ÉCOLE POLYTECHNIQUE
ParisTech

CMAP

École Doctorale : Mathématiques et Informatique

THÈSE

pour obtenir le titre de

Docteur de l'École Polytechnique

Spécialité : MATHÉMATIQUES APPLIQUÉES

présentée par

Nabil KAZI-TANI

Analysis of Backward SDEs with Jumps and Risk Management Issues

---

Étude des EDS rétrogrades avec sauts et problèmes de gestion du risque

Composition du Jury :

Directrices de Thèse :	Jocelyne BION-NADAL	-	École Polytechnique
	Nicole EL KAROUI	-	Université de Paris 6
Rapporteurs :	Laurent DENIS	-	Université d'Evry Val d'Essone
	Mete SONER	-	ETH Zürich
Examinateurs :	Guillaume GORGE	-	AXA - Belgique
	Etienne PARDOUX	-	Université de Provence
	Nizar TOUZI	-	École Polytechnique

Ecole École Polytechnique
Centre de Mathématiques Appliquées
UMR CNRS 7641
91128 Palaiseau Cedex

Remerciements

Je tiens en premier lieu à remercier mes deux directrices de thèse : Jocelyne Bion-Nadal et Nicole El Karoui, pour leur disponibilité et les précieux conseils qu'elles ont pu me donner. Jocelyne m'a accueilli au CMAP, sa rigueur et ses larges connaissances dans de nombreux domaines des mathématiques restent pour moi un exemple. Nicole a spontanément accepté de m'encadrer et m'a constamment encouragé, avoir eu la chance de travailler à ses côtés m'a permis de faire des progrès considérables et pour tout cela je l'en remercie.

Laurent Denis et Mete Soner ont accepté la lourde tâche de rapporter ce manuscrit et de le lire en détail, leurs remarques ont contribué à son amélioration. En un mot, merci ! (Or in two words for Mete: Thank you!)

Cette thèse doit beaucoup à Guillaume Gorge, qui, il y a trois ans, s'est tout de suite intéressé à mon projet, m'a présenté Nicole et m'a permis grâce au financement Axa de le réaliser dans de bonnes conditions. Je le remercie d'accepter de faire partie de mon jury.
Le lecteur attentif remarquera sans peine que toutes les parties de cette thèse sont inspirées d'articles et d'idées de Nizar Touzi, je lui suis extrêmement redevable et le remercie également de participer à la soutenance.
Etienne Pardoux est l'un des chercheurs ayant initié la théorie des EDSR, c'est donc un honneur pour moi de l'avoir dans mon jury, je le remercie chaleureusement de venir de loin pour assister à ma présentation.

Une grande partie de cette thèse a été écrite en collaboration avec Dylan Possamaï et Chao Zhou. Je garde en mémoire les super moments passés entre Zürich, Metabief, l'IHP, Roscoff, Singapour ou Sydney. Je n'oublie pas non plus la remarque après plusieurs semaines de travail intensif, un long silence et puis... "Plus rien ne marche, qu'est-ce qu'on fait ?"
Merci à Chao et Xiaobo pour leur accueil chaleureux à Singapour.

Le Centre de Mathématiques Appliquées à l'école Polytechnique est un endroit idéal pour travailler. Je remercie l'ensemble des membres du labo et les organisateurs du groupe de travail. Les discussions et échanges avec les autres thésards du CMAP ont toujours été agréables et encourageants, je pense en particulier à Xiaolu Tan, Guillaume Royer, Tarik Ben Zineb, Nicolas Landon et Romain Bompis. Je remercie Trung Lap Nguyen pour ses schémas mathématiques décrivant la stratégie à adopter pendant le tournoi de foot inter-labo, c'est une leçon pour beaucoup d'entraineurs dans le monde !

Je remercie également Nasséra Naar, Nathalie Hurel et Alexandra Noiret pour leur aide indispensable, en particulier pour les déplacements en conférence.

Je suis reconnaissant à Ayman Moussa et Rafik Imekraz pour leurs conseils mathématiques avisés. Je remercie également leurs remerciements de thèse pour m'avoir fait comprendre, à moi aussi, ce qu'est un autocuiseur.

J'ai eu la chance de réaliser une partie de ce travail au sein du Group Risk Management à Axa, un environnement dynamique et stimulant. Merci à Jean-Sébatien Lagacé pour son accueil dans l'équipe

et pour nos très nombreuses discussions. Je remercie les doctorants du GRM pour les bons moments passés ensemble, je pense à Théodora Dupont-Courtade, Madeleine-Sophie Deroche, Christophe Dutang, Fatima Jouad, Lihang Wang et Saad Mouti. Cette thèse n'aurait pas été la même sans les discussions et échanges avec Elisabeth Golovina-Benois, Edern Leroux, Thierry Cohignac, Charles-Erwin Eliachar, Pierre-Edouard Chaix, Mathieu Choux, Julio Theluska ou encore Pierre Folk, je les remercie pour leurs idées, leur bonne humeur et leurs encouragements. Merci également à Valérie Letombe pour son aide précieuse.

J'ai une pensée pour ma famille et mes amis, qui ont toujours su m'encourager dans ce projet.

Je remercie enfin l'inégalité de Jensen, qui a bien voulu avoir le bon goût de tomber dans le sens utile une fois dans les dernières semaines de la thèse.

Étude des EDS rétrogrades avec sauts et problèmes de gestion du risque

Résumé : Cette thèse traite d'une part, de questions de gestion, de mesure et de transfert du risque et d'autre part, de problèmes d'analyse stochastique à sauts avec incertitude de modèle.

Le premier chapitre est consacré à l'analyse des intégrales de Choquet, comme mesures de risque monétaires non nécessairement invariantes en loi. Nous établissons d'abord un nouveau résultat de représentation des mesures de risque comonotones, puis un résultat de représentation des intégrales de Choquet en introduisant la notion de *distorsion locale*. Ceci nous permet de donner ensuite une forme explicite à l'inf-convolution de deux intégrales de Choquet, avec des exemples illustrant l'impact de l'absence de la propriété d'invariance en loi.

Nous nous intéressons ensuite à un problème de tarification d'un contrat de réassurance non proportionnelle, contenant des clauses de reconstitution. Après avoir défini le prix d'indifférence relatif à la fois à une fonction d'utilité et à une mesure de risque, nous l'encadrons par des valeurs facilement implémentables.

Nous passons alors à un cadre dynamique en temps. Pour cela, nous montrons, en adoptant une approche par point fixe, un théorème d'existence de solutions bornées pour une classe d'équations différentielles stochastiques rétrogrades (EDSRs dans la suite) avec sauts et à croissance quadratique. Sous une hypothèse additionnelle classique dans le cadre à sauts, ou sous une hypothèse de convexité du générateur, nous établissons un résultat d'unicité grâce à un principe de comparaison. Nous analysons les propriétés des espérances non linéaires correspondantes. En particulier, nous obtenons une décomposition de Doob-Meyer des surmartingales non-linéaires ainsi que leur régularité en temps. En conséquence, nous en déduisons facilement un principe de comparaison inverse. Nous appliquons ces résultats à l'étude des mesures de risque dynamiques associées, sur une filtration engendrée à la fois par un mouvement brownien et par une mesure aléatoire à valeurs entières, à leur repésentation duale, ainsi qu'à leur inf-convolution, avec des exemples explicites.

La seconde partie de cette thèse concerne l'analyse de l'incertitude de modèle, dans le cas particulier des EDSRs du second ordre avec sauts. Nous imposons que ces équations aient lieu au sens presque-sûr, pour toute une famille non dominée de mesures de probabilités qui sont solution d'un problème de martingales sur l'espace de Skorohod.

Nous étendons d'abord la définition des EDSRs du second ordre, telles que définies par Soner, Touzi et Zhang [112], au cas avec sauts. Pour ce faire, nous démontrons un résultat d'agrégation au sens de [111] sur l'espace des trajectoires càdlàg. Ceci nous permet, entre autres, d'utiliser une version quasi-sûre du compensateur de la mesure des sauts du processus canonique. Nous montrons alors un résultat d'existence et d'unicité pour notre classe d'EDSRs du second ordre. Ces équations sont affectées par l'incertitude portant à la fois sur la volatilité et sur les sauts du processus qui les dirige.

Mots-clés : mesures de risque convexes, intégrales de Choquet, capacités sous-modulaires, invariance en loi, inf-convolution, réassurance non proportionnelle, clauses de reconstitution, tarification par indifférence, équations différentielles stochastiques rétrogrades avec sauts, générateur à croissance quadratique, décomposition de Doob-Meyer non linéaire, principe de comparaison inverse, équations différentielles stochastiques rétrogrades du second ordre, agrégation sur l'espace de Skorohod, analyse stochastique quasi-sûre avec sauts.

Analysis of Backward SDEs with Jumps and Risk Management Issues

Abstract: This PhD dissertation deals with issues in management, measure and transfer of risk on the one hand and with problems of stochastic analysis with jumps under model uncertainty on the other hand.

The first chapter is dedicated to the analysis of Choquet integrals, as non necessarily law invariant monetary risk measures. We first establish a new representation result of convex comonotone risk measures, then a representation result of Choquet integrals by introducing the notion of *local distortion*. This allows us then to compute in an explicit manner the inf-convolution of two Choquet integrals, with examples illustrating the impact of the absence of the law invariance property.

Then we focus on a non-proportional reinsurance pricing problem, for a contract with reinstatements. After defining the indifference price with respect to both a utility function and a risk measure, we prove that is is contained in some interval whose bounds are easily calculable.

Then we pursue our study in a time dynamic setting. We prove the existence of bounded solutions of quadratic backward stochastic differential equations (BSDEs for short) with jumps, using a direct fixed point approach as in Tevzadze [117]. Under an additional standard assumption, or under a convexity assumption of the generator, we prove a uniqueness result, thanks to a comparison theorem. Then we study the properties of the corresponding non-linear expectations, we obtain in particular a non linear Doob-Meyer decomposition for g-submartingales and their regularity in time. As a consequence of this results, we obtain a converse comparison theorem for our class of BSDEs. We give applications for dynamic risk measures and their dual representation, and compute their inf-convolution, with some explicit examples, when the filtration is generated by both a Brownian motion and an integer valued random measure.

The second part of this PhD dissertation is concerned with the analysis of model uncertainty, in the particular case of second order BSDEs with jumps. These equations hold $\mathbb{P}$-almost surely, where $\mathbb{P}$ lies in a wide family of probability measures, corresponding to solutions of some martingale problems on the Skorohod space of càdlàg paths.

We first extend the definition given by Soner, Touzi and Zhang [112] of second order BSDEs to the case with jumps. For this purpose, we prove an aggregation result, in the sense of [111], on the Skorohod space. This allows us to use a quasi-sure version of the canonical process jump measure compensator. Then we prove a wellposedness result for our class of second order BSDEs. These equations include model uncertainty, affecting both the volatility and the jump measure of the canonical process.

Keywords : Convex risk measures, Choquet integrals, submodular capacities, law invariance, inf-convolution, non-proportional reinsurance, reinstatement clause, indifference pricing, backward stochastic differential equations with jumps, quadratic growth generator, non-linear Doob-Meyer decomposition, converse comparison theorem, second order BSDEs, aggregation on Skorohod space, quasi-sure stochastic analysis with jumps.

Contents

1 Introduction générale **1**

1.1 Problèmes récents de mesure et de gestion du risque . . . 1

1.1.1 Des mesures de risque aux intégrales de Choquet . . . 1

1.1.2 Une problématique de tarification en Réassurance . . . 7

1.2 Une étude des EDSRs quadratiques à sauts . . . 11

1.2.1 Préliminaires sur les EDSRs . . . 11

1.2.2 Une approche par point fixe . . . 14

1.2.3 Applications aux mesures de risque dynamiques . . . 18

1.3 Un saut dans les EDSRs du second ordre . . . 21

1.3.1 EDSRs et incertitude de modèle . . . 21

1.3.2 Résultats nouveaux . . . 24

I Risk Measures, BSDEs and Applications **27**

2 Choquet Integrals and Applications **29**

2.1 Introduction . . . 29

2.1.1 Choquet integrals . . . 29

2.1.2 Reinsurance and risk transfer . . . 30

2.2 Preliminaries on Risk Measures . . . 32

2.2.1 Introduction . . . 32

2.2.2 VaR Representation of Convex Risk Measures . . . 36

2.2.3 Optimal risk transfer and inf-convolution . . . 39

2.3 The case of Comonotonic Risk Measures . . . 39

2.3.1 Representation results on L^p . . . 40

2.3.2 A quantile weighting result for general Choquet integrals . . . 41

2.3.3 Inf-Convolution of Choquet Integrals . . . 43

2.3.4 An extension to subadditive comonotonic functionals . . . 48

2.4 Examples . . . 50

2.4.1 AVaR Examples . . . 50

2.4.2 Epsilon-contaminated capacities . . . 51

2.4.3 Volatility uncertainty . . . 54

2.4.4 Involving the Entropic Risk Measure . . . 56

3 A Reinsurance Pricing Problem **61**

3.1 Introduction . . . 61

3.1.1 The collective risk theory . . . 61

3.1.2 Links with the existing literature . . . 62

3.2 Pricing of Layers with Reinstatements . . . 62

3.2.1 The contract payoff . . . 62

3.2.2 The pricing bounds . . . 65

4 Quadratic Backward SDEs with Jumps 69

4.1 Introduction . . . 69
4.2 An existence and uniqueness result . . . 72
4.2.1 Notations . . . 72
4.2.2 Standard spaces and norms . . . 73
4.2.3 A word on càdlàg BMO martingales . . . 74
4.2.4 The non-linear generator . . . 77
4.2.5 A priori estimates . . . 78
4.2.6 Existence and uniqueness for a small terminal condition . . . 82
4.2.7 Existence for a bounded terminal condition . . . 88
4.2.8 A uniqueness result . . . 92
4.2.9 A priori estimates and stability . . . 98
4.3 Quadratic g-martingales with jumps . . . 100
4.3.1 Non-linear Doob Meyer decomposition . . . 103
4.3.2 Upcrossing inequality . . . 109
4.4 Dual Representation and Inf-Convolution . . . 113
4.4.1 Dual Representation of the g-expectation . . . 113
4.4.2 Inf-Convolution of g-expectations . . . 116
4.4.3 Examples of inf-convolution . . . 118

II Model Uncertainty and 2BSDEs with Jumps 125

5 Second Order Backward SDEs with Jumps 127

5.1 Introduction . . . 127
5.2 Issues related to aggregation . . . 129
5.2.1 The stochastic basis . . . 129
5.2.2 The main problem . . . 130
5.2.3 Characterization by martingale problems . . . 131
5.2.4 Notations and definitions . . . 133
5.3 Preliminaries on 2BSDEs . . . 142
5.3.1 The non-linear Generator . . . 142
5.3.2 The Spaces and Norms . . . 145
5.3.3 Formulation . . . 146
5.3.4 Connection with standard backward SDEs with jumps . . . 147
5.3.5 Connection with G-expectations and G-Lévy processes . . . 148
5.4 Uniqueness result . . . 150
5.4.1 Representation of the solution . . . 150
5.4.2 A priori estimates and uniqueness of the solution . . . 154

6 Existence of 2BSDEs with Jumps and Applications 159
6.1 Introduction . . . 159
6.2 A direct existence argument . . . 160
6.2.1 Notations . . . 160
6.2.2 Existence when ξ is in $\mathrm{UC_b}(\Omega)$ . . . 163
6.2.3 Main result . . . 172
6.2.4 An extension of the representation formula . . . 172
6.3 Application to a robust utility maximization problem . . . 175
6.3.1 The Market . . . 175
6.3.2 Solving the optimization problem with a Lipschitz 2BSDEJ . . . 177
6.3.3 A link with a particular quadratic 2BSDEJ . . . 180

A Technical Results for Quadratic BSDEs with Jumps 185

B Technical Results for 2BSDEs with Jumps 187
B.1 The measures $\mathbb{P}^{\alpha,\nu}$ . . . 187
B.2 Other properties . . . 192
B.2.1 Upward directed set . . . 192
B.2.2 L^r-Integrability of exponential martingales . . . 193

Bibliography 199

CHAPTER 1

Introduction générale

Contents

1.1 Problèmes récents de mesure et de gestion du risque 1

1.1.1 Des mesures de risque aux intégrales de Choquet 1

1.1.2 Une problématique de tarification en Réassurance 7

1.2 Une étude des EDSRs quadratiques à sauts 11

1.2.1 Préliminaires sur les EDSRs 11

1.2.2 Une approche par point fixe 14

1.2.3 Applications aux mesures de risque dynamiques 18

1.3 Un saut dans les EDSRs du second ordre 21

1.3.1 EDSRs et incertitude de modèle 21

1.3.2 Résultats nouveaux 24

1.1 Problèmes récents de mesure et de gestion du risque

1.1.1 Des mesures de risque aux intégrales de Choquet

Dans la première partie de cette thèse, nous aurons recours à de nombreuses reprises à des espérances non linéaires, objets qui correspondent, au signe près, soit à des mesures de risque, soit à des fonctions d'utilité monétaires.

Nous commencerons par étudier les mesures de risque, en nous focalisant sur le cas particulier des intégrales de Choquet, pour analyser des problèmes de transferts de risque formulés à l'aide de ces fonctionnelles.

Nous analyserons ensuite en section 1.1.2 un problème de tarification par indifférence, faisant intervenir à la fois une fonction d'utilité monétaire et une mesure de risque.

1.1.1.1 Mesures de risque

Ces dernières années, une attention croissante a été portée aux méthodes permettant d'associer une mesure chiffrée à un risque, ce risque pouvant être issu d'une position financière ou d'aléas assurés par une compagnie d'assurance, ces derniers incluant des risques majeurs variés comme ceux des catastrophes naturelles, des pandémies, ou encore des risques industriels.

Motivés par certaines imperfections de mesures traditionnelles de risque comme la Value-at-Risk (VaR), Artzner et al [2] ont conduit une analyse axiomatique du capital de solvabilité requis et ont introduit la notion de mesure de risque cohérente sur un espace de probabilité.

Dans l'espace $\mathcal{X}$ de toutes les positions financières, il existe un sous-ensemble $\mathcal{A}$ de positions qui sont considérées comme acceptables, par exemple, du point de vue d'une autorité de régulation. Supposons que si une position Y est acceptable et $X \leq Y$ alors X est acceptable. On définit alors le risque $\rho(X)$ comme la somme minimale à adjoindre à la perte X et à investir de façon sûre, pour que X devienne acceptable, on a donc :

$$\rho(X) = \inf\{m \in \mathbb{R} | X - m \in \mathcal{A}\}.$$

Adjoindre $m \in \mathbb{R}$ à X revient à diminuer les pertes du montant deterministe m.

Il s'ensuit que ρ est une application définie sur l'ensemble $\mathcal{X}$ telle que $\rho(X + c) = \rho(X) + c$ pour toute constante c, c'est la propriété de *cash invariance*, appellée également *invariance par translation*. De plus, le risque augmente quand les pertes éventuelles augmentent, $\rho(X) \leq \rho(Y)$ si $X \leq Y$.

Nous avons plus précisément la définition suivante. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité fixé.

Définition .1 *Une application $\rho : \mathcal{X} \to \mathbb{R}$ est dite :*

- *[MO] Monotone si $X \leq Y$ p.s. implique $\rho(X) \leq \rho(Y)$.*
- *[CI] Invariante par translation si $\forall c \in \mathbb{R}, \ \rho(X + c) = \rho(X) + c$.*
- *[CO] Convexe si $\forall \ 0 \leq \lambda \leq 1, \ \rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$.*
- *[PH] Positivement homogène si $\forall \ \lambda \geq 0, \ \rho(\lambda X) = \lambda\rho(X)$.*

Une application vérifiant les propriétés [MO], [CI] est appellée *mesure de risque monétaire*. Si elle satisfait de plus [CO] alors elle est appellée *mesure de risque convexe* et si elle satisfait également [PH] alors nous parlerons de *mesure de risque cohérente*.

Nous faisons l'interprétation *d'orientation* suivante : les variables aléatoires que nous manipulerons représentent des pertes lorsqu'elles prennent des valeurs positives. Nous considérons alors ici des mesures de risque monotones croissantes, qui sont invariantes par translation avec un signe +.

Remarque .1 *L'espérance mathématique relative à une mesure de probabilité $\mathbb{P}$ fixée vérifie les propriétés [MO], [CI], [CO] et [PH], du fait des propriétés de positivité et de linéarité de l'espérance. Une mesure de risque cohérente ρ vérifie les mêmes propriétés que l'espérance, sauf la linéarité. Dans notre contexte, la définition d'espérance non linéaire coïncide avec celle de mesure de risque cohérente.*

Les normes réglementaires européennes, comme Bâle 2, ou Solvabilité 2, préconisent l'utilisation de la "Value at Risk" (VaR) comme mesure de risque. Elle consiste à définir une position X comme acceptable si la probabilité d'une perte est en dessous d'un seuil fixé β, i.e $\mathbb{P}(X > 0) \leq \beta$. La mesure de risque correspondante est alors donnée par :

$$\text{VaR}(X) = \inf\{m \in \mathbb{R} | \mathbb{P}(X - m > 0) \leq \beta\}.$$

C'est donc un quantile de la distribution de X. La VaR est une mesure de risque trés populaire mais qui a l'inconvénient de ne pas être convexe, cela signifie que la VaR peut pénaliser la diversification d'une position bien qu'elle soit désirable.

On définit formellement l'ensemble $\mathcal{A}$ des positions acceptables de la manière suivante :

$$\mathcal{A} := \{X \in \mathcal{X} \,|\, \rho(X) \leq 0\}$$

Si l'ensemble $\mathcal{A}$ est convexe, alors l'application ρ est une *mesure de risque convexe*. Si de plus la classe $\mathcal{A}$ est un cône, alors ρ est une *mesure de risque cohérente*. La propriété de cône de $\mathcal{A}$ signifie qu'une position X reste acceptable si on la multiplie par un facteur positif, même très grand.

Des représentations duales de type "Legendre-Fenchel" ont été démontrées pour les mesures cohérentes ainsi que pour les mesures convexes de risque. Une mesure de risque convexe s'écrira alors :

$$\rho(X) = \sup_{\mathbf{Q} \in \mathcal{M}_1} \{\mathbf{E}_{\mathbf{Q}}[-X] - \alpha(\mathbf{Q})\}$$

où $\mathcal{M}_1$ est l'ensemble de toutes les mesures de probabilité sur $(\Omega, \mathcal{F})$. Une mesure de risque convexe s'exprime donc toujours comme la "pire" des pertes moyennes à laquelle on retranche une pénalité, sur une famille de scénarios.

Ces représentations duales ont été prouvées par Artzner & al [2] dans les cas des mesures de risque cohérentes et par Frittelli et Rosazza Gianin [55] dans le cas convexe. Nous faisons référence à l'ouvrage [54] pour plus de détails sur la théorie des mesures de risque.

1.1.1.2 Inf-convolution et transfert de risque optimal

Supposons que deux agents économiques, notés 1 et 2 pour simplifier, doivent se partager un risque total représenté par une variable aléatoire X. Supposons également que ces agents utilisent respectivement les mesures de risque convexes ρ_1 et ρ_2. Nous supposerons, pour fixer les idées, que les agents en question sont un assureur et un réassureur, mais les résultats énoncés sont évidemment également valables dans un cadre plus général d'un "acheteur" et d'un "émetteur". Le partage optimal sera donné par la résolution de l'inf-convolution des mesures ρ_1 et ρ_2.

Plus précisément, plaçons nous comme dans [5] et [68] dans le cadre d'un modèle principal-agent. L'exposition de l'assureur est modélisée par une variable aléatoire positive X. Si l'assureur souscrit un contrat de réassurance, il prendra en charge les pertes $X - F$ et transfèrera au réassureur une quantité F ; pour cela, il doit s'acquitter d'une prime $\pi(F)$.

Nous supposons que l'assureur minimise son risque, sous la contrainte qu'une transaction ait lieu, il résout donc :

$$\inf_{F,\pi} \{\rho_1(X - F + \pi)\} \text{ sous la contrainte } \rho_2(F - \pi) \leq \rho_2(0) = 0. \tag{1.1.1}$$

La saturation de cette contrainte et l'invariance par translation de ρ_2 nous donnent le prix optimal $\pi = \rho_2(F) - \rho_2(0) = \rho_2(F)$. Il s'agit d'un prix d'indifférence pour le réassureur, c'est-à-dire du prix auquel il est indifférent, du point de vue de la mesure de risque ρ_2, entre entrer et ne pas entrer dans la transaction.

En remplaçant $\pi = \rho_2(F)$ dans (1.1.1) et en utilisant la propriété d'invariance par translation de ρ_1, nous obtenons que le programme de l'assureur est équivalent à

$$\inf_{F} \{\rho_1(X - F) + \rho_2(F)\} =: \rho_1 \square \rho_2(X) \tag{1.1.2}$$

$\rho_1 \square \rho_2$ ainsi défini, est l'inf-convolution des mesures de risque ρ_1 et ρ_2.

Barrieu and El Karoui [6] ont montré que si l'assureur et le réassureur utilisent la même mesure de risque, mais avec des coefficients d'aversion pour le risque différents, alors le partage optimal de risque est *proportionnel*.

Plus précisément, soient ρ une mesure de risque convexe et γ_1, γ_2 deux paramètres réels strictement positifs. Nous définissons ρ_1 et ρ_2 par $\rho_i(X) := \gamma_i \rho(\frac{1}{\gamma_i} X)$. Nous avons alors le résultat suivant :

$$\rho_1 \square \rho_2(X) = \rho_1 \left(\frac{\gamma_2}{\gamma_1 + \gamma_2} X \right) + \rho_2 \left(\frac{\gamma_1}{\gamma_1 + \gamma_2} X \right).$$

En d'autres termes, l'inf-convolution est exacte et est atteinte en $F^* := \frac{\gamma_1}{\gamma_1+\gamma_2} X$. Il s'agit là d'un contrat classique en réassurance, qui prévoit que l'assureur prenne en charge une proportion fixe aX des pertes et transfère au réassureur le risque résiduel $(1-a)X$ avec $0 < a < 1$. Ici $a = \frac{\gamma_2}{\gamma_1+\gamma_2} = \frac{\frac{1}{\gamma_1}}{\frac{1}{\gamma_1}+\frac{1}{\gamma_2}}$. Si le paramètre γ_1 modélise l'aversion pour le risque de l'assureur, $\frac{1}{\gamma_1}$ représente alors la *tolérance* pour le risque ou *l'appétit* au risque de l'assureur. Nous voyons alors que plus la tolérance relative pour le risque d'un agent est grande, plus la part du risque qu'il prend en charge est importante.

Les transferts optimaux ne sont cependant pas toujours de type proportionnel. C'est notamment le cas lorsque l'on relaxe l'hypothèse que les agents utilisent la même mesure de risque avec des coefficients d'aversion pour le risque différents. Cette recherche s'intéresse tout particulièrement aux transferts non proportionnels qui sont notamment explorés dans [68], sous les hypothèses de comonotonie et d'invariance en loi.

Définition .2 *Une mesure de risque ρ est dite invariante en loi si $\rho(X) = \rho(Y)$ pour toutes variables aléatoires X et Y ayant la même distribution.*

Pour définir les mesures de risque comonotones, nous devons d'abord définir des variables comonotones :

Définition .3 *Deux variables aléatoires X et Y sur $(\Omega, \mathcal{F})$ sont dites comonotones si*

$$\forall (\omega, \omega') \in \Omega^2, \ \big(X(\omega) - X(\omega')\big) \big(Y(\omega) - Y(\omega')\big) \geq 0$$

Autrement dit, dans un langage économique, X et Y sont comonotones si on ne peut pas utiliser une variable comme une *couverture* pour l'autre. Denneberg [40] a montré que X et Y sont comonotones si et seulement si il existe une troisième variable aléatoire Z telle que X et Y s'ecrivent comme des fonctions croissantes de Z.

Les variables X_1 et X_2 correspondant aux pertes à la charge de l'assureur et du réassureur dans les contrats usuels de réassurance sont des variables comonotones. A titre d'exemple, $X_1 = aX$ et $X_2 = (1-a)X$ sont comonotones. C'est aussi le cas de $X_1 = X \wedge k$ et $X_2 = (X-k)^+$, où k est un réel.

Définition .4 *Une mesure de risque ρ est dite comonotone si*

$$\rho(X+Y) = \rho(X) + \rho(Y), \ \textit{pour toutes variables aléatoires } X, Y \textit{ qui sont comonotones.}$$

Une mesure comonotone est donc additive pour des variables aléatoires comonotones. L'interprétation économique est la suivante : si les variables X et Y ne peuvent être utilisées comme une couverture l'une pour l'autre, alors combiner X et Y ne doit pas apporter de diversification. Dans ce cas, on exige de la mesure comonotone de ne pas être sous-additive mais simplement additive.

Jouini, Schachermayer et Touzi [68] ont prouvé que le partage de risque optimal entre des agents utilisant des mesures de risque comonotones et invariantes en loi est donné par une somme de contrats non proportionnels. Il s'agit d'une somme de contrats "stop-loss" dans la terminologie actuarielle.

Il est possible de relaxer l'hypothèse de monotonie des mesures de risque dans l'étude des transferts optimaux.

Filipović et Svindland [53] ont montré que lorsque les mesures de risque sont invariantes par translation et invariantes en loi, il existe un partage optimal se faisant par un contrat dont le *payoff* est défini comme une fonction lipschitzienne du risque total X. Les mesures de risque considérées par Filipović et Svindland ne sont pas supposées monotones. Acciaio [1] traite également le problème du transfert de risque optimal pour des fonctionnelles non nécessairement monotones, sans faire l'hypothèse de l'invariance en loi, en utilisant la meilleure approximation monotone d'une mesure de risque non monotone introduite dans [85].

Nous nous référons à Burgert et Ruschendorf [20] et aux références qui y sont mentionnées pour un aperçu de l'importante littérature sur le partage optimal de risque.

1.1.1.3 Résultats nouveaux

Les résultats nouveaux de cette partie de la thèse concernent le cas particulier où ρ est définie par une intégrale de Choquet.

Définition .5 *Une fonction d'ensemble $c : \mathcal{F} \to [0,1]$ sera dite*

1. *Normalisée si $c(\emptyset) = 0$ et $c(\Omega) = 1$.*
2. *Monotone si $c(A) \leq c(B)$ pour tout $A \subset B$, A, $B \in \mathcal{F}$.*
3. *Sous-modulaire si $c(A \cup B) + c(A \cap B) \leq c(A) + c(B)$, $\forall\ A$, $B \in \mathcal{F}$.*

Les fonctions d'ensemble normalisées et monotones sont appelées *capacités.*

Définition .6 *L'intégrale de Choquet d'une variable aléatoire X, par rapport à une capacité c est définie par :*

$$\int X d\,c := \int_{-\infty}^{0} (c(X > x) - 1)\, dx + \int_{0}^{\infty} c(X > x) dx$$

Rappellons que nous avons fixé un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

Notons déjà que si $c = Q$ est une mesure de probabilité sur $(\Omega, \mathcal{F})$ alors l'intégrale de Choquet par rapport à c coïncide avec l'espérance usuelle par rapport à la mesure Q.

Si nous définissons une mesure de risque ρ par $\rho(X) := \int X d\,c$, alors ρ est monotone, invariante par translation et positivement homogène.

D'autre part, si ρ est invariante en loi, alors il existe une fonction ψ définie sur $[0,1]$ telle que la capacité c s'écrive $c(A) = \psi(\mathbb{P}(A))$ pour tout A dans $\mathcal{F}$. Si de plus ρ est convexe, alors la capacité c est sous-modulaire et la fonction ψ est concave. Ces propriétés des intégrales de Choquet sont développées dans [54].

Nous commençons par montrer que toute mesure de risque comonotone et convexe définie sur L^p, qui a des valeurs finies, s'écrit comme une intégrale de Choquet par rapport à une certaine capacité. Ce résultat est montré sans l'hypothèse de convexité dans [54], dans le cas des mesures définies sur L^∞, c'est à dire avec des mesures de risques automatiquement à valeurs finies, et uniformément continues sur L^∞. Nous faisons l'hypothèse supplémentaire de convexité, pour utiliser l'extension par Biagini et Frittelli [10] du théorème de Namioka-Klee : toute mesure de risque $\rho : L^p \to \mathbb{R}$, convexe, monotone et à valeurs finies est continue sur L^p. Ce résultat nous permet d'utiliser le même type de démonstration que [54].

Définition .7 *Une fonction croissante $\psi : [0,1] \mapsto [0,1]$ telle que $\psi(0) = 0$ et $\psi(1) = 1$ est appelée fonction de distorsion.*

Notre résultat le plus utile dans ce premier chapitre peut être formulé ainsi : soit X une variable aléatoire donnée, si ρ est une intégrale de Choquet par rapport à une capacité c, et si la fonction $x \mapsto c(X > x)$ hérite des intervalles de constance de $x \mapsto \mathbb{P}(X > x)$, alors $\rho(X)$ coïncide avec $\rho_{\psi_X}(X)$ où ρ_{ψ_X} est une intégrale de Choquet par rapport à une capacité qui s'écrit comme une distorsion de $\mathbb{P}$.

Autrement dit, même si ρ n'est pas une intégrale de Choquet par rapport à une distorsion de probabilité, il est possible de faire coïncider ρ, en chaque variable X avec une intégrale de Choquet par rapport à une distorsion. Cependant, la distorsion dépend de X et nous l'appellerons *distorsion locale*.

Ce résultat nous permet d'abord d'écrire une mesure de risque comonotone comme intégrale sur $[0,1]$ de quantiles pondérés. Ce type de décomposition avait été obtenu par Kusuoka [76] dans le cas invariant en loi, la pondération des quantiles était alors donnée en fonction de la distorsion. Dans notre cas, la pondération est donnée en fonction de la distorsion locale.

Enfin, ces résultats nous permettent de donner une représentation explicite de l'inf-convolution de deux intégrales de Choquet par rapport à des capacités différentes. Le contrat optimal est donné par une somme de tranches de réassurance et les niveaux des tranches correspondent à des quantiles du risque total. Ceci est obtenu en comparant les distorsions locales sur l'intervalle $[0,1]$. Les points où elles sont égales correspondent à des points d'accord entre les deux agents et c'est en ces points que sont calculés les quantiles permettant de définir les niveaux des tranches optimales.

1.1.2 Une problématique de tarification en Réassurance

1.1.2.1 Motivations

Le deuxième chapitre de cette première partie de thèse traite d'un problème de tarification d'un contrat de réassurance non proportionnelle contenant des reconstitutions.

Commençons par introduire des éléments de contexte et quelques définitions.

Soit u_0 la richesse initiale d'une entreprise d'assurance et soit β son taux instantané de primes reçues. Nous représenterons la situation nette de la compagnie à une date t, $R(t)$, différence entre ce qu'elle doit et ce qui lui est dû, comme suit :

$$R(t) = u_0 + \beta t - \sum_{i=1}^{N_t} X_i$$

où $\{N_t,\ t \geq 0\}$ est un processus de sauts, à valeurs entières, modélisant le nombre de sinistres survenus jusqu'à la date t et $\{X_i,\ i \in \mathbb{N}^*\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées représentant les montants de sinistres. $\sum_{i=1}^{N_t} X_i$ est alors le processus du montant cumulé des sinistres que l'entreprise a réglés jusqu'à la date t.

Un contrat de réassurance de type "Excess-of-Loss", que nous noterons simplement XL, prévoit que le réassureur couvre les pertes liées à chaque sinistre qui sont comprises entre l et $l+m$. l est la *rétention* du contrat et m la *limite*. La valeur m correspond à ce que le réassureur rembourse au maximum à chaque sinistre. Plus précisément, les parts respectives Z_i et C_i à la charge du réassureur et de l'assureur sont données par

$$Z_i = (X_i - l)^+ - (X_i - l - m)^+ = \begin{cases} 0 \text{ si } X_i \leq l \\ X_i - l, \text{ si } l \leq X_i \leq l + m \\ m \text{ si } X_i \geq l + m \end{cases} \tag{1.1.3}$$

et

$$C_i = (X_i \wedge l) + (X_i - l - m)^+ = \begin{cases} X_i \text{ si } X_i \leq l \\ l, \text{ si } l \leq X_i \leq l + m \\ X_i - m \text{ si } X_i \geq l + m \end{cases} \tag{1.1.4}$$

Nous avons bien-sûr $Z_i + C_i = X_i$.

Soit $Z(t)$ la somme totale que devra rembourser le réassureur

$$Z(t) := \sum_{i=1}^{N_t} Z_i$$

Cette valeur est générallement limitée à un montant maximal M. Lorsque M est un multiple du remboursement maximal m de chaque sinistre, c'est-à-dire lorsque $M = (k+1)m$, nous disons que le contrat contient k reconstitutions. Dans ce cas, le remboursement total maximal du réassureur est donné par $R_k(t) := Z(t) \wedge ((k+1)m)$.

La couverture offerte par la j-ième reconstitution est donnée par

$$r_j := \min\left[\max(Z(t) - jm, 0), m\right]$$

Pour acquérir cette couverture, la cédante paye une prime P_j donnée par

$$P_j := c_j p_0 \frac{1}{m} r_{j-1} = c_j p_0 \frac{1}{m} \min\left[\max(Z(t) - (j-1)m, 0), m\right]$$

où p_0 est la prime initiale permettant d'entrer dans la transaction et où les coefficients (c_j) sont compris entre 0 et 1 et spécifiés initialement dans le contrat.
Les reconstitutions successives se paient donc au pro-rata de la partie de la tranche qui a été consommée précedemment.

Prenons un exemple concret, permettant de mieux comprendre le fonctionnement de ces contrats. Supposons que l'entreprise d'assurance achète un contrat XL à maturité un an, comprenant deux reconstitutions et couvrant les pertes comprises entre 100 et 200 ($l = 100$ et $m = 100$). La prime initiale requise par le réassureur s'élève à $p_0 = 10$, la première reconstitution est payable à 80% de la prime initiale et la seconde à 50% ($c_1 = 0.8$ et $c_2 = 0.5$).

Supposons que les quatre premiers sinistres qui se produisent correspondent aux montants suivants :

1. $X_1 = 150$, on a alors $C_1 = 100$ et $Z_1 = 50$. La cédante doit reconstituer la moitié de la tranche, pour cela elle s'acquitte du montant $50 * \frac{1}{100} * 80\% * 10 = 4$.

2. $X_2 = 190$, on a alors $C_1 = 100$ et $Z_1 = 90$. La cédante doit reconstituer 90% de la tranche. Pour cela, elle utilise ce qui reste de la première reconstitution (50) et une partie de la seconde reconstitution (40). Elle paye donc :
$50 * \frac{1}{100} * 80\% * 10 + 40 * \frac{1}{100} * 50\% * 10 = 6$

3. $X_3 = 200$, il faut reconstituer toute la tranche, mais seulement 60% de la seconde reconstitution sont disponibles, la tranche est reconstituée seulement à cette hauteur et la cédante paye pour cela : $60 * \frac{1}{100} * 50\% * 10 = 3$.

4. $X_4 = 200$, la compagnie d'assurance n'était plus couverte que pour les pertes comprises entre 100 et 160, on a donc ici $Z_4 = 60$ et $C_4 = 140$. Toutes les reconstitutions ont été consommées, le contrat s'arrête et la cédante n'est plus réassurée pour le reste de l'année.

La prime totale payée par la cédante s'élève dans cet exemple à $\overline{P} = 10+4+6+3 = 23$. Ce montant aurait pu être plus faible si toutes les reconstitutions n'avaient pas été consommées avant la maturité du contrat.

La prime totale aléatoire payée par la cédante sur la période considérée s'exprime donc comme suit

$$\overline{P}(t) := p_0 \left(1 + \frac{1}{m} \sum_{j=1}^{k} c_j \min\left[\max(Z(t) - (j-1)m, 0), m\right] \right)$$

1.1.2.2 Bornes de prix d'indifférence

Nous proposons une méthodologie permettant à la compagnie d'assurance de calculer des bornes de prix d'indifférence pour les contrats avec reconstitutions. Cela fait intervenir à la fois une fonction d'utilité et une mesure de risque de la cédante.

Uniquement pour les besoins de cette section, nous faisons l'hypothèse d'*orientation* opposée à celle qui a été prise dans le chapitre précedent sur les intégrales de Choquet. En effet, il sera plus simple de manipuler ici des variables aléatoires qui représentent des gains nets plutôt que des pertes. Une mesure de risque *monétaire* vérifiera donc les deux propriétés suivantes

- $X \leq Y$ p.s. implique $\rho(X) \geq \rho(Y)$.
- $\forall c \in \mathbb{R},\ \rho(X+c) = \rho(X) - c$.

Les propriétés de convexité et de positive homogéneité ne sont pas modifiées.

Comme cela est expliqué dans le premier chapitre, une mesure de risque ρ est utilisée par la compagnie d'assurance pour calculer son montant minimal de réserves $\rho(X)$. Si on définit une fonction U par $U(X) := -\rho(X)$, où ρ est une mesure de risque convexe, on obtient alors une fonction d'utilité croissante et concave, qui vérifie la propriété d'invariance par translation avec un signe + :

- $X \leq Y$ p.s. implique $U(X) \leq U(Y)$.
- $\forall c \in \mathbb{R},\ U(X+c) = U(X) + c$.

L'ensemble des positions acceptables $\mathcal{A}_U$, du point de vue de la fonction d'utilité U est constitué des variables qui donnent une utilité positive

$$\mathcal{A}_U := \{X \mid U(X) \geq 0\}.$$

Cependant, si agent économique utilise une mesure ρ pour évaluer ses risques, rien ne l'oblige à utiliser la fonction $U = -\rho$ pour évaluer ses gains. U peut généralement être différente de $-\rho$.

Le prix d'indifférence d'une tranche de réassurance, avec ou sans reconstitutions, relatif à une fonction d'utilité monétaire U, est le prix auquel la compagnie d'assurance est indifférente entre acheter ou ne pas acheter la tranche du point de vue de la fonction U.

Cependant, l'achat de la tranche a également un impact sur les réserves minimales de l'entreprise d'assurance. Supposons que le coût du capital est donné par un taux $0 < \overline{c} < 1$ et notons R la situation nette de la cédante, sans une tranche XL donnée et R^{XL} le gain net avec la tranche. Nous définissons alors la prime initiale d'indifférence $\widetilde{P}_0$ relative au couple (U, ρ) comme la solution de l'équation

$$U\left[R(t) - \overline{c}\,\rho(R(t))\right] = U\left[R^{XL}(t) - \overline{c}\,\rho(R^{XL}(t))\right] \tag{1.1.5}$$

Ce prix d'indifférence tient donc compte à la fois de la modification de l'utilité du fait de l'ajout de la tranche et de la modification des réserves, par le biais du coût du capital.

Dans le cas d'une tranche XL avec reconstitutions, la prime totale payée par la compagnie d'assurance pendant la durée de vie du contrat est une variable aléatoire, on ne peut alors pas utiliser l'invariance par translation de U et de ρ pour isoler $\widetilde{P}_0$ dans l'équation (1.1.5).

Cependant nous parvenons, en utilisant la concavité de U et la convexité de ρ, à borner le prix d'indifférence $\widetilde{P}_0$ par deux valeurs implémentables p_1 et p_2. Les prix obtenus, pour des exemples classiques de mesures de risque et de fonctions d'utilité sont proches des prix observés dans les échanges réels de ce type de contrats de réassurance.

Notons que lorsque $\bar{c} = 0$ et $U(X) = \mathbb{E}^{\mathbb{P}}[X]$, alors les bornes sont égales et le prix ainsi obtenu correspond à la *prime pure* dans la terminologie actuarielle, c'est-à-dire au prix égal à la moyenne de ce qui sera à la charge du réassureur. Sundt [114] a obtenu une formule pour la prime pure dans le cas de contrats avec reconstitution ; cette formule coincide avec la notre dans ce cas particulier. Dans [114] et [119], des formules sont également proposées pour d'autres principes de calcul de primes, comme le principe de l'écart-type ou celui de la PH-transform. Par analogie avec le principe de la prime pure, Mata [86] propose un prix de contrats avec reconstitutions qui doit être solution de l'équation $\rho(-\overline{P}(t)) = \rho(-R_k(t))$. Dans le cas particulier où $\bar{c} = 0$, alors ce prix correspond dans notre cadre à p_2. En effet, l'approximation $\rho(-X) \approx -\rho(X)$ est faite dans [86] en amont de l'équation qui est posée. Si ρ est convexe et normalisée, nous nous contentons de remarquer que $\rho(-X) \geq -\rho(X)$.

En plus de tenir compte de l'effet sur les réserves, à travers le coût du capital, les bornes que nous proposons dépendent également de ce que la compagnie d'assurance a déjà en position.

Pour les exemples numériques, nous effectuons des calculs sur la base de simulations de pertes. Nous renvoyons à l'article [63] pour des approximations de la loi des pertes au delà d'un seuil, lorsque des reconstitutions sont possibles.

1.2 Une étude des EDSRs quadratiques à sauts

L'objectif principal de ce chapitre est l'étude des équations différentielles stochastiques rétrogrades (EDSRs dans la suite) avec sauts dont le générateur est à croissance quadratique. Avant de présenter en détail ce sujet, nous introduisons le cadre classique d'analyse des EDSRs.

1.2.1 Préliminaires sur les EDSRs

1.2.1.1 En filtration continue

Nous nous plaçons sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, dont la filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ est engendrée par un mouvement brownien B à valeurs dans $\mathbb{R}^d$ et qui est augmentée des négligeables pour $\mathbb{P}$. Résoudre une EDSR avec générateur g et condition terminale ξ, consiste alors à trouver un couple (Y, Z) de processus $\mathbb{F}$-adaptés tels que

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \ \mathbb{P} - a.s, \ t \in [0, T] \tag{1.2.1}$$

où les données (g, ξ) de l'équation sont telles que

- $g : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ est une fonction mesurable par rapport à $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}^d)$, où $\mathcal{P}$ désigne la tribu prévisible de $[0, T] \times \Omega$ et où T est un horizon temporel fini et déterministe.
- ξ est une variable aléatoire $\mathcal{F}_T$-mesurable à valeurs dans $\mathbb{R}$.

En s'intéressant aux méthodes de dualité et aux principes du maximum pour des problèmes de contrôle optimal stochastique, Bismut [15] a étudié une équation différentielle stochastique rétrograde avec générateur linéaire. Pardoux et Peng [97] ont ensuite généralisé ces équations au cas d'un générateur non linéaire et lipschitzien et en ont montré l'existence et l'unicité dans le théorème suivant

Théorème .1 (Pardoux et Peng [97]) *Supposons que le générateur g soit une fonction lipschitzienne en (y,z), uniformément en (s,ω) et*

$$\mathbb{E}^{\mathbb{P}}\left[|\xi|^2+\int_0^T |g_s(0,0)|^2\,ds\right]<+\infty.$$

Alors l'EDSR (1.2.1) admet une solution et une seule (Y,Z) telle que Z soit un processus de carré intégrable.

Depuis ce résultat fondamental, une attention croissante a été portée aux EDSRs et à leurs applications, non seulement au contrôle stochastique, mais aussi aux jeux différentiels stochastiques, à l'économie théorique ou aux mathématiques financières.

L'objet de très nombreux articles, sur lesquels nous reviendrons en partie, a été d'affaiblir les conditions portant sur le générateur g pour que l'équation (1.2.1) soit encore bien posée.

1.2.1.2 En filtration discontinue

Supposons maintenant que la filtration de l'espace probabilisé filtré $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ est engendrée par un mouvement brownien B à valeurs dans $\mathbb{R}^d$ et une mesure aléatoire indépendante à valeurs entières $\mu(\omega,dt,dx)$ définie sur $\mathbb{R}^+\times E$ avec compensateur $\lambda(\omega,dt,dx)$, où $E=\mathbb{R}^d\backslash\{0\}$. Nous supposons $\widetilde{\Omega}:=\Omega\times\mathbb{R}^+\times E$ muni de la tribu $\widetilde{\mathcal{P}}:=\mathcal{P}\times\mathcal{B}(E)$, où $\mathcal{P}$ désigne la tribu prévisible sur $\Omega\times\mathbb{R}^+$.

Afin de garantir l'existence du compensateur $\lambda(\omega,dt,dx)$, nous supposons que pour tout A dans $\mathcal{B}(E)$ et tout ω dans Ω, le processus $X_t:=\mu(\omega,A,[0,t])\in\mathcal{A}^+_{loc}$, ce qui signifie qu'il existe une suite croissante de temps d'arrêts (T_n) telle que $T_n\to+\infty$ presque sûrement et les processus arrêtés $X_t^{T_n}$ sont croissants, càdlàg, adaptés et tels que $\mathbb{E}[X_\infty]<+\infty$.

Nous ferons l'hypothèse dans tout le chapitre que λ est absolument continu par rapport à la mesure de Lebesgue dt, i.e. $\lambda(\omega,dt,dx)=\nu_t(\omega,dx)\,dt$. Enfin, nous notons $\widetilde{\mu}$ la mesure des sauts compensée :

$$\widetilde{\mu}(\omega,dx,dt)=\mu(\omega,dx,dt)-\nu_t(\omega,dx)\,dt.$$

Le processus Y défini par l'équation (1.2.1), fournit une généralisation possible de l'espérance conditionnelle de ξ, puisque lorsque g est la fonction nulle, on a $Y_t=\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t]$, et dans ce cas, Z est le processus apparaissant dans la propriété de représentation prévisible de la $(\mathcal{F}_t)$-martingale continue $\{\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t]\,,\,t\geq 0\}$.

Dans le cas d'un espace de probabilité filtré dont la filtration est engendrée à la fois par le mouvement brownien B et la mesure aléatoire μ avec compensateur ν, la représentation des martingales pour $\{\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t]\,,\,t\geq 0\}$ devient

$$\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t]=\int_0^t Z_s dB_s+\int_0^t\int_{\mathbb{R}^d\backslash\{0\}}\psi_s(x)(\mu-\nu)(ds,dx),$$

où Z est un processus prévisible tel que

$$\int_0^T |Z_t|^2 \, dt < +\infty, \ \mathbb{P} - a.s.$$

et $\psi : \widetilde{\Omega} \mapsto \mathbb{R}$ est une fonction prévisible telle que

$$\int_0^T \int_{\mathbb{R}^d \setminus \{0\}} |U_s(x)|^2 \, \nu_t(dx) ds < +\infty, \ \mathbb{P} - a.s.$$

Cela conduit à généraliser de manière naturelle l'équation (1.2.1) au cas à sauts de la manière suivante : nous dirons qu'un triplet (Y, Z, U) de processus $\mathbb{F}$-adaptés, est solution de l'EDSR avec générateur g et condition terminale ξ, si, pour tout $t \in [0, T]$, on a $\mathbb{P}$- presque sûrement

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s, U_s) ds - \int_t^T Z_s dB_s - \int_t^T \int_{\mathbb{R}^d \setminus \{0\}} U_s(x)(\mu - \nu)(ds, dx). \tag{1.2.2}$$

où les données (g, ξ) de l'équation sont telles que

- $g : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \times L^2(\nu) \mapsto \mathbb{R}$ est une fonction mesurable par rapport à $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(L^2(\nu))$.
- ξ est une variable aléatoire $\mathcal{F}_T$-mesurable à valeurs dans $\mathbb{R}$.

Dans ce contexte avec sauts, le générateur g dépend donc à la fois de Z et de U. U joue ici un rôle analogue à celui de la variation quadratique dans le cas continu. Cependant, des différences sont notables, en particulier pour chaque t, U_t est une fonction de E dans $\mathbb{R}$, c'est pourquoi le traitement de la dépendance en u du générateur sera différent du traitement de la dépendance en z dans les hypothèses que nous ferons.

Li et Tang [116] ont été les premiers à prouver l'existence et l'unicité d'une solution à (1.2.2) dans le cas d'un générateur g qui est Lipschitz en (y, z, u), uniformément en t et ω, généralisant ainsi le résultat et la méthode de Pardoux et Peng [97]. Le théorème de Li et Tang est démontré dans le cas où μ est une mesure aléatoire de Poisson, ainsi dans ce cas le compensateur ν ne dépend ni de t, ni de ω.

Hypothèse .1
- *Le générateur* $g : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \times L^2(\nu) \mapsto \mathbb{R}$ *est une fonction* $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(L^2(\nu))$*-mesurable et satisfait*

$$\mathbb{E}\left[\int_0^T |g_t(0,0,0)|^2 \, dt\right] < +\infty$$

- *Il existe une constante* $K > 0$ *telle que*

$$\left|g_t(y, z, u) - g_t(y', z', u')\right| \leq K \left(\left|y - y'\right| + \left|z - z'\right| + \left\|u - u'\right\|\right),$$

pour tout $0 \leq t \leq T$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$, $u, u' \in L^2(\nu)$.

Théorème .2 (Li et Tang [116]) *Supposons que le générateur g satisfasse l'hypothèse .1, il existe alors un unique triplet (Y, Z, U) solution de l'équation (1.2.2) qui vérifie*

$$\|Y\|_{\mathcal{S}^2}^2 := \left\| \sup_{0 \le t \le T} |Y_t| \right\|_{L^2(\mathbb{P})} < +\infty \tag{1.2.3}$$

$$\|Z\|_{\mathbb{H}^2}^2 := \mathbb{E}\left[\int_0^T |Z_t|^2 \, dt\right] < +\infty \tag{1.2.4}$$

$$\|U\|_{\mathbb{J}^2}^2 := \mathbb{E}\left[\int_0^T \int_E U_t(x)^2 \nu(dx) dt\right]^{1/2} < +\infty. \tag{1.2.5}$$

Nous noterons $\mathcal{B}^2 := \mathcal{S}^2 \times \mathbb{H}^2 \times \mathbb{J}^2$ l'ensemble des triplets (Y, Z, U) vérifiant (1.2.3), (1.2.4) et (1.2.5), muni des normes ainsi définies. $\mathcal{B}^2$ est alors un espace de Banach.

1.2.2 Une approche par point fixe

La preuve des théorèmes .1 et .2 se fait selon le schéma simplifié suivant, qui consiste à construire une application ψ de $\mathcal{B}^2$ dans lui-même, de sorte que $(Y, Z, U) \in \mathcal{B}^2$ est solution de l'EDSR (1.2.2) si et seulement si (Y, Z, U) est un point fixe de ψ. Nous esquissons ce schéma dans le cadre avec sauts :

(i) Pour $(y, z, u) \in \mathcal{B}^2$, on définit $(Y, Z, U) = \psi(y, z, u)$ comme étant la solution de l'EDSR

$$Y_t = \xi + \int_t^T g_s(y_s, z_s, u_s) ds - \int_t^T Z_s dB_s - \int_t^T \int_{\mathbb{R}^d \setminus \{0\}} U_s(x)(\mu - \nu)(ds, dx).$$

Le générateur de cette équation étant indépendant de (Y, Z, U), une solution appartenant à $\mathcal{B}^2$ est obtenue par une application du théorème de représentation des martingales.

(ii) En utilisant la formule d'Itô, ainsi que l'hypothèse que g est lipschitzienne, on montre que ψ est une contraction dans l'espace $\mathcal{B}^2$ muni d'une norme à poids bien choisie, équivalente à la norme initiale. ψ admet donc un unique point fixe, fournissant ainsi une solution unique à l'EDSR (1.2.2) dans $\mathcal{B}^2$.

Dans notre cas, où le générateur g est à croissance quadratique, nous suivrons le même type de schéma de preuve par point fixe, avec certaines différences sensibles. Nous étendons en fait le résultat de Tevzadze [117], obtenu dans le cas quadratique en filtration continue. Plus exactement, nous ferons l'hypothèse suivante sur le générateur :

Hypothèse .2 (Croissance quadratique) *Il existe des constantes $(\alpha, \beta, \gamma) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+^*$ telles que pour tout (ω, t, y, z, u),*

$$-\alpha - \beta |y| - \frac{\gamma}{2} |z|^2 - \frac{1}{\gamma} j_t(-\gamma u) \le g_t(\omega, y, z, u) \le \alpha + \beta |y| + \frac{\gamma}{2} |z|^2 + \frac{1}{\gamma} j_t(\gamma u),$$

où pour tout t, la fonction $j_t : L^2(\nu_t) \cap \mathbb{L}^\infty(\nu_t) \mapsto \mathbb{R}$ est définie par

$$j_t(u) := \int_E \left(e^{u(x)} - 1 - u(x)\right) \nu_t(dx).$$

Nous autoriserons, dans le chapitre 4, la constante α à être un processus aléatoire.

Cette inégalité est l'extension naturelle de l'hypothèse de croissance quadratique formulée dans le cas continu lorsque g ne dépend pas de u. En effet, si l'on considère l'EDSR (1.2.2) avec $g_t(y,z,u) = \frac{\gamma}{2}|z|^2 + \frac{1}{\gamma}j_t(\gamma u)$, alors par une application de la formule d'Itô, on a

$$Y_t = \frac{1}{\gamma}\ln\left(\mathbb{E}_t^{\mathbb{P}}\left[\exp(\gamma\xi)\right]\right),\ \mathbb{P}-a.s.$$

et nous retrouvons alors l'expression de la mesure de risque entropique conditionnelle, qui correspond dans le cas continu, à l'EDSR avec générateur $\frac{\gamma}{2}|z|^2$.

