delta t(\alpha_j(t^{n+1}) + \beta_j(t^{n+1}) + \hat{R}(t^{n+1},z_j,\pi[a]^{n+1}) + \eta) - \\ &\Delta t\alpha_j(t^{n+1})\tilde{b}_{j-1}^{n+1} - \Delta t\beta_j(t^{n+1})\tilde{b}_{j+1}^{n+1} + \Delta t\sum_{|i|>k}(\hat{W}(t^{n+1},z_j,y_i,\pi[a]^{n+1}))^-\tilde{b}_{j+i}^{n+1} \\ &-\Delta t\sum_{|i|>k}(\hat{W}(t^{n+1},z_j,y_i,\pi[a]^{n+1}))^+\tilde{b}_{j+i}^{n} = 0 \end{aligned} \tag{8.18}$$

where $\tilde{b}_j^n$ is the approximated value of $\tilde{b}$ at (t_n, z_j), and α_j, β_j positive weights given by :

$$\begin{aligned}\alpha_{j,central}(t,\pi) &= \frac{D(t,z_j)}{2\Delta z^2} + \frac{\tilde{V}(t,.) - \pi D(t,z) - \mu(t,z_j)}{2\Delta z}\\ \beta_{j,central}(t,\pi) &= \frac{D(t,z_j)}{2\Delta z^2} - \frac{\tilde{V}(t,.) - \pi D(t,z) - \mu(t,z_j)}{2\Delta z}\end{aligned}$$

if $\alpha_{j,central}(t,\pi)$ or $\beta_{j,central}(t,\pi)$ is negative, we use

$$\begin{aligned}\alpha_{j,forward/backward}(t,\pi) &= \frac{D(t,z_j)}{2\Delta z^2} + \left(\frac{\tilde{V}(t,.) - \pi D(t,z) - \mu(t,z_j)}{\Delta z}\right)^+\\ \beta_{j,forward/backward}(t,\pi) &= \frac{D(t,z_j)}{2\Delta z^2} + \left(\frac{\tilde{V}(t,.) - \pi D(t,z) - \mu(t,z_j)}{\Delta z}\right)^-\end{aligned}$$

Proposition 8.6. *For a space discretization accurate enough (Δz small enough), taking*

$$\eta = (\Pi + \epsilon)\,\|\mu\|_\infty + 2\int_{|y|>1}(1 + \Pi|e^y - 1|)\nu(dy)$$

the scheme (8.18) *is consistent and stable so it converges to the viscosity solution of the PIDE* (8.16)*.*

Again, the proof of the above result is given in De Franco, Tankov, and Warin (2012). Remark that this result is not the same of Theorem 8.4: the function b depends on a through the optimal control π, for which we do not control the sign. This has an important impact on the convergence of the Scheme for b.

8.4 A numerical example

In this last paragraph we will study the problem (8.5) when L in (8.1) is a Normal Inverse Gaussian process with parameter $\alpha, \beta, \delta, \mu$: $L_t \sim NIG(\alpha, \beta, \delta t, \mu t)$.

Remark 8.7. *Remark that α should not be mistaken for the parameter in Lemma 8.1; similarly μ is not the drift function given in the same Lemma. We use this notation because it is standard in the literature.*

We can write then

$$L_t = \left(\mu + \frac{\beta\delta}{\alpha^2 - \beta^2} + \int_{|y|\geq 1} y\nu(dy)\right)t + \int_0^t \int_{\mathbb{R}} y\bar{J}(dyds)$$

where J is a Poisson random measure with intensity

$$\nu(dy) = \frac{\alpha\delta}{\pi|y|}K_1(\alpha|y|)e^{\beta y}dy \qquad \nu(dy) \overset{y\to 0}{\sim} \frac{1}{|y|^2}dy$$

$$\nu(dy) \overset{y\to+\infty}{\sim} \frac{1}{|y|^{3/2}}e^{-(\alpha-\beta)y}dy \qquad \nu(dy) \overset{y\to-\infty}{\sim} \frac{1}{|y|^{3/2}}e^{-(\alpha+\beta)|y|}dy$$

where K_1 is the modified Bessel function of the second kind (paragraph 4.4.3 in Cont and Tankov (2004)).

Remark 8.8. *The NIG is a infinite variation Lévy process with stable-like behavior of small jumps, and since the Blumenthal-Getoor index is equal to* 1*, we cannot apply Lemma 8.1 and then Theorems 7.14–7.17. It is nevertheless a case of interest because the NIG model is popular among practitioners.*

We want to solve problem (8.5) for European options with maturity $T = 1$ week and duration d equal to 7 days of the week. We recall that the future contract in this case is given by

$$F_{7days,1week,t} = \frac{1}{7}\int_7^{14} \psi(0,s)e^{g(s)A_t}ds \tag{8.19}$$

and A_t is given in (8.2) relatively to the NIG process L given above.

In Table 8.1 we give the forward curve for the week, whereas the discount factor c is taken equal to 0.19. As we said in Remark 8.8, we cannot apply Lemma 8.1 and Theorems 7.14–7.17 to solve problem (8.5) by the mean of the PIDEs (8.6)–(8.7). It is nevertheless interesting to see what we get when we compute numerically these PIDEs in the case of the NIG process. We obtain the dynamics of the process Z by using the same computations of Lemma 8.1, and in this numerical experiment we use the following parameters of the NIG process: $\mu = 0.08$, $\alpha = 6.23$, $\beta = 0.06$, $\delta = 0.1027$. Remark that the Lévy measure $\nu(dy)$ relative to these parameters satisfies, at least, the regularity and integrability conditions $i) - ii) - iii)$ of the above mentioned Lemma. Remark also that in this case Z is not a Lévy process: we cannot use the method given in Hubalek et al. (2006) to obtain the optimal strategy in problem (8.5).

Day	s	Price ($\psi(0,s)$)
Monday	$s \in [7,8)$	80
Tuesday	$s \in [8,9)$	90
Wednesday	$s \in [9,10)$	70
Thursday	$s \in [10,11)$	90
Friday	$s \in [11,12)$	80
Saturday	$s \in [12,13)$	70
Sunday	$s \in [13,14]$	60

Table 8.1: The forward curve. Prices are given in Eur.

For all numerical experiments we suppose that $\hat{Z} = 12$, $\underline{Z} = 8$, we take a number of meshes equal to 800 and a number of time step equal to 800. The value k, used to define the domains $\hat{\Omega}_i$ in (8.9), is taken equal to 3. We already discussed on the fact that it is numerically better to take $k \geq 1$ when the jump activity of the process is important, as in the NIG case.
All the MonteCarlo calculations are carried out with 2 million particles. We can now apply the scheme (8.15) to compute the function a. Figure 8.5 shows what we obtain by using a sufficiently accurate discretization procedure. Although we cannot apply Theorem 7.14, we observe that, numerically, the function a is sufficiently smooth. How to explain this regularity? The reason must be sought in the discretization procedure proposed in Paragraph 8.3.1, in where we replaced the non local first order operator $\mathcal{B}^{tr}a$ with a second order term of the form $D(t,z)\partial^2_{zz}a +$ first order terms.

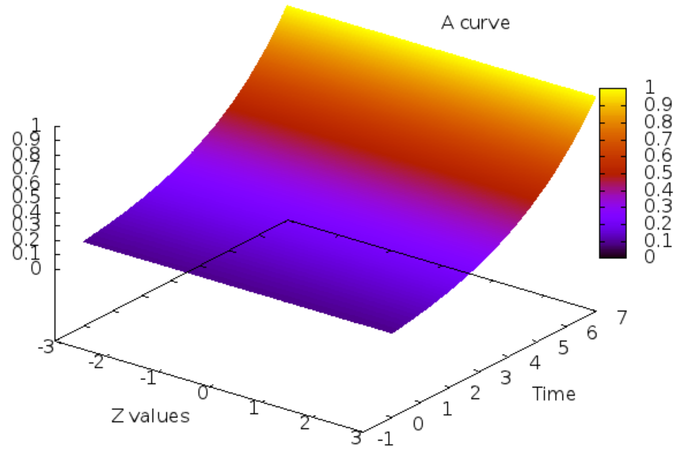

Figure 8.5: The value function $(t, z) \to a(t, z)$ for the NIG process.

Heuristically this is equivalent to replace the small jumps of the process Z with a Brownian motion. The second order term $D(t, z)\partial^2_{zz}a$ transform the PIDE (8.6) into a non degenerate parabolic second order PIDE with artificial boundary conditions (the PIDE (8.12)). As we will explain in Chapter 9, this PIDE has a unique smooth solution, and this explain why we obtain a smooth numerical approximation of the function a. This is one of the reasons that motivated us to test our PIDEs for the NIG process, even if, as we said, the behavior of small jumps does not fulfill our initial Assumptions 7.1.

The optimal control $\pi^*(t, z)$ is shown in Figure 8.6. We now use the function a and the optimal control π^* to solve the PIDE (8.7) by means of the scheme (8.18). In Figure 8.7 we present the result for an at-the-money call option with strike 1 on the future contract introduced in (8.19), whereas Figure 8.8 shows the profile of an at-the-money put option.

8.4.1 The martingale case

The quadratic hedge problem is relatively easy when the underlying stock price is a martingale. Practitioners usually compute the hedge strategy by supposing that the underlying stock price process is a martingale. Assuming that F is a martingale means that we should have

$$F_{d,T,t} := \frac{1}{d} \int_T^{T+d} \psi(0, s) \exp\left(M(s, t) + l(s)A_t\right) ds$$

for some M that makes F a martingale under the historical probability $\mathbb{P}$. From the definition of L and Itô's formula it is easy to prove that M defined as follows

$$-dM(s, t) := e^{-c(s-t)} \left(\mu + \frac{\delta\beta}{\alpha^2 - \beta^2}\right) dt + \int \left(\exp(e^{-c(s-t)}y) - 1 - e^{-c(s-t)}y\right) \nu(dy)dt$$

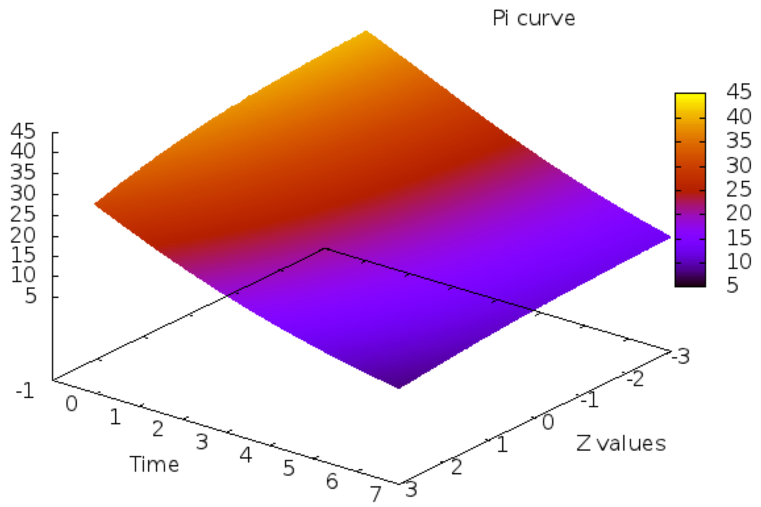

Figure 8.6: The optimal control $(t, z) \to \pi^*(t, z)$ for the NIG process.

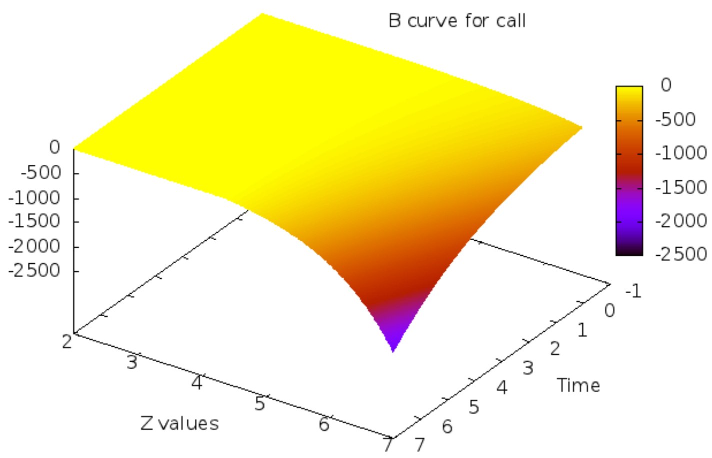

Figure 8.7: The value function $b(t, z)$ for an at-the-money call option.

makes F a martingale. If we want to solve problem (8.5) with the above F, then, as usual, we have to compute the functions a and b, solutions of the PIDEs (8.6)–(8.7). From Remark 5.13 we know that, in the martingale case, the function a is equal to 1. The optimal strategy is then completely determined by the function b, which is the solution of a linear PIDE. We compute numerically this strategy and compare it with the one previously found, when we considered the model (8.19), i.e. when F was not supposed to be a martingale. We propose to evaluate the loss of efficiency when using these two hedge strategies on call and put options with

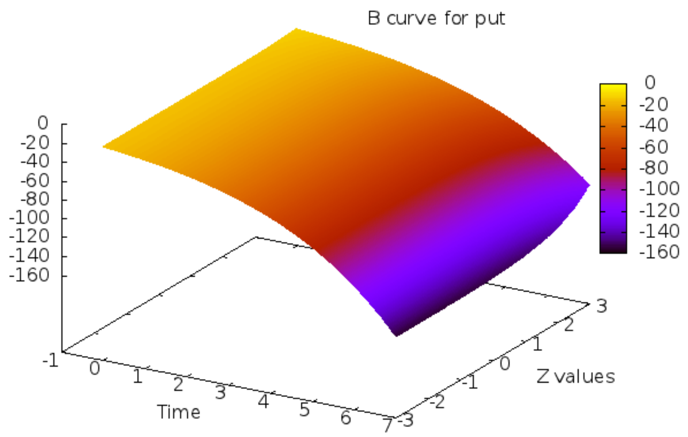

Figure 8.8: The value function $b(t, z)$ for an at-the-money put option.

different moneyness. Our efficiency comparison criterion is the following: if H is the option (call/put) and $\theta^{true}, \theta^{mart}$ are, respectively, the quadratic and the martingale hedge strategies, then the efficiency is measured in terms of the standard deviation of the hedged portfolios:

$$\text{efficiency}(\theta^{true})^2 := \text{variance}\left(H(F_{d,T,t}) - x^{true} - \int_t^T \theta_{r-}^{true} dF_{d,T,r}\right)$$

where x^{true} is the true optimal price given in (5.42). Similarly

$$\text{efficiency}(\theta^{mart})^2 := \text{variance}\left(H(F_{d,T,t}) - x^{mart} - \int_t^T \theta_{r-}^{mart} dF_{d,T,r}\right)$$

where x^{mart} is the price given in (5.42) when one uses the function a and b relative to the martingale model., i.e. x^{mart} is the risk neutral price of H. Table 8.2 resumes our analysis when $t = 0$. The numerical experiment proves that one loses efficiency when using the martingale hedge strategy. This is coherent with the fact that θ^{true} achieves the minimum un problem (8.5), so then it overperforms the strategy θ^{mart}, which is just an admissible strategy in the previous mentioned problem.

Option H	Moneyness	Option value (x^{true})	efficiency(θ^{true})	efficiency(θ^{mart})
Call	1	4.199	1.085	1.316
Put	1	4.213	1.087	1.315
Call	1.5	0.120	0.168	0.212
Put	1.5	38.80	0.175	0.34

Table 8.2: Pricing and standard deviation of hedged portfolio

Chapter 9

The PIDE truncation effect

The chapter is organized as follows: we start by explaining why it is important to study the semi linear PIDE verified by the value function of the pure investment problem, when one truncates the domain of solvability (Section 9.1). For this, we first provide an approximation of this value function by cutting the small jumps of the Lévy measure and replace them with a term involving the second derivative, and then we prove the convergence of this approximation to the true value function (Section 9.2). We then prove that the truncated PIDE relative to this approximation has a unique smooth solution (Section 9.3), and we finally prove an estimate on the error due to the truncation of the domain (Section 9.4).

Contents

9.1 Motivations 177
9.2 A first approximation 179
9.3 The semi linear PIDE on a bounded domain 182
9.3.1 Existence and Uniqueness in the viscosity sense 182
9.3.2 Existence and Uniqueness in the viscosity sense 184
9.4 Estimate on the truncated PIDE 191

9.1 Motivations

In the example proposed in Chapter 8 we showed how to solve the PIDEs (8.6)–(8.7) to obtain the value function and the optimal control of problem (8.5). Our theoretical results ensured that these PIDEs have a smooth classical solution on the domain $[0, T] \times \mathbb{R}$. Nevertheless, in order to implement a numerical scheme, we needed to truncate the solvability domain of these PIDEs: we consider (8.6)–(8.7) in a bounded domain of the form $[0, T] \times [-\underline{Z}, \underline{Z}]$ with artificial Dirichlet conditions at the boundary. It is then natural to ask whenever these truncated PIDEs still have a smooth solution and how the artificial boundary conditions affect the solution.

To simplify the presentation we consider the model of Chapter 6 and we assume that the process Z in (5.5) does not depend on U. We will also concentrate on the PIDE verified by the function a, since a similar approach can be used for b. Remark that, from a practical point of view, one only needs of a and b to compute

the optimal strategy of problem (5.11), which is given in (5.50).
We recall that if a denotes the function in (5.13), then $(t,z) \to a(t,z)e^{\eta t}$ is the unique solution of

$$0 = -\frac{\partial a}{\partial t} + \eta a + \mathcal{A}_t a - \mathcal{B}_t a - \mathcal{H}[a], \qquad a(T,z) = e^{\eta T}$$

and it belongs to $C^{(1-\delta)/2+1,2+(1-\delta)}([0,T]\times\mathbb{R})$, for any $\eta \geq 0$. From a numerical point of view one is instead interested in

$$\begin{cases} 0 = -\frac{\partial a}{\partial t} + \eta a + \mathcal{A}_t a - \mathcal{B}_t a - \mathcal{H}_t[a] & (t,z) \in [0,T) \times (-\underline{Z},\, \underline{Z}) \\ a(T,z) = e^{\eta T} & z \in (-\underline{Z},\, \underline{Z}) \\ a(t,z) = e^{\eta t} q(t,z) & (t,z) \in [0,T] \times (-\underline{Z},\, \underline{Z})^c \end{cases} \tag{9.1}$$

for some artificial Dirichlet boundary condition q. Remark that it is possible to consider a more general domain of the form (Z_1, Z_2): in general the choice of the domain depends on the particular needs of the numerical discretization. However, it is important for the sequel that the domain has a smooth boundary: in the one dimensional case this is trivially true if one considers open intervals. For the multidimensional case one should consider, for example, bounded cylinders. The above truncated PIDE naturally arises when one wants to compute numerically the function a. Our aim is then to prove that it has a unique solution, which also is smooth inside the domain, and give some estimate of the error between the function a and the solution of the above PIDE. The analysis of this PIDE may be very difficult when the intensity measure is not finite. This is related to the behavior of the non local operator

$$(t,z) \to \mathcal{B}a(t,z) := \int \left(a(t,z+\gamma(t,z,y)) - a(t,z) - \gamma(t,z,y)\frac{\partial a}{\partial z}(t,z)\mathbb{1}_{\{|y|\leq 1\}} \right) \nu(dy)$$

when z reaches the boundary of the truncated domain. Heuristically, we can remark that the operator $\mathcal{B}$ behaves as $\partial^2_{zz}a(t,z)+$ "first order operator". Since there is no hope to prove that the solution of PIDE (9.1) is twice continuously differentiable at the boundary $[0,T] \times \{-\underline{Z}, \underline{Z}\}$, the operator $\mathcal{B}$ will be not properly defined on the above boundary, and the map $(t,z) \to \mathcal{B}a(t,z)$, for $|z| \leq \underline{Z}$, may fail to be uniformly Hölder continuous. Much more tractable is the case when the intensity measure $\nu(dy)$ is finite. For this, before considering the truncated PIDE (9.1), we show how to transform the initial parameters of the model (5.5), $(\mu, \sigma^2, \gamma, \nu(dy))$ in order to obtain a finite intensity measure. As we will see in Section 9.2, this is done by cutting the small jumps of the process Z and "replace" them with a Brownian motion, as we did in Paragraph 8.3.1. This replacing transforms, at the same time, the jump measure of the process into a finite intensity measure and the volatility function σ. Our task then reduces to considering the value function of the pure investment problem when the process Z has finite jumps. In the sequel, we will denote this new function a^h, where $h > 0$ is the level at which we cut the small jumps. A first result is to prove that this new value function a^h converges to the

function a when $h \to 0$. In Section 9.3 we then study the truncated PIDE which characterizes the function a^h, which has the same structure as PIDE (9.1), except for the fact that the non local operators are all of order zero.
Throughout the chapter we denote $\Omega^T := [0,T] \times [-\underline{Z}, \underline{Z}]$ and $\mho^T := [0,T] \times (-\underline{Z}, \underline{Z})^c$. Moreover, for a function φ defined on Ω^T, $\|\varphi\|_{l/2,l,\Omega^T}$ denotes the Hölder norm of φ relatively to Ω^T: for example,

$$\|\varphi\|_{\infty,\Omega^T} := \sup_{t\le T, |z|\le \underline{Z}} |\varphi(t,z)| \quad \text{and} \quad \langle\varphi\rangle^{(l)}_{z,\Omega^T} := \sup_{(t,z),(t,z')\in\Omega^T} \frac{|\varphi(t,z)-\varphi(t,z')|}{|z-z'|^l}$$

The same convention stands in force for the Hölder norm of type 2, and for the Hölder norm on the domain $\mho^T$.