Pour que ceci soit cohérent, il faut bien-sûr que la fonction j_t soit bien définie, pour tout t. Remarquons que l'inégalité de Taylor nous garantit que si la fonction $x \mapsto u(x)$ est bornée $d\nu_t$-presque partout, pour tout $0 \leq t \leq T$, alors nous avons

$$0 \leq e^{u(x)} - 1 - u(x) \leq Cu^2(x),\ d\nu_t - p.p. \text{ pour tout } 0 \leq t \leq T$$

et j_t est alors bien définie sur $L^2(\nu_t) \cap \mathbb{L}^\infty(\nu_t)$, pour tout $0 \leq t \leq T$.

Nous utilisons le schéma de preuve par point fixe mentionné plus haut pour montrer l'existence de solutions (Y,Z,U) avec Y borné, les espaces de martingales BMO jouent alors un rôle particulier dans ce cadre. Plus précisément, nous avons $Y \in \mathcal{S}^\infty$ où $\mathcal{S}^\infty$ désigne l'espace des processus Y', $\mathcal{F}_t$ progressivement mesurables, à valeurs réelles, tels que

$$\left\|Y'\right\|_{\mathcal{S}^\infty} := \sup_{0\leq t\leq T}\left\|Y'_t\right\|_\infty < +\infty.$$

Introduisons également les espaces de processus suivants.

Nous noterons BMO l'espace des martingales M càdlàg, de carré intégrable, à valeurs dans $\mathbb{R}^d$, telles que

$$\|M\|_{\text{BMO}} := \sup_{\tau\in\mathcal{T}_0^T}\left\|\mathbb{E}_\tau^{\mathbb{P}}\left[(M_T - M_{\tau^-})^2\right]\right\|_\infty < +\infty,$$

où $\mathcal{T}_0^T$ désigne l'ensemble des $\mathcal{F}_t$ temps d'arrêt à valeurs dans $[0,T]$.

$\mathbb{J}^2_{\text{BMO}}$ désigne l'espace des processus $U : \Omega \times [0,T] \times E$, prévisibles et $\mathcal{E}$-mesurables tels que

$$\|U\|^2_{\mathbb{J}^2_{\text{BMO}}} := \left\|\int_0^\cdot\int_E U_s(x)\widetilde{\mu}(ds,dx)\right\|_{\text{BMO}} < +\infty.$$

$\mathbb{H}^2_{\text{BMO}}$ est l'espace des processus Z $\mathcal{F}_t$-progressivement mesurables à valeurs dans $\mathbb{R}^d$ et tels que

$$\|Z\|^2_{\mathbb{H}^2_{\text{BMO}}} := \left\|\int_0^\cdot Z_s dB_s\right\|_{\text{BMO}} < +\infty.$$

L'utilisation de ces espaces est motivée par le fait que si (Y,Z,U) est une solution de l'EDSR (1.2.2) avec g vérifiant l'hypothèse .2 de croissance quadratique, $Y \in \mathcal{S}^\infty$ et

$$\int_0^T |Z_t|^2\,dt < +\infty \quad \text{et} \quad \int_0^T\int_{\mathbb{R}^d\backslash\{0\}} |U_s(x)|^2\,\nu_t(dx)ds < +\infty,\ \mathbb{P}-a.s.$$

alors $Z \in \mathbb{H}^2_{\text{BMO}}$ et $U \in \mathbb{J}^2_{\text{BMO}} \cap L^\infty(\nu)$.

Nous notons $\mathcal{B} := \mathcal{S}^\infty \times \mathbb{H}^2_{\text{BMO}} \times \mathbb{J}^2_{\text{BMO}} \cap L^\infty(\nu)$.

Revenons maintenant au schéma de preuve mentionné plus haut. Nous remarquons déjà que l'étape (i) nécessite l'application du théorème de représentation des martingales. Dans notre cas, puisque la filtration $\mathbb{F}$ est engendrée à la fois par un mouvement brownien et une mesure aléatoire générale, nous ferons l'hypothèse que la représentation des $\mathbb{F}$-martingales a lieu. Ceci sera vrai dans le cas particulier d'une mesure aléatoire de Poisson d'une part ; nous exhiberons d'autre part dans le chapitre sur les EDSR du second ordre, des exemples de compensateur ν dépendant des variables (t, ω) pour lesquels la représentation a lieu.

Dans l'étape (ii) nous ne pouvons plus utiliser le fait que g est lipschitzienne, mais seulement l'hypothèse .2 de croissance quadratique. Nous ne pouvons alors plus directement garantir que l'application ψ est contractante. Cependant, en se restreignant à une boule $\mathcal{B}_R \subset \mathcal{B}$ de rayon R bien choisi et à une condition terminale de norme assez petite, nous pouvons montrer que $\psi(\mathcal{B}_R) \subset \mathcal{B}_R$ et que ψ est une contraction dans cette boule de rayon R. Ceci nous donne l'existence et l'unicité d'une solution, pour une condition terminale de norme assez petite.

Pour prouver l'existence d'une solution pour une condition terminale bornée de norme quelconque, nous nous inspirons à nouveau de Tevzadze [117]. Nous commençons par décomposer ξ comme suit : $\xi = \sum_{i=1}^n \xi_i$, où chaque variable aléatoire ξ_i a une norme $\mathbb{L}^\infty$ assez petite. Nous avons donc l'existence d'une solution de chaque EDSR avec condition terminale ξ_i. Nous procédons ensuite à un recollement des ces solutions pour obtenir une solution à l'EDSR initiale, avec condition terminale ξ. Ce procédé de recollement utilise des changements absolument continus de probabilité, dont la dérivée de Radon-Nykodim est donnée par une martingale exponentielle de Doléans-Dade faisant intervenir les dérivées en z et en u du générateur. Nous avons donc besoin, pour définir correctement ces changements de probabilité, d'imposer des hypothèses supplémentaires, en particulier que g est C^2 en z et deux fois différentiable au sens de Fréchet en u.

Nous parvenons donc à montrer l'existence d'une solution à l'EDSR 1.2.2, lorsque le générateur g est à croissance quadratique et vérifie certaines conditions sur sa dérivée en z et sa dérivée de Fréchet en u. Ce faisant, nous perdons le caractère unique de la solution. Pour avoir une solution unique, nous utilisons un théorème de comparaison.

Dans ce contexte avec sauts, nous devons encore imposer une hypothèse additionnelle pour montrer un théorème de comparaison. Ce théorème permet de comparer les solutions Y_1 et Y_2 de deux EDSR, lorsque l'on sait comparer les conditions terminales et les générateurs. La preuve de théorème repose encore sur un changement de probabilité, via une exponentielle de Doléans-Dade. Or dans le cas continu, l'exponentielle de Doléans-Dade d'une martingale BMO est une vraie martingale uniformément intégrable, et surtout strictement positive par définition. La définition de l'exponentielle de Doléans-Dade dans le cas avec sauts n'implique pas la positivité. En effet, si M est une martingale locale nulle en 0, la formule exponentielle est définie par

$$\mathcal{E}(M)_t := \exp\left(M_t - \frac{1}{2} < M^c, M^c >_t\right) \prod_{s \le t} (1 + \Delta M_s) \exp(-\Delta M_s).$$

On voit donc que si les sauts de M ne sont pas toujours supérieurs à -1, $\mathcal{E}(M)$ n'est pas nécessairement positive. Pour garantir que le changement de probabilité utile dans la preuve du théorème de comparaison est bien défini, on impose l'hypothèse suivante, introduite par Royer [106] et par Barles, Buckdahn et Pardoux [4] sous la forme d'une égalité.

Hypothèse .3 *Pour tout (y, z, u, u'), il existe un processus (γ_t) prévisible et $\mathcal{E}$-mesurable tel que*

$$g_t(y, z, u) - g_t(y, z, u') \leq \int_E \gamma_t(x)(u - u')(x)\nu_t(dx),$$

et tel qu'il existe des constantes $C_2 > 0$ et $C_1 \geq -1 + \delta$ pour un certain $\delta > 0$, indépendantes de (y, z, u, u') telles que

$$C_1(1 \wedge |x|) \leq \gamma_t(x) \leq C_2(1 \wedge |x|).$$

Cette hypothèse n'est pas seulement artificielle et due à la méthode de démonstration puisque Barles, Buckdahn et Pardoux [4] donnent un contre exemple du principe de comparaison lorsque cette hypothèse n'est pas satisfaite. Cette hypothèse est également suffisante pour obtenir le théorème de comparaison dans notre cadre quadratique.

Ceci nous permet d'obtenir, sous cette hypothèse, l'unicité de la solution à l'EDSR 1.2.2.

Cependant, cette hypothèse n'est pas nécessaire, puisque nous parvenons à montrer un théorème de comparaison dans le cas où le générateur g est convexe en (z, u). Nous prouvons donc un résultat d'unicité dans le cas où le générateur vérifie soit l'hypothèse .3, soit une hypothèse de convexité. Cette dernière présente l'intérêt de garder le caractère quadratique en z et en u, contrairement à l'hypothèse .3 qui implique que le générateur est Lipschitz en u.

1.2.3 Applications aux mesures de risque dynamiques

1.2.3.1 Une décomposition de Doob-Meyer non linéaire

Commençons par une définition générale.

Définition .8 *Soient $\xi \in \mathbb{L}^\infty(\mathbb{P})$ et g tels que l'EDSR avec condition terminale ξ et générateur g admet une solution unique et vérifie un principe de comparaison. Alors pour tout $t \in [0, T]$, on définit la g-espérance conditionnelle de ξ comme suit*

$$\mathcal{E}_t^g(\xi) := Y_t$$

où (Y, Z, U) est solution de l'EDSR

$$Y_t = \xi + \int_t^T g_t(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(ds, dx), \ \ \mathbb{P} - a.s.$$

La théorie des g-espérances insiste donc sur le point de vue des EDSR comme généralisation possible de l'espérance conditionnelle linéaire. Elle a été introduite par Peng [99] comme exemple d'espérance non linéaire. Depuis, de nombreux auteurs ont généralisé ses résultats, les étendant notamment au cas d'un générateur quadratique (Ma et Yao [84]). Une extension au cas d'une filtration discontinue a été obtenue par Royer [106] et par Lin [83]. En particulier, Royer [106] donne des conditions de dominations suffisantes sous lesquelles une espérance non linéaire s'écrit comme une g-espérance.

Il est alors naturel d'utiliser ces espérances non linéaires pour définir des martingales non linéaires de la manière suivante.

Définition .9 *$X \in \mathcal{S}^\infty$ est appelé une g-sousmartingale (resp. surmartingale) si*

$$\mathcal{E}^g_s[X_t] \geq \;\; (resp. \leq) X_s, \;\; \mathbb{P}-a.s., \;\; pour\ tout\ 0 \leq s \leq t \leq T.$$

X est appelé une g-martingale si ce processus est à la fois une g-sousmartingale et une g-surmartingale.

Peng [100] a prouvé un théorème de structure des g-sousmartingales dans le cas continu : toute g-sous martingale se décompose comme la somme d'une g-martingale et d'un processus croissant. Cela s'apparente donc à une décomposition de Doob-Meyer non linéaire.

Nous en donnons ici une extension dans le cadre à sauts avec générateur quadratique. Si Y est une g-sousmartingale avec g soit convexe en (z, u), soit satisfaisant l'hypothèse .3, et, vérifiant de plus une hypothèse de croissance quadratique, alors il existe un processus croissant A, prévisible, nul en 0 et des processus (Z, U) (vérifiant les hypothèses minimales pour que l'équation ci-dessous soit bien définie) tels que pour tout $t \in [0, T]$

$$Y_t = Y_T + \int_t^T g_s(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(ds, dx) - A_T + A_t, \;\; \mathbb{P}-a.s.$$

Au-delà de l'intérêt théorique que ce résultat présente en lui-même, il permet d'obtenir comme corollaire simple un principe de comparaison inverse. Ce principe stipule que s'il est possible de comparer les composantes Y_1 et Y_2 de deux solutions (Y_1, Z_1, U_1) et (Y_2, Z_2, U_2) à deux EDSR avec comme générateurs g_1 et g_2 et même condition terminale, alors nous pouvons également comparer les générateurs. Ce résultat est valide sous les hypothèses sur les fonctions g_i qui permettent d'obtenir la décomposition de Doob-Meyer des g_i-sousmartingales. Nous supposerons de plus que g_i est continue à droite en t. (voir le corollaire 4.3.1 pour un énoncé précis).

1.2.3.2 Représentation duale et inf-convolution

Dans cette sous-partie, insiprés par Barrieu et El Karoui [7], nous adoptons le point de vue des mesures de risque dynamiques sur les opérateurs $(\mathcal{E}^g_t)_{t\geq 0}$.

Certaines propriétés sont inhérentes à $\mathcal{E}^g_t$, tandis que d'autres sont héritées du générateur g. Pour toute fonction g telle que $\mathcal{E}^g_t$ est bien défini, $\mathcal{E}^g_t$ est une mesure de risque dynamique sur $\mathbb{L}^\infty$ qui est monotone et consistante en temps. Plus spécifiquement,

- $\xi_1 \geq \xi_2$, $\mathbb{P}$-presque sûrement implique $\mathcal{E}^g_t(\xi_1) \geq \mathcal{E}^g_t(\xi_1)$, $\mathbb{P}$-presque sûrement, pour tout $0 \leq t \leq T$.

- Pour tous temps d'arrêt bornés $R \leq S \leq T$ et toute variable aléatoire ξ, $\mathcal{F}_T$-mesurable,

$$\mathcal{E}^g_R\left(\mathcal{E}^g_S(\xi)\right) = \mathcal{E}^g_R(\xi), \;\; \mathbb{P}-p.s.$$

$\mathcal{E}^g$ hérite notamment des propriétés suivantes du générateur g :

- Si g ne dépend pas de la variable y, alors $\mathcal{E}^g$ est invariante par translation.
- Si g est convexe en (y,z,u), alors $\mathcal{E}^g$ est convexe.

Ces propriétés permettent de construire de nombreux exemples de mesures de risque. Pour un générateur indépendant de y et convexe en (z,u), $\mathcal{E}^g$ est une mesure de risque dynamique, monétaire, convexe et consistante en temps. Entre autres exemples, le fait de choisir $g(z,u) := \frac{\gamma}{2}|z|^2 + \frac{1}{\gamma}j_t(\gamma u)$, où la fonction j est définie plus haut, nous donne la mesure de risque entropique conditionnelle, sur la filtration engendrée par un mouvement brownien et une mesure aléatoire.

Dans la suite, nous traitons les deux points suivants : d'une part, l'analyse de la forme de la représentation duale de ces mesures de risque dynamiques obtenues via des équations rétrogrades et d'autre part, l'étude de l'inf-convolution de deux mesures de risque de ce type, obtenues via deux générateurs différents.

Sur le premier point, nous étendons la représentation de type Legendre-Fenchel obtenue par Barrieu et El Karoui dans le cas quadratique continu. Nous montrons que, pour une variable ξ_T, $\mathcal{F}_T$-mesurable et bornée, $\mathcal{E}^g_T(\xi_T)$ s'écrit encore comme la "pire" espérance (conditionnelle) moins une pénalité. Cette dernière est obtenue à partir de la transformée de Legendre-Fenchel du générateur. Plus précisément,

$$\mathcal{E}^g_t(\xi_T) = \operatorname*{ess\,sup}_{(\mu,v)\in\mathbb{H}^2_{\mathrm{BMO}}\times\mathcal{A}} \left\{ \mathbb{E}_t^{Q^{\mu,v}}\left[\xi_T - \int_t^T G_s(\mu_s,v_s)ds\right]\right\}$$

où $Q^{\mu,v}$ est la mesure de probabilité définie par

$$\frac{dQ^{\mu,v}}{d\mathbb{P}} = \mathcal{E}\left(\int_0^\cdot \mu_s dB_s + \int_0^\cdot\int_E v_s(x)\tilde{\mu}(ds,dx)\right),$$

où $\mathcal{A}$ désigne l'espace des applications v dans $\mathbb{J}^2_{BMO}\cap\mathbb{L}^\infty(\nu)$ telles qu'il existe $\delta>0$ tel que $v_t(x)\geq -1+\delta$, $\mathbb{P}\times dt\times d\nu_t$ -p.s. et où enfin, G_t est la transformée de Legendre-Fenchel de g définie comme suit

$$G_t(\mu,v) := \sup_{z\in\mathbb{R}^n,\, u\in L^2(\nu_t)} \left\{ <\mu,z>_{\mathbb{R}^n} + <v,u>_{L^2(\nu_t)} - g_t(z,u)\right\},\quad \mu\in\mathbb{R}^n,\ v\in L^2(\nu).$$

Une fois cette représentation duale obtenue, nous nous concentrons sur le problème de l'inf-convolution des mesures $\mathcal{E}^{g^1}$ et $\mathcal{E}^{g^2}$ obtenues à partir de deux générateurs g^1 et g^2. Là encore, nous étendons le résultat valable dans le cas quadratique continu et qui peut se résumer de la manière suivante : l'inf-convolution de $\mathcal{E}^{g^1}$ et $\mathcal{E}^{g^2}$ est encore solution d'une EDSR quadratique dont le générateur est donné par l'inf-convolution des générateurs g^1 et g^2. Donc, plutôt que de chercher une solution dans l'ensemble des variables dans $\mathbb{L}^\infty$, il est possible de se ramener au problème plus simple de l'inf-convolution des générateurs déterministes.

Ceci permet de traiter plusieurs exemples pour lesquels nous donnons une représentation explicite de l'inf-convolution. Comme dans le premier chapitre de cette thèse, cela donne une repérentation du partage de risque optimal, mais cette fois-ci dans un cadre dynamique en temps, entre deux agents utilisant les mesures de risque $\mathcal{E}^{g^1}$ et $\mathcal{E}^{g^2}$. Par comparaison avec les résultats du premier chapitre, nous donnons ici des exemples d'échange qui ont une forme ni linéaire ni non-proportionnelle.

1.3 Un saut dans les EDSRs du second ordre

1.3.1 EDSRs et incertitude de modèle

Toute modélisation mathématique d'un phénomène réel, qu'il soit de nature physique, biologique, économique ou humain constitue intrinsèquement une simplification de la réalité. Tout modèle comporte donc de manière inhérente une erreur, liée à l'écart entre ce que fournit le modèle et la réalité. Est-il alors possible de mesurer le risque d'un écart jugé trop important et, dans quel sens ? Considérer l'incertitude de modèle consiste à proposer des modèles dits "robustes", comprenant leur propre incertitude de paradigme.

Récemment, des contributions importantes ont été apportées dans le cas particulier de la modélisation stochastique, avec incertitude de volatilité (voir les références [100], [37] et [112]). Soner, Touzi et Zhang [112] ont défini les équations différentielles stochastiques rétrogrades du second ordre, en imposant qu'une équation soit vérifiée $\mathbb{P}$-presque sûrement, pour tout $\mathbb{P}$ appartenant à une famille non dominée de mesures de probabilité. Chaque mesure $\mathbb{P}$ représentant un scénario possible de volatilité, la solution est alors robuste en termes d'erreur de modélisation de la volatilité.

Nous montrons dans cette partie comment étendre les travaux de Soner, Touzi et Zhang au cas discontinu, et modéliser ainsi l'incertitude sur la mesure des sauts, à travers des EDSRs du second ordre avec sauts.

Avant de nous intéresser véritablement aux EDSR du second ordre, nous introduisons un problème annexe extrêmement important dans ce cadre, celui de l'agrégation de processus aléatoires.

1.3.1.1 La question de l'agrégation

Une question cruciale dans la définition des EDSR du second ordre (2EDSR dans la suite) donnée dans [112] est celle de l'agrégation de la variation quadratique du processus canonique B sur une large famille de mesures de probabilité.

Soit $\mathcal{P}$ un ensemble non nécessairement dominé de mesures de probabilité et soit $\{X^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$ une famille de processus stochastiques indéxés par $\mathcal{P}$. On peut prendre par exemple une famille d'intégrales stochastiques $X_t^{\mathbb{P}} := {}^{(\mathbb{P})}\int_0^t H_s dB_s$, où $\{H_t, t \geq 0\}$ est un processus prévisible. La dépendance en $\mathbb{P}$ de X vient du fait que l'intégrale d'Itô ${}^{(\mathbb{P})}\int_0^t H_s dB_s$ est définie comme une limite en probabilité sous $\mathbb{P}$.

Le problème de l'agrégation de la famille $\{X^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$ consiste alors à en trouver un représentant universel au sens presque-sûr. Plus précisément,

Définition .10 *Un représentant universel de la famille* $\{X^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$ *est un processus* $\hat{X}$ *tel que*

$$\hat{X} = X^{\mathbb{P}},\ \mathbb{P} - p.s,\ pour\ tout\ \mathbb{P} \in \mathcal{P}.$$

Bichteler [11], Karandikar [69], ou plus récemment Nutz [95], ont montré sous différentes hypothèses qu'il est possible de trouver un représentant universel des intégrales stochastiques ${}^{(\mathbb{P})}\!\int_0^t H_s dB_s$. Une conséquence directe en est la possibilité d'agréger la variation quadratique $\{[B,B]^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$. En effet, grâce à la formule d'Itô, nous pouvons écrire

$$[B,B]_t^{\mathbb{P}} = B_t B_t^T - 2 \cdot {}^{(\mathbb{P})}\!\!\int_0^t B_{s^-} dB_s^T$$

et puisque les intégrales stochastiques peuvent être construites trajectoriellement, nous obtenons l'existence d'un processus $[B,B]$ tel que $[B,B] = [B,B]^{\mathbb{P}}$, $\mathbb{P}$-presque sûrement, pour toute mesure $\mathbb{P}$ dans $\mathcal{P}$.

Afin de définir correctement la notion d'EDSR du second ordre, nous avons besoin d'agréger non seulement la variation quadratique de la partie martingale continue du processus canonique B, mais aussi le compensateur de la mesure des sauts de B. Le compensateur dépend explicitement de la mesure de probabilité sous laquelle il est construit et il n'est pas évident d'en trouver un représentant universel de manière directe, comme pour la variation quadratique. Pour résoudre ce problème, nous restreignons l'ensemble des mesures de probabilité, en suivant la même approche que Soner, Touzi et Zhang [111].

En particulier, nous imposons une forme particulière aux mesures des sauts que nous considérons, forme analogue à celle de la densité des variations quadratiques dans [111]. Partant d'une classe simple de coefficients $\mathcal{A}_0$, nous construisons une autre classe $\mathcal{A}$ de coefficients s'écrivant de la manière suivante :

$$\nu_t(\omega) = \sum_{n=0}^{+\infty} \sum_{i=1}^{+\infty} \nu_t^{n,i}(\omega) \mathbf{1}_{\widetilde{E}_n^i}(\omega) \mathbf{1}_{[\tilde{\tau}_n(\omega), \tilde{\tau}_{n+1})}(t),$$

où les mesures de sauts $(\nu_t^{n,i})$ appartiennent à $\mathcal{A}_0$, $(\tilde{\tau}_n)$ est une suite de temps d'arrêt prenant au plus un nombre dénombrable de valeurs et pour chaque n, $(\widetilde{E}_n^i)$ est une partition de Ω. Nous dirons alors que la classe $\mathcal{A}$ est engendrée par la classe $\mathcal{A}_0$.

Les familles de mesures de probabilité $\mathcal{P}$, le long desquelles nous cherchons à construire des représentants universels, sont composées de solution à des problèmes de martingales. En effet, un problème de martingale nous permet de construire de manière naturelle une mesure de probabilité sous laquelle le processus canonique B a une densité de variation quadratique α et un compensateur de mesures des sauts ν qui sont prescrits. Pour être à même de traiter le problème de l'agrégation, nous devons limiter la forme des coefficients (α, ν) à la forme décrite plus haut.

Nous nous limitons aux coefficients (α, ν) dans $\mathcal{A}_0$ pour lesquels le problème de martingales correspondant a une solution unique. Nous montrons ensuite que ceci implique que le problème des martingales avec des coefficients dans une classe $\mathcal{A}$, qui est engendrée par $\mathcal{A}_0$ a encore une solution unique. Nous prouvons alors que tout procesus, qui satisfait une condition de cohérence minimale, a un représentant universel. Cette condition de cohérence correspond à l'hypothèse de *compatibilité avec la base stochastique* présentée dans le premier chapitre de [45].

1.3.1.2 Volatilité et mesures de sauts incertaines

Dans la continuité des travaux de Soner, Touzi et Zhang dans le cas continu, nous définissons une classe d'EDSR avec sauts, en demandant que ces équations soient vraies $\mathbb{P}$-presque sûrement, pour tout $\mathbb{P}$ dans une famille de mesures de probabilité construite à partir de solutions de problèmes de martingales. Chaque mesure $\mathbb{P}$ représente un scénario de volatilité et de mesures des sauts.

Nous étudions des EDSR du second ordre prenant la forme suivante

$$\begin{aligned} Y_t = \xi + \int_t^T F_s(Y_s, Z_s, U_s, \hat{\alpha}, \hat{\nu})ds - \int_t^T Z_s dB_s^c \\ - \int_t^T \int_E U_s(x)(\mu_{B^d} - \hat{\nu})(ds, dx) + K_T - K_t, \;\; \mathbb{P}^{\alpha,\nu}\text{-p.s }, \forall \mathbb{P}^{\alpha,\nu} \in \mathcal{P}, \end{aligned} \tag{1.3.1}$$

où $(\hat{\alpha}, \hat{\nu}) = (\alpha, \nu), \;\mathbb{P}^{\alpha,\nu}\text{-p.s }, \forall \mathbb{P}^{\alpha,\nu} \in \mathcal{P}$.

Nous mentionnons dans la suite quelques différences fondamentales entre les EDSRs classiques et le cas du second ordre, ainsi que les particularités liées à la présence de sauts.

(i) Premièrement, dans le cas du second ordre avec sauts, nous introduisons une équation similaire à celle du cas classique, mais dont le générateur dépend à la fois de la variation quadratique de B^c d'une part et du compensateur de la mesure des sauts de B^d d'autre part, où B^c et B^d désignent respectivement la partie martingale continue et la partie purement discontinue de B. Rappellons que le générateur d'une EDSR classique ne dépend pas de la variation quadratique du processus canonique, et que dans le cas du second ordre continu, le générateur ne dépend que de la variation quadratique de la partie martingale brownienne. Dans notre cas, puisque nous avons un terme de martingale à sauts additionnel dans l'équation, nous aurons une dépendence additionnelle dans le générateur.

(ii) Une deuxième différence importante avec le cas classique est, comme nous l'avons indiqué plus tôt, que l'EDSR doit être vraie au sens $\mathbb{P}$-presque sûr, pour tout $\mathbb{P}$ dans une large famille de mesures de probabilité. Sous chaque $\mathbb{P}$, B^c et B^d ont respectivement une variation quadratique et un compensateur de mesure des sauts qui sont prescrits. C'est pourquoi nous pouvons comprendre intuitivement les EDSR du second ordre avec sauts comme des rétrogrades classiques avec incertitude de modèle, où l'incertitude affecte à la fois la volatilité et la mesure des sauts du processus qui dirige l'équation.

(iii) Une troisième différence majeure dans le cas du second ordre est la présence d'un processus croissant additionnel K dans l'équation. Notons que la solution Y d'une EDSR du second ordre se représente comme un supremum essentiel de solutions $y^{\mathbb{P}}$ de rétrogrades classiques. Le processus K maintient Y au dessus de toute solution classique $y^{\mathbb{P}}$ avec variation quadratique et compensateur de mesure des sauts donnés sous $\mathbb{P}$. Le processus K est alors formellement analogue au processus croissant apparaissant dans les EDSRs réfléchies (telles que définies dans [46] par exemple).

1.3.2 Résultats nouveaux

1.3.2.1 Existence, unicité et représentation

L'étude de l'équation 1.3.1 que nous menons correspond au cas d'un générateur F qui est Lipschitz en (y, z, u). Nous commençons par montrer que si une solution existe, alors elle s'écrit comme un $\mathbb{P}$-supremum essentiel de solutions d'EDSR classiques, pour chaque $\mathbb{P}$ dans la famille de mesure de probabilités que nous considérons : pour tout $0 \leq t \leq T$ et tout $\mathbb{P}$ dans $\mathcal{P}$,

$$Y_t = \underset{\mathbb{P}' \in \mathcal{P}(t,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}} \left\{ y_t^{\mathbb{P}'}(\xi) \right\}, \ \mathbb{P} - a.s., \tag{1.3.2}$$

où $\mathcal{P}(t, \mathbb{P})$ désigne les probabilités $\mathbb{P}'$ qui coincident avec $\mathbb{P}$ sur $\mathcal{F}_t$ et où $y_t^{\mathbb{P}'}(\xi)$ est la valeur en t de la solution de l'EDSR classique sous $\mathbb{P}'$ avec générateur $\widehat{F}$ et condition terminale ξ, $\widehat{F}$ étant défini par

$$\widehat{F}(y, z, u) := F(y, z, u, \widehat{a}, \widehat{\nu}).$$

Cette représentation est l'analogue du théorème 4.4 [112] dans le cas continu. Elle décrit la solution Y d'une EDSR du second ordre comme le représentant universel des $\mathbb{P}$-supremum essentiels de solutions d'EDSR classiques.

Cette représentation, couplée à certaines estimations a priori, nous permet d'établir que si une solution à notre classe d'EDSR du second ordre existe, alors cette solution est unique.

Pour la preuve d'existence, nous adoptons la même démarche que Soner, Touzi et Zhang [112]. L'outil principal utilisé est la distribution conditionnelle de probabilité régulière (ou regular conditional probability distribution, R.C.P.D) introduite par Stroock et Varadhan [113]. Lorsque le générateur est nul, la composante Y d'une solution d'une EDSR peut être écrite comme une espérance conditionnelle. Par analogie avec ce cas simple, les r.c.p.d nous permettent de construire des solutions d'EDSR trajectoriellement, contournant ainsi les problèmes subtils liés aux ensembles négligeables quand les mesures de probabilité sont singulières.

1.3.2.2 Une application à la maximisation d'utilité en modèle incertain

Nous considérons dans cette section la situation où un agent économique a accès à un marché financier. Il doit trouver une répartition optimale de sa richesse entre un actif non risqué et un actif risqué dont le prix (S_t) est supposé gouverné par une diffusion.

Nous notons π le processus qui à chaque instant représente le montant de capital investi dans l'actif risqué. C'est ce processus π qui est contrôlé, il représente la stratégie de l'agent économique. Le processus de richesse est alors donnée par :

$$X_t^{x,\pi} := x + \int_0^t \pi_u \frac{dS_u}{S_u},$$

où x représente la richesse initiale.

Une stratégie optimale, si elle existe, est obtenue par la maximisation de la fonction d'utilité U de l'agent, supposée croissante et concave. Il s'agit donc de résoudre le problème de maximisation suivant

$$V(x) := \sup_{\pi \in \mathcal{A}} \mathbb{E}^{\mathbb{P}} \left[U \left(X_T^{x,\pi} \right) \right],$$

où $\mathcal{A}$ désigne l'ensemble des stratégies admissibles et T l'horizon du problème.

Dans le cas particulier où U est la fonction d'utilité puissance, S un mouvement brownien géométrique et où aucune contrainte n'est appliquée sur les stratégies possibles, Merton [88] a résolu ce problème de maximisation par des techniques classiques de contrôle stochastique. Ce problème a ensuite progressivement été étendu au cas de fonctions d'utilité générales par Pliska [103], d'ajout de coûts de transaction par Constantinides et Magill [31], de contraintes sur les stratégies admissibles par Cvitanić et Karatzas [34] et Zariphopoulou [121], ou encore au cas de modèles généraux de semi-martingales par Kramkov et Schachermayer [75].

C'est seulement au début des années 2000 qu'un lien est établi entre ce problème de maximisation d'utilité et la théorie des EDSR, dans les travaux de El Karoui et Rouge [51], qui étudient un problème lié à celui de la maximisation d'utilité, dans le cas d'une fonction d'utilité exponentielle et de stratégies admissibles contraintes à demeurer dans un ensemble convexe et fermé. Hu, Imkeller et Müller [] étendent ensuite ces résultats au cas des utilités puissance et logarithme. Dans ces travaux, un lien est établi entre la fonction valeur du problème de maximisation et des EDSR quadratiques.

Une généralisation possible des travaux précédents consiste à introduire de l'incertitude sur le modèle utilisé. Ceci peut être fait en considérant non pas une mesure de probabilité $\mathbb{P}$ de référence comme dans les approches précédentes, mais une famille de mesures de probabilité. Ceci signifie que l'agent ne connait pas de manière parfaite le modèle qui gouverne la dynamique de l'actif risqué S. Dans ce contexte, le problème de maximisation devient le suivant

$$V(x) := \sup_{\pi \in \mathcal{A}} \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}} \left[U \left(X_T^{x,\pi} \right) \right],$$

où $\mathcal{P}$ représente l'ensemble des mesures de probabilité possibles. L'agent se place donc dans la pire des situations pour lui en termes de modèle, puis recherche une stratégie optimale dans ce cadre. Lorsque l'ensemble $\mathcal{P}$ n'est pas dominé, le problème devient techniquement difficile. Denis et Kervarec [38] considèrent une famille de mesures de probabilité correspondant à des modèles de semi-martingales avec des scénarios différents de drift et de volatilité. En supposant alors la fonction d'utilité de l'agent bornée, ils obtiennent d'une part, des résultats de dualité pour les fonctions valeur robustes qu'ils analysent et, d'autre part, l'existence d'une stratégie optimale.

Matoussi, Possamai et Zhou [87] étudient aussi un problème de maximisation d'utilité robuste avec incertitude sur la volatilité. Ils étendent les techniques initiées par El Karoui et Rouge au cadre des martingales non linéaires obtenues naturellement par l'utilisation de toute une famille de mesures de probabilité. Ces travaux établissent donc un lien entre le problème de maximisation d'utilité robuste et une EDSR du second ordre particulière.

Nous étendons un des résultats prouvés dans [87] au cadre avec sauts. En particulier, nous supposons que l'actif risqué S vérifie la dynamique suivante, pour tout $\mathbb{P}$ dans $\mathcal{P}$

$$\frac{dS_t}{S_{t^-}} = b_t dt + dB_t^c + \int_E \beta_t(x) \mu_{B^d}(dt, dx), \quad \mathbb{P} - p.s.$$

où b et β sont des processus prévisibles. Remarquons que l'écriture précédente comporte l'incertitude de modèle sur la diffusion de S puisque $d < B^c, B^c >_t^{\mathbb{P}} = a_t^{\mathbb{P}} dt$, $\mathbb{P}$-p.s. D'autre part, le compensateur de la mesure des sauts du processus canonique change avec la mesure de probabilité considérée. Ceci nous permet donc de modéliser l'incertitude portant à la fois sur la volatilité, sur les instants et sur la taille des sauts.

Dans le cas de la fonction d'utilité exponentielle définie par $U(x) := -\exp(-\eta x)$, avec $\eta > 0$, nous montrons que la fonction valeur V de notre problème de maximisation d'utilité robuste s'écrit $V(x) = -\exp(-\eta x)Y_0$, où (Y, Z, U) est la solution d'une EDSR avec sauts du second ordre particulière, avec générateur Lipschitz. Nous montrons par ailleurs l'existence d'une stratégie π^* optimale. Pour obtenir ce résultat, nous restreignons l'ensemble des stratégies admisibles à un ensemble compact. Ceci nous permet de nous ramener, comme dans [82], à une EDSR avec générateur Lipschitz.

Part I

Risk Measures, BSDEs and Applications

Chapter 2

Choquet Integrals and Applications

Contents

2.1 Introduction . . . **29**

2.1.1 Choquet integrals . . . 29

2.1.2 Reinsurance and risk transfer . . . 30

2.2 Preliminaries on Risk Measures . . . **32**

2.2.1 Introduction . . . 32

2.2.2 VaR Representation of Convex Risk Measures . . . 36

2.2.3 Optimal risk transfer and inf-convolution . . . 39

2.3 The case of Comonotonic Risk Measures . . . **39**

2.3.1 Representation results on L^p . . . 40

2.3.2 A quantile weighting result for general Choquet integrals . . . 41

2.3.3 Inf-Convolution of Choquet Integrals . . . 43

2.3.4 An extension to subadditive comonotonic functionals . . . 48

2.4 Examples . . . **50**

2.4.1 AVaR Examples . . . 50

2.4.2 Epsilon-contaminated capacities . . . 51

2.4.3 Volatility uncertainty . . . 54

2.4.4 Involving the Entropic Risk Measure . . . 56

2.1 Introduction

2.1.1 Choquet integrals

One possible generalization of the Lebesgue integral with respect to a (probability) measure is the Choquet integral with respect to a capacity (definition 2.2.5). Thus, it provides in particular a generalization of the mathematical expectation. It was first introduced by Choquet in 1953 ([29]).

Choquet asked the following natural question: what are the properties of a measure μ that preserve the interior regular approximation property: for each borelian B,

$$\mu(B) = \sup\{\mu(K),\ \text{K compact},\ K \subset B\}.$$

He showed in [29] that additivity properties do not play any role, and that by contrast monotonicity and continuity properties are crucial. That is how the additivity property $\mu(A \cap B) + \mu(A \cup B) =$

$\mu(A) + \mu(B)$ is dropped and replaced by property 3 in definition 2.2.4. Then Choquet introduced a notion of integral with respect to set functions not having additivity properties (see [40] for a monography).

Capacities and Choquet integrals were first developped and used in the potential theory. Then Choquet integrals have been applied intensively as criteria in economic decisions under uncertainty after the demonstrations by Greco [57] and Schmiedler [108] of Choquet integral representations of increasing and comonotonic functionals, under additional continuity assumptions. Since then, an important literature has developped on the study of economic beliefs and preference relations ([56]), their representations using non-additive capacities ([43], [44], [26]) and the equilibria that it generate ([24], [25], [21]).

We adopt in this chapter the point of view of convex risk measures and focus on risk measures that we can represent as Choquet integrals, which are not necessarily law invariant. In [68], Jouini & al. showed that the optimal risk sharing between two agents having law invariant and comonotonic monetary risk measures is given by a sum of stop loss contracts. We extend their result in several directions, since we consider law dependent, and cash sub-additive risk measures, and show that the bounds of the stop loss contracts correspond to quantile values of the total risk. Moreover, we treat the case of L^p random variables, which is more suitable to real cases and theoretical examples than bounded random variables. To achieve this, we first prove a quantile weighting result for general choquet integrals, and using this result, we compute explicitely the inf-convolution of two Choquet integrals and characterize the inf-convolution operator as a Choquet integral with respect to some capacity.

2.1.2 Reinsurance and risk transfer

Reinsurance is a mecanism allowing an insurer to transfer a part of its subscribed risk to a reinsurer.

There are several types of reinsurance contracts. The insurer can transfer its risk in a *proportional* or *linear* way, which means that he gives away a fixed percentage of its losses and of its premiums to the reinsurer: for a loss represented by a positive random variable X, he leaves $F = aX$ with $0 < a < 1$ to the reinsurer and supports $X - F = (1 - a)X$.

The insurer can also buy a *non-proportional* contract, which allows him to put a ceilling on its losses since he will be exposed to $X \wedge k$, the reinsurer taking the remaining part $(X - k)^+$, with $k \in \mathbb{R}^+$.

A very natural question for the insurer is how to optimally transfer its risk ?

Barrieu and El Karoui [6] studied the optimal contract design problem using inf-convolution techniques, in particular they find again in a very simple framework a result of Borch [16], stating that the optimal risk transfer is obtained through a linear quota-share contract. The insurer has to stay exposed to a proportional part of its subscribed risk, with a coefficient equal to its relative risk aversion coefficient over the total risk aversion of the two market players.

Attrition Versus Extreme Risk

In practice, a great majority of major risks reinsurance exchanges are done in a non-proportional way.

To understand this fact in view of the previous proportional optimality result, we can distinguish the risk that the insurer is facing, considering the following two components: an "attrition risk" which is basically a high frequency / low severity risk, and the "extreme risk" that is encountered for example in natural or industrial catastrophes and which is on the contrary a low frequency / high severity risk.

The attrition risk is often considered as the heart of the insurer activity, this is the risk that he knows best and he is in a position to develop a methodology and tools to manage it. In particular, the premium that the insurer receives for the attrition risk is a major tool to manage it, as well as proportional-type reinsurance contracts if the insurer wants to reduce it.

The extreme risk is the one that the insurer really wants to reduce, this risk is considered as the heart of the reinsurer activity. The insurer and reinsurer do not have the same level of information and knowledge on the extreme risk, and this is what justifies the exchange, which is done through non proportional contracts in the majority of cases. See [22].

The criterion choice

A major issue for an insurance company in the design of its reinsurance policy is the choice of its risk retention, that is to say the parameter k appearing in the payoff function $X \wedge k$. In the literature, two main criteria are used: maximizing the cumulative expected discounted dividends (see Taksar [115] for a survey) or minimizing the ruin probability.

De Finetti ([35]) and then many authors have studied the first criterion (among them [109], [61], [118], [60] and the references therein). The goal in these papers is to find, from the point of view of an insurance company, an optimal strategy concerning both the reinsurance buying process and the dividends distribution process. One of their main findings - using stochastic control techniques - is the existence of a feedback optimal strategy for reinsurance on the one hand, and a "barrier" type optimal strategy for dividends on the other hand, in which the insurance company does not pay any dividends when its wealth process exceeds a certain known level. However in practice, dividends decision are often separated from reinsurance decisions, the main issue of dividends being that it sends a strong signal to the insurance industry and to financial markets concerning the company's financial health, making the barrier strategy hardly applicable.

The ruin theory is a standard branch of actuarial science, but it ignores the insurance company economic value, and frictional costs generated by economic capital management. The ruin probability minimization does not appear in the legal solvency II framework. There is a need for insurers to introduce modern risk theory in the reinsurance optimization process.

We focus in this chapter on the non proportional reinsurance optimization, using convex risk measures as a criterion. Indeed, convex risk measures constitute a flexible tool for optimization, and it is suitable to the inf-convolution operator since the inf-convolution of monetary convex risk measures is again a monetary convex risk measure (see [6]). We will discuss the criteria choices, related to the market participants risk aversion.

2.2 Preliminaries on Risk Measures

In this section we will give some definitions, properties and representation of convex risk measures, and then state the reinsurance optimization problem using the introduced convex risk measures as criterion.

2.2.1 Introduction

2.2.1.1 Axioms

Motivated by some imperfections of traditional risk measures such as value-at-risk (which is recommended by the solvency II european regulatory requirement for insurance companies), Artzner & al. [2] and then Follmer and Schied [54] and Frittelli and Rosazza-Gianin [55] introduced the notions of coherent and convex risk measures. We will now recall their definitions, and some key properties. We state the given results for a generic space $\mathcal{X}$ of random variables, typically we will take $\mathcal{X} = L^\infty(\mathbb{P})$ or $\mathcal{X} = L^p(\mathbb{P})$ with $p \geq 1$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a fixed probability space. We will skip the subscript $\mathbb{P}$ in the definition of $\mathcal{X}$ when there is no possible confusion.

Definition 2.2.1 *An application $\rho : \mathcal{X} \to \mathbb{R}$ is called:*

- *[MO] Monotonic if $X \leq Y$ a.s. implies $\rho(X) \leq \rho(Y)$.*
- *[CI] Cash-invariant if $\forall c \in \mathbb{R},\ \rho(X + c) = \rho(X) + c$.*
- *[CO] Convex if $\forall\ 0 \leq \lambda \leq 1,\ \rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$.*
- *[PH] Positively homogeneous if $\forall\ \lambda \geq 0,\ \rho(\lambda X) = \lambda\rho(X)$.*

An application verifying the axioms [MO], [CI] is called a monetary risk measure. If it furthermore satisfies [CO] then it is called a convex risk measure, and if it also satisfies the axiom [PH] then we say it is a coherent risk measure.

The cash-invariance property, also called translation invariance, gives to the value $\rho(X)$ a monetary unit: $\rho(X)$ may be intuitively interpretated as the minimal capital amount to add to X, and to invest in a risk free manner, to make the position X acceptable or risk free. Indeed $\rho(X - \rho(X)) = \rho(X) - \rho(X) = 0$.

The convexity axiom will describe the preference for diversity: it means that diversification should not increase risk.

Note that here positive random variables represent losses, then we consider *increasing* risk measures, and cash-invariant with a "+" sign.

2.2.1.2 Dual representation

An important kind of results are the Fenchel-Legendre type dual representation theorems for risk measures. It has been given in Artzner & al. [2] for coherent risk measures and Frittelli and Rosazza-Gianin [55] for convex risk measures defined on L^∞, and after that by Biagini and Frittelli [10] in the case of convex risk measures defined on L^p.

Since the dual space of L^∞ can be identified with the space M^{ba} of finitely additive set functions with finite total variation, the application of general duality theorems only gives the dual representation with respect to the subspace of finitely additive set functions normalized to 1. For the representation to hold with probability measures, we need the following additional property:

Definition 2.2.2 *ρ satisfies the Fatou property if it is lower semi-continuous with respect to the bounded pointwise convergence, i.e if for any bounded sequence X_n of L^∞ converging pointwise to X, we have*

$$\rho(X) \leq \liminf \rho(X_n).$$

Remark 2.2.1 *The Fatou property is equivalent with continuity from above (see [54]).*

We can now state the representation result:

Theorem 2.2.1 ([55]) *A convex risk measure ρ on L^∞, satisfying the Fatou property, admits the following dual representation:*

$$\rho(X) = \sup_{\mathbf{Q}\in\mathcal{M}_1} \{E^{\mathbf{Q}}[X] - \alpha(\mathbf{Q})\} \tag{2.2.1}$$

where $\mathcal{M}_1$ denotes the set of all probability measures on $\mathcal{F}$ and $\alpha : \mathcal{M}_1 \to \mathbb{R} \cup \{+\infty\}$ is a penalty function given by:

$$\alpha(\mathbf{Q}) = \sup_{X\in L^\infty(\mathbb{P})} \{E^{\mathbf{Q}}[X] - \rho(X)\}$$

If we interpret any probability measure $\mathbf{Q}$ on $(\Omega, \mathcal{F})$ as a possible model, or possible weighting of events, this representation tells us that every convex risk measure that is upper semi-continuous can be seen as a worst case expectation over a family of models minus a penalty, the penalty representing the likelihood of the given models.

In the case of risk measures ρ defined on L^p, we need the additional property of finiteness of ρ on L^p. Indeed, Biagini and Frittelli [10] proved that any convex and finite risk measure on L^p is continuous with respect to the L^p norm. Then a dual representation holds as in (2.2.1) in the classic $L^p - L^q$ duality.

2.2.1.3 Examples

The Value-at-Risk

The Value-at-Risk is the main risk measure used in practice and it is the measure recommended by the Solvency II european regulator for insurance companies. It consists in computing an α-quantile of the risk X:

$$VaR_\alpha(X) := \overline{q}_X(\alpha) = \inf\{x \in \mathbb{R} \text{ such that } \overline{F}_X(x) \leq \alpha\},$$

where $\alpha \in [0,1]$ and $\overline{F}_X$ denotes the tail cumulative distribution function of X. α is considered as an "acceptable" bankruptcy probability, and then $VaR_\alpha(X)$ represents the losses that are attained only with that probability.

The measure $\rho = VaR_\alpha$ is not convex, and then do not satisfy a dual representation, but the following Average Value-at-Risk will.

The Average Value-at-Risk

To overcome the non-convexity of the monetary risk measure VaR_α, we consider the following Average Value-at-Risk at level $\alpha \in (0,1]$ of a position $X \in \mathcal{X}$, given by:

$$AVaR_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \overline{q}_X(u)du$$

The name Average Value-at-Risk is justified by the following equality, which holds if X has a continuous distribution, it says that the quantity $AVaR_\alpha(X)$ is an average of the losses greater or equal to $VaR_\alpha(X)$:

$$AVaR_\alpha(X) = \mathbb{E}\left[X|X \geq VaR_\alpha(X)\right].$$

$AVaR_\alpha$ is a coherent risk measure and its dual representation is given as follows:

$$AVaR_\alpha(X) = \sup_{Q\in\mathcal{M}_1}\{E^Q(X) - \alpha(Q)\},\ X \in \mathcal{X}$$

where the penalty function α only takes the values $+\infty$ or 0, it is given by the indicator function in the sense of convex analysis of the following set:

$$\mathcal{Q} := \{Q \in \mathcal{M}_1 \mid \frac{dQ}{d\mathbb{P}} \leq \frac{1}{\alpha}\}$$

Note that we have:

$$AVaR_\alpha(X) = \overline{q}_X(\alpha) + \frac{1}{\alpha}\mathbb{E}\left[(X - \overline{q}_X(\alpha))^+\right]$$

The Entropic risk measure

As mentionned before, the penalty function α appearing in the representation (2.2.1) account for the likelihood of a given model $\mathbf{Q}$. Suppose that we measure the likelihood of $\mathbf{Q}$ using the information theoretic entropic function $H(\mathbf{Q}|\mathbb{P})$ defined by

$$H(\mathbf{Q}|\mathbb{P}) := \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbf{Q}}{d\mathbb{P}}\log(\frac{d\mathbf{Q}}{d\mathbb{P}})\right].$$

The corresponding risk measure

$$\rho_e(X) = \sup_{\mathbf{Q}\in\mathcal{M}_1(\mathbb{P})}\{E^{\mathbf{Q}}[X] - H(\mathbf{Q}|\mathbb{P})\}$$

is called the entropic risk measure.

The variational principle for the entropic function H gives

$$\rho_e(X) = \log\mathbb{E}_{\mathbb{P}}[e^X].$$

The Worst case risk measure

The worst case risk measure is defined on L^∞ by

$$\rho_{max}(X) := \sup_{\omega\in\Omega} X(\omega).$$

It is a coherent risk measure, whose dual representation is given by

$$\rho_{max}(X) = \sup_{\mathbf{Q}\in\mathcal{M}_1} E^{\mathbf{Q}}[X].$$

It is the most conservative coherent risk measure. It represents in an insurance context, the total insurable value of an insurance company if X denotes its total exposition.

The Semi-deviation risk measure

For $0 \leq \delta \leq 1$ and $1 \leq p \leq \infty$, we define the semi-deviation risk measure as follows :

$$\rho(X) := \mathbb{E}[X] + \delta\mathbb{E}\left[(\mathbb{E}(X) - X)^p\right]^{1/p}.$$

ρ is a positively homogeneous law-invariant monetary risk measure. If we define $U(X) := -\rho(-X)$, we obtain the semi-deviation utility function, whose properties are studied in [68] for example. We will use this utility function in the next chapter.

2.2.1.4 Inf-convolution

Given two convex risk measures ρ_1 and ρ_2 with penalty functions α_1 and α_2, we define their inf-convolution in the following way:

$$\rho_1\Box\rho_2(X) := \inf_{F\in\mathcal{X}}\{\rho_1(X-F) + \rho_2(F))\} \tag{2.2.2}$$

Barrieu and El Karoui [5] proved under the assumption $\rho_1\Box\rho_2(0) > -\infty$ that $\rho_1\Box\rho_2$ is a finite convex risk measure. If ρ_1 is continuous from below, then $\rho_1\Box\rho_2$ is continuous from below and we have the following dual representation :

$$\rho_1\Box\rho_2(X) := \sup_{\mathbf{Q}\in\mathcal{M}_1}\{E^{\mathbf{Q}}[X] - (\alpha_1+\alpha_2)(\mathbf{Q})\} \tag{2.2.3}$$

The inf-convolution operator appears naturally in optimal risk transfer problems as we will see through the example of reinsurance optimization.

2.2.2 VaR Representation of Convex Risk Measures

Assumption 2.2.1 *We suppose in the rest of the chapter that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we work with is atomless.*

This assumption induces that our probability space supports a random variable with a continuous distribution (see for instance Proposition A.27 in [54]). Its usefullness is mentioned in Example 2.2.1, it is also important in the proof of Proposition 2.3.2.

The representation theorems stated above can be reformulated making use of Choquet integrals and distortion functions, that we now introduce.

Definition 2.2.3 *Every nondecreasing function $\psi : [0,1] \to [0,1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ will be called a distortion function.*

For a given right continuous distortion function ψ, we consider a $[0,1]$-valued random variable Λ_ψ whose law is given by:

$$\mathbb{P}(\Lambda_\psi \leq x) = \begin{cases} 0 \text{ if } x \leq 0 \\ \psi(x), \text{ if } 0 \leq x \leq 1 \\ 1 \text{ if } x \geq 1 \end{cases} \tag{2.2.4}$$

We will use random variables in the form of Λ_ψ to give an interpretation of the Choquet integral and of convex risk measures.

Definition 2.2.4 *A set function $c : \mathcal{F} \to [0,1]$ is called*

1. *Normalized if $c(\emptyset) = 0$ and $c(\Omega) = 1$.*
2. *Monotone if $c(A) \leq c(B)$ whenever $A \subset B$, A, $B \in \mathcal{F}$.*
3. *Submodular if $c(A \cup B) + c(A \cap B) \leq c(A) + c(B)$, $\forall$ A, $B \in \mathcal{F}$.*
4. *Outer regular if for every decreasing sequence (A_n) of elements of $\mathcal{F}$, we have*

$$c\left(\bigcap_n A_n\right) = \lim c(A_n). \tag{2.2.5}$$

Normalized and monotone set functions are called capacities.

Definition 2.2.5 *The Choquet integral of a random variable X, with respect to a capacity c is defined by:*

$$\int X d\,c := \int_{-\infty}^{0} (c(X > x) - 1)\, dx + \int_0^\infty c(X > x) dx$$

Example 2.2.1 *Using a distortion function ψ, we can define a capacity c_ψ given by $c_\psi(A) = \psi(\mathbb{P}(A))$, $\forall A \in \mathcal{F}$. For $\psi(x) = x$, the Choquet integral $\int X d\,c_\psi$ is just the expectation of X under the probability measure $\mathbb{P}$. The function ψ is used to distort the expectation operator $\mathbb{E}_\mathbb{P}$ into the non-linear functional ρ_ψ. It is known (see for example [54]) that when the probability space is atomless, ψ is concave if and only if ρ_ψ is a sub-linear risk measure.*

If we define $\psi(u) := \phi\left(\phi^{-1}(u) + \alpha\right)$, with $\alpha > 0$, where ϕ denotes the standard gaussian cumulative distribution function, we obtain the Wang transform risk measure, which is popular in insurance pricing (see Wang [120] for details on this particular Choquet integral).

Definition 2.2.6 *We will say that a risk measure ρ is law invariant, if $\rho(X) = \rho(Y)$ whenever X and Y have the same law under $\mathbb{P}$.*

The Choquet integrals with respect to capacities constructed from distortion functions as in the previous example can be used as building blocks for any convex and law invariant risk measure which is continuous from below. This is the result of Corollary 4.72 in [54]. This means that we can use Choquet integrals instead of the usual linear expectation in the dual representation (2.2.1) of any convex, law invariant and continuous from below risk measure ρ as follows

$$\rho(X) = \sup_{\psi \in \mathcal{C}} \left\{ \int X \, dc_\psi - \gamma(\psi) \right\}, \tag{2.2.6}$$

where $\mathcal{C}$ denotes the set of concave distortion functions and

$$\gamma(\psi) := \sup_{X \in L^\infty(\mathbb{P})} \left\{ \int X \, dc_\psi - \rho(X) \right\}.$$

In [76] and [54], a dual representation of law invariant convex risk measures is given in terms of $AVaR$. We have the same representation in terms of VaR risk measures, at random levels of the form Λ_ψ:

Proposition 2.2.1 *A convex risk measure ρ is law invariant and continous from above if and only if*

$$\rho(X) = \sup_{\psi \in \mathcal{C}} \left(\mathbb{E}_\mathbb{P} \left[VaR_{\Lambda_\psi}(X) \right] - \gamma(\psi) \right)$$

with

$$\gamma(\psi) = \sup_{X \in L^\infty(\mathbb{P})} \left(\mathbb{E}_\mathbb{P} \left[VaR_{\Lambda_\psi}(X) \right] - \rho(X) \right).$$

Proof.
We just notice that since $\{\overline{q}_X(\Lambda_\psi) > x\} = \{\Lambda_\psi \leq \mathbb{P}(X > x)\}$,

$$\mathbb{P}(\overline{q}_X(\Lambda_\psi) > x) = \mathbb{P}(\Lambda_\psi \leq \mathbb{P}(X > x)) = \psi(\mathbb{P}(X > x)) = c_\psi(X > x).$$

Using this, we can rewrite the Choquet integral as :

$$\int X d\, c_\psi = \int_{-\infty}^0 \left(\mathbb{P}(\overline{q}_X(\Lambda_\psi) > x) - 1 \right) dx + \int_0^\infty \mathbb{P}(\overline{q}_X(\Lambda_\psi) > x) dx = \mathbb{E}_\mathbb{P} \left[VaR_{\Lambda_\psi}(X) \right]$$

Now using the representation (2.2.6), we get the desired result. □

This proof enlights in particular that the Choquet integral of a r.v. X is a non linear operator, that can be rewritten as a classic linear expectation of a quantile of X, evaluated at a random level.

Remark 2.2.2 *(Link with AVaR representations) The dual representation of law invariant convex risk measures, with AVaR as building blocks, is as follows:*
ρ is convex and law invariant if and only if

$$\rho(X) = \sup_{\mu \in \mathcal{M}_1((0,1])} \left(\int_0^1 AVaR_\alpha(X) \mu(d\alpha) - \gamma(\mu) \right)$$

where

$$\gamma(\mu) = \sup_{X \in L^\infty(\mathbb{P})} \left(\int_0^1 AVaR_\alpha(X)\mu(d\alpha) - \rho(X) \right)$$

and $\mathcal{M}_1((0,1])$ *denotes the set of probability measures on* $(0,1]$.
These two representations with VaR and AVaR risk measures, evaluated at random levels, lead to no contradiction since

$$\int_0^1 AVaR_\alpha(X)\mu(d\alpha) = \int X d\, c_\psi = \mathbb{E}_\mathbb{P}\left[VaR_{\Lambda_\psi}(X)\right]$$

where ψ *is the distortion function defined by* $\frac{d\psi}{dt} = \int_t^1 \frac{1}{s}\mu(ds)$, *with* μ *a given probability measure on* $(0,1]$.

2.2.3 Optimal risk transfer and inf-convolution

Following the framework of [68] and [5], we consider the problem of optimal risk sharing between an insurer and a reinsurer in a principal-agent framework.
The exposure of the insurer (the agent) is modeled by a positive random variable X. Both the insurer and the reinsurer assess their risk using an increasing law invariant convex monetary risk measure (resp. ρ_1 and ρ_2). For a given loss level X, the insurer will take in charge $X - F$ and transfer to the reinsurer a quantity F, and for this he will pay a premium $\pi(F)$.

The insurer (the agent) minimizes his risk under the constraint that a transaction takes place, he solves :

$$\inf_{F,\pi}\{\rho_1(X - F + \pi)\} \text{ under the constraint } \rho_2(F - \pi) \leq \rho_2(0) = 0 \tag{2.2.7}$$

Binding this last constraint and using the cash-additivity property for ρ_2 gives the optimal price $\pi = \rho_2(F) - \rho_R(0) = \rho_R(F)$. This is an indifference pricing rule for the reinsurer, that is to say the price at which he is indifferent (from a risk perspective) between entering and not entering into the transaction.
Replacing $\pi = \rho_2(F)$ in (2.2.7) and using the cash-additivity property of ρ_1, the insurer program becomes equivalent to the following one:

$$\inf_F\{\rho_A(X - F) + \rho_R(F)\} =: \rho_A \square \rho_R(X)$$

We are left with the inf-convolution of ρ_1 and ρ_2, problem for which we give some explicit solutions in Theorem 2.3.3 and in Section 2.4.

2.3 The case of Comonotonic Risk Measures

We start this section with some definitions and properties related to the important notion of comonotonicity.
Let us begin with the definition of comonotonicity for random variables:

Definition 2.3.1 *Two random variables* X *and* Y *on* $(\Omega, \mathcal{F})$ *are called comonotone if*

$$\forall\, (\omega, \omega') \in \Omega^2,\ \left(X(\omega) - X(\omega')\right)\left(Y(\omega) - Y(\omega')\right) \geq 0$$

Otherwise said, X and Y are comonotone if we cannot use one variable as a hedge against the other. Denneberg [40] proved that two variables X and Y are comonotone if there exist a third random variable Z such that X and Y can be written as nondecreasing functions of Z.

The variables ξ_1 and ξ_2 corresponding to what is payable by the insurer and the reinsurer in most of common reinsurance contracts correspond to comonotone random variables. For instance, if the total loss is represented by a positive random variable X, $\xi_1 = \alpha X$ and $\xi_2 = (1-\alpha)X$ are comonotone. This is also true for $\xi_1 = X \wedge k$ and $\xi_2 = (X-k)^+$.

The risk evaluation corresponding to an agregate position $X + Y$ of comonotone variables X and Y is then just the sum of the individual risks:

Definition 2.3.2 *A risk measure ρ is called comonotonic if*

$$\rho(X+Y) = \rho(X) + \rho(Y), \text{ whenever } X, Y \text{ are comonotone random variables.}$$

2.3.1 Representation results on L^p

In [54], it is proved that a monetary risk measure ρ on $L^\infty(\mathbb{P})$ is comonotonic if and only if there exists a capacity c such that

$$\rho(X) = \int X\, dc, \tag{2.3.1}$$

and in this case, c is given by $c(A) = \rho(\mathbf{1}_A)$.

We can extend the representation (2.3.1) to risk measures defined on L^p spaces, but we will need the additional requirement that ρ is convex. Indeed, we will need the continuity of the Choquet integral on L^p to prove the representation on L^p, and it will be given by the following extended Namioka-Klee theorem proved in [10] for convex functionals:

Theorem 2.3.1 (Biagini - Frittelli [10]) *Any finite, convex and monotone increasing risk measure $\rho : L^p \to \mathbb{R}$ is continuous on L^p.*

Using this continuity result, we can state a representation on L^p, which is a direct consequence of Theorem 4.82 in [54]. Note that Kervarec [73], in the 3rd chapter of her PhD dissertation, also proves a continuity result for coherent risk measures defined on a space analogous to L^p in a situation with model uncertainty, with a non dominated family of probability measures.

Theorem 2.3.2 *Let ρ be a convex and comonotonic monetary risk measure on L^p, and let $C(X) := \int X\, dc$, where c is a capacity given by $c(A) = \rho(\mathbf{1}_A)$. If ρ and C are finite on L^p, then*

$$\rho(X) = C(X) \text{ for any r.v } X \text{ in } L^p. \tag{2.3.2}$$

Proof.

Of course Theorem 2.3.1 applies and ρ is L^p-continuous.

First note that since ρ is comonotonic and since the pair (X, X) is comonotone, then $\rho(2X) = 2\rho(X)$, by induction $\rho(nX) = n\rho(X)$ for $n \in \mathbb{N}^*$, and this easily implies that $\rho(qX) = q\rho(X)$ for any $q \in \mathbb{Q}^+$. We can use the L^p-continuity of ρ to deduce that $\rho(\lambda X) = \lambda\rho(X)$ for any $\lambda \in \mathbb{R}^+$. We have obtained that ρ is positively homogeneous, in particular $\rho(0) = 0$, and if we define $c(A) := \rho(\mathbf{1}_A)$ for any A in $\mathcal{F}$, then c is a capacity (monotone and normalized), which is also submodular by convexity of ρ.

We can define the Choquet integral $C(X) := \int X\, dc$. C is a convex (since c is submodular) and comonotonic monetary risk measure, so using the finiteness of C on L^p and Theorem 2.3.1, we obtain the continuity of C on L^p.

Now, by Theorem 4.82 in [54], we know that C and ρ coincide on L^∞. By the continuity of ρ and C on L^p and density of L^∞ in L^p, we have the desired result.

□

2.3.2 A quantile weighting result for general Choquet integrals

We denote by ρ_ψ the Choquet integral with respect to the capacity c_ψ constructed from a distortion function ψ. Recall that $\mathcal{X} = L^\infty$ or $\mathcal{X} = L^p$.

Assumption 2.3.1 *For any interval $[a, b]$ such that $x \to \mathbb{P}(X > x)$ is constant on $[a, b]$, the function $x \to c(X > x)$ is also constant on $[a, b]$. We will say that $\overline{c}_X$ inherits the constant intervals of $\overline{F}_X$.*

Remark 2.3.1 *If Q is a probability measure, then $x \to Q(X > x)$ inherits the constant intervals of $\overline{F}_X$ if and only if the law of X under Q is absolutely continuous with respect to the law of X under $\mathbb{P}$.*

Lemma 2.3.1 *Let c be a capacity on $\mathcal{F}$. Let X be a random variable in $\mathcal{X}$ such that $\overline{c}_X$ inherits the constant intervals of $\overline{F}_X$, then there exists a distortion function ψ_X such that the Choquet integral with respect to c coincides with ρ_{ψ_X} at X :*

$$\int X dc = \int X dc_{\psi_X}$$

Furthermore, if the capacity c is outer regular then ψ_X is right-continuous.

Proof. Recall that $\overline{q}_X(u) = \inf\{x \in \mathbb{R} \,|\, \overline{F}_X(x) \leq u\}$. In particular $\overline{q}_X(\overline{F}_X(x)) \leq x$.

Step 1: We first prove that for any real number x,

$$c\{X > x\} = c\{X > \overline{q}_X\left(\overline{F}_X(x)\right)\}.$$

For any x such that $\overline{F}_X(x)$ is a point of continuity of $\overline{q}_X$, we have $\overline{q}_X\left(\overline{F}_X(x)\right) = x$, so the claim is obvious in that case. Then let x be such that $\overline{F}_X(x)$ is a discontinuity point for the function $\overline{q}_X$.

We have $\overline{q}_X(\overline{F}_X(x)) < x$, and $\overline{F}_X$ is constant on the interval $[\overline{q}_X(\overline{F}_X(x)), x]$, then by Assumption 2.3.1, $\bar{c}_X$ is also constant on this interval which is the desired claim.

Step 2: For a given random variable X and a capacity c, we define

$$\psi_X(u) := c(X > \overline{q}_X(u)). \tag{2.3.3}$$

ψ_X is a distortion function since c is monotone and normalized, and using step 1 above, we have

$$\psi_X(\mathbb{P}(X > x)) = c\left(X > \overline{q}_X(\mathbb{P}(X > x))\right) = c(X > x).$$

And hence,

$$\int_{-\infty}^{0} \left(c(X > x) - 1\right) dx + \int_{0}^{\infty} c(X > x) dx = \int_{-\infty}^{0} \left(c_{\psi_X}(X > x) - 1\right) dx + \int_{0}^{\infty} c_{\psi_X}(X > x) dx,$$

which is the desired result.

The proof that the outer regularity of c implies the right continuity of ψ_X is as follows:
Since $\overline{q}_X$ is right-continuous and non increasing, for every non increasing sequence (u_n) converging to u, $\overline{q}_X(u_n)$ goes to $\overline{q}_X(u)$ and $\{X > \overline{q}_X(u_{n+1})\} \subset \{X > \overline{q}_X(u_n)\}$. Then the application of the property (2.2.5) with $A_n = \{X > \overline{q}_X(u_n)\}$ gives the right continuity of ψ_X.

$\square$

In [76], Kusuoka provided a representation of *law invariant* risk measures as an integral of weighted quantiles. More precisely, he proved that ρ is a monetary risk measure which is law invariant and comonotonic if and only if there exists a distortion function ψ such that for every X in $L^\infty(\mathbb{P})$,

$$\rho(X) = \int_0^1 \overline{q}_X(u)\psi'(u) du \tag{2.3.4}$$

We can still write a quantile representation for a monetary risk measure which is comonotonic but not necessarily law invariant:

Proposition 2.3.1 *Let ρ be a monetary risk measure satisfying the Fatou property of Definition 2.2.2, which is comonotonic, and let X be in L^∞. If $x \to \rho(\mathbf{1}_{\{X>x\}})$ inherits the constant intervals of $\overline{F}_X$, then there exists a distortion function ψ_X such that*

$$\rho(X) = \int_0^1 \overline{q}_X(u)\psi_X'(u) du. \tag{2.3.5}$$

Proof.
Since ρ is comonotonic, there exists a capacity c such that $\rho(X) = \int X dc$. The Fatou property of ρ is equivalent with its continuity from above (Lemma 4.20 in [54]), and this implies that $c(A) = \rho(\mathbf{1}_A)$ is outer regular. At the point X, ρ coincides by Lemma 2.3.1 with a right-continuous ψ_X-distortion risk measure, and we re-write it as follows:

$$\int_{-\infty}^{0} \left(c_{\psi_X}(X > x) - 1\right) dx + \int_{0}^{\infty} c_{\psi_X}(X > x) dx = \mathbb{E}_{\mathbb{P}}\left[\overline{q}_X(\Lambda_{\psi_X})\right]$$

and the last quantity, by definition of Λ_{ψ_X}, is equal to $\int_0^1 \overline{q}_X(u)\psi'_X(u)du$, which ends the proof. Remark that Λ_{ψ_X} is well defined since ψ_X is right-continuous.

□

Remark 2.3.2 *The same proof shows that the result is true for risk measures defined on L^p, provided that $C(X) = \int X\,dc$ is finite, where $c(A) := \rho(\mathbf{1}_A)$.*

In view of these results, the inf-convolution of comonotonic risk measures reduces to the inf-convolution of Choquet integrals, the study of which will require some preliminary results that we enonciate and prove in section 2.3.2. We will represent (2.3.8) the inf-convolution of two comonotonic and cash additive functionnals as a Choquet integral in section 2.3.3.

2.3.3 Inf-Convolution of Choquet Integrals

As shown in the following proposition, the Kusuoka-type quantile representation (2.3.5) is stable under inf-convolution. This extends a result in [67] to non necessarily law invariant measures.

Proposition 2.3.2 *Let ρ_1 and ρ_2 be convex comonotonic risk measures satisfying the Fatou property and let $X \in \mathcal{X}$. Assume that $\rho_1 \square \rho_2(0) > -\infty$ and the functions $x \to \rho_i(\mathbf{1}_{\{X>x\}})$ inherit the constant intervals of $\overline{F}_X$.*

If $\mathcal{X} = L^p$, we also assume that ρ_i and its associated comonotonic risk measure is finite on L^p, $i = 1, 2$.

Then there exist two distortion functions g_1^X and g_2^X such that

$$\rho_1 \square \rho_2(X) = \int \overline{q}_X(u)(g_1^X \wedge g_2^X)'(u)\,du \tag{2.3.6}$$

where $(g_1^X \wedge g_2^X)(x) := \min\big(g_1^X(x), g_2^X(x)\big)$.

Proof. Since ρ_1 and ρ_2 are convex and comonotonic, using either (2.3.1) or (2.3.2) we have the existence of two submodular capacities c_1 and c_2 such that ρ_1, ρ_2 are the associated Choquet integrals. Let $c_{1,2}(A) := \min(c_1(A), c_2(A))$, $\forall A \in \mathcal{F}$. The dual representation of ρ_1 and ρ_2 is given by:

$$\rho_i(X) = \sup_{Q \in \mathcal{M}_{1,f}} \{E_Q(X) - \alpha_i(Q)\} \tag{2.3.7}$$

with $\alpha_i(Q) = 0$ if $Q \in \mathcal{C}_i$ and $+\infty$ if not.
$\mathcal{C}_i$ is the core of c_i defined by $\mathcal{C}_i = \{Q \in \mathcal{M}_{1,f} | Q(A) \leq c_i(A),\ \forall A \in \mathcal{F}\}$ and $\mathcal{M}_{1,f}$ denotes the set of finitely additive set functions. Now theorem 3.6 in [5] says that $\rho_1 \square \rho_2$ is a convex risk measure with a penalty function given by $\tilde{\alpha}(Q) = \alpha_1(Q) + \alpha_2(Q)$, which equals 0 if $Q \in \tilde{\mathcal{C}} = \mathcal{C}_1 \cap \mathcal{C}_2 = \{Q \in \mathcal{M}_{1,f} | Q(A) \leq \min(c_1(A), c_2(A)),\ \forall A \in \mathcal{F}\}$ and $+\infty$ if not. $\tilde{\alpha}(Q)$ is the core of $c_{1,2}$. $c_{1,2}$ is obviously a capacity, and using theorem 4.88 in [54], $c_{1,2}$ is a submodular capacity (here the Assumption **??** of atomless probability space is important) and we have

$$\rho_1 \square \rho_2(X) = \int X dc_{1,2} \tag{2.3.8}$$

Now since $c_{1,2}(X > x)$ inherits the constant intervals of $\overline{F}_X(x)$ (since c_1 and c_2 do), we can use Lemma 2.3.1 to provide the existence of a distortion function $\tilde{g}$ such that $\int X dc_{1,2}$ coincides with the Choquet integral with respect to the capacity $\tilde{g}(\mathbb{P}(.))$ at X. We have

$$\tilde{g}(u) = c_{1,2}(X > \overline{q}_X(u)) = c_1(X > \overline{q}_X(u)) \wedge c_2(X > \overline{q}_X(u)) = \min\left(g_1^X(u), g_2^X(u)\right),$$

with $g_i^X(u) = c_i(X > \overline{q}_X(u))$, $i = 1, 2$. This implies, as in the demonstration of Proposition 2.3.1, that $\rho_1 \Box \rho_2(X) = \int \overline{q}_X(u)(g_1^X \wedge g_2^X)'(u)\, du$ (this is where the Fatou property is needed).

$\square$

Remark 2.3.3 *The formula (2.3.8) in the previous proof tells us that the inf-convolution $\rho_1 \Box \rho_2(X)$ is again a Choquet integral with respect to the capacity*

$$c_{1,2}(A) = \min(c_1(A), c_2(A)).$$

In the particular case where c_1 and c_2 are constructed with distortion functions ψ_1 and ψ_2 as in Example 2.2.1, then $c_{1,2}$ is also a distorted probability with respect to the function $\psi_{1,2}(u) := \min(\psi_1(u), \psi_2(u))$.

If at a given point u^* in $[0, 1]$, we have $\psi_1(u^*) < \psi_2(u^*)$, then we can say that at the given quantile level u^*, agent 1 is less risk averse than agent 2, since he uses a less severe distortion of the a-priori probability $\mathbb{P}$ at this point u^*. Now the form of $\psi_{1,2}$ suggests that the inf-convolution splits the total risk X in many parts, which are asigned to the agent being the less risk averse on the corresponding quantile intervals (see Theorem 2.3.3 and Section 2.4 for more details).

Using this result, we are able to solve explicitely the inf-convolution of convex Choquet integrals.

Recall that the core $\mathcal{C}_i$ of the capacity c_i is defined by

$$\mathcal{C}_i = \{Q \in \mathcal{M}_{1,f} \,|\, Q(A) \leq c_i(A),\ \forall A \in \mathcal{F}\}$$

where $\mathcal{M}_{1,f}$ denotes the set of finitely additive set functions on Ω.

In order to correctly formulate our next theorem, we need the following definitions.

Definition 2.3.3 *We will say that a subset B of $[0, 1]$ is densely ordered if for any x, y in B with $x < y$, there exists z in B such that $x < z < y$.*

For instance the set $\mathbb{Q} \cap [0, 1]$ is numerable and densely ordered.

Definition 2.3.4 *Let f be a real valued function defined on $[0, 1]$. We define the changing sign set of f by those u in $(0, 1)$ such that for every sequence (u_n) increasing towards u and every sequence (v_n) decreasing towards u, there is an integer N, such that $\forall n \geq N$, $sgn[f(u_n)] = -sgn[f(v_n)]$, where as usual, $sgn(x) = -1$ if $x < 0$ and $sgn(x) = 1$ if $x > 0$.*

The two previous definitions are used to formulate the following assumption on two distortion functions.

Assumption 2.3.2 *Let ψ_1 and ψ_2 be two distortion functions and denote $f := \psi_1 - \psi_2$. The set of changing sign points of f contains no densely ordered subset.*

The economic interpretation of the previous assumption is as follows: it means that $\psi_1 - \psi_2$ does not change sign too often. Recall from the previous remark that these functions represent uncertainty weighting, if $\psi_1(u) \geq \psi_2(u)$ for a certain point u, an economic agent using ψ_1 as a distortion function asigns more uncertainty to the point u than an agent 2 using ψ_2. Now when the sign of $\psi_1 - \psi_2$ changes, it means that the order of uncertainty weighting between agent 1 and agent 2 changes. If it changes very often in the sense that Assumption 2.3.2 is not satisfied, this means that the agents relative agreement on the probability weighting changes very often.

A sufficient condition for Assumption 2.3.2 to hold is that the distortions ψ_1 and ψ_2 are continuous and have a finite number of crossing points. This will be the case in all the examples we consider.

Theorem 2.3.3 *Let X be in L^∞. Let ρ_1 and ρ_2 be two comonotonic risk measures satisfying the assumptions of Proposition 2.3.2. Suppose also that the associated local distortions g_1^X and g_2^X are continuous and satisfy Assumption 2.3.2.*
Then there exists a possibly infinite increasing sequence $\{k_n, n \leq N\}$ of real numbers corresponding to quantile values of X, such that

$$\rho_1 \Box \rho_2(X) = \rho_1 (X - Y^*) + \rho_2(Y^*)$$

where Y^* is given by:

$$Y^* = \sum_{p=0}^{N} (X - k_{2p})^+ - (X - k_{2p+1})^+.$$

Remark 2.3.4 *If the sequence $\{k_n, n \geq 0\}$ only takes two non zero values k_1 and k_2, then the first agent (the insurer in that case) takes in charge $X \wedge k_1 + (X - k_2)^+$ and agent two (the reinsurer) is exposed to $(X - k_1)^+ - (X - k_2)^+$. This is exactly the definition of the Excess-of-Loss contract, denoted "A-XS-B", that we find in a large majority of catastrophe reinsurance market exchanges (here $B = k_1$ and $A = k_2 - k_1$).*

Remark 2.3.5 *This result means that the inf-convolution of law invariant and comonotonic risk measures is given by a generalization of the Excess-of-Loss contract, with more treshold values. The domain $\mathbb{R}^+$ of attainable losses is divided in "ranges", and each range is alternatively at the charge of one of the two agents.*

The case where N is infinite is of course impossible in practice, but it gives an additional conceptual insight to the optimal structure of the risk transfer.
Proof.

We start by writing the representation (2.3.6) for the inf-convolution and divide it in two parts. For $X \in \mathcal{X}$, let $A := \{x \in [0,1] : g_1^X(x) \leq g_2^X(x)\}$ (we will write g_1 and g_2 instead of g_1^X and g_2^X for convenience),

$$\begin{aligned}\rho_1\Box\rho_2(X) &= \int_0^1 \overline{q}_X(u)(g_1\wedge g_2)'(u)\,du\\ &= \int_0^1 \overline{q}_X(u)g_1'(u)\mathbf{1}_A(u)\,du + \int_0^1 \overline{q}_X(u)g_2'(u)\mathbf{1}_{A^c}(u)\,du.\end{aligned}$$

Since g_1 and g_2 satisfy Assumption 2.3.2 and are continuous, the following non increasing sequence $\{u_n,\, n\geq 1\}$ is well defined: it corresponds to the points of equality of the functions g_1 and g_2 defined as follows :

$$u_0 = 1;\; u_{n+1} = \sup\{x < u_n \;\text{ such that }\; g_1(x) = g_2(x) \;\text{ and }\; g_1(x^-) \neq g_2(x^-)\}$$

For each n we define $k_n := \overline{q}_X(u_n)$.

Then we can write the set A as the numerable union of disjoint intervals $A = \bigcup_{n\geq 0}[u_{n+1}, u_n]$, where $u_0 = 1$, $\lim_{n\to+\infty} u_n = 0$ and possibly there are integers $k\in\mathbb{N}$ such that $g_1(x) = g_2(x)$ for any $x\in[u_{k+1}, u_k]$. The non increasing sequence $\{u_n,\, n\geq 1\}$ corresponds to points of equality of the functions g_1 and g_2. We can rewrite the inf-convolution

$$\begin{aligned}\rho_1\Box\rho_2(X) =& \sum_{p\geq 1}\int_0^1 \overline{q}_X(u)\mathbf{1}_{(u_{2p+1}\leq u\leq u_{2p})}g_1'(u)\,du\\ &+ \sum_{p\geq 1}\int_0^1 \overline{q}_X(u)\mathbf{1}_{(u_{2p+2}\leq u\leq u_{2p+1})}g_1'(u)\,du.\end{aligned} \tag{2.3.9}$$

Then we remark that

$$\begin{aligned}\overline{q}_X(u)\mathbf{1}_{(u_{i+1}\leq u\leq u_i)} =& (\overline{q}_X(u) - \overline{q}_X(u_i))^+ - (\overline{q}_X(u) - \overline{q}_X(u_{i+1}))^+\\ &- \overline{q}_X(u_{i+1})\mathbf{1}_{\{u\leq u_{i+1}\}} + \overline{q}_X(u_i)\mathbf{1}_{\{u\leq u_i\}},\end{aligned}$$

We define the increasing sequence (k_i) by $k_i := \overline{q}_X(u_i)$, and the random variables $Z := (X-k_i)^+ - (X-k_{i+1})^+$ and $Z_i = \mathbf{1}_{\{X>k_i\}}$, then we have

$$\overline{q}_X(u)\mathbf{1}_{(u_{i+1}\leq u\leq u_i)} = \overline{q}_Z(u) - k_{i+1}\overline{q}_{Z_{i+1}}(u) + k_i\overline{q}_{Z_i}(u).$$

We can plug this equality into (2.3.9) to obtain

$$\begin{aligned}\rho_1\Box\rho_2(X) =& \rho_1\left(\sum_{p\geq 1}(X-k_{2p})^+ - (X-k_{2p+1})^+\right)\\ &+ \rho_2\left(\sum_{p\geq 1}(X-k_{2p+1})^+ - (X-k_{2p+2})^+\right)\\ &- \sum_{p\geq 1}\Big(k_{2p+1}\rho_1(X>k_{2p+1}) - k_{2p}\rho_1(X>k_{2p})\Big)\\ &- \sum_{p\geq 1}\Big(k_{2p+2}\rho_2(X>k_{2p+2}) - k_{2p+1}\rho_2(X>k_{2p+1})\Big)\end{aligned}$$

Where we denoted $\rho(X > k) := \rho(\mathbf{1}_{\{X>k\}})$.

Finally, by noticing that for any i, $\rho_1(\mathbf{1}_{(X>k_i)}) = g_1(u_i) = g_2(u_i) = \rho_2(\mathbf{1}_{(X>k_i)})$, we obtain the desired claim. □

The last proof shows in particular that the points k_i defining the risk transfer are such that $\rho_1(X > k_i) = \rho_2(X > k_i)$. This is the counterpart on $\mathbb{R}$ of the equality at the points u_i belonging to $[0, 1]$ of the local distortion functions g_1^X and g_2^X.

In Figure 2.1 below, we represented two possible functions g_1^X and g_2^X crossing twice on $(0, 1)$, and thus generating a layer as the optimal structure for the risk sharing. Assume that an insurer uses the local distortion at X corresponding to the red curve, and that a reinsurer uses the one corresponding to the green curve.

In the interval $(w, 1)$, the reinsurer uses a more severe probability distortion, and thus all the risk is assumed by the insurer in that region, since $(w, 1)$ corresponds to losses X belonging to $(0, k_1)$, with $k_1 = \overline{q}_X(w)$. The same thing happens for the interval $(0, v)$ corresponding to losses belonging to $(k_2, +\infty)$, with of course $k_2 = \overline{q}_X(v)$. Finally, the reinsurer assumes all the losses belonging to (k_1, k_2), since he has a less severe distortion on the interval (v, w).

With the notations of the last proof, k_1 and k_2 are the only two finite points such that we have the "agreement" $\rho_1(X > k_i) = \rho_2(X > k_i)$.

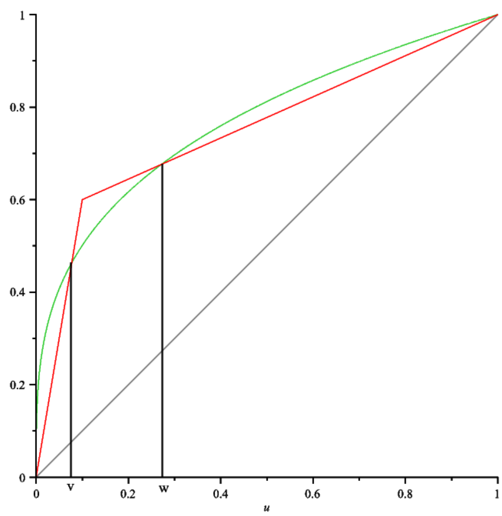

Figure 2.1: Example of configuration leading to a layer as optimal risk transfer

2.3.4 An extension to subadditive comonotonic functionals

2.3.4.1 Representation on L^∞ and L^p

Proposition 2.3.3 *ρ is an increasing, comonotonic and cash sub-additive functional on $L^\infty(\mathbb{P})$, if and only if there exist a set function c such that for each X in $L^\infty(\mathbb{P})$, $\rho(X) = \int_\Omega X dc$, where for*

$A \in \mathcal{F}$, c is defined by $c(A) = \nu.\tilde{c}(A)$ and $\tilde{c}$ is a capacity on $\mathcal{F}$. In that case $\nu = \rho(1) = c(\Omega)$.

Proof.

First suppose that ρ is an increasing, comonotonic and cash sub-additive risk measure. Using the monotony and cash sub-additivity of ρ, we have that ρ is continuous with respect to the supremum norm since

$$X \leq Y + \|X - Y\|_\infty \Rightarrow \rho(X) \leq \rho(Y + \|X - Y\|_\infty) \leq \rho(Y) + \|X - Y\|_\infty,$$

Then $|\rho(X) - \rho(Y)| \leq \|X - Y\|_\infty$. Lipschitz continuity and comonotonicity imply that ρ is positively homogeneous (see lemma 4.77 in [54]). Then $\rho(0) = 0$, and by comonotonicity and positive homogeneity,

$$\rho(X + m) = \rho(X) + \rho(m) = \rho(X) + m\rho(1)$$

and $\rho(1) \leq 1$ by cash sub-additivity. So ρ is linear cash-subadditive. Set $\nu := \rho(1)$ and define $\tilde{\rho}(X) := \rho(\frac{1}{\nu}X)$, then $\tilde{\rho}$ is increasing, comonotonic, and cash additive, using (2.3.1), we know that there exist a capacity $\tilde{c}$ such that $\tilde{\rho}(X) = \int X d\tilde{c}$. Setting $c(A) = \nu.\tilde{c}(A)$, we obtain $\rho(X) = \int X dc$.

Now suppose that ρ is the Choquet integral on $L^\infty(\mathbb{P})$ with respect to a monotone set function c such that $c(\Omega) \leq 1$, then it is proved in [40] that ρ is increasing and comonotonic. Furthermore, $\rho(X + m) = \int X dc + mc(\Omega) \leq \rho(X) + m$.

□

Using the same arguments as in the proofs of Theorem 2.3.2 and Proposition 2.3.3, we can show the following:

Corollary 2.3.1 *Let ρ be an increasing, convex, comonotonic and cash sub-additive functional on L^p. Let $C(X) := \int X\, dc$ where c is the capacity given by $c(A) := \rho(\mathbf{1}_A)$. Assume that ρ and C are finite on L^p, then*

$$\rho(X) = \int X\, dc,\ \forall X \in L^p.$$

2.3.4.2 Sub-additive measures and intrinsic discounting factor

Definition 2.3.5 *An application ρ is said to be **cash sub-additive** if for each m in $\mathbb{R}^+$, $\rho(X + m) \leq \rho(X) + m$ and $\rho(X - m) \geq \rho(X) - m$.*
We will say that ρ is linear cash sub-additive if there is a fixed parameter ν in $[0, 1]$ such that $\rho(X + m) - \rho(X) = \nu m$.

As stated in [50], the economic interest of such a property is to take into account the time value of money. If $\rho(X)$ represents the minimal capital required to hedge the risky position X with maturity T, then the minimal capital needed to make $(X - m)$ acceptable is $\rho(X) - e^{-rT}m$, where m is a deterministic amount and r is the riskless interest rate. In this case ρ is such that $\rho(X - m) \geq \rho(X) - m$.

We can associate to every cash sub-additive risk measure an intrinsic discounting factor given by $\nu(m) := \frac{\rho(X+m)-\rho(X)}{m}$. Indeed, for each m,

$$\rho(X+m) = \rho(X) + \nu(m)m,$$

and $0 \leq \nu(m) \leq 1$.

For linear cash sub-additive risk measures, $\nu(m) = \nu$ is constant. For instance, if $\rho = e^{-rT}\tilde{\rho}$, where $T > 0$, r is a risk free rate and $\tilde{\rho}$ is cash additive, then $\nu(m) = e^{-rT}$.

For general cash sub-additive risk measures as in definition 2.3.5, $\nu(m)$ can depend on m. This means implicitely that the interest rate depends on the amount on which it is applied : for a fixed maturity, capitalizing during the considered period an amount m is multiplying it by $\frac{1}{\nu(m)} = \frac{m}{\rho(X+m)-\rho(X)}$, that is to say the ratio between the deterministic amount m, and the excess of regulatory capital saved today. In this sense $\nu(m)$ represents the internal cost of capital, for a company using ρ as risk measure.

If $\rho(X+m)-\rho(X)$ is a linear function of m, the capitalization factor $\frac{1}{\nu(m)}$ is independent of m, this will be the case for comonotonic cash sub-additive risk measures. Otherwise, $\nu(m)$ depends on m, and on the current position X, to which an economic agent is already exposed. Said otherwise, we can model an intrinsic capitalisation factor $\frac{1}{\nu}$, depending on the risk measure of the agent considered, on what he is already exposed to, and on the amount invested.

Example 2.3.1 *We consider here random variables in $L^2(\mathbb{P})$. Let us suppose that an agent has an initial economic endowment X following a centered standard gaussian law $N(0,1)$, and that he uses $\rho(X) := \mathbb{E}_{\mathbb{P}}(X^+)$ as risk criterion. It is a classic example of cash sub-additive risk measure ([50]). Following this particular risk vision, to have m at the end of the period, the agent needs the amount $\nu(m).m$ today, with*

$$\nu(m) = \frac{\rho(X+m)-\rho(X)}{m} = \phi(m) - \frac{1-e^{-\frac{m^2}{2}}}{m},$$

where ϕ denotes the cumulative distribution function of a standard gaussian law. We clearly see the dependence of $\nu(m)$ on the choice of ρ, on the law of X, and on the amount m.

Note that this coincides with the interpretation of El Karoui and Ravanelli [50] of ρ as a measure with interest rate ambiguity since

$$\rho(X) = \sup_{D \in L^2(\mathbb{P})} \left(\mathbb{E}_{\mathbb{P}}(D.X) \mid 0 \leq D \leq 1\right).$$

2.4 Examples

2.4.1 AVaR Examples

1. Let α and β be in $(0,1)$ and consider $\rho_1 = AVaR_\alpha$ and $\rho_2 = AVaR_\beta$. Then we have

$$(AVaR_\alpha \Box AVaR_\beta)(X) = AVaR_{\alpha\wedge\beta}(X),\ \forall X \in L^\infty.$$

Indeed, $AVaR_\alpha$ is the Choquet integral with respect to the distorted probability $c(A) = \psi_\alpha(\mathbb{P}(A))$ where

$$\psi_\alpha(u) = \frac{u}{\alpha} \wedge 1.$$

It is easy to see that if $\beta < \alpha$, then $\psi_\alpha(u) \leq \psi_\beta(u)$ for any u in $[0,1]$. This means that the distortion ψ_β is always more severe than the distortion ψ_α, and from Theorem 2.3.3 we know that the agent using the less severe distortion assumes all the risk.

2. A classic measure used in reinsurance as a risk measure or as a pricing functional is the *Proportional Hazard* transform, or PH-transform. It is defined by the following Choquet integral

$$\rho_{PH}(X) = \int X\, dc, \quad \text{with } c(X > x) = \mathbb{P}(X > x)^r,\ 0 < r < 1.$$

Using Theorem 2.3.3, and the fact that the functions $\psi_\alpha(u) = \frac{u}{\alpha} \wedge 1$ and $\psi_{PH}(u) = u^r$ only coincide at the point $u^* = \alpha^{\frac{1}{1-r}}$, we obtain that

$$\rho_{PH} \square AVaR_\alpha(X) = \rho_{PH}(X \wedge k^*) + AVaR_\alpha((X - k^*)^+),$$

with $k^* = \overline{q}_X(\alpha^{\frac{1}{1-r}})$.

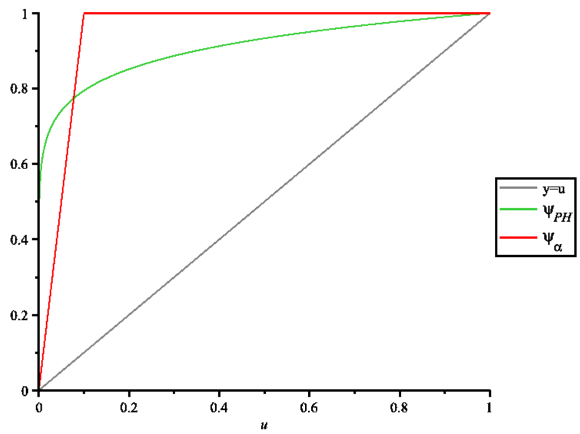

Figure 2.2: PH-transform and $AVaR_\alpha$ distortions

We can take for example $\alpha = r = \frac{1}{10}$, in that case it leads to the distortion functions represented in Figure 2.2 and $u^* = \alpha^{\frac{1}{1-r}} = 0.0774$. If X has a beta distribution with density $f_X(x) = \frac{x^{v-1}(1-x)^{w-1}}{\beta(v,w)}$, $v = 4$ and $w = 8$, then $k^* = 0.53$.

Assume that an insurance company is facing a peril corresponding to a random destruction rate which has a $\beta(4,8)$ distribution on $[0,1]$.If the total insurable value is given by $M > 0$, then an insurer using the PH-transform Choquet integral will cede every losses above the value $0.53M$ to a reinsurer using the $AVaR_\alpha$ as risk measure.

2.4.2 Epsilon-contaminated capacities

Let ν be the capacity given as the *Epsilon-contamination* of a given probability measure Q, which means that ν is the submodular capacity defined by

$$\nu(A) := (1-\varepsilon)Q(A) + \varepsilon, \quad 0 < \varepsilon < 1 \quad \text{for} \quad A \neq \{\emptyset\} \quad \text{and } \nu(\{\emptyset\}) = 0$$

Let X be a bounded random variable, denote $X^* := \sup X(\omega)$ and assume that $X^* > 0$. Applying the definition of the Choquet integral of X with respect to ν, we get

$$\begin{aligned}
\int X\, d\nu &= \int_{-\infty}^{0} (\nu(X > x) - 1)\, dx + \int_{0}^{X^*} \nu(X > x) dx + \int_{X^*}^{+\infty} \nu(X > x) dx \\
&= \int_{-\infty}^{0} (1-\varepsilon)\, (Q(X > x) - 1)\, dx + \int_{0}^{X^*} [(1-\varepsilon)Q(X > x) + \varepsilon]\, dx \\
&= (1-\varepsilon)E^Q[X] + \varepsilon X^*.
\end{aligned}$$

In this case the Choquet integral is affected by the supremum of X.

By analogy with Dow and Werlang [43], we define the uncertainty aversion of a capacity ν at $A \in \mathcal{F}$ by

$$e(\nu, A) := \nu(A) + \nu(A^c) - 1.$$

It is easy to check that an Epsilon contamination ν of Q is a constant uncertainty aversion capacity, i.e $e(\nu, A) = \varepsilon$, for any $A \neq \{\emptyset\}$, $A \neq \Omega$.

If two agents are using respectively an ε_1 and ε_2-contaminated capacities ν_1 and ν_2, we will say that agent one is more uncertainty averse than agent two if $\varepsilon_1 > \varepsilon_2$.

Remark 2.4.1 *We give here slightly different definitions than Carlier, Dana and Shahidi [23] or Dow and Werlang [43]. This is due to the facts that we consider submodular instead of supermodular capacities and that a positive random variable X represents losses for us, instead of gains, so a risk averse agent seeks for instance to increase rather than decrease the expectation of X with respect to the a priori probabilty measure.*

Remark 2.4.2 *A capacity ν, given as an Epsilon-contamination of a probability measure Q is not a distortion of the probability Q in the sense of Example 2.2.1. In particular, if we define a capacity $\widetilde{\nu}$ by*

$$\widetilde{\nu}(A) := \psi(Q(A)), \quad \textit{with } \psi \textit{ given by} \quad \psi(0) = 0 \textit{ and } \psi(u) = (1-\varepsilon)u + \varepsilon,\ 0 < u \leq 1,$$

then $\nu \neq \widetilde{\nu}$. Indeed, for a Q-negligible set $B \neq \{\emptyset\}$, we have

$$\widetilde{\nu}(B) = 0 \quad \textit{and} \quad \nu(B) = \varepsilon > 0.$$

We will nonetheless write the Choquet integral with respect to ν using local distortion functions.

Since $\int X\,d\nu$ is affected by sup $X(\omega)$, we only consider bounded random variables X, but the simple uniform distribution on $[0,1]$ already provides an interesting example.

Recall that our a priori given probability space is $(\Omega, \mathcal{F}, \mathbb{P})$. We will write $\nu(X > x)$ as $\psi_X^\nu(\mathbb{P}(X > x))$ for some local distortion function ψ_X^ν.

We assume that agent one uses a Choquet integral with respect to the capacity ν_1, given as an ε_1-contamination of a probabilty measure Q, such that $Q << \mathbb{P}$. Agent two uses a Choquet integral with respect to ν_2, the ε_2-contamination of $\mathbb{P}$. Let us take $\varepsilon_1 < \varepsilon_2$. In other word we have

$$\begin{aligned}
&\nu_1(A) = (1-\varepsilon_1)Q(A) + \varepsilon_1, \quad \text{with} \quad Q << \mathbb{P} \quad \text{for} \quad A \neq \{\emptyset\} \quad \text{and } \nu_1(\{\emptyset\}) = 0,\\
&\nu_2(A) = (1-\varepsilon_2)\mathbb{P}(A) + \varepsilon_2 \quad \text{for} \quad A \neq \{\emptyset\} \quad \text{and } \nu_2(\{\emptyset\}) = 0.
\end{aligned}$$

Notice that ν_1 and ν_2 are two different examples of submodular capacities which are not given by a distortion of the probability measure $\mathbb{P}$.

The following lemma gives an example of two different optimal risk transfers, with two random variables having the same law.

Lemma 2.4.1 *Let Z and Z' be two independent and uniformly distributed random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $[0,1]$. Let $\theta > 0$ be a given parameter and define Q as the θ-Esscher transform of $\mathbb{P}$, i.e $\frac{dQ}{d\mathbb{P}} = \frac{e^{\theta Z}}{\mathbb{E}^{\mathbb{P}}[e^{\theta Z}]}$. We have*

$$\begin{aligned}
\int Z\,d\nu_1 &= \int_0^1 \psi_Z^{\nu_1}[\mathbb{P}(Z > x)]dx \quad \textit{where} \quad \psi_Z^{\nu_1}(u) = \varepsilon_1 + (1-\varepsilon_1)\left(\frac{e^\theta - e^{(1-u)\theta}}{e^\theta - 1}\right),\\
\int Z\,d\nu_2 &= \int_0^1 \psi_Z^{\nu_2}[\mathbb{P}(Z > x)]dx \quad \textit{where} \quad \psi_Z^{\nu_2}(u) = \varepsilon_2 + (1-\varepsilon_2)u,\\
\int Z'\,d\nu_1 &= \int_0^1 \psi_{Z'}^{\nu_1}[\mathbb{P}(Z' > x)]dx \quad \textit{where} \quad \psi_{Z'}^{\nu_1}(u) = \varepsilon_1 + (1-\varepsilon_1)u,\\
\int Z'\,d\nu_2 &= \int_0^1 \psi_{Z'}^{\nu_2}[\mathbb{P}(Z' > x)]dx \quad \textit{where} \quad \psi_{Z'}^{\nu_2}(u) = \psi_Z^{\nu_2}(u) = \varepsilon_2 + (1-\varepsilon_2)u.
\end{aligned}$$

As a consequence, we have the following optimal risk transfer structures:

$$\rho_1 \Box \rho_2(Z) = \rho_1\Big((Z-k^*)^+\Big) + \rho_2(Z \wedge k^*)$$

with $k^* = \overline{q}_X^{\mathbb{P}}(u^*)$, where $u^* \in (0,1)$ is the solution of the equation $\psi_Z^{\nu_1}(u) = \psi_Z^{\nu_2}(u)$. Furthermore, we have

$$\rho_1 \Box \rho_2(Z') = \rho_1(Z'),$$

where we denoted by ρ_1 and ρ_2 the Choquet integrals with respect to ν_1 and ν_2.

Proof. The functions $\nu_1(Z > x)$ and $\nu_2(Z > x)$ for $x \in (0,1)$ inherit the constant intervals of $\mathbb{P}(Z > x)$ since $\mathbb{P}(Z > x)$ is never constant on $(0,1)$, so there is nothing to check, and this is also true for Z' instead of Z.

Then Lemma 2.3.1 gives that

$$\int Z\,d\nu_i = \int_0^1 \psi_Z^{\nu_i}[\mathbb{P}(Z > x)]dx, \quad \text{with} \quad \psi_Z^{\nu_i}(u) = \nu_i\big(Z > \overline{q}_Z^{\mathbb{P}}(u)\big),\ i = 1, 2.$$

We calculate for instance, for $0 < u \leq 1$

$$\nu_1\big(Z > \overline{q}_Z^{\mathbb{P}}(u)\big) = (1-\varepsilon_1)Q\big(Z > \overline{q}_Z^{\mathbb{P}}(u)\big) + \varepsilon_1 \quad \text{and} \quad \psi_Z^{\nu_1}(0) = 0.$$

Moreover

$$Q\big(Z > \overline{q}_Z^{\mathbb{P}}(u)\big) = \mathbb{E}^{\mathbb{P}}\left[\frac{e^{\theta Z}}{\mathbb{E}^{\mathbb{P}}[e^{\theta Z}]}\mathbf{1}_{\{Z > \overline{q}_Z^{\mathbb{P}}(u)\}}\right] = \frac{e^{\theta} - e^{(1-u)\theta}}{e^{\theta} - 1},$$

since $\overline{q}_Z^{\mathbb{P}}(u) = 1 - u$. This proves the first claim of the Lemma. The functions $\psi_Z^{\nu_2}$, $\psi_{Z'}^{\nu_1}$, $\psi_{Z'}^{\nu_2}$ are computed similarly, and using the fact that Z is independent of Z'. The different optimal risk transfer structures are direct applications of Theorem 2.3.3.

$\square$

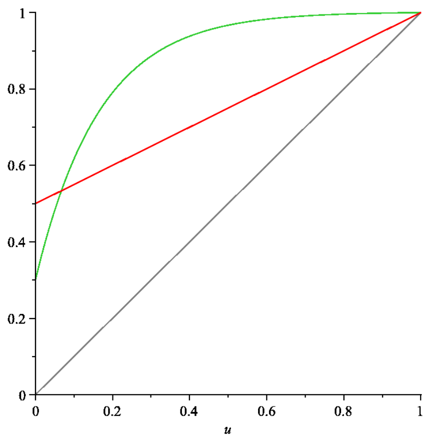

Figure 2.3: The functions $\psi_Z^{\nu_1}$ (green curve) and $\psi_Z^{\nu_2}$(red curve)

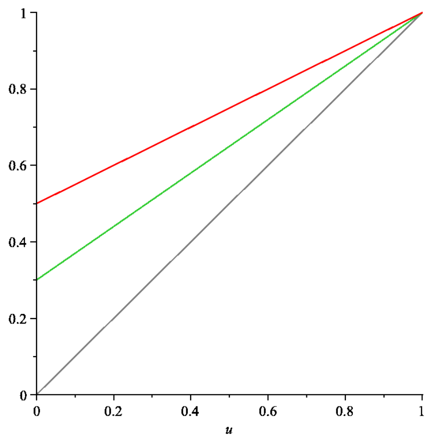

Figure 2.4: The functions $\psi_{Z'}^{\nu_1}$ (green curve) and $\psi_{Z'}^{\nu_2}$ (red curve)

In Figures 2.3 and 2.4 above, we represented the local distortions $\psi_Z^{\nu_i}$ and $\psi_{Z'}^{\nu_i}$, $i = 1, 2$, with $(\varepsilon_1, \varepsilon_2, \theta) = (0.3, 0.5, 6)$.

We clearly see in Figure 2.4 that for the r.v. Z', there is no risk transfer, agent one will assume all the risk, since he is less risk averse than agent two. Whereas Figure 2.3 shows that there is a unique agreement point k^* such that $\rho_1(Z > k^*) = \rho_2(Z > k^*)$, and k^* is approximatevily given by $\overline{q}_Z^{\mathbb{P}}(0.067) = 0.932$.

2.4.3 Volatility uncertainty

Computing the Choquet integral of a r.v X with respect to a submodular capacity is a way to inflate the $\mathbb{P}$-expectation of X, and hence to take into account the model uncertainty we have relatively to the probability measure $\mathbb{P}$.

In the recent literature on model uncertainty, we can find important contributions to the particular case of volatility uncertainty ([101], [37] or [112]). These papers all use probabilities of a certain form that we now describe.

Let $\Omega = \mathcal{C}(\mathbb{R}^+, \mathbb{R}^d)$ be the canonical space of $\mathbb{R}^d$-valued continuous paths defined on $\mathbb{R}^+$, equipped with its Borel σ-field, B the canonical process on Ω and $\mathbb{P}_0$ the Wiener measure. Let $\alpha : \mathbb{R}^+ \to \mathbb{S}_d^{>0}$ be a deterministic function such that $\int_0^1 |\alpha_t|\, dt < +\infty$, where $\mathbb{S}_d^{>0}$ denotes the space of all $d \times d$ real valued positive definite matrices.

We define the probability measure $\mathbb{P}^\alpha$ on Ω by

$$\mathbb{P}^\alpha = \mathbb{P}_0 \circ (X^\alpha)^{-1}, \quad \text{where} \quad X_t^\alpha := \int_0^t \alpha^{1/2}(s)\, dB_s,\ t \in [0,1],\ \mathbb{P}_0 - a.s.$$

which means that $\mathbb{P}^\alpha$ is the law under $\mathbb{P}_0$ of the process X^α. Then a key observation is that the $\mathbb{P}^\alpha$-distribution of B is equal to the $\mathbb{P}_0$-distribution of X^α. In particular, B has a $\mathbb{P}^\alpha$-quadratic variation which is absolutely continuous with respect to the Lebesgue measure dt with density α. Then for different functions α, the measures $\mathbb{P}^\alpha$ offer a way to model the quadratic variation uncertainty, or volatility uncertainty, for the canonical process B.

Assume that two agents use the same distortion function ψ, but disagree on the volatility process. More precisely, let c_1 and c_2 be two submodular capacities on $\mathcal{F}$ given by

$$c_i(A) := \psi[\mathbb{P}^{\alpha_i}(A)],\ i = 1, 2,\ \ A \in \mathcal{F},$$

where ψ is a given distortion function, and α_1, α_2 are two different volatility processes. Let us consider the corresponding Choquet integrals evaluated at B_1:

$$\rho_i(B_1) = \int_{-\infty}^0 [\psi(\mathbb{P}^{\alpha_i}(B_1 > x)) - 1]dx + \int_0^{+\infty} \psi(\mathbb{P}^{\alpha_i}(B_1 > x))dx,\ i = 1, 2.$$

Since the probabilities are different we cannot use this form to find an optimal risk transfer between agent 1 and agent two. Notice that the measures $\mathbb{P}_0, \mathbb{P}^{\alpha_1}$ and $\mathbb{P}^{\alpha_2}$ are mutually singular. But we can use local distortion functions, and do as if $\mathbb{P}_0$ is a reference measure.

Lemma 2.4.2 *We have*

$$\rho_1 \square \rho_2(B_1) = \rho_1(-B_1^-) + \rho_2(B_1^+).$$

where, as usual, $X^- = \max(-X, 0)$ *and* $X^+ = \max(X, 0)$.

Proof.

We can use Lemma 2.3.1, and write everything in terms of $\mathbb{P}_0(B_1 > x)$ as follows.

Since B_1 under $\mathbb{P}_0$ is a standard gaussian r.v, the function $x \to \mathbb{P}_0(B_1 > x)$ has no constant intervals. By Lemma 2.3.1, there exist local distortion functions $\psi_{B_1}^i$ such that

$$\psi(\mathbb{P}^{\alpha_i}(B_1 > x)) = \psi_{B_1}^i(\mathbb{P}_0(B_1 > x)),\ \ x \in \mathbb{R}\ i = 1, 2.$$

We have for example for agent one ($i = 1$)

$$\psi_{B_1}^1(u) = c_1(B_1 > \overline{q}(u)) = \psi[\mathbb{P}^{\alpha_1}(B_1 > x)],$$

where we denoted $\overline{q}(u) := \overline{q}^{\mathbb{P}_0}_{B_1}(u)$ for simplicity. Moreover, using the fact that B is a $\mathbb{P}_0$-Brownian motion, and the Dubins-Schwarz theorem, we get

$$\mathbb{P}^{\alpha_1}(B_1 > x) = \mathbb{P}_0\left(X_1^{\alpha_1} > x\right) = \mathbb{P}_0\left(\int_0^1 \alpha_1^{1/2}(s)\,dB_s > x\right) = \overline{F}\left(\frac{\overline{q}(u)}{\left\|\alpha_1^{1/2}\right\|_{L^2}}\right),$$

where $\overline{F}$ is the tail distribution function of the standard gaussian law. So we found that

$$\psi^1_{B_1}(u) = \psi\left[\overline{F}\left(\frac{\overline{q}(u)}{\left\|\alpha_1^{1/2}\right\|_{L^2}}\right)\right], \quad \text{and of course} \quad \psi^2_{B_1}(u) = \psi\left[\overline{F}\left(\frac{\overline{q}(u)}{\left\|\alpha_2^{1/2}\right\|_{L^2}}\right)\right].$$

The functions $\psi^1_{B_1}$ and $\psi^2_{B_1}$ are equal on $(0,1)$ only at $u^* = 1/2$. Then using Proposition 2.3.2, we get

$$\rho_1 \Box \rho_2(B_1) = \rho_1(B_1 \wedge k^*) + \rho_2((X - k^*)^+),$$

with $k^* = \overline{q}(u^*) = 0$. This ends the proof.

□

In Figure 2.5 below, we represented these two local distortions, in the particular case where $\psi(u) = \sqrt{u}$, and the functions α_1 and α_2 are constants respectively equal to 4 and 9.

Notice that the Choquet integrals ρ_1 and ρ_2 are convex, but the local distortions $\psi^1_{B_1}$ and $\psi^2_{B_1}$ are not concave. This cannot happen for classic distorted probabilities since the Choquet integral with respect to a distorted probability is convex if and only if the distortion is concave, provided the probability space is atomless.

2.4.4 Involving the Entropic Risk Measure

Let us assume here that ρ_1 is the entropic risk measure, defined by:

$$\rho_A(X) = \eta \ln \mathbb{E}_{\mathbb{P}}[\exp(\frac{1}{\eta}X)],\ \eta \in \mathbb{R}^+$$

And define $\rho_2 = AVaR_\alpha$ with α in $(0,1)$.

We cannot use theorem 2 to compute the inf-convolution $\rho_1 \Box \rho_2$ since the entropic risk measure is not comonotonic, however it is shown in [68] that $\rho_1 \Box \rho_2(X) = \rho_1(X \wedge k) + \rho_2((X-k)^+)$ for a certain real value k. We retrieve here this result, and we extend it by computing explicitely the value of k.

Proposition 2.4.1

$$\rho_1 \Box \rho_2(X) = \rho_1(X \wedge k) + \rho_2((X-k)^+), \tag{2.4.1}$$

where $k = \eta - \eta \ln \alpha + c$*, and* c *is a constant defined by* $\mathbb{E}_{\mathbb{P}}\left(\exp(\frac{1}{\eta}(X-c) - 1) \wedge (\frac{1}{\alpha})\right) = 1$.