9.2 A first approximation

Let $h > 0$ and consider

$$\gamma^h(t,z) := \int_{|y|\le h} \gamma^2(t,z,y)\nu(dy)$$

Assumptions 5.1–6.1 on γ show that γ^h is bounded, Lipschitz continuous w.r.t. t and z and that $\gamma^h \to 0$ when $h \to 0$, uniformly in t, z. Since

$$\begin{aligned}\mathcal{B}\varphi(t,z) = &\int_0^1 d\theta \int_0^\theta d\theta' \int_{|y|\le h} \gamma^2 \frac{\partial^2\varphi}{\partial z^2}(t, z+\theta'\gamma(t,z,y))\nu(dy) \\ &+ \int_{|y|>h} \left(\varphi(t,z+\gamma(t,z,y)) - \varphi(t,z) - \gamma(t,z,y)\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)\end{aligned}$$

then we could replace the above operator with

$$\frac{\gamma^h(t,z)}{2}\frac{\partial^2\varphi}{\partial z^2}(t,z) + \int_{|y|>h} \left(\varphi(t,z+\gamma(t,z,y)) - \varphi(t,z) - \gamma(t,z,y)\frac{\partial\varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)$$

This is equivalent to consider the pure investment problem (5.13) with the new parameters

$$\left(\mu,\, \sigma^2 + \gamma^h, \gamma,\, \nu(dy)\mathbb{1}_{\{|h|<y\}}\right) \tag{9.2}$$

If a^h denotes the value function of the pure investment problem corresponding to these initial parameters, then from Theorems 5.11–6.8 we obtain that the map $(t,z) \to a^h(t,z)e^{\eta t}$ is the unique solution of

$$0 = -\frac{\partial a^h}{\partial t} + \eta a^h + \mathcal{A}^h_t a^h - \mathcal{B}^h_t a^h - \mathcal{H}^h[a^h], \qquad a^h(T,z) = e^{\eta T} \tag{9.3}$$

provided that $T < T^{*,h}$, where $T^{*,h}$ is the maximal time given in Theorem 5.4 relatively to the new parameters. The above operators are given by

$$\begin{aligned}
\mathcal{A}_t^h a^h :=& -\mu\frac{\partial a^h}{\partial z} - \frac{1}{2}\left(\sigma^2+\gamma^h\right)\frac{\partial^2 a^h}{\partial z^2}\\
\mathcal{B}_t^h a^h :=& \int_{h<|y|}\left(a^h(t,z+\gamma) - a^h(t,z) - \gamma\frac{\partial a^h}{\partial z}\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)\\
\mathcal{Q}_t^h a^h :=& \tilde{\mu}^h a^h + \left(\sigma^2+\gamma^h\right)\frac{\partial a^h}{\partial z} + \int_{h<|y|}(e^\gamma - 1)\left(a^h(t,z+\gamma) - a^h(t,z)\mathbb{1}_{\{|y|\le 1\}}\right)\nu(dy)\\
\mathcal{G}_t^h a^h :=& \left(\sigma^2+\gamma^h\right)a^h + \int_{h<|y|}(e^\gamma-1)^2 a^h(t,z+\gamma)\nu(dy)\\
\mathcal{H}_t^h[a^h] :=& \inf_{|\pi|\le\bar{\Pi}^h}\left\{2\pi\mathcal{Q}_t^h a^h + \pi^2\mathcal{G}_t^h a^h\right\}
\end{aligned}\tag{9.4}$$

where $\tilde{\mu}^h$ is given in (5.6) with the new parameters and $\bar{\Pi}^h$ in (5.38) is the a priori bound corresponding to the new parameters. The main advantage when considering the value function a^h is due to the fact that the intensity measure is now finite and then we could write, for example,

$$\mathcal{B}_t^h a^h = \int_{h<|y|}\left(a^h(t,z+\gamma) - a^h(t,z)\right)\nu(dy) - \frac{\partial a^h}{\partial z}\int_{h<|y|\le 1}\gamma(t,z,y)\nu(dy)$$

It is relatively easy to check the regularity of the right hand side when z approaches the boundary of the truncated domain. Our aim now is to prove that the function a^h converges in some functional space to a and the optimal control $(\pi^h)^*$ also converges to π^*, the optimal control relative to the function a. For this we will suppose that

$$T < T^* \wedge T^{*,h} \tag{9.5}$$

in order to guarantee that both a and a^h are smooth solutions of their respective PIDEs.

Theorem 9.1. *Let Assumptions 5.1–6.1 hold true together with the condition* (9.5)*. Then*

i). $\left\|a - a^h\right\|_{2-\delta,H} \le M_\eta \varrho(h)$

ii). $\left\|\pi^* - (\pi^h)^*\right\|_{1-\delta,H} \le M_\eta \varrho(h)$

where $(\pi^h)^*$ *is the optimal control given in* (5.47) *relatively to* a^h*,* $M(\eta)$ *is a positive constant that depends on* $\eta > 0$ *but not on* h *and* $\varrho(h)$ *is the function introduced in Lemma 6.2:*

$$\varrho(h) := \int_{|y|\le h}\tau^2(y)\nu(dy) \to 0, \text{ when } h \to 0$$

Proof.
From Proposition 6.5 we know that there exists a sequence $\varphi^n \in C^{1+(1-\delta)/2,2+(1-\delta)}([0,T]\times\mathbb{R})$ that converges in $H^{2-\delta}([0,T]\times\mathbb{R})$ to a:

$$\|\varphi^n - a\|_{2-\delta,H} \to 0,\ n\to\infty$$

and the same holds true when considering the function a^h:

$$\left\|\varphi^{n,h} - a^h\right\|_{2-\delta,H} \to 0,\, n \to \infty$$

where $\varphi^{n,h}$ is another sequence that also belongs to $C^{1+(1-\delta)/2,2+(1-\delta)}([0,T]\times\mathbb{R})$. Since

$$\left\|a - a^h\right\|_{2-\delta,H} \leq \|\varphi^n - a\|_{2-\delta,H} + \left\|\varphi^n - \varphi^{n,h}\right\|_{2-\delta,H} + \left\|\varphi^{n,h} - a^h\right\|_{2-\delta,H} \tag{9.6}$$

we can concentrate on the middle term $\left\|\varphi^n - \varphi^{n,h}\right\|_{2-\delta,H}$. Let M denote a positive constant that does not depend on η or h whereas M_η is a positive constant that does depend on η but not on h. They may change from line to line.

If $\Delta_h^n := \varphi^n - \varphi^{n,h}$ then we can write

$$\begin{aligned}
&-\frac{\partial}{\partial t}\triangle_h^{n+1} + \mathcal{A}_t^h\triangle_h^{n+1} + \eta\triangle_h^{n+1} = \mathcal{B}_t^h\triangle_h^n + \mathcal{H}_t[\varphi^n] - \mathcal{H}_t^h[\varphi^{n,h}] \\
&-\frac{1}{2}\gamma^h\frac{\partial^2\varphi^{n+1}}{\partial^2 z} + \int_{|y|\leq h}\left(\varphi^n(t,z+\gamma) - \varphi^n(t,z) - \gamma\frac{\partial\varphi^n}{\partial z}\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy) \\
&\triangle_h^{n+1}(T,z) = 0
\end{aligned}$$

where we basically used the definition of these sequences stated in (6.3).

First we can estimate the second derivative appearing in the above PDE:

$$\begin{aligned}
&\left\|\frac{1}{2}\gamma^h\frac{\partial^2\varphi^{n+1}}{\partial^2 z} - \int_{|y|\leq h}\left(\varphi^n(t,z+\gamma) - \varphi^n(t,z) - \gamma\frac{\partial\varphi^n}{\partial z}\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)\right\|_\infty \\
\leq & M\sup_n \|\,\varphi^n\|_{2,H}\int_{|y|\leq h}\tau^2(y)\nu(dy) \leq M_\eta\varrho(h)
\end{aligned} \tag{9.7}$$

by using Lemma 6.9. We can proceed as in the proof of Lemma 6.2 to prove that for any $r, \epsilon > 0$

$$\left\|\mathcal{B}^h\triangle_h^n\right\|_\infty \leq M\left(\left(\varrho(r) + \varsigma(r)\epsilon^{1-\delta}\right)\|\triangle_h^n\|_{2-\delta,H} + \varsigma(r)\epsilon^{-1}\|\triangle_h^n\|_\infty\right) \tag{9.8}$$

where $\varsigma(r) \to \infty$ if the function τ is not integrable around zero. To conclude, we readapt the proof of Lemma 6.3 and obtain

$$\begin{aligned}
\left\|\mathcal{H}[\varphi^n] - \mathcal{H}^h[\varphi^{n,h}]\right\|_\infty \leq & M\left(\left\|\mathcal{Q}\varphi^n - \mathcal{Q}^h\varphi^{n,h}\right\|_\infty + \left\|\mathcal{G}\varphi^n - \mathcal{G}^h\varphi^{n,h}\right\|_\infty\right) \\
\leq & M\left(\|\triangle_h^n\|_{1+\lambda,H} + \varrho(h)\|\varphi^n\|_{2,H}\right)
\end{aligned}$$

for some $0 < \lambda < 1-\delta$, where we used the structure of the operators $\mathcal{Q},\mathcal{G}$ and $\mathcal{Q}^h,\mathcal{G}^h$. Apply then Proposition C.3 to deduce

$$\left\|\mathcal{H}[\varphi^n] - \mathcal{H}^h[\varphi^{n,h}]\right\|_\infty \leq M\left(\epsilon^{1-\delta-\lambda}\|\triangle_h^n\|_{2-\delta,H} + \epsilon^{-(1+\lambda)}\|\triangle_h^n\|_\infty + \varrho(h)\|\varphi^n\|_{2,H}\right)$$

We now use the above estimation and (9.7)–(9.8) as in the proof of Proposition 6.5 to obtain, for some η big enough and some $\beta \in (0,1)$

$$\|\triangle_h^{n+1}\|_{2-\delta,H,\eta} \leq \beta\|\triangle_h^n\|_{2-\delta,H,\eta} + M(\eta)\varrho(h)$$

where, as usual, $\| \|_{2-\delta,H,\eta} := \| \|_{2-\delta,H} + \eta \| \|_{\infty}$. It follows then

$$\left\|\triangle_h^{n+1}\right\|_{2-\delta,H,\eta} \leq M(\eta)\left(\beta^{n+1} + \varrho(h)\right)$$

Finally, we use the above estimate in (9.6) and let $n \to \infty$ to prove part i). For part ii) we have

$$\left\|\pi^* - (\pi^h)^*\right\|_{1-\delta,H} = \left\|\frac{\mathcal{Q}a}{\mathcal{G}a} - \frac{\mathcal{Q}^h a^h}{\mathcal{G}^h a^h}\right\|_{1-\delta,H}$$

Using the estimation given in i) and the structure of the above operators it is not difficult to prove that

$$\left\|\pi^* - (\pi^h)^*\right\|_{1-\delta,H} \leq M(\eta)\varrho(h)$$

Remark that the optimal control depends on the first derivative of a: this explains why we can just give an estimate on the Hölder norm of order $1-\delta$.

□

This result proves that we can approximate with arbitrary precision the value function a and the optimal control π^* by taking h small enough. This allows us to study the truncated PIDE for a^h instead of a.

9.3 The semi linear PIDE on a bounded domain

9.3.1 Existence and Uniqueness in the viscosity sense

From now on we assume fixed $h > 0$. We start by making some simplifications in the definition of the differential operators in (9.4): define first $\mu^h(t,z) := \mu(t,z) - \int_{h<|h|\leq 1} \gamma(t,z,y)\nu(dy)$, then redefine the operators as follows:

$$\begin{aligned}
\mathcal{A}_t^h a^h &\overset{redef}{:=} -\mu^h \frac{\partial a^h}{\partial z} - \frac{1}{2}\left(\sigma^2 + \gamma^h\right)\frac{\partial^2 a^h}{\partial z^2} \\
\mathcal{B}_t^h a^h &\overset{redef}{:=} \int_{h<|y|} \left(a^h(t,z+\gamma) - a^h(t,z)\right)\nu(dy) \\
\mathcal{Q}_t^h a^h &\overset{redef}{:=} \left(\mu^h + \frac{1}{2}\left(\sigma^2 + \gamma^h\right)\right)a^h + \left(\sigma^2 + \gamma^h\right)\frac{\partial a^h}{\partial z} + \int_{h<|y|} \left(e^\gamma - 1\right) a^h(t,z+\gamma)\nu(dy) \\
\mathcal{G}_t^h a^h &\overset{redef}{:=} \left(\sigma^2 + \gamma^h\right)a^h + \int_{h<|y|} (e^\gamma - 1)^2 a^h(t,z+\gamma)\nu(dy) \\
\mathcal{H}_t^h[a^h] &\overset{redef}{:=} \inf_{|\pi|\leq \bar{\Pi}^h} \left\{2\pi \mathcal{Q}_t^h a^h + \pi^2 \mathcal{G}_t^h a^h\right\}
\end{aligned} \tag{9.9}$$

The function $(t,z) \to a^h(t,z)e^{\eta t}$ verifies

$$0 = -\frac{\partial a^h}{\partial t} + \eta a^h + \mathcal{A}_t^h a^h - \mathcal{B}_t^h a^h - \mathcal{H}^h[a^h], \qquad a^h(T,z) = e^{\eta T} \tag{9.10}$$

From the above definition we remark that the non local operators are all of order zero, which will substantially simplify our computations. Remark however that at this point we cannot let anymore $h \to 0$, since the above coefficients and operators may not be properly defined if the Lévy measure ν is infinite. According to Theorem 9.1, we assume then that $h > 0$ is selected and fixed in order to have a suitably small error when replacing the function a with a^h.

Remark 9.2. *From now on, all the constants appearing in our estimates may depend on h.*

As we already anticipated in Section 9.1, let $\underline{Z} > 0$ and consider the problem (9.10) in the bounded parabolic domain $[0,T] \times [-\underline{Z}, \underline{Z}]$:

$$\begin{cases} 0 = -\dfrac{\partial a^{tr}}{\partial t} + \eta a^{tr} + \mathcal{A}_t^h a^{tr} - \mathcal{B}_t^h a^{tr} - \mathcal{H}^h[a^{tr}] & (t,z) \in [0,T) \times (-\underline{Z}, \underline{Z}) \\ a^{tr}(T,z) = e^{\eta T} & z \in (-\underline{Z}, \underline{Z}) \\ a^{tr}(t,z) = e^{\eta t} q(t,z) & (t,z) \in [0,T] \times (-\underline{Z}, \underline{Z})^c \end{cases} \tag{9.11}$$

Here the superscript tr stands for *truncated.* In the rest of the Chapter we will prove that the above PIDE has a unique solution which is smooth in the domain $[0,T] \times [-\underline{Z}, \underline{Z}]$ and give an estimate on the error between a^{tr} and a^h. One can then use Theorem 9.1 to deduce an estimate on the error between a^{tr} and a.

Note that, due to the non local component, the Dirichlet condition q has to be specified on the entire domain $[0,T] \times (-\underline{Z}, \underline{Z})^c$. We assume that q has the same regularity as the function a^h outside the domain[1]: $q \in C^{(1-\delta)/2+1,2+(1-\delta)}(\mho^T)$ and

$$\lim_{t \to T, |z| \geq \underline{Z}} q(t,z) = 1 \tag{9.12}$$

to ensure continuity at time $t = T$.

Lemma 9.3. *Under Assumptions 5.1–6.1 there exists a unique viscosity solution of PIDE* (9.11)*. If* (9.12) *also holds true then it is continuous and it assumes the boundary condition in the classical sense.*

Proof.
The result is a direct application[2] of Theorem 3 in Barles, Chasseigne, and Imbert (2008). Remark that we would have a unique viscosity solution even under less constraining Assumptions.

□

[1]This is actually not necessary and one could consider some other Dirichlet boundary condition that belongs to $C^{(1-\delta')/2+1,2+(1-\delta')}(\mho^T)$ for some $\delta' \in (0,1)$. We prefer to take q in the same space of a^h for sake of coherence.

[2]Theorem 3 in Barles, Chasseigne, and Imbert (2008) proves existence and uniqueness for elliptic differential problems. However, as the authors precise in Section 4.3 of the same paper, the proof can be easily adapted to the parabolic case.

9.3.2 Existence and Uniqueness in the viscosity sense

In this paragraph we will first prove that the unique viscosity solution of PIDE (9.11) belongs to $C^{2-\delta}(\Omega^T)$. This will allow us to remove the non linearity in the PIDE (9.11) and finally prove that this unique viscosity solution belongs to $C^{1+(1-\delta)/2,2+(1-\delta)}(\Omega^T)$.

For this, we will need to assume that the function q verifies a compatibility condition at time $t = T$. Condition (9.12) gives the continuity of q, but we also have to impose that the derivative w.r.t. t of the function a^{tr}, which can be computed from the equation and the initial conditions, is equal to the derivative of the Dirichlet boundary condition. In particular one must have

$$\frac{\partial a^{tr}}{\partial t}(T,z) = \eta e^{\eta T} - \mathcal{H}^h[a^{tr}(T,z)]$$

from (9.11). If we impose that

$$\frac{\partial e^{\eta t} q}{\partial t}(T,z) = \frac{\partial a^{tr}}{\partial t}(T,z), \quad z \in \{-\underline{Z}, \underline{Z}\}$$

then we should have

$$\frac{\partial q}{\partial t}(T,z) = \frac{\left(\mathcal{Q}^h a^{tr}(T,z)\right)^2}{\mathcal{G}^h a^{tr}(T,z)}, \quad z \in \{-\underline{Z}, \underline{Z}\}$$

By using the terminal condition of a^{tr} and the definition of $\mathcal{A}^h$ and $\mathcal{G}^h$, we will obtain

$$\begin{aligned}&\frac{\partial q}{\partial t}(T,z)\\ &= \frac{\left(\left(\mu + \frac{1}{2}\sigma^2\right) + \int_{\mathbb{R}} \left(\left(e^\gamma - 1 - \gamma \mathbb{1}_{\{|y|\le 1\}}\right) \mathbb{1}_{\{|h|<y\}} + \frac{1}{2}\gamma^2 \mathbb{1}_{\{|y|\le h\}}\right)\nu(dy)\right)^2}{\sigma^2 + \int_{h<|y|} \left((e^\gamma - 1)^2 \mathbb{1}_{\{|h|<y\}} + \gamma^2 \mathbb{1}_{\{|y|\le h\}}\right)\nu(dy)}\end{aligned} \tag{9.13}$$

where the coefficients are evaluated at the point (T,z) and $z \in \{-\underline{Z}, \underline{Z}\}$. We also need to assume some regularity on the function γ and σ:

$$\text{There exists } \delta' \in (0,1) \text{ such that } \partial_z\sigma, \partial_z\gamma \in H^{\delta'}([0,T]\times\mathbb{R}), \text{ for any } t, y. \tag{9.14}$$

We will explain later where these assumptions are needed. Define then

$$\kappa := \min(1-\delta,\, \delta') \tag{9.15}$$

As we already did many times (see Sections 6.2.2–7.4.1), we prove that the unique viscosity solution of PIDE (9.11) belongs to $C^{1+(1-\delta)/2,2+(1-\delta)}(\Omega^T)$ by introducing a sequence of smooth function φ^n and proving that it converges to the unique viscosity solution in the above mentioned Hölder space. Let us start with some $\tilde{\varphi}_0 \in H^{\kappa/2+1,2+\kappa}([0,T]\times\mathbb{R})$ and consider $(\tilde{\varphi}^n)_{n\in\mathbb{N}}$ defined by

$$\tilde{\varphi}^{n+1}(t,z) := \begin{cases} \varphi^{n+1}(t,z) & (t,z) \in [0,T]\times[-\underline{Z},\, \underline{Z}] \\ e^{\eta t} q(t,z) & (t,z) \in [0,T]\times[-\underline{Z},\, \underline{Z}]^c \end{cases} \tag{9.16}$$

where φ^{n+1} verifies

$$\begin{cases} -\dfrac{\partial \varphi^{n+1}}{\partial t} + \eta\varphi^{n+1} + \mathcal{A}^h\varphi^{n+1} = \mathcal{B}^h_t\tilde{\varphi}^n - \mathcal{H}^h[\tilde{\varphi}^n] & (t,z)\in[0,T)\times(-\underline{Z},\,\underline{Z}) \\ \varphi^{n+1}(T,z) = e^{\eta T} & z\in(-\underline{Z},\,\underline{Z}) \\ \varphi^{n+1}(t,z) = e^{\eta t}q(t,z) & (t,z)\in[0,T]\times\{-\underline{Z},\,\underline{Z}\} \end{cases} \tag{9.17}$$

Remark 9.4. *From now on we will adopt the following notation: for any function $\psi : \Omega^T \to \mathbb{R}$, we denote with $\tilde{\psi}$ the extension of this map on the domain $[0,T]\times\mathbb{R}$ with the Dirichlet boundary condition:*

$$\tilde{\psi}^{n+1}(t,z) := \begin{cases} \psi^{n+1}(t,z) & (t,z)\in[0,T]\times[-\underline{Z},\,\underline{Z}] \\ e^{\eta t}q(t,z) & (t,z)\in[0,T]\times[-\underline{Z},\,\underline{Z}]^c \end{cases}$$

Our first objective is to prove that the sequence in (9.16) is well defined. For this we need a preliminary result on the properties on the operators $\mathcal{B}^h$ and $\mathcal{H}^h$:

Lemma 9.5. *Let Assumptions 5.1–6.1 hold true. There exists some positive constant $M > 0$ such that for all $\psi \in C^{\kappa/2+1,2+\kappa}(\Omega^T)$ we have*

$$\begin{aligned} \left\|\mathcal{B}^h\tilde{\varphi}^n\right\|_{\infty,\Omega^T} &\le M\left(\|\varphi^n\|_{\infty,\Omega^T} + e^{\eta T}\|q\|_{\infty,\mho^T}\right) \\ \left\|\mathcal{H}^h[\tilde{\varphi}^n]\right\|_{\infty,\Omega^T} &\le M\left(\|\varphi^n\|_{1,H,\Omega^T} + e^{\eta T}\|q\|_{\infty,\mho^T}\right) \end{aligned}$$

and

$$\begin{aligned} \left\|\mathcal{B}^h\tilde{\varphi}^n\right\|_{\kappa/2,\kappa,\Omega^T} &\le M\left(\|\varphi^n\|_{\kappa/2,\kappa,\Omega^T} + e^{\eta T}\|q\|_{\kappa/2,\kappa,\mho^T}\right) \\ \left\|\mathcal{H}^h[\tilde{\varphi}^n]\right\|_{\kappa/2,\kappa,\Omega^T} + \left\|\mathcal{Q}^h\tilde{\varphi}^n\right\|_{\kappa/2,\kappa,\Omega^T} &+ \left\|\mathcal{G}^h\tilde{\varphi}^n\right\|_{\kappa/2,\kappa,\Omega^T} \le \\ &M\left(\|\varphi^n\|_{(\kappa+1)/2,1+\kappa,\Omega^T} + e^{\eta T}\|q\|_{\kappa/2,\kappa,\mho^T}\right) \end{aligned}$$

where $\tilde{\psi}$ denotes the extension of ψ according to the Remark 9.4. The constant M does not depend on η.