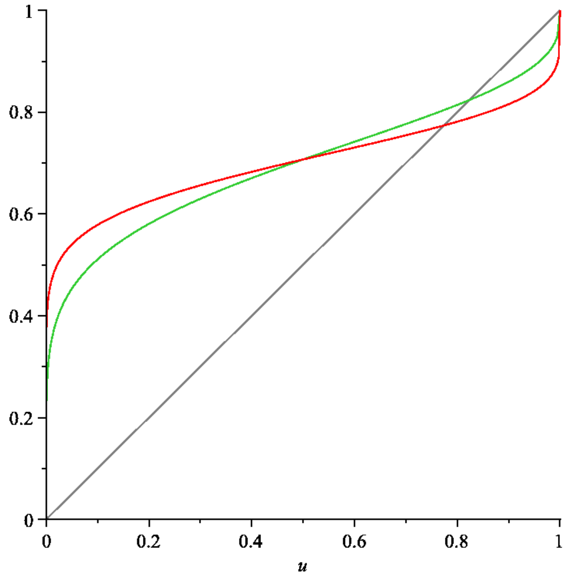

Figure 2.5: The functions $\psi^1_{B_1}$(red curve) and $\psi^2_{B_1}$ (green curve)

Proof.
We write the dual representation of $\rho_1\Box\rho_2(X)$ using as penalty function the sum of penalty functions of the dual representations of ρ_1 and ρ_2:

$$\begin{aligned}\rho_1\Box\rho_2(X) &= \sup\left(\mathbb{E}_{\mathbb{P}}(ZX) - \eta\mathbb{E}_{\mathbb{P}}(Z\ln Z);\ Z \leq \frac{1}{\alpha},\ \mathbb{E}_{\mathbb{P}}(Z) = 1\right)\\ &= \sup\left(\mathbb{E}_{\mathbb{P}}\left(ZX - \eta Z\ln Z - cZ\right);\ Z \leq \frac{1}{\alpha}\right)\\ &= \mathbb{E}_{\mathbb{P}}\left[\sup\left(ZX - \eta Z\ln Z - cZ;\ Z \leq \frac{1}{\alpha}\right)\right]\\ &= \mathbb{E}_{\mathbb{P}}\left(Z_cX - \eta Z_c\ln Z_c\right)\end{aligned}$$

where

$$Z_c = \exp(\frac{1}{\eta}(X-c)-1)\wedge(\frac{1}{\alpha}) \text{ and } c \text{ is such that } \mathbb{E}_{\mathbb{P}}(Z_c) = 1$$

We denote by Q_X the probability measure such that $\frac{dQ_X}{d\mathbb{P}} = Z_c$, and by $\mathrm{H}(Q_X|\mathbb{P})$ the relative entropy of Q_X with respect to $\mathbb{P}$. Since

$$\rho_1\Box\rho_2(X) = E_{Q_X}(X) - \eta\mathrm{H}(Q_X|\mathbb{P}) \text{ with } \frac{dQ_X}{d\mathbb{P}} \leq \frac{1}{\alpha},$$

we have to find a r.v Φ such that $Q_X \in \partial^-\rho_1(\Phi)$, i.e $E_{Q_X}(\Phi) = \mathrm{AVaR}_\alpha(\Phi)$, where we have denoted ∂^- the subgradient operator.
Since

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}(Z_c\Phi) &= \mathbb{E}_{\mathbb{P}}\left[\Phi\left(\exp(\frac{1}{\eta}(X-c)-1)\wedge(\frac{1}{\alpha})\right)\right]\\ &= \mathbb{E}_{\mathbb{P}}\left[\Phi\exp(\frac{1}{\eta}(X-c)-1)\mathbf{1}_{(X\leq k_{c,\eta,\alpha})} + \frac{1}{\alpha}\Phi\mathbf{1}_{(X>k_{c,\eta,\alpha})}\right]\end{aligned}$$

with $k_{c,\eta,\alpha} = \eta - \eta \ln \alpha + c$, we set $\Phi = X \vee k$, $k \in \mathbb{R}$ and we search k such that $E_{Q_X}(X \vee k) = \mathrm{AVaR}_\alpha(X \vee k)$.

$$\mathbb{E}_\mathbb{P}(Z_c(X \vee k_{c,\eta,\alpha})) = \mathbb{E}_\mathbb{P}\left[k_{c,\eta,\alpha} \exp(\frac{1}{\eta}(X-c)-1)\mathbf{1}_{(X \leq k_{c,\eta,\alpha})} + \frac{1}{\alpha} X \mathbf{1}_{(X > k_{c,\eta,\alpha})}\right].$$

Using the fact that $\mathbb{E}_\mathbb{P}(Z_c) = 1$, we deduce that

$$\mathbb{E}_\mathbb{P}\left[\exp(\frac{1}{\eta}(X-c)-1)\mathbf{1}_{(X \leq k_{c,\eta,\alpha})}\right] = 1 - \frac{1}{\alpha}\bar{F}_X(k_{c,\eta,\alpha}).$$

where $\bar{F}_X$ denotes the survival function associated to X (i.e $\bar{F}_X = 1 - F_X$ where F_X is the cumulative distribution function of X).

$$\mathbb{E}_\mathbb{P}(Z_c(X \vee k_{c,\eta,\alpha})) = k_{c,\eta,\alpha}\left(1 - \frac{1}{\alpha}\bar{F}_X(k_{c,\eta,\alpha})\right) + \mathbb{E}_\mathbb{P}\left(\frac{1}{\alpha} X \mathbf{1}_{(X > k_{c,\eta,\alpha})}\right).$$

Let us now compute the quantity $\mathrm{AVaR}_\alpha(X \vee k)$. We easily have that $\forall u \in [0,1]$, $q_{X \vee k}(u) = q_X(u) \vee k$, so that:

$$\begin{aligned}
\mathrm{AVaR}_\alpha(X \vee k) &= \frac{1}{\alpha}\int_{1-\alpha}^{1} q_X(u) \vee k \, du \\
&= \frac{1}{\alpha}\,\mathbb{E}_\mathbb{P}\left[X \vee k \mathbf{1}_{(X > q_X(1-\alpha))}\right] \\
&= \frac{1}{\alpha} k(\alpha - \bar{F}_X(k)) + \mathbb{E}_\mathbb{P}\left[\frac{1}{\alpha} X \mathbf{1}_{(X > k \vee q_X(1-\alpha))}\right].
\end{aligned}$$

Given that

$$\bar{F}_X(k_{c,\eta,\alpha}) = \alpha\left(1 - \mathbb{E}_\mathbb{P}\left[\exp(\frac{1}{\eta}(X-c)-1)\mathbf{1}_{(X \leq k_{c,\eta,\alpha})}\right]\right) \leq \alpha,$$

we have $k_{c,\eta,\alpha} \geq q_X(1-\alpha)$. Using this, we can finally conclude that

$$E_{Q_X}(X \vee k_{c,\eta,\alpha}) = \mathrm{AVaR}_\alpha(X \vee k_{c,\eta,\alpha})$$

We denote $\tilde{k} := k_{c,\eta,\alpha}$. As pointed out in [68], the couple $(X - X \vee \tilde{k}, X \vee \tilde{k}) = (-(X-\tilde{k})^-, X \vee \tilde{k})$ is an optimal risk transfer. Using the cash-additivity property of ρ_A and ρ_R, we see that for any constant $c \in \mathbb{R}$, if $(X - F, F)$ is an optimal risk transfer then $(X - F + c, F - c)$ is again optimal. Since $-(X-\tilde{k})^- + \tilde{k} = X \wedge \tilde{k}$ and $X \vee \tilde{k} - \tilde{k} = (X - \tilde{k})^+$, the couple $(X \wedge \tilde{k}, (X-\tilde{k})^+)$ is optimal. Now rewriting $\mathbb{E}_\mathbb{P}(Z_c) = 1$ using that $c = \tilde{k} - \eta + \eta \ln \alpha$ gives (2.4.2) and the proof is complete.

□

We could actually directly check that $\rho_1 \Box \rho_2(X) = \rho_1(X \wedge k_{c,\eta,\alpha}) + \rho_2([X - k_{c,\eta,\alpha}]^+)$.

Remark 2.4.3 *$k \in \mathbb{R}$ is such that*

$$\rho_A(X \wedge k) = \eta \ln(\alpha) + k. \tag{2.4.2}$$

Chapter 3

A Reinsurance Pricing Problem

Contents

3.1 Introduction **61**

3.1.1 The collective risk theory 61

3.1.2 Links with the existing literature 62

3.2 Pricing of Layers with Reinstatements **62**

3.2.1 The contract payoff 62

3.2.2 The pricing bounds 65

3.1 Introduction

3.1.1 The collective risk theory

There are two different possible ways to apprehend the total loss of an insurance company in a given period of time (typically one year). First, one can consider the *individual* claim amounts S_i, that is to say the loss corresponding to the policy of the i-th insuree, and then, take the sum $I = \sum_{i=1}^{n} S_i$, where n is the number of policies. This kind of individual model of risk dates back to Cramér [32].

This individual model is numerically difficult to implement. Let us however mention the De Pril algorithm, which allows theoretically to obtain the law of I, but it is not widely used in practice (see [98], page 129).

A second possible way to proceed is to consider a random number of claims N in the given period and individual losses X_i, regarless to which policy it is attached. Then the sum $X = \sum_{i=1}^{N} X_i$ also denotes the total loss of the insurance company. It is the traditional *collective model* way to see the total loss (see [41] for a detailed account).

The collective risk model can be analysed dynamically in time, by writting

$$X_t = \sum_{i=1}^{N_t} X_i,$$

where N is a Poisson process counting the number of claims in the interval $[0, t]$. The variables (X_i) are traditionally assumed independent and identically distributed. In that case, X is a pure jump Levy process, for which we have powerful analytic tools.

The collective risk model is widely used among practitioners mainly because of its tractability. Panjer's recursion [96] is a popular technique to approximate the distribution function of the total loss in the collective model.

In this chapter, we will see how this risk modelization is affected by the presence of non-proportional reinsurance contracts.

3.1.2 Links with the existing literature

The indifference price p of a reinsurance layer is solution of an equation indicating that the utility of an insurance company, when it buys the contract and pays the price p, is equal to its utility when it does not enter the transaction. When the contract contains reinstatements, the total premium paid is random (see equation (3.2.3)). This can complicate the calculation of the indifference price. In the next sections we will bound this price by two easily computable values.

Despite the huge literature on reinsurance problems, there are little references concerning the pricing of reinsurance contracts containing reinstatements. Sundt [114] obtained a formula for the pure premium of contracts with reinstatements ; this formula coincide with ours in this particular case. Sundt [114] and Wahlin [119] also give formulas corresponding to other premium principles, such that the standard deviation principle or the PH-transform. Mata [86] studies the properties of different pricing principles. Now that we better understand the properties and axioms of convex monetary risk measures, we can compare it with the properties stated by Mata, who defines indifference prices with respect to various comonotonic risk measures, but using the approximation $\rho(-X) \approx -\rho(X)$. If ρ is convex and normalized, we can only say that $\rho(-X) \geq -\rho(X)$. That is why the values computed in [86] correspond, in the particular case where there is no cost of capital, to our upper bound (p_2 appearing in equation (3.2.6)).

3.2 Pricing of Layers with Reinstatements

3.2.1 The contract payoff

Let u_0 be the initial surplus of an insurance company, and let β be the gross premium income per time unit, so that the surplus process R without reinsurance is given at time t by

$$R(t) = u_0 + \beta t - \sum_{i=1}^{N_t} X_i$$

where $\{N_t,\ t \geq 0\}$ is a non decreasing integer valued pure jump process modelling the number of claims up to time t, and $\{X_i,\ i \in \mathbb{N}^*\}$ is a sequence of independent and identically distributed (i.i.d.) random variables modelling the individual claim sizes of the insurance portfolio.

In an XL reinsurance contract with retention l and limit m, the reinsurer covers the part of each claim that exceeds l, up to the upper bound $l + m$. The reinsurer's part Z_i and the insurer's part C_i related to a claim X_i are then given by

$$Z_i = (X_i - l)^+ - (X_i - l - m)^+ = \begin{cases} 0 \text{ if } X_i \le l \\ X_i - l, \text{ if } l \le X_i \le l + m \\ m \text{ if } X_i \ge l + m \end{cases} \tag{3.2.1}$$

and

$$C_i = (X_i \wedge l) + (X_i - l - m)^+ = \begin{cases} X_i \text{ if } X_i \le l \\ l, \text{ if } l \le X_i \le l + m \\ X_i - m \text{ if } X_i \ge l + m \end{cases} \tag{3.2.2}$$

Let

$$Z(t) := \sum_{i=1}^{N_t} Z_i$$

be the aggregate liability of the reinsurer. It is usually limited to a maximal amount M.

When M is a multiple of the individual maximum cover m, i.e when $M = (k+1)m$, we say that the reinsurance contract contains k reinstatements. In that case the aggregate liability of the reinsurer is given by $R_k(t) := Z(t) \wedge ((k+1)m)$.

The cover of the j-th reinstatement is given by

$$r_j := \min\left[\max(Z(t) - jm, 0), m\right]$$

For the j-th reinstatement, the cedent pays a premium P_j given by

$$P_j := c_j p_0 \frac{1}{m} \min\left[\max(Z(t) - (j-1)m, 0), m\right]$$

Let us consider a concrete example, allowing to better understand these contracts payoffs. Assume that the insurance company enters an XL contract with one year maturity, including two reinstatements and covering the losses between 100 and 200 ($l = 100$ and $m = 100$). The initial premium required by the reinsurer is given by $p_0 = 10$. The costs of the first and second reinstatements correspond respectively to 80% and 50% of the initial premium ($c_1 = 0.8$ and $c_2 = 0.5$).

Assume the first four losses correspond to the following amounts:

1. $X_1 = 150$, we have then $C_1 = 100$ and $Z_1 = 50$. The insurance company has to reinstate half of the layer, to do so it pays $50 * \frac{1}{100} * 80\% * 10 = 4$.

2. $X_2 = 190$, we have then $C_1 = 100$ and $Z_1 = 90$. The ceding company has to reinstate 90% of the layer. For this, it uses what is left of the first reinstatement (50) and a part of the second one (40). So it pays :
$50 * \frac{1}{100} * 80\% * 10 + 40 * \frac{1}{100} * 50\% * 10 = 6$

3. $X_3 = 200$, in that case all the layer has to be reinstated, but only 60% of the second reinstatement is available, the layer is then reinstated only up to this limit and the company pays: $60 * \frac{1}{100} * 50\% * 10 = 3$.

4. $X_4 = 200$, the insurance company was only covered for the losses between 100 and 160, we have then here $Z_4 = 60$ and $C_4 = 140$. All the reinstatements have been used, the contract stops and the ceding company is not reinsured for the rest of the year.

The total premium paid in this example amounts to $\overline{P} = 10 + 4 + 6 + 3 = 23$. This value could have been lower if all the reinstatements had not been used before the end of the contract.

We can write the random total premium paid by the cedent during the considered period in the following way:

$$\overline{P}(t) := p_0 \left(1 + \frac{1}{m} \sum_{j=1}^{k} c_j \min\left[\max(Z(t) - (j-1)m, 0), m\right] \right) \tag{3.2.3}$$

and the wealth process of the cedent is given by

$$\begin{aligned} R(t) &= u_0 + \beta t - \sum_{i=1}^{N_t} C_i - \overline{P}(t) \\ &= u_0 + \beta t - \sum_{i=1}^{N_t} X_i + \sum_{i=1}^{N_t} Z_i - \overline{P}(t). \end{aligned}$$

In what follows, we will denote for simplicity

$$\widetilde{N}(t) := \frac{1}{m} \sum_{j=1}^{k} c_j \min\left[\max(Z(t) - (j-1)m, 0), m\right]$$

Finally, let us now consider the more realistic situation where the cedent has already bought reinsurance contracts which are still running, and is considering to add a given layer $[l, l+m]$ with k reinstatements to its reinsurance program. Then we will denote by $R(t)$ the reference net result of the cedent (without the additional layer) and we denote F^{XL} the different quantities concerning the additional XL layer.

We have for instance

$$\begin{aligned} R^{XL}(t) &= u_0 + \beta t - \sum_{i=1}^{N_t} X_i + \sum_{i=1}^{N_t} Z_i - \overline{P} + \sum_{i=1}^{N_t} Z_i^{XL} - \overline{P}^{XL} \\ &= R(t) + \sum_{i=1}^{N_t} Z_i^{XL} - \overline{P}^{XL}, \end{aligned} \tag{3.2.4}$$

where $\overline{P}^{XL}$ is the total premium paid by the cedent's company for the additional XL layer.

3.2.2 The pricing bounds

We assume that the cedent uses a concave monetary utility function U, and a convex monetary risk measure ρ to compute its needed regulatory capital.

Definition 3.2.1 *We say that $\widetilde{P}_0$ is the indifference price of a given XL layer relatively to the pair (U, ρ), if $\widetilde{P}_0$ solves the equation*

$$U\left(R(t) - \overline{c}\,\rho(R(t))\right) = U\left(R^{XL}(t) - \overline{c}\,\rho(R^{XL}(t))\right) \tag{3.2.5}$$

where $\overline{c}$ is a given cost of capital.

When the insurance company does not buy the XL layer, its net result is given by $R(t)$, and its regulatory needed capital is expressed by $\rho(R(t))$. If the XL layer is added to the current reinsurance program, then the net result becomes $R^{XL}(t)$ and there is possibly a regulatory capital saving thanks to the layer that we can write as $\rho(R(t)) - \rho(R^{XL}(t))$.

Remark 3.2.1 *If we define $\overline{U}(R) := U(R - \overline{c}\,\rho(R))$, then we have*

$$\overline{U}(R+m) = \overline{U}(R) + (1+\overline{c})\,m \geq \overline{U}(X) + m,$$

and $\overline{U}$ is a cash super-additive utility function. This is analogous to the presence of a capitalization factor, which is here given by $\overline{c}$.

We have the following estimate for the indifference price of an XL layer with reinstatements

Proposition 3.2.1 *If $\widetilde{P}_0$ is the indifference price of a given XL layer relatively to the pair (U, ρ), then*

$$p_1 \leq \widetilde{P}_0 \leq p_2, \tag{3.2.6}$$

where

$$p_1 := \frac{A}{-\overline{U}(-1-\widetilde{N})}, \qquad p_2 := \frac{A}{\overline{U}(1+\widetilde{N})},$$

and

$$A := \overline{U}\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) - \overline{U}(R(t)).$$

Proof.

We use the decomposition (3.2.4) of R^{XL} and the cash invariance property to deduce that equation (3.2.5) is equivalent to

$$\begin{aligned} &U\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) - \overline{c}\,\rho\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) \\ &= U\left(R(t)\right) - \overline{c}\rho(R(t)). \end{aligned} \tag{3.2.7}$$

We can use the fact that U is super-additive whereas ρ is subadditive, and that both are positively homegeneous to write

$$U\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) \geq U\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) + \widetilde{P}_0\, U\left(-(1+\widetilde{N}(t))\right),$$

$$\rho\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) \leq \rho\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) + \widetilde{P}_0\, \rho\left(-(1+\widetilde{N}(t))\right).$$

By plugging this into (3.2.7) we obtain

$$\begin{aligned}
&\widetilde{P}_0 \left[U\left(-(1+\widetilde{N}(t))\right) - \bar{c}\,\rho\left(-(1+\widetilde{N}(t))\right)\right] \\
&\leq U(R(t)) - U\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) - \bar{c}\left[\rho(R(t)) - \rho\left(R(t) + \sum_{i=1}^{N_t} Z_i\right)\right]
\end{aligned} \tag{3.2.8}$$

We notice that $-(1+\widetilde{N}(t)) < 0$ a.s. and since U is monotone increasing and ρ is monotone decreasing, we have

$$U\left(-(1+\widetilde{N}(t))\right) - \bar{c}\,\rho\left(-(1+\widetilde{N}(t))\right) < 0.$$

Then we deduce from inequality (3.2.8) that $p_1 \leq \widetilde{P}_0$.

To prove the other inequality, we start again from equation (3.2.7), and notice this time that since U is super-additive whereas ρ is subadditive, we have $U(X-Y) \leq U(X) - U(Y)$ and $\rho(X-Y) \geq \rho(X) - \rho(Y)$, for any bounded random variables X and Y. This gives

$$\begin{aligned}
U\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) &\leq U\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) - \widetilde{P}_0\, U\left(1+\widetilde{N}(t)\right), \\
\rho\left(R(t) + \sum_{i=1}^{N_t} Z_i - \widetilde{P}_0(1+\widetilde{N}(t))\right) &\geq \rho\left(R(t) + \sum_{i=1}^{N_t} Z_i\right) - \widetilde{P}_0\, \rho\left(1+\widetilde{N}(t)\right).
\end{aligned}$$

By plugging this again into (3.2.7), and noticing that

$$U\left(1+\widetilde{N}(t)\right) - \bar{c}\,\rho\left(1+\widetilde{N}(t)\right) > 0,$$

we obtain $\widetilde{P}_0 \leq p_2$.

□

Remark 3.2.2 *The monotonicity properties of U and ρ implies that $0 \leq p_1$. Furthermore, the concavity and convexity properties of U and ρ implies that we have indeed $p_1 \leq p_2$.*

In figures 3.1 and 3.2 below, we represented the upper and lower bounds p_1 and p_2 obtained for the layers with retention equal to $10^8(n+1)$ euros and limit equal to 10^8 euros, for n varying from 0 to 32. In other words, we computed the bound prices for 33 XL contracts, with the same limit value, for a retention ranging from one hundred million euros to 3.3 billion euros. All the contracts contain $k=4$ possible reconstitutions, and each reconstitution when it happens, is paid at $c=100\%$ of the initial price. The prices are expressed as a percentage of $m=10^8$, which correspond to rates on line in the reinsurance terminology.

These reisurance contracts cover an insurance portfolio of risk of natural catastrophes type. The calculationshave been made using 25000 years of simulations of losses from a portfolio of Axa.

We took as a matter of example a utility function given by the semi-deviation utility with parameter $\delta = \frac{1}{2}$ and as risk measures, the $AVaR_\alpha$ with $\alpha = 1/200$ in Figure 3.1 and the Wang transform risk measure with parameter $\gamma = -2$ in Figure 3.2. Those functionals are defined in Chapter 2.

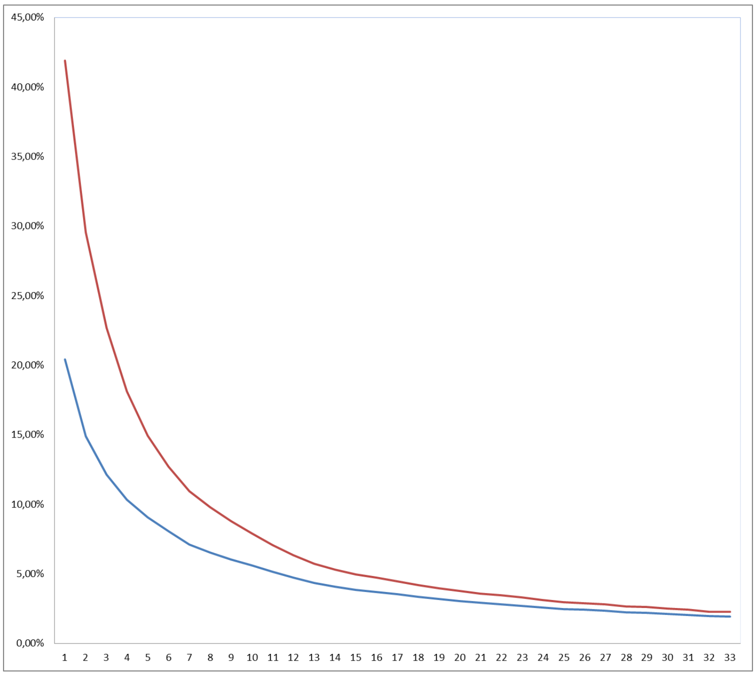

Figure 3.1: Semi-deviation utility function and $AVaR_\alpha$ risk measure

We can see that the gap between p_1 and p_2 is important for the first contracts. This is due to the fact that the first contracts correspond to low retentions, for which all the reconstitutions are systematically used. More precisely, if $k = 4$, the value $\widetilde{N}$ appearing in the definitions of p_1 and p_2 will be we very often equal to 4 for the contracts with a low retention, and the different quantities $-\overline{U}(-1-4)$ and $U(1+4)$ will contribute considerably to the gap between p_1 and p_2.

However for the contracts with a high retention, begining from 1.2 billion, the bounds are rather satisfying.

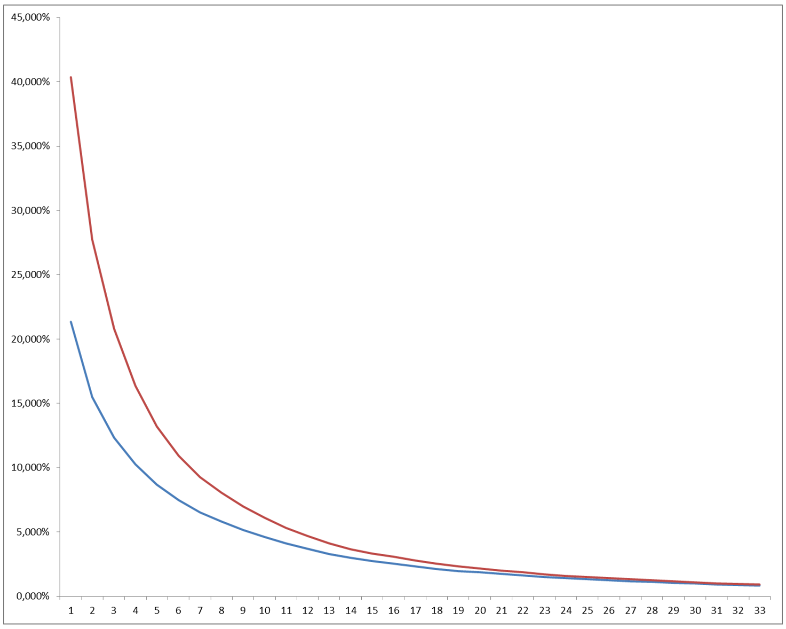

Figure 3.2: Semi-deviation utility function and Wang transform risk measure

Chapter 4

Quadratic Backward SDEs with Jumps

Contents

4.1 Introduction **69**

4.2 An existence and uniqueness result **72**

4.2.1 Notations 72

4.2.2 Standard spaces and norms 73

4.2.3 A word on càdlàg BMO martingales 74

4.2.4 The non-linear generator 77

4.2.5 A priori estimates 78

4.2.6 Existence and uniqueness for a small terminal condition 82

4.2.7 Existence for a bounded terminal condition 88

4.2.8 A uniqueness result 92

4.2.9 A priori estimates and stability 98

4.3 Quadratic g-martingales with jumps **100**

4.3.1 Non-linear Doob Meyer decomposition 103

4.3.2 Upcrossing inequality 109

4.4 Dual Representation and Inf-Convolution **113**

4.4.1 Dual Representation of the g-expectation 113

4.4.2 Inf-Convolution of g-expectations 116

4.4.3 Examples of inf-convolution 118

4.1 Introduction

Motivated by duality methods and maximum principles for optimal stochastic control, Bismut studied in [15] a linear backward stochastic differential equation (BSDE). In their seminal paper [97], Pardoux and Peng generalized such equations to the non-linear Lipschitz case and proved existence and uniqueness results in a Brownian framework. Since then, a lot of attention has been given to BSDEs and their applications, not only in stochastic control, but also in theoretical economics, stochastic differential games and financial mathematics.

In this context, the generalization of Backward SDEs to a setting with jumps enlarges again the scope of their applications, for instance to insurance modeling, in which jumps are inherent. Li and Tang [81] first proved the wellposedness of Lipschitz BSDEs with jumps, using a fixed point approach simililar to the one used in [97].

Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ generated by an $\mathbb{R}^d$-valued Brownian motion B and a random measure μ with compensator ν, solving a BSDE with generator g, and terminal condition ξ consists in finding a triple of progressively measurable processes (Y, Z, U) such that for all $t \in [0, T]$, $\mathbb{P} - a.s.$

$$Y_t = \xi + \int_t^T g_s(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_{\mathbb{R}^d \setminus \{0\}} U_s(x)(\mu - \nu)(ds, dx). \tag{4.1.1}$$

See Section 4.2.1 for more precise definitions and notations.

In this chapter, g will be supposed to satisfy a Lipschitz-quadratic growth property. More precisely, g will be Lipschitz in y, and will satisfy the quadratic growth condition (iii) of Assumption 4.2.1.

When the filtration is generated only by a Brownian motion, the existence and uniqueness of quadratic BSDEs with a bounded terminal condition has been first treated in Kobylanski [74], by means of an exponential transform method, allowing to fall into the scope of BSDEs with a coefficient having a linear growth, then the result for quadratic BSDEs is obtained by an approximation operation. Tevzadze [117] has also given a direct proof for the wellposedness in the Lipschitz-quadratic setting. His methodology is fundamentally different, since he uses fixed-point approach to obtain existence of a solution for small terminal condition, and then pastes solutions together in the general bounded case. More recently, Briand and Hu [18] went beyond the bounded case to obtain an existence result for quadratic BSDEs with a terminal condition having finite exponential moments. In a similar vein, but using a forward point of view and stability results for special class of quadratic semimartingales, Barrieu and El Karoui [7] generalized the above results.

Nonetheless, when it comes to quadratic BSDEs in a discontinuous setting, the literature is less abounding. Until very recently, the only existing results concerned particular cases of quadratic BSDEs, mainly with links to applications to utility maximization or indifference pricing problems. Thus, Becherer [9] studied first bounded solutions to BSDEs with jumps in a finite activity setting, and his general results were improved by Morlais [91], who proved existence of the solution to a special quadratic BSDE with jumps, using the same type of techniques as Kobylanski. More recently, Ngoupeyou [94], in his PhD thesis, extended partly the approach of [7] to the jump case. Laeven and Stadje [77] proved recently a general existence result for BSDEJs with convex generators, using verification arguments. We emphasize that our approach is very different and do not need any convexity assumption, even though the two results do not imply each other. Finally, El Karoui, Matoussi and Ngoupeyou [47] extended the approach of [7] to the jump case, and hence obtained complete results for unbounded terminal conditions.

Our first aim in this chapter is to extend the fixed-point methodology of Tevzadze [117] to the case of discontinuous filtration. We first obtain our result for a terminal condition ξ having a $\|\cdot\|_\infty$-norm which is small enough. Then the result for any ξ in $\mathbb{L}^\infty$ follows by splitting ξ in pieces having a small enough norm, and then pasting the obtained solutions to a single equation. Since we deal with bounded solutions, the space of BMO martingales will play a particular role in our setting, this will be the natural space for the continuous and the pure jump martingale terms appearing in the BSDE (4.1.1), when Y is bounded.

In this framework with jumps, we need additional assumptions on the generator g for a comparison theorem to hold. Namely, we will use the Assumption 4.2.5, first introduced by Royer [106] in order to ensure the validity of a comparison theorem for Lipschitz BSDEs with jumps. We extend here this comparison theorem to our setting (Proposition 4.2.7), and then use it to give a uniqueness result.

This wellposedness result for bounded quadratic BSDEs with jumps opens the way to many possible applications. Barrieu and El Karoui [6] used quadratic BSDEs to define time consistent convex risk measures and studied their properties. We extend here some of these results to the case with jumps. When the generator g is independent of y and convex in (z, u), we can define through the solution of the BSDE a convex operator acting on the terminal condition. This operator, called g-expectation, has been first studied by Peng [99], and then extended to the case of quadratic coefficients by Ma and Yao [84], or to discontinuous filtrations by Royer [106] and Lin [83].

In this chapter, we go further in the study of quadratic BSDEs with jumps by proving a non-linear Doob Meyer decomposition for g-submartingales. As a consequence, we obtain a converse comparison theorem.

The last two results are dedicated to an explicit dual representation of the solution Y, when g is independent of y and convex in (z, u). This allows to study some particular risk measures on a discontinuous filtration, as the *entropic risk measure*, corresponding to the solution of a quadratic BSDE. Finally, we prove an explicit representation for the inf-convolution of quadratic BSDEs, thus giving the form of the optimal risk transfer between two agents using quadratic convex g-expectations as risk measures. The inf-convolution is again a convex operator, solving a particular BSDE. We give a sufficient condition for this BSDE to have a coefficient satisfying a quadratic growth property.

The rest of this chapter is organized as follows. In Section 4.2, we obtain our general existence and uniqueness result, then in Section 4.3, we study general properties of quadratic g-martingales with jumps, such as regularity in time and the Doob-Meyer decomposition. Finally, Section 4.4 is devoted to the two applications mentioned above.

4.2 An existence and uniqueness result

4.2.1 Notations

Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ generated by a $\mathbb{R}^d$-valued Brownian motion B, solving a BSDE with generator g, and terminal condition ξ consists in finding a pair of progressively measurable processes (Y, Z) such that

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s,\ \mathbb{P} - a.s,\ t \in [0, T]. \tag{4.2.1}$$

The process Y we define this way is a possible generalization of the conditional expectation of ξ, since when g is the null function, we have $Y_t = \mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t]$, and in that case, Z is the process appearing in the $(\mathcal{F}_t)$-martingale representation property of $\{\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t], t \geq 0\}$.

In the case of a filtered probability space generated by both a Brownian motion B and a Poisson random measure μ with compensator ν, the martingale representation for $\{\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t], t \geq 0\}$ becomes

$$\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_t] = \int_0^t Z_s dB_s + \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} U_s(x) \widetilde{\mu}(dx, ds),$$

where U is a predictable function, and $\widetilde{\mu} = \mu - \nu$.

This leads to the following natural generalization of equation (4.2.1) to the case with jumps.

Definition 4.2.1 *Let ξ be a $\mathcal{F}_T$-measurable random variable. A solution of the BSDEJ with terminal condition ξ and generator g is a triple (Y, Z, U) of progressively measurable processes such that*

$$Y_t = \xi + \int_t^T g_s(Y_s, Z_s, U_s) ds - \int_t^T Z_s dB_s - \int_t^T \int_{\mathbb{R}^d \setminus \{0\}} U_s(x) \widetilde{\mu}(dx, ds), \ 0 \leq t \leq T, \ \mathbb{P} - a.s. \quad (4.2.2)$$

where $g : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{A}(E) \to \mathbb{R}$ is a given application and

$$\mathcal{A}(E) := \left\{ u : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}, \, \mathcal{B}(\mathbb{R}^d \setminus \{0\}) - \text{measurable} \right\}.$$

The processes Z and U are supposed to satisfy the minimal assumptions so that the quantities in (4.2.2) are well defined, namely $(Z, U) \in \mathcal{Z} \times \mathcal{U}$, where $\mathcal{Z}$ denotes the space of all $\mathbb{F}$-predictable $\mathbb{R}^d$-valued processes Z with

$$\int_0^T |Z_t|^2 \, dt < +\infty, \ \mathbb{P} - a.s.$$

and $\mathcal{U}$ denotes the space of all $\mathbb{F}$-predictable functions U with

$$\int_0^T \int_{\mathbb{R}^d \setminus \{0\}} |U_s(x)|^2 \, \nu_t(dx) ds < +\infty, \ \mathbb{P} - a.s.$$

Notice that in this discontinuous setting, the generator g depends on both Z and U. Here U plays a role analogous to the quadratic variation in the continuous case. However, there are some notable differences, since for each t, U_t is a function from E into $\mathbb{R}^d \setminus \{0\}$, and that is why the treatment of the dependence in u in the assumptions for the generator is not symmetric to the treatment of the dependence in z, and in particular we deal with Fréchet derivatives with respect to u. See for instance Assumption 4.2.3.

When we pass from the continuous setting to the discontinuous one, the natural analogue of the integral with respect to the Brownian motion would be the integral with respect to a compensated Poisson random measure. The compensator would then not depend on t, neither on ω, by comparison to the non dependence in ω of the quadratic variation of the Brownian motion.

However we will allow the compensator of the jump measure that we will consider to depend on both t and ω. This will only slightly increase the complexity of our proofs, provided that the martingale representation property of Assumption 4.2.2 holds true.

We consider in all this chapter a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P}\right)$, whose filtration satisfies the usual hypotheses of completeness and right-continuity. We suppose that this filtration is generated by a d-dimensional Brownian motion B and an independent integer valued random measure

$\mu(\omega, dt, dx)$ defined on $\mathbb{R}^+ \times E$, with compensator $\lambda(\omega, dt, dx)$, where $E := \mathbb{R}^d \backslash \{0\}$. $\widetilde{\Omega} := \Omega \times \mathbb{R}^+ \times E$ is equipped with the sigma-field $\widetilde{\mathcal{P}} := \mathcal{P} \times \mathcal{E}$, where $\mathcal{P}$ denotes the predictable σ-field on $\Omega \times \mathbb{R}^+$ and $\mathcal{E}$ is the Borel σ-field on E.

To guarantee the existence of the compensator $\lambda(\omega, dt, dx)$, we assume that for each A in $\mathcal{B}(E)$ and each ω in Ω, the process $X_t := \mu(\omega, A, [0, t]) \in \mathcal{A}^+_{loc}$, which means that there exists an increasing sequence of stopping times (T_n) such that $T_n \to +\infty$ a.s. and the stopped processes $X_t^{T_n}$ are increasing, cádlág, adapted and satisfy $\mathbb{E}[X_\infty] < +\infty$.

We assume in all this chapter that λ is absolutely continuous with respect to the Lebesgue measure dt, i.e. $\lambda(\omega, dt, dx) = \nu_t(\omega, dx)dt$. Finally, we denote $\widetilde{\mu}$ the compensated jump measure:

$$\widetilde{\mu}(\omega, dx, dt) = \mu(\omega, dx, dt) - \nu_t(\omega, dx)\, dt.$$

Notation 4.2.1 *In the rest of the chapter, we will denote $\mathbb{E}_t(\cdot)$ the conditional expectation with respect to the filtration $(\mathcal{F}_t)$.*

4.2.2 Standard spaces and norms

We introduce the following norms and spaces for any $p \geq 1$.

$\mathcal{S}^p$ is the space of $\mathbb{R}$-valued càdlàg and $\mathcal{F}_t$-progressively measurable processes Y such that

$$\|Y\|^p_{\mathcal{S}^p} := \mathbb{E}^{\mathbb{P}}\left[\sup_{0 \leq t \leq T} Y_t^p\right] < +\infty.$$

$\mathcal{S}^\infty$ is the space of $\mathbb{R}$-valued càdlàg and $\mathcal{F}_t$-progressively measurable processes Y such that

$$\|Y\|_{\mathcal{S}^\infty} := \sup_{0 \leq t \leq T} \|Y_t\|_\infty < +\infty.$$

$\mathbb{H}^p$ is the space of $\mathbb{R}^d$-valued and $\mathcal{F}_t$-progressively measurable processes Z such that

$$\|Z\|^p_{\mathbb{H}^p} := \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |Z_t|^2\, dt\right)^{\frac{p}{2}}\right] < +\infty.$$

The three spaces above are the classical ones in the BSDE theory in continuous filtrations. We introduce finally a space which is specific to the jump case, and which plays the same role for U as $\mathbb{H}^p$ for Z.

$\mathbb{J}^p$ is the space of predictable and $\mathcal{E}$-measurable applications $U : \Omega \times [0, T] \times E$ such that

$$\|U\|^p_{\mathbb{J}^p} := \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T \int_E |U_s(x)|^2\, \nu_s(dx)ds\right)^{\frac{p}{2}}\right] < +\infty.$$

4.2.3 A word on càdlàg BMO martingales

The recent literature on quadratic BSDEs is very rich on remarks and comments about the very deep theory of continuous BMO martingales. However, it is clearly not as well documented when it comes to càdlàg BMO martingales, whose properties are crucial ingredients in this chapter. Indeed, apart from some remarks in the book by Kazamaki [72], the extension to the càdlàg case of the classical results of BMO theory, cannot always be easily found. Our main goal in this short subsection is to give a rapid overview of the existing literature and results concerning BMO martingales with càdlàg trajectories, with an emphasis where the results differ from the continuous case. Let us start by recalling some notations and definitions.

BMO is the space of square integrable càdlàg $\mathbb{R}^d$-valued martingales M such that

$$\|M\|_{\mathrm{BMO}} := \sup_{\tau\in\mathcal{T}_0^T} \left\|\mathbb{E}^{\mathbb{P}}_\tau\left[(M_T - M_{\tau^-})^2\right]\right\|_\infty < +\infty,$$

where $\mathcal{T}_0^T$ is the set of $\mathcal{F}_t$ stopping times taking their values in $[0,T]$.

$\mathbb{J}^2_{\mathrm{BMO}}$ is the space of predictable and $\mathcal{E}$-measurable applications $U:\Omega\times[0,T]\times E$ such that

$$\|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} := \left\|\int_0^\cdot \int_E U_s(x)\widetilde{\mu}(dx,ds)\right\|_{\mathrm{BMO}} < +\infty.$$

$\mathbb{H}^2_{\mathrm{BMO}}$ is the space of $\mathbb{R}^d$-valued and $\mathcal{F}_t$-progressively measurable processes Z such that

$$\|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} := \left\|\int_0^\cdot Z_s dB_s\right\|_{\mathrm{BMO}} < +\infty.$$

As soon as the process $<M>$ is defined for a martingale M, which is the case if for instance M is locally square integrable, then it is easy to see that $M\in\mathrm{BMO}$ if the jumps of M are uniformly bounded in t by some positive constant C and

$$\sup_{\tau\in\mathcal{T}_0^T}\left\|\mathbb{E}^{\mathbb{P}}_\tau\left[<M>_T - <M>_\tau\right]\right\|_\infty \leq C.$$

Furthermore the BMO norm of M is then smaller than $2C$.

We also recall the so called energy inequalities (see [72] and the references therein). Let $Z\in\mathbb{H}^2_{\mathrm{BMO}}$, $U\in\mathbb{J}^2_{\mathrm{BMO}}$ and $p\geq 1$. Then we have

$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |Z_s|^2\,ds\right)^p\right] \leq 2p!\left(4\,\|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}}\right)^p \tag{4.2.3}$$

$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T\int_E U_s^2(x)\nu_s(dx)ds\right)^p\right] \leq 2p!\left(4\,\|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}}\right)^p. \tag{4.2.4}$$

Let us now turn to more precise properties and estimates for BMO martingales. It is a classical result (see [72]) that the Doléans-Dade exponential of a continuous BMO martingale is a uniformly integrable martingale. Things become a bit more complicated in the càdlàg case, and more assumptions are needed. Let us first define the Doléans-Dade exponential of a square integrable martingale X, denoted $\mathcal{E}(x)$. This is as usual (see for instance Protter [105]) the unique solution Z of the SDE

$$Z_t = 1 + \int_0^t Z_{s^-} dX_s, \ \mathbb{P} - a.s.,$$

and is given by the formula

$$\mathcal{E}(X)_t = e^{X_t - \frac{1}{2} <X^c>_t} \prod_{0<s\leq t} (1 + \Delta X_s) e^{-\Delta X_s}.$$

One of the first results concerning Doléans-Dade exponential of BMO martingales was proved by Doléans-Dade and Meyer [42]. They showed that

Proposition 4.2.1 *Let M be a càdlàg BMO martingale such that $\|M\|_{\text{BMO}} < 1/8$. Then $\mathcal{E}(M)$ is a strictly positive uniformly integrable martingale.*

The constraint on the norm of the martingale being rather limiting for applications, this result was subsequently improved by Kazamaki [70], where the constraints is now on the jumps of the martingale

Proposition 4.2.2 *Let M be a càdlàg BMO martingale such that there exists $\delta > 0$ with $\Delta M_t \geq -1 + \delta$, for all $t \in [0, T]$, $\mathbb{P} - a.s.$ Then $\mathcal{E}(M)$ is a strictly positive uniformly integrable martingale.*

Furthermore, we emphasize, as recalled in the counter-example of Remark 2.3 in [72], that a complete generalization to the càdlàg case is not possible. We also refer the reader to Lépingle and Mémin [79] and [80] for general sufficient conditions for the uniform integrability of Doléans-Dade exponentials of càdlàg martingales. This also allows us to obtain immediately a Girsanov Theorem in this setting, which will be extremely useful throughout the chapter.

Proposition 4.2.3 *Let us consider the following càdlàg martingale M*

$$M_t := \int_0^t \varphi_s dB_s + \int_0^t \int_E \gamma_s(x) \widetilde{\mu}(dx, ds), \ \mathbb{P} - a.s.,$$

where $(\varphi, \gamma) \in \mathbb{H}^2_{\text{BMO}} \times \mathbb{J}^2_{\text{BMO}} \cap L^\infty(\nu)$ and where there exists $\delta > 0$ with $\gamma_t \geq -1 + \delta$, $\mathbb{P} \times d\nu_t - a.e.$, for all $t \in [0, T]$.

Then, the probability measure $\mathbb{Q}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}(M_.)$, is indeed well-defined and starting from any $\mathbb{P}$-martingale, by, as usual, changing adequately the drift and the jump intensity, we can obtain a $\mathbb{Q}$-martingale.

We now address the question of the so-called reverse Hölder inequality, which implies in the continuous case that if M is a BMO martingale, there exists some $r > 1$ such that $\mathcal{E}(M)$ is L^r-integrable. As for the previous result on uniform integrability, this was extended to the càdlàg case first in [42] and [71], with the additional assumption that the BMO norm or the jumps of M are sufficiently small. The following generalization is taken from [64]

Proposition 4.2.4 *Let M be a càdlàg BMO martingale such that there exists $\delta > 0$ with $\Delta M_t \geq -1+\delta$, for all $t \in [0,T]$, $\mathbb{P}-a.s.$ Then $\mathcal{E}(M)$ is in L^r for some $r > 1$.*

We finish this short survey by considering the so-called Muckenhoupt condition. In [72], Kazamaki showed that the BMO property a of continuous martingale M was equivalent to the existence of some $p > 1$ such that $\mathbb{P}-a.s.$

$$\sup_{0\leq t\leq T} \mathbb{E}_t^{\mathbb{P}}\left[\left(\frac{\mathcal{E}(M)_t}{\mathcal{E}(M)_T}\right)^{\frac{1}{p-1}}\right] \leq C_p, \tag{4.2.5}$$

for some constant C_p depending only on p. The generalization to the càdlàg case has once more been obtained by Kazamaki [71] who showed that

Proposition 4.2.5 *Let M be a càdlàg martingale. Then*

(i) *If $-1 < \Delta M_t \leq C$ for some constant C and for every t, and if M satisfies (4.2.5) for some $p > 1$, then M is BMO.*

(ii) *If M is BMO, then there exists some $a > 0$ such that $-1 < a\Delta M_t \leq C_a$ and such that aM satisfies (4.2.5) for some $p > 1$.*

4.2.4 The non-linear generator

Recall the Definition 4.2.1 of backward SDEs with jumps. We need now to specify in more details the assumptions we make on the generator g. The most important assumption in our setting in our setting will be the quadratic assumption of Assumption 4.2.1 (ii) below. This assumption appears in the paper by El Karoui, Matoussi and Ngoupeyou [47]. It is the natural generalization to the jump case of the usual quadratic growth assumption in z. Before proceeding further, let us define the following function

$$j_t(u) := \int_E \left(e^{u(x)} - 1 - u(x)\right)\nu_t(dx).$$

This function $j(u)$ plays the same role for the u variable as the square function for the variable z. In order to understand this, let us consider the following "simplest" quadratic BSDE with jumps

$$y_t = \xi + \int_t^T \left(\frac{\gamma}{2}|z_s|^2 + \frac{1}{\gamma}j_s(\gamma u_s)\right)ds - \int_t^T z_s dB_s - \int_t^T \int_E u_s(x)\widetilde{\mu}(dx,ds),\ \mathbb{P}-a.s.$$

Then a simple application of Itô's formula gives formally

$$e^{\gamma y_t} = e^{\gamma\xi} - \int_t^T e^{\gamma y_s} z_s dB_s - \int_t^T \int_E e^{\gamma y_{s^-}}\left(e^{\gamma u_s(x)} - 1\right)\widetilde{\mu}(dx,ds),\ \mathbb{P}-a.s.$$

Still formally, taking the conditional expectation above gives finally

$$y_t = \frac{1}{\gamma}\ln\left(\mathbb{E}_t^{\mathbb{P}}\left[e^{\gamma\xi}\right]\right),\ \mathbb{P}-a.s.,$$

and we recover the so-called entropic risk measure which in the continuous case corresponds to a BSDE with generator $\frac{\gamma}{2}|z|^2$.

Of course, for the above to make sense, the function j must at the very least be well defined. A simple application of Taylor's inequalities shows that if the function $x \mapsto u(x)$ is bounded $d\nu_t - a.e.$ for every $0 \leq t \leq T$, then we have for some constant $C > 0$

$$0 \leq e^{u(x)} - 1 - u(x) \leq Cu^2(x),\ d\nu_t - a.e. \text{ for every } 0 \leq t \leq T.$$

Hence, if we introduce for $1 < p \leq +\infty$ the spaces

$$L^p(\nu) := \{u,\ \mathcal{E}\text{-mesurable, such that } u \in L^p(\nu_t) \text{ for all } 0 \leq t \leq T\},$$

then j is well defined on $L^2(\nu) \cap L^\infty(\nu)$.

We now give our quadratic growth assumption on the generator g

Assumption 4.2.1 *[Quadratic growth]*

(i) For fixed (y, z, u), g is $\mathbb{F}$-progressively measurable.

(ii) g satisfies the following integrability condition. There exists some $C > 0$ such that

$$\underset{\tau \in \mathcal{T}_0^T}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_\tau^{\mathbb{P}}\left[\left(\int_\tau^T |g_t(0,0,0)|\,dt\right)^2\right] = \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |g_t(0,0,0)|\,dt\right)^2\right] \leq C. \tag{4.2.6}$$

(iii) g has the following growth property. There exists $(\beta, \gamma) \in \mathbb{R}_+ \times \mathbb{R}_+^*$ and a positive predictable process α satisfying the same integrability condition (4.2.6) as $g_t(0,0,0)$, such that for all (ω, t, y, z, u)

$$-\alpha_t - \beta\,|y| - \frac{\gamma}{2}\,|z|^2 - \frac{1}{\gamma} j_t(-\gamma u) \leq g_t(\omega, y, z, u) - g_t(0,0,0) \leq \alpha_t + \beta\,|y| + \frac{\gamma}{2}\,|z|^2 + \frac{1}{\gamma} j_t(\gamma u). \tag{4.2.7}$$

Remark 4.2.1 *We emphasize that unlike the usual quadratic growth assumptions for continuous BSDEs, condition (4.2.7) is not symmetric. It is mainly due to the fact that unlike the functions $|.|$ and $|.|^2$, the function j is not even. Moreover, by having this non-symmetric condition, it is easily seen that if Y is a solution to equation (4.2.2) with a generator satisfying the condition (4.2.7), then $-Y$ is also a solution to a BSDE whose generator satisfy the same condition (4.2.7). More precisely, if (Y, Z, U) solves equation (4.2.2), then $(-Y, -Z, -U)$ solves the BSDEJ with terminal condition $-\xi$ and generator $\widetilde{g}_t(y, z, u) := -g_t(-y, -z, -u)$ which clearly also satisfies (4.2.7). This will be important for the proof of Lemma 4.2.1.*

We also want to insist on the structure which appears in (4.2.7). Indeed, the constant γ in front of the quadratic term in z is the same as the one appearing in the term involving the function j. As already seen for the entropic risk-measure above, if the constants had been different, say respectively γ_1 and γ_2, the exponential transformation would have failed. Moreover, since the function $\gamma \mapsto \gamma^{-1} j_t(\gamma u)$ is not monotone, then we cannot increase or decrease γ_1 and γ_2 to recover the desired estimate (4.2.7).

4.2.5 A priori estimates

We first prove a result which shows a link between the BMO spaces and quadratic BSDEs with jumps. We emphasize that only Assumption 4.2.1 is necessary to obtain it. Before proceeding, we denote for all $t \in [0,T]$, $\mathcal{T}_t^T$ the collection of stopping times taking values in $[t,T]$. We also define for every $x \in \mathbb{R}$ and every $\eta \neq 0$

$$h_\eta(x) := \frac{e^{\eta x} - 1 - \eta x}{\eta}.$$

The function h_η already appears in our growth Assumption 4.2.1(ii), and the following trivial property that it satisfies is going to be crucial for us

$$h_{2\eta}(x) = \frac{\left(e^{\eta x} - 1\right)^2}{2\eta} + h_\eta(x). \tag{4.2.8}$$

We also give the two following inequalities which are of the utmost importance in our jump setting. We emphasize that the first one is trivial, while the second one can be proved using simple but tedious algebra.

$$2 \leq e^x + e^{-x}, \text{ for all } x \in \mathbb{R} \tag{4.2.9}$$

$$x^2 \leq a\left(e^x - 1\right)^2 + \frac{\left(1 - e^{-x}\right)^2}{a}, \text{ for all } (a,x) \in \mathbb{R}_+^* \times \mathbb{R}. \tag{4.2.10}$$

We then have

Lemma 4.2.1 *Let Assumption 4.2.1 hold. Assume that (Y,Z,U) is a solution of the BSDE (4.2.2) such that $(Z,U) \in \mathcal{Z} \times \mathcal{U}$, the jumps of Y are bounded and*

$$\operatorname*{ess\,sup}_{\tau \in \mathcal{T}_0^T}{}^{\mathbb{P}} \ \mathbb{E}_\tau^{\mathbb{P}}\left[\exp\left(2\gamma \sup_{\tau \leq t \leq T} \pm Y_t\right) \vee \exp\left(4\gamma \sup_{\tau \leq t \leq T} \pm Y_t\right)\right] < +\infty, \ \mathbb{P}-a.s. \tag{4.2.11}$$

Then $Z \in \mathbb{H}^2_{\mathrm{BMO}}$ and $U \in \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$.

Proof. First of all, since the size of the jumps of Y is bounded, there exists a version of U, that is to say that there exists a predictable function $\widetilde{U}$ such that for all $t \in [0,T]$

$$\int_E \left|\widetilde{U}_t(x) - U_t(x)\right|^2 \nu_t(dx) = 0,$$

and such that

$$\left|\widetilde{U}_t(x)\right| \leq C, \text{ for all } t.$$

For the sake of simplicity, we will always consider this version and we still denote it U. For the proof of this result, we refer to Morlais [91].

Let us consider the following processes

$$\int_0^T e^{2\gamma Y_t} Z_t dB_t \text{ and } \int_0^T e^{2\gamma Y_{t^-}} \left(e^{2\gamma U_t(x)} - 1\right) \widetilde{\mu}(dx,dt).$$

We will first show that they are local martingales. Indeed, we have

$$\int_0^T e^{4\gamma Y_t} Z_t^2 dt \leq \exp\left(4\gamma \sup_{0\leq t\leq T} Y_t\right) \int_0^T Z_t^2 dt < +\infty,\ \mathbb{P}-a.s.,$$

since $Z \in \mathcal{Z}$ and (4.2.11) holds.

Similarly, we have

$$\int_0^T \int_E e^{4\gamma Y_t} U_t^2(x)\nu_t(dx)dt \leq \exp\left(4\gamma \sup_{0\leq t\leq T} Y_t\right) \int_0^T \int_E U_t^2(x)\nu_t(dx)dt < +\infty,\ \mathbb{P}-a.s.,$$

since $U \in \mathcal{U}$ and (4.2.11) holds.

Let now $(\tau_n)_{n\geq 1}$ be a localizing sequence for the $\mathbb{P}$-local martingales above. By Itô's formula under $\mathbb{P}$ applied to $e^{2\gamma Y_t}$, we have for every $\tau \in \mathcal{T}_0^T$

$$\begin{aligned}
&\frac{4\gamma^2}{2}\int_\tau^{\tau_n} e^{2\gamma Y_t}\left|Z_t\right|^2 dt + 2\gamma\int_\tau^{\tau_n}\int_E e^{2\gamma Y_t} h_{2\gamma}\left(U_t(x)\right)\nu_t(dx)dt \\
&= e^{2\gamma Y_{\tau_n}} - e^{-2\gamma Y_\tau} + 2\gamma\int_\tau^{\tau_n} e^{2\gamma Y_t} g_t(Y_t,Z_t,U_t)dt - 2\gamma\int_\tau^{\tau_n} e^{2\gamma Y_t} Z_t dB_t \\
&\quad - \int_\tau^{\tau_n}\int_E e^{2\gamma Y_{t^-}}\left(e^{2\gamma U_t(x)}-1\right)\widetilde{\mu}(dx,dt) \\
&\leq e^{2\gamma Y_{\tau_n}} - e^{-2\gamma Y_\tau} + 2\gamma\int_\tau^{\tau_n} e^{2\gamma Y_t}\left(\alpha_t + |g_t(0,0,0)| + \beta\left|Y_t\right| + \frac{\gamma}{2}\left|Z_t\right|^2 + \int_E h_\gamma\left(U_t(x)\right)\nu_t(dx)\right)dt \\
&\quad - 2\gamma\int_\tau^{\tau_n} e^{2\gamma Y_t} Z_t dB_t - \int_\tau^{\tau_n}\int_E e^{2\gamma Y_{t^-}}\left(e^{2\gamma U_t(x)}-1\right)\widetilde{\mu}(dx,dt).
\end{aligned}$$

Now the situation is going to be different from the continuous case, and the property (4.2.8) is going to be important. Indeed, we can take conditional expectation and thus obtain

$$\begin{aligned}
&\mathbb{E}_\tau^{\mathbb{P}}\left[\gamma^2\int_\tau^{\tau_n} e^{2\gamma Y_t}\left|Z_t\right|^2 dt + \int_\tau^{\tau_n}\int_E e^{2\gamma Y_t}\left(e^{\gamma U_t(x)}-1\right)^2\nu_t(dx)dt\right] \\
&\leq \mathbb{E}_\tau^{\mathbb{P}}\left[\left(\int_\tau^{\tau_n}\left(\alpha_t + |g_t(0,0,0)|\right)dt\right)^2 + \exp\left(2\gamma\sup_{\tau\leq t\leq T} Y_t\right) + \left(C_0\beta+\gamma^2\right)\exp\left(4\gamma\sup_{\tau\leq t\leq T} Y_t\right)\right] \\
&\leq C\left(1 + \mathbb{E}_\tau^{\mathbb{P}}\left[\exp\left(2\gamma\sup_{\tau\leq t\leq T} Y_t\right)\vee\exp\left(4\gamma\sup_{\tau\leq t\leq T} Y_t\right)\right]\right),
\end{aligned}$$

where we used the inequality $2ab \leq a^2+b^2$, the fact that for all $x\in\mathbb{R}$, $|x|\,e^x \leq C_0 e^{2x}$ for some constant $C_0>0$ and the fact that Assumption 4.2.1(ii) and (iii) hold.

Using Fatou's lemma and the monotone convergence Theorem, we obtain

$$\begin{aligned}
&\mathbb{E}_\tau^{\mathbb{P}}\left[\gamma^2\int_\tau^{T} e^{2\gamma Y_t}\left|Z_t\right|^2 dt + \int_\tau^{T}\int_E e^{2\gamma Y_t}\left(e^{\gamma U_t(x)}-1\right)^2\nu_t(dx)dt\right] \\
&\leq C\left(1 + \underset{\tau\in\mathcal{T}_0^T}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_\tau^{\mathbb{P}}\left[\exp\left(2\gamma\sup_{\tau\leq t\leq T} Y_t\right)\vee\exp\left(4\gamma\sup_{\tau\leq t\leq T} Y_t\right)\right]\right). \qquad (4.2.12)
\end{aligned}$$

Now, we apply the above estimate for the solution $(-Y,-Z,-U)$ of the BSDEJ with terminal condition $-\xi$ and generator $\widetilde{g}_t(y,z,u) := -g_t(-y,-z,-u)$, which still satisfies Assumption 4.2.1 (see Remark 4.2.1)

$$\begin{aligned}&\mathbb{E}_\tau^{\mathbb{P}}\left[\gamma^2\int_\tau^T e^{-2\gamma Y_t}\left|Z_t\right|^2 dt+\int_\tau^T\int_E e^{-2\gamma Y_t}\left(e^{-\gamma U_t(x)}-1\right)^2\nu_t(dx)dt\right]\\&\leq C\left(1+\underset{\tau\in\mathcal{T}_0^T}{\operatorname{ess\,sup}}^{\mathbb{P}}\ \mathbb{E}_\tau^{\mathbb{P}}\left[\exp\left(2\gamma\sup_{\tau\leq t\leq T}(-Y_t)\right)\vee\exp\left(4\gamma\sup_{\tau\leq t\leq T}(-Y_t)\right)\right]\right).\end{aligned}\tag{4.2.13}$$

Let us now sum the inequalities (4.2.12) and (4.2.13). We obtain

$$\begin{aligned}&\mathbb{E}_\tau^{\mathbb{P}}\left[\int_\tau^T\gamma^2\left(e^{2\gamma Y_t}+e^{-2\gamma Y_t}\right)|Z_t|^2+\int_E e^{2\gamma Y_t}\left(e^{\gamma U_t(x)}-1\right)^2+e^{-2\gamma Y_t}\left(e^{-\gamma U_t(x)}-1\right)^2\nu_t(dx)dt\right]\\&\leq C\left(1+\underset{\tau\in\mathcal{T}_0^T}{\operatorname{ess\,sup}}^{\mathbb{P}}\ \mathbb{E}_\tau^{\mathbb{P}}\left[\sup_{\tau\leq t\leq T}\left\{\exp\left(2\gamma Y_t\right)\vee\exp\left(4\gamma Y_t\right)+\exp\left(2\gamma\left(-Y_t\right)\right)\vee\exp\left(4\gamma\left(-Y_t\right)\right)\right\}\right]\right).\end{aligned}$$

Finally, from the inequalities (4.2.9) and (4.2.10), this shows the desired result. □

Remark 4.2.2 *In the above Proposition, if we only assume that*

$$\mathbb{E}^{\mathbb{P}}\left[\exp\left(2\gamma\sup_{0\leq t\leq T}\pm Y_t\right)\vee\exp\left(4\gamma\sup_{0\leq t\leq T}\pm Y_t\right)\right]<+\infty,$$

then the exact same proof would show that $(Z,U)\in\mathbb{H}^2\times\mathbb{J}^2$. Moreover, using the Neveu-Garsia Lemma in the same spirit as [7], we could also show that $(Z,U)\in\mathbb{H}^p\times\mathbb{J}^p$ for all $p>1$.

We emphasize that the results of this Proposition highlight the fact that we do not necessarily need to consider solutions with a bounded Y in the quadratic case to obtain *a priori* estimates. It is enough to assume the existence of some exponential moments. This generalizes to the jump case some of the ideas developed in [18] and [7]. Nonetheless, our proof of existence will rely heavily on BMO properties of the solution, and the simplest condition to obtain the estimate (4.2.11) is to assume that Y is indeed bounded. The aim of the following Proposition is to show that we can control the $\mathcal{S}^\infty$ norm of Y by the L^∞ norm of ξ. Since the proof is very similar to the proof Lemma 1 in [18], we will omit it.

Proposition 4.2.6 *Let $\xi\in\mathbb{L}^\infty$. Let Assumption 4.2.1 hold and assume that $g(0,0,0)$ and α are bounded by some constant M. Let $(Y,Z,U)\in\mathcal{S}^\infty\times\mathbb{H}^2\times\mathbb{J}^2$ be a solution of the BSDE (4.2.2). Then we have*

$$|Y_t|\leq\gamma M\frac{e^{\beta(T-t)}-1}{\beta}+\gamma e^{\beta(T-t)}\left\|\xi\right\|_{\mathbb{L}^\infty}.$$

4.2.6 Existence and uniqueness for a small terminal condition

The aim of this Section is to obtain an existence and uniqueness result for BSDEJ with quadratic growth when the terminal condition is small enough. However, we will need more assumptions for our proof to work. First, we assume from now on that we have the following martingale representation property. We need this assumption since we will rely on the existence results in [4] or [81] which need the martingale representation.

Assumption 4.2.2 *Any local martingale M has the predictable representation property, that is to say, that there exists a unique predictable process H and a unique predictable function U such that $(H,U) \in \mathcal{Z} \times \mathcal{U}$ and*

$$M_t = M_0 + \int_0^t H_s dB_s + \int_0^t \int_E U_s(x)\widetilde{\mu}(dx,ds), \ \mathbb{P}-a.s.$$

Remark 4.2.3 *This martingale representation property holds for instance when the compensator ν does not depend on ω, i.e when ν is the compensator of the counting measure of an additive process in the sense of Sato [107]. It also holds when ν has the particular form described in Chapter 5, in which case ν depends on ω.*

Of course, we also need to assume more properties for our generator g.

Assumption 4.2.3 *[Lipschitz assumption]*

Let Assumption 4.2.1(i),(ii) hold and assume furthermore that

(i) g is uniformly Lipschitz in y.

$$\left|g_t(\omega,y,z,u) - g_t(\omega,y',z,u)\right| \leq C\left|y-y'\right| \text{ for all } (\omega,t,y,y',z,u).$$

(ii) $\exists\ \mu > 0$ and $\phi \in \mathbb{H}^2_{\mathrm{BMO}}$ such that for all (t,y,z,z',u)

$$\left|g_t(\omega,y,z,u) - g_t(\omega,y,z',u) - \phi_t.(z-z')\right| \leq \mu\left|z-z'\right|\left(|z|+\left|z'\right|\right).$$

(iii) $\exists\ \mu > 0$ and $\psi \in \mathbb{J}^2_{\mathrm{BMO}}$ such that for all (t,x)

$$C_1(1\wedge|x|) \leq \psi_t(x) \leq C_2(1\wedge|x|),$$

where $C_2 > 0$, $C_1 \geq -1+\delta$ where $\delta > 0$. Moreover, for all (ω,t,y,z,u,u')

$$\left|g_t(\omega,y,z,u) - g_t(\omega,y,z,u') - < \psi_t, u-u' >_t\right| \leq \mu\left\|u-u'\right\|_{L^2(\nu_t)}\left(\|u\|_{L^2(\nu_t)} + \left\|u'\right\|_{L^2(\nu_t)}\right),$$

where $< u_1, u_2 >_t := \int_E u_1(x)u_2(x)\nu_t(dx)$ is the scalar product in $L^2(\nu_t)$.

Remark 4.2.4 *Let us comment on the above assumptions. The first one concerning Lipschitz continuity in the variable y is classical in the BSDE theory. The two others may seem a bit complicated, but they are almost equivalent to saying that the function g is locally Lipschitz in z and u. In the case of the variable z for instance, those two properties would be equivalent if the process ϕ were bounded. Here we allow something a bit more general by letting ϕ be unbounded but in $\mathbb{H}^2_{\mathrm{BMO}}$. Once again, since these assumptions allow us to apply the Girsanov property of Proposition 4.2.3, we do not need to bound the processes and BMO type conditions are sufficient. Moreover, Assumption 4.2.3 also implies a weaker version of Assumption 4.2.1. Indeed, it implies clearly that*

$$|g_t(y,z,u) - g_t(0,0,0) - \phi_t.z - < \psi_t, u >_t| \leq C|y| + \mu\left(|z|^2 + \|u\|^2_{L^2(\nu_t)}\right).$$

Then, for any $u \in L^2(\nu) \cap L^\infty(\nu)$ and for any $\gamma > 0$, we have using the mean value Theorem

$$\frac{\gamma}{2} e^{-\gamma \|u\|_{L^\infty(\nu)}} \|u\|^2_{L^2(\nu_t)} \leq \frac{1}{\gamma} j_t(\pm \gamma u) \leq \frac{\gamma}{2} e^{\gamma \|u\|_{L^\infty(\nu)}} \|u\|^2_{L^2(\nu_t)}.$$

Therefore we deduce using the Cauchy-Schwartz inequality and the trivial inequality $2ab \leq a^2 + b^2$

$$g_t(y,z,u) - g_t(0,0,0) \leq \frac{|\phi_t|^2}{2} + \frac{\|\psi_t\|^2_{L^2(\nu_t)}}{2} + C|y| + \left(\mu + \frac{1}{2}\right)\left(|z|^2 + \frac{2e^{\gamma \|u\|_{L^\infty(\nu)}}}{\gamma^2} j_t(\gamma u)\right)$$
$$g_t(y,z,u) - g_t(0,0,0) \geq -\frac{|\phi_t|^2}{2} - \frac{\|\psi_t\|^2_{L^2(\nu_t)}}{2} - C|y| - \left(\mu + \frac{1}{2}\right)\left(|z|^2 - \frac{2e^{\gamma \|u\|_{L^\infty(\nu)}}}{\gamma^2} j_t(-\gamma u)\right).$$

Then, since Assumption 4.2.1(ii) holds and since $(\phi, \psi) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}}$, we deduce from the energy inequalities (4.2.3)

$$\operatorname*{ess\,sup}_{\tau \in \mathcal{T}_0^T}{}^{\mathbb{P}} \ \mathbb{E}^{\mathbb{P}}_\tau \left[\left(\int_\tau^T \left(\frac{|\phi_t|^2}{2} + \frac{\|\psi_t\|^2_{L^2(\nu_t)}}{2}\right) dt\right)^2\right] \leq \frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |\phi_t|^2 dt\right)^2 + \left(\int_0^T \|\psi_t\|^2_{L^2(\nu_t)} dt\right)^2\right]$$
$$< +\infty.$$

Hence, we have obtained a growth property which is similar to (4.2.7)*, the only difference being that the constants appearing in the quadratic term in z and the term involving the function j cannot, a priori, be the same. This prevents us from recovering the structure already mentioned in Remark 4.2.1.*

We now show that if we can solve the BSDEJ (4.2.2) for a generator g satisfying Assumption 4.2.3 with $\phi = 0$ and $\psi = 0$, we can immediately obtain the existence for general ϕ and ψ. This will simplify our subsequent proof of existence. Notice that the result relies essentially on the Girsanov Theorem of Proposition 4.2.3.

Lemma 4.2.2 *Define*

$$\overline{g}_t(\omega, y, z, u) := g_t(\omega, y, z, u) - < \phi_t(\omega), z >_{\mathbb{R}^d} - < \psi_t(\omega), u >_{L^2(\nu_t)}.$$

Then (Y, Z, U) is a solution of the BSDEJ with generator g and terminal condition ξ under $\mathbb{P}$ if and only if (Y, Z, U) is a solution of the BSDEJ with generator $\overline{g}$ and terminal condition ξ under $\mathbb{Q}$ where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(\int_0^T \phi_s dB_s + \int_0^T \int_E \psi_s(x) \widetilde{\mu}(dx, ds)\right).$$

Proof. We have clearly

$$Y_t = \xi + \int_t^T g_s(Y_s, Z_s, U_s) ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x) \widetilde{\mu}(dx, ds)$$
$$\Leftrightarrow Y_t = \xi + \int_t^T \overline{g}_s(Y_s, Z_s, U_s) ds - \int_t^T Z_s(dB_s - \phi_s ds)$$
$$- \int_t^T \int_E U_s(x)(\widetilde{\mu}(dx, ds) - \psi_s(x)\nu_s(dx) ds).$$

Now, by our BMO assumptions on ϕ and ψ and the fact that we assumed that $\psi \geq -1+\delta$, we can apply Proposition 4.2.3 and $\mathbb{Q}$ is well defined. Then by Girsanov Theorem, we know that $dB_s - \phi_s ds$ and $\widetilde{\mu}(dx, ds) - \psi_s(x)\nu(dx)ds$ are martingales under $\mathbb{Q}$. Hence the desired result. $\square$

Remark 4.2.5 *It is clear that if g satisfies Assumption 4.2.3, then $\overline{g}$ defined above satisfies Assumption 4.2.3 with $\phi = \psi = 0$.*

Following Lemma 4.2.2 we assume for the time being that $g(0,0,0) = \phi = \psi = 0$. Our first result is the following

Theorem 4.2.1 *Assume that*

$$\|\xi\|_\infty \leq \frac{1}{2\sqrt{15}\sqrt{2670}\mu e^{\frac{3}{2}CT}},$$

where C is the Lipschitz constant of g in y, and μ is the constant appearing in Assumption 4.2.3. Then under Assumption 4.2.3 with $\phi = 0$, $\psi = 0$ and $g(0,0,0) = 0$, there exists a unique solution $(Y, Z, U) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$ of the BSDEJ (4.2.2).

Remark 4.2.6 *Notice that in the above Theorem, we do not need Assumption 4.2.1(iii) to hold. This is linked to the fact that, as discussed in Remark 4.2.4, Assumption 4.2.3 implies a weak version of Assumption 4.2.1(iii), which is sufficient for our purpose here.*

Proof. We first recall that we have with Assumption 4.2.3 when $g(0,0,0) = \phi = \psi = 0$

$$|g_t(y,z,u)| \leq C\,|y| + \mu\,|z|^2 + \mu\,\|u\|^2_{L^2(\nu_t)}. \tag{4.2.14}$$

Consider now the map $\Phi : (y,z,u) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu) \to (Y,Z,U)$ defined by

$$Y_t = \xi + \int_t^T g_s(Y_s, z_s, u_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(dx, ds). \tag{4.2.15}$$

The above is nothing more than a BSDE with jumps whose generator depends only on Y and is Lipschitz. Besides, since $(z,u) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$, using (4.2.14) and the energy inequalities (4.2.3) we clearly have

$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |g_s(0, z_s, u_s)|\,ds\right)^2\right] < +\infty.$$

Hence, the existence of $(Y,Z,U) \in \mathcal{S}^2 \times \mathbb{H}^2 \times \mathbb{J}^2$ is ensured by the results of Barles, Buckdahn and Pardoux [4] or Li and Tang [81] for Lipschitz BSDEs with jumps. Of course, we could have let the generator in (4.2.15) depend on (y_s, z_s, u_s) instead. The existence of (Y,Z,U) would then have been a consequence of the predictable martingale representation Theorem. However, the form that we have chosen will simplify some of the following estimates.

Step 1: We first show that $(Y,Z,U) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$.

Recall that by the Lipschitz hypothesis in y, there exists a bounded process λ such that

$$g_s(Y_s, z_s, u_s) = \lambda_s Y_s + g_s(0, z_s, u_s).$$

Let us now apply Itô's formula to $e^{\int_t^s \lambda_u du}|Y_s|$. We obtain easily from Assumption 4.2.3

$$\begin{aligned}
Y_t &= \mathbb{E}_t^{\mathbb{P}}\left[e^{\int_t^T \lambda_s ds}\xi + \int_t^T e^{\int_t^s \lambda_u du}(\lambda_s Y_s + g_s(0,z_s,u_s))ds - \int_t^T \lambda_s e^{\int_t^s \lambda_u du} Y_s ds\right] \\
&\leq \mathbb{E}_t^{\mathbb{P}}\left[e^{\int_t^T \lambda_s ds}\xi + \mu\int_t^T e^{\int_t^s \lambda_u du}\left(|z_s|^2 + \int_E u_s^2(x)\right)ds\right] \\
&\leq \|\xi\|_\infty + C\left(\|z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|u\|^2_{\mathbb{J}^2_{\mathrm{BMO}}}\right).
\end{aligned}$$

Therefore Y is bounded and consequently, since its jumps are also bounded, we know that there is a version of U such that

$$\|U\|_{L^\infty(\nu)} \leq 2\|Y\|_{\mathcal{S}^\infty}.$$

Let us now prove that $(Z,U) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}}$. Applying Itô's formula to $e^{\eta t}|Y_t|^2$ for some $\eta > 0$, we obtain for any stopping time $\tau \in \mathcal{T}_0^T$

$$\begin{aligned}
&e^{\eta\tau}|Y_\tau|^2 + \mathbb{E}_\tau^{\mathbb{P}}\left[\int_\tau^T e^{\eta s}|Z_s|^2 ds + \int_\tau^T\int_E e^{\eta s}U_s^2(x)\nu_s(dx)ds\right] \\
&= \mathbb{E}_\tau^{\mathbb{P}}\left[e^{\nu T}\xi^2 + 2\int_\tau^T e^{\eta s}Y_s g_s(Y_s,z_s,u_s)ds - \eta\int_\tau^T e^{\eta s}|Y_s|^2 ds\right] \\
&\leq \mathbb{E}_\tau^{\mathbb{P}}\left[e^{\eta T}\xi^2 + (2C-\eta)\int_\tau^T e^{\eta s}|Y_s|^2 ds + 2\|Y\|_{\mathcal{S}^\infty}\int_\tau^T e^{\eta s}|g_s(0,z_s,u_s)|\,ds\right].
\end{aligned}$$

Choosing $\eta = 2C$, and using the elementary inequality $2ab \leq \frac{a^2}{\varepsilon} + \varepsilon b^2$, we obtain

$$\begin{aligned}
&|Y_\tau|^2 + \mathbb{E}_\tau^{\mathbb{P}}\left[\int_\tau^T |Z_s|^2 ds + \int_\tau^T\int_E U_s^2(x)\nu_s(dx)ds\right] \\
&\leq \mathbb{E}_\tau^{\mathbb{P}}\left[e^{\eta T}\xi^2 + \varepsilon\|Y\|^2_{\mathcal{S}^\infty} + \frac{e^{2\eta T}}{\varepsilon}\left(\int_\tau^T |g_s(0,z_s,u_s)|\,ds\right)^2\right].
\end{aligned}$$

Hence,

$$(1-\varepsilon)\|Y\|^2_{\mathcal{S}^\infty} + \|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} \leq e^{\eta T}\|\xi\|^2_\infty + 64\mu^2\frac{e^{2\eta T}}{\varepsilon}\left(\|z\|^4_{\mathbb{H}^2_{\mathrm{BMO}}} + \|u\|^4_{\mathbb{J}^2_{\mathrm{BMO}}}\right).$$

And finally, choosing $\varepsilon = 1/2$

$$\|Y\|^2_{\mathcal{S}^\infty} + \|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} \leq 2e^{\eta T}\|\xi\|^2_\infty + 256\mu^2 e^{2\eta T}\left(\|z\|^4_{\mathbb{H}^2_{\mathrm{BMO}}} + \|u\|^4_{\mathbb{J}^2_{\mathrm{BMO}}}\right).$$

Our problem now is that the norms for Z and U in the left-hand side above are to the power 2, while they are to the power 4 on the right-hand side. Therefore, it will clearly be impossible for us to prevent an explosion if we do not first start by restricting ourselves in some ball with a well chosen radius. This is exactly what we are going to do. Define therefore $R = \frac{1}{2\sqrt{2670}\mu e^{\nu T}}$, and assume that $\|\xi\|_\infty \leq \frac{R}{\sqrt{15}e^{\frac{1}{2}\eta T}}$ and that

$$\|y\|^2_{\mathcal{S}^\infty} + \|z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|u\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} + \|u\|^2_{L^\infty(\nu)} \leq R^2.$$

Denote $\Lambda := \|Y\|^2_{\mathcal{S}^\infty} + \|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} + \|U\|^2_{L^\infty(\nu)}$. We have, since $\|U\|^2_{L^\infty(\nu)} \leq 4\|Y\|^2_{\mathcal{S}^\infty}$

$$\begin{aligned}\Lambda \leq 5\|Y\|^2_{\mathcal{S}^\infty} + \|Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} &\leq 10e^{\nu T}\|\xi\|^2_\infty + 1280\mu^2 e^{2\nu T}\left(\|z\|^4_{\mathbb{H}^2_{\mathrm{BMO}}} + \|u\|^4_{\mathbb{J}^2_{\mathrm{BMO}}}\right)\\ &\leq \frac{2R^2}{3} + 3560\mu^2 e^{2\nu T}R^4\\ &= \frac{2R^2}{3} + \frac{R^2}{3} = R^2.\end{aligned}$$

Hence if $\mathcal{B}_R$ is the ball of radius R in $\mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$, we have shown that $\Phi(\mathcal{B}_R) \subset \mathcal{B}_R$.

Step 2: We show that Φ is a contraction in this ball of radius R.

For $i = 1, 2$ and $(y^i, z^i, u^i) \in \mathcal{B}_R$, we denote $(Y^i, Z^i, U^i) := \Phi(y^i, z^i, u^i)$ and

$$\begin{aligned}&\delta y := y^1 - y^2,\ \delta z := z^1 - z^2,\ \delta u := u^1 - u^2,\ \delta Y := Y^1 - Y^2\\ &\delta Z := Z^1 - Z^2,\ \delta U := U^1 - U^2,\ \delta g := g(0, z^1, u^1) - g(0, z^2, u^2).\end{aligned}$$

Arguing as above, we obtain easily

$$\|\delta Y\|^2_{\mathcal{S}^\infty} + \|\delta Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|\delta U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} \leq 4e^{2\eta T}\sup_{\tau \in \mathcal{T}_0^T}\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T |\delta g_s|\,ds\right]\right)^2.$$

We next estimate that

$$\begin{aligned}\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T |\delta g_s|\,ds\right]\right)^2 &\leq 2\mu^2\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T |\delta z_s|\left(|z^1_s| + |z^2_s|\right)ds\right]\right)^2\\ &\quad + 2\mu^2\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T \|\delta u_s\|_{L^2(\nu_s)}\left(\|u^1_s\|_{L^2(\nu_s)} + \|u^2_s\|_{L^2(\nu_s)}\right)ds\right]\right)^2\\ &\leq 2\mu^2\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T |\delta z_s|^2\,ds\right]\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T (|z^1_s| + |z^2_s|)^2 ds\right]\right.\\ &\quad \left.+\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T \|\delta u_s\|^2_{L^2(\nu_s)}\,ds\right]\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T \left(\|u^1_s\|_{L^2(\nu_s)} + \|u^2_s\|_{L^2(\nu_s)}\right)^2 ds\right]\right)\\ &\leq 4R^2\mu^2\left(\mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T |\delta z_s|^2\,ds\right] + \mathbb{E}^{\mathbb{P}}_\tau\left[\int_\tau^T\int_E \delta u^2_s(x)\nu(dx)ds\right]\right)\\ &\leq 32R^2\mu^2\left(\|\delta z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|\delta u\|^2_{\mathbb{J}^2_{\mathrm{BMO}}}\right)\end{aligned}$$

From these estimates, we get

$$\begin{aligned}\|\delta Y\|^2_{\mathcal{S}^\infty} + \|\delta Z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|\delta U\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} + \|\delta U\|^2_{L^\infty(\nu)} &\leq 20 \times 32R^2\mu^2 e^{2\nu T}\left(\|\delta z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|\delta u\|^2_{\mathbb{J}^2_{\mathrm{BMO}}}\right)\\ &= \frac{16}{267}\left(\|\delta z\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \|\delta u\|^2_{\mathbb{J}^2_{\mathrm{BMO}}}\right).\end{aligned}$$

Therefore Φ is a contraction which has a unique fixed point. $\square$

Then, from Lemma 4.2.2, we have immediately the following Corollary

Corollary 4.2.1 *Assume that*

$$\|\xi\|_\infty \leq \frac{1}{2\sqrt{15}\sqrt{2670}\mu e^{\frac{3}{2}CT}},$$

where C is the Lipschitz constant of g in y, and μ is the constant appearing in Asssumption 4.2.3. Then under Assumption 4.2.3 with $g(0,0,0)=0$, there exists a unique solution $(Y,Z,U)\in\mathcal{S}^\infty\times\mathbb{H}^2_{\mathrm{BMO}}\times\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$ of the BSDEJ (4.2.2).

We now show how we can get rid off the assumption that $g_t(0,0,0)=0$.

Corollary 4.2.2 *Assume that*

$$\|\xi\|_\infty + D\left\|\int_0^T |g_t(0,0,0)|\,dt\right\|_\infty \leq \frac{1}{2\sqrt{15}\sqrt{2670}\mu e^{\frac{3}{2}CT}},$$

where C is the Lipschitz constant of g in y, μ is the constant appearing in Asssumption 4.2.3 and D is a large enough positive constant. Then under Assumption 4.2.3, there exists a solution $(Y,Z,U)\in\mathcal{S}^\infty\times\mathbb{H}^2_{\mathrm{BMO}}\times\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$ of the BSDEJ (4.2.2).

Proof. By Corollary 4.2.1, we can show the existence of a solution to the BSDEJ with generator $\widetilde{g}_t(y,z,u) := g_t(y-\int_0^t g_s(0,0,0)ds, z, u) - g_t(0,0,0)$ and terminal condition $\overline{\xi} := \xi + \int_0^T g_t(0,0,0)dt$. Indeed, even though $\overline{g}$ is not null at $(0,0,0)$, it is not difficult to show with the same proof as in Theorem 4.2.1 that a solution $(\overline{Y},\overline{Z},\overline{U})$ exists (the same type of arguments are used in [117]). More precisely, $\widetilde{g}$ still satisfies Assumption 4.2.3(i) and when ϕ and ψ in Assumption 4.2.3 are equal to 0, we have the estimate

$$|\widetilde{g}_t(y,z,u)| \leq C\left\|\int_0^T |g_s(0,0,0)|\,ds\right\|_\infty + C\,|y| + \mu\,|z|^2 + \mu\,\|u\|^2_{L^2(\nu_t)},$$

which is the counterpart of (4.2.14). Thus, since the constant term in the above estimate is assumed to be small enough, it will play the same role as $\|\xi\|_\infty$ in the first Step of the proof of Theorem 4.2.1.

For the Step 2, everything still work thanks to the following estimate

$$\begin{aligned}\left|\widetilde{g}_t(0,z^1,u^1)-\widetilde{g}_t(0,z^2,u^2)\right| &\leq \mu\left|z^1-z^2\right|\left(\left|z^1\right|+\left|z^2\right|\right)\\ &\quad+\mu\left\|u^1-u^2\right\|_{L^2(\nu_t)}\left(\left\|u^1\right\|_{L^2(\nu_t)}+\left\|u^2\right\|_{L^2(\nu_t)}\right).\end{aligned}$$

Then, if we define

$$(Y_t,Z_t,U_t) := \left(\overline{Y}_t - \int_0^t g_s(0,0,0)ds, \overline{Z}_t, \overline{U}_t\right),$$

it is clear that it is a solution to the BSDEJ with generator g and terminal condition ξ. □

Remark 4.2.7 *We emphasize that the above proof of existence extends readily to a terminal condition which is in $\mathbb{R}^n$ for any $n\geq 2$.*

4.2.7 Existence for a bounded terminal condition

We now show that we can still prove existence of a solution for any bounded terminal condition. Nonetheless, we will need to strengthen once more our assumptions on the generator, mainly by assuming more regularity.

Assumption 4.2.4 (i) *g is uniformly Lipschitz in y.*

$$\left|g_t(\omega,y,z,u)-g_t(\omega,y',z,u)\right|\leq C\left|y-y'\right| \text{ for all } (\omega,t,y,y',z,u).$$

(ii) *g is C^2 in z and there is a constant $\theta>0$ and a process $(r_t)_{0\leq t\leq T}\in\mathbb{H}^2_{\mathrm{BMO}}$, such that for all (t,ω,y,z,u),*

$$|D_z g_t(\omega,y,z,u)|\leq r_t+\theta\,|z|\,,\ |D^2_{zz}g_t(\omega,y,z,u)|\leq\theta.$$

(iii) *g is twice Fréchet differentiable in the Banach space $L^2(\nu)$ and there are constants θ, $\delta>0$, $C_1\geq -1+\delta$, $C_2\geq 0$ and a predictable function $m\in\mathbb{J}^2_{\mathrm{BMO}}$ such that for all (t,ω,y,z,u,x),*

$$|D_u g_t(\omega,y,z,u)|\,(x)\leq m_t(x)+\theta\,|u(x)|\,,\ C_1(1\wedge|x|)\leq(D_u g_t(\omega,y,z,u))(x)\leq C_2(1\wedge|x|)$$

$$\left\|D^2_u g_t(\omega,y,z,u)\right\|_{L^2(\nu_t)}\leq\theta.$$

Remark 4.2.8 *The assumptions (ii) and (iii) above are generalizations to the jump case of the assumptions considered by Tevzadze [117]. They will only be useful in our proof of existence and is tailor-made to allow us to apply the Girsanov transformation of Proposition 4.2.3. Notice also that since the space $L^2(\nu)$ is clearly a Banach space, there is no problem to define the Fréchet derivative.*

We emphasize here that Assumption 4.2.4 is stronger than Assumption 4.2.3. Indeed, we have the following result

Lemma 4.2.3 *If Assumption 4.2.4(ii) and (iii) hold, then so do Assumption 4.2.3(ii) and (iii).*

Proof. We will only show that if Assumption 4.2.4(iii) holds, so does Assumption 4.2.3(iii), the proof being similar for Assumption 4.2.4(ii). Since g is twice Fréchet differentiable in u, we introduce the process $\psi_t:=D_u g_t(y,z,0)$ which is bounded from above by m and from below by $C_1\geq -1+\delta$ by assumption. Thus, $\psi\in\mathbb{J}^2_{\mathrm{BMO}}$. By the mean value theorem, we compute that for some $\lambda\in[0,1]$ and with $u_\lambda:=\lambda u+(1-\lambda)u'$

$$\begin{aligned}\left|g_t(y,z,u)-g_t(y,z,u')-<\psi_t,u-u'>_t\right|&\leq\|D_u g_t(y,z,u_\lambda)-\psi_t\|\left\|u-u'\right\|_{L^2(\nu_t)}\\&\leq\theta\left\|\lambda u+(1-\lambda)u'\right\|_{L^2(\nu_t)}\left\|u-u'\right\|_{L^2(\nu_t)},\end{aligned}$$

by the bound on $D^2_u g$. The result now follows easily. □

We can now state our main existence result.

Theorem 4.2.2 *Let $\xi\in\mathbb{L}^\infty$. Under Assumptions 4.2.1 and 4.2.4, there exists a solution $(Y,Z,U)\in\mathcal{S}^\infty\times\mathbb{H}^2_{\mathrm{BMO}}\times\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$ of the BSDEJ (4.2.2).*

The idea of the proof is to find a "good" splitting of the BSDE into the sum of BSDEs for which the terminal condition is small and existence holds. Then we paste everything together. This is during this pasting step that the regularity of the generator in z and u in Assumption 4.2.4 is going to be important.

Proof.

(i) We first assume that $g_t(0,0,0)=0$.

Consider an arbitrary decomposition of ξ

$$\xi=\sum_{i=1}^n \xi_i \text{ such that } \|\xi_i\|_\infty \leq \frac{1}{2\sqrt{15}\sqrt{2670}\mu e^{\frac{3}{2}CT}},\text{ for all } i.$$

We will now construct a solution to (4.2.2) recursively.

Step 1 We define $g^1:=g$ and (Y^1,Z^1,U^1) as the unique solution of

$$Y_t^1=\xi_1+\int_t^T g_s^1(Y_s^1,Z_s^1,U_s^1)ds-\int_t^T Z_s^1dB_s-\int_t^T\int_E U_s^1(x)\widetilde{\mu}(ds,dx),\ \mathbb{P}-a.s. \tag{4.2.16}$$

Let us show why this solution exists. Since g^1 satisfies Assumption 4.2.4, we know by Lemma 4.2.3 that it satisfies Assumption 4.2.3 with $\phi_t:=D_zg_t(y,0,u)$ and $\psi_t:=D_ug_t(y,z,0)$, these processes being respectively in $\mathbb{H}^2_{\mathrm{BMO}}$ and $\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$ by assumption. Furthermore, we have $\psi_t(x)\geq C_1(1\wedge|x|)$ with $C_1\geq -1+\delta$. Thanks to Theorem 4.2.1 and with the notations of Lemma 4.2.2, we can then define the solution to the BSDEJ with driver $\overline{g}^1$ and terminal condition ξ_1 under the probability measure $\mathbb{Q}^1$ defined by

$$\frac{d\mathbb{Q}^1}{d\mathbb{P}}=\mathcal{E}\left(\int_0^T\phi_sdB_s+\int_0^T\int_E\psi_s(x)\widetilde{\mu}(dx,ds)\right).$$

Thanks to Lemma 4.2.2, this gives us a solution (Y^1,Z^1,U^1) to (4.2.16) with Y^1 bounded, which in turn implies with Lemma 4.2.1 that $(Z^1,U^1)\in\mathbb{H}^2_{\mathrm{BMO}}\times\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$.

Step 2 We assume that we have constructed similarly $(Y^j,Z^j,U^j)\in\mathcal{S}^\infty\times\mathbb{H}^2_{\mathrm{BMO}}\times\mathbb{J}^2_{\mathrm{BMO}}\cap L^\infty(\nu)$ for $j\leq i-1$. We then define the generator

$$g_t^i(y,z,u):=g_t\left(\overline{Y}_t^{i-1}+y,\overline{Z}_t^{i-1}+z,\overline{U}_t^{i-1}+u\right)-g_t\left(\overline{Y}_t^{i-1},\overline{Z}_t^{i-1},\overline{U}_t^{i-1}\right),$$

where

$$\overline{Y}_t^{i-1}:=\sum_{j=1}^{i-1}Y_t^j,\ \overline{Z}_t^{i-1}:=\sum_{j=1}^{i-1}Z_t^j,\ \overline{U}_t^{i-1}:=\sum_{j=1}^{i-1}U_t^j.$$

Notice that since g satisfies Assumption 4.2.1(iii), we have the estimate

$$\begin{aligned}
g_t^i(y,z,u)&\leq 2\alpha_t+\beta\left|y+\overline{Y}_t^{i-1}\right|+\beta\left|\overline{Y}_t^{i-1}\right|+\gamma\left|z+\overline{Z}_t^{i-1}\right|^2+\gamma\left|\overline{Z}_t^{i-1}\right|^2\\
&\quad+\frac{1}{\gamma}j_t\left(\gamma\left(u+\overline{U}_t^{i-1}\right)\right)+\frac{1}{\gamma}j_t\left(\gamma\overline{U}_t^{i-1}\right)\\
&\leq 2\alpha_t+2\beta\left|\overline{Y}_t^{i-1}\right|+3\gamma\left|\overline{Z}_t^{i-1}\right|^2+\frac{1}{\gamma}j_t\left(\gamma\overline{U}_t^{i-1}\right)+\frac{1}{2\gamma}j_t\left(2\gamma\overline{U}_t^{i-1}\right)\\
&\quad+\beta\left|y\right|+2\gamma\left|z\right|^2+\frac{1}{2\gamma}j_t\left(2\gamma u\right),
\end{aligned}$$

where we used the inequalities $(a+b)^2 \leq a^2+b^2$ and (4.4.10) (with $\gamma_1 = \gamma_2 := 1/(2\gamma)$).

Then, since $(\overline{Y}^{i-1}, \overline{Z}^{i-1}, \overline{U}^{i-1}) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$, we know that the term which does not depend on (y,z,u) above satisfies the same integrability condition as $g_t(0,0,0)$ in (4.2.6). Therefore, since $g^i(0,0,0) = 0$, we have one side of the inequality in Assumption 4.2.1(iii), and the other one can be proved similarly. This yields that g^i satisfies Assumption 4.2.1.

Similarly as in Step 1, we will now show that there exists a solution $(Y^i, Z^i, U^i) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$ to the BSDEJ

$$Y^i_t = \xi_i + \int_t^T g^i_s(Y^i_s, Z^i_s, U^i_s)ds - \int_t^T Z^i_s dB_s - \int_t^T \int_E U^i_s(x)\widetilde{\mu}(dx,ds),\ \mathbb{P}-a.s. \tag{4.2.17}$$

Since g satisifies Assumptions 4.2.4, we can define

$$\phi^i_t := D_z g^i_t(y,0,u) = D_z g_t(y, \overline{Z}^{i-1}_t, u),\ \psi^i_t := D_u g^i_t(y,z,0) = D_u g_t(y,z,\overline{U}^{i-1}_t).$$

We then know that

$$\left|\phi^i_t\right| \leq r_t + \theta\left|\overline{Z}^{i-1}_t\right|,\ \left|\psi^i_t\right| \leq m_t + \theta\left|\overline{U}^{i-1}_t\right|,\ \psi^i_t(x) \geq C_1(1\wedge|x|) \geq -1+\delta.$$

Since by hypothesis $(\overline{Z}^{i-1}, \overline{U}^{i-1}) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$, we can define a probability measure $\mathbb{Q}^i$ by

$$\frac{d\mathbb{Q}^i}{d\mathbb{P}} = \mathcal{E}\left(\int_0^T \phi^i_s dB_s + \int_0^T \int_E \psi^i_s(x)\widetilde{\mu}(dx,ds)\right).$$

Now, using the notations of Lemma 4.2.2, we define a generator $\bar{g}^i$ from g^i. It is then easy to check that $\bar{g}^i$ satisfies Assumption 4.2.3. Therefore, by Theorem 4.2.1, we obtain the existence of a solution to the BSDE with generator $\bar{g}^i$ and terminal condition ξ^i under $\mathbb{Q}^i$. Using Lemma 4.2.2, this provides a solution (Y^i, Z^i, U^i) with Y^i bounded to the BSDE (4.2.17). By Lemma 4.2.1 and since g^i satisfies Assumption 4.2.1, the boundedness of Y^i implies that $(Z^i, U^i) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$ and therefore that $(\overline{Y}^i, \overline{Z}^i, \overline{U}^i) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$.

Step 3 Finally, by summing the BSDEJs (4.2.17), we obtain

$$\overline{Y}^n = \xi + \int_t^T g_s(\overline{Y}^n_s, \overline{Z}^n_s, \overline{U}^n_s)ds - \int_t^T \overline{Z}^n_s dB_s - \int_t^T \int_E \overline{U}^n_s(x)\widetilde{\mu}(dx,ds).$$

Since $\overline{Y}^n$ is bounded (because the Y^i are all bounded), Lemma 4.2.1 implies that $(\overline{Z}^n, \overline{U}^n) \in \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$, which ends the proof.

(ii) In the general case $g_t(0,0,0) \neq 0$, we can argue exactly as in Corollary 4.2.2 (see also Proposition 2 in [117]) to obtain the result. □

4.2.8 A uniqueness result

We emphasize that the above Theorems provide an existence result for every bounded terminal condition, but we only have uniqueness when the infinite norm of ξ is small enough. In order to have a general uniqueness result, we add the following assumptions, which were first introduced by Royer [106] and Briand and Hu [19]

Assumption 4.2.5 *For every (y,z,u,u') there exists a predictable and $\mathcal{E}$-mesurable process (γ_t) such that*

$$g_t(y,z,u) - g_t(y,z,u') \leq \int_E \gamma_t(x)(u-u')(x)\nu_t(dx),$$

where there exists constants $C_2 > 0$ and $C_1 \geq -1+\delta$ for some $\delta > 0$, independent of (y,z,u,u') such that

$$C_1(1 \wedge |x|) \leq \gamma_t(x) \leq C_2(1 \wedge |x|).$$

Assumption 4.2.6 *g is jointly convex in (z,u).*

We then have the following result

Theorem 4.2.3 *Assume that $\xi \in \mathbb{L}^\infty$, and that the generator g satisfies either*

(i) *Assumptions 4.2.1, 4.2.4(i),(ii) and 4.2.5.*

(ii) *Assumptions 4.2.1, 4.2.4 and 4.2.6, and that $g(0,0,0)$ and the process α appearing in Assumption 4.2.1(iii) are bounded by some constant $M > 0$.*

Then there exists a unique solution to the BSDEJ (4.2.2).

In order to prove this Theorem, we will use the following comparison Theorem for solutions of BSDEJs.

Proposition 4.2.7 *Let ξ^1 and ξ^2 be two $\mathcal{F}_T$-measurable random variables. Let g^1 be a function satisfying either*

(i) *Assumptions 4.2.1, 4.2.3(i),(ii) and 4.2.5.*

(ii) *Assumptions 4.2.1, 4.2.3(i) and 4.2.6, and that $g^1(0,0,0)$ and the process α appearing in Assumption 4.2.1(iii) are bounded by some constant $M > 0$.*

Let g^2 be another function and for $i = 1,2$, let (Y^i, Z^i, U^i) be the solution of the BSDEJ with terminal condition ξ^i and generator g^i (we assume that existence holds in our spaces), that is to say for every $t \in [0,T]$

$$Y_t^i = \xi^i + \int_t^T g_s^i(Y_s^i, Z_s^i, U_s^i)ds - \int_t^T Z_s^i dBs - \int_t^T \int_E U_s^i(x)\widetilde{\mu}(dx,ds), \ \mathbb{P}-a.s.$$

Assume further that $\xi^1 \leq \xi^2$, $\mathbb{P}-a.s.$ and $g_t^1(Y_t^2, Z_t^2, U_t^2) \leq g_t^2(Y_t^2, Z_t^2, U_t^2)$, $\mathbb{P}-a.s.$ Then $Y_t^1 \leq Y_t^2$, $\mathbb{P}-a.s.$

Moreover in case (i), if in addition we have $Y_0^1 = Y_0^2$, then for all t, $Y_t^1 = Y_t^2$, $Z_t^1 = Z_t^2$ and $U_t^1 = U_t^2$, $\mathbb{P}-a.s.$

Remark 4.2.9 *Of course, we can replace the convexity property in Assumption 4.2.5 by concavity without changing the results of Proposition 4.2.7. Indeed, if Y is a solution to the BSDE with convex generator g and terminal condition ξ, then $-Y$ is a solution to the BSDE with concave generator $\widetilde{g}(y,z,u) := -g(-y,-z,-u)$ and terminal condition $-\xi$. then we can apply the results of Proposition 4.2.7.*

Proof.

Step 1 In order to prove (i), let us note

$$\delta Y := Y^1 - Y^2,\ \delta Z := Z^1 - Z^2,\ \delta U := U^1 - U^2,\ \delta\xi := \xi^1 - \xi^2$$
$$\delta g_t := g_t^1(Y_t^2, Z_t^2, U_t^2) - g_t^2(Y_t^2, Z_t^2, U_t^2).$$

Using Assumption 4.2.3(i), (ii), we know that there exists a bounded process λ and a process η with

$$|\eta_s| \leq \mu\left(\left|Z_s^1\right| + \left|Z_s^2\right|\right),$$

such that

$$\begin{aligned}\delta Y_t = \delta\xi &+ \int_t^T \delta g_s ds + \int_t^T \lambda_s \delta Y_s ds + \int_t^T (\eta_s + \phi_s)\delta Z_s ds \\ &+ \int_t^T g_s^1(Y_s^1, Z_s^1, U_s^1) - g_s^1(Y_s^1, Z_s^1, U_s^2) ds - \int_t^T \int_E \delta U_s(x)\gamma_s(x)\nu_s(dx)ds \\ &+ \int_t^T \int_E \delta U_s(x)\gamma_s(x)\nu_s(dx)ds - \int_t^T \delta Z_s dB_s - \int_t^T \int_E \delta U_s(x)\widetilde{\mu}(dx,ds),\end{aligned} \tag{4.2.18}$$

where γ is the predictable process appearing in the right hand side of Assumption 4.2.5.

Define for $s \geq t$, $e^{\Lambda_s} := e^{\int_t^s \lambda_u du}$, and

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \mathcal{E}\left(\int_t^s (\eta_s + \phi_s)dB_s + \int_t^s \int_E \gamma_s(x)\widetilde{\mu}(dx,ds)\right).$$

Since the Z^i are in $\mathbb{H}^2_{\mathrm{BMO}}$, so is η and by our assumption on γ_s the above stochastic exponential defines a true strictly positive uniformly integrable martingale (see Kazamaki [70]). Then applying Itô's formula and taking conditional expectation under the probability measure $\mathbb{Q}$, we obtain

$$\begin{aligned}\delta Y_t = \mathbb{E}_t^{\mathbb{Q}}&\left[e^{\Lambda_T}\delta\xi + \int_t^T e^{\Lambda_s}\delta g_s ds\right] \\ &+ \mathbb{E}_t^{\mathbb{P}}\left[\int_t^T e^{\Lambda_s}\left(g_s^1(Y_s^1, Z_s^1, U_s^1) - g_s^1(Y_s^1, Z_s^1, U_s^2) - \int_E \gamma_s(x)\delta U_s(x)\nu_s(dx)\right)ds\right] \leq 0,\end{aligned} \tag{4.2.19}$$

using Assumption 4.2.5.

Step 2 The proof of the comparison result when (ii) holds is a generalization of Theorem 5 in [19]. However, due to the presence of jumps our proof is slightly different. For the convenience of the reader, we will highlight the main differences during the proof.

For any $\theta \in (0,1)$ let us denote

$$\delta Y_t := Y_t^1 - \theta Y_t^2,\ \delta Z_t := Z_t^1 - \theta Z_t^2,\ \delta U_t := U_t^1 - \theta U_t^2,\ \delta\xi := \xi^1 - \theta\xi^2.$$

First of all, we have for all $t \in [0,T]$

$$\delta Y_t = \delta\xi + \int_t^T G_s ds - \int_t^T \delta Z_s dB_s - \int_t^T \int_E \delta U_s(x)\widetilde{\mu}(dx,ds),\ \mathbb{P}-a.s.,$$

where

$$G_t := g_t^1(Y_t^1, Z_t^1, U_t^1) - \theta g_t^2(Y_t^2, Z_t^2, U_t^2).$$

We emphasize that unlike in [19], we have not linearized the generator in y using the Assumption 4.2.3(i). It will be clear later on why.

We will now bound G_t from above. First, we rewrite it as

$$G_t = G_t^1 + G_t^2 + G_t^3,$$

where

$$\begin{aligned}
&G_t^1 := g_t^1(Y_t^1, Z_t^1, U_t^1) - g_t^1(Y_t^2, Z_t^1, U_t^1),\ G_t^2 := g_t^1(Y_t^2, Z_t^1, U_t^1) - \theta g_t^1(Y_t^2, Z_t^2, U_t^2)\\
&G_t^3 := \theta\left(g_t^1(Y_t^2, Z_t^2, U_t^2) - g_t^2(Y_t^2, Z_t^2, U_t^2)\right).
\end{aligned}$$

Then, we have using Assumption 4.2.3(i)

$$\begin{aligned}
G_t^1 &= g_t^1(Y_t^1, Z_t^1, U_t^1) - g_t^1(\theta Y_t^2, Z_t^1, U_t^1) + g_t^1(\theta Y_t^2, Z_t^1, U_t^1) - g_t^1(Y_t^2, Z_t^1, U_t^1)\\
&\leq C\left(|\delta y_t| + (1-\theta)\left|y_t^2\right|\right).
\end{aligned} \tag{4.2.20}$$

Next, we estimate G^2 using Assumption 4.2.1 and the convexity in (z,u) of g^1

$$\begin{aligned}
g_t^1(Y_t^2, Z_t^1, U_t^1) &= g_t^1\left(Y_t^2, \theta Z_t^2 + (1-\theta)\frac{Z_t^1-\theta Z_t^2}{1-\theta}, \theta U_t^2 + (1-\theta)\frac{U_t^1-\theta U_t^2}{1-\theta}\right)\\
&\leq \theta g_t^1(Y_t^2, Z_t^2, U_t^2) + (1-\theta) g_t^1\left(Y_t^2, \frac{\delta Z_t}{1-\theta}, \frac{\delta U_t}{1-\theta}\right)\\
&\leq \theta g_t^1(Y_t^2, Z_t^2, U_t^2) + (1-\theta)\left(2M + \beta\left|Y_t^2\right|\right) + \frac{\gamma}{2(1-\theta)}\left|\delta Z_t\right|^2 + \frac{1-\theta}{\gamma} j_t\left(\frac{\gamma}{1-\theta}\delta U_t\right).
\end{aligned}$$

Hence

$$G_t^2 \leq (1-\theta)\left(2M + \beta\left|Y_t^2\right|\right) + \frac{\gamma}{2(1-\theta)}\left|\delta Z_t\right|^2 + \frac{1-\theta}{\gamma} j_t\left(\frac{\gamma}{1-\theta}\delta U_t\right). \tag{4.2.21}$$

Finally, G^3 is negative by assumption. Therefore, using (4.2.20) and (4.2.21), we obtain

$$G_t \leq C\left|\delta Y_t\right| + (1-\theta)\left(\alpha + (\beta + C)\left|Y_t^2\right|\right) + \frac{\gamma}{2(1-\theta)}\left|Z_t\right|^2 + \frac{1-\theta}{\gamma} j_t\left(\frac{\gamma}{1-\theta}\delta U_t\right). \tag{4.2.22}$$

Now we will get rid off the quadratic and exponential terms in z and u using a classical exponential change. Let us then denote for some $\nu > 0$

$$P_t := e^{\nu\delta Y_t},\ Q_t := \nu e^{\nu\delta Y_t}\delta Z_t,\ R_t(x) := e^{\nu\delta Y_{t^-}}\left(e^{\nu\delta U_t(x)} - 1\right).$$

By Itô's formula we obtain for every $t \in [0,T]$, $\mathbb{P}-a.s.$

$$P_t = P_T + \int_t^T \nu P_s \left(G_s - \frac{\nu}{2} |\delta Z_s|^2 - \frac{1}{\nu} j_s(\nu \delta U_s) \right) ds - \int_t^T Q_s dB_s - \int_t^T \int_E R_s(x) \widetilde{\mu}(dx, ds).$$

Now choose $\nu = \gamma/(1-\theta)$. We emphasize that this is here that the presence of jumps forces us to change a bit our proof in comparison with the one in [19]. Indeed, if we had immediately linearized in y then we could not have chosen ν constant such that the quadratic and exponentials terms in (4.2.22) would disappear. This is not a problem in [19], since they can choose ν of the form $M/(1-\theta)$ with M large enough and still make the quadratic term in z disappear. However, in the jump case, the application $\gamma \mapsto \gamma^{-1} j_t(\gamma u)$ is not always increasing, and this trick does not work. Nonetheless, we now define the strictly positive and continuous process

$$D_t := \exp\left(\gamma \int_0^t \left(2M + (\beta + C) \left| Y_t^2 \right| + \frac{C}{1-\theta} |\delta Y_s| \right) ds \right).$$

Applying Itô's formula to $D_t P_t$, we obtain

$$\begin{aligned} d(D_s P_s) = & - \nu D_s P_s \left(G_s - \frac{\nu}{2} |\delta Z_s|^2 - \frac{1}{\nu} j_s(\nu \delta U_s) - C |\delta Y_s| + (1-\theta) \left(2M + (\beta + C) \left| Y_s^2 \right| \right) \right) ds \\ & + D_s Q_s dB_s + \int_E D_{s^-} R_s(x) \widetilde{\mu}(dx, ds). \end{aligned}$$

Hence, using the inequality (4.2.22), we deduce

$$D_t P_t \leq \mathbb{E}_t^{\mathbb{P}} \left[D_T P_T \right], \ \mathbb{P}-a.s.,$$

which can be rewritten

$$\delta Y_t \leq \frac{1-\theta}{\gamma} \ln \left(\mathbb{E}_t^{\mathbb{P}} \left[\exp \left(\gamma \int_t^T \left(2M + (\beta + C) \left| Y_t^2 \right| + \frac{C}{1-\theta} |\delta Y_s| \right) ds + \frac{\gamma}{1-\theta} \delta\xi \right) \right] \right), \ \mathbb{P}-a.s.$$

Next, we have

$$\delta\xi = (1-\theta)\xi^1 + \theta \left(\xi^1 - \xi^2 \right) \leq (1-\theta) \left| \xi^1 \right|.$$

Consequently, we have for some constant $C_0 > 0$, independent of θ, using the fact that Y^2 and ξ^1 are bounded $\mathbb{P}-a.s.$

$$\delta Y_t \leq \frac{1-\theta}{\gamma} \left(\ln(C_0) + \ln \left(\mathbb{E}_t^{\mathbb{P}} \left[\exp \left(\frac{C}{1-\theta} \int_t^T |\delta Y_s| \, ds \right) \right] \right) \right), \ \mathbb{P}-a.s. \tag{4.2.23}$$

We finally argue by contradiction. More precisely, let

$$\mathcal{A} := \left\{ \omega \in \Omega, Y_t^1(\omega) > Y_t^2(\omega) \right\},$$

and assume that $\mathbb{P}(\mathcal{A}) > 0$. Let us then call $\mathcal{N}$ the $\mathbb{P}$-negligible set outside of which (4.2.23) holds. Since $\mathcal{A}$ has a strictly positive probability, $\mathcal{B} := \mathcal{A} \cap (\Omega \backslash \mathcal{N})$ is not empty and also has a strictly positive probability. Then, we would have from (4.2.23) that for every $\omega \in \mathcal{B}$

$$\delta Y_t(\omega) \leq \frac{1-\theta}{\gamma} \ln(C_0) + \frac{C}{\gamma} \int_t^T \|\delta Y_s\|_{\infty, \mathcal{B}} \, ds, \tag{4.2.24}$$

where $\|\cdot\|_{\infty,\mathcal{B}}$ is the usual infinite norm restricted to $\mathcal{B}$.

Now, using the dominated convergence theorem, we can let $\theta \uparrow 1^-$ in (4.2.24) to obtain that for any $\omega \in \mathcal{B}$

$$Y_t^1(\omega) - Y_t^2(\omega) \leq \frac{C}{\gamma} \int_t^T \left\|Y_s^1 - Y_s^2\right\|_{\infty,\mathcal{B}} ds,$$

which in turns implies, since $\mathcal{B} \subset \mathcal{A}$

$$\left\|Y_t^1 - Y_t^2\right\|_{\infty,\mathcal{B}} \leq \frac{C}{\gamma} \int_t^T \left\|Y_s^1 - Y_s^2\right\|_{\infty,\mathcal{B}} ds.$$

But with Gronwall's lemma this implies that $\left\|Y_t^1 - Y_t^2\right\|_{\infty,\mathcal{B}} = 0$ and the desired contradiction. Hence the result.