Proof.
The proof can be completed with the type of computations we made for Lemmas 6.2–6.3, by using the fact that outside the domain Ω^T, the map $\tilde{\psi}$ is equal to $e^{\eta t}q$. The only difference arises when one has to estimate $\langle\mathcal{B}^h\tilde{\varphi}^n\rangle^{(\kappa)}_{z,\Omega^T}$ (or $\langle\mathcal{H}^h\tilde{\varphi}^n\rangle^{(\kappa)}_{z,\Omega^T}$ and so on), since one has to take into account the particular form of $\tilde{\psi}$. We detail this computation for $\langle\mathcal{B}^h\tilde{\varphi}^n\rangle^{(\kappa)}_{z,\Omega^T}$: we first have

$$\langle\mathcal{B}^h\tilde{\varphi}^n\rangle^{(\kappa)}_{z,\Omega^T} \le \int_{h<|y|}\langle\tilde{\varphi}^n(t,z+\gamma(t,z,y)) - \tilde{\varphi}^n(t,z)\rangle^{(\kappa)}_{z,\Omega^T}\nu(dy)$$

Fix $z, z' \in [-\underline{Z}, \underline{Z}]$: it follows

$$\begin{aligned}&\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t,z) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y)) + \tilde{\varphi}^n(t,z')\right| \\ \leq &\langle \varphi^n \rangle^{(\kappa)}_{z,\Omega^T} |z-z'|^\kappa + \left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right|\end{aligned}$$

If $z+\gamma(t,z,y), z'+\gamma(t,z',y) \in (-\underline{Z}, \underline{Z})$ then

$$\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right| \leq M(1+\rho(y)^\kappa)|z-z'|^\kappa \langle \varphi^n \rangle^{(\kappa)}_{z,\Omega^T}$$

If instead $z+\gamma(t,z,y) \in (-\underline{Z}, \underline{Z})$ but $z'+\gamma(t,z',y) \notin (-\underline{Z}, \underline{Z})$, then we can find some $\lambda \in [0,1]$ such that $z_\lambda + \gamma(t, z_\lambda, y) \in [0,T] \times \{\underline{Z}, \underline{Z}\}$, where $z_\lambda := z + \lambda(z'-z)$. This is due to the continuity of the map $z \to \gamma(t,z,y)$. We deduce then

$$\begin{aligned}&\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right| \\ \leq &\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z_\lambda+\gamma(t,z_\lambda,y))\right| \\ + &\left|\tilde{\varphi}^n(t, z_\lambda+\gamma(t,z_\lambda,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right| \\ \leq &M(1+\rho(y)^\kappa)|z-z_\lambda|^\kappa \left(\langle \varphi^n \rangle^{(\kappa)}_{z,\Omega^T} + \langle q \rangle^{(\kappa)}_{\mho^T}\right) \\ \leq &M(1+\rho(y)^\kappa)|z-z'|^\kappa \left(\langle \varphi^n \rangle^{(\kappa)}_{z,\Omega^T} + e^{\eta T}\langle q \rangle^{(\kappa)}_{\mho^T}\right)\end{aligned}$$

Finally, if both are outside the interval $(-\underline{Z}, \underline{Z})$ then, we can distinguish two cases

- both $z+\gamma(t,z,y)$ and $z'+\gamma(t,z',y)$ are on the same side w.r.t. $(-\underline{Z}, \underline{Z})$, and then

$$\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right| \leq M(1+\rho(y)^\kappa)e^{\eta T}|z-z'|^\kappa \langle q \rangle^{(\kappa)}_{\mho^T}$$

- one is bigger than $\underline{Z}$ whereas the other is smaller that $-\underline{Z}$: in this case the triangular inequality yields

$$\begin{aligned}&\left|\tilde{\varphi}^n(t, z+\gamma(t,z,y)) - \tilde{\varphi}^n(t, z'+\gamma(t,z',y))\right| \\ \leq &M(1+\rho(y)^\kappa)|z-z'|^\kappa \left(\langle \varphi^n \rangle^{(\kappa)}_{z,\Omega^T} + e^{\eta T}\langle q \rangle^{(\kappa)}_{\mho^T}\right)\end{aligned}$$

Adding up all the above estimates we obtain

$$\langle \mathcal{B}^h \tilde{\varphi}^n \rangle^{(\kappa)}_{z,\Omega^T} \leq M\left(\|\varphi^n\|_{\kappa/2,\kappa,\Omega^T} + e^{\eta T}\|q\|_{\kappa/2,\kappa,\mho^T}\right)$$

The same argument can be used to estimate $\langle \mathcal{B}^h \tilde{\varphi}^n \rangle^{((\kappa)/2)}_{t,\Omega^T}$.

□

The above Lemma allows us to prove that the sequence in (9.16) is well defined. Assume that the compatibility conditions (9.12)–(9.13) hold true. By recurrence, if $\varphi^n \in C^{\kappa/2+1,2+\kappa}(\Omega^T)$ is well defined then Lemma 9.5 yields

$$\mathcal{B}^h \tilde{\varphi}^n + \mathcal{H}^h[\tilde{\varphi}^n] \in C^{\kappa/2,\kappa}(\Omega^T)$$

We can use Theorem 5.2, Chapter IV in Ladyzenskaja et al. (1967) to deduce that PIDE (9.17) has a unique solution which belongs to $C^{\kappa/2+1,2+\kappa}(\Omega^T)$. This proves that φ^{n+1} is also well defined. Furthermore

$$\left\|\varphi^{n+1}\right\|_{\frac{\kappa}{2}+1,2+\kappa,\Omega^T} \leq M\left(e^{\eta T} + \left\|\mathcal{B}^h\tilde{\varphi}^n + \mathcal{H}^h[\tilde{\varphi}^n]\right\|_{\kappa/2,\kappa,\Omega^T} + \left\|e^{\eta t}q\right\|_{\frac{\kappa}{2},2+\kappa,\mho^T}\right)$$

Remark 9.6. *The function $\tilde{\varphi}^n$ is Lipschitz continuous at the boundary $[0,T]\times\{-\underline{Z},\underline{Z}\}$: since $\varphi^n \in C^{\kappa/2+1,2+\kappa}(\Omega^T)$ then $\left\|D_z\varphi^{n+1}\right\|_{\infty,\Omega^T}$ is finite. But $\left\|D_z(e^{\eta t}q)\right\|_{\infty,\mho^T}$ is also finite for any $t \leq T$. This means that the function $z \to \tilde{\varphi}^{n+1}(t,z)$ is continuous and has bounded left and right derivatives on the boundary $[0,T]\times\{-\underline{Z},\underline{Z}\}$, which proves that $\tilde{\varphi}^{n+1}$ is Lipschitz continuous on $[0,T]\times\{-\underline{Z},\underline{Z}\}$.*

The analog of Proposition 6.5 is this case it the following:

Proposition 9.7. *Let Assumptions 5.1–6.1 hold true together with the compatibility conditions* (9.12)–(9.13). *Assume also that condition* (9.14) *stands in force. For η big enough, the sequence φ^n defined in* (9.16) *converges to some $\varphi^* \in H^{2-\delta}(\Omega^T)$. Furthermore for any $\upsilon \in (0,1)$ there exists some positive constant M_υ which depends on υ such that*

$$\left|\varphi^*(t,z) - \varphi^*(t',z)\right| \leq M_\upsilon|t-t'|^\upsilon, \quad \textit{for any } t,t',z$$

and

$$\left|D_z\varphi^*(t,z) - D_z\varphi^*(t',z)\right| \leq M_\upsilon|t-t'|^{\upsilon/2}, \quad \textit{for any } t,t',z$$

Proof.
If $\triangle^{n+1} := \varphi^{n+1} - \varphi^n$ and $\tilde{\triangle}^{n+1} := \tilde{\varphi}^{n+1} - \tilde{\varphi}^n$ then

$$\begin{cases} -\dfrac{\partial\triangle^{n+1}}{\partial t} + \eta\triangle^{n+1} + \mathcal{A}^h\triangle^{n+1} = \mathcal{B}^h_t\tilde{\triangle}^n + \mathcal{H}^h[\tilde{\varphi}^n] - \mathcal{H}^h[\tilde{\varphi}^{n-1}] & (t,z)\in[0,T)\times(-\underline{Z},\,\underline{Z}) \\ \varphi^{n+1}(T,z) = 0 & z\in(-\underline{Z},\,\underline{Z}) \\ \varphi^{n+1}(t,z) = 0 & (t,z)\in[0,T]\times\{-\underline{Z},\,\underline{Z}\} \end{cases}$$

The unique solution of the above linear PDE is explicitly given in Ladyzenskaja et al. (1967), Section §16,

$$\triangle^{n+1}(t,z) = \int_t^T e^{-\eta(s-t)}ds\int_{-\underline{Z}}^{\underline{Z}} G(T-t,z,T-s,\xi)R^n(y,\xi)d\xi \tag{9.18}$$

where $R^n := \mathcal{B}^h\tilde{\triangle}^n + \mathcal{H}^h[\tilde{\varphi}^n] - \mathcal{H}^h[\tilde{\varphi}^{n-1}]$ and G the Green's function given in Theorem 16.3 of Ladyzenskaja et al. (1967), Chapter IV, §16, which verifies

i). For $2r+s \leq 2$,

$$\left|D_t^iD_y^jG(t,z,s,\xi)\right| \leq m_1(t-s)^{-\frac{1+2i+j}{2}}\exp\left(-m_2\frac{|z-\xi|^2}{t-s}\right)$$

ii). For $2i+j=2$ and any $\iota \in (0,1)$

$$\left|D_t^i D_z^j G(t,z,s,\xi) - D_t^i D_z^j G(t,z',s,\xi)\right| \leq m_1 |z-z'|^\iota (t-s)^{-\frac{3+\iota}{2}} \exp\left(-m_2 \frac{|z''-\xi|^2}{t-s}\right)$$

Here z'' is the closest point to ξ which belongs to the segment $\overrightarrow{zz'}$.

iii). For $2i+j=1,2$ and $s<t'<t$.

$$\begin{aligned}\left|D_t^i D_z^j G(t,z,s,\xi) - D_t^i D_z^j G(t',z,s,\xi)\right| \leq & m_1 |t'-s|^{-\frac{3+\iota}{2}} |t-t'|^{\frac{2-2i-j+\iota}{2}} \\ & \exp\left(-m_2 \frac{|w-\xi|^2}{s-t}\right)\end{aligned}$$

for $0 \leq s < t$ and positive constants m_1, m_2.

We can follow the scheme of the proof of Proposition 6.5 to deduce that $(\varphi^n)_n$ is a Cauchy sequence in $H^{2-\delta}(\Omega^T)$. There exists then some $\varphi^* \in H^{2-\delta}(\Omega^T)$ such that $\varphi_n \to \varphi^*$.

Let us now prove the regularity of φ^* w.r.t. t. From (9.18) we can write, for $t'<t$,

$$\begin{aligned}& \left|\triangle^{n+1}(t,z) - \triangle^{n+1}(t',z)\right| \\ \leq & \|R^n\|_{\infty,\Omega^T} \int_t^T e^{-\eta(s-t)} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t,z,T-s,\xi) - G(T-t',z,T-s,\xi)\right| d\xi ds \\ & + \|R^n\|_{\infty,\Omega^T} \int_{t'}^t e^{-\eta(s-t')} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t',z,T-s,\xi)\right| d\xi ds \\ & + \|R^n\|_{\infty,\Omega^T} \left|e^{\eta(t'-t)} - 1\right| \int_t^T e^{-\eta(s-t)} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t',z,T-s,\xi)\right| d\xi ds\end{aligned}$$

By using Lemma 6.4 we have

$$\int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t,z,T-s,\xi) - G(T-t',z,T-s,\xi)\right| d\xi \leq M$$

but also

$$\int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t,z,T-s,\xi) - G(T-t',z,T-s,\xi)\right| d\xi \leq M|t-t'||s-t|^{-1}$$

so that

$$\int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t,z,T-s,\xi) - G(T-t',z,T-s,\xi)\right| d\xi \leq M_\upsilon |t-t'|^\upsilon |s-t|^{-\upsilon}$$

The first term in the right hand side is then estimated with

$$\int_t^T e^{-\eta(s-t)} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t,z,T-s,\xi) - G(T-t',z,T-s,\xi)\right| d\xi ds \leq M_\upsilon |t-t'|^\upsilon$$

For the second it is straightforward

$$\int_{t'}^{t} e^{-\eta(s-t')} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t', z, T-s, \xi)\right| d\xi ds \leq M|t-t'|$$

whereas for the third one, there is some $\lambda \in [0,1]$ such that

$$\begin{aligned}&\left|e^{\eta(t'-t)} - 1\right| \int_{t}^{T} e^{-\eta(s-t)} \int_{-\underline{Z}}^{\underline{Z}} \left|G(T-t', z, T-s, \xi)\right| d\xi ds \\ \leq & M|t-t'|\eta e^{\lambda\eta(t'-t)} \int_{t}^{T} e^{-\eta(s-t)} ds \leq M|t-t'|\end{aligned}$$

since $t' < t$. We conclude then

$$\left|\triangle^{n+1}(t,z) - \triangle^{n+1}(t',z)\right| \leq M_\upsilon \left\|R^n\right\|_{\infty,\Omega^T} |t-t'|^\upsilon$$

for some positive M_υ which depends on υ. We can finally readapt the argument of the proof of Proposition 6.5 to prove that the sequence of functions $t \to \varphi^n(t,.)$ is also a Cauchy sequence in the Hölder space $H^\upsilon\left([0,T]\right)$, and then its limit, φ^*, has to belong to this space. In particular

$$\left|\varphi^*(t,z) - \varphi^*(t',z)\right| \leq M_\upsilon |t-t'|^\upsilon, \quad \text{for any } t, t', z$$

With the same type of computations we prove that the sequence of functions $t \to D_z\varphi^n(t,z)$ is a Cauchy sequence in the Hölder space $H^{\upsilon/2}\left([0,T]\right)$ for any $\upsilon \in (0,1)$. It follows that also $D_z\varphi^*$ has to belongs to this space.

□

Corollary 9.8. *Let Assumptions 5.1–6.1 hold true together with the compatibility conditions* (9.12)–(9.13)*. Let* φ^* *be the limit in* $H^{2-\delta}(\Omega^T)$ *of the sequence* φ_n*. Its extension to* $[0,T]\times\mathbb{R}$*,* $\tilde{\varphi}^*$ *is the unique viscosity solution of PIDE* (9.11)*:*

$$\tilde{\varphi}^*(t,z) = a^{tr}(t,z) \tag{9.19}$$

Proof. We define

$$\psi(t,z) := \mathbb{E}\left[\int_t^{\hat{\beta}^{t,z}} e^{-\eta(s-t)} \left(\mathcal{B}_t^h\tilde{\varphi}^n - \mathcal{H}^h[\tilde{\varphi}^n]\right)(s, \hat{Z}_s^{t,z}) + e^{\eta\beta^{t,z}} q\left(\beta^{t,z}, \hat{Z}_{\hat{\beta}^{t,z}}^{t,z}\right)\right]$$

where $\hat{\beta}^{t,z}$ is the hitting time of the boundary:

$$\hat{\beta}^{t,z} := T \wedge \inf\left\{s > t : \left|\hat{Z}_s^{t,z}\right| \geq \underline{Z}\right\} \tag{9.20}$$

and

$$d\hat{Z}_s^{t,z} := \mu^h(s, \hat{Z}_s^{t,z})ds + \left(\sqrt{\sigma^2+\gamma^h}\right)(s, \hat{Z}_s^{t,z})dW_s, \quad \hat{Z}_t^{t,z} = z$$

Remark that from the Markov property of the process $\hat{Z}$ we have

$$\hat{\beta}^{t,z} = \hat{\beta}^{\theta, \hat{Z}_\theta^{t,z}} \text{ for any stopping time } t \leq \theta \leq \hat{\beta}^{t,z}$$

In particular, according to Remark 9.4, we can write

$$\tilde{\psi}(t,z) := \mathbb{E}\left[\int_t^\theta e^{-\eta(s-t)}\left(\mathcal{B}_t^h\tilde{\varphi}^n - \mathcal{H}^h[\tilde{\varphi}^n]\right)(s,\tilde{Z}_s^{t,z})ds + e^{\eta\theta}\tilde{\psi}\left(\theta, Z_\theta^{t,z}\right)\right]$$

As in Step 3 of the proof of Theorem 6.8, we deduce that ψ is the unique viscosity solution[3] of (9.17). Since φ^{n+1} is the unique smooth solution of (9.17), we deduce $\psi = \varphi^{n+1}$, and then

$$\tilde{\varphi}^{n+1}(t,z) := \mathbb{E}\left[\int_t^\theta e^{-\eta(s-t)}\left(\mathcal{B}_t^h\tilde{\varphi}^n - \mathcal{H}^h[\tilde{\varphi}^n]\right)(s,\tilde{Z}_s^{t,z})ds + e^{\eta\theta}\tilde{\varphi}^{n+1}\left(\theta, Z_\theta^{t,z}\right)\right]$$

Since $\varphi^n \to \varphi^*$ in $H^{2-\delta}(\Omega^T)$ (Proposition 9.7) and the operators $\mathcal{B}^h$ and $\mathcal{H}^h$ are bounded and continuous, we can let $n \to \infty$ and obtain

$$\tilde{\varphi}^*(t,z) := \mathbb{E}\left[\int_t^\theta e^{-\eta(s-t)}\left(\mathcal{B}_t^h\tilde{\varphi}^* - \mathcal{H}^h[\tilde{\varphi}^*]\right)(s,\hat{Z}_s^{t,z}) + e^{\eta\theta}\tilde{\varphi}^*\left(\theta, Z_\theta^{t,z}\right)\right]$$

For $t = T$ we have $\tilde{\varphi}^*(T,z) = \varphi^*(T,z) = e^{\eta T}$, by construction of the sequence φ^n, which converges to φ^*. If instead $|z| \geq \underline{Z}$, i.e., we are outside the domain, then trivially $\beta^{t,z} = t$, from which we deduce $\tilde{\varphi}(t,z) = e^{\eta t}q(t,z)$. Again by following the Step 3 of the proof of Theorem 6.8 we deduce that $\tilde{\varphi}^*$ is a viscosity solution of PIDE (9.11): form the uniqueness of this solution (Lemma 9.3) we deduce

$$\tilde{\varphi}^*(t,z) = a^{tr}(t,z)$$

□

We are now able to prove the main result of this chapter:

Theorem 9.9. *Let Assumptions 5.1–6.1 stand in force and assume also that the functions σ and γ verify* (9.14)*. If the Dirichlet boundary condition q belongs to $C^{(1-\delta)/2+1,2+(1-\delta)}(\mho^T)$ and verifies the compatibility conditions* (9.12)*–*(9.13)*, then the PIDE* (9.11) *has a unique solution a^{tr} which verifies*

$$a^{tr}\Big|_{\Omega^T} \in C^{\kappa/2+1,\kappa+2}(\Omega^T)$$

where κ is given in (9.15)*. Trivially*

$$a^{tr}\Big|_{\mho^T} \in C^{\kappa/2+1,\kappa+2}(\mho^T)$$

since q belongs to this space, and $z \to a^{tr}(t,z)$ is Lipschitz continuous for all $t \leq T$.

Proof.
The only thing we need to prove is that $a^{tr}\Big|_{\Omega^T} \in C^{\kappa/2+1,\kappa+2}(\Omega^T)$. Remark that

$$\mathcal{H}^h[a^{tr}] = 2\pi^{tr}\mathcal{Q}^h a^{tr} + (\pi^{tr})^2\mathcal{G}a^{tr}$$

[3]According to the definition of viscosity solution, one replaces the function ψ with a twice continuously differentiable test function, for which we can apply Itô's formula. Then we do not need to take care of the non smoothness of $\tilde{\psi}$ at the boundary $(0,T) \times \{-\underline{Z}, \underline{Z}\}$. In other words, the Feynman-Kac formula holds true and gives a probabilistic representation of the function $\tilde{\varphi}^{n+1}$.

where

$$\pi^{tr} := -\bar{\Pi}^h \vee -\frac{\mathcal{Q}^h a^{tr}}{\mathcal{G} a^{tr}} \wedge \bar{\Pi}^h \tag{9.21}$$

from the definition of $\mathcal{H}^h[a^{tr}]$ in (9.9). Proposition 9.7, Corollary 9.8 and the regularity of q imply that $\pi^{tr} \in C^{(1-\delta)/2,1-\delta}([0,T] \times \mathbb{R})$ and in particular, from (9.15), we have $\pi^{tr} \in C^{\kappa/2,\kappa}([0,T] \times \mathbb{R})$. This allows us to linearize the PIDE (9.11): as already seen many times, we define the map $\Xi_\eta(\psi)$ as follows

$$\begin{cases} -\dfrac{\partial \Xi_\eta(\psi)}{\partial t} + \eta \Xi_\eta(\psi) + \mathcal{A}^h \Xi_\eta(\psi) = \mathcal{B}_t^h \tilde{\psi} + 2\pi^{tr} \mathcal{Q}^h \tilde{\psi} + (\pi^{tr})^2 \mathcal{G}^h \tilde{\psi} & [0,T) \times (-\underline{Z}, \underline{Z}) \\ \Xi_\eta(\psi)(T,z) = e^{\eta T} & z \in (-\underline{Z}, \underline{Z}) \\ \Xi_\eta(\psi)(t,z) = e^{\eta t} q(t,z) & [0,T] \times \{-\underline{Z}, \underline{Z}\} \end{cases}$$

for $\psi \in C^{\kappa/2+1,2+\kappa}(\Omega^T)$. Theorem 5.2, Chapter IV in Ladyzenskaja et al. (1967) guarantees that this map is well defined in $C^{\kappa/2+1,2+\kappa}(\Omega^T)$. We can now proceed as in Step 2 of the proof of Theorem 6.8 to deduce that, for η big enough, the map Ξ_η is a contraction in n $C^{\kappa/2+1,2+\kappa}(\Omega^T)$[4]. If call ψ^* its unique fixed then, according to Remark 9.4, it verifies

$$\begin{cases} -\dfrac{\partial \tilde{\psi}^*}{\partial t} + \eta \tilde{\psi}^* + \mathcal{A}^h \tilde{\psi}^* = \mathcal{B}_t^h \tilde{\psi}^* + 2\pi^{tr} \mathcal{Q}^h \tilde{\psi}^* + (\pi^{tr})^2 \mathcal{G}^h \tilde{\psi}^* & (t,z) \in [0,T) \times (-\underline{Z}, \underline{Z}) \\ \tilde{\psi}^*(T,z) = e^{\eta T} & z \in (-\underline{Z}, \underline{Z}) \\ \tilde{\psi}^*(t,z) = e^{\eta t} q(t,z) & (t,z) \in [0,T] \times \{-\underline{Z}, \underline{Z}\} \end{cases}$$

Since a^{tr} is the unique viscosity solution of the above PIDE (Lemma 9.3) we deduce that $\tilde{\psi}^* = a^{tr}$, i.e. the restriction to Ω^T of a^{tr} belongs to $C^{\kappa/2+1,2+\kappa}(\Omega^T)$. This proves that PIDE (9.11) has a unique solution, whose restriction to Ω^T belongs to $C^{\kappa/2+1,2+\kappa}(\Omega^T)$. It is clear that this solution a^{tr} is Lipschitz continuous at the boundary: we already know that it was continuous (Lemma 9.3), and since the derivative w.r.t. z of a^{tr} and q are bounded, we deduce that a^{tr} is Lipschitz continuous at the boundary. Clearly $a^{tr} = q$ outside the boundary.