Step 3 Let us now assume that $Y_0^1 = Y_0^2$ and that we are in the same framework as in Step 1. Using this in (4.2.19) above when $t = 0$, we obtain

$$\begin{aligned}0 &= \mathbb{E}^{\mathbb{Q}}\left[e^{\Lambda_T}\delta\xi + \int_0^T e^{\Lambda_s}\delta g_s ds\right] \\ &+ \mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{\Lambda_s}\left(g_s^1(Y_s^1, Z_s^1, U_s^1) - g_s^1(Y_s^1, Z_s^1, U_s^2) - \int_E \delta U_s(x)\gamma_s(x)\nu_s(dx)\right) ds\right] \leq 0. \quad (4.2.25)\end{aligned}$$

Hence, since all the above quantities have the same sign, this implies in particular that

$$e^{\Lambda_T}\delta\xi + \int_0^T e^{\Lambda_s}\delta g_s ds = 0, \ \mathbb{P} - a.s.$$

Moreover, we also have $\mathbb{P} - a.s.$

$$\int_0^T e^{\Lambda_s}\left(g_s^1(Y_s^1, Z_s^1, U_s^1) - g_s^1(Y_s^1, Z_s^1, U_s^2)\right) ds = \int_0^T e^{\Lambda_s}\left(\int_E \delta U_s(x)\gamma_s(x)\nu_s(dx)\right) ds.$$

Using this result in (4.2.18), we obtain with Itô's formula

$$\begin{aligned}\delta Y_t &= \int_0^T e^{\Lambda_s}\left(\int_E \delta U_s(x)\gamma_s(x)\nu_s(dx)\right) ds - \int_t^T e^{\Lambda_s}\delta Z_s(dB_s - (\eta_s + \phi_s)ds) \\ &- \int_t^T \int_E e^{\Lambda_{s^-}}\delta U_s(x)\widetilde{\mu}(dx, ds). \quad (4.2.26)\end{aligned}$$

The right-hand side is a martingale under $\mathbb{Q}$ with null expectation. Thus, since $\delta Y_t \leq 0$, this implies that $Y_t^1 = Y_t^2$, $\mathbb{P} - a.s.$ Using this in (4.2.26), we obtain that the martingale part must be equal to 0, which implies that $\delta Z_t = 0$ and $\delta U_t = 0$. □

Remark 4.2.10 *In the above proof of the comparison theorem in case (i), we emphasize that it is actually sufficient that, instead of Assumption 4.2.5, the generator g satisfies*

$$g_s^1(Y_s^1, Z_s^1, U_s^1) - g_s^1(Y_s^1, Z_s^1, U_s^2) \leq \int_E \gamma_s(x)\delta U_s(x)\nu_s(dx),$$

for some γ_s such that

$$C_1(1 \wedge |x|) \leq \gamma_s(x) \leq C_2(1 \wedge |x|).$$

Besides, this also holds true for the comparison Theorem for Lipschitz BSDEs with jumps proved by Royer (see Theorem 2.5 in [106]).

We can now prove Theorem 4.2.3

Proof. [Proof of Theorem 4.2.3] First let us deal with the question of existence.

(i) If g satisfies Assumptions 4.2.1, 4.2.4(i),(ii) and 4.2.5, the existence part can be obtained exactly as in the previous proof, starting from a small terminal condition, and using the fact that Assumption 4.2.5 implies that g is Lipschitz in u. Thus we omit it.

(ii) If g satisfies Assumptions 4.2.1 and 4.2.4, then we already proved existence for bounded terminal conditions.

The uniqueness is then a simple consequence of the above comparison theorem. $\square$

Remark 4.2.11 *As a consequence of the nonlinear Doob-Meyer decomposition that we prove in the next Section, we will obtain a reverse comparison Theorem in Corollary 4.3.1.*

4.2.9 A priori estimates and stability

In this subsection, we show that under our hypotheses, we can obtain *a priori* estimates for quadratic BSDEs with jumps. We have the following results

Proposition 4.2.8 *Let $(\xi^1, \xi^2) \in \mathbb{L}^\infty \times \mathbb{L}^\infty$ and let g be a function satisfying Assumptions 4.2.1, 4.2.3(i),(ii) and 4.2.5. Let us consider for $i = 1, 2$ the solutions $(Y^i, Z^i, U^i) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\mathrm{BMO}} \times \mathbb{J}^2_{\mathrm{BMO}}$ of the BSDEs with generator g and terminal condition ξ^i (once again existence is assumed). Then we have for some constant $C > 0$*

$$\begin{aligned}
&\left\|Y^1 - Y^2\right\|_{\mathcal{S}^\infty} + \left\|U^1 - U^2\right\|_{L^\infty(\nu)} \leq C \left\|\xi^1 - \xi^2\right\|_\infty \\
&\left\|Z^1 - Z^2\right\|^2_{\mathbb{H}^2_{\mathrm{BMO}}} + \left\|U^1 - U^2\right\|^2_{\mathbb{J}^2_{\mathrm{BMO}}} \leq C \left\|\xi^1 - \xi^2\right\|_\infty .
\end{aligned}$$

Proof. Following exactly the same arguments as in Step 1 of the proof Proposition 4.2.7, we obtain with the same notations

$$Y^1_t - Y^2_t = \mathbb{E}^{\mathbb{Q}}_t \left[e^{\Lambda_T}(\xi^1 - \xi^2)\right] \leq C \left\|\xi^1 - \xi^2\right\|_\infty , \ \mathbb{P} - a.s.$$

Notice then that this implies as usual that there is a version of $(U^1 - U^2)$ (still denoted $(U^1 - U^2)$ for simplicity) which is bounded by $2\left\|Y^1 - Y^2\right\|_{\mathcal{S}^\infty}$. This gives easily the first estimate.

Let now $\tau \in \mathcal{T}_{0,T}$ be a stopping time. Denote also

$$\delta g_s := g_s(Y^1_s, Z^1_s, U^1_s) - g_s(Y^2_s, Z^2_s, U^2_s).$$

By Itô's formula, we have using standard calculations

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}_\tau \left[\int_\tau^T |Z_s|^2 ds + \int_\tau^T \int_E U^2_s(x)\nu_s(dx)ds\right] &\leq \mathbb{E}^{\mathbb{P}}_\tau \left[\left|\xi^1 - \xi^2\right|^2 + 2\int_\tau^T (Y^1_s - Y^2_s)\delta g_s ds\right] \\
&\leq \left\|\xi^1 - \xi^2\right\|^2_\infty + 2\left\|Y^1 - Y^2\right\|_{\mathcal{S}^\infty} \mathbb{E}^{\mathbb{P}}_\tau \left[\int_\tau^T |\delta g_s|\, ds\right].
\end{aligned} \tag{4.2.27}$$

Then, using Assumption 4.2.1, we estimate

$$\begin{aligned}|\delta g_t| &\leq C\left(|g_t(0,0,0)| + \alpha_t + \sum_{i=1,2}\left|Y_t^i\right| + \left|Z_t^i\right|^2 + j_t\left(\gamma U^i\right) + j_t\left(-\gamma U^i\right)\right)\\ &\leq C\left(|g_t(0,0,0)| + \alpha_t + \sum_{i=1,2}\left|Y_t^i\right| + \left|Z_t^i\right|^2 + \left\|U_t^i\right\|^2_{L^2(\nu)}\right),\end{aligned}$$

where we used the fact that for every x in a compact subset of $\mathbb{R}$, $0 \leq e^x - 1 - x \leq Cx^2$. Using this estimate and the integrability assumed on $g_t(0,0,0)$ and α_t in (4.2.27) entails

$$\begin{aligned}&\mathbb{E}_\tau^{\mathbb{P}}\left[\int_\tau^T |Z_s|^2\,ds + \int_\tau^T\int_E U_s^2(x)\nu_s(dx)ds\right]\\ &\leq \left\|\xi^1-\xi^2\right\|_\infty^2 + C\left\|\xi^1-\xi^2\right\|_\infty\left(1 + \sum_{i=1,2}\left\|Y^i\right\|_{\mathcal{S}^\infty} + \left\|Z^i\right\|_{\mathbb{H}^2_{\rm BMO}} + \left\|U^i\right\|_{\mathbb{J}^2_{\rm BMO}}\right)\\ &\leq C\left\|\xi^1-\xi^2\right\|_\infty,\end{aligned}$$

which ends the proof. □

Proposition 4.2.9 *Let $(\xi^1,\xi^2) \in \mathbb{L}^\infty \times \mathbb{L}^\infty$ and let g be a function satisfying Assumptions 4.2.1, 4.2.3(i) and 4.2.6 and such that $g(0,0,0)$ and the process α appearing in Assumption 4.2.1(iii) are bounded by some constant $M > 0$. Let us consider for $i = 1,2$ the solutions $(Y^i,Z^i,U^i) \in \mathcal{S}^\infty \times \mathbb{H}^2_{\rm BMO} \times \mathbb{J}^2_{\rm BMO}$ of the BSDEs with generator g and terminal condition ξ^i (once again existence is assumed). Then we have for some constant $C > 0$*

$$\begin{aligned}&\left\|Y^1-Y^2\right\|_{\mathcal{S}^\infty} + \left\|U^1-U^2\right\|_{L^\infty(\nu)} \leq C\left\|\xi^1-\xi^2\right\|_\infty\\ &\left\|Z^1-Z^2\right\|^2_{\mathbb{H}^2_{\rm BMO}} + \left\|U^1-U^2\right\|^2_{\mathbb{J}^2_{\rm BMO}} \leq C\left\|\xi^1-\xi^2\right\|_\infty.\end{aligned}$$

Proof. Following Step 2 of the proof of Proposition 4.2.7, we obtain for any $\theta \in (0,1)$

$$\frac{Y_t^1-\theta Y_t^2}{1-\theta} \leq \frac{1}{\gamma}\ln\left(\mathbb{E}_t^{\mathbb{P}}\left[\exp\left(\gamma\int_t^T\left(2M + (\beta+C)\left|Y_s^2\right| + \frac{C\left|Y_s^1-\theta Y_s^2\right|}{1-\theta}\right)ds + \frac{\gamma(\xi^1-\theta\xi^2)}{1-\theta}\right)\right]\right),$$

and of course by symmetry, the same holds if we interchange the roles of the exponents 1 and 2.

Since all the quantites above are bounded, we obtain easily after some calculations and after letting $\theta \uparrow 1^-$

$$Y_t^1 - Y_t^2 \leq C\left(\left\|\xi^1-\xi^2\right\|_\infty + \int_t^T \left\|Y_s^1-Y_s^2\right\|_\infty ds\right),\ \mathbb{P}-a.s.,$$

and symmetrically

$$Y_t^2 - Y_t^1 \leq C\left(\left\|\xi^1-\xi^2\right\|_\infty + \int_t^T \left\|Y_s^1-Y_s^2\right\|_\infty ds\right),\ \mathbb{P}-a.s.$$

Hence, we can use Gronwall's lemma to obtain

$$\left\|Y^1-Y^2\right\|_{\mathcal{S}^\infty} \leq C\left\|\xi^1-\xi^2\right\|_\infty.$$

All the other estimates can then be obtained as in the proof of Proposition 4.2.8. □

4.3 Quadratic g-martingales with jumps

The theory of g-expectations was introduced by Peng in [99] as an example of non-linear expectations. Since then, numerous authors have generalized his results, extending them notably to the case of quadratic coefficients (see Ma and Yao [84]). An extension to discontinuous filtrations was obtained by Royer [106] and Lin [83]. In particular, Royer [106] gave domination conditions under which we can write a non-linear expectation as a g-expectation. We refer the interested reader to these papers for more details about these filtration-consistent operators, and we recall for simplicity some of their general properties below.

Let us start with a general definition.

Definition 4.3.1 *Let $\xi \in \mathbb{L}^\infty$ and let g be such that the BSDEJ with generator g and terminal condition ξ has a unique solution and such that comparison in the sense of Proposition 4.2.7 holds (for instance g could satisfy any of the conditions in Theorem 4.2.3). Then for every $t \in [0,T]$, we define the conditional g-expectation of ξ as follows*

$$\mathcal{E}^g_t[\xi] := Y_t,$$

where (Y,Z,U) solves the following BSDEJ

$$Y_t = \xi + \int_t^T g_t(Y_s,Z_s,U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(dx,ds).$$

Remark 4.3.1 *Notice that $\mathcal{E}^g : \mathbb{L}^\infty(\Omega,\mathcal{F}_T,\mathbb{P}) \to \mathbb{L}^\infty(\Omega,\mathcal{F}_t,\mathbb{P})$ does not define a true operator. Indeed, to each bounded $\mathcal{F}_T$-measurable random variable ξ, we associate the value Y_t, which is defined $\mathbb{P}$-a.s., that is to say outside a $\mathbb{P}$-negligible set N, but this set N depends on ξ. We cannot a priori find a common negligible set for all variables in $\mathbb{L}^\infty$, and then define an operator $\mathcal{E}^g$ on a fixed domain, except if we only consider a countable set of variables ξ on which act $\mathcal{E}^g$.*

We have a notion of g-martingales and g-sub(super)martingales.

Definition 4.3.2 *$X \in \mathcal{S}^\infty$ is called a g-submartingale (resp. g-supermartingale) if*

$$\mathcal{E}^g_s[X_t] \geq \text{ (resp. } \leq) X_s,\ \mathbb{P}-a.s., \text{ for any } 0 \leq s \leq t \leq T.$$

X is called a g-martingale if it is both a g-sub and supermartingale.

The following results are easy generalizations of the classical arguments which can be found in [99] or [6], and are consequences of the comparison theorem. We therefore omit the proofs.

Lemma 4.3.1 *$\{\mathcal{E}^g_t\}_{t\geq 0}$ is monotonic increasing and time consistent, i.e.*

- *$\xi_1 \geq \xi_2$, $\mathbb{P}$-a.s. implies that $\mathcal{E}^g_t(\xi_1) \geq \mathcal{E}^g_t(\xi_2)$, $\mathbb{P}$-a.s., $\forall t \geq 0$.*
- *For any bounded stopping times $R \leq S \leq \tau$ and $\mathcal{F}_\tau$-measurable random variable ξ_τ,*

$$\mathcal{E}^g_R(\mathcal{E}^g_S(\xi_\tau)) = \mathcal{E}^g_R(\xi_\tau)\ \mathbb{P}\text{-a.s.} \tag{4.3.1}$$

Definition 4.3.3 *We will say that $\mathcal{E}^g$ is*

(i) Constant additive, if for any stopping times $R \leq S$, any $\mathcal{F}_R$-measurable random variable η_R and any $\mathcal{F}_S$-measurable random variable ξ_S,

$$\mathcal{E}^g_R(\xi_S + \eta_R) = \mathcal{E}^g_R(\xi_S) + \eta_R, \ \mathbb{P}\text{-}a.s.$$

(ii) Positively homogeneous, if for any stopping times $R \leq S$, and any positive $\mathcal{F}_R$-measurable random variable λ,

$$\mathcal{E}^g_R(\lambda \xi_S) = \lambda \mathcal{E}^g_R(\xi_S).$$

(iii) Convex, if for any stopping times $R \leq S$, any random variables (ξ^1_S, ξ^2_S) and any $\lambda \in [0,1]$,

$$\mathcal{E}^g_R(\lambda \xi^1_S + (1-\lambda)\xi^2_S) \leq \lambda \mathcal{E}^g_R(\xi^1_S) + (1-\lambda)\mathcal{E}^g_R(\xi^2_S).$$

The next Lemma shows that the operator $\mathcal{E}^g$ inherits the above properties from the generator g.

Lemma 4.3.2 *(i) If g does not depend on y, then $\mathcal{E}^g$ is constant additive.*

(ii) If g is positively homogeneous in (y,z,u), then $\mathcal{E}^g$ is positively homogeneous.

(iii) If g is moreover right continuous on $[0,T)$ and continuous at T, then the reverse implications of (i) and (ii) are also true.

(iv) $\mathcal{E}^g$ is convex if g is convex in (y,z,u).

(v) If $g^1 \leq g^2$, $\mathbb{P}$-a.s., then $\mathcal{E}^{g^1} \leq \mathcal{E}^{g^2}$. If g^1 and g^2 are moreover right continuous on $[0,T)$ and continuous at T, then the reverse is also true.

Proof.

We adapt the ideas of the proofs in [6] to our context with jumps.

(i) The proof of the first property is exactly the same as the proof of Theorem 6.7.b2 in [6], so we omit it.

(ii) Let $g^\lambda(t,y,z,u) := \frac{1}{\lambda} g_t(\lambda y, \lambda z, \lambda u)$. Then $\left\{\frac{1}{\lambda}\mathcal{E}^g_t(\lambda \xi_S)\right\}_{t\geq 0}$ is a solution of the BSDEJ with coefficient g^λ and terminal condition ξ_S. If $g = g^\lambda$ then

$$\frac{1}{\lambda}\mathcal{E}^g_t(\lambda \xi_S) = \mathcal{E}^g_t(\xi_S),$$

which is the desired result.

(iii) The reverse implications in (i) and (ii) are direct consequences of Corollary 4.3.1.

(iv) Suppose that g is convex in (y,z,u). Let (Y^i, Z^i, U^i) be the unique solution of the BSDE with coefficients (g, ξ^i_S), $i=1,2$, and set

$$\widetilde{Y}_t = \lambda Y^1_t + (1-\lambda) Y^2_t, \ \widetilde{Z}_t = \lambda Z^1_t + (1-\lambda) Z^2_t \text{ and } \widetilde{U}_t(\cdot) = \lambda U^1_t(\cdot) + (1-\lambda) U^2_t(\cdot).$$

We have

$$\begin{aligned}
-d\widetilde{Y}_t &= \left[\lambda g_t(Y_t^1, Z_t^1, U_t^1) + (1-\lambda) g_t(Y_t^2, Z_t^2, U_t^2)\right] dt - \left(\lambda Z_t^1 + (1-\lambda) Z_t^2\right) dB_t \\
&\quad - \int_E (\lambda U_t^1(x) + (1-\lambda) U_t^2(x)) \widetilde{\mu}(dt, dx) \\
&= \left[g_t(\widetilde{Y}_t, \widetilde{Z}_t, \widetilde{U}_t) + k(t, Y_t^1, Y_t^2, Z_t^1, Z_t^2, U_t^1, U_t^2, \lambda)\right] dt - \widetilde{Z}_t dB_t - \int_E \widetilde{U}_t(x) \widetilde{\mu}(dt, dx),
\end{aligned}$$

where

$$k(t, Y_t^1, Y_t^2, Z_t^1, Z_t^2, U_t^1, U_t^2, \lambda) := \lambda g_t(Y_t^1, Z_t^1, U_t^1) + (1-\lambda) g_t(Y_t^2, Z_t^2, U_t^2) - g_t(\widetilde{Y}_t, \widetilde{Z}_t, \widetilde{U}_t),$$

is a non negative function. Then using Proposition 4.2.7 we obtain in particular

$$\mathcal{E}_t^g(\lambda \xi_S^1 + (1-\lambda)\xi_S^2) \le \widetilde{Y}_t = \lambda \mathcal{E}_t^g(\xi_S^1) + (1-\lambda) \mathcal{E}_t^g(\xi_S^2).$$

(v) This last property is a direct consequence of the comparison Theorem 4.2.7. The reverse implication is again a consequence of Corollary 4.3.1. □

Example 4.3.1 *These easy properties allow us to construct examples of time consistent dynamic convex risk measures, by appropriate choices of generator g.*

- *As already pointed out in Remark 4.2.4, defining $g_t(z,u) := \frac{\gamma}{2} |z_t|^2 + \frac{1}{\gamma} j_t(\gamma u_t)$, we obtain the so called entropic risk measure on our particular filtration.*
- *As proved in [106], if we define*

$$g_t(z,u) := \eta |z| + \eta \int_E (1 \wedge |x|) u^+(x) \nu_t(dx) - C_1 \int_E (1 \wedge |x|) u^-(x) \nu_t(dx),$$

  where $\eta > 0$ and $-1 < C_1 \le 0$, then $\mathcal{E}^g$ is a convex risk measure with the following representation $\mathcal{E}_0^g(\xi) = \sup_{\mathbb{Q} \in \mathbf{Q}} \mathbb{E}^{\mathbb{Q}}[\xi]$, with

$$\begin{aligned}
\mathbf{Q} := \Bigg\{ \mathbb{Q}, \frac{d\mathbb{Q}}{d\mathbb{P}} | \mathcal{F}_t = \mathcal{E}\left(\int_0^t \mu_s dB_s + \int_0^t \int_E v_s(x) \widetilde{\mu}(ds, dx) \right) \\
\text{with } \mu \text{ and } v \text{ predictable, } |\mu_s| \le \eta,\ v_s^+(x) \le \eta(1 \wedge x),\ v_s^-(x) \le C_1(1 \wedge x) \Bigg\}.
\end{aligned}$$

- *If we define a linear generator g by*

$$g_t(z,u) := \alpha z + \beta \int_E (1 \wedge |x|) u(x) \nu_t(dx),\ \alpha \in \mathbb{R},\ \beta \ge -1 + \delta \text{ for some } \delta > 0,$$

  then we obtain a linear risk measure, since $\mathcal{E}^g$ will only consist of a linear expectation with respect to the probability measure $\mathbb{Q}$, whose Radon-Nikodym derivative is equal to

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(\alpha B_t + \int_0^t \int_E \beta(1 \wedge |x|) \widetilde{\mu}(ds, dx) \right).$$

In the rest of this section, we will provide important properties of quadratic g-expectations and the associated g-martingales in discontinuous filtrations, which generalize the known results in simpler cases.

4.3.1 Non-linear Doob Meyer decomposition

We start by proving that the non-linear Doob Meyer decomposition first proved by Peng in [100] still holds in our context. We have two different sets of assumptions under which this result holds, and they are both related to the assumptions under which our comparison theorem 4.2.7 holds. From a technical point of view, our proof consists in approximating our generator by a sequence of Lipschitz generators. However, the novelty here is that because of the dependence of the generator in u, we cannot use the classical exponential transformation and then use some truncation arguments, as in [74] and [84]. Indeed, since u lives in an infinite dimensional space, those truncation type arguments no longer work *a priori*. Instead, inspired by [8], we will only use regularizations by inf-convolution, which are known to work in any Banach space.

Theorem 4.3.1 *Let Y be a càdlàg g-submartingale (resp. g-supermartingale) in $\mathcal{S}^\infty$ (we assume that existence and uniqueness for the BSDEs with generator g hold for any bounded terminal condition). Assume further either one of these conditions*

(i) *Assumptions 4.2.1 and 4.2.5 hold, with the addition that the process γ does not depend on (y,z) and that $|g(0,0,0)|+\alpha\leq M$, where α is the process appearing in Assumption 4.2.1(iii) and $M>0$ is constant.*

(ii) *Assumptions 4.2.1, 4.2.3(i) hold, g is concave (resp. convex) in (z,u), $|g(0,0,0)|+\alpha\leq M$, where α is the process appearing in Assumption 4.2.1(iii) and $M>0$ is constant.*

Then there exists a predictable non-decreasing (resp. non-increasing) process A null at 0 and processes $(Z,U)\in\mathbb{H}^2\times\mathbb{J}^2$ such that

$$Y_t=Y_T+\int_t^T g_s(Y_s,Z_s,U_s)ds-\int_t^T Z_s dB_s-\int_t^T\int_E U_s(x)\widetilde{\mu}(dx,ds)-A_T+A_t,\ t\in[0,T].$$

Remark 4.3.2 *We emphasize that the two assumptions in the above theorem are not of the same type. Indeed, Assumption 4.2.5 implies that the generator g is uniformly Lipschitz in u, which is a bit disappointing if we want to work in a quadratic context. This is why we also considered the convexity hypothesis on g, which allows us to retrieve a generator which is quadratic in both (z,u). We do not know whether those two assumptions are necessary or not to obtain the result, but we remind the reader that our theorem encompasses the case of the so-called entropic generator, which has quadratic growth and is convex in (z,u). To the best of our knowledge, this particular case which was already proved in [94], was the only result available in the literature up until now.*

Proof.

First of all, if Y is g-supermartingale, then $-Y$ is a g^--submartingale where

$$g_t^-(y,z,u):=-g_t(-y,-z,-u).$$

Since g^- satisfies exactly the same Assumptions as g, and given that g^- is convex when g is concave, it is clear that we can without loss of generality restrict ourselves to the case of g-submartingales. We start with the first result.

Step 1: Assumptions 4.2.1 and 4.2.5 hold.

We will approximate the generator g by a sequence of functions (g^n) which are uniformly Lipschitz in (y,z) (recall that under the assumed assumptions, g is already Lipschitz in u). We emphasize that unlike most of the literature on quadratic BSDEs, with the notable exception of [8] and [7], we will not use any exponential change in our proof.

Building upon the results of Lepeltier and San Martin [78], we would like to use a sup-convolution to regularize our generator. However, due to the quadratic growth assumption in z, such a sup-convolution is not always well defined. Therefore, we will first use a truncation argument to bound our generator from above by a function with linear growth. Let us thus define for all $n\geq 0$

$$\widetilde{g}_t^n(y,z,u):=g_t(y,z,u)\wedge\left(M+n\left|z\right|-\frac{\gamma}{2}\left|z\right|^2\right),$$

where the constants (α,γ) are the ones appearing in Assumption 4.2.4(ii).

It is clear that we have the following estimates

$$-M-\beta\left|y\right|-\frac{\gamma}{2}\left|z\right|^2-\frac{1}{\gamma}j_t\left(-\gamma u\right)\leq\widetilde{g}_t^n(y,z,u)\leq M+\beta\left|y\right|+n\left|z\right|+\frac{1}{\gamma}j_t\left(\gamma u\right),$$

and that $\widetilde{g}^n$ decreases pointwise to g.

We now define for all $p\geq n\vee\beta$

$$\widetilde{g}_t^{n,p}(y,z,u):=\sup_{(w,v)\in\mathbb{Q}^{d+1}}\left\{\widetilde{g}_t^n(w,v,u)-p\left|y-w\right|-p\left|z-v\right|\right\}.$$

This function is indeed well-defined, since we have for $p\geq n$

$$\begin{aligned}\widetilde{g}_t^{n,p}(y,z,u)&\leq M+\frac{1}{\gamma}j_t(\gamma u)+\sup_{(w,v)\in\mathbb{Q}^{d+1}}\left\{\beta\left|w\right|+n\left|v\right|-p\left|y-w\right|-p\left|z-v\right|\right\}\\&=M+\beta\left|y\right|+n\left|z\right|+\frac{1}{\gamma}j_t(\gamma u).\end{aligned}$$

Moreover, by the results of Lepeltier and San Martin [78], we know that $\widetilde{g}^{n,p}$ is uniformly Lipschitz in (y,z) and that $\widetilde{g}^{n,p}(y,z,u)\downarrow g_t(y,z,u)$ as n and p go to $+\infty$. Finally, we define

$$g_t^n(y,z,u):=\widetilde{g}_t^{n,n}(y,z,u).$$

Then the g^n are uniformly Lipschitz in (y,z,u) and decrease pointwise to g. Now, we want somehow to use the fact that we know that the non-linear Doob-Meyer decomposition holds when the underlying generator is Lipschitz. But this was shown by Royer only when the generator also satisfies Assumption 4.2.5. Therefore, we will now verify that g^n inherits Assumption 4.2.5 from g. First of all, we show that this is true for $\widetilde{g}^n$.

Let $u^1,u^2\in L^\infty(\nu)\cap L^2(\nu)$ and fixe some $(y,z)\in\mathbb{R}^{d+1}$. Then if we have

$$g_t(y,z,u^1)\leq M+n\left|z\right|-\frac{\gamma}{2}\left|z\right|^2\text{ and }g_t(y,z,u^2)\leq M+n\left|z\right|-\frac{\gamma}{2}\left|z\right|^2,$$

then

$$\widetilde{g}_t^n(y,z,u^1)-\widetilde{g}_t^n(y,z,u^2)=g_t(y,z,u^1)-g_t(y,z,u^2),$$

and the result is clear with the same process γ as the one for g.

Similarly, if

$$g_t(y,z,u^1) \geq M + n|z| - \frac{\gamma}{2}|z|^2 \text{ and } g_t(y,z,u^2) \geq M + n|z| - \frac{\gamma}{2}|z|^2,$$

then

$$\widetilde{g}_t^n(y,z,u^1) - \widetilde{g}_t^n(y,z,u^2) = 0,$$

and the desired result also follows by choosing the process γ in Assumption 4.2.5 to be 0.

Finally, if (the remaining case can be treated similarly)

$$g_t(y,z,u^1) \geq M + n|z| - \frac{\gamma}{2}|z|^2 \text{ and } g_t(y,z,u^2) \leq M + n|z| - \frac{\gamma}{2}|z|^2,$$

then

$$\widetilde{g}_t^n(y,z,u^1) - \widetilde{g}_t^n(y,z,u^2) \leq M + n|z| - \frac{\gamma}{2}|z|^2 - g_t(y,z,u^2) \leq g_t(y,z,u^1) - g_t(y,z,u^2),$$

and the desired result follows once more with the same process γ as the one for g.

Next, we show that $\widetilde{g}^{n,p}$ inherits Assumption 4.2.5 from $\widetilde{g}^n$. Indeed, we have

$$\widetilde{g}_t^{n,p}(y,z,u^1) - \widetilde{g}_t^{n,p}(y,z,u^2) \leq \sup_{(w,v)\in\mathbb{Q}^{d+1}} \left\{\widetilde{g}_t^n(w,v,u^1) - \widetilde{g}_t^n(w,v,u^2)\right\},$$

which implies the desired result since the process γ in Assumption 4.2.5 does not depend on (y,z).

Let now Y be a g-submartingale. We will now show that it is also a g^n-submartingale for all $n \geq 0$. Let now $\mathcal{Y}$ (resp. $\mathcal{Y}^n$) be the unique solution of the BSDE with terminal condition Y_T and generator g (resp. g^n). Since g^n satisfies Assumption 4.2.5 and is uniformly Lipschitz in (y,z,u), we can apply the comparison theorem for Lipschitz BSDEJs (see [106]) to obtain

$$Y_t \leq \mathcal{Y}_t \leq \mathcal{Y}_t^n, \ \mathbb{P}-a.s.$$

Hence Y is a g^n-submartingale. We can therefore apply the Doob-Meyer decomposition in the Lipschitz case (see Theorem 1.1 in Lin [83] or Theorem 4.1 in Royer [106]) to obtain the existence of $(Z^n, U^n) \in \mathbb{H}^2 \times \mathbb{J}^2$ and of a predictable non-decreasing process A^n null at 0 such that

$$Y_t = Y_T + \int_t^T g_t^n(Y_s, Z_s^n, U_s^n)ds - \int_t^T Z_s^n dB_s - \int_t^T \int_E U_s^n(x)\widetilde{\mu}(dx,ds) - A_T^n + A_t^n. \tag{4.3.2}$$

Since Y does not depend on n, the martingale part of (4.3.2) neither, which entails that Z^n and U^n are independent of n. We can rewrite (4.3.2) as

$$Y_t = Y_T + \int_t^T g_t^n(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(dx,ds) - A_T^n + A_t^n. \tag{4.3.3}$$

Since g^n converges pointwise to g, the dominated convergence Theorem implies that

$$\int_0^T \left(g_s^n(Y_s,Z_s,U_s) - g_s(Y_s,Z_s,U_s)\right)ds \to 0, \ \mathbb{P}-a.s.$$

Hence, it holds $\mathbb{P}-a.s.$ that for all $s\in[0,T]$

$$A_s^n \to A_s := Y_s - Y_0 + \int_0^s g_r(Y_r,Z_r,U_r)dr - \int_0^s Z_r dB_r - \int_0^s \int_E U_r(x)\widetilde{\mu}(dx,dr).$$

Furthermore, it is easy to see that A is still a predictable non-decreasing process null at 0.

Step 2: The concave case.

We have seen in the above proof that the main ingredients to obtain the desired decomposition are the comparison theorem and the non-linear Doob-Meyer decomposition in the Lipschitz case. As we have already seen in our comparison result of Proposition 4.2.7, Assumption 4.2.5 plays, at least formally, the same role as the concavity/convexity assumption 4.2.6. Moreover, we show in the Appendix (see Proposition A.0.3) that the non linear Doob-Meyer decomposition also holds in the Lipschitz case under Assumption 4.2.6 instead of Assumption 4.2.5. We are therefore led to proceed exactly as in the previous step. Define thus

$$\widetilde{g}_t^n(y,z,u) := g_t(y,z,u) \wedge \left(M + n|z| + n\|u\|_{L^2(\nu_t)} - \frac{\gamma}{2}|z|^2 - \frac{1}{\gamma}j_t(\gamma u)\right).$$

Then $\widetilde{g}^n$ is still concave as the minimum of two concave functions, converges pointwise to g and verifies

$$-M-\beta|y| - \frac{\gamma}{2}|z|^2 - \frac{1}{\gamma}j_t(-\gamma u) \leq \widetilde{g}_t^n(y,z,u) \leq M + \beta|y| + n|z| + n\|u\|_{L^2(\nu_t)}.$$

Thanks to this estimate the following sup-convolution is well defined for $p\geq \beta\vee n$

$$\widetilde{g}_t^{n,p}(y,z,u) := \sup_{(w,v,r)\in\mathbb{Q}^{d+1}\times L^2(\nu_t)} \left\{\widetilde{g}_t^n(w,v,r) - p|y-w| - p|z-v| - p\|u-r\|_{L^2(\nu_t)}\right\},$$

and is still concave as the sup-convolution of concave functions.

We can then finish the proof exactly as in Step 1, using the comparison theorem of Proposition 4.2.7 and the non-linear Doob-Meyer decomposition given by Proposition A.0.3. □

Remark 4.3.3 *Following the obtention of this non-linear Doob-Meyer decomposition, it is interesting to wonder whether we can say anything about the non-decreasing process A (apart from saying that it is predictable). For instance, since we are working with bounded g-supermartingales, we may think that K can also be bounded. However, it is already known for classical supermartingales (corresponding to the case $g=0$) that this is not true. Indeed, let X be a supermartingale and let A be the predictable non-decreasing process appearing in its Doob-Meyer decomposition. Then, the inequality $|X_t|\leq M$ for all t only implies that*

$$\mathbb{E}^{\mathbb{P}}\left[(A_t)^p\right] \leq p!M^p, \text{ for all } p\geq 1.$$

Since we have

$$\mathbb{E}_t^{\mathbb{P}}\left[X_t - X_T\right] = \mathbb{E}_t^{\mathbb{P}}\left[A_T - A_t\right],$$

we may then wonder if there could exists another non-decreasing process C_t bounded but not necessarily adapted such that

$$\mathbb{E}_t^{\mathbb{P}}\left[X_t - X_T\right] = \mathbb{E}_t^{\mathbb{P}}\left[C_T - C_t\right]. \tag{4.3.4}$$

This result is then indeed true, and as shown by Meyer [90], if X is càdlàg, positive, bounded by some constant M, then if we denote $\dot{X}$ the predictable projection of X, the non-decreasing process C in (4.3.4) is given by

$$C_T - C_t = M\left(1 - \exp\left(-\int_t^T \frac{dA_s^c}{M - \dot{X}_s}\right)\prod_{t<s\leq T}\left(1 - \frac{\Delta A_s}{M - \dot{X}_s}\right)\right), \tag{4.3.5}$$

where A^c is the continuous part of A.

If we now consider a g-supermartingale Y satisfying either one of the assumptions in Theorem 4.3.1, then a simple application of Itô's formula shows that

$$\widetilde{Y}_t := \exp\left(\gamma Y_t + \gamma M t + \gamma\beta\int_0^t |Y_s|\,ds\right),$$

is a bounded classical supermartingale, which therefore admits the following decomposition

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \widetilde{Z}_s dB_s + \int_0^t\int_E \widetilde{U}_s(x)\widetilde{\mu}(dx,ds) + \widetilde{A}_t,\ \mathbb{P}-a.s., \tag{4.3.6}$$

for some $(\widetilde{Z},\widetilde{U}) \in \mathbb{H}^2 \times \mathbb{J}^2$ and some predictable non-decreasing process $\widetilde{A}$. We can the apply Meyer's result to obtain

$$\mathbb{E}_t^{\mathbb{P}}\left[\widetilde{Y}_t - \widetilde{Y}_T\right] = \mathbb{E}_t^{\mathbb{P}}\left[D_T - D_t\right],$$

where D is given by (4.3.5).

Then, applying Itô's formula to $\ln(\widetilde{Y}_t)$ in (4.3.6), we can show after some calculations that

$$Y_t = Y_T + \int_t^T g_t(Y_s,Z_s,U_s)ds - \int_t^T Z_s dB_s - \int_t^T\int_E U_s(x)\widetilde{\mu}(dx,ds) + A_T - A_t,$$

where $(Z,U) \in \mathbb{H}^2 \times \mathbb{J}^2$ and A is a predictable process with finite variation, and where (Z,U,A) can be computed explicitly from $(\widetilde{Z},\widetilde{U},\widetilde{A})$.

By uniqueness of the non-linear Doob-Meyer decomposition fro Y, A is actually non-decreasing, and we have a result somehow similar to that of Meyer, using the relation between $\widetilde{A}$ and A. It would of course be interesting to pursue further this study.

We end this section with a converse comparison result for our class of quadratic BSDEs, which is a consequence of the previous Doob-Meyer decomposition.

Corollary 4.3.1 *Let g^1 be a function satisfying either one of the assumptions in Theorem 4.3.1 and g^2 be another function. We furthermore suppose that $t \mapsto g_t^i(\cdot,\cdot,\cdot)$ is right continuous in $t \in [0,T)$ and continuous at T, for $i = 1,2$. For any $\xi \in L^\infty$, denote for $i = 1,2$, $Y_t^{i,\xi}$ the solution of the BSDEJ with generator g^i and terminal condition ξ (existence and uniqueness are assumed to hold in our spaces). If we have*

$$Y_t^{1,\xi} \leq Y_t^{2,\xi},\ t \in [0,T],\ \forall \xi \in L^\infty,\ \mathbb{P}-a.s.,$$

then we have

$$g_t^1(y,z,u) \leq g_t^2(y,z,u),\ \forall (t,y,z,u),\ \mathbb{P}-a.s.$$

Proof. For any $\xi \in L^\infty$, the assumption of the Corollary is equivalent to saying that $Y^{2,\xi}$ is a g^1-supermartingale. Given the assumptions on g^1, we can apply Theorem 4.3.1 to obtain the existence of $(\widetilde{Z}^{2,\xi}, \widetilde{U}^{2,\xi}, A^{2,\xi})$ such that for any $0 \leq s < t \leq T$

$$Y_s^{2,\xi} = Y_t^{2,\xi} + \int_s^t g_r^1\left(Y_r^{2,\xi}, \widetilde{Z}_r^{2,\xi}, \widetilde{U}_r^{2,\xi}\right) dr - \int_s^t \widetilde{Z}_r^{2,\xi} dB_r - \int_s^t \int_E \widetilde{U}_r^{2,\xi}(x)\widetilde{\mu}(dx,dr) + A_t^{2,\xi} - A_s^{2,\xi},\ \mathbb{P}-a.s. \tag{4.3.7}$$

Moreover, if we denote $(Y^{2,\xi}, Z^{2,\xi}, U^{2,\xi})$ the solution of the BSDEJ with generator g^2 and terminal condition ξ, we also have by definition

$$Y_s^{2,\xi} = Y_t^{2,\xi} + \int_s^t g_r^2\left(Y_r^{2,\xi}, Z_r^{2,\xi}, U_r^{2,\xi}\right) dr - \int_s^t Z_r^{2,\xi} dB_r - \int_s^t \int_E U_r^{2,\xi}(x)\widetilde{\mu}(dx,dr),\ \mathbb{P}-a.s. \tag{4.3.8}$$

Identifying the martingale parts in (4.3.7) and (4.3.8), we obtain that $\mathbb{P}-a.s.$, $\widetilde{Z}^{2,\xi} = Z^{2,\xi}$ and $\widetilde{U}^{2,\xi} = U^{2,\xi}$. Furthermore, this implies by taking the expectation that

$$\frac{1}{t-s}\int_s^t \mathbb{E}^{\mathbb{P}}\left[g_r^1\left(Y_r^{2,\xi}, Z_r^{2,\xi}, U_r^{2,\xi}\right)\right] dr \leq \frac{1}{t-s}\int_s^t \mathbb{E}^{\mathbb{P}}\left[g_r^2\left(Y_r^{2,\xi}, Z_r^{2,\xi}, U_r^{2,\xi}\right)\right] dr.$$

Now, we finish using the same argument as in Chen [27]. Let $\xi = X_T$ where for a given (s, y_0, z_0, u_0), X is the solution of the SDE (existence and uniqueness are classical, see for instance Jacod [66])

$$X_t = y_0 - \int_s^t g_r^2(X_r, z_0, u_0) dr + \int_s^t z_0 dB_r + \int_s^t \int_E u_0(x)\widetilde{\mu}(dx,dr).$$

Letting $t \longrightarrow s^+$, we obtain $g_s^1(y_0, z_0, u_0) \leq g_s^2(y_0, z_0, u_0)$, which is the desired result. $\square$

4.3.2 Upcrossing inequality

In this subsection, we prove the so-called upcrossing inequality for quadratic g-submartingales, which is similar to the one obtained by Ma and Yao [84] in the case without jumps. This property is essential for the study of path regularity of g-submartingales.

Theorem 4.3.2 *Let (X_t) be a g-submartingale (reps. g-supermartingale) and assume that either one of the following holds (as usual we assume existence and uniqueness for the solutions of BSDEJ driven by g with any bounded terminal condition)*

(i) *Assumptions 4.2.1, 4.2.3(i),(ii) and 4.2.5 hold, with the addition that $|g(0,0,0)| + \alpha \leq M$, where α is the process appearing in Assumption 4.2.1(iii) and $M > 0$ is constant.*

(ii) *Assumptions 4.2.1, 4.2.3(i) hold, g is concave (reap. convex), with the addition that $|g(0,0,0)| + \alpha \leq M$, where α is the process appearing in Assumption 4.2.1(iii) and $M > 0$ is constant.*

Set

$$J := \gamma M \frac{e^{\beta T} - 1}{\beta} + \gamma e^{\beta T} \|X\|_{\mathbb{D}^\infty}.$$

Denote for any $\theta \in (0,1)$

$$\widetilde{X}_t := X_t + k(J+1)t,\ \widehat{X}_t := \exp\left(k_\theta(1+J)t + \frac{\gamma\theta}{1-\theta}X_t\right) t \in [0,T],$$

where k and k_θ are a well-chosen constants depending on θ, C, M, β and γ, the constants in Assumption 4.2.1. Let $0 = t_0 < t_1 < ... < t_n = T$ be a subdivision of $[0, T]$ and let $a < b$, we denote $U_a^b[\widetilde{X}, n]$, the number of upcrossings of the interval $[a, b]$ by $(\widetilde{X}_{t_j})_{0\leq j\leq n}$. Then

- *If (i) above holds, there exists a BMO process $(\lambda_t^n,\ t \in [0, T])$ such that*

$$\mathbb{E}^{\mathbb{P}}\left[U_a^b[X, n]\mathcal{E}\left(\int_0^T (\lambda_s^n + \phi_s)\,dB_s + \int_0^T \int_E \gamma_s(x)\widetilde{\mu}(dx, ds)\right)\right] \leq \frac{\|X\|_{\mathcal{S}^\infty} + 2k(J + 1)T + |a|}{b - a},$$

  where ϕ and γ are defined in Assumption 4.2.3(ii) and 4.2.5, and such that

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^{t_n} |\lambda_s^n|^2\,ds\right] \leq C_1,$$

  a constant independent of the choice of the subdivision.

- *If (ii) above holds, then for any $\theta \in (0, 1)$*

$$\mathbb{E}^{\mathbb{P}}\left[U_a^b[X, n]\right] \leq \frac{\exp\left(\frac{\gamma\theta}{1-\theta}\|X\|_{\mathcal{S}^\infty}\right) + \exp\left(\frac{\gamma\theta}{1-\theta}|a|\right)}{\exp\left(\frac{\gamma\theta}{1-\theta}b\right) - \exp\left(\frac{\gamma\theta}{1-\theta}a\right)}.$$

Proof. As usual, we can restrict ourselves to the g-submartingale case.

Step 1: When (i) holds.

For any $j \in 1, \cdots, n$, we consider the following BSDE with jumps

$$Y_t^j = X_{t_j} + \int_t^{t_j} g_s(Y_s^j, Z_s^j, U_s^j)ds - \int_t^{t_j} Z_s^j dB_s - \int_t^{t_j}\int_E U_s^j(x)\widetilde{\mu}(dx, ds),\ 0 \leq t \leq t_j,\ \mathbb{P}-a.s. \quad (4.3.9)$$

From Proposition 4.2.6 one has

$$\left\|Y^j\right\|_{\mathcal{S}^\infty} \leq \gamma M \frac{e^{\beta(t_j - t_{j-1})} - 1}{\beta} + \gamma e^{\beta(t_j - t_{j-1})}\left\|X_{t_j}\right\|_{\mathcal{S}^\infty} \leq J. \quad (4.3.10)$$

We can rewrite (4.3.9) as follows

$$\begin{aligned}
Y_t^j = X_{t_j} &+ \int_t^{t_j} \left[g_s(Y_s^j, Z_s^j, U_s^j) - g_s(Y_s^j, 0, U_s^j)\right] ds \\
&+ \int_t^{t_j} \left[g_s(Y_s^j, 0, U_s^j) - g_s(0, 0, U_s^j)\right] ds + \int_t^{t_j} \left[g_s(0, 0, U_s^j) - g_s(0, 0, 0)\right] ds \\
&+ \int_t^{t_j} g_s(0, 0, 0)ds - \int_t^{t_j} Z_s^j dB_s - \int_t^{t_j}\int_E U_s^j(x)\widetilde{\mu}(dx, ds),\ 0 \leq t \leq t_j,\ \mathbb{P} - a.s.
\end{aligned}$$

Then by Assumption 4.2.3(ii), there exist a bounded process η^n and BMO processes ϕ, λ^n with

$$|\lambda_t^n| \leq \mu\left|Z_t^j\right|,\ \mathbb{P} - a.s., \forall t \in [t_{j-1}, t_j],$$

such that

$$\begin{aligned}
Y_t^j &= X_{t_j} + \int_t^{t_j} \left[(\lambda_s^n + \phi_s) Z_s^j + \eta_s^n Y_s^j \right] ds + \int_t^{t_j} \left[g_s(0,0,U_s^j) - g_s(0,0,0) \right] ds \\
&\quad + \int_t^{t_j} g_s(0,0,0) ds - \int_t^{t_j} Z_s^j dB_s - \int_t^{t_j} \int_E U_s^j(x) \widetilde{\mu}(dx,ds) \\
&= X_{t_j} + \int_t^{t_j} \left[\eta_s^n Y_s^j + g_s(0,0,0) \right] ds + \int_t^{t_j} \left[g_s(0,0,U_s^j) - g_s(0,0,0) \right] ds \\
&\quad - \int_t^{t_j} \int_E \gamma_s(x) U_s^j(x) \nu_s(dx) ds - \int_t^{t_j} Z_s^j \left(dB_s - (\lambda_s^n + \phi_s) ds \right) \\
&\quad - \int_t^{t_j} \int_E U_s^j(x) \left[\widetilde{\mu}(dx,ds) - \gamma_s(x) \nu_s(dx) ds \right] \\
&\leq X_{t_j} + k(J+1)(t_j - t) - \int_t^{t_j} Z_s^j dB_s^n - \int_t^{t_j} \int_E U_s^j(x) \widetilde{\mu}_1(ds,dx),
\end{aligned}$$

for some positive constant k and where

$$B_t^n := B_t - \int_0^t (\lambda_s^n + \phi_s) ds \text{ and } \widetilde{\mu}_1(ds,dx) = \widetilde{\mu}(dx,ds) - \gamma_s(x)\nu_s(dx)ds.$$

With our Assumptions, we can once more use Girsanov's theorem and define an equivalent probability measure $\mathbb{P}^n$ such that

$$\frac{d\mathbb{P}^n}{d\mathbb{P}} = \mathcal{E}\left(\int_0^{\cdot} (\lambda_s^n + \phi_s) dB_s + \int_0^{\cdot} \int_E \gamma_s(x) \widetilde{\mu}(dx,ds) \right)_{t_n}.$$

Taking the conditional expectation on both sides of the above inequality, we obtain

$$\mathcal{E}_t^g \left[X_{t_j} \right] = Y_t^j \leq \mathbb{E}_t^{\mathbb{P}^n} \left[X_{t_j} \right] + k(J+1)(t_j - t), \ \mathbb{P} - a.s., \forall t \in [t_{j-1}, t_j].$$

In particular, taking $t = t_{j-1}$ we have

$$X_{t_{j-1}} \leq \mathcal{E}_{t_{j-1}}^g \left[X_{t_j} \right] \leq \mathbb{E}_{t_{j-1}}^{\mathbb{P}^n} \left[X_{t_j} \right] + k(J+1)(t_j - t_{j-1}), \ \mathbb{P} - a.s.$$

Hence $(\widetilde{X}_{t_j})_{j=0..n}$ is a $\mathbb{P}^n$-submartingale. Define now the quantities

$$u_t := b + k(J+1)t \text{ and } l_t := a + k(J+1)t.$$

Then, we can apply the classical upcrossing inequality for $\widetilde{X}$, u and l

$$\mathbb{E}^{\mathbb{P}^n}[U_l^u[\widetilde{X},n]] \leq \frac{\mathbb{E}^{\mathbb{P}^n}\left[\left(\widetilde{X}_T - l_T \right)^+ \right]}{u_T - l_T} \leq \frac{\|X\|_{\mathcal{S}^\infty} + 2k(J+1)T + |a|}{b-a}.$$

Notice then finally that $U_l^u[\widetilde{X},n] = U_a^b[X,n]$, which implies the desired result.

Step 2: When (ii) holds.

Using the same arguments as in the proof of (ii) of Proposition 4.2.7, we can show, using the concavity of g and Assumption 4.2.1, that for any $\theta \in (0,1)$

$$\begin{aligned}\theta g_t(y,z,u) &\leq g_t(0,0,0) + C\theta\,|y| + (1-\theta)M + \frac{\gamma}{1-\theta}\theta^2\,|z|^2 + \frac{1-\theta}{\gamma} j_t\left(\frac{\gamma\theta u}{1-\theta}\right) \\ &\leq C\theta\,|y| + (2-\theta)M + \frac{\gamma}{1-\theta}\theta^2\,|z|^2 + \frac{1-\theta}{\gamma} j_t\left(\frac{\gamma\theta u}{1-\theta}\right).\end{aligned}$$

Hence, considering as in Step 1 for any $j = 0...n$ the solution Y^j of (4.3.9), we can use the same exponential transformation as in Step 2 of the proof of Proposition 4.2.7 to obtain

$$\begin{aligned}\exp\left(\frac{\gamma\theta}{1-\theta}Y_t^j\right) &\leq \mathbb{E}_t^{\mathbb{P}}\left[\exp\left(\frac{\gamma\theta}{1-\theta}X_{t_j} + \gamma\frac{2-\theta}{1-\theta}M(t_j-t) + \frac{\gamma C}{1-\theta}\int_t^{t_j}\left|Y_s^j\right| ds\right)\right] \\ &\leq \exp\left(\gamma\left(\frac{CJ}{1-\theta} + M(2-\theta)\right)(t_j-t)\right)\mathbb{E}_t^{\mathbb{P}}\left[\exp\left(\frac{\gamma\theta}{1-\theta}X_{t_j}\right)\right] \\ &\leq \exp\left(k_\theta(1+J)(t_j-t)\right)\mathbb{E}_t^{\mathbb{P}}\left[\exp\left(\frac{\gamma\theta}{1-\theta}X_{t_j}\right)\right],\end{aligned}$$

for some constant k_θ depending on γ, C, M and θ.

As in Step 1, choosing $t = t_{j-1}$ and using the fact that X is a g-submartingale, we deduce that $(\widehat{X}_{t_j})_{j=0..n}$ is a $\mathbb{P}$-submartingale, where

$$\widehat{X}_t := \exp\left(k_\theta(1+J)t + \frac{\gamma\theta}{1-\theta}X_t\right).$$

Define now the quantities

$$u_t^\theta := \exp\left(k_\theta(1+J)t + \frac{\gamma\theta}{1-\theta}b\right) \text{ and } l_t^\theta := \exp\left(k_\theta(1+J)t + \frac{\gamma\theta}{1-\theta}a\right).$$

We apply the classical upcrossing inequality for $\widehat{X}$, u^θ and l^θ

$$\mathbb{E}^{\mathbb{P}}[U_{l^\theta}^{u^\theta}[\widehat{X},n]] \leq \frac{\mathbb{E}^{\mathbb{P}}\left[\left(\widehat{X}_T - l_T^\theta\right)^+\right]}{u_T^\theta - l_T^\theta} \leq \frac{\exp\left(\frac{\gamma\theta}{1-\theta}\|X\|_{\mathcal{S}^\infty}\right) + \exp\left(\frac{\gamma\theta}{1-\theta}|a|\right)}{\exp\left(\frac{\gamma\theta}{1-\theta}b\right) - \exp\left(\frac{\gamma\theta}{1-\theta}a\right)},$$

which ends the proof, noticing that $U_{l^\theta}^{u^\theta}[\widehat{X},n] = U_a^b[X,n]$. □

With this upcrossing inequality in hand, we can argue exactly as in [84] (see Corollary 5.6) to obtain

Corollary 4.3.2 *Let g be as in Theorem 4.3.2. Then any g-sub(super)martingale X admits a càdlàg modification and furthermore for any countable dense subset $\mathcal{D}$ of $[0,T]$, it holds for all $t \in [0,T]$ that the two following limits exist $\mathbb{P}-a.s.$*

$$\lim_{r\uparrow t,\ r\in\mathcal{D}} X_r \text{ and } \lim_{r\downarrow t,\ r\in\mathcal{D}} X_r.$$

4.4 Dual Representation and Inf-Convolution

We generalize in this section some results of Barrieu and El Karoui [6] to the case of quadratic BSDEs with jumps. We give a dual representation of the related g-expectations, viewed as convex dynamic risk measures and then we compute in an explicit manner the inf-convolution of two convex g-expectations.

4.4.1 Dual Representation of the g-expectation

We will assume in this section that $g_t(y,z,u) = g_t(z,u)$ is independent of y and that the function g is convex. We will prove a dual Legendre-Fenchel type representation for the functional $\mathcal{E}^g$, making use of the Legendre-Fenchel transform of g. This problem has been treated by Barrieu and El Karoui [6] in the case of quadratic BSDEs, we extend it here to the case of quadratic BSDEs with jumps.

In this section, $\mathcal{E}^g$ will correspond to a time consistent dynamic convex risk measures. Hence $\mathcal{E}^g$ admits a dual representation, as in [6]. In this particular case of risk measures constructed from backward SDEs, the penalty function appearing in the dual representation is an integral of the Legendre-Fenchel transform of the generator g. The operator $\mathcal{E}^g$, viewed as a time-consistent dynamic convex risk measure has interesting economic applications in insurance.

For $\mu \in \mathbb{R}^d$ and $v \in L^2(\nu)$, define the Legendre-Fenchel transform of g in (z,u) as follows

$$G_t(\mu, v) := \sup_{(z,u)\in\mathbb{R}^d\times L^2(\nu_t)} \left\{ < \mu, z >_{\mathbb{R}^d} + < v, u >_{L^2(\nu_t)} - g_t(z,u) \right\}.$$

Let $\mathcal{A}$ denote the space of applications $v \in \mathbb{J}^2_{\mathrm{BMO}} \cap L^\infty(\nu)$ such that there exists a constant $\delta > 0$ with $v_t(x) \geq -1 + \delta$, $\mathbb{P} \times dt \times d\nu_t$ -a.e.

Theorem 4.4.1 *Let g be a given convex function in (z,u) and let Assumptions 4.2.1 and 4.2.4 hold ; assume that $g(0,0,0)$ and the process α appearing in Assumption 4.2.1(iii) are bounded by some constant $M > 0$ (then, the existence and uniqueness of the solution of the BSDEJ with generator g and terminal condition $\xi \in L^\infty$ hold by Theorem 4.2.3 (ii)). We then have*

(i) *For any $\xi_T \in L^\infty$,*

$$\mathcal{E}^g_t(\xi_T) = \operatorname*{ess\,sup}^{\mathbb{P}}_{(\mu,v)\in\mathbb{H}^2_{\mathrm{BMO}}\times\mathcal{A}} \left\{ \mathbb{E}^{\mathbb{Q}^{\mu,v}}_t \left[\xi_T - \int_t^T G_s(\mu_s, v_s) ds \right] \right\}, \ \mathbb{P} - a.s.,$$

where $\mathbb{Q}^{\mu,v}$ is the probability measure defined by

$$\frac{d\mathbb{Q}^{\mu,v}}{d\mathbb{P}} = \mathcal{E}\left(\int_0^\cdot \mu_s dB_s + \int_0^\cdot \int_E v_s(x) \widetilde{\mu}(ds,dx) \right).$$

(ii) *Moreover, there exist measurable functions $\overline{\mu}(w,t)$ and $\overline{v}(\omega,t,\cdot)$ such that*

$$\mathcal{E}^g_t(\xi_T) = \mathbb{E}^{\mathbb{Q}^{\overline{\mu},\overline{v}}}_t \left[\xi_T - \int_t^T G_s(\overline{\mu}_s, \overline{v}_s) ds \right], \ \mathbb{P} - a.s. \tag{4.4.1}$$

Proof. Thanks to the Kazamaki criterion (see for instance Lemma 4.1 in [91]), we know that if $\mu \in \mathbb{H}^2_{\mathrm{BMO}}$ and $v \in \mathbb{J}^2_{\mathrm{BMO}}$, then $\Gamma_{\mu,v} := \frac{d\mathbb{Q}^{\mu,v}}{d\mathbb{P}}$ is a true martingale and the probability measure $\mathbb{Q}^{\mu,v}$ is well defined.

$\mathcal{E}^g_t(\xi_T)$ is by definition solution of

$$\begin{aligned}\mathcal{E}^g_t(\xi_T) &= \xi_T + \int_t^T g_s(Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_E U_s(x)\widetilde{\mu}(ds, dx) \\ &= \xi_T + \int_t^T \left[g_s(Z_s, U_s) - < \mu_s, Z_s >_{\mathbb{R}^d} - < v_s, U_s >_{L^2(\nu_s)}\right] ds \\ &\quad - \int_t^T Z_s dB^\mu_s - \int_t^T \int_E U_s(x)\widetilde{\mu}^v(ds, dx),\ \mathbb{P} - a.s.,\end{aligned} \tag{4.4.2}$$

where $B^\mu_t := B_t - \int_0^t \mu_s ds$ is a $\mathbb{Q}^{\mu,v}$-Brownian motion and

$$\widetilde{\mu}^v([0,t], A) := \widetilde{\mu}([0,t], A) - \int_0^t \int_A v_s(x)\nu_s(dx)ds \text{ is a } \mathbb{Q}^{\mu,v} - \text{martingale.}$$

By Lemma 4.2.1, $Z \in \mathbb{H}^2_{\mathrm{BMO}}$ and $U \in \mathbb{J}^2_{\mathrm{BMO}}$. Let us prove that we also have $(Z, J) \in \mathbb{H}^2(\mathbb{Q}^{\mu,v}) \times \mathbb{J}^2(\mathbb{Q}^{\mu,v})$. Indeed, using the number $r > 1$ given by Proposition 4.2.4

$$\begin{aligned}\mathbb{E}^{Q^{\mu,v}}\left[\int_0^T |Z_s|^2 ds\right] &= \mathbb{E}^{\mathbb{P}}\left[\Gamma_{\mu,v}\int_0^T |Z_s|^2 ds\right] \\ &\leq \left(\mathbb{E}^{\mathbb{P}}\left[\Gamma^r_{\mu,v}\right]\right)^{1/r}\left(\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |Z_s|^2 ds\right)^q\right]\right)^{1/q} < +\infty,\end{aligned}$$

where $1/r + 1/q = 1$ and where we used Proposition 4.2.4 and the energy inequality (4.2.3). The proof for J is the same.

Moreover,

$$-G_t(\mu, v) = -\sup_{(z,u)\in\mathbb{R}^d\times L^2(\nu_t)} \left\{\langle \mu, z\rangle_{\mathbb{R}^d} + < v, u >_{L^2(\nu_t)} - g_t(z, u)\right\} \leq g_t(0, 0),$$

which means that $-G_t(\mu, v)$ is $\mathbb{Q}^{\mu,v} \times dt$-integrable. Using these integrability properties and the definition of G, we take the conditional expectation in (4.4.2) to obtain

$$\mathcal{E}^g_t(\xi_T) \leq \mathbb{E}^{\mathbb{Q}^{\mu,v}}_t\left[\xi_T - \int_t^T G_s(\mu_s, v_s)ds\right]. \tag{4.4.3}$$

By our assumptions, g is C^2 in z and twice Fréchet differentiable in u, then $\partial g(Z_t, U_t)$ contains a unique element, where the subdifferential ∂g is defined by

$$\begin{aligned}\partial g(Z_t, U_t) = \Big\{(\mu, v) \in \mathbb{R}^d \times L^2(\nu_t) \text{ s.t } g_t(z', u') \geq g_t(Z_t, U_t) - < \mu, z' - Z_t >_{\mathbb{R}^d} \\ - < v, u' - U_t >_{L^2(\nu_t)},\ \forall (z', u')\Big\}.\end{aligned}$$

We take $(\overline{\mu}, \overline{v}) \in \partial g(Z_t, U_t)$. We have

$$g_t(Z_t, U_t) = < \overline{\mu}_t, Z_t >_{\mathbb{R}^d} + < \overline{v}_t, U_t >_{L^2(\nu_t)} - G_t(\overline{\mu}_t, \overline{v}_t).$$

We refer to [6] for the measurability of $\overline{\mu}$ and $\overline{v}$ with respect to the variable ω. We use Assumption 4.2.3 (which holds, since it is implied by Assumption 4.2.4) to write

$$\begin{aligned}
|g_t(z,u)| &\leq |g_t(z,0)| + \|D_u g_t(z,0)\|^2_{L^2(\nu_t)} \|u\|^2_{L^2(\nu_t)} + C\|u\|^2_{L^2(\nu_t)} \\
&\leq |g_t(0,0)| + C|z|^2 + C\left(1 + \|u\|^2_{L^2(\nu_t)}\right) \\
&\leq C|z|^2 + C\left(1 + \|u\|^2_{L^2(\nu_t)}\right),
\end{aligned}$$

where C is a constant whose value may vary from line to line. Putting the above estimation in G leads to

$$\begin{aligned}
G_t(\overline{\mu}_t, \overline{v}_t) &= \sup_{(z,u)\in\mathbb{R}^d\times L^2(\nu_t)} \left\{ < \overline{\mu}_t, z >_{\mathbb{R}^d} + < \overline{v}_t, u >_{L^2(\nu_t)} - g_t(z,u) \right\} \\
&\geq \sup_{u\in L^2(\nu_t)} \left\{ < \overline{v}_t, u >_{L^2(\nu_t)} - C - C\|u\|^2_{L^2(\nu_t)} \right\} + \sup_{z\in\mathbb{R}^d} \left\{ < \overline{\mu}_t, z >_{\mathbb{R}^d} - \frac{\gamma}{2}|z|^2 \right\} \\
&= \frac{1}{4C}\|\overline{v}_t\|^2_{L^2(\nu_t)} - C + \frac{1}{4C}|\overline{\mu}_t|^2 .
\end{aligned}$$

From this, we deduce that for $\varepsilon < \frac{1}{4C}$,

$$\begin{aligned}
\left(\frac{1}{4C} - \varepsilon\right)\left(\|\overline{v}_t\|^2_{L^2(\nu_t)} + |\overline{\mu}_t|^2\right) &\leq G_t(\overline{\mu}_t, \overline{v}_t) + C - \varepsilon\|\overline{v}_t\|^2_{L^2(\nu_t)} - \varepsilon|\overline{\mu}_t|^2 \\
&= C - g_t(Z_t, U_t) + < \overline{v}_t, U_t >_{L^2(\nu)} - \varepsilon\|\overline{v}_t\|^2_{L^2(\nu)} \\
&\quad + < \overline{\mu}_t, Z_t > - \varepsilon|\overline{\mu}_t|^2 \\
&\leq C - g_t(Z_t, U_t) + \frac{1}{4\varepsilon}\|U_t\|^2_{L^2(\nu)} + \frac{1}{4\varepsilon}|Z_t|^2 .
\end{aligned}$$

Since $|g_t(Z_t,U_t)|^{\frac{1}{2}}$ and U are respectively in $\mathbb{H}^2_{\mathrm{BMO}}$ and $\mathbb{J}^2_{\mathrm{BMO}}$, using the fact that

$$\left(\frac{1}{4C} - \varepsilon\right)\|\overline{v}_t\|^2_{L^2(\nu_t)} \leq \left(\frac{1}{4C} - \varepsilon\right)\left(\|\overline{v}_t\|^2_{L^2(\nu)} + |\overline{\mu}_t|^2\right) \text{ and,}$$

$$\left(\frac{1}{4C} - \varepsilon\right)|\overline{\mu}_t|^2 \leq \left(\frac{1}{4C} - \varepsilon\right)\left(\|\overline{v}_t\|^2_{L^2(\nu_t)} + |\overline{\mu}_t|^2\right),$$

we obtain that $\overline{v}$ is in $\mathbb{J}^2_{\mathrm{BMO}}$ and $\overline{\mu}$ is in $\mathbb{H}^2_{\mathrm{BMO}}$.

Furthermore, by our assumptions, $\overline{v} = D_u g \geq -1 + \delta$ and $\overline{v}$ is bounded, then $\overline{v} \in \mathcal{A}$. The inequality (4.4.3) is thus an equality, and the representation (4.4.1) holds true. □

4.4.2 Inf-Convolution of g-expectations

Let $g^1_t(z,u)$ and $g^2_t(z,u)$ be two convex functions such that

$$(g^1 \square g^2)(t,0,0) = \inf_{(\mu,v)\in\mathbb{R}^d\times L^2(\nu_t)} \left\{ g^1_t(\mu,v) + g^2_t(-\mu,-v) \right\} > 0. \tag{4.4.4}$$

The aim of this Section is to compute the optimal risk transfer between two agents using $\mathcal{E}^{g^1}$ and $\mathcal{E}^{g^2}$ as risk measures. The total risk is modeled by a $\mathcal{F}_T$-measurable random variable ξ_T. The optimal risk transfer will be given through the inf-convolution of the risk measures $\mathcal{E}^{g^1}$ and $\mathcal{E}^{g^2}$. We will show that, provided that all the quantities considered behave well enough and are in the right spaces, we can identify the inf-convolution of $\mathcal{E}^{g^1}$ and $\mathcal{E}^{g^2}$ as the solution of a BSDE whose generator is the inf-convolution of g^1 and g^2. Furthermore, we will explicitly construct two $\mathcal{F}_T$-measurable random variables $F_T^{(1)}$ and $F_T^{(2)}$ such that $F_T^{(1)} + F_T^{(2)} = \xi_T$ and

$$(\mathcal{E}^{g^1} \Box \mathcal{E}^{g^2})(\xi_T) = \mathcal{E}^{g^1}(F_T^{(1)}) + \mathcal{E}^{g^2}(F_T^{(2)}).$$

We will say that $(F_T^{(1)}, F_T^{(2)})$ is the optimal risk transfer between the agents 1 and 2.

For this purpose, and for the sake of simplicity, we will assume throughout this section that the solutions to all the considered BSDEs exists. Notice that this is not such a stringent assumption. Indeed, when it comes to the growth condition of Assumption 4.2.1, if we assume that g^1 has quadratic growth in z and u and is strongly convex in (z, u), that is to say that there exists some constant $C > 0$ such that

$$g_t^1(z, u) - \frac{C}{2}\left(|z|^2 + \|u\|^2_{L^2(\nu_t)}\right),$$

is convex, then, because g^2 is convex, it is a classical result that $g^1 \Box g^2$ also has quadratic growth.

Furthermore, we are convinced that as in the classical results by Kobylanski [74] in the continuous case, this growth condition should be enough to obtain existence of maximal and minimal solutions to the corresponding BSDEs.

Remark 4.4.1 *Notice that since our generators are defined on $\Omega \times [0, T] \times \mathbb{R}^d \times L^2(\nu) \cap L^\infty(\nu)$, a quadratic growth condition on u is equivalent to the exponential growth assumed in Assumption 4.2.1.*

Theorem 4.4.2 *Let g^1 and g^2 be two given generators satisfying the assumptions of Theorem 4.4.1. Denote $(\mathcal{E}_t^{1,2}(\xi_T), Z_t, U_t)$ the solution of the BSDE with generator $g^1 \Box g^2$ and terminal condition ξ_T, and let $(\hat{Z}_t^{(1)}, \hat{U}_t^{(1)})$ and $(\hat{Z}_t^{(2)}, \hat{U}_t^{(2)})$ be four predictable processes such that*

$$(g^1 \Box g^2)(t, Z_t, U_t) = g_t^1(\hat{Z}_t^{(1)}, \hat{U}_t^{(1)}) + g_t^2(\hat{Z}_t^{(2)}, \hat{U}_t^{(2)})\ dt \times \mathbb{P} - a.s. \tag{4.4.5}$$

Then

(i) *For any $\mathcal{F}_T$-measurable r.v $F \in L^\infty$,*

$$\mathcal{E}_t^{1,2}(\xi_T) \leq \mathcal{E}_t^{g^1}(\xi_T - F) + \mathcal{E}_t^{g^2}(F),\ \mathbb{P} - a.s.,\ \forall t \in [0, T]. \tag{4.4.6}$$

(ii) *Define*

$$F_T^{(2)} := -\int_0^T g_s^2(\hat{Z}_s^{(2)}, \hat{U}_s^{(2)})ds + \int_0^T \hat{Z}_s^{(2)} dB_s + \int_0^T \int_E \hat{U}_s^{(2)}(x)\widetilde{\mu}(ds, dx),$$

and assume that the BSDEs with generators g^1 and g^2 and terminal conditions $\xi_T - F_T^{(2)}$ and $F_T^{(2)}$ have a solution.

If furthermore $\hat{Z}^{(i)} \in \mathbb{H}^2_{\mathrm{BMO}}$ and $\hat{U}^{(i)} \in \mathbb{J}^2_{\mathrm{BMO}}$, $i = 1, 2$, then

$$(\mathcal{E}^{g^1} \Box \mathcal{E}^{g^2})_t(\xi_T) = \mathcal{E}_t^{1,2}(\xi_T) = \mathcal{E}_t^{g^1}(\xi_T - F_T^{(2)}) + \mathcal{E}_t^{g^2}(F_T^{(2)}). \tag{4.4.7}$$

Proof. $(\hat{Z}^{(2)}, \hat{U}^{(2)})$ is well defined and predictable thanks to Proposition 8.1 in [6]. For $F \in L^\infty$, $\mathcal{E}_t^{g^1}(\xi_T - F) + \mathcal{E}_t^{g^2}(F)$ is solution of

$$\begin{aligned}
d(\mathcal{E}_t^{g^1}(\xi_T - F) + \mathcal{E}_t^{g^2}(F)) &= \left(g_t^1(Z_t^1, U_t^1) + g_t^2(Z_t^2, U_t^2)\right) dt - (Z_t^1 + Z_t^2) dB_t \\
&\quad - \int_E (U_t^1(x) + U_t^2(x)) \widetilde{\mu}(dt, dx) \\
&= \left(g_t^1(Z_t - Z_t^2, U_t - U_t^2) + g_t^2(Z_t^2, U_t^2)\right) dt - Z_t dB_t \\
&\quad - \int_E U_t(x) \widetilde{\mu}(dt, dx)
\end{aligned}$$

and $\mathcal{E}_T^{g^1}(\xi_T - F) + \mathcal{E}_T^{g^2}(F) = \xi_T$.

Since g^1 satisfies the convexity Assumption 4.2.6, then the function $(z,u) \mapsto g_t^1(z - Z_t^2, u - U_t^2) - g_t^2(Z_t^2, U_t^2)$ is also convex, and we can apply the comparison theorem which directly implies inequality (4.4.6).

Assume now that $\hat{Z}^{(i)} \in \mathbb{H}^2_{\mathrm{BMO}}$ and $\hat{U}^{(i)} \in \mathbb{J}^2_{\mathrm{BMO}}$, $i = 1, 2$, and define $F_t^{(i)}$ by by the following forward equations

$$\begin{aligned}
F_t^{(1)} &:= \mathcal{E}_0^{1,2}(\xi_T) - \int_0^t g_s^1(\hat{Z}_s^{(1)}, \hat{U}_s^{(1)}) ds + \int_0^t \hat{Z}_s^{(1)} dB_s + \int_0^t \int_E \hat{U}_s^{(1)}(x) \widetilde{\mu}(ds, dx), \\
F_t^{(2)} &:= - \int_0^t g_s^2(\hat{Z}_s^{(2)}, \hat{U}_s^{(2)}) ds + \int_0^t \hat{Z}_s^{(2)} dB_s + \int_0^t \int_E \hat{U}_s^{(2)}(x) \widetilde{\mu}(ds, dx).
\end{aligned}$$

Then we have

$$F_t^{(i)} = F_T^{(i)} + \int_t^T g_s^i(\hat{Z}_s^{(i)}, \hat{U}_s^{(i)}) ds - \int_t^T \hat{Z}_s^{(i)} dB_s - \int_t^T \int_E \hat{U}_s^{(i)}(x) \widetilde{\mu}(ds, dx),$$

and by uniqueness, $F_t^{(i)} = \mathcal{E}_t^{g^i}(F_T^{(i)})$. Moreover, by definition, we have $F_T^{(1)} + F_T^{(2)} = \xi_T$.

Since

$$(g^1 \square g^2)(t, Z_t, U_t) = g_t^1(\hat{Z}_t^{(1)}, \hat{U}_t^{(1)}) + g_t^2(\hat{Z}_t^{(2)}, \hat{U}_t^{(2)}) \; dt \times \mathbb{P} - \text{a.s.}, \tag{4.4.8}$$

we have the equality

$$\mathcal{E}_t^{1,2}(\xi_T) = \mathcal{E}_t^{g^1}(F_T^{(1)}) + \mathcal{E}_t^{g^2}(F_T^{(2)}).$$

We can conclude that the processes $\mathcal{E}_t^{g^1}(F_T^{(1)}) + \mathcal{E}_t^{g^2}(F_T^{(2)})$ and $\mathcal{E}_t^{1,2}(\xi_T)$ are solution of the BSDE with coefficients $(g^1 \square g^2, \xi_T)$, by uniqueness we have that equality (4.4.7) holds. $\square$

Remark 4.4.2 *As already mentionned in Chapter 2, the optimal structure $(F_T^{(1)}, F_T^{(2)})$ is defined up to a constant, more precisely, $(F_T^{(1)} + m, F_T^{(2)} - m)$ with $m \in \mathbb{R}$ is again an optimal structure. Indeed, the cash-additivity property implies that*

$$\mathcal{E}_t^{g^1}(F_T^{(1)} + m) + \mathcal{E}_t^{g^2}(F_T^{(2)} - m) = \mathcal{E}_t^{g^1}(F_T^{(1)}) + \mathcal{E}_t^{g^2}(F_T^{(2)}).$$

4.4.3 Examples of inf-convolution

In this Section, we use the previous result on the inf-convolution of g-expectations to treat several particular examples.

4.4.3.1 Quadratic and Quadratic

We first study the inf-convolution of two dynamic entropic risk-measure. This example is treated by Barrieu and El Karoui [6] by a direct method, they find that the optimal risk transfer is *proportional* in the sense that there exists $a \in (0,1)$ such that

$$(\mathcal{E}^{g^1} \Box \mathcal{E}^{g^2})(\xi_T) = \mathcal{E}^{g^1}(a\xi_T) + \mathcal{E}^{g^2}((1-a)\xi_T).$$

We retrieve here this result using Theorem 4.4.2. For this, we first need to study the inf-convolution of the two corresponding generators g^i, $i = 1, 2$

$$g_t^i(z,u) := \frac{1}{2\gamma_i}|z|^2 + \gamma_i \int_E \left(e^{\frac{u(x)}{\gamma_i}} - 1 - \frac{u(x)}{\gamma_i} \right) \nu_t(dx), \tag{4.4.9}$$

where $(\gamma_1, \gamma_2) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$.

Lemma 4.4.1 *Let g^1 and g^2 be the two convex generators defined in equation (4.4.9). For any bounded $\mathcal{F}_T$-measurable random variable ξ_T, we have,*

$$(\mathcal{E}^{g^1} \Box \mathcal{E}^{g^2})(\xi_T) = \mathcal{E}^{g^1}\left(\frac{\gamma_1}{\gamma_1+\gamma_2}\xi_T \right) + \mathcal{E}^{g^2}\left(\frac{\gamma_2}{\gamma_1+\gamma_2}\xi_T \right).$$

Proof.

We can calculate

$$\begin{aligned}
&(g^1 \Box g^2)(t,z,u) \\
&= \inf_v \left\{ \frac{1}{2\gamma_1}|v|^2 + \frac{1}{2\gamma_2}|z-v|^2 \right\} \\
&+ \inf_w \left\{ \gamma_1 \int_E \left(e^{\frac{w(x)}{\gamma_1}} - 1 - \frac{w(x)}{\gamma_1} \right) \nu_t(dx) + \gamma_2 \int_E \left(e^{\frac{u(x)-w(x)}{\gamma_2}} - 1 - \frac{u(x)-w(x)}{\gamma_2} \right) \nu_t(dx) \right\}.
\end{aligned}$$

The first infimum above is easy to calculate and is attained for

$$v^* := \frac{\gamma_1}{\gamma_1+\gamma_2} z.$$

For the second one, we postulate similarly that it should be attained for

$$w^* := \frac{\gamma_1}{\gamma_1+\gamma_2} u.$$

In order to verify this result, it is sufficient to prove that for all $(x,y) \in \mathbb{R}^2$

$$(\gamma_1+\gamma_2)\left(e^{\frac{x+y}{\gamma_1+\gamma_2}} - 1 - \frac{x+y}{\gamma_1+\gamma_2} \right) \leq \gamma_1 \left(e^{\frac{x}{\gamma_1}} - 1 - \frac{x}{\gamma_1} \right) + \gamma_2 \left(e^{\frac{y}{\gamma_2}} - 1 - \frac{y}{\gamma_2} \right). \tag{4.4.10}$$

Set $\lambda := \gamma_1/(\gamma_1+\gamma_2)$, $a := x/\gamma_1$, $b := y/\gamma_2$ and $h(x) := e^x - 1 - x$, this is equivalent to

$$h(\lambda a + (1-\lambda)b) \leq \lambda h(a) + (1-\lambda)h(b),$$

which is clear by convexity of the function h.

Therefore, we finally obtain

$$(g^1 \square g^2)(t,z,u) = \frac{1}{2(\gamma_1+\gamma_2)}|z|^2 + (\gamma_1+\gamma_2)\int_E \left(e^{\frac{u(x)}{\gamma_1+\gamma_2}} - 1 - \frac{u(x)}{\gamma_1+\gamma_2} \right) \nu_t(dx).$$

Using the notations of Theorem 4.4.2, we can compute the quantity $F_T^{(2)}$, giving the optimal risk transfer

$$\begin{aligned} F_T^{(2)} &= \frac{\gamma_2}{\gamma_1+\gamma_2}\Big(-\int_0^T \left[\frac{1}{2(\gamma_1+\gamma_2)}|Z_t|^2 + (\gamma_1+\gamma_2)\int_E \left(e^{\frac{U_t(x)}{\gamma_1+\gamma_2}} - 1 - \frac{U_t(x)}{\gamma_1+\gamma_2} \right)\nu_t(dx) \right] dt \\ &\quad + \int_0^T Z_t dB_t + \int_0^T \int_E U_t(x)\widetilde{\mu}(dt,dx) \Big), \\ &= \frac{\gamma_2}{\gamma_1+\gamma_2}\left(\xi_T - \mathcal{E}_0^{1,2}(\xi_T) \right). \end{aligned}$$

We calculate similarly $F_T^{(1)}$ and obtain

$$F_T^{(1)} = \frac{\gamma_1}{\gamma_1+\gamma_2}\xi_T + \frac{\gamma_2}{\gamma_1+\gamma_2}\mathcal{E}_0^{1,2}(\xi_T).$$

Now using Remark 4.4.2 with $m = \frac{\gamma_2}{\gamma_1+\gamma_2}\mathcal{E}_0^{1,2}(\xi_T)$, we obtain that the proportional structure $\left(\frac{\gamma_1}{\gamma_1+\gamma_2}\xi_T, \frac{\gamma_2}{\gamma_1+\gamma_2}\xi_T\right)$ is optimal.

□

4.4.3.2 Linear and Quadratic

Here, we assume that $d = 1$. We study the inf-convolution of a dynamic entropic risk-measure with a linear one corresponding to a linear BSDE. In this case, we want to calculate the inf-convolution of the two corresponding generators g^1 and g^2 given by

$$g_t^1(z,u) := \frac{1}{2\gamma}|z|^2 + \gamma \int_E \left(e^{\frac{u(x)}{\gamma}} - 1 - \frac{u(x)}{\gamma} \right)\nu_t(dx),$$

and

$$g_t^2(z,u) := \alpha z + \beta \int_E (1 \wedge |x|)u(x)\nu_t(dx),$$

where $(\gamma,\alpha,\beta) \in \mathbb{R}_+^* \times \mathbb{R} \times [-1+\delta, +\infty)$ for some $\delta > 0$.

Lemma 4.4.2 *Let g^1 and g^2 be defined in the two previous equations. We have, for any bounded $\mathcal{F}_T$-measurable random variable ξ_T,*

$$(\mathcal{E}^{g^1} \square \mathcal{E}^{g^2})(\xi_T) = \mathcal{E}^{g^1}(F_T^{(1)}) + \mathcal{E}^{g^2}(F_T^{(2)}),$$

where

$$F_T^{(2)} = \xi_T + \frac{1}{2}\alpha^2\gamma T + \gamma \int_0^T \int_E (\beta(1\wedge|x|) - \ln(1+\beta(1\wedge|x|)))\nu_t(dx)dt - \alpha\gamma B_T$$
$$- \gamma \int_0^T \int_E \ln(1+\beta(1\wedge|x|))\widetilde{\mu}(dt,dx),$$

and $F_T^{(1)} = \xi_T - F_T^{(2)}$.

Remark 4.4.3 *Notice that $F_T^{(2)}$ has no longer the linear form with respect to ξ_T obtained in the previous example. Now, the agent 2 receives the value ξ_T perturbed by a random value only depending on the data contained in the filtration, i.e the Brownian motion B and the random measures μ and (ν_t).*

Proof. We start by computing the inf-convolution in (z,u) of the generators:

$$(g^1\Box g^2)(t,z,u) = \inf_v \left\{ \frac{1}{2\gamma}|v|^2 + \alpha(z-v) \right\}$$
$$+ \inf_w \left\{ \gamma \int_E \left(e^{\frac{w(x)}{\gamma}} - 1 - \frac{w(x)}{\gamma} \right) \nu_t(dx) + \beta \int_E (1\wedge|x|)\,(u(x)-w(x))\,\nu_t(dx) \right\}.$$

The first infimum above is easy to calculate and is attained for

$$v^* := \alpha\gamma.$$

Similarly, it is easy to show that the function $w \to \gamma\left(e^{\frac{w}{\gamma}} - 1 - \frac{w}{\gamma}\right) + \beta(1\wedge|x|)\,(u(x)-w)$ attains its minimum at $w^* := \gamma\ln(1+\beta(1\wedge|x|))$. Therefore, we finally obtain

$$(g^1\Box g^2)(t,z,u) = \alpha z - \frac{\alpha^2\gamma}{2} + \int_E \beta(1\wedge|x|)u(x)\nu_t(dx)$$
$$+ \int_E \gamma\left[\beta(1\wedge|x|) - (1+\beta(1\wedge|x|))\ln(1+\beta(1\wedge|x|))\right]\nu_t(dx).$$

Notice that all the quantities appearing in $g^1\Box g^2$ are finite. Indeed, we first have for any $u \in L^2(\nu)\cap L^\infty(\nu)$

$$|\beta(1\wedge|x|)u(x)| \le \beta^2(1\wedge|x|^2) + (u(x))^2, \tag{4.4.11}$$

and this quantity is therefore ν_t-integrable for all t. Then, since $\beta \ge -1+\delta$, it is also clear that for some constant $C_\delta > 0$

$$0 \le (1+\beta(1\wedge|x|))\ln(1+\beta(1\wedge|x|)) - \beta(1\wedge|x|) \le C_\delta(1\wedge|x|^2), \tag{4.4.12}$$

and thus the second integral is also finite.

We can now compute for instance the quantity $F_T^{(2)}$, which is given after some easy calculations by

$$F_T^{(2)} = -\int_0^T \left(\alpha Z_t + \beta \int_E (1\wedge|x|)U_t(x)\nu_t(dx) \right) dt + \int_0^T Z_t dB_t + \int_0^T \int_E U_t(x)\widetilde{\mu}(dt,dx)$$
$$+ \alpha^2\gamma T + \gamma\beta \int_0^T \int_E (1\wedge|x|)\ln(1+\beta(1\wedge|x|))\nu_t(dx)dt - \alpha\gamma B_T$$
$$- \gamma \int_0^T \int_E \ln(1+\beta(1\wedge|x|))\widetilde{\mu}(dt,dx).$$

Recall that (Z,U) is part of the solution of the BSDE with generator $g^1\square g^2$ and terminal condition ξ_T.

Similarly, we can compute the value $F_T^{(1)}$, and using the fact that $F_T^{(1)}+F_T^{(2)}=\xi_T$, we obtain

$$\begin{aligned}\xi_T =&\frac{1}{2}\alpha^2\gamma T+\gamma\int_0^T\int_E\left((1+\beta(1\wedge|x|))\ln(1+\beta(1\wedge|x|))-\beta(1\wedge|x|)\right)\nu_t(dx)dt\\&-\int_0^T g_t^2(Z_t,U_t)dt+\int_0^T Z_t dB_t+\int_0^T\int_E U_t(x)\widetilde{\mu}(dt,dx).\end{aligned}$$

And finally, we can conclude that the optimal risk transfer takes the form

$$\begin{aligned}F_T^{(2)} =&\xi_T+\frac{1}{2}\alpha^2\gamma T+\gamma\int_0^T\int_E(\beta(1\wedge|x|)-\ln(1+\beta(1\wedge|x|)))\nu_t(dx)dt-\alpha\gamma B_T\\&-\gamma\int_0^T\int_E\ln(1+\beta(1\wedge|x|))\widetilde{\mu}(dt,dx).\end{aligned}$$

□

4.4.3.3 Absolute value and Quadratic

Here, we study again the example of the entropic risk measure, i.e. we keep the generator g^1 given by

$$g_t^1(z,u):=\frac{1}{2\gamma}|z|^2+\gamma\int_E\left(e^{\frac{u(x)}{\gamma}}-1-\frac{u(x)}{\gamma}\right)\nu_t(dx),$$

where $\gamma>0$. The generator g^2 correspond to the second example of risk measure given in Examples 4.3.1, i.e.

$$g_t^2(z,u):=\eta|z|+\eta\int_E(1\wedge|x|)u^+(x)\nu_t(dx)-C_1\int_E(1\wedge|x|)u^-(x)\nu_t(dx),$$

where $\eta>0$ and $-1+\delta\le C_1\le 0$ for some $\delta>0$.

Then

$$\begin{aligned}&(g^1\square g^2)(t,z,u)=\inf_v\left\{\frac{1}{2\gamma}|v|^2+\eta|z-v|\right\}\\&+\inf_w\Big\{\gamma\int_E\left(e^{\frac{w(x)}{\gamma}}-1-\frac{w(x)}{\gamma}\right)\nu_t(dx)+\eta\int_E(1\wedge|x|)\left(u(x)-w(x)^+\right)\nu_t(dx)\\&-C_1\int_E(1\wedge|x|)\left(u(x)-w(x)^-\right)\nu_t(dx)\Big\}.\end{aligned}$$

These infima are easy to calculate and are attained at (v^*,w^*) defined by

$$\begin{aligned}&v^*=\mu\gamma\mathbf{1}_{\{z>\mu\gamma\}}-\mu\gamma\mathbf{1}_{\{z<-\mu\gamma\}}+z\mathbf{1}_{\{-\mu\gamma\le z\le\mu\gamma\}}\\&w^*(x)=\psi(\eta,x)\mathbf{1}_{\{u(x)>\psi(\eta,x)\}}+\psi(C_1,x)\mathbf{1}_{\{u(x)\le\psi(C_1,x)\}}+u(x)\mathbf{1}_{\{\psi(C_1,x)<u(x)\le\psi(\eta,x)\}}.\end{aligned}$$

where $\psi(r,x):=\gamma\ln(1+r(1\wedge|x|))$. We finally obtain

$$(g^1\square g^2)(t,z,u)=A(z)+B(u),$$

where

$$A(z) := \left(\mu z - \frac{1}{2}\mu^2\gamma\right)\mathbf{1}_{\{z>\mu\gamma\}} - \left(\mu z + \frac{1}{2}\mu^2\gamma\right)\mathbf{1}_{\{z<-\mu\gamma\}} + \frac{1}{2\gamma}z^2\mathbf{1}_{\{-\mu\gamma\le z\le\mu\gamma\}}$$

$$B(u) := \int_E \left[C_1(1\wedge|x|)(\gamma+u(x)) - \psi(C_1,x)(1+C_1(1\wedge|x|))\mathbf{1}_{\{u(x)\le\psi(C_1,x)\}}\right]\nu_t(dx)$$

$$+\int_E \left[\eta(1\wedge|x|)(\gamma+u(x)) - \psi(\eta,x)(1+\eta(1\wedge|x|))\mathbf{1}_{\{u(x)>\psi(\eta,x)\}}\right]\nu_t(dx)$$

$$+\int_E \left[\gamma\left(e^{\frac{u(x)}{\gamma}} - 1 - \frac{u(x)}{\gamma}\right)\mathbf{1}_{\{\psi(C_1,x)<u(x)\le\psi(\eta,x)\}}\right]\nu_t(dx),$$

that is to say this inf-convolution is equal to a linear function outside a given set, and is quadratic inside the set (in z and u). Notice that all the quantities appearing in $B(u)$ are finite, thanks to the inequalities (4.4.11) and (4.4.12) given in the previous example.

Now to have the optimal risk transfer, we calculate $F_T^{(1)}$ and $F_T^{(2)}$ using their definition. We obtain for example for $F_T^{(2)}$

$$\begin{aligned}
F_T^{(2)} = &\int_0^T (\mu^2\gamma - \mu Z_t)\mathbf{1}_{\{|Z_t|>\mu\gamma\}}dt \\
&- \int_0^T \int_E \Big[\eta(1\wedge|x|)\left(\psi(\eta,x)\mathbf{1}_{\{U_t(x)>\psi(\eta,x)\}} + U_t(x)\mathbf{1}_{\{0<U_t(x)<\psi(\eta,x)\}}\right) \\
&+ C_1(1\wedge|x|)\left(\psi(C_1,x)\mathbf{1}_{\{U_t(x)<\psi(C_1,x)\}} + U_t(x)\mathbf{1}_{\{\psi(C_1,x)<U_t(x)<0\}}\right)\Big]\nu_t(dx)dt \\
&+ \int_0^T \int_E \Big(U_t(x) - \psi(\eta,x)\mathbf{1}_{\{U_t(x)>\psi(\eta,x)\}} + \psi(C_1,x)\mathbf{1}_{\{U_t(x)\le\psi(C_1,x)\}} \\
&+ U_t(x)\mathbf{1}_{\{\psi(C_1,x)<U_t(x)\le\psi(\eta,x)\}}\Big)\widetilde{\mu}(dt,dx),
\end{aligned}$$

where (Z,U) is part of the solution of the BSDE with generator $g^1\Box g^2$ and terminal condition ξ_T.

And we have of course $F_T^{(1)} = \xi_T - F_T^{(2)}$.

Remark 4.4.4 *In this example, the optimal transfer has again no longer the proportional form found in example 1. Moreover, this time we are only able to calculate $F_T^{(1)}$ and $F_T^{(2)}$ in function of (Z,U), where, as defined before, $(\mathcal{E}_t^{1,2}(\xi_T), Z_t, U_t)$ denotes the solution of the BSDE with generator $g^1\Box g^2$ and terminal condition ξ_T.*

Part II

Model Uncertainty and 2BSDEs with Jumps

Chapter 5

Second Order Backward SDEs with Jumps

Contents

5.1 Introduction **127**

5.2 Issues related to aggregation **129**

5.2.1 The stochastic basis 129

5.2.2 The main problem 130

5.2.3 Characterization by martingale problems 131

5.2.4 Notations and definitions 133

5.3 Preliminaries on 2BSDEs **142**

5.3.1 The non-linear Generator 142

5.3.2 The Spaces and Norms 145

5.3.3 Formulation 146

5.3.4 Connection with standard backward SDEs with jumps 147

5.3.5 Connection with G-expectations and G-Lévy processes 148

5.4 Uniqueness result **150**

5.4.1 Representation of the solution 150

5.4.2 A priori estimates and uniqueness of the solution 154

5.1 Introduction

Recall from the previous chapter that (Y, Z, U) is a solution of the BSDE with jumps, with generator F and terminal condition ξ if for all $t \in [0, T]$, we have $\mathbb{P} - a.s.$

$$Y_t = \xi + \int_t^T F(s, Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s - \int_t^T \int_{\mathbb{R}^d \backslash \{0\}} U_s(x)(\mu - \nu)(ds, dx), \tag{5.1.1}$$

where μ is an integer valued random measure with compensator ν.

Our aim in this chapter is to generalize (5.1.1) to the second order, as introduced recently by Soner, Touzi and Zhang [112]. Their key idea in the definition of the second order BSDE is that the equation defining the solution has to hold $\mathbb{P}$-almost surely, for every $\mathbb{P}$ in a class of non dominated probability measures. Furthermore, they prove a uniqueness result using a representation result of the 2BSDEs as essential supremum of standard BSDEs. We will extend these definitions and properties to the case with jumps. For the sake of clarity, we will now briefly outline the main differences and difficulties due to the presence of jumps.

(i) First, we will introduce an equation similar to (5.1.1), for which the driver F depends also on the quadratic variation of B^c on the one hand, and on the compensator of the jump measure of B^d on the other hand, where B^c and B^d denote respectively the continuous and the purely discontinuous parts of B. Recall that the generator of standard BSDEs does not depend on the quadratic variation of the canonical process, and that in the continuous second order case it only depends on the quadratic variation of the Brownian martingale part. In our case, since we have an additional jump martingale term in the equation, we also have an additional dependence in the driver.

(ii) A second major difference with (5.1.1) in the second order case, is, as we recalled earlier, that the BSDE has to hold $\mathbb{P}$-almost surely for every probability measure $\mathbb{P}$ lying in a wide family of probability measures. Under each $\mathbb{P}$, B^c and B^d have a prescribed quadratic variation and jump measure compensator. This is why we can intuitively understand the second order BSDE with jumps (2BSDEJ in the sequel) (5.3.3) as a backward SDE with model uncertainty, where the uncertainty affects both the quadratic variation and the jump measure of the process driving the equation.

(iii) The last major difference with (5.1.1) in the second order case is the presence of an additional increasing process K in the equation. To have an intuition for K, one has to have in mind the representation (5.4.1) that we prove in Theorem 5.4.1, stating that the Y part of a solution of a 2BSDEJ is an essential supremum of solutions of standard backward SDEs with jumps. The process K maintains Y above any standard solution $y^{\mathbb{P}}$ of a BSDE with jumps, with given quadratic variation and jump measure under $\mathbb{P}$. The process K is then formally analogous to the increasing process appearing in reflected BSDEs (as defined in [46] for instance).

There are many other possible approaches in the literature to handle volatility and / or jump measure uncertainty in stochastic models ([3], [39], [37], [14]). Among them, Peng [101] introduced a notion of Brownian motion with uncertain variance structure, called G-Brownian motion. This process is defined without making reference to a given probability measure. It refers instead to the G-Gaussian law, defined by a partial differential equation (see [102] for a detailed exposition and references).

Finally, recall that Pardoux and Peng [97] proved that if the randomness in g and ξ is induced by the current value of a state process defined by a forward stochastic differential equation, the solution to the BSDE (4.2.1) could be linked to the solution of a semilinear PDE by means of a generalized Feynman-Kac formula. Similarly, Soner, Touzi and Zhang [112] showed that 2BSDEs generalized the point of view of Pardoux and Peng, in the sense that they are connected to the larger class of fully non linear PDEs. In this context, the 2BSDEJs are the natural candidates for a probabilistic solution of fully non linear integro-differential equations. Note that in a recent preprint, Bion-Nadal [13], by taking a dual point of view, connects the probabilistic construction of time consistent convex dynamic procedures in a non dominated setting, to a class of non local and fully non linear second order partial differential equations.

The rest of this chapter is organised as follows. In Section 5.2, we introduce the set of probability measures on the Skorohod space $\mathbb{D}$ that we will work with. Using the notion of martingale problems on $\mathbb{D}$, we construct probability measures under which the canonical process has given characteristics. Then we prove an aggregation result under this family. In Section 5.3, we define the notion of 2BSDEJs and show how it is linked to standard BSDEs with jumps. Section 5.4 is devoted to our uniqueness result and some a priori estimates.

5.2 Issues related to aggregation

5.2.1 The stochastic basis

Let $\Omega := \mathbb{D}([0,T],\mathbb{R}^d)$ be the space of càdlàg paths defined on $[0,T]$ with values in R^d and such that $w(0)=0$, equipped with the Skorokhod topology, so that it is a complete, separable metric space (see [12] for instance).

We denote B the canonical process, $\mathbb{F} := \{\mathcal{F}_t\}_{0\leq t\leq T}$ the filtration generated by B, $\mathbb{F}^+ := \left\{\mathcal{F}_t^+\right\}_{0\leq t\leq T}$ the right limit of $\mathbb{F}$ and for any $\mathbb{P}$, $\mathcal{F}_t^{\mathbb{P}} := \mathcal{F}_t^+ \vee \mathcal{N}^{\mathbb{P}}(\mathcal{F}_t^+)$ where

$$\mathcal{N}^{\mathbb{P}}(\mathcal{G}) := \left\{E\in\Omega,\ \text{there exists } \widetilde{E}\in\mathcal{G} \text{ such that } E\subset\widetilde{E} \text{ and } \mathbb{P}(\widetilde{E})=0\right\}.$$

As usual, for any filtration $\mathbb{G}$ and any probability measure $\mathbb{P}$, $\overline{\mathcal{G}}^{\mathbb{P}}$ will denote the corresponding completed filtration. We then define as in [112] a local martingale measure $\mathbb{P}$ as a probability measure such that B is a $\mathbb{P}$-local martingale. Since we are working in the Skorohod space, we can then define the continuous martingale part of B, noted B^c, and its purely discontinuous part, noted B^d, both being local martingales under each local martingale measures (see [65]). We then associate to the jumps of B a counting measure μ_{B^d}, which is a random measure on $\mathcal{B}(\mathbb{R}^+)\times E$ (where $E := \mathbb{R}^r\backslash\{0\}$ for some $r\in\mathbb{N}^*$), defined pathwise by

$$\mu_{B^d}([0,t],A) := \sum_{0<s\leq t}\mathbf{1}_{\{\Delta B_s^d\in A\}},\ \forall t\geq 0,\ \forall A\subset E. \tag{5.2.1}$$

We also denote by $\nu_t^{\mathbb{P}}(dt,dx)$ the compensator of $\mu_{B^d}(dt,dx)$, which is a predictable random measure, under $\mathbb{P}$ and by $\widetilde{\mu}_{B^d}^{\mathbb{P}}(dt,dx)$ the corresponding compensated measure.

We then denote $\overline{\mathcal{P}}_W$ the set of all local martingale measures $\mathbb{P}$ such that $\mathbb{P}$-a.s.

(i) The quadratic variation of B^c is absolutely continuous with respect to the Lebesgue measure dt and its density takes values in $\mathbb{S}_d^{>0}$.

(ii) The compensator $\nu_t^{\mathbb{P}}(dt,dx)$ under $\mathbb{P}$ is absolutely continuous with respect to the Lebesgue measure dt.

In this discontinuous setting, we will say that a probability measure $\mathbb{P}\in\overline{\mathcal{P}}_W$ satisfies the martingale representation property if for any $(\overline{\mathbb{F}}^{\mathbb{P}},\mathbb{P})$-local martingale M, there exists a unique $\overline{\mathbb{F}}^{\mathbb{P}}$-predictable processes H and a unique $\overline{\mathbb{F}}^{\mathbb{P}}$-predictable function U such that $(H,U)\in\mathbb{H}^2_{loc}(\mathbb{P})\times\mathbb{J}^2_{loc}(\mathbb{P})$ (those spaces are defined later) and

$$M_t = M_0 + \int_0^t H_s dB_s^c + \int_0^t\int_E U_s(x)\widetilde{\mu}_{B^d}^{\mathbb{P}}(ds,dx).$$

We now follow [111] and introduce their so-called universal filtration. For this we let $\mathcal{P}$ be a given subset of $\overline{\mathcal{P}}_W$, we define

Definition 5.2.1 (i) *A property is said to hold $\mathcal{P}$-quasi-surely ($\mathcal{P}$-q.s. for short), if it holds $\mathbb{P}-a.s.$ for all $\mathbb{P} \in \mathcal{P}$.*

(ii) *We call $\mathcal{P}$-polar sets the elements of $\mathcal{N}_{\mathcal{P}} := \cap_{\mathbb{P}\in\mathcal{P}}\mathcal{N}^{\mathbb{P}}(\mathcal{F}_\infty)$.*

Then, we define as in [111]

$$\widehat{\mathbb{F}}^{\mathcal{P}} := \left\{\widehat{\mathcal{F}}_t^{\mathcal{P}}\right\}_{t\geq 0} \text{ where } \widehat{\mathcal{F}}_t^{\mathcal{P}} := \bigcap_{\mathbb{P}\in\mathcal{P}} \left(\mathcal{F}_t^{\mathbb{P}} \vee \mathcal{N}_{\mathcal{P}}\right).$$

Finally, we let $\mathcal{T}$ and $\widehat{\mathcal{T}}^{\mathcal{P}}$ the sets of all $\mathbb{F}$ and $\widehat{\mathbb{F}}^{\mathcal{P}}$ stopping times, and we recall that thanks to Lemma 2.4 in [111] we do not have to worry about the universal filtration not being complete under each $\mathbb{P} \in \mathcal{P}$.

5.2.2 The main problem

A crucial issue in the definition of the 2BSDEs in [112] is the aggregation of the quadratic variation of the canonical process B under a wide family of probability measures.

Let $\mathcal{P} \subset \overline{\mathcal{P}}_W$ be a set of non necessarily dominated probability measures and let $\{X^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$ be a family of random variables indexed by $\mathcal{P}$. One can think for example of the stochastic integrals $X_t^{\mathbb{P}} := {}^{(\mathbb{P})}\int_0^t H_s dB_s$, where $\{H_t, t \geq 0\}$ is a predictable process.

Definition 5.2.2 *An aggregator of the family $\{X^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}\}$ is a r.v. $\hat{X}$ such that*

$$\hat{X} = X^{\mathbb{P}},\ \mathbb{P} - a.s, \text{ for every } \mathbb{P} \in \mathcal{P}.$$

Bichteler [11], Karandikar [69], or more recently Nutz [95] all showed in different contexts, and under different assumptions, that it is possible to find an aggregator for the Itô stochastic integrals ${}^{(\mathbb{P})}\!\int_0^t H_s dB_s$. A direct consequence of this result is the possibility to aggregate the quadratic variation process $\{[B, B]_t, t \geq 0\}$. Indeed, using Itô's formula, we can write

$$[B, B]_t = B_t B_t^T - 2\int_0^t B_{s^-} dB_s^T$$

and the aggregation of the stochastic integrals automatically yields the aggregation of the bracket $\{[B, B]_t, t \geq 0\}$.

This also allows us to give a pathwise definition of the process $\widehat{a}$, which is an aggregator for the density of the quadratic variation of the continuous part of B, by

$$\widehat{a}_t := \limsup_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\left(\langle B^c\rangle_t - \langle B^c\rangle_{t-\varepsilon}\right),$$

Soner, Touzi and Zhang [111], motivated by the study of stochastic target problems under volatility uncertainty, obtained an aggregation result for a family of probability measures corresponding to the laws of some continuous martingales on the canonical space $\Omega = \mathcal{C}(\mathbb{R}^+, \mathbb{R}^d)$, under a *separability* assumption on the quadratic variations (see their definition 4.8) and under an additional *consistency* condition (which is usually only necessary) for the family to aggregate.

To define correctly the notion of second order backward SDEs with jumps, we need to aggregate not only the quadratic variation $[B, B]$ of the canonical process, but also its compensated jump measure. However, this predictable compensator is usually obtained thanks to the Doob-Meyer decomposition of the submartingale $[B, B]$. It is therefore obvious that this compensator depends explicitely on the underlying probability measure, and it is not clear at all whether an aggregator always exists or not. This is a first main difference with the continuous case. In order to solve this problem, we follow the spirit of [111] and restrict our set of probability measures (by adding an analogous separability condition for jump measures) so as to generalize some of their results to the case of processes with jumps.

After these first notations, in the following subsection, in order to construct a probability measure under which the canonical process has a given quadratic variation and a given jump measure, we will use the notion of martingale problem for semimartingales with general characteristics, as defined in the book by Jacod and Shiryaev [65] to which we refer.

5.2.3 Characterization by martingale problems

In this section, we extend the connection between diffusion processes and probability measures established in [111] thanks to weak solutions of SDEs, to our general jump case with the more general notion of martingale problems.

Let $\mathcal{N}$ be the set of $\mathbb{F}$-predictable random measures ν on $\mathcal{B}(E)$ satisfying

$$\int_0^t \int_E (1 \wedge |x|^2)\nu_s(dx)ds < +\infty \text{ and } \int_0^t \int_{|x|>1} x\nu_s(dx)ds < +\infty, \ \forall \omega \in \Omega, \tag{5.2.2}$$

and let $\mathcal{D}$ be the set of $\mathbb{F}$-predictable processes α taking values in $\mathbb{S}_d^{>0}$ with

$$\int_0^T |\alpha_t| dt < +\infty, \text{ for every } \omega \in \Omega.$$

We define a martingale problem as follows

Definition 5.2.3 *For $\mathbb{F}$-stopping times τ_1 and τ_2, for $(\alpha, \nu) \in \mathcal{D} \times \mathcal{N}$ and for a probability measure $\mathbb{P}_1$ on $\mathcal{F}_{\tau_1}$, we say that $\mathbb{P}$ is a solution of the martingale problem $(\mathbb{P}_1, \tau_1, \tau_2, \alpha, \nu)$ if*

(i) $\mathbb{P} = \mathbb{P}_1$ *on* $\mathcal{F}_{\tau_1}$.

(ii) *The canonical process B on $[\tau_1, \tau_2]$ is a semimartingale under $\mathbb{P}$ with characteristics*

$$\left(-\int_{\tau_1}^{\cdot} \int_E x \mathbf{1}_{|x|>1} \nu_s(dx)ds, \int_{\tau_1}^{\cdot} \alpha_s ds, \ \nu_s(dx)ds \right).$$

Remark 5.2.1 *We refer to Theorem II.2.21 in [65] for the fact that $\mathbb{P}$ is a solution of the martingale problem $(\mathbb{P}_1, \tau_1, \tau_2, \alpha, \nu)$ if and only if the following properties hold*

(i) $\mathbb{P} = \mathbb{P}_1$ *on* $\mathcal{F}_{\tau_1}$.

(ii) *The processes M, J and L defined below are $\mathbb{P}$-local martingales on $[\tau_1, \tau_2]$*

$$M_t := B_t - \sum_{\tau_1 \le s \le t} \mathbf{1}_{|\Delta B_s| > 1} \Delta B_s + \int_{\tau_1}^t \int_E x \mathbf{1}_{|x|>1} \nu_s(dx) ds, \ \tau_1 \le t \le \tau_2$$

$$J_t := M_t^2 - \int_{\tau_1}^t \alpha_s ds - \int_{\tau_1}^t \int_E x^2 \nu_s(dx) ds, \ \tau_1 \le t \le \tau_2$$

$$Q_t := \int_{\tau_1}^t \int_E g(x) \mu_B(ds, dx) - \int_{\tau_1}^t \int_E g(x) \nu_s(dx) ds, \ \tau_1 \le t \le \tau_2, \ \forall g \in \mathcal{C}^+(\mathbb{R}^d),$$

where $\mathcal{C}^+(\mathbb{R}^d)$ is discriminating family of bounded Borel functions (see Remark II.2.20 in [65] for more details).

We say that the martingale problem associated to (α, ν) has a unique solution if, for every stopping times τ_1, τ_2 and for every probability measure $\mathbb{P}_1$, the martingale problem $(\mathbb{P}_1, \tau_1, \tau_2, \alpha, \nu)$ has a unique solution.

Let now $\overline{\mathcal{A}}_W$ be the set of $(\alpha, \nu) \in \mathcal{D} \times \mathcal{N}$, such that there exists a solution to the martingale problem $(\mathbb{P}^0, 0, +\infty, \alpha, \nu)$, where $\mathbb{P}^0$ is such that $\mathbb{P}^0(B_0 = 0) = 1$.

We also denote by $\mathcal{A}_W$ the set of $(\alpha, \nu) \in \overline{\mathcal{A}}_W$ such that there exists a unique solution to the martingale problem $(\mathbb{P}_1, 0, +\infty, \alpha, \nu)$, where $\mathbb{P}_1$ is such that $\mathbb{P}_1(B_0 = 0) = 1$. Denote $\mathbb{P}_\nu^\alpha$ this unique solution. Finally we set

$$\mathcal{P}_W := \{\mathbb{P}_\nu^\alpha, \ (\alpha, \nu) \in \mathcal{A}_W\}.$$

Remark 5.2.2 *We take here as an initial condition that $B_0 = 0$. This does not generate a loss of generality, since at the end of the day, the probability measures under which we are going to work will all satisfy the $0-1$ Blumethal law. Hence, B_0 will have to be a constant and we choose 0 for simplicity.*

5.2.4 Notations and definitions

Following [111], for $a, b \in \mathcal{D}$ and $\nu, \beta \in \mathcal{N}$, we define the first disagreement times as follows

$$\theta^{a,b} := \inf\left\{t \ge 0, \ \int_0^t a_s ds \ne \int_0^t b_s ds\right\},$$

$$\theta_{\nu,\beta} := \inf\left\{t \ge 0, \ \int_0^t \int_E x \nu_s(dx) ds \ne \int_0^t \int_E x \beta_s(dx) ds\right\}$$

$$\theta_{\nu,\beta}^{a,b} := \theta^{a,b} \wedge \theta_{\nu,\beta}.$$

For every $\hat{\tau}$ in $\hat{\mathcal{F}}^{\mathcal{P}}$, we define the following event

$$\Omega_{\hat{\tau}}^{a,\nu,b,\beta} := \left\{\hat{\tau} < \theta_{\nu,\beta}^{a,b}\right\} \cup \left\{\hat{\tau} = \theta_{\nu,\beta}^{a,b} = +\infty\right\}.$$

Finally, we introduce the following notion inspired by [111]

Definition 5.2.4 *$\mathcal{A}_0 \subset \mathcal{A}_W$ is a generating class of coefficients if*

(i) *$\mathcal{A}_0$ is stable for the concatenation operation, i.e. if $(a,\nu),(b,\beta) \in \mathcal{A}_0 \times \mathcal{A}_0$ then for each t,*

$$\left(a\mathbf{1}_{[0,t]} + b\mathbf{1}_{[t,+\infty)}, \nu\mathbf{1}_{[0,t]} + \beta\mathbf{1}_{[t,+\infty)}\right) \in \mathcal{A}_0.$$

(ii) *For every $(a,\nu),(b,\beta) \in \mathcal{A}_0 \times \mathcal{A}_0$, $\theta^{a,b}_{\nu,\beta}$ is a constant. Or equivalently, for each t, $\Omega_t^{a,\nu,b,\beta}$ equals Ω or $\emptyset$.*

Definition 5.2.5 *We say that $\mathcal{A}$ is a separable class of coefficients generated by $\mathcal{A}_0$ if $\mathcal{A}_0$ is a generating class of coefficients and if $\mathcal{A}$ consists of all processes a and random measures ν of the form*

$$\begin{aligned} a_t(\omega) &= \sum_{n=0}^{+\infty}\sum_{i=1}^{+\infty} a_t^{n,i}(\omega)\mathbf{1}_{E_n^i}(\omega)\mathbf{1}_{[\tau_n(\omega),\tau_{n+1}(\omega))}(t) \\ \nu_t(\omega) &= \sum_{n=0}^{+\infty}\sum_{i=1}^{+\infty} \nu^{n,i}\mathbf{1}_{\widetilde{E}_n^i}(\omega)\mathbf{1}_{[\tilde{\tau}_n(\omega),\tilde{\tau}_{n+1})}(t), \end{aligned} \tag{5.2.3}$$

where for each i and for each n, $(a^{n,i},\nu^{n,i}) \subset \mathcal{A}_0$, τ_n and $\tilde{\tau}_n$ are $\mathbb{F}$-stopping times with $\tau_0 = 0$, such that

(i) *$\tau_n < \tau_{n+1}$ on $\{\tau_n < +\infty\}$ and $\tilde{\tau}_n < \tilde{\tau}_{n+1}$ on $\{\tilde{\tau}_n < +\infty\}$.*

(ii) $\inf\{n \geq 0,\ \tau_n = +\infty\} + \inf\{n \geq 0,\ \tilde{\tau}_n = +\infty\} < \infty$.

(iii) *τ_n and $\tilde{\tau}_n$ take countably many values in some fixed $I_0 \subset [0,T]$ which is countable and dense in $[0,T]$.*

(iv) *For each n, $(E_i^n)_{i\geq 1} \subset \mathcal{F}_{\tau_n}$ and $(\tilde{E}_i^n)_{i\geq 1} \subset \mathcal{F}_{\tilde{\tau}_n}$ form a partition of Ω.*

Remark 5.2.3 *If we refine the subdivisions, we can always take a common sequence of stopping times $(\tau_n)_{n\geq 0}$ and common sets $(E_i^n)_{i\geq 1, n\geq 0}$ for a and for ν. This will be used throughout this section. Moreover, the definition indeed depends on the countable subset I_0 introduced above. We acknowledge that as in [111] this set could be changed, but for the sake of clarity, it will be fixed throughout the chapter. We emphasize that will prove in Chapter 6 (see Section 6.2.4) that this has only a limited impact on our results. For practical purposes, one could take for instance $I_0 = \mathbb{Q} \cap [0,T]$.*

Remark 5.2.4 *The form for a and ν in Definition 5.2.5 is directly inspired by the so-called properties of stability by concatenation and by bifurcation in the theory of stochastic control. As shown for instance in [45], [48] or [17] (see Remark 3.1), this property of control processes is tailor-made to be able to retrieve the dynamic programming principle, and is the least that once can expect. In our case, 2BSDEs can be seen formally as a weak version of a stochastic control problem for which the controls are a and ν, and we will see in Proposition 6.2.2 in Chapter 6 that this form is exactly what we need to recover the dynamic programming equation for our solution of a 2BSDE with jumps.*

Example 5.2.1 *$\widetilde{\mathcal{A}}_0$ composed of deterministic processes a and ν form a generating class of coefficients.*

The following Proposition generalizes Proposition 4.11 of [111] and shows that a separable class of coefficients inherits the "good" properties of its generating class.