□

9.4 Estimate on the truncated PIDE

Theorem 9.9 tells us that there exists a unique solution of PIDE (9.11), which is smooth in Ω^T. For the remainder of this section we will concentrate on the error estimation between the function a^h introduced in Section 9.2 and the function a^{tr}.

[4] This is done by writing the Feynman-Kac formula for $\Xi_\eta(\psi)$, and we explained in the proof of Corollary 9.8 how it can be done.

For this, we will prove that the function a^{tr} is the value function of a stochastic optimization problem. Let us start with the definition of a^h:

$$x^2 a^h(t,z) := \inf_{\theta \in \mathcal{X}(t,z,x)} \mathbb{E}\left[\left(X_T^{t,z,x,\theta}\right)^2\right]$$

where X is the wealth process given in (5.9) corresponding to the new parameters in (9.2), and $\mathcal{X}(t,z,x)$ is the set of admissible strategies in (5.10). The function a^h belongs to $C^{\kappa/2+1,2+\kappa}([0,T]\times\mathbb{R})$ and the optimal control $\pi^{*,h}$ of the above stochastic problem is given in (5.47). In particular

$$x^2 a^h(t,z) := \inf_{\theta \in \mathcal{E}^T(t,z,x)} \mathbb{E}\left[\left(X_T^{t,z,x,\theta}\right)^2\right]$$

where

$$\mathcal{E}^\alpha(t,z,x) := \left\{ \theta_s := \theta\left(s, Z_{s-}^{t,z}, X_{s-}^{t,z,x,\theta}\right) \;\middle|\; \begin{array}{l} \theta(t,z,x) := e^{-z}\pi(t,z)x \quad \text{and} \\ \\ \pi \in C^{(1+\kappa)/2,1+\kappa}([0,T]\times\mathbb{R}),\ t \le s \le \alpha \end{array} \right\}$$

since the optimal control $\pi^{*,h}$ belongs to the above subset of admissible strategies. Furthermore for any stopping times $t \le \alpha \le T$ we have

$$x^2 a^h(t,z) := \inf_{\theta \in \mathcal{E}^\alpha(t,z,x)} \mathbb{E}\left[\left(X_\alpha^{t,z,x,\theta}\right)^2 a^h\left(\alpha, Z_\alpha^{t,z}\right)\right] \tag{9.22}$$

Remark that this is the dynamic programming principle when we restrict ourselves to the Markovian strategies in $\mathcal{E}(t,z,x)$. The following result proves that a^{tr} can also be represented as in (9.22).

Lemma 9.10. *Let the Assumptions of Theorem 9.9 hold true. Then*

$$x^2 a^{tr}(t,z) = \inf_{\theta \in \mathcal{E}^\beta(t,z,x)} \mathbb{E}\left[\left(X_{\beta^{t,z}}^{t,z,x,\theta}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right]$$

where

$$\beta^{t,z} := T \wedge \inf\left\{s > t;\ \left|Z_s^{t,z}\right| \ge \underline{Z}\right\} \tag{9.23}$$

and Z *is the process given in* (5.5) *when using the parameters in*(9.2).

Proof.
Let us define the Markovian strategy

$$\theta_s^{tr} := \pi^{tr}(s, Z_{s-}^{t,z}) e^{-Z_{s-}^{t,z}} X_{s-}^{t,z,x,\theta^{tr}}, \quad X_s^{t,z,x,\theta^{tr}} := x + \int_t^s \theta_{s-}^{tr} de^{-Z_{s-}^{t,z}} \tag{9.24}$$

where π^{tr} is the optimal control given in (9.21) and

$$w(t,z,x) := \mathbb{E}\left[\left(x + \int_t^{\beta^{t,z}} \theta_{r-}^{tr} de^{-Z_{s-}^{t,z}}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right] := x^2 \check{\varphi}(t,z)$$

The flow $(t,z) \to Z_{\cdot}^{t,z}$ is continuous in the topology of uniform convergence on compacts (Theorem 37, Chapter V, Protter (2004)), so then $(t,z) \to \beta^{t,z}$ is at least lower semi-continuous (Theorem 38, Chapter V, Protter (2004)). It follows that $(t,z,x) \to w(t,z,x)$ is continuous w.r.t. x and at least measurable w.r.t (t,z). Let now $\alpha \in \left[t, \beta^{t,z}\right]$ be a stopping time: it follows

$$\begin{aligned}
w(t,z,x) =& \mathbb{E}\left[\mathbb{E}^{\mathcal{F}_\alpha}\left[\left(x + \int_t^\alpha \theta_{r-}^{tr} de^{-Z_{s-}^{t,z}} + \int_\alpha^{\beta^{t,z}} \theta_{r-}^{tr} de^{-Z_{s-}^{t,z}}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right]\right] \\
=& \mathbb{E}\left[\mathbb{E}^{\mathcal{F}_\alpha}\left[\left(X_\alpha^{t,z,x,\theta^{tr}} + \int_\alpha^{\beta^{\alpha, Z_\alpha^{t,z}}} \theta_{r-}^{tr} de^{-Z_{s-}^{t,z}}\right)^2 q\left(\beta^{\alpha, Z_\alpha^{t,z}}, Z_{\beta^{\alpha, Z_\alpha^{t,z}}}^{t,z}\right)\right]\right] \\
=& \mathbb{E}\left[w\left(\alpha, Z_\alpha^{t,z}, X_\alpha^{t,z,x,\theta}\right)\right]
\end{aligned}$$

since the strategy θ^{tr} is Markovian. As in the proof of Theorem 6.8 we deduce that $\check{\varphi}$ is a viscosity solution of PIDE (9.11) for $\eta = 0$, and then, from the uniqueness, $\check{\varphi} = a^{tr}$. It follows then

$$\begin{aligned}
x^2 a^{tr}(t,z) =& \mathbb{E}\left[\left(1 + \int_t^{\beta^{t,z}} \theta_{r-}^{tr} de^{-Z_{s-}^{t,z}}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right] \\
\geq& \inf_{\theta \in \mathcal{E}^\beta(t,z,x)} \mathbb{E}\left[\left(X_{\beta^{t,z}}^{t,z,x,\theta}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right]
\end{aligned}$$

If now

$$x^2 a^{tr}(t,z) > \inf_{\theta \in \mathcal{E}^\beta(t,z,x)} \mathbb{E}\left[\left(X_{\beta^{t,z}}^{t,z,x,\theta}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right]$$

then there would exist some $\hat{\pi} \in C^{(1+\kappa)/2, 1+\kappa}([0,T] \times \mathbb{R})$ such that

$$x^2 a^{tr}(t,z) > \mathbb{E}\left[\left(\hat{X}_{\beta^{t,z}}^{t,z,x}\right)^2 q\left(\beta^{t,z}, Z_{\beta^{t,z}}^{t,z}\right)\right] := x^2 \hat{\psi}(t,z)$$

where $d\hat{X}_s^{t,z,x} := \hat{X}_{s-}^{t,z,x} \hat{\pi}(s, Z_{s-}^{t,z}) e^{-Z_{s-}^{t,z}} de^{Z_s^{t,z}}$, $\hat{X}_t^{t,z,x} = x$. It is not complicated to prove that $\hat{\psi}$ is the unique solution of

$$\begin{cases}
-\dfrac{\partial \hat{\psi}}{\partial t} + \mathcal{A}^h \hat{\psi} = \mathcal{B}_t^h \hat{\psi} + 2\hat{\pi} \mathcal{Q}^h \hat{\psi} + \hat{\pi}^2 \mathcal{G}^h \hat{\psi} & (t,z) \in [0,T) \times (-\underline{Z}, \underline{Z}) \\
\hat{\psi}(T,z) = 1 & z \in (-\underline{Z}, \underline{Z}) \\
\hat{\psi}(t,z) = q(t,z) & (t,z) \in [0,T] \times (-\underline{Z}, \underline{Z})^c
\end{cases}$$

and it belongs to $C^{1+\kappa/2, 2+\kappa}(\Omega^T)$. See for example the computations we did in the proof of Theorem 9.9. Since a^{tr} verifies the PIDE (9.11) we deduce that $\varphi := \hat{\psi} - a^{tr}$ should verify

$$\begin{cases} -\dfrac{\partial \varphi}{\partial t} + \mathcal{A}^h \varphi - \mathcal{B}_t^h \varphi - 2\hat{\pi}\mathcal{Q}^h \varphi - \hat{\pi}^2 \mathcal{G}^h \varphi = F & (t,z) \in [0,T) \times (-\underline{Z},\, \underline{Z}) \\ \varphi(T,z) = 0 & z \in (-\underline{Z},\, \underline{Z}) \\ \varphi(t,z) = 0 & (t,z) \in [0,T] \times (-\underline{Z},\, \underline{Z})^c \end{cases}$$

where $F(t,z) := \left(2\hat{\pi}\mathcal{Q}^h a^{tr} + \hat{\pi}^2 \mathcal{G}^h a^{tr} - \mathcal{H}^h[a^{tr}]\right)(t,z) \geq 0$ for all (t,z), simply from the definition of $\mathcal{H}^h[a^{tr}]$ in (9.9). We can also rearrange the terms in the left hand side of the above PIDE in the following way:

$$\begin{aligned} \mu_{new}(t,z) :=& -\left(\mu^h + 2\hat{\pi}\left(\sigma^2 + \gamma^h\right)\right) \\ \sigma_{new}(t,z) :=& \sqrt{\sigma^2 + \gamma^h} \\ r_{new}(t,z) :=& -2\hat{\pi}\left(\mu^h + \frac{1}{2}\left(\sigma^2 + \gamma^h\right) + \int_{h<|y|} \left(e^{\gamma} - 1\right)\nu(dy)\right) \\ & - \hat{\pi}^2\left(\sigma^2 + \gamma^h + \int_{h<|y|} \left(e^{\gamma} - 1\right)^2 \nu(dy)\right) \\ \nu_{new}(t,z,dy) :=& \left(1 + \hat{\pi}\left(e^{\gamma} - 1\right)\right)^2 \nu(dy) \mathbb{1}_{\{h<|y|\}} \end{aligned}$$

where μ^h stands for $\mu^h(t,z)$ and so on. With these notations we can write

$$\begin{cases} -\dfrac{\partial \varphi}{\partial t} - \mu_{new}\dfrac{\partial \varphi}{\partial z} - \dfrac{1}{2}\sigma_{new}^2 \dfrac{\partial^2 \varphi}{\partial z^2} - \mathcal{B}^{new}\varphi - r_{new}\varphi = F & (t,z) \in [0,T) \times (-\underline{Z},\, \underline{Z}) \\ \varphi(T,z) = 0 & z \in (-\underline{Z},\, \underline{Z}) \\ \varphi(t,z) = 0 & (t,z) \in [0,T] \times (-\underline{Z},\, \underline{Z}) \end{cases}$$

where $\mathcal{B}^{new}$ is the non local operator when one uses the Lévy measure ν_{new}. The Feynman-Kac formula[5] yields

$$\varphi(t,z) = \mathbb{E}\left[\int_t^{\beta^{t,z}} F(s, Z_s^{new,t,z}) \exp\left(\int_t^s r_{new}(u, Z_u^{new,t,z}) du\right) ds\right] \tag{9.25}$$

where

$$\begin{aligned} dZ_s^{new,t,z} =& \mu_{new}(s, Z_s^{new,t,z}) ds + \sigma_{new}(s, Z_s^{new,t,z}) dW_s \\ & + \int \gamma(s, Z_{s-}^{new,t,z}, y) \bar{\mathcal{J}}(s, Z_{s-}^{new,t,z}, dyds) \end{aligned}$$

and $\mathcal{J}$ is a Poisson random measure such that, for all Borel set $\mathcal{O}$ whose closure does not contain zero and any t, z, the process

$$\xi_s := \mathcal{J}(t,z,[0,s] \times \mathcal{O}) - \int_t^s \int_{\mathcal{O} \cap \{|y|>h\}} \left(1 + \hat{\pi}(t,z)\left(e^{\gamma(t,z,y)} - 1\right)\right)^2 \nu(dy)$$

[5]See the proof of Corollary 9.8 to justify the Feynman-Kac formula for PIDE in truncated domains.

is a martingale. From (9.25) and $F \geq 0$ we deduce $\varphi := \hat{\psi} - a^{tr} \geq 0$, which is a contradiction since we assumed that $a^{tr}(t,z) > \hat{\psi}(t,z)$. It follows then that

$$x^2 a^{tr}(t,z) \leq \inf_{\theta \in \mathcal{E}^{\beta}(t,z,x)} \mathbb{E}\left[\left(X^{t,z,x,\theta}_{\beta^{t,z}}\right)^2 q\left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right)\right]$$

which concludes the proof.

□

We conclude the chapter with an estimate on the error $a^h - a^{tr}$:

Theorem 9.11. *Let the Assumptions of Theorem9.9 hold true. Then there exists a positive constant M_1 which only depends on the model parameters such that*

$$\left|a^h(t,z) - a^{tr}(t,z)\right| \leq M_1 \left\|a^h - q\right\|_{\infty,\mho^T} \mathbb{P}\left(\beta^{t,z} < T\right)^{1/2}$$

where $\beta^{t,z}$ is the hitting time of the process Z introduced in (9.23)*. Moreover, there exists a positive constant M_2 which only depends on the parameters of the process Z such that*

$$\mathbb{P}\left(\beta^{t,z} < T\right) \leq \frac{M_2}{\underline{Z}^2}(1+z^2)$$

Proof.
From (9.22) and Lemma 9.10 we have

$$x^2\left(a^h(t,z) - a^{tr}(t,z)\right) \leq \mathbb{E}\left[\left(X^{t,z,x,\theta^{tr}}_{\beta^{t,z}}\right)^2 \left(a^h - q\right)\left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right)\right]$$

where θ^{tr} is the optimal strategy given in (9.24). But also

$$x^2\left(a^h(t,z) - a^{tr}(t,z)\right) \geq \mathbb{E}\left[\left(X^{t,z,x,\theta^{h}}_{\beta^{t,z}}\right)^2 \left(a^h - q\right)\left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right)\right]$$

where θ^h is the optimal strategy associated to the value function a^h. Cauchy-Schwarz inequality and standard estimate on the wealth process imply

$$\left|a^h(t,z) - a^{tr}(t,z)\right| \leq M_1 \mathbb{E}\left[\left|a^h - q\right|^2 \left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right)\right]^{1/2}$$

for some positive $M_1 > 0$ that only depends on the parameters market. Remark that this is possible since θ^{tr} and θ^h belong to the space $\mathcal{E}(t,z,x)$ and then the related π^{tr} and π^h are bounded. We conclude by remarking that

$$\mathbb{E}\left[\left|a^h - q\right|^2 \left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right)\right] = \mathbb{E}\left[\left|a^h - q\right|^2 \left(\beta^{t,z}, Z^{t,z}_{\beta^{t,z}}\right) \mathbb{1}_{\{\beta^{t,z} < T\}}\right]$$

because on the set $\{\beta^{t,z} = T\}$ the compatibility condition (9.12) yields $q(T,.) = 1 = a^h(T,.)$. It follows

$$\left|a^h(t,z) - a^{tr}(t,z)\right| \leq M_1 \left\|a^h - q\right\|_{\infty,\mho^T} \mathbb{P}\left(\beta^{t,z} < T\right)^{1/2}$$

We conclude the proof with the estimate on $\mathbb{P}\left(\beta^{t,z} < T\right)$, using Chebyshev type inequality : from

$$\mathbb{E}\left[\left(Z^{t,z}_{\beta^{t,z}}\right)^2\right] \geq \mathbb{E}\left[\left(Z^{t,z}_{\beta^{t,z}}\right)^2 \mathbb{1}_{\{\beta^{t,z}<T\}}\right]$$

we deduce

$$\mathbb{P}\left(\beta^{t,z} < T\right) \leq \underline{Z}^{-2}\mathbb{E}\left[\left(Z^{t,z}_{\beta^{t,z}}\right)^2\right] \leq 2\left(z^2 + \mathbb{E}\left[\left(Z^{t,z}_{\beta^{t,z}} - z\right)^2\right]\right)$$

We can readapt the proof of Lemma A.1 to deduce

$$\mathbb{P}\left(\beta^{t,z} < T\right) \leq M_2 \underline{Z}^{-2}(1 + z^2)$$

for some positive constant M_2 which only depends on the parameters of Z.

□

We remark that the estimate in Theorem 9.11 makes only use of the $\mathbb{L}^\infty$-norm of $a^h - q$ outside the boundary. It also makes explicit the dependence of the error with respect to the probability that the process Z exits from the domain of truncation $[-\underline{Z}, \underline{Z}]$. Nevertheless this estimate cannot be used to control the error between the optimal strategies associated to a^h and a^{tr}, since these optimal strategies depend on the first derivative of their respective value functions, which are not taken into account in the estimate of the above Theorem.

Part III

Appendix

Appendix A

Doléans-Dade exponential and other estimations

The main objective of this appendix is to give an exhaustive description of the processes used in Chapter 5.

Notations.
In this Appendix, $\mathbb{E}$ denotes the expectation under the historical probability $\mathbb{P}$, $t \in [0,T]$ and $(u,z,x) \in \mathbb{R}^3$. Also, ϑ_h denotes a positive function which only depends on $h > 0$ and $\vartheta_h \to 1$ when $h \to 0^+$. M denotes a positive constant, which may change from line to line of our proofs, which does not depend on h, t or (u,z,x).

We will systematically omit high order terms of the form $o(h)$ in our estimations.

We start by giving a short list of properties for the processes U and Z given in (5.5).

Lemma A.1. *Let Assumptions 5.1-[**C**, **I1**] hold true. For all $t \in [0,T)$, $h > 0$ and $z, u \in \mathbb{R}^2$*

$$\mathbb{E}\left[\left(Z_{t+h}^{t,u,z} - z\right)^2\right] \leq C_{z,2} h \vartheta_h, \qquad C_{z,2} = 2(\sigma_{max}^2 + \|\tau\|_{2,\nu}^2)$$

$$\mathbb{E}\left[\left(U_{t+h}^{t,u} - u\right)^2\right] \leq C_{u,2} h \vartheta_h, \qquad C_{u,2} = 2\left(\sigma_{max}^2 + \left\|\tau^U\right\|_{2,\nu_n}^2\right)$$

and for all $\varepsilon > 0$ one has

$$\mathbb{E}\left[\left|Z_{t+h}^{t,u+\eta,z+\varepsilon} - Z_{t+h}^{t,u,z}\right|^2\right] \leq M\left(\eta^2 + \varepsilon^2\right)\vartheta_h \tag{A.1}$$

$$\mathbb{E}\left[\left|U_{t+h}^{t,u+\eta} - U_{t+h}^{t,u}\right|^2\right] \leq M \eta^2 \vartheta_h \tag{A.2}$$

where $M > 0$ does not depend on $u, z, \varepsilon, \eta, t$ or h

Proof.

The first estimation can be obtained by the Itô-Lévy isometry:

$$\begin{aligned}\mathbb{E}\left[\left(Z_{t+h}^{t,u,z}-z\right)^2\right] &\leq 2\int_t^{t+h}\mathbb{E}\left[\sigma^2(r,U_r^{t,u},Z_r^{t,u,z})+\int_{\mathbb{R}}\gamma^2(r,U_{r-}^{t,u},Z_{r-}^{t,u,z},y)\nu(dy)\right]dr \\ &\leq 2\left(\sigma_{max}^2+\|\tau\|_{2,\nu}^2\right)h\end{aligned}$$

The same type of computations can be done for the process U to prove the result. For (A.1)–(A.2) one can apply Lemma 3.1 in Pham (1998).

□

The other two processes appearing in Chapter 5, $\exp(Z)$ and the derivative of flow DZ, are two examples of stochastic exponentials: let L be a $\mathbb{R}$-valued semimartingale on a filtered probability space $(\Omega,\mathbb{P},(\mathcal{F}_t)_{t\leq T})$ driven by a Brownian motion W and an independent Poisson random measure J whose Lévy measure is denoted by $\nu(dy)$:

$$L_s^{t,l}:=l+\int_t^s b(r)dr+\int_t^s a(r)dW_r+\int_t^s\int_{\mathbb{R}}s(r,y)\bar{J}\left(dydr\right)$$

where b and a are caglad, bounded and measurable real valued processes. We set

$$\beta:=\sup_{r\in[0,T],\omega\in\Omega}|b(r,\omega)| \qquad \text{and} \qquad \alpha:=\sup_{r\in[0,T],\omega\in\Omega}|a(r,\omega)|$$

We assume that s also is a caglad adapted real valued process such that $g(y):=\sup_{r,\omega}|s(r,\omega,y)|$ verifies $g\in\mathbb{L}^2\left(\mathbb{R},\nu(dy)\right)$. We define

$$\Psi_1=\int_{|y|\geq 1}g\left(y\right)\nu\left(dy\right) \qquad \Psi_2=\int_{\mathbb{R}}g^2\left(y\right)\nu\left(dy\right)$$

The stochastic exponential or Doléans-Dade exponential (DDE) (Doléans-Dade, 1970; Ash and Doléans-Dade, 2000) of the process L, usually denoted by $\mathcal{E}(L)$, is the unique semimartingale solution of $dV_r^{t,l}=V_{r-}^{t,l}dL_r^{t,l}$, $V_t^{t,l}=1$. The solution of this SDE can be explicitly given in terms of L (Protter, 2004):

$$V_r^{t,l}=\exp\left(L_r^{t,l}-\frac{1}{2}\left[L^{t,l},L^{t,l}\right]_r^c\right)\prod_{t\leq u\leq r}\left(1+\Delta L_u^{t,l}\right)e^{-\Delta L_u^{t,l}} \tag{A.3}$$

where $[L,L]^c$ stands for the continuous part of the quadratic variation of L. The next Lemma gives some classical estimations on the moments of the process V.

Lemma A.2. *Let $t\in[0,T)$, $l\in\mathbb{R}$ and $h\geq 0$. Then*

$$\begin{aligned}\mathbb{E}\left[\left(V_{t+h}^{t,l}\right)^2\right] &\leq e^{C_{V,1}h}, & C_{V,1} &= 2\beta+\alpha^2+2\Psi^1+\Psi_2 \\ \mathbb{E}\left[\left(V_{t+h}^{t,l}-1\right)^2\right] &\leq C_{V,2}h\vartheta_h, & C_{V,2} &= 2(\alpha^2+\Psi_2) \\ \left|\mathbb{E}\left[V_{t+h}^{t,l}-1\right]\right| &\leq C_{V,3}h\vartheta_h, & C_{V,3} &= \beta+\Psi^1\end{aligned}$$

Proof.
The proof is basically based on the Itô's formula and the Gronwall's inequality. First we have

$$\begin{aligned}\mathbb{E}\left[\left(V_{t+h}^{t,l}\right)^2\right] =&1+\mathbb{E}\left[\int_t^{t+h}\left(V_r^{t,l}\right)^2\left(2b_r+a_r^2+\int_{\mathbb{R}}\left(s^2(r,y)+2s(r,y)\mathbb{1}_{\{|y|\leq 1\}}\right)\nu(dy)\right)dr\right]\\ \leq&1+C_{V,1}\int_t^{t+h}\mathbb{E}\left[\left(V_r^{t,l}\right)^2\right]dr\end{aligned}$$

From the Gronwall's inequality we conclude $\mathbb{E}\left[\left(V_{t+h}^{t,l}\right)^2\right]\leq e^{C_{V,1}h}$. Convexity and Itô-Lévy isometry yield

$$\begin{aligned}&\mathbb{E}\left[\left(V_{t+h}^{t,l}-1\right)^2\right]\leq 2\mathbb{E}\left[\left(\int_t^{t+h}V_r^{t,l}\left(b_r+\int_{|y|\geq 1}s(r,y)\nu(dy)\right)dr\right)^2\right]\\ &+2\mathbb{E}\left[\int_t^{t+h}\left(V_r^{t,l}\right)^2\left(a_r^2+\int_{\mathbb{R}}s^2(r,y)\nu(dy)\right)dr\right]\leq C_{V,2}\int_t^{t+h}\mathbb{E}\left[\left(V_r^{t,l}\right)^2\right]dr+o(h)\end{aligned}$$

and then $\mathbb{E}\left[\left(V_{t+h}^{t,l}-1\right)^2\right]\leq C_{V,2}h\vartheta_h+o(h)$. For the last inequality we have

$$\begin{aligned}&\left|E\left[V_{t+h}^{t,l}-1\right]\right|=\left|\mathbb{E}\left[\int_t^{t+h}V_r^{t,l}\left(b_rdr+\int_{|y|\geq 1}s(r,y)\nu(dy)\right)dr\right]\right|\\ &\leq\int_t^{t+h}\mathbb{E}\left[\left|V_r^{t,l}\right|\left|b_r+\int_{|y|\geq 1}s(r,y)\nu(dy)\right|\right]dr\leq C_{V,3}\int_t^{t+h}\mathbb{E}\left[\left(V_r^{t,l}\right)^2\right]^{\frac{1}{2}}dr\end{aligned}$$

so that $\left|E\left[V_{t+h}^{t,l}-1\right]\right|\leq C_{V,3}h\vartheta_h$

□

The process $\exp(Z)$**.** By using Itô's formula one has

$$de^{Z_r^{t,u,z}} = e^{Z_{r-}^{t,u,z}}\left(\tilde{\mu}dr+\sigma dW_r^1+\int_{\mathbb{R}}\left(e^{\gamma}-1\right)\bar{J}(dydr)\right),\quad e^{Z_t^{t,u,z}}=e^z \tag{A.4}$$

where $\tilde{\mu}=\tilde{\mu}\left(r,U_r^{t,u},Z_r^{t,u,z}\right)$ is defined in (5.6). We deduce that $\exp(Z_r-z)$ is a stochastic exponential.

Corollary A.3. *Suppose that Assumptions 5.1-*[**C**, **I1**] *hold true. For all* $t,\in[0,T)$,

$h \geq 0$ and $z, u \in \mathbb{R}$ one has

$$\begin{aligned}
\mathbb{E}\left[\left(e^{Z^{t,u,z}_{t+h}-z}\right)^2\right] &\leq e^{C_{e,1}h}, & C_{e,1} &= 2\|\tilde{\mu}\| + \sigma^2_{max} + 2\|\tau\|_{1,\nu} + \|\tau\|^2_{2,\nu} \\
\mathbb{E}\left[\left(e^{Z^{t,u,z}_{t+h}-z}-1\right)^2\right] &\leq C_{e,2}h\vartheta_h, & C_{e,2} &= 2(\sigma^2_{max} + \|\tau\|^2_{2,\nu}) \\
\left|\mathbb{E}\left[e^{Z^{t,u,z}_{t+h}-z}-1\right]\right| &\leq C_{e,3}h\vartheta_h, & C_{e,3} &= \|\tilde{\mu}\| + \|\tau\|_{1,\nu} \\
\mathbb{E}\left[\left(e^{Z^{t,u,z}_{t+h}-z}-1\right)^2\right] &\geq C_{e,5}h, & C_{e,5} &= (\sigma^2_{min} \vee |\Gamma|)
\end{aligned}$$

Moreover

$$\mathbb{E}\left[\left(e^{Z^{t,u+\eta,z+\varepsilon}_{t+h}} - e^{Z^{t,u,z}_{t+h}}\right)^2\right] \leq M(e^z)^2(\eta^2+\varepsilon^2)\vartheta_\eta\vartheta_\varepsilon\vartheta_h \tag{A.5}$$

where $M > 0$ does not depend on u, z, h, η or ε. If Assumption 5.1-[**I2**] *also holds true then*

$$\mathbb{E}\left[\left(e^{Z^{t,u,z}_{t+h}-z}-1\right)^4\right] \leq C_{e,4}h\vartheta_h, \qquad C_{e,4} = \left(2\|\tau\|^4_{4,\nu} + \sigma^2_{max} + 2\|\tau\|^2_{2,\nu}\right)h\vartheta_h$$

Proof.
For the first three estimations we use Lemma A.2 with $\beta = \|\tilde{\mu}\|$, $\alpha = \sigma_{max}$ and $g = \tau$. For the fourth inequality we remark that

$$\mathbb{E}\left[\left(e^{Z^{t,u,z}_{t+h}-z}-1\right)^2\right] = 2\mathbb{E}\left[\int_t^{t+h}\left(e^{Z^{t,u,z}_{r}-z}-1\right)d(e^{Z^{t,u,z}_{r}-z})\right] + \mathbb{E}\left[\left[e^{Z^{t,u,z}_{t+h}}-1\right]_{t+h}\right]$$

Since

$$2\mathbb{E}\left[\int_t^{t+h}\left(e^{Z^{t,u,z}_{r-}-z}-1\right)d(e^{Z^{t,u,z}_{r}-z})\right] \geq -Mh^{\frac{3}{2}}$$

for some $M > 0$, we can omit this high order term. By using Jensen's inequality we get

$$\begin{aligned}
\mathbb{E}\left[\left[e^{Z^{t,u,z}_{t+h}-z}-1\right]_{t+h}\right] \geq& (\sigma^2_{min} \vee |\Gamma|)\int_t^{t+h}\mathbb{E}\left[(e^{Z^{t,u,z}_{r}-z})^2\right]dr \\
\geq& (\sigma^2_{min} \vee |\Gamma|)\int_t^{t+h} e^{2\mathbb{E}[Z^{t,u,z}_{r}-z]}dr
\end{aligned} \tag{A.6}$$

and

$$\mathbb{E}[Z^{t,u,z}_r - z] = \int_t^r \mathbb{E}\left[\mu(s, U^{t,u}_s, Z^{t,u,z}_s) + \int_{|y|\geq 1}\gamma(s, U^{t,u}_{s-}, Z^{t,u,z}_{s-}, y)\nu(dy)\right]ds$$

From Assumptions 5.1 we have $\mu(t,u,z) \geq -\|\mu\|_\infty$ whereas

$$\int_{|y|\geq 1}\gamma(s,u,z,y)\nu(dy) \geq -\int_{|y|\geq 1}|\gamma(s,u,z,y)|\nu(dy) \geq -\|\tau\|_{1,\nu}$$

so then $\mathbb{E}[Z_r^{t,u,z} - z] \geq -M(r-t)$ for some positive M. From (A.6) we conclude

$$\mathbb{E}\left[\left(e^{Z_{t+h^{t,u,z}}-z} - 1\right)^2\right] \geq (\sigma_{min}^2 \vee |\Gamma|)h$$

For estimation (A.5) we apply Lemma 3.1 in Pham (1998).

For the last inequality we first apply Itô's formula to $(e^{Z-z}-1)^2$ to obtain de decomposition

$$(e^{Z_{t+h}^{t,u,z}-z} - 1)^2 = \int_t^{t+h} \alpha_s^{(1)} ds + \int_t^{t+h} \alpha_s^{(2)} dW_s + \int_t^{t+h}\int_{\mathbb{R}} \alpha_s^{(3)} \tilde{J}(dyds)$$

Since we are interested to $(e^{Z_{t+h}^{t,u,z}-z} - 1)^4$ we can forget the process $\alpha^{(1)}$ since it will contribute with a high order term $o(h)$. We are left with

$$\begin{aligned}
\alpha_s^{(2)} =& 2\sigma e^{Z_s^{t,u,z}-z}\left(e^{Z_s^{t,u,z}-z} - 1\right)\\
\alpha_s^{(3)} =& \left(e^{Z_s^{t,u,z}-z+\gamma} - 1\right)^2 - \left(e^{Z_s^{t,u,z}-z} - 1\right)^2\\
=& e^{Z_s^{t,u,z}-z}(e^\gamma - 1)\left(\left(e^{Z_s^{t,u,z}-z}\right)(e^\gamma-1) + 2\left(e^{Z_s^{t,u,z}-z} - 1\right)\right)
\end{aligned}$$

The Itô-Lévy isometry yields

$$\begin{aligned}
&\mathbb{E}\left[\left(e^{Z_{t+h}^{t,u,z}-z} - 1\right)^4\right] \leq 2\left\|\tau\right\|_{4,\nu}^4 \int_t^{t+h} \mathbb{E}\left[\left(e^{Z_s^{t,u,z}-z}\right)^4\right] ds\\
&+\left(4\sigma_{max}^2 + 8\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h} \mathbb{E}\left[\left(e^{Z_s^{t,u,z}-z}\right)^4\right]^{1/2} \mathbb{E}\left[\left(e^{Z_{t+h}^{t,u,z}-z} - 1\right)^4\right]^{1/2} ds
\end{aligned}$$

Since $E\left[\left(e^{Z_s^{t,u,z}-z}\right)^4\right] \leq e^{2M(s-t)}$ for some $M > 0$ and $\sqrt{x} \leq x + 1/4$ for all $x \geq 0$ we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(e^{Z_{t+h}^{t,u,z}-z} - 1\right)^4\right] \leq& 2\left\|\tau\right\|_{4,\nu}^4 h\vartheta_h\\
&+\left(4\sigma_{max}^2 + 8\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h} e^{M(s-t)}\left(\frac{1}{4} + \mathbb{E}\left[\left(e^{Z_s^{t,u,z}-z} - 1\right)^4\right]\right)ds\\
\leq& \left(2\left\|\tau\right\|_{4,\nu}^4 + \sigma_{max}^2 + 2\left\|\tau\right\|_{2,\nu}^2\right) h\vartheta_h
\end{aligned}$$

□

The derivative of flow DZ. The second example of stochastic exponential associated to Z is the so called derivative of flow , defined by

$$DZ_s^{t,u,z} = 1 + \int_t^s DZ_{r-}^{t,u,z}\left(\frac{\partial\mu_r}{\partial z}dr + \frac{\partial\sigma_r}{\partial z}dW_r^1 + \int_{\mathbb{R}}\frac{\partial\gamma_r(y)}{\partial z}\bar{J}(dydr)\right) \qquad \text{(A.7)}$$

where $\partial\mu_r/\partial u$ stands for $\partial\mu/\partial z\left(r, U_r^{t,u}, Z_r^{t,u,z}\right)$ and so on.

Corollary A.4. *Suppose that Assumptions 5.1-*[$\mathbf{C},\mathbf{I1}$] *hold true. For all* $t \in [0,T)$ *,* $h \geq 0$ *and* $u,z \in \mathbb{R}$

$$\mathbb{E}\left[\left(DZ^{t,u,z}_{t+h}\right)^2\right] \leq e^{C_{dz,1}h}, \qquad C_{dz,1} = K_{max}\left(2(1+\|\tau\|_{1,\nu}) + K_{max}(1+\|\tau\|^2_{2,\nu})\right)$$

$$\mathbb{E}\left[\left(DZ^{t,u,z}_{t+h}-1\right)^2\right] \leq C_{dz,2}h\vartheta_h, \qquad C_{dz,2} = 2K^2_{max}(1+\|\tau\|^2_{2,\nu})$$

$$\left|\mathbb{E}\left[DZ^{t,u,z}_{t+h}-1\right]\right| \leq C_{dz,3}h\vartheta_h, \qquad C_{dz,3} = K_{max}(1+\|\tau\|_{1,\nu})$$

If we define for $\varepsilon > 0$

$$D_\sharp Z^{\varepsilon,u,z}_s := \varepsilon^{-1}\left(Z^{t,u,z+\varepsilon}_s - Z^{t,u,z}_s\right) \tag{A.8}$$

then

$$\mathbb{E}\left[\left|D_\sharp Z^{\varepsilon,t,u,z}_s - DZ^{t,u,z}_s\right|\right] \to 0 \quad \text{when } \varepsilon \to 0$$

Proof.
For the first three inequalities we apply Lemma A.2 with $\beta = \alpha = K_{max}$ and $g = K_{max}\tau$. For $D_\sharp Z$, we remark that Lemma A.1 gives $\mathbb{E}\left[\left|D_\sharp Z^{\varepsilon,t,u,z}_s\right|^2\right] \leq M\vartheta_{s-t}$, so then the family $\varepsilon \to D_\sharp Z^{\varepsilon,t,u,z}_s$ is uniformly bounded in $\mathbb{L}^2(\mathbb{P},\mathcal{F}_s)$, which in particular means that it is uniformly integrable. If we prove that $D_\sharp Z^{\varepsilon,t,u,z}_s \xrightarrow{\mathbb{P}} DZ^{t,u,z}_s$ then dominated convergence applies and we obtain the result. We first remark that the process $D_\sharp Z^{\varepsilon,t,u,z}$ is a Doléans-Dade exponential: $dD_\sharp Z^{\varepsilon,t,u,z}_s = D_\sharp Z^{\varepsilon,t,u,z}_{s-}dP^{\epsilon,t,u,z}_s$ where

$$\begin{aligned}
P^{\epsilon,t,u,z}_r &:= \int_t^r v^{1,\varepsilon}_s ds + \int_t^r v^{2,\varepsilon}_s dW^1_s + \int_t^r \int_{\mathbb{R}} v^{3,\varepsilon}_s(y)\bar{J}(dyds) \\
v^{1,\varepsilon}_s &:= \int_0^1 \partial_z\mu\left(s, U^{t,u}_s, Z^{t,u,z}_s + x\left(Z^{t,u,z+\varepsilon}_s - Z^{t,u,z}_s\right)\right)dx \\
v^{2,\varepsilon}_s &:= \int_0^1 \partial_z\sigma\left(s, U^{t,u}_s, Z^{t,u,z}_s + x\left(Z^{t,u,z+\varepsilon}_s - Z^{t,u,z}_s\right)\right)dx \\
v^{3,\varepsilon}_s(y) &:= \int_0^1 \partial_z\gamma\left(s, U^{t,u}_{s-}, Z^{t,u,z}_{s-} + x\left(Z^{t,u,z+\varepsilon}_{s-} - Z^{t,u,z}_{s-}\right), y\right)dx
\end{aligned}$$

Let us define

$$P^{*,t,u,z}_s := \int_t^s \partial_z\mu_r dr + \int_t^s \partial_z\sigma_r dW^1_r + \int_t^r \int_{\mathbb{R}} \partial_z\gamma_r(y)\bar{J}(dydr)$$

where $\partial_z\mu_r$ stands for $\partial_z\mu(r, U^{t,z}_r, Z^{t,u,z}_r)$ and so on. From Assumptions 5.1, Lemma A.1 and Doob's inequalities it is not complicate to prove that

$$\mathbb{E}\left[\sup_{t\leq s\leq T}\left(P^{\epsilon,t,u,z}_s - P^{*,t,u,z}_s\right)^2\right] \to 0, \quad \epsilon \to 0$$

and then, by using the Itô's formula and Gronwall's inequality

$$\mathbb{E}\left[\sup_{t\le s\le T}\left(D_\sharp Z_s^{\varepsilon,t,u,z}-DZ_s^{t,u,z}\right)^2\right]\to 0,\quad \epsilon\to 0$$

We deduce then that $D_\sharp Z^{\varepsilon,t,u,z}\overset{ucp}{\to} DZ^{t,u,z}$ when $\varepsilon\to 0$, which concludes our proof.

□

Another inequality used in the proof of Theorem 5.4 mixes the exponential e^Z and $|DZ|$

Lemma A.5. *Let Assumption 5.1-*[$\mathbf{C},\mathbf{I1},\mathbf{I2}$] *hold true. Then for all* $l\in\mathbb{R}$ *and* $(t,u,z\in[0,T)\times\mathbb{R}^2)$

$$\mathbb{E}\left[\left(1-l\left(e^{Z_{t+h}^{t,u,z}-z}-1\right)\right)^2\left|DZ_{t+h}^{t,u,z}\right|\right]\le 1+C_{dz,e}(l)h\vartheta_h$$

where

$$\begin{aligned} C_{dz,e}(l) &= (C_{dz,3}+C_{dz,2})+2|l|\left(C_{e,3}+K_{max}(\sigma_{max}+3\|\tau\|_{2,\nu}^2)\right)\\ &+\frac{|l|^2}{2}\left(3C_{e,2}+K_{max}C_{e,4}\right)\end{aligned}$$

Proof.
In order to simplify our notations, we will always omit arguments in the coefficients μ,σ and γ and, when there is no ambiguity, we also omit the superscript (t,u,z) in the processes e^Z and DZ.
From (A.3) we can derive the explicit solution of the SDE (A.7). In particular it is straightforward to prove that the process $R_s:=|DZ_s|$ verifies

$$\frac{dR_s}{R_s}=\partial_z\mu ds+\partial_z\sigma dW_s^1+\int_{A_s^-}(-2-\partial_z\gamma)\bar{J}(dyds)+\int_{A_s^+}\partial_z\gamma\bar{J}(dydu)$$

where

$$\begin{aligned} A_s^- &:= A^-(s,u,z):=\{y\in\mathbb{R}\,|\,\partial_z\gamma(s,u,z,y)<-1\}\\ A_s^+ &:= A^+(s,u,z):=\{y\in\mathbb{R}\,|\,\partial_z\gamma(s,u,z,y)\ge-1\}\end{aligned}$$

and from Assumptions 5.1 we have

$$A^-(s,u,z):=\{y\in\mathbb{R}\,|\,\partial_z\gamma(s,u,z,y)<-1\}\subseteq\{y\in\mathbb{R}\,|\,K_{max}\tau(y)>1\}$$

For sake of compactness let us define

$$\alpha(s,u,z,y):=\partial_z\gamma(s,u,z,y)\mathbb{1}_{A^+(s,u,z)}-(2+\partial_z\gamma(s,u,z,y))\mathbb{1}_{A^-(s,u,z)}\qquad\text{(A.9)}$$

which trivially verifies

$$\sup_{s,u,z}|\alpha(s,u,z,y)|\le K_{max}\tau(y)+2\mathbb{1}_{\{K_{max}\tau(y)>1\}}\qquad\text{(A.10)}$$

Remark that

$$\nu(\{y \in \mathbb{R} \mid K_{max}\tau(y) > 1\}) \leq K_{max}^2 \|\tau\|_{2,\nu}^2 \tag{A.11}$$

If now $V_s := le^{Z_s - z} - (l+1)$ then our aim is to estimate $\mathbb{E}\left[V_{t+h}^2 R_{t+h}\right]$. Let us denote

$$\Delta V_s := V_s - V_{s-} \quad \text{and} \quad \Delta R_s := R_s - R_{s-}$$

The Itô's formula yields:

$$\begin{aligned}
V_s^2 R_u =& 1 + G_s + I_s + \text{ martingales} \\
dG_s :=& \left(V_s^2 R_s \partial_z \mu + 2V_s R_s l e^{Z_s - z}\tilde{\mu} + R_s l^2 e^{2(Z_s - z)}\sigma^2 + 2V_s\left(le^{Z_s - z}\sigma R_s \partial_z \sigma\right)\right) ds \\
dI_s :=& \int_{|y|\geq 1} \left(2V_{s-}R_{s-}\Delta V_s + V_{s-}^2 \Delta R_s\right)\nu(dy)ds \\
&+ \int_{\mathbb{R}} \left(\Delta V_s^2 R_{s-} + \Delta V_s^2 \Delta R_s + 2V_{s-}\Delta V_s \Delta R_s\right)\nu(dy)ds
\end{aligned}$$

where $G_t = I_t = 0$. Elementary estimations yield

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_t^{t+h} dG_s\right]\right| \leq K_{max}\int_t^{t+h} \mathbb{E}\left[V_s^2 R_s\right] ds + l^2\sigma_{max}^2 \int_t^{t+h} \mathbb{E}\left[e^{2(Z_s - z)}R_s\right] ds \\
&+ 2|l|\left(\|\tilde{\mu}\| + K_{max}\sigma_{max}\right)\int_t^{t+h} \mathbb{E}\left[V_s^2 R_s\right]^{1/2} \mathbb{E}\left[R_s e^{2(Z_s - s)}\right]^{1/2} ds
\end{aligned}$$

where we used the estimations given in Assumptions 5.1. From Corollaries A.3 and A.4 we also have $\mathbb{E}\left[e^{2(Z_s - z)}R_s\right] \leq e^{M(s-t)}$ for some positive $M > 0$: it follows

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_t^{t+h} dG_s\right]\right| \leq K_{max}\int_t^{t+h} \mathbb{E}\left[V_s^2 R_s\right] ds + l^2\sigma_{max}^2 h\vartheta_h \\
&+ 2|l|\left(\|\tilde{\mu}\| + K_{max}\sigma_{max}\right)\int_t^{t+h} \mathbb{E}\left[V_s^2 R_s\right]^{1/2} e^{M(s-t)/2} du
\end{aligned}$$

For the process I we first have $\Delta V_s = le^{Z_{s-} - z}(e^{\gamma} - 1)$ and $\Delta R_s = R_{s-}\alpha_s$ where α is defined in (A.9). It follows

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_t^{t+h} dI_s\right]\right| \leq \int_t^{t+h} \mathbb{E}\left[V_{s-}^2 R_{s-}\int_{|y|\geq 1} |\alpha_s|\nu(dy)\right] ds \\
&+ 2|l|\int_t^{t+h} \mathbb{E}\left[|V_{s-}|R_{s-}e^{Z_{s-} - z}\int_{\mathbb{R}} (|e^{\gamma} - 1|\mathbb{1}_{\{|y|\leq 1\}} + |\alpha_{s-}||e^{\gamma} - 1|)\nu(dy)\right] ds \\
&+ |l|^2\int_t^{t+h} \mathbb{E}\left[e^{2(Z_{s-} - z)}R_{s-}\int_{\mathbb{R}} (e^{\gamma} - 1)^2(1 + |\alpha_{s-}|)\nu(dy)\right] ds
\end{aligned}$$

Using Assumptions 5.1 and estimations (A.11)–(A.10) we obtain

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_t^{t+h} dI_s\right]\right| \leq \left(K_{max}\|\tau\|_{1,\nu} + 2K_{max}^2\|\tau\|_{2,\nu}^2\right)\int_t^{t+h} \mathbb{E}\left[V_{s-}^2 R_{s-}\right] ds \\
&+ 2|l|\left(\|\tau\|_{1,\nu} + 3K_{max}\|\tau\|_{2,\nu}^2\right)\int_t^{t+h} \mathbb{E}\left[V_s^2 R_s\right]^{1/2} \mathbb{E}\left[R_{s-}e^{2(Z_{s-} - z)}\right]^{1/2} ds \\
&+ |l|^2\left(3\|\tau\|_{2,\nu}^2 + K_{max}\|\tau\|_{3,\nu}^3\right)\int_t^{t+h} \mathbb{E}\left[e^{2(Z_{s-} - z)}R_{s-}\right] ds
\end{aligned}$$

As for the process G, after simplifications, the above reduces to

$$\left|\mathbb{E}\left[\int_t^{t+h} dI_s\right]\right| \leq \left(K_{max}\left\|\tau\right\|_{1,\nu} + 2K_{max}^2\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h}\mathbb{E}\left[V_{s-}^2R_{s-}\right]ds$$
$$+ 2|l|\left(\left\|\tau\right\|_{1,\nu} + 3K_{max}\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h}\mathbb{E}\left[V_s^2R_s\right]^{1/2}e^{M(s-t)/2}ds$$
$$+ |l|^2\left(\sigma_{max}^2 + 3\left\|\tau\right\|_{2,\nu}^2 + K_{max}\left\|\tau\right\|_{3,\nu}^3\right)h\vartheta_h$$

By adding the above estimation with the one we found for the process G we prove that

$$\mathbb{E}\left[V_{t+h}^2R_{t+h}\right] \leq 1 + \left(K_{max} + K_{max}\left\|\tau\right\|_{1,\nu} + 2K_{max}^2\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h}\mathbb{E}\left[V_s^2R_s\right]ds$$
$$+ 2|l|\left(\|\tilde{\mu}\| + K_{max}\sigma_{max} + \left\|\tau\right\|_{1,\nu} + 3K_{max}\left\|\tau\right\|_{2,\nu}^2\right)\int_t^{t+h}\mathbb{E}\left[V_s^2R_s\right]^{1/2}e^{M(s-t)/2}ds$$
$$+ |l|^2\left(\sigma_{max}^2 + 3\left\|\tau\right\|_{2,\nu}^2 + K_{max}\left\|\tau\right\|_{3,\nu}^3\right)h\vartheta_h$$

We can simplify the above estimate by using the constants introduced in Corollaries A.3–A.4: since

$$K_{max} + K_{max}\left\|\tau\right\|_{1,\nu} := C_{dz,3} \qquad 2K_{max}^2\left\|\tau\right\|_{2,\nu}^2 \leq C_{dz,2}$$
$$\|\tilde{\mu}\| + \left\|\tau\right\|_{1,\nu} := C_{e,3} \qquad \sigma_{max}^2 + 3\left\|\tau\right\|_{2,\nu}^2 \leq \frac{3}{2}C_{e,2}$$

and

$$K_{max}\left\|\tau\right\|_{3,\nu}^3 \leq K_{max}\left(\int_{\{y:\tau(y)\leq 1\}}\tau^2(y)\nu(dy) + \int_{\{y:\tau(y)>1\}}\tau^4(y)\nu(dy)\right) \leq \frac{K_{max}}{2}C_{e,4}$$

we obtain

$$\mathbb{E}\left[V_{t+h}^2R_{t+h}\right] \leq 1 + (C_{dz,3} + C_{dz,2})\int_t^{t+h}\mathbb{E}\left[V_s^2R_s\right]ds$$
$$+ 2|l|\left(C_{e,3} + K_{max}(\sigma_{max} + 3\left\|\tau\right\|_{2,\nu}^2)\right)\int_t^{t+h}\mathbb{E}\left[V_s^2R_s\right]^{1/2}e^{M(s-t)/2}ds$$
$$+ \frac{|l|^2}{2}\left(3C_{e,2} + K_{max}C_{e,4}\right)h\vartheta_h$$

Also $\int_t^{t+h} e^{M(s-t)/2} = h + o(h)$ and $\sqrt{x} \leq px + (4p)^{-1}$ for any $p > 0$: it follows then

$$\mathbb{E}\left[V_{t+h}^2R_{t+h}\right] \leq 1 + \frac{1}{2p}|l|\left(C_{e,3} + K_{max}(\sigma_{max} + 3\left\|\tau\right\|_{2,\nu}^2)\right)h\vartheta_h$$
$$+\frac{|l|^2}{2}\left(3C_{e,2} + K_{max}C_{e,4}\right)h\vartheta_h$$
$$+\int_t^{t+h}\left(C_{dz,3} + C_{dz,2} + 2p|l|\left(C_{e,3} + K_{max}(\sigma_{max} + 3\left\|\tau\right\|_{2,\nu}^2)\vartheta_h\right)\right)\mathbb{E}\left[V_s^2R_s\right]ds$$

We can finally apply Gronwall's inequality to deduce

$$\mathbb{E}\left[V_{t+h}^2 R_{t+h}\right] \leq 1 + \frac{1}{2p}|l|\left(C_{e,3} + K_{max}(\sigma_{max} + 3\left\|\tau\right\|_{2,\nu}^2)\right)h\vartheta_h$$
$$+\frac{|l|^2}{2}\left(3C_{e,2} + K_{max}C_{e,4}\right)h\vartheta_h$$
$$+\left(C_{dz,3} + C_{dz,2}\right)h\vartheta_h + 2p|l|\left(C_{e,3} + K_{max}(\sigma_{max} + 3\left\|\tau\right\|_{2,\nu}^2)\vartheta_h\right)h\vartheta_h$$

We minimize over $p > 0$ and find the optimal estimate for $p = 1/2$, which concludes our proof.

□

Appendix B

About a cubic ordinary differential equation

In this part we want to study the following ODE

$$-L'(t) = \Lambda + \Psi L(t)(L^2(t)+1),\; L(T) = 0,\; t \in [0,T] \tag{B.1}$$

where $\Psi \geq 0$ and $\Lambda \geq 0$ are given constants. We can assume that $\Lambda > 0$, otherwise the solution will be null on $[0,T]$. It follows that L is non increasing and then positive. In particular

$$L(0) = \sup_{t\in[0,T]} L(t)$$

We are interested in the behavior of L in the neighborhood of zero, and in finding conditions on Λ and Ψ under which the function L has no explosion at $t = 0$, i.e. $L(0) < +\infty$.

Lemma B.1. *Let $y^* < 0$ be the unique real root of $l \to \Psi l^3 + \Psi l + \Lambda$ and $\triangle := \sqrt{3(y^*)^2+4}$. Define*

$$Q(L) := \frac{1}{2(3(y^*)^2+1)}\left[\log\left(\frac{(L-y^*)^3}{L^3+L+\Lambda/\Psi}\right) - \frac{6y^*}{\triangle}\arctan\left(\frac{2L+y^*}{\triangle}\right)\right]$$

which depends only on the ratio Λ/Ψ. The solution of the ODE in (B.1) *is implicitly given by $Q(L(t)) = \Psi(T-t)+Q(0)$. If we set*

$$T(y^*) := -\frac{1}{\Psi}\left(\frac{3y^*\pi}{2(3(y^*)^2+1)\triangle} + Q(0)\right) \tag{B.2}$$

then $L(0) < +\infty$ if and only if $T < T(y^*)$. Moreover $T(y^*) \to +\infty$ when $\Lambda \to 0$ and Ψ remains fixed.

Proof.
Let $0 < \lambda := \Lambda/\Psi$ and $q(l) := l^3 + l + \lambda$. Since $q' > 0$ and $q(0) = \lambda > 0$ we deduce that there exists a unique $y^* < 0$ such that $q(y^*) = 0$ and then $q(l) = (l-y^*)(l^2+y^*l+1+(y^*)^2)$. Finally remark that y^* only depends on the ratio Λ/Ψ. One may use Cardano's formulae to find y^* exactly.

The ODE (B.1) can be integrated separately so then, at least formally, we can write

$$\int \frac{dL}{L^3 + L + \lambda} = \Psi(\zeta - t), \text{ for some } \zeta \in \mathbb{R}$$

In particular if $Q(l)$ is a primitive of $1/q$ then the solution of (B.1) will be given by $Q(L(t)) = \Psi(\zeta - t)$ with ζ such that $L(T) = 0$. Elementary computations give us this primitive Q:

$$Q(L) = \frac{1}{2(3(y^*)^2+1)}\left[\log\left(\frac{(L-y^*)^3}{L^3+L+\lambda}\right) - \frac{6y^*}{\triangle}\arctan\left(\frac{2L+y^*}{\triangle}\right)\right]$$

According to the initial condition $L(T) = 0$ we have $Q(L(t)) = Q(0) + \Psi(T - t)$ where

$$Q(0) = \frac{1}{2(3(y^*)^2+1)}\left[\log\left(\frac{(y^*)^2}{1+(y^*)^2}\right) - \frac{6y^*}{\triangle}\arctan\left(\frac{y^*}{\triangle}\right)\right] \leq 0$$

since $\lambda = -y^*(1 + (y^*)^2)$. If we can invert this primitive then the solution of ODE (B.1) is given by $L(t) = Q^{-1}(\Psi(T - t) + Q(0))$. But remark that $Q'(L) > 0$ if $L > 0$. Furthermore

$$\lim_{L\to+\infty} Q(L) = -\frac{3y^*\pi}{2(3(y^*)^2+1)\triangle}$$

It follows then that we can invert the primitive if and only if

$$Q(0) \leq \Psi(T-t) + Q(0) \leq -\frac{3y^*\pi}{2(3(y^*)^2+1)\triangle}$$

for all $t \in [0, T]$, and this is possible if and only if

$$T \leq -\frac{1}{\Psi}\left(Q(0) + \frac{3y^*\pi}{2(3(y^*)^2+1)\triangle}\right) := T^*(y^*)$$

Remark that since $y^* < 0$ and $Q(0) \leq 0$ it follows that $T^*(y^*) > 0$. In particular if $T < T(y^*)$ then $L(0) = \sup_t L(t) < +\infty$. To conclude remark that when $\Lambda \to 0$ we have $y^* \to 0$ and then $Q(0) \to -\infty$, so $T(y^*) \to +\infty$.

$\square$

Appendix C

Hölder spaces

C.1 Introduction

We give here a complete definition of the functional spaces used in Chapters 6 and 7. We call elliptic those spaces of functions which take values in $\mathbb{R}^n$, whereas parabolic are those spaces of functions defined in $[0,T]\times\mathbb{R}^n$. The difference between Hölder spaces of type 1 and type 2 arises in their parabolic version. This distinction is needed since the natural space in which one has to work is not the same if one deals with processes leaded by a Brownian motion and a Poisson random measure (Chapter 6) or only by a Poisson random measure (Chapter 7).

For any $l > 0$ we define

$$
\begin{aligned}
l &= \lfloor l \rfloor + \{l\}^- \qquad \text{where} \qquad \{l\}^- \in [0,1), \quad \lfloor l \rfloor \in \mathbb{N} \\
l &= \lceil l \rceil + \{l\}^+ \qquad \text{where} \qquad \{l\}^+ \in (0,1], \quad \lceil l \rceil \in \mathbb{N}
\end{aligned}
$$

From now on M denotes a positive constant which may change from line to line and $\varphi : \mathbb{R}^n \to E$ is a measurable map, where E equipped with $\| \, \|$ is a Banach space. Often E is some $\mathbb{R}^n$ or $\mathbb{S}_n(\mathbb{R})$, the space of symmetric matrices. For $\beta \in (0,1]$ we define

$$
\langle \varphi \rangle^{(\beta)} := \sup_{|x-x'|\le 1} \frac{|\varphi(x) - \varphi(x')|}{|x-x'|^\beta} \tag{C.1}
$$

We start with elliptic Hölder spaces: for a non negative l let $C^{\lceil l \rceil}(\mathbb{R}^n)$ denote the space of differentiable functions on $\mathbb{R}^n$ which are continuous together with their derivative of all order $j \le \lceil l \rceil$. On this space define the norm

$$
\|\varphi\|_{l,e} := \sum_{j=0}^{\lceil l \rceil} \sum_{(j)} \left\| D_x^j \varphi \right\|_\infty + \sum_{(\lceil l \rceil)} \langle D_x^{\lceil l \rceil} \varphi \rangle^{(\{l\}^+)} \tag{C.2}
$$

The elliptic Hölder space of order l is defined as the subset of $C^{\lceil l \rceil}(\mathbb{R}^n)$ of functions with finite norm:

$$
H_e^l(\mathbb{R}^n) := C^{\lceil l \rceil}(\mathbb{R}^n) \cap \left\{ \|\varphi\|_{l,e} < \infty \right\}
$$

We make the convention that $H_e^0(\mathbb{R}^n) = \mathbb{L}_m^\infty(\mathbb{R}^n)$, the space of bounded and measurable functions. Equipped with the above norm they all are Banach spaces.

C.2 Parabolic Hölder spaces of type 1

In this paragraph we (shortly) describe the parabolic Hölder spaces used in Chapter 6. For a complete review see for example Chapter I in Ladyzenskaja et al. (1967). Let $Q_T := (0,T) \times \mathbb{R}^n$ and $\bar{Q}_T$ its closure. For $\varphi : [0,T] \times \mathbb{R}^n \to E$ and $\beta \in (0,1)$ we define

$$\langle \varphi \rangle_{x,Q_T}^{(\beta)} := \sup_{t \le T, x, x' : |x-x'| \le 1} \frac{|\varphi(t,x) - \varphi(t,x')|}{|x-x'|^\beta}$$
$$\langle \varphi \rangle_{t,Q_T}^{(\beta)} := \sup_{x \in \mathbb{R}^n, t, t' : |t-t'| \le 1} \frac{|\varphi(t,x) - \varphi(t',x)|}{|t-t'|^\beta}$$

Let l be a positive non integer real number: $l \in \mathbb{R}^+ \setminus \mathbb{N}^*$ and $C^{\lfloor l/2 \rfloor, \lfloor l \rfloor}(\mathbb{R}^n)$ denote the space of continuously differentiable functions on Q_T which are continuous up the boundary together with their mixed derivative of the form $D_t^r D_x^s$ for all $2r + s \le \lfloor l \rfloor$. On this functional space we introduce the following norm

$$\|\varphi\|_{l/2,l} := \sum_{j=0}^{\lfloor l \rfloor} \sum_{2r+s=j} \|D_t^r D_x^s \varphi\|_\infty + \langle \varphi \rangle_{Q_T}^{(l)} \tag{C.3}$$
$$\langle \varphi \rangle_{Q_T}^{(l)} = \langle \varphi \rangle_{x,Q_T}^{(l)} + \langle \varphi \rangle_{t,Q_T}^{(l/2)}$$

$$\langle \varphi \rangle_{x,Q_T}^{(l)} = \sum_{2r+s=\lfloor l \rfloor} \langle D_t^r D_x^s \varphi \rangle_{x,Q_T}^{\{l\}^-}, \qquad \langle \varphi \rangle_{t,Q_T}^{(l/2)} = \sum_{0 < l-2r-s < 2} \langle D_t^r D_x^s \varphi \rangle_{x,Q_T}^{(l-2r-s)/2}$$

The parabolic Hölder space of type 1 is then defined as

$$C^{l/2,l}([0,T] \times \mathbb{R}^n) := C^{\lfloor l/2 \rfloor, \lfloor l \rfloor}(\mathbb{R}^n) \cap \left\{ \|\varphi\|_{l/2,l} < \infty \right\}$$

and it is a Banach space. There are is no ambiguity to call it $C^{l/2,l}([0,T] \times \mathbb{R}^n)$ since l is always non integer and we can then distinguish it from $C^{\lfloor l/2 \rfloor, \lfloor l \rfloor}(\mathbb{R}^n)$.

We list here the parabolic Hölder spaces of type 1 which are used in Chapter 6 with their relative norm. For $\beta \in (0,1)$:
Parabolic Hölder space of order β**:** $C^{\beta/2,\beta}([0,T] \times \mathbb{R}^n)$. For $\varphi \in C^0([0,T] \times \mathbb{R}^n)$:

$$\|\varphi\|_{\beta/2,\beta} := \|\varphi\|_\infty + \langle \varphi \rangle_{x,Q_T}^{(\beta)} + \langle \varphi \rangle_{t,Q_T}^{(\beta/2)}$$

Parabolic Hölder space of order $1+\beta$**:** $C^{(1+\beta)/2,1+\beta}([0,T] \times \mathbb{R}^n)$. For $\varphi \in C^{0,1}([0,T] \times \mathbb{R}^n)$:

$$\|\varphi\|_{(1+\beta)/2,1+\beta} := \|\varphi\|_\infty + \|D_x \varphi\|_\infty + \langle \varphi \rangle_{t,Q_T}^{(1+\beta)/2} + \langle D_x \varphi \rangle_{x,Q_T}^{(\beta)} + \langle D_x \varphi \rangle_{t,Q_T}^{(\beta/2)}$$

Parabolic Hölder space of order $2+\beta$**:** $C^{1+\beta/2,2+\beta}([0,T] \times \mathbb{R}^n)$. For $\varphi \in C^{1,2}([0,T] \times \mathbb{R}^n)$

$$\|\varphi\|_{1+\beta/2,2+\beta} := \|\varphi\|_\infty + \|D_x \varphi\|_\infty + \left\|D_x^2 \varphi\right\|_\infty + \|D_t \varphi\|_\infty + \langle D_x \varphi \rangle_{t,Q_T}^{(1+\beta)/2}$$
$$+ \langle D_t \varphi \rangle_{x,Q_T}^{(\beta)} + \langle D_t \varphi \rangle_{t,Q_T}^{(\beta/2)} + \langle D_x^2 \varphi \rangle_{x,Q_T}^{(\beta)} + \langle D_x^2 \varphi \rangle_{t,Q_T}^{(\beta/2)}$$

We now prove an important property involving the norms defined above. Let us start with this useful result:

Lemma C.1. *If $p \in (0,1)$ and $a, b > 0$ then*

$$a^p b^{1-p} \le \frac{1}{M(p)}\left(\epsilon a + \frac{1}{\epsilon^{p^*}} b\right) \text{ for all } \epsilon > 0.$$

where $p^* = \frac{p}{1-p}$ *and* $M(p) = \left(\frac{p}{1-p}\right)^{1-p} + \left(\frac{1-p}{p}\right)^p$

Proof.
The result can be obtained by minimizing the function $\epsilon \to \frac{1}{M(p)}\left(\epsilon a + \frac{1}{\epsilon^{p^*}} b\right)$.

□

We can now prove the following:

Proposition C.2. *Let $l, \upsilon \in (0,2)$ with $l < \upsilon$ and $l, \upsilon \neq 1$. There exists a constant $M > 0$ only depending on β, n, υ and T such that*

$$\|\varphi\|_{(1+l)/2,1+l} \le \begin{cases} M\left(\epsilon^{(\upsilon-l)/2}\|\varphi\|_{(1+\upsilon)/2,1+\upsilon} + \epsilon^{-(1+l)}\|\varphi\|_\infty\right) & \text{if } l < \upsilon < 1 \\ M\left(\epsilon^{(1-l)/2}\|\varphi\|_{(1+\upsilon)/2,1+\upsilon} + \epsilon^{-(1+l)}\|\varphi\|_\infty\right) & \text{if } l < 1 < \upsilon \\ M\left(\epsilon^{(\upsilon-l)/2}\|\varphi\|_{(1+\upsilon)/2,1+\upsilon} + \epsilon^{-(2+l)}\|\varphi\|_\infty\right) & \text{if } 1 < l < \upsilon \end{cases}$$

for all $\varphi \in C^{(1+\upsilon)/2,1+\upsilon}([0,T] \times \mathbb{R}^n)$ and all $\epsilon \in (0,1)$.

Proof.
To lighten our computations we will write $\| \, \|_l := \| \, \|_{l/2,l}$ when there is no need to highlight both the subscripts.

Take $\epsilon \in (0,1)$ and denote with M a positive constant which may change from line to line and that only depends on l, υ, n and T but not on φ or ϵ.

We distinguish several cases:
Case $l, \upsilon \in (0,1)$. In this case we have

$$\|\varphi\|_{(1+l)/2,1+l} := \|\varphi\|_\infty + \|D_x\varphi\|_\infty + \langle\varphi\rangle^{(1+l)/2}_{t,Q_T} + \langle D_x\varphi\rangle^{(l)}_{x,Q_T} + \langle D_x\varphi\rangle^{(l/2)}_{t,Q_T}$$

Form Lemma C.1 with $p = 1/2$ we obtain

$$\|\varphi\|_\infty \le \|\varphi\|^{1/2}_\infty \|\varphi\|^{1/2}_\upsilon \le M\left(\epsilon\|\varphi\|_{1+\upsilon} + \epsilon^{-1}\|\varphi\|_\infty\right) \tag{C.4}$$

Also

$$\begin{aligned}\langle\varphi\rangle^{((1+l)/2)}_{t,Q_T} &= \sup_{x\in\mathbb{R}^n}\sup_{|t-t'|\le 1}\frac{|\varphi(t,x)-\varphi(t',x)|}{|t-t'|^{(1+l)/2}} \\ &\le \sup_{x\in\mathbb{R}^n}\sup_{|t-t'|\le\epsilon}\frac{|\varphi(t,x)-\varphi(t',x)|}{|t-t'|^{(1+\upsilon)/2}}|t-t'|^{(\upsilon-l)/2} + \sup_{x\in\mathbb{R}^n}\sup_{\epsilon<|t-t'|\le 1}\frac{|\varphi(t,x)-\varphi(t',x)|}{|t-t'|^{(1+l)/2}} \\ &\le \epsilon^{(\upsilon-l)/2}\|\varphi\|_{1+\upsilon} + \epsilon^{-(1+l)/2}\|\varphi\|_\infty\end{aligned} \tag{C.5}$$

Let now

$$\triangle^j[\epsilon,\varphi](t,x) := \frac{\varphi(t,x+\epsilon e_j)-\varphi(t,x)}{\epsilon}, \qquad \triangle[\epsilon,\varphi](t,x) = \left(\triangle^j[\epsilon,\varphi](t,x)\right)_{j\le n}$$

where e_j is the $j-$th element of the canonical base of $\mathbb{R}^n$. It follows, for some $y\in[0,1]^n$,

$$\begin{aligned}|D_x\varphi(t,x)| \le& |D_x\varphi(t,x)-\triangle[\epsilon,\varphi](t,x)| + |\triangle[\epsilon,\varphi](t,x)|\\ \le& |D_x\varphi(t,x)-D_x\varphi(t,x+y\epsilon))| + |\triangle[\epsilon,\varphi](t,x)|\\ \le& \epsilon^v \|\varphi\|_{1+v} + \epsilon^{-1}\|\varphi\|_\infty\end{aligned}$$

By taking the supremum over $(t,x)\in Q_T$

$$\|D_x\varphi\|_\infty \le M\left(\epsilon^v\|\varphi\|_{1+v} + \epsilon^{-1}\|\varphi\|_\infty\right) \tag{C.6}$$

For $\langle D_x\varphi\rangle^{(l)}_{x,Q_T}$ and $\langle D_x\varphi\rangle^{(l/2)}_{x,Q_T}$ we can proceed as in (C.5)

$$\begin{aligned}\langle D_x\varphi\rangle^{(l)}_{x,Q_T} \le& \epsilon^{v-l}\|\varphi\|_{1+v} + \epsilon^{-l}\|D_x\varphi\|_\infty \le M\left(\epsilon^{v-l}\|\varphi\|_{1+v} + \epsilon^{-(l+1)}\|\varphi\|_\infty\right)\\ \langle D_x\varphi\rangle^{(l/2)}_{t,Q_T} \le& \epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{-l/2}\|D_x\varphi\|_\infty \le M\left(\epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{-1-l/2}\|\varphi\|_\infty\right)\end{aligned}$$

where we used (C.6) in the last inequalities. By adding up the above estimations together with (C.4)–(C.5) and (C.6) we obtain

$$\|\varphi\|_{(1+l)/2,1+l} \le M\left(\epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{-(1+l)}\|\varphi\|_\infty\right)$$

Case $l<1<v$. In this case the function φ is twice differentiable w.r.t. x and once w.r.t. t: we can use the same methods by modifying the exponents of ϵ. For (C.5) we shall have

$$\langle\varphi\rangle^{((1+l)/2)}_{t,Q_T} \le \epsilon^{(1-l)/2}\|\varphi\|_{1+v} + \epsilon^{-(1+l)/2}\|\varphi\|_\infty$$

and (C.6) becomes

$$\|D_x\varphi\|_\infty \le M\left(\epsilon\|\varphi\|_{1+v} + \epsilon^{-1}\|\varphi\|_\infty\right)$$

whereas the estimations for $\langle D\varphi\rangle$ are modified into

$$\begin{aligned}\langle D_x\varphi\rangle^{(l)}_{x,Q_T} \le& M\left(\epsilon^{1-l}\|\varphi\|_{1+v} + \epsilon^{-(l+1)}\|\varphi\|_\infty\right)\\ \langle D_x\varphi\rangle^{(l/2)}_{t,Q_T} \le& M\left(\epsilon^{(1-l)/2}\|\varphi\|_{1+v} + \epsilon^{-(1+l/2)}\|D_x\varphi\|_\infty\right)\end{aligned}$$

so in conclusion

$$\|\varphi\|_{(1+l)/2,1+l} \le M\left(\epsilon^{(1-l)/2}\|\varphi\|_{1+v} + \epsilon^{-(1+l)}\|\varphi\|_\infty\right)$$

Case $l\in(1,2)$. Let $r=l-1$ and $u=v-1$. From the definition of the Hölder norm of order $1+l>2$ we have

$$\begin{aligned}\|\varphi\|_{(1+l)/2,1+l} =& \|\varphi\|_{r/2+1,r+2}\\ =& \|\varphi\|_\infty + \|D_t\varphi\|_\infty + \langle D_t\varphi\rangle^{(l-1)}_{x,Q_T} + \langle D_t\varphi\rangle^{((l-1)/2)}_{t,Q_T} + \|D_x\varphi\|_{(1+r)/2+1,r+1}\end{aligned}$$

Remark first that (C.4) still holds true and also

$$\|D_x\varphi\|_\infty \leq M\left(\epsilon\|\varphi\|_{1+v} + \epsilon^{-1}\|\varphi\|_\infty\right)$$

As we did for the derivative w.r.t. x in the previous cases

$$\begin{aligned}
\|D_t\varphi\|_\infty &\leq M\left(\epsilon^{v-1}\|\varphi\|_{1+v} + \epsilon^{-1}\|\varphi\|_\infty\right)\\
\langle D_t\varphi\rangle_{x,Q_T}^{(l-1)} &\leq M\left(\epsilon^{v-l}\|\varphi\|_{1+v} + \epsilon^{1-l}\|D\varphi\|_\infty\right)\\
&\leq M\left((\epsilon^{v-l} + \epsilon^{2-l})\|\varphi\|_{1+v} + \epsilon^{-l}\|\varphi\|_\infty\right)\\
\langle D_t\varphi\rangle_{t,Q_T}^{((l-1)/2)} &\leq M\left(\epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{(1-l)/2}\|D\varphi\|_\infty\right)\\
&\leq M\left((\epsilon^{(v-l)/2} + \epsilon^{(3-l)/2})\|\varphi\|_{1+v} + \epsilon^{-(1+l)/2}\|D\varphi\|_\infty\right)
\end{aligned}$$

where we used the estimation given above on $\|D_x\varphi\|_\infty$. For the last estimation we can use the result given in the first case:

$$\begin{aligned}
\|D\varphi\|_{(1+r)/2+1,r+1} &\leq M\left(\epsilon^{(u-r)/2}\|D\varphi\|_{1+u} + \epsilon^{-(1+r)}\|D\varphi\|_\infty\right)\\
&= M\left(\epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{-l}\|D\varphi\|_\infty\right)
\end{aligned}$$

If we use the estimation on $\|D_x\varphi\|_\infty$ with ϵ^2 we obtain

$$\|D\varphi\|_{(1+r)/2+1,r+1} \leq M\left((\epsilon^{(v-l)/2} + \epsilon^{2-l})\|\varphi\|_{1+v} + \epsilon^{-(2+l)}\|\varphi\|_\infty\right)$$

We can sum up the above estimation to get

$$\|\varphi\|_{(1+l)/2,1+l} \leq M\left(\epsilon^{(v-l)/2}\|\varphi\|_{1+v} + \epsilon^{-(2+l)}\|\varphi\|_\infty\right)$$

which concludes our proof.

□

C.3 Parabolic Hölder spaces of type 2

We now define parabolic Hölder spaces of type 2 on $Q_T := (0,T)\times\mathbb{R}^n$: for $l > 0$ let $C^{0,\lceil l\rceil}(\mathbb{R}^n)$ be the space of functions on Q_T which are continuous up the boundary together with their derivative in the space variable D_x^j of all order $j \leq \lceil l\rceil$ and measurable w.r.t. t. On this space define the norm

$$\|\varphi\|_{l,H} := \sum_{j=0}^{\lceil l\rceil}\sum_{(j)}\left\|D_x^j\varphi\right\|_\infty + \sum_{(\lceil l\rceil)}\langle D_x^{\lceil l\rceil}\varphi\rangle_{x,Q_T}^{(\{l\}^+)} \qquad \text{(C.7)}$$

The parabolic Hölder space of order l is then defined as

$$H^l([0,T]\times\mathbb{R}^n) := C^{0,\lceil l\rceil}(\mathbb{R}^n)\cap\left\{\|\varphi\|_{l,H} < \infty\right\}$$

We also define $H^0([0,T]\times\mathbb{R}^n)$ as the space of bounded and measurable functions on $[0,T]\times\mathbb{R}^n$. These spaces are all Banach spaces equipped with their respective norms. For more details see for example Triebel (1992); Gilbarg and Trudinger (2001); Adams and Fournier (2009). In the literature these are also called Hölder-Zygmund spaces.

Remark that $H^l([0,T]\times\mathbb{R}^n) = \mathbb{L}^\infty_m([0,T]\to H^l_e(\mathbb{R}^n))$, the space of bounded and measurable functions taking values in the elliptic Hölder space $H^l_e(\mathbb{R}^n)$. Furthermore in this definition we do not impose any condition on the regularity of t but the measurability with respect to the Borel sets of $[0,T]$, which guarantees that we can always write $\int\varphi(t,x)dt$.

Finally remark also that the parabolic Hölder space of type 1 is only defined for non integer positive l whereas the one of type 2 is defined for any positive l.

We give here the analogous of Proposition C.2:

Proposition C.3. *Let $\beta,\upsilon\in[0,2)$, $\beta<\upsilon$. There exists a constant $M>0$ only depending on β, υ, n and T such that for all $\varphi\in H^{1+\upsilon}([0,T]\times\mathbb{R}^n)$ and all $0<\epsilon<1$*

$$\|\varphi\|_{\beta,H}\le\begin{cases} M\left(\epsilon^{1-\beta}\|\varphi\|_{\upsilon,H}+\epsilon^{-1}\|\varphi\|_\infty\right) & \text{if}\quad \beta<1<\upsilon\\ M\left(\epsilon^{\upsilon-\beta}\|\varphi\|_{\upsilon,H}+\epsilon^{-\max(\beta,1)}\|\varphi\|_\infty\right) & \text{otherwise}\end{cases}$$

Proof.
We can use the same ideas as in the proof of Proposition (C.2). From Lemma C.1 we deduce

$$\|\varphi\|_\infty\le\|\varphi\|_\infty^{1/2}\|\varphi\|_{\upsilon,H}^{1/2}\le M\left(\epsilon\|\varphi\|_{\upsilon,H}+\epsilon^{-1}\|\varphi\|_\infty\right)\tag{C.8}$$

If $\beta<1$ then

$$\begin{aligned}&\sup_{t,z,0<|h|\le1}\frac{|\varphi(t,z+h)-\varphi(t,z)|}{|h|^\beta}\\ \le&\sup_{t,z,0<|h|\le\epsilon}\frac{|\varphi(t,z+h)-\varphi(t,z)|}{|h|^\beta}+\sup_{t,z,\epsilon<|h|\le1}\frac{|\varphi(t,z+h)-\varphi(t,z)|}{|h|^\beta}\\ \le&M\left(\epsilon^{\min(\upsilon,1)-\beta}\sup_{t,0<|h|\le1}\frac{|\varphi(t,z+h)-\varphi(t,z)|}{|h|^{\min(\upsilon,1)}}+\epsilon^{-\beta}\|\varphi\|_\infty\right)\\ \le&M\left(\epsilon^{\min(\upsilon,1)-\beta}\|\varphi\|_{\upsilon,H}+\epsilon^{-\beta}\|\varphi\|_\infty\right)\end{aligned}$$

Using the above estimation and (C.8) we obtain

$$\|\varphi\|_{\beta,H}\le M\left(\epsilon^{\min(\upsilon,1)-\beta}\|\varphi\|_{\upsilon,H}+\epsilon^{-1}\|\varphi\|_\infty\right)$$

If $\beta=1$ we can estimate $\|D\varphi\|_\infty$ since $\upsilon>1$: we can use the same technique as in the proof of Proposition C.2 to get

$$\|D\varphi\|_\infty\le M\left(\epsilon^{\{\upsilon\}^+}\|\varphi\|_{\upsilon,H}+\epsilon^{-1}\|\varphi\|_\infty\right)$$

and then again

$$\|\varphi\|_{1,H} \leq M \left(\epsilon^{\upsilon-1} \|\varphi\|_{\upsilon,H} + \epsilon^{-1} \|\varphi\|_{\infty} \right)$$

since $\{\upsilon\}^+ = \upsilon - 1$. Finally if $\beta > 1$ we can repeat the argument as in the case $\beta < 1$:

$$\langle D\varphi \rangle_{z,Q_T}^{(\beta-1)} \leq M \left(\epsilon^{\upsilon-\beta} \|\varphi\|_{\upsilon,H} + \epsilon^{1-\beta} \|D\varphi\|_{\infty} \right) \leq M \left(\epsilon^{\upsilon-\beta} \|\varphi\|_{\upsilon,H} + \epsilon^{-\beta} \|\varphi\|_{\infty} \right)$$

by using the estimation on $\|D\varphi\|_{\infty}$ so then

$$\|\varphi\|_{\beta,H} \leq M \left(\epsilon^{\upsilon-\beta} \|\varphi\|_{\upsilon,H} + \epsilon^{-\beta} \|\varphi\|_{\infty} \right)$$

which concludes our proof.

□

Appendix D

Itô's formula for pure jump processes

In the proof of Theorem 5.11–5.14, we used Itô's formula for continuously differentiable functions, when (U, Z, P) is a pure jump process. In Theorem 32, Chapter II of Protter (2004), the Itô's formula is stated for twice continuously differentiable functions:

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s^c + \sum_{0<s\leq t} \big(f(X_s) - f(X_{s-}) - \Delta X_s f'(X_{s-})\big)$$

where $f \in \mathcal{C}^2$, X is a real valued semimartingale and $[X,X]^c$ is the continuous part of its quadratic variation. The purpose of this appendix is to prove the above formula under weaker assumptions when the semimartingale X is a pure jump process.

Theorem D.1. *Let X be a $\mathbb{R}^n$-valued semimartingale for which there exists $\eta \in (1,2)$ verifying*

$$\sum_{0<s\leq t:|\Delta X_s|\leq 1} |\Delta X_s|^\eta < \infty, \text{ a.s.}$$

Let also $f : [0,t] \times \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function whose partial derivative $\partial_x f$ satisfies the Hölder condition

$$\sup_{t\leq T} \sup_{|x-y|\leq 1} \frac{|\partial_x f(t,x) - \partial_x f(t,y)|}{|x-y|^{\eta-1+\delta}} < \infty$$

for $\delta > 0$. Then the following formula holds:

$$f(t,X_t) - f(0,X_0) = \int_0^t \partial_t f(s,X_s)ds + \int_{0+}^t \partial_x f(s,X_{s-})dX_s + \sum_{0<s\leq t} (f(s,X_s) - f(s,X_{s-}) - \Delta X_s \partial_x f(s,X_{s-}))$$

Proof.
We follow the proof of Theorem 32, Chapter II in Protter (2004). For sake of

simplicity, we assume that f does not depend on t. Define

$$\mathbb{T} := \{s \in (0,t] \mid \Delta X_s \neq 0\}$$

the set of the jump times of X. For given $0 < \epsilon < 1$ let also $A(\epsilon,\omega)$ and $B(\epsilon,\omega)$ be two subsets of $\mathbb{T}$ verifying

$$1)\, A \cup B = \mathbb{T} \qquad 2) \sum_{s\in B(\epsilon)} |\Delta X_s|^{\eta+\delta'} \leq \epsilon \qquad 3)\, A \quad \text{is finite}$$

where $0 < \delta' < \delta$. Remark in particular that all the jump times corresponding to the big jumps ($|\Delta X_s| > 1$) belong to A. For a partition $0 = T_0^n \leq \cdots \leq T_n^n = t$ verifying $\sup_i |T_i^n - T_{i-1}^n| \to 0$ a.s. when $n \to \infty$, we can write

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_i \left(f(X_{T_{i+T}^n}) - f(X_{T_i^n})\right) \\ &= \sum_{i,A(\epsilon)} \left(f(X_{T_{i+T}^n}) - f(X_{T_i^n})\right) + \sum_{i,B(\epsilon)} \left(f(X_{T_{i+T}^n}) - f(X_{T_i^n})\right) \end{aligned}$$

where $\sum_{i,A}$ stands for the summation over times T_i^n such that $A(\epsilon)\cap(T_i^n, T_{i+1}^n] \neq \emptyset$. Since A is almost surely finite we deduce

$$\lim_{n\to\infty} \sum_{i,A(\epsilon)} \left(f(X_{T_{i+T}^n}) - f(X_{T_i^n})\right) = \sum_{s\in A(\epsilon)} \left(f(X_s) - f(X_{s-})\right), \, a.s.$$

For the summation over the jumps in B, we need a non standard Taylor expansion of f:

$$f(y) = f(x) + \partial_x f(x)(y-x) + R(x,y)$$

where

$$R(y,x) = (y-x)' \int_0^1 \left(\partial_x f(x + \theta(y-x)) - \partial_x f(x)\right) d\theta$$

Since $\partial_x f$ is locally Hölder , then for $|y-x| \leq 1$

$$|R(x,y)| \leq M|x-y|^{\eta+\delta} = Mr(|y-x|)|x-y|^{\eta+\delta'}$$

for some positive constant M, where $r(u) = u^{\delta-\delta'}$. It follows

$$\begin{aligned} \sum_{i,B(\epsilon)} \left(f(X_{T_{i+T}^n}) - f(X_{T_i^n})\right) &= \sum_i \partial_x f(X_{T_i^n}) \left(X_{T_{i+T}^n} - X_{T_i^n}\right) \\ &\quad - \sum_{i,A(\epsilon)} \partial_x f(X_{T_i^n}) \left(X_{T_{i+T}^n} - X_{T_i^n}\right) + \sum_{i,B(\epsilon)} R\left(X_{T_{i+T}^n}, X_{T_i^n}\right) \end{aligned}$$

By letting $n \to \infty$ we obtain

$$\begin{aligned} &\sum_i \partial_x f(X_{T_i^n}) \left(X_{T_{i+T}^n} - X_{T_i^n}\right) \to \int_{0^+}^t \partial_x f(X_{s-}) dX_s, \, a.s \\ &\sum_{i,A(\epsilon)} \partial_x f(X_{T_i^n}) \left(X_{T_{i+T}^n} - X_{T_i^n}\right) \to \sum_{s\in A(\epsilon)} \partial_x f(X_{s-}) \Delta X_s, \, a.s. \end{aligned}$$

whereas, from the estimate on the function R and the definition of the set B

$$\limsup_{n\to\infty} \sum_{i,B(\epsilon)} R\left(X_{T^n_{i+T}}, X_{T^n_i}\right) \le M' r(\epsilon)$$

for some positive constant M'. By adding up all the above terms we deduce that, for any $\epsilon > 0$,

$$f(X_t) - f(X_0) = \int_{0^+}^t \partial_x f(s, X_{s-}) dX_s + \sum_{s\in A(\epsilon)} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})) + M' r(\epsilon)$$

The last thing we need to control is the convergence of the right hand side when $\epsilon \to 0$. The function r goes to zero since $\delta > \delta'$, so the only thing we need to check is the convergence of the above series. Since

$$\left| \sum_{0<s\le t} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})) - \sum_{s\in A(\epsilon)} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})) \right|$$
$$= \left| \sum_{s\in B(\epsilon)} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})) \right| \le M \sum_{s\in B(\epsilon)} |\Delta X_s|^{\eta+\delta'} \le \epsilon$$

then

$$\sum_{i,A(\epsilon)} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})) \overset{\epsilon\to 0}{\to} \sum_{0<s\le t} (f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-}))$$

and the proof is complete if we prove that the above series is absolutely convergent. As pointed our in Theorem 32, Chapter II of Protter (2004), it suffices to prove the convergence for the semimartingale $X\mathbb{1}_{V_k}$, where

$$V_k := \inf\{s > 0, |X_s| \ge k\},\ k > 0$$

We can then assume that the semimartingale X takes values in a compact set of $\mathbb{R}^n$. Since f' is bounded on compact sets we deduce

$$\begin{aligned} &\sum_{0<s\le t} |f(X_s) - f(X_{s-}) - \Delta X_s \partial_x f(X_{s-})| \\ \le& M \left(\sum_{0<s\le t,\, |\Delta X_s|\le 1} |\Delta X_s|^{\eta+\delta} + \sum_{0<s\le t,\, |\Delta X_s|>1} |\Delta X_s| \right) < \infty \end{aligned}$$

which concludes the proof.

□

As a consequence of the above theorem, we can prove the Feynman-Kac formula for pure jump processes:

Corollary D.2. *Let μ, γ and $\nu(dy)$ verify the Assumptions 5.1-*[$\mathbf{C}, \mathbf{I1}$] *and Assumption 7.6-*[$\mathbf{L}, \mathbf{I}$]. *Assume that $\varphi \in H^{\alpha+\delta}([0,T] \times \mathbb{R}^n)$ is the unique solution of*

$$\begin{aligned} & -\frac{\partial \varphi}{\partial t} - \mu \partial_x \varphi - \int \left(\varphi(t, x+\gamma) - \varphi(t,x) - \gamma \partial_x \varphi(t,x)\right) \nu(dy) = F \\ & \varphi(T, .) = G(.) \end{aligned} \tag{D.1}$$

where $F \in H^{\delta}([0,T] \times \mathbb{R}^n)$ and $G \in H^{\alpha+\delta}([0,T] \times \mathbb{R}^n)$. Then

$$\varphi(t,x) = \mathbb{E}\left[G\left(X_T^{t,x}\right) + \int_t^T F\left(s, X_s^{t,x}\right) ds\right] \tag{D.2}$$

where

$$X_s^{t,x} = x + \int_t^s \mu(r, X_r^{t,x}) dr + \int_t^s \int_{\mathbb{R}} \gamma(t, X_{r-}^{t,x}, y) \tilde{J}(dydr)$$

Remark D.3. *The above corollary can be stated under mild assumptions: in particular one can allow non smooth terminal condition G, or unbounded coefficients μ of γ, provided that they are Lipschitz continuous.*

Proof.
Let ψ denote the right hand side of (D.2). The Markov property of the process X yields

$$\psi(t,x) = \mathbb{E}\left[\int_t^{t+h} F\left(s, X_s^{t,x}\right) ds + \psi\left(t+h, X_{t+h}^{t,x}\right)\right]$$

for $h > 0$. As in the proof of Theorem 6.8, we can prove that ψ is the unique viscosity solution of PIDE (D.1). From the uniqueness of the solution of PIDE (D.1), we deduce that $\psi = \varphi$, and then (D.2) holds true.

□

Appendix E

Density of an α-stable Lévy process

In this appendix we want to give some estimations on the density of the Lévy process associated to the operator introduced in (7.16). Let us recall its definition:

$$\mathcal{B}_t^{st}\varphi(z) := \int_{\mathbb{R}} \left(\varphi(t, z+y) - \varphi(t,z) - y\frac{\partial \varphi}{\partial z}(t,z)\mathbb{1}_{\{|y|\leq 1\}} \right) \nu^{st}(y)dy$$

where

$$\nu^{st}(y) := \frac{g(0^+)}{|y|^{1+\alpha}}\mathbb{1}_{\{0<y\}} + \frac{g(0^-)}{|y|^{1+\alpha}}\mathbb{1}_{\{y<0\}}$$

This is the differential operator corresponding to the Lévy process $dL_t := \int y\bar{J}^\alpha(dydt)$, where J^α is a Poisson random measure whose Lévy measure is given by $\nu^{st}(dy)$, $\alpha \in (1,2)$. The characteristic triplet of L is given by $(0, \nu^{st}, c)$, $c := \int_{|y|>1} y\nu^{st}(dy)$. It follows that its characteristic function is $\Phi_t(w) := \mathbb{E}\left[\exp(iwL_t)\right] := \exp(tl(w))$ and it is a well known result that

$$l(w) = -\sigma^\alpha |w|^\alpha \left(1 - i\beta \text{sign}(\omega)\tan\frac{\pi\alpha}{2} + ic\omega\right) \tag{E.1}$$

where

$$\sigma := \left[-\left(g(0^+) + g(0^-)\right)\Gamma(-\alpha)\cos\left(\frac{\pi\alpha}{2}\right)\right]^{1/\alpha}$$
$$\beta := \frac{g(0^+) - g(0^-)}{g(0^+) + g(0^-)}$$

See for example Proposition 28.3 in Sato (1999) or Section 3.7 in Cont and Tankov (2004). It follows that the process L has a infinitely differentiable density which can be expressed in terms of inverse Fourier transform of its characteristic functions:

$$m_t(\xi) := \frac{1}{2\pi}\int e^{-i\xi w}\Phi_t(w)dw \tag{E.2}$$

The objective of this appendix is to prove the following:

Lemma E.1. *Let $m_t(\xi)$, $t \in (0,T]$, be the density of a Lévy process L with characteristic triplet $(0, \nu^{st}, c)$ where $\alpha \in (1,2)$. There exists positive constant $M = M(\alpha, T) > 0$, only depending on α, T and the characteristics of the Lévy process such that*

$$\int_{\mathbb{R}} \left| D_\xi^k m_t(\xi) \right| d\xi \le M t^{-\frac{k}{\alpha}}$$

and

$$\int_{\mathbb{R}} \left| \frac{\partial m}{\partial t}(t,\xi) \right| d\xi \le M t^{-\frac{2}{\alpha}}$$

Proof.
In this proof M represents some positive constant only depending on T, α, k and the characteristics of the Lévy process. It may change from line to line. By changing the variable we can write

$$\int_{\mathbb{R}} \left| D_\xi^k m_t(\xi) \right| d\xi = t^{-\frac{k}{\alpha}} \int_{\mathbb{R}} t^{\frac{k+1}{\alpha}} \left| D_\xi^k m_t(\xi t^{\frac{1}{\alpha}}) \right| d\xi \tag{E.3}$$

Our goal is to prove that $\int_{\mathbb{R}} t^{\frac{k+1}{\alpha}} \left| D_\xi^k m_t(\xi t^{\frac{1}{\alpha}}) \right| d\xi$ is bounded uniformly in t. Cauchy-Schwarz inequality yields

$$\begin{aligned}\int_{\mathbb{R}} t^{\frac{k+1}{\alpha}} \left| D_\xi^k m_t(\xi t^{\frac{1}{\alpha}}) \right| d\xi \le& \left(\int_{\mathbb{R}} \frac{1}{1+\xi^2} d\xi \right)^{1/2} \left(\int_{\mathbb{R}} (1+\xi^2) t^{\frac{2(k+1)}{\alpha}} |D_\xi^k m_t(\xi t^{\frac{1}{\alpha}})|^2 d\xi \right)^{1/2} \\ \le& M \left(\int_{\mathbb{R}} (1+\xi^2) t^{\frac{2(k+1)}{\alpha}} |D_\xi^k m_t(\xi t^{\frac{1}{\alpha}})|^2 d\xi \right)^{1/2}\end{aligned}$$

Also

$$D_\xi^k m_t(\xi t^{\frac{1}{\alpha}}) = \frac{(-i)^k}{2\pi} \int_{\mathbb{R}} w^k e^{-iw\xi t^{\frac{1}{\alpha}}} \Phi_t(w) dw = t^{-\frac{k+1}{\alpha}} \frac{(-i)^k}{2\pi} \int_{\mathbb{R}} w^k e^{-iw\xi} \Phi_t(wt^{-\frac{1}{\alpha}}) dw$$

where Φ is the Fourier transform of L: the $k-$th derivative of m_t is simply the Fourier transform of the function $w \to t^{-\frac{k+1}{\alpha}} (-i)^k w^k \Phi_t(wt^{-\frac{1}{\alpha}})$. Standard properties of the Fourier transform yield

$$\begin{aligned}&\int_{\mathbb{R}} t^{\frac{k+1}{\alpha}} \left| D_\xi^k m_t(\xi t^{\frac{1}{\alpha}}) \right| d\xi \\ \le& M \left(\int_{\mathbb{R}} (1+\xi^2) t^{\frac{2(k+1)}{\alpha}} |D_\xi^k m_t(\xi t^{\frac{1}{\alpha}})|^2 d\xi \right)^{1/2} \\ \le& M t^{\frac{(k+1)}{\alpha}} \left(\int_{\mathbb{R}} \left(t^{-2\frac{k+1}{\alpha}} |\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}})|^2 + t^{-2\frac{k+1}{\alpha}} |\frac{d}{d\xi}(\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}}))|^2 \right) d\xi \right)^{1/2} \\ \le& M \left(\int_{\mathbb{R}} \left(|\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}})|^2 + |\frac{d}{d\xi}(\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}}))|^2 \right) d\xi \right)^{1/2}\end{aligned} \tag{E.4}$$

If we prove that the above is bounded then we are done. From (E.1) it follows

$$\left(\int |\xi^k \Phi_t(t^{-\frac{1}{\alpha}} \xi)|^2 \right)^{1/2} = \left(\int |\xi|^{2k} e^{-2\sigma^\alpha |\xi|^\alpha} \right)^{1/2} \le M \tag{E.5}$$

For the second term, we first remark that

$$\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}}) = \xi^k \exp\left(-\sigma^\alpha |\xi|^\alpha \left(1 - i\beta \tan(\frac{\alpha\pi}{2})\mathrm{sign}(\xi)\right) + ict^{1-1/\alpha}\xi\right)$$

from which we deduce

$$|\frac{d}{d\xi}(\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}}))| \leq M\left(1+\xi^k\right)\exp\left(-\sigma^\alpha|\xi|^\alpha\right)$$

so then

$$\left(\int |\frac{d}{d\xi}(\xi^k \Phi_t(\xi t^{-\frac{1}{\alpha}}))|^2 d\xi\right)^{1/2} \leq M$$

since $1 < \alpha < 2$. We use the above estimations and (E.5) in (E.4) to prove the first part of the Lemma.

For the second part, we can remark that the density function of the Lévy process $L_s^{t,l}$, $\xi \to m(s-t, \xi - l)$, verifies

$$\frac{\partial m}{\partial t}(s-t, \xi - l)$$
$$+ \int \left(m(s-t, \xi - l - y) - m(s-t, \xi - l) - y\frac{\partial m}{\partial \xi}(s-t, \xi - l)\mathbb{1}_{\{|y|\leq 1\}}\right)\nu^{st}(dy) = 0,$$
$$m(0, \xi) = \delta_l(\xi)$$

where $\delta_l(\xi)$ is the Dirac mass at the point l. It follows

$$\left|\frac{\partial m}{\partial t}(t, \xi - l)\right| \leq \frac{1}{2}\int_{|y|\leq 1} |y|^2 \int_0^1 d\theta \left|\frac{\partial^2 m}{\partial \xi^2}(t, \xi - l - \theta y)\right| \nu^{st}(dy)$$
$$+ \int_{|y|>1} |m(t, \xi - l - y) - m(t, \xi - l)|\, \nu^{st}(dy)$$

We now integrate over ξ, and by using the previous estimate, we obtain

$$\int_{\mathbb{R}} \left|\frac{\partial m}{\partial t}(t, \xi)\right| d\xi \leq M t^{-\frac{2}{\alpha}}$$

□

Notations

$(\Omega, \mathcal{F}, \mathbb{P})$	Probability space
$\omega \in \Omega$	Scenario of randomness
$\mathbb{P}$, $\mathbb{Q}$	(eventually signed) probability measures
$\mathbb{Q} \ll \mathbb{P}$	$\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$
$\mathbb{E}^{\mathbb{P}}$	Expectation operator under $\mathbb{P}$
$\mathbb{L}^p(E, \mu)$, $p \geq 1$	The Banach space of p-integrable functions on E with respect to μ
$\delta_x(dy)$	The Dirac distribution with mass at x
W_t, B_t	Standard Brownian motions
$(J, \nu(dy))$, $(N, \nu_n(dy))$	Poisson random measures and their Lévy measures
$\mathbb{L}(S)$	The space of integrands with respect to the semi-martingale S
$C^k(\mathbb{R}^n)$	The space of $\mathbb{R}$-valued continuously differentiable functions up to the order $m \leq k$ taking values in $\mathbb{R}^n$.
$C^{h,k}([0,T] \times \mathbb{R}^n)$	The space of $\mathbb{R}$-valued functions which are continuously differentiable in the space variable up to the order $m \leq k$ and in the time variable up to the order $m \leq h$, taking values in $[0,T] \times \mathbb{R}^n$.
$H^l_e(\mathbb{R}^n)$	The elliptic Hölder space of order $l \in \mathbb{R}^+$
$C^{l/2,l}([0,T] \times \mathbb{R}^n)$, $l \in \mathbb{R}^+ \setminus \{\mathbb{N}\}$	The parabolic Hölder space of type 1 of order l
$H^l([0,T] \times \mathbb{R}^n)$	The parabolic Hölder space of type 2 of order l
$D_t\varphi$ or $\partial\varphi/\partial t$	The derivative w.r.t. t for $\varphi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$.
$D^{\lvert k\rvert}_x\varphi$, $\partial^{\lvert k\rvert}_x\varphi$ or $\partial^{\lvert k\rvert}\varphi/\partial x^{\lvert k\rvert}$	For a multi index $k \in \mathbb{N}^n$, $\lvert k\rvert = k_1 + \cdots + k_n$ and $\varphi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ $D^{\lvert k\rvert}_x\varphi = \partial^{\lvert k\rvert}\varphi/\partial x^{k_1} \cdots \partial x^{k_n}$
$\mathbb{S}_n(\mathbb{R})$	The space of real valued semi-definite positive symmetric matrices
ϑ_h	For $h \geq 0$ it defines a positive locally bounded function such that $\vartheta_h \to 1$ when $h \to 0$
For any $l \geq 0$ $l = \lfloor l \rfloor + \{l\}^-$	where $\{l\}^- \in [0,1)$ and $\lfloor l \rfloor \in \mathbb{N}$
For any $l \geq 0$ $l = \lceil l \rceil + \{l\}^+$	where $\{l\}^+ \in (0,1]$ and $\lceil l \rceil \in \mathbb{N}$

Index

Acceptance set, 33

BSDEs with jumps, 115

Copulae, 57
CVaR, 36, 60

Density of a α-stable Lévy process, 223
Derivative of flow, 85, 203
Dirichlet boundary condition, 178, 183
Doléans-Dade Exponential, 199
Dynamic programming principle, 98

Electricity markets, 155
Entropy, 37, 51

Future contracts, 158

G-divergence, 37, 61
Galtchouk-Kunita-Watanabe decomposition, 68
Generalized inverse distribution, 35, 56
Green's function, 187

Hölder spaces, 211
Hitting time, 189

Linear PDE fundamental solution, 106

Normal Inverse Gaussian process, 22, 171

Optimal quadratic hedge price, 89, 96

Portfolio insurance, 40, 63
Pure investment problem, 69, 73, 74, 124, 147, 179

Quadratic hedge, 68

Risk measures, 31, 41
  spectral, 35
  coherent, 33
  continuous from above, 34
  continuous from below, 34
  convex, 33
  law invariant, 32, 40
  minimal penalty function, 34, 45, 51
  on $\mathbb{L}^p$ spaces, 38
  representation, 33
  spectral, 36, 56

Sobolev space, 116

Utility-based hedging, 68

VaR, 35
Variance-optimal signed martingale measure, 74
Viscosity solution, 96, 140, 150, 183, 191

Bibliography

Adams, R. and J. Fournier (2009). *Sobolev spaces, II edition*, Volume 65. Academic press New York.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical finance 9*(3), 203–228.

Ash, R. and C. Doléans-Dade (2000). *Probability and measure theory*. Academic Press.

Barles, G., R. Buckdahn, and E. Pardoux (1997). Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes 60*(1-2), 57–83.

Barles, G., E. Chasseigne, and C. Imbert (2008). The dirichlet problem for second-order elliptic integro-differential equations. *Indiana University Mathematics Journal 57*(1), 213–146.

Barles, G. and C. Imbert (2008). Second-order elliptic integro-differential equations: viscosity solutions' theory revisited. *25*(3), 567–585.

Barrieu, P. and N. El Karoui (2004). Optimal derivatives design under dynamic risk measures. *351*, 13.

Barrieu, P. and N. El Karoui (2008). Pricing, hedging and optimally designing derivatives via minimization of risk measures. *Indifference Pricing*, 77–141.

Basak, S. and A. Shapiro (2001). Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial Studies 14*(2), 371.

Basel Committee(1996). Overview of the amendment to the capital accord to incorporate market risks. *www.bsi.org*.

Basel Committee(2004). International convergence of capital measurement and capital standards: a revised framework. *www.bsi.org*.

Bensoussan, A. and J. Lions (1984). *Impulse control and quasi-variational inequalities*, Volume 216. Gauthier-Villars Paris.

Biagini, S. and M. Frittelli (2009). On the extension of the namioka-klee theorem and on the fatou property for risk measures. *Optimality and Risk-Modern Trends in Mathematical Finance*, 1–28.

Bion-Nadal, J. (2008). Dynamic risk measures: time consistency and risk measures from bmo martingales. *Finance and Stochastics 12*(2), 219–244.

Bion-Nadal, J. (2009). Time consistent dynamic risk processes. *Stochastic Processes and their Applications 119*(2), 633–654.

Black, F. and A. Perold (1992). Theory of constant proportion portfolio insurance. *Journal of Economic Dynamics and Control 16*(3-4), 403–426.

Bouchard, B., R. Elie, and C. Imbert (2010). Optimal control under stochastic target constraints. *SIAM Journal on Control and Optimization 48*, 5.

Bouchard, B., R. Elie, and N. Touzi (2009). Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization 48*(5), 3123–3150.

Boyle, P. and W. Tian (2007). Portfolio management with constraints. *Mathematical Finance 17*(3), 319–343.

Cannarsa, P. and C. Sinestrari (2004). *Semiconcave functions, Hamilton-Jacobi equations, and optimal control.* Birkhauser.

Černỳ, A. and J. Kallsen (2007). On the structure of general mean-variance hedging strategies. *The Annals of probability 35*(4), 1479–1531.

Černỳ, A. and J. Kallsen (2008). Mean–Variance Hedging and Optimal Investment in Heston's Model with Correlation. *Mathematical Finance 18*(3), 473–492.

Clewlow, L. and C. Strickland (2000). *Energy derivatives: pricing and risk management*, Volume 124. Lacima publications Warwick.

Cont, R. and P. Tankov (2004). *Financial modelling with jump processes.* Chapman & Hall.

Cont, R. and P. Tankov (2009). Constant proportion portfolio insurance in the presence of jumps in asset prices. *Mathematical Finance 19*(3), 379–401.

Cont, R., P. Tankov, and E. Voltchkova (2007). Hedging with options in models with jumps. *Stochastic analysis and applications*, 197–217.

Crandall, M., H. Ishii, and P. Lions (1992). *User's guide to viscosity solutions of second order partial differential equations*, Volume 27. American Mathematical Society.

Crépey, S. and A. Matoussi (2008). Reflected and doubly reflected bsdes with jumps: a priori estimates and comparison. *The Annals of Applied Probability 18*(5), 2041–2069.

Csiszar, I. (1967). On topological properties of f-divergences. *Studia Sci. Math. Hungarica 2*, 329–339.

De Franco, C. and P. Tankov (2011). Portfolio insurance under a risk-measure constraint. *Insurance: Mathematics and Economics 49*(3), 361–370.

De Franco, C., P. Tankov, and X. Warin (2012). Quadratic hedge in electricity markets. *working paper*.

Delbaen, F., P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker (2002). Exponential hedging and entropic penalties. *Mathematical Finance 12*(2), 99–123.

Doléans-Dade, C. (1970). Quelques applications de la formule de changement de variables pour les semimartingales. *Probability Theory and Related Fields 16*(3), 181–194.

Doob, J. (1994). *Measure theory*, Volume 143. Springer.

Ekeland, I., A. Galichon, and M. Henry (2009). Comonotonic measures of multivariate risks. *Mathematical Finance*.

Ekeland, I. and W. Schachermayer (2011). Law invariant risk measures on $\mathbb{R}^n$. *Working paper*.

El Karoui, N., M. Jeanblanc, and V. Lacoste (2005). Optimal portfolio management with american capital guarantee. *Journal of Economic Dynamics and Control 29*(3), 449–468.

El Karoui, N., S. Peng, and M. Quenez (1997). Backward stochastic differential equations in finance. *Mathematical finance 7*(1), 1–71.

El Karoui, N. and R. Rouge (2000). Pricing via utility maximization and entropy. *Mathematical Finance 10*(2), 259–276.

Emmer, S., C. Klüppelberg, and R. Korn (2001). Optimal portfolios with bounded capital at risk. *Mathematical Finance 11*(4), 365–384.

Föllmer, H. and P. Leukert (1999). Quantile hedging. *Finance and Stochastics 3*(3), 251–273.

Föllmer, H. and A. Schied (2004). Stochastic finance. an introduction in discrete time. *de Gruyter Studies in Mathematics*, 648.

Forsyth, P., Y. d'Halluin, and K. Vetzal (2005). Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis 25*(1), 87–112.

Forsyth, P. and G. Labahn (2007). Numerical methods for controlled hamilton-jacobi-bellman pdes in finance. *Journal of Computational Finance 11*(2).

Forsyth, P., J. Wan, and I. Wang (2007). Robust numerical valuation of european and american options under the cgmy process. *Journal of Computational Finance 10*(4), 31.

Friedman, A. (1964). *Partial differential equations of parabolic type*, Volume 196. Prentice-Hall Englewood Cliffs, NJ.

Frittelli, M. and E. Gianin (2004). Dynamic convex risk measures. *Risk measures for the 21st century*, 227–248.

Galtchouk, L. (1976). Représentation des martingales engendrées par un processus à accroissements indépendants (cas des martingales de carré intégrable). *Annales de Henri Poincaré (B) 12*, 199–211.

Geman, H. and A. Roncoroni (2006). Understanding the fine structure of electricity prices. *Journal of Business 79*(3).

Gilbarg, D. and N. Trudinger (2001). *Elliptic partial differential equations of second order.* Springer Verlag.

Gundel, A. and S. Weber (2007). Robust utility maximization with limited downside risk in incomplete markets. *Stochastic Processes and Their Applications 117*(11), 1663–1688.

He, X. and X. Zhou (2010). Portfolio choice via quantiles. *Preprint*.

Hubalek, F., J. Kallsen, and L. Krawczyk (2006). Variance-optimal hedging for processes with stationary independent increments. *The Annals of Applied Probability 16*(2), 853–885.

Ishii, H. (1989). On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic pde's. *Communications on pure and applied mathematics 42*(1), 15–45.

Jacod, J. and A. Shiryaev (2003). *Limit Theorems for Stochastic Processes.* Berlin: Springer.

Jakobsen, E. and K. Karlsen (2005). Continuous dependence estimates for viscosity solutions of integro-pdes. *Journal of Differential Equations 212*(2), 278–318.

Jakobsen, E. and K. Karlsen (2006). A maximum principle for semicontinuous functions applicable to integro-partial differential equations. *NoDEA: Nonlinear Differential Equations and Applications 13*(2), 137–165.

Jeanblanc, M., M. Mania, M. Santacroce, and M. Schweizer (2011). Mean-Variance Hedging via Stochastic Control and BSDEs for General Semimartingales. *Working paper*.

Jensen, R. (1988). The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Archive for Rational Mechanics and Analysis 101*(1), 1–27.

Jin, H. and X. Yu Zhou (2008). Behavioral portfolio selection in continuous time. *Mathematical Finance 18*(3), 385.

Jouini, E., M. Meddeb, and N. Touzi (2004). Vector-valued coherent risk measures. *Finance and Stochastics 8*(4), 531–552.

Jouini, E., W. Schachermayer, and N. Touzi (2006). Law invariant risk measures have the fatou property. *Advances in mathematical economics*, 49–71.

Kaina, M. and L. Rüschendorf (2009). On convex risk measures on l^p-spaces. *Mathematical methods of operations research 69*(3), 475–495.

Kallsen, J. and R. Vierthauer (2009). Quadratic hedging in affine stochastic volatility models. *Review of Derivatives Research 12*(1), 3–27.

Karatzas, I. and S. Shreve (1998). *Methods of mathematical finance.* Springer Verlag.

Kunita, H. and S. Watanabe (1967). On square integrable martingales. *Nagoya Mathematical Journal 30*, 209–245.

Ladyzenskaja, O., V. Solonnikov, and N. Ural'ceva (1967). *Linear and quasilinear equations of parabolic type, Translations of Mathematical Monographs, Vol. 23.*

Laurent, J. and H. Pham (1999). Dynamic programming and mean-variance hedging. *Finance and Stochastics 3*(1), 83–110.

Liese, F. and I. Vajda (1987). Convex statistical distances. *Teubner-Texte zur Mathematik 95. Teubner Verlagsgesellschaft, Leipzig,.*

Lions, P. (1983). Optimal control of diffusion processes and hamilton–jacobi–bellman equations. *Communications in partial differential equations 8*(11), Part I 1101–1134, Part II 1229–1276.

Meyer-Brandis, T. and P. Tankov (2008). Multi-factor jump-diffusion models of electricity prices. *International Journal of Theoretical and Applied Finance 11*(5), 503–528.

Mikulevicius, R. and H. Pragarauskas (2009). On Hölder solutions of the integro-differential Zakai equation. *Stochastic Processes and their Applications 119*(10).

Mikulevicius, R. and H. Pragarauskas (2011). On the cauchy problem for integro differential operators in Hölder classes and the uniqueness of the martingale problem. *preprint, arXiv:*1103.3492.

Pham, H. (1998). Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *Journal of Mathematical Systems, Estimation and Control 8*(1), 1–27.

Pham, H. (2000). On quadratic hedging in continuous time. *Mathematical Methods of Operations Research 51*(2), 315–339.

Pham, H. (2007). *Optimisation et contrôle stochastique appliqués à la finance.* Springer Verlag.

Protter, P. (2004). *Stochastic integration and differential equations.* Springer Verlag.

Rockafellar, R. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of risk 2*, 21–42.

Rogers, L. (2009). Optimal and robust contracts for a risk-constrained principal. *Mathematics and Financial Economics 2*(3), 151–171.

Rong, S. (1997). On solutions of backward stochastic differential equations with jumps and applications. *Stochastic Processes and their Applications 66*(2), 209–236.

Royer, M. (2006). Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic processes and their applications 116*(10), 1358–1376.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge Univ Pr.

Schweizer, M. (2001). A guided tour through quadratic hedging approaches . *In: Option Pricing, Interest Rates and Risk Management (eds: Jouini E, Cvitanic J and Musiela M)*, 538–574.

Triebel, H. (1992). Theory of function spaces II. *Modern Birkhäuser Classics*.