Proposition 5.2.1 *Let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$. Then*

(i) *If $\mathcal{A}_0 \subset \mathcal{A}_W$, then $\mathcal{A} \subset \mathcal{A}_W$.*

(ii) *$\mathcal{A}$-quasi surely is equivalent to $\mathcal{A}_0$-quasi surely.*

(iii) *If every $\mathbb{P} \in \{\mathbb{P}^{\alpha}_{\nu},\ (\alpha,\nu) \in \mathcal{A}_0\}$ satisfies the martingale representation property, then every $\mathbb{P} \in \{\mathbb{P}^{\alpha}_{\nu},\ (\alpha,\nu) \in \mathcal{A}\}$ also satisfies the martingale representation property.*

(iv) *If every $\mathbb{P} \in \{\mathbb{P}^{\alpha}_{\nu},\ (\alpha,\nu) \in \mathcal{A}_0\}$ satisfies the Blumenthal $0-1$ law, then every $\mathbb{P} \in \{\mathbb{P}^{\alpha}_{\nu},\ (\alpha,\nu) \in \mathcal{A}\}$ also satisfies the Blumenthal $0-1$ law.*

As in [111], to prove this result, we need two Lemmas. The first one is a straightforward generalization of Lemma 4.12 in [111], so we omit the proof. The second one is analogous to Lemma 4.13 in [111].

Lemma 5.2.1 *Let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$. For any $(a,\nu) \in \mathcal{A}$, and any $\mathbb{F}$-stopping time $\tau \in \mathcal{T}$, there exist $\tilde{\tau} \in \mathcal{T}$ with $\tilde{\tau} \geq \tau$, a sequence $(a^n,\nu^n)_{n\geq 1} \subset \mathcal{A}_0$ and a partition $(E_n)_{n\geq 1} \subset \mathcal{F}_\tau$ of Ω such that $\tilde{\tau} > \tau$ on $\{\tau < +\infty\}$ and*

$$a_t(\omega) = \sum_{n\geq 1} a^n_t(\omega)\mathbf{1}_{E_n}(\omega) \text{ and } \nu_t(\omega) = \sum_{n\geq 1} \nu^n_t(\omega)\mathbf{1}_{E_n}(\omega),\ t < \tilde{\tau}. \tag{5.2.4}$$

In particular, $E_n \subset \Omega^{a,\nu,a^n,\nu^n}_{\hat{\tau}}$ which implies that $\cup_n \Omega^{a,\nu,a^n,\nu^n}_{\hat{\tau}} = \Omega$. Finally, if a and ν take the form 5.2.3 and $\tau \geq \tau_n$, then we can choose $\tilde{\tau} \geq \tau_{n+1}$.

Proof. We refer to the proof of Lemma 4.12 in [111]. □

Lemma 5.2.2 *Let $\tau_1, \tau_2 \in \mathcal{T}$ be two stopping times such that $\tau_1 \leq \tau_2$, and $(a^i,\nu^i)_{i\geq 1} \subset \overline{\mathcal{A}}_W$ and let $\{E_i,\ i \geq 1\} \subset \mathcal{F}_{\tau_1}$ be a partition of Ω. Finally let $\mathbb{P}^0$ be a probability measure on $\mathcal{F}_{\tau_1}$ and let $\{\mathbb{P}^i,\ i \geq 1\}$ be a sequence of probabilty measures such that for each i, $\mathbb{P}^i$ is a solution of the martingale problem $(\mathbb{P}^0, \tau_1, \tau_2, a^i, \nu^i)$. Define*

$$\mathbb{P}(E) := \sum_{i\geq 1} \mathbb{P}^i(E \cap E_i) \text{ for all } \mathcal{F}_{\tau_2},$$

$$a_t := \sum_{i\geq 1} a^i_t \mathbf{1}_{E_i} \text{ and } \nu_t := \sum_{i\geq 1} \nu^i_t \mathbf{1}_{E_i},\ t \in [\tau_1, \tau_2].$$

Then $\mathbb{P}$ is a solution of the martingale problem $(\mathbb{P}^0, \tau_1, \tau_2, a, \nu)$.

Proof. By definition, $\mathbb{P} = \mathbb{P}^0$ on $\mathcal{F}_{\tau_1}$. In view of Remark 5.2.1, it is enough to prove that M, J and Q are $\mathbb{P}$-local martingales on $[\tau_1, \tau_2]$. By localizing if necessary, we may assume as usual that all these processes are actually bounded. For any stopping times $\tau_1 \leq R \leq S \leq \tau_2$, and any bounded $\mathcal{F}_R$-measurable random variable η, we have

$$\mathbb{E}^{\mathbb{P}}\left[(M_S - M_R)\eta\right] = \sum_{i\geq 1} \mathbb{E}^{\mathbb{P}^i}\left[(M_S - M_R)\eta \mathbf{1}_{E_i}\right] = \sum_{i\geq 1} \mathbb{E}^{\mathbb{P}^i}\left[\mathbb{E}^{\mathbb{P}^i}[(M_S - M_R)|\mathcal{F}_R]\eta \mathbf{1}_{E_i}\right] = 0.$$

Thus M is a $\mathbb{P}$-local martingale on $[\tau_1, \tau_2]$. We can prove in exactly the same manner that J and Q are also $\mathbb{P}$-local martingales on $[\tau_1, \tau_2]$ and the proof is complete. □

Proof of Proposition 5.2.1. The proof follows closely the proof of Proposition 4.11 in [111] and we give it for the conenience of the reader.

(i) We take $(a, \nu) \in \mathcal{A}$, let us prove that $(a, \nu) \in \mathcal{A}_W$. We fix two stopping times θ_1, θ_2 in $\mathcal{T}$. We define a sequence $(\tilde{\tau}_n)_{n\geq 0}$ as follows:

$$\tilde{\tau}_0 := \theta_1 \text{ and } \tilde{\tau}_n := (\tau_n \vee \theta_1) \wedge \theta_2, \ n \geq 1.$$

To prove that the martingale problem $(\mathbb{P}^0, \theta_1, \theta_2, a, \nu)$ has a unique solution, we prove by induction on n that the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_n, a, \nu)$ has a unique solution.

Step 1 of the induction: Let $n = 1$, and let us first construct a solution to the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_1, a, \nu)$. For this purpose, we apply Lemma 5.2.1 with $\tau = \tilde{\tau}_0$ and $\tilde{\tau} = \tilde{\tau}_1$, which leads to $a_t = \sum_{i\geq 1} a^i_t \mathbf{1}_{E_i}$ and $\nu_t = \sum_{i\geq 1} \nu^i_t \mathbf{1}_{E_i}$ for all $t < \tilde{\tau}_1$, where $(a^i, \nu^i) \in \mathcal{A}_0$ and $\{E_i,\ i \geq 1\} \subset \mathcal{F}_{\tilde{\tau}_0}$ form a partition of Ω. For $i \geq 1$, let $\mathbb{P}^{0,i}$ be the unique solution of the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_1, a_i, \nu_i)$ and define

$$\mathbb{P}^{0,a}(E) := \sum_{i\geq 1} \mathbb{P}^{0,i}(E \cap E_i) \text{ for all } E \in \mathcal{F}_{\tilde{\tau}_1}.$$

Lemma 5.2.2 tells us that $\mathbb{P}^{0,a}$ solves the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_1, a, \nu)$. Now let $\mathbb{P}$ be an arbitrary solution of the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_1, a, \nu)$, and let us prove that $\mathbb{P} = \mathbb{P}^{0,a}$. We first define

$$\mathbb{P}^i(E) := \mathbb{P}(E \cap E_i) + \mathbb{P}^{0,i}(E \cap E_i^c),\ \forall E \in \mathcal{F}_{\tilde{\tau}_1}.$$

Using Lemma 5.2.2, and the facts that $a^i = a\mathbf{1}_{E_i} + a^i \mathbf{1}_{E_i^c}$ and $\nu^i = \nu \mathbf{1}_{E_i} + \nu^i \mathbf{1}_{E_i^c}$, we conclude that $\mathbb{P}^i$ solves the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_1, a^i, \nu^i)$. This problem having a unique solution, we have $\mathbb{P}^i = \mathbb{P}^{0,i}$ on $\mathcal{F}_{\tilde{\tau}_1}$. This implies that for each $i \geq 1$ and for each $E \in \mathcal{F}_{\tilde{\tau}_1}$, $\mathbb{P}^i(E \cap E_i) = \mathbb{P}^{0,i}(E \cap E_i)$, and finally

$$\mathbb{P}^{0,a}(E) = \sum_{i\geq 1} \mathbb{P}^{0,i}(E \cap E_i) = \sum_{i\geq 1} \mathbb{P}^i(E \cap E_i) = \mathbb{P}(E),\ \forall E \in \mathcal{F}_{\tilde{\tau}_1}.$$

Step 2 of the induction: We assume that the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_n, a, \nu)$ has a unique solution denoted $\mathbb{P}^n$. Using the same reasoning as above, we see that the martingale problem $(\mathbb{P}^n, \tilde{\tau}_n, \tilde{\tau}_{n+1}, a, \nu)$ has a unique solution, denoted $\mathbb{P}^{n+1}$. Then the processes M, J and Q defined in Remark 5.2.1 are $\mathbb{P}^{n+1}$-local martingales on $[\tilde{\tau}_n, \tilde{\tau}_{n+1}]$, and since $\mathbb{P}^{n+1}$ coincides with $\mathbb{P}^n$ on $\mathcal{F}_{\tilde{\tau}_n}$, M, J and Q are also $\mathbb{P}^{n+1}$-local martingales on $[\tilde{\tau}_0, \tilde{\tau}_n]$. And hence $\mathbb{P}^{n+1}$ solves the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_{n+1}, a, \nu)$. We suppose now that $\mathbb{P}$ is another arbitrary solution to the martingale problem $(\mathbb{P}^0, \tilde{\tau}_0, \tilde{\tau}_{n+1}, a, \nu)$. By the induction assumption, $\mathbb{P}^n = \mathbb{P}$ on $\mathcal{F}_{\tilde{\tau}_n}$, then $\mathbb{P}$ solves the martingale problem $(\mathbb{P}^n, \tilde{\tau}_n, \tilde{\tau}_{n+1}, a, \nu)$, and by uniqueness $\mathbb{P} = \mathbb{P}^{n+1}$ on $\mathcal{F}_{\tilde{\tau}_{n+1}}$. The induction is now complete.

Remark that $\mathcal{F}_{\theta_2} = \vee_{n\geq 1}\mathcal{F}_{\tilde{\tau}_n}$. Indeed, since $\inf\{n \geq 1,\ \tau_n = +\infty\} < +\infty$, then $\inf\{n \geq 1,\ \tilde{\tau}_n = \theta_2\} < +\infty$. This allows to define $\mathbb{P}^\infty(E) := \mathbb{P}^n(E)$ for $E \in \mathcal{F}_{\tilde{\tau}_n}$ and to extend it uniquely to $\mathcal{F}_{\theta_2}$. Now using again Remark 5.2.1, we conclude that $\mathbb{P}^\infty$ solves $(\mathbb{P}^0, \theta_1, \theta_2, a, \nu)$ and is unique.

(ii) We now prove that $\mathcal{A}$-quasi surely is equivalent to $\mathcal{A}_0$-quasi surely.

We take $(a, \nu) \in \mathcal{A}$ and we apply Lemma 5.2.1 with $\tau = +\infty$ to write $a_t = \sum_{i\geq 1} a_t^i \mathbf{1}_{E_i}$ and $\nu_t = \sum_{i\geq 1} \nu_t^i \mathbf{1}_{E_i}$ for all $t \geq 0$, where $(a^i, \nu^i) \in \mathcal{A}_0$ and $\{E_i,\ i \geq 1\} \subset \mathcal{F}_\infty$ form a partition of Ω. Take a set E such that $\mathbb{P}_{\tilde{\nu}}^{\tilde{a}}(E) = 0$ for every $(\tilde{a}, \tilde{\nu}) \in \mathcal{A}_0$, then

$$\mathbb{P}_\nu^a(E) = \sum_{i\geq 1} \mathbb{P}_\nu^a(E \cap E_i) = \sum_{i\geq 1} \mathbb{P}_{\nu^i}^{a^i}(E \cap E_i) = 0.$$

(iii) By (i), since $\mathcal{A} \subset \mathcal{A}_W$, for any $(a, \nu) \in \mathcal{A}$, the corresponding martingale problem has a unique solution, which is therefore an extremal point in the set of solutions. Hence, we can apply Theorem III.4.29 in [65] to obtain immediately the predictable representation property.

(iv) Take $(a, \nu) \in \mathcal{A}$ of the form (5.2.3), in which we can take $\tau_0 = 0$ without loss of generality. $\mathbb{P}_\nu^a$ is the law on $[0, \tau_1]$ of a semimartingale with characteristics

$$\left(-\int_0^t \int_E x \mathbf{1}_{|x|>1} \tilde{\nu}_s(dx)ds,\ \int_0^t \tilde{a}_s ds,\ \tilde{\nu}_s(dx)ds \right),$$

where

$$\tilde{a}_t := \sum_{i\geq 1} a^{0,i} \mathbf{1}_{E_0^i} \text{ and } \tilde{\nu}_t := \sum_{i\geq 1} \nu^{0,i} \mathbf{1}_{E_0^i},$$

where $\{E_0^i,\ i \geq 1\} \subset \mathcal{F}_0$ is a partition of Ω. Since $\mathcal{F}_0$ is trivial, the partition is only composed of Ω and $\emptyset$, and then

$$\tilde{a}_t := a_t^{0,1} \text{ and } \tilde{\nu}_t = \nu_t^{0,1}.$$

Then for $E \in \mathcal{F}_{0^+}$, $\mathbb{P}_\nu^a(E) = \mathbb{P}_{\tilde{\nu}}^{\tilde{a}}(E) = 0$, since $\mathbb{P}_{\tilde{\nu}}^{\tilde{a}}$ satisfies the Blumenthal $0 - 1$ law by hypothesis. □

Remark 5.2.5 *If $\mathcal{A}_0$ consists of deterministic mappings as in example 5.2.1, then $\mathbb{P}_\nu^a$ is the law on $[0, \tau_1]$ of an additive process with non random characteristics, for which the Blumenthal $0 - 1$ law holds (see for instance [107]).*

We now state the following Proposition which tells us that our probability measure coincide until their first time of disagreement.

Proposition 5.2.2 *Let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$, let $\mathcal{P}_{\mathcal{A}} := \{\mathbb{P}^a_\nu,\ (a,\nu) \in \mathcal{A}\}$ and let $(a,\nu) \times (b,\beta) \in \mathcal{A} \times \mathcal{A}$.*

(i) *$\theta^{a,b}_{\nu,\beta}$ is an $\mathbb{F}$-stopping time taking countably many values.*

(ii) *Moreover, we have the following coherence condition*

$$\mathbb{P}^a_\nu\left(E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}}\right) = \mathbb{P}^b_\beta\left(E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}}\right),\ \forall \hat{\tau} \in \hat{\mathcal{T}}^{\mathcal{P}_{\mathcal{A}}},\ \forall E \in \hat{\mathcal{F}}^{\mathcal{P}_{\mathcal{A}}}_{\hat{\tau}}.$$

Proof. (i) Let us prove that $\left\{\theta^{a,b}_{\nu,\beta} \leq t_1\right\} \in \mathcal{F}_{t_1}$, for any $t_1 \geq 0$.

We apply Lemma 5.2.1 for (a,ν) and (b,β) with $\tau = t_1$ to obtain the existence of (a^n_t, ν^n_t), (b^n_t, β^n_t) in $\mathcal{A}_0$ and $E_n \subset \mathcal{F}_{t_1}$ forming a partition of Ω, such that $(a^n, b^n, \nu^n, \beta^n)$ coincide with (a,b,ν,β) on E_n for $t < \tilde{\tau}$, where $\tilde{\tau} > t_1$. Then

$$\left\{\theta^{a,b}_{\nu,\beta} \leq t_1\right\} = \bigcup_{n\geq 1}\left\{\theta^{a^n,b^n}_{\nu^n,\beta^n} \leq t_1\right\} \cap E_n$$

By the constant disagreement times property of $\mathcal{A}_0$, $\left\{\theta^{a^n,b^n}_{\nu^n,\beta^n} \leq t_1\right\}$ is either Ω or $\emptyset$, and since $E_n \in \mathcal{F}_{t_1}$, then

$$\left\{\theta^{a,b}_{\nu,\beta} \leq t_1\right\} \in \mathcal{F}_{t_1}.$$

To show that $\theta^{a,b}_{\nu,\beta}$ takes countably many values, we apply again Lemma 5.2.1 with $\tau = \theta^{a,b}_{\nu,\beta}$, to obtain the existence of (a^n_t, ν^n_t), (b^n_t, β^n_t) in $\mathcal{A}_0$ and $E_n \subset \mathcal{F}_\tau$ forming a partition of Ω, such that $(a^n, b^n, \nu^n, \beta^n)$ coincide with (a,b,ν,β) on E_n for $t < \tilde{\tau}$, where $\tilde{\tau} > \tau$. Since $\theta^{a^n,b^n}_{\nu^n,\beta^n}$ is a constant and given that $\theta^{a,b}_{\nu,\beta} = \theta^{a^n,b^n}_{\nu^n,\beta^n}$ on E_n, we have the desired result.

(ii) We write that for any $E \in \hat{\mathcal{F}}^{\mathcal{P}_{\mathcal{A}}}_{\hat{\tau}}$

$$\begin{aligned}E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}} \cap \left\{\theta^{a,b}_{\nu,\beta} \leq t\right\} &= E \cap \left\{\hat{\tau} < \theta^{a,b}_{\nu,\beta}\right\} \cap \left\{\theta^{a,b}_{\nu,\beta} \leq t\right\} \\ &= \bigcup_{m\geq 1}\left(E \cap \left\{\hat{\tau} < \theta^{a,b}_{\nu,\beta}\right\} \cap \left\{\hat{\tau} \leq t - \frac{1}{m}\right\} \cap \left\{\theta^{a,b}_{\nu,\beta} \leq t\right\}\right).\end{aligned}$$

Since $\left\{\theta^{a,b}_{\nu,\beta} \leq t\right\} \in \mathcal{F}_t$, we get that for any $m \geq 1$,

$$E \cap \left\{\hat{\tau} < \theta^{a,b}_{\nu,\beta}\right\} \cap \left\{\hat{\tau} \leq t - \frac{1}{m}\right\} \in \hat{\mathcal{F}}^{\mathcal{P}_{\mathcal{A}}}_{t-\frac{1}{m}} \subset \mathcal{F}^+_{t-\frac{1}{m}} \vee \mathcal{N}^{\mathbb{P}^a_\nu}(\mathcal{F}_\infty) \subset \mathcal{F}_t \vee \mathcal{N}^{\mathbb{P}^a_\nu}(\mathcal{F}_\infty),$$

and then

$$E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}} \in \mathcal{F}_{\theta^{a,b}_{\nu,\beta}} \vee \mathcal{N}^{\mathbb{P}^a_\nu}(\mathcal{F}_\infty).$$

From this last assertion, we deduce that there exist measurable sets $E^{a,\nu}_i$, $E^{b,\beta}_i$ belonging to $\mathcal{F}_{\theta^{a,b}_{\nu,\beta}}$, $i = 1,2$, such that

$$E^{a,\nu}_1 \subset E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}} \subset E^{a,\nu}_2,\ E^{b,\beta}_1 \subset E \cap \Omega^{a,\nu,b,\beta}_{\hat{\tau}} \subset E^{b,\beta}_2$$
$$\mathbb{P}^a_\nu\left(E^{a,\nu}_2 \backslash E^{a,\nu}_1\right) = \mathbb{P}^b_\beta\left(E^{b,\beta}_2 \backslash E^{b,\beta}_1\right) = 0.$$

We set $E^1 := E_1^{a,\nu} \cup E_1^{b,\beta}$ and $E^2 := E_2^{b,\beta} \cap E_1^{b,\beta}$, then

$$E^1,\, E^2 \in \mathcal{F}_{\theta^{a,b}_{\nu,\beta}},\; E^1 \subset E \cap \Omega_{\hat{\tau}}^{a,\nu,b,\beta} \subset E^2 \text{ and } \mathbb{P}_\nu^a\left(E^2 \backslash E^1\right) = \mathbb{P}_\beta^b\left(E^2 \backslash E^1\right) = 0.$$

This implies that

$$\mathbb{P}_\nu^a\left(E \cap \Omega_{\hat{\tau}}^{a,\nu,b,\beta}\right) = \mathbb{P}_\nu^a(E^2) \text{ and } \mathbb{P}_\beta^b\left(E \cap \Omega_{\hat{\tau}}^{a,\nu,b,\beta}\right) = \mathbb{P}_\beta^b(E^2),$$

but the solutions of the martingale problems $(\mathbb{P}^0, 0, \theta^{a,b}_{\nu,\beta}, a, \nu)$ and $(\mathbb{P}^0, 0, \theta^{a,b}_{\nu,\beta}, b, \beta)$ are equal by definition. And since $E^2 \in \mathcal{F}_{\theta^{a,b}_{\nu,\beta}}$, we have

$$\mathbb{P}_\nu^a(E^2) = \mathbb{P}_\beta^b(E^2)$$

which gives the desired result. □

We now have every tools we need to state and prove the main resut of this section, which generalizes the aggregation result of Theorem 5.1 in [111]. For this purpose, we use the more general aggregation result of Cohen [30], that does not concern only volatility or jump measure uncertainty.

Theorem 5.2.1 *Let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$ and $\mathcal{P}_\mathcal{A}$ the corresponding probability measures. Let*

$$\{X^{a,\nu},\, (a,\nu) \in \mathcal{A}\}\,,$$

be a family of $\hat{\mathbb{F}}^\mathbb{P}$-progressively measurable processes.

Then the following two conditions are equivalent

(i) $\{X^{a,\nu},\, (a,\nu) \in \mathcal{A}\}$ *satisfies the following consistency condition*

$$X^{a,\nu} = X^{b,\beta},\; \mathbb{P}_\nu^a\text{-a.s. on } [0, \theta^{a,b}_{\nu,\beta}) \text{ for any } (a,\nu) \in \mathcal{A} \text{ and } (b,\beta) \in \mathcal{A}.$$

(ii) *There exists a $\mathcal{P}$-q.s unique process X such that*

$$X = X^{a,\nu},\; \mathbb{P}_\nu^a\text{-a.s.},\; \forall (a,\nu) \in \mathcal{A}.$$

Proof. We first prove that (i) implies (ii). Using Lemma 3 in [30], we see that the definition of the generating classes, together with Proposition 5.2.2, implies that the family $\mathcal{P}$ satisfies the Hahn property defined in [30]. Now Theorem 4 of [30] gives the result. The fact that (ii) implies (i) is a consequence of the unicity of the solution of the martingale problem $(\mathbb{P}^0, 0, \theta^{a,b}_{\nu,\beta}, a, \nu)$ on $[0, \theta^{a,b}_{\nu,\beta})$. □

Now that we have Theorem 5.2.1, we can answer our first issue concerning the aggregation of the predictable compensators associated to the jump measure μ_{B^d} of the canonical process. Indeed, let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$ and $\mathcal{P} := \{\mathbb{P}_\nu^a,\, (a,\nu) \in \mathcal{A}\}$. Then, for each Borel set $A \in \mathcal{B}(E)$ and for each $t \in [0,T]$ the family $\{\nu_t^{\mathbb{P}_\nu^a}(A)\}_{(a,\nu)\in\mathcal{A}}$ clearly satisfies the consistency condition above (because it is defined through the Doob-Meyer decomposition, see Proposition 6.6 in [111]), and therefore there exists a process $\widehat{\nu}$ such that

$$\widehat{\nu}_t(A) = \nu_t^\mathbb{P}(A), \text{ for every } \mathbb{P} \in \mathcal{P}_\mathcal{A}. \tag{5.2.5}$$

We then denote

$$\widetilde{\mu}_{B^d}(dt, dx) := \mu_{B^d}(dt, dx) - \widehat{\nu}_t(dx)dt.$$

5.2.4.1 The strong formulation

In this Section, we will concentrate on a subset of $\mathcal{P}_W$. For this purpose, we define

$$\mathcal{V} := \{\nu \in \mathcal{N},\ (I_d, \nu) \in \mathcal{A}_W\}.$$

For each $\nu \in \mathcal{V}$, we denote $\mathbb{P}^\nu := \mathbb{P}_\nu^{I_d}$ and for each $\alpha \in \mathcal{D}$, we define

$$\mathbb{P}^{\alpha,\nu} := \mathbb{P}^\nu \circ (X_\cdot^\alpha)^{-1}, \text{ where } X_t^\alpha := \int_0^t \alpha_s^{1/2} dB_s^c + B_t^d,\ \mathbb{P}^\nu - a.s. \tag{5.2.6}$$

Let us now define,

$$\mathcal{P}_S := \{\mathbb{P}^{\alpha,\nu},\ (\alpha,\nu) \in \mathcal{A}_W\}.$$

Then α is the quadratic variation density of the continuous part of X^α and

$$dB_s^c = \alpha_s^{-1/2} dX_s^{\alpha,c},$$

under $\mathbb{P}^\nu$. Moreover, $\nu_t(dx)dt$ is the compensator of the measure associated to the jumps of X^α and $\Delta X_s^\alpha = \Delta B_s$ under $\mathbb{P}^\nu$. We also define for each $\mathbb{P} \in \mathcal{P}_W$ the following process

$$L_t^{\mathbb{P}} := W_t^{\mathbb{P}} + B_t^d,\ \mathbb{P} - a.s., \tag{5.2.7}$$

where $W_t^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion defined by

$$W_t^{\mathbb{P}} := \int_0^t \widehat{a}_s^{-1/2} dB_s^c.$$

Then, $\mathcal{P}_S$ is a subset of $\mathcal{P}_W$ and we have by definition

$$\text{the } \mathbb{P}^{\alpha,\nu}\text{-distribution of } (B, \widehat{a}, \widehat{\nu}, L^{\mathbb{P}^{\alpha,\nu}}) \text{ is equal to the } \mathbb{P}^\nu\text{-distribution of } (W^\alpha, \alpha, \nu, B). \tag{5.2.8}$$

We also have the following characterization in terms of filtrations, which is similar to Lemma 8.1 in [111]

Lemma 5.2.3 $\mathcal{P}_S = \left\{\mathbb{P} \in \mathcal{P}_W,\ \overline{\mathbb{F}^{L^{\mathbb{P}}}}^{\mathbb{P}} = \overline{\mathbb{F}}^{\mathbb{P}}\right\}$

Proof. By the above remarks, it is clear that α and B are $\overline{\mathbb{F}^{X^\alpha}}^{\mathbb{P}^\nu}$-progressively measurable. But by definition, $\mathbb{F}$ is generated by B, thus we conclude easily that $\overline{\mathbb{F}}^{\mathbb{P}^\nu} \subset \overline{\mathbb{F}^{X^\alpha}}^{\mathbb{P}^\nu}$. The other inclusion being clear by definition, we have

$$\overline{\mathbb{F}}^{\mathbb{P}^\nu} = \overline{\mathbb{F}^{X^\alpha}}^{\mathbb{P}^\nu}.$$

Now we can use (5.2.8) to obtain that $\overline{\mathbb{F}^{L^{\mathbb{P}^{\alpha,\nu}}}}^{\mathbb{P}^{\alpha,\nu}} = \overline{\mathbb{F}}^{\mathbb{P}^{\alpha,\nu}}$.

Conversely, let $\mathbb{P} \in \mathcal{P}_W$ be such that $\overline{\mathbb{F}^{L^{\mathbb{P}^{\alpha,\nu}}}}^{\mathbb{P}^{\alpha,\nu}} = \overline{\mathbb{F}}^{\mathbb{P}^{\alpha,\nu}}$. Then, there exists some measurable function β such that $B_\cdot = \beta(L_\cdot^{\mathbb{P}})$. Let ν be the compensator of the measure associated to the jumps of B under $\mathbb{P}$. Define then,

$$\alpha_t := \frac{d < \beta(B), \beta(B) >_t^c}{dt},$$

we conclude then that $\mathbb{P} = \mathbb{P}^{\alpha,\nu}$. □

Define now $\mathcal{A}_S := \{(\alpha,\nu) \in \mathcal{A}_W,\ \mathbb{P}^\alpha_\nu \in \mathcal{P}_S\}$. It is important to notice that in our framework, it is not clear whether all the probability measures in $\mathcal{P}_S$ satisfy the martingale representation property and the Blumenthal $0-1$ law. Indeed, this is due to the fact that the processus $L^{\mathbb{P}}$ does not necessarily satisfy them. This is a major difference with [111]. Nonetheless, if we restrict ourselves to a subset of $\mathcal{P}_S$, we are going to see that we can still recover them. First, we have the following generalization of Proposition 8.3 of [111].

Proposition 5.2.3 *Let $\mathcal{A}$ be a separable class of coefficients generated by $\mathcal{A}_0$. If $\mathcal{A}_0 \subset \mathcal{A}_S$, then $\mathcal{A} \subset \mathcal{A}_S$.*

Proof. This is a straightforward generalization of the proof of Proposition 8.3 in [111], using the same kind of modifications as in our previous proofs, so we omit it. □

Let us now consider the set introduced above in Example 5.2.1

$$\tilde{\mathcal{A}}_0 := \{(\alpha,\nu) \in \mathcal{D} \times \mathcal{N} \text{ which are deterministic}\},\ \mathcal{P}_{\tilde{\mathcal{A}}_0} := \left\{\mathbb{P}^\alpha_\nu,\ (\alpha,\nu) \in \tilde{\mathcal{A}}_0\right\}.$$

$\tilde{\mathcal{A}}_0$ is a generating class of coefficients, and it is a well known result that $\tilde{\mathcal{A}}_0 \subset \mathcal{A}_W$ (see Theorem III.2.16 in [65]) and that every probability measure in $\mathcal{P}_{\tilde{\mathcal{A}}_0}$ satisfies the martingale representation property and the Blumenthal $0-1$ law, since the canonical process is actually an additive process under them. Moreover we also have

Lemma 5.2.4 *We have $\mathcal{P}_{\tilde{\mathcal{A}}_0} \subset \mathcal{P}_S$.*

Proof. Let $\mathbb{P} := \mathbb{P}^\alpha_\nu$ be a probability measure in $\mathcal{P}_{\tilde{\mathcal{A}}_0}$. As argued previously, we have $\mathbb{P}-a.s.$

$$B^c_t = \int_0^t \alpha_s^{1/2} dL^{\mathbb{P},c}_t \text{ and } \Delta L^{\mathbb{P}}_t = \Delta B_t.$$

Since α is deterministic, it is clear that we have $\overline{\mathbb{F}}^{\mathbb{P}} = \overline{\mathbb{F}^{L^{\mathbb{P}}}}^{\mathbb{P}}$, which implies the result. □

Finally, we consider $\tilde{\mathcal{A}}$ the separable class of coefficients generated by $\tilde{\mathcal{A}}_0$ and $\mathcal{P}_{\tilde{\mathcal{A}}}$ the corresponding set of probability measures. Then, using the above results and Propositions 5.2.1 and 5.2.3, we have

Proposition 5.2.4 *$\mathcal{P}_{\tilde{\mathcal{A}}} \subset \mathcal{P}_S$ and every probability measure in $\mathcal{P}_{\tilde{\mathcal{A}}}$ satisfies the martingale representation property and the Blumenthal $0-1$ law.*

Proof. Once we know that the augmented filtration generated by $L^{\mathbb{P}}$ satisfies the martingale representation property and the Blumenthal $0-1$ law for every $\mathbb{P} \in \mathcal{P}_{\tilde{\mathcal{A}}_0}$, we can argue exactly as in the proof of Lemma 8.2 to obtain the results for $\mathcal{P}_{\tilde{\mathcal{A}}_0}$. The result for $\mathcal{P}_{\tilde{\mathcal{A}}}$ then comes easily from Proposition 5.2.1. □

Remark 5.2.6 *In our jump framework, we need to impose this separability structure on both α and ν, in order to be able to retrieve not only the aggregation result of Theorem 5.2.1 but also the property that all our probability measures satisfy the Blumenthal $0-1$ law and the martingale representation* *property. However, if one is only interested in being able to consider standard BSDEs with jumps, then we do not need the aggregation result and we can work with a larger set of probability measures without restrictions on the α. Namely, let us define*

$$\overline{\mathcal{P}}_{\widetilde{\mathcal{A}}} := \left\{\mathbb{P}^{\alpha,\nu},\ \alpha \in \mathcal{D},\ (I_d, \nu) \in \widetilde{\mathcal{A}}\right\}.$$

Then we can show as above that $\overline{\mathcal{P}}_{\widetilde{\mathcal{A}}} \subset \mathcal{P}_S$ and that all the probability measures in $\overline{\mathcal{P}}_{\widetilde{\mathcal{A}}}$ satisfy the Blumenthal $0-1$ law and the martingale representation property. This is going to be important in Chapter 6.

5.3 Preliminaries on 2BSDEs

5.3.1 The non-linear Generator

In this subsection we will introduce the function which will serve as the generator of our 2BSDE with jumps. Let us define the spaces

$$\hat{L}^2 := \bigcap_{\nu\in\mathcal{N}} L^2(\nu) \text{ and } \hat{L}^1 := \bigcap_{\nu\in\mathcal{N}} L^1(\nu).$$

For any $\mathcal{C}^1$ function v with bounded gradient, any $\omega \in \Omega$ and any $0 \leq t \leq T$, we denote $\tilde{v}$ the function

$$\tilde{v}(e) := v(e+\omega(t)) - v(\omega(t)) - \mathbf{1}_{\{|e|\leq 1\}}\ e.(\nabla v)(\omega(t)),\ \text{for } e \in E.$$

The hypothesis on v ensure that $\tilde{v}$ is an element of $\hat{L}^1$. We then consider a map

$$H_t(\omega, y, z, u, \gamma, \tilde{v}) : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \times \hat{L}^2 \times D_1 \times D_2 \to \mathbb{R},$$

where $D_1 \subset \mathbb{R}^{d\times d}$ is a given subset containing 0 and $D_2 \subset \hat{L}^1 \cap \mathcal{N}^*$, where $\mathcal{N}^*$ denotes the topological dual of $\mathcal{N}$.

Define the following conjugate of H with respect to γ and v by

$$F_t(\omega, y, z, u, a, \nu) := \sup_{\{\gamma,\tilde{v}\}\in D_1\times D_2} \left\{\frac{1}{2}\mathrm{Tr}(a\gamma) + <\tilde{v}, \nu> - H_t(\omega, y, z, u, \gamma, \tilde{v})\right\},$$

for $a \in \mathbb{S}_d^{>0}$ and $\nu \in \mathcal{N}$, and where $<\tilde{v}, \nu>$ is defined by

$$<\tilde{v}, \nu> := \int_E \tilde{v}(e)\nu(de). \tag{5.3.1}$$

The quantity $<\tilde{v}, \nu>$ will not appear again in this chapter, since we formulate the needed hypothesis for the backward equation generator directly on the function F. But the particular form of $<\tilde{v}, \nu>$ comes from the intuition that the second order BSDE with jumps is an essential supremum of classical BSDEs with jumps (BSDEJs). Indeed, solutions to markovian BSDEJs provide viscosity solutions to some parabolic partial integro-differential equations whose non local operator is given by a quantity similar to $<\tilde{v}, \nu>$ (see [4] for more details). We define

$$\widehat{F}_t(y, z, u) := F_t(y, z, u, \widehat{a}_t, \widehat{\nu}_t) \text{ and } \widehat{F}_t^0 := \widehat{F}_t(0,0,0),\ \mathcal{P}_{\tilde{\mathcal{A}}}\text{-q.s.} \tag{5.3.2}$$

We denote by $D^1_{F_t(y,z,u)}$ the domain of F in a and by $D^2_{F_t(y,z,u)}$ the domain of F in ν, for a fixed (t, ω, y, z, u). As in [112] we fix a constant $\kappa \in (1,2]$ and restrict the probability measures in $\mathcal{P}_H^\kappa \subset \mathcal{P}_{\tilde{\mathcal{A}}}$

Definition 5.3.1 *$\mathcal{P}_H^\kappa$ consists of all $\mathbb{P}\in\mathcal{P}_{\tilde{\mathcal{A}}}$ such that*

(i) $\mathbb{E}^{\mathbb{P}}\left[\int_0^T\int_E x^2\widehat{\nu}_t(dx)dt\right]<+\infty.$

(ii) $\underline{a}^{\mathbb{P}}\leq\widehat{a}\leq\bar{a}^{\mathbb{P}}$, *$dt\times d\mathbb{P}-as$ for some* $\underline{a}^{\mathbb{P}},\bar{a}^{\mathbb{P}}\in\mathbb{S}_d^{>0}$, *and* $\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T\left|\widehat{F}_t^0\right|^\kappa dt\right)^{\frac{2}{\kappa}}\right]<+\infty.$

Remark 5.3.1 *The above conditions assumed on the probability measures in $\mathcal{P}_H^\kappa$ ensure that under any $\mathbb{P}\in\mathcal{P}_H^\kappa$, the canonical process B is actually a true càdlàg martingale. This will be important when we will define standard BSDEs under each of those probability measures.*

We now state our main assumptions on the function F which will be our main interest in the sequel

Assumption 5.3.1 (i) The domains $D^1_{F_t(y,z,u)}=D^1_{F_t}$ and $D^2_{F_t(y,z,u)}=D^2_{F_t}$ are independent of (ω,y,z,u).

(ii) For fixed (y,z,a,ν), F is $\mathbb{F}$-progressively measurable in $D^1_{F_t}\times D^2_{F_t}$.

(iii) The following uniform Lipschitz-type property holds. For all $(y,y',z,z',u,t,a,\nu,\omega)$

$$\left|F_t(\omega,y,z,u,a,\nu)-F_t(\omega,y',z',u,a,\nu)\right|\leq C\left(\left|y-y'\right|+\left|a^{1/2}\left(z-z'\right)\right|\right).$$

(iv) For all $(t,\omega,y,z,u^1,u^2,a,\nu)$, there exist two processes γ and γ' such that

$$\int_E\delta^{1,2}u(x)\gamma'_t(x)\nu(dx)\leq F_t(\omega,y,z,u^1,a,\nu)-F_t(\omega,y,z,u^2,a,\nu)\leq\int_E\delta^{1,2}u(x)\gamma_t(x)\nu(dx),$$

where $\delta^{1,2}u:=u^1-u^2$ and $c_1(1\wedge|x|)\leq\gamma_t(x)\leq c_2(1\wedge|x|)$ with $-1+\delta\leq c_1\leq 0,\ c_2\geq 0$, and $c'_1(1\wedge|x|)\leq\gamma'_t(x)\leq c'_2(1\wedge|x|)$ with $-1+\delta\leq c'_1\leq 0,\ c'_2\geq 0$, for some $\delta>0$.

(v) F is uniformly continuous in ω for the Skorohod topology, that is to say that there exists some modulus of continuity ρ such that for all $(t,\omega,\omega',y,z,u,a,\nu)$

$$\left|F_t(\omega,y,z,u,a,\nu)-F_t(\omega',y,z,u,a,\nu)\right|\leq\rho\left(d_S(\omega_{.\wedge t},\omega'_{.\wedge t})\right),$$

where d_S is the Skorohod metric and where $\omega_{.\wedge t}(s):=\omega(s\wedge t)$.

Remark 5.3.2 *The assumptions (i) and (ii) are classic in the second order framework ([112]). The Lipschitz assumption (iii) is standard in the BSDE theory since the paper [97]. The hypothesis (iv) allows to have a comparison theorem in the framework with jumps, it has been introduced in [106], it is also present in [4] in the form of an equality. The last hypothesis (v) of uniform continuity in ω is also proper to the second order framework, it is linked to our intensive use of regular conditional probability distributions in the next chapter to construct our solutions pathwise, thus avoiding complex issues related to negligeable sets.*

Remark 5.3.3 (i) *For $\kappa_1 < \kappa_2$, applying Hölder's inequality gives us*

$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T \left|\widehat{F}_t^0\right|^{\kappa_1} dt\right)^{\frac{2}{\kappa_1}}\right] \leq C\mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T \left|\widehat{F}_t^0\right|^{\kappa_2} dt\right)^{\frac{2}{\kappa_2}}\right],$$

where C is a constant. Then it is clear that $\mathcal{P}_H^\kappa$ is decreasing in κ.

(ii) *The Assumption 5.3.1, together with the fact that $\widehat{F}_t^0 < +\infty$, $\mathbb{P}^{\alpha,\nu}$-a.s for every $\mathbb{P}^{\alpha,\nu} \in \mathcal{P}_H^\kappa$, implies that $\widehat{a}_t \in D^1_{F_t}$ and $\widehat{\nu}_t \in D^2_{F_t}$ $dt \times \mathbb{P}^{\alpha,\nu}$-a.s., for all $\mathbb{P}^{\alpha,\nu} \in \mathcal{P}_H^\kappa$.*

5.3.2 The Spaces and Norms

We now define as in [112], the spaces and norms which will be needed for the formulation of the second order BSDEs.

As in the previous chapter, we will use the following notation:

Notation 5.3.1 *We will denote $\mathbb{E}_t(\cdot)$ the conditional expectation with respect to the filtration $\mathbb{F}$.*

For $p \geq 1$, $L_H^{p,\kappa}$ denotes the space of all $\mathcal{F}_T$-measurable scalar r.v. ξ with

$$\|\xi\|^p_{L_H^{p,\kappa}} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[|\xi|^p\right] < +\infty.$$

$\mathbb{H}_H^{p,\kappa}$ denotes the space of all $\mathbb{F}^+$-predictable $\mathbb{R}^d$-valued predictable processes Z with

$$\|Z\|^p_{\mathbb{H}_H^{p,\kappa}} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |\widehat{a}_t^{1/2} Z_t|^2 dt\right)^{\frac{p}{2}}\right] < +\infty.$$

$\mathbb{D}_H^{p,\kappa}$ denotes the space of all $\mathbb{F}^+$-progressively measurable $\mathbb{R}$-valued processes Y with

$$\mathcal{P}_H^\kappa - q.s. \text{ càdlàg paths, and } \|Y\|^p_{\mathbb{D}_H^{p,\kappa}} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\sup_{0\leq t\leq T} |Y_t|^p\right] < +\infty.$$

$\mathbb{J}_H^{p,\kappa}$ denotes the space of all $\mathbb{F}^+$-predictable functions U with

$$\|U\|^p_{\mathbb{J}_H^{p,\kappa}} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T \int_E |U_s(x)|^2\, \widehat{\nu}_t(dx)ds\right)^{\frac{p}{2}}\right] < +\infty.$$

For each $\xi \in L_H^{1,\kappa}$, $\mathbb{P} \in \mathcal{P}_H^\kappa$ and $t \in [0,T]$ denote

$$\mathbb{E}_t^{H,\mathbb{P}}[\xi] := \operatorname*{ess\,sup^{\mathbb{P}}}_{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})} \mathbb{E}_t^{\mathbb{P}'}[\xi] \text{ where } \mathcal{P}_H^\kappa(t^+,\mathbb{P}) := \left\{\mathbb{P}' \in \mathcal{P}_H^\kappa : \mathbb{P}' = \mathbb{P} \text{ on } \mathcal{F}_t^+\right\}.$$

Then we define for each $p \geq \kappa$,

$$\mathbb{L}_H^{p,\kappa} := \left\{\xi \in L_H^{p,\kappa} : \|\xi\|_{\mathbb{L}_H^{p,\kappa}} < +\infty\right\} \text{ where } \|\xi\|^p_{\mathbb{L}_H^{p,\kappa}} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\operatorname*{ess\,sup^{\mathbb{P}}}_{0\leq t\leq T} \left(\mathbb{E}_t^{H,\mathbb{P}}[|\xi|^\kappa]\right)^{\frac{p}{\kappa}}\right].$$

Remark 5.3.4 *$\mathbb{L}_H^{p,\kappa}$ aside, the definitions of the previous spaces are classic, but the second order framework induces the presence of an essential supremum over our family of probability measures. As for $\mathbb{L}_H^{p,\kappa}$, it appears naturally in the a priori estimates, we refer to [112] for more details.*

Finally, we denote by $\mathrm{UC}_b(\Omega)$ the collection of all bounded and uniformly continuous maps $\xi : \Omega \to \mathbb{R}$ with respect to the Skorohod distance d_S (defining the Skorohod topology), and we let

$$\mathcal{L}_H^{p,\kappa} := \text{the closure of } \mathrm{UC}_b(\Omega) \text{ under the norm } \|\cdot\|_{\mathbb{L}_H^{p,\kappa}}, \text{ for every } 1 < \kappa \leq p.$$

For a given probability measure $\mathbb{P} \in \mathcal{P}_H^\kappa$, the spaces $L^p(\mathbb{P})$, $\mathbb{D}^p(\mathbb{P})$, $\mathbb{H}^p(\mathbb{P})$ and $\mathbb{J}^p(\mathbb{P})$ correspond to the above spaces when the set of probability measures is only the singleton $\{\mathbb{P}\}$. Finally, we have $\mathbb{H}^p_{loc}(\mathbb{P})$ denotes the space of all $\mathbb{F}^+$-predictable $\mathbb{R}^d$-valued processes Z with

$$\left(\int_0^T \left|\widehat{a}_t^{1/2} Z_t\right|^2 dt\right)^{\frac{p}{2}} < +\infty, \ \mathbb{P} - a.s.$$

$\mathbb{J}^p_{loc}(\mathbb{P})$ denotes the space of all $\mathbb{F}^+$-predictable functions U with

$$\left(\int_0^T \int_E |U_s(x)|^2 \, \widehat{\nu}_t(dx)ds\right)^{\frac{p}{2}} < +\infty, \ \mathbb{P} - a.s.$$

5.3.3 Formulation

We shall consider the following second order backward SDE with jumps (2BSDEJ for short), for $0 \leq t \leq T$ and $\mathcal{P}_H^\kappa$-q.s.

$$Y_t = \xi + \int_t^T \widehat{F}_s(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s^c - \int_t^T \int_E U_s(x)\tilde{\mu}_{B^d}(ds, dx) + K_T - K_t. \tag{5.3.3}$$

Definition 5.3.2 *We say $(Y, Z, U) \in \mathbb{D}_H^{2,\kappa} \times \mathbb{H}_H^{2,\kappa} \times \mathbb{J}_H^{2,\kappa}$ is a solution to 2BSDEJ (5.3.3) if*

- *$Y_T = \xi$, $\mathcal{P}_H^\kappa$-q.s.*
- *For all $\mathbb{P} \in P_H^\kappa$ and $0 \leq t \leq T$, the process $K^\mathbb{P}$ defined below is predictable and has non-decreasing paths $\mathbb{P} - a.s.$*

$$K_t^\mathbb{P} := Y_0 - Y_t - \int_0^t \widehat{F}_s(Y_s, Z_s, U_s)ds + \int_0^t Z_s dB_s^c + \int_0^t \int_E U_s(x)\tilde{\mu}_{B^d}(ds, dx). \tag{5.3.4}$$

- *The family $\left\{K^\mathbb{P}, \mathbb{P} \in \mathcal{P}_H^\kappa\right\}$ satisfies the minimum condition*

$$K_t^\mathbb{P} = \underset{\mathbb{P}' \in \mathcal{P}_H(t^+, \mathbb{P})}{\operatorname{ess\,inf}^\mathbb{P}} \ \mathbb{E}_t^{\mathbb{P}'}\left[K_T^{\mathbb{P}'}\right], \ 0 \leq t \leq T, \ \mathbb{P} - a.s., \ \forall \mathbb{P} \in \mathcal{P}_H^\kappa. \tag{5.3.5}$$

Moreover if the family $\left\{K^\mathbb{P}, \mathbb{P} \in \mathcal{P}_H^\kappa\right\}$ can be aggregated into a universal process K, we call (Y, Z, U, K) a solution of the 2BSDEJ (5.3.3).

Remark 5.3.5 *Since with our set $\mathcal{P}_H^\kappa$ we have the aggregation property of Theorem 5.2.1, and since the minimum condition (5.3.5) implies easily that the family $\left\{K^{\mathbb{P}}, \mathbb{P}\in\mathcal{P}_H^\kappa\right\}$ satisfies the consistency condition, we can apply Theorem 5.2.1 and find an aggregator for the family. This is different from [112], because we are working with a smaller set of probability measures. Therefore, from now on, we will supress the dependence in $\mathbb{P}$ of K, when it will be clear that we are dealing with a solution to a 2BSDE.*

Following [112], in addition to Assumption 5.3.1, we will always assume

Assumption 5.3.2 (i) $\mathcal{P}_H^\kappa$ is not empty.

(ii) The process $\widehat{F}^0$ satisfy the following integrability condition

$$\phi_H^{2,\kappa} := \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\operatorname*{ess\,sup}_{0\leq t\leq T}^{\mathbb{P}}\left(\mathbb{E}_t^{H,\mathbb{P}}\left[\int_0^T |\hat{F}_s^0|^\kappa ds\right]\right)^{\frac{2}{\kappa}}\right] < +\infty \tag{5.3.6}$$

5.3.4 Connection with standard backward SDEs with jumps

Let us assume that H is linear in γ and $\tilde{v}$, in the following sense

$$H_t(y,z,u,\gamma,\tilde{v}) := \frac{1}{2}\mathrm{Tr}\left[I_d\gamma\right] + \langle \tilde{v},\nu^*\rangle - f_t(y,z,u), \tag{5.3.7}$$

where $\nu^*\in\mathcal{N}$. We then have the following result

Lemma 5.3.1 *If H is of the form (5.3.7), then $D^1_{F_t}=\{I_d\}$, $D^2_{F_t}=\{\nu^*\}$ and*

$$F_t(\omega,y,z,u,a,\nu) = F_t(\omega,y,z,u,Id,\nu^*) = f_t(y,z,u).$$

Proof. First notice that

$$\begin{aligned}H_t(\omega,y,z,u,\gamma,\tilde{v}) = &\sup_{(a,\nu)\in\mathbb{S}_d^{>0}\times\mathcal{N}}\left\{\frac{1}{2}\mathrm{Tr}(a\gamma) + \int_0^T\int_E \tilde{v}(e)\nu_s(\omega)(ds,de) - \delta_{Id}(a) - \delta_{\nu^*}(\nu)\right\}\\ &- f_t(y,z,u).\end{aligned}$$

where δ_A denotes the characteristic function of a subset A in the convex analysis sense.

By definition of F, we get

$$F_t(\omega,y,z,u,a,\nu) = f_t(y,z,u) + H^{**}(a,\nu),$$

where H^{**} is the double Fenchel-Legendre transform of the function

$$(a,\nu)\mapsto \delta_{Id}(a)+\delta_{\nu^*}(\nu),$$

which is convex and lower-semicontinuous.

This then implies that

$$F_t(\omega,y,z,u,a,\nu) = f_t(y,z,u) + \delta_{Id}(a)+\delta_{\nu^*}(\nu),$$

which is the desired result. □

If we further assume that $\mathbb{E}^{\mathbb{P}_{\nu^*}}\left[\int_0^T |f_t(0,0,0)|^2 dt\right] < +\infty$, then $\mathcal{P}_H^\kappa = \{\mathbb{P}_{\nu^*}\}$ and the minimality condition on $K = K^{\mathbb{P}_{\nu^*}}$ implies that $0 = \mathbb{E}^{\mathbb{P}_{\nu^*}}[K_T]$, which means that $K = 0$, $\mathbb{P}_{\nu^*}$-a.s. and the 2BSDEJ is reduced to a classical backward SDE with jumps.

5.3.5 Connection with G-expectations and G-Lévy processes

5.3.5.1 Reminder on G-Lévy processes

In a recent paper [62], Hu and Peng introduced a new class of processes with independent and stationary increments, called G-Lévy processes. This processes are defined intrinsiquely, i.e. without making reference to any probability measure.

Let $\widetilde{\Omega}$ be a given set and let $\mathcal{H}$ be a linear space of real valued functions defined on $\widetilde{\Omega}$, containing the constants and such that $|X| \in \mathcal{H}$ if $X \in \mathcal{H}$. A sublinear expectation is a functional $\hat{\mathbb{E}} : \mathcal{H} \to \mathbb{R}$ which is monotone increasing, constant preserving, sub-additive and positively homogeneous. We refer to Definition 1.1 of [102] for more details. The triple $(\widetilde{\Omega}, \mathcal{H}, \hat{\mathbb{E}})$ is called a sublinear expectation space.

Definition 5.3.3 *A d-dimensional cadlag process $\{X_t,\, t \geq 0\}$ defined on a sublinear expectation space $(\widetilde{\Omega}, \mathcal{H}, \hat{\mathbb{E}})$ is called a G-Lévy process if:*

(i) $X_0 = 0$.

(ii) *X has independent increments: $\forall s, t > 0$, the random variable $(X_{t+s} - X_t)$ is independent from $(X_{t_1}, \dots, X_{t_n})$, for each $n \in \mathbb{N}$ and $0 \leq t_1 < \cdots < t_n \leq t$. The notion of independence used here corresponds to definition 3.10 in [102].*

(iii) *X has stationary increments: $\forall s, t > 0$, the distribution of $(X_{t+s} - X_t)$ does not depend on t. The notion of distribution used here corresponds to the definition given in §3 of [102].*

(iv) *For each $t \geq 0$, there exists a decomposition $X_t = X_t^c + X_t^d$, where $\{X_t^c,\, t \geq 0\}$ is a continuous process and $\{X_t^d,\, t \geq 0\}$ is a pure jump process.*

(v) *(X_t^c, X_t^d) is a $2d$-dimensional process satisfying conditions (i), (ii) and (iii) of this definition and*

$$\lim_{t \to 0^+} \frac{1}{t} \hat{\mathbb{E}}\left(|X_t^c|^3\right) = 0, \;\; \hat{\mathbb{E}}\left(\left|X_t^d\right|\right) \leq Ct,\, t \geq 0$$

for a real constant C.

In [62], Hu and Peng proved the following Lévy-Khintchine representation for G-Lévy processes:

Theorem 5.3.1 (Hu et Peng [62]) *Let $\{X_t,\, t \geq 0\}$ be a G-Lévy process. Then for each Lipschitz and bounded function φ, the function u defined by $u(t,x) := \hat{\mathbb{E}}\left(\varphi[x + X_t]\right)$ is the unique viscosity* *solution of the following partial integro-differential equation:*

$$\partial_t u(t,x) - \sup_{(b,\alpha,\nu)\in\mathcal{U}} \left\{ \int_E [u(t,x+z)-u(t,x)]\nu(dz) + < Du(t,x), B > \right.$$
$$\left. +\frac{1}{2}\mathrm{Tr}\left[D^2u(t,x)\alpha\alpha^T\right] \right\} = 0$$

where $\mathcal{U}$ is a subset of $\mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathcal{M}_R^+$ satisfying

$$\sup_{(b,\alpha,\nu)\in\mathcal{U}} \left\{ \int_{\mathbb{R}^d} |z|\,\nu(dz) + |b| + \mathrm{Tr}\left[\alpha\alpha^T\right] \right\} < +\infty$$

and where $\mathcal{M}_R^+$ denotes the set of positive Radon measures on E.

Notice that Hu and Peng study the case of G-Lévy processes with a discontinuous part that is of finite variation.

5.3.5.2 A connection with a particular 2BSDEJ

In our framework, we know that B^d is a purely discontinuous semimartingale of finite variation under $\mathbb{P}_\nu$ if $\int_0^T \int_{|x|\leq 1} |x|\,\nu_s(dx)ds < +\infty$, $\mathbb{P}_\nu$-a.s. We give a function H below, that is the natural candidate to retrieve the example of G-Lévy processes in our context.

Let $\tilde{\mathcal{N}}$ be any subset of $\mathcal{N}$ that is convex and closed for the weak topology on $\mathcal{M}_R^+$. We define

$$H_t(\omega,\gamma,\tilde{v}) := \sup_{(a,\nu)\in\mathbb{S}_d^{>0}\times\mathcal{N}} \left\{ \frac{1}{2}\mathrm{Tr}(a\gamma) + \int_0^T \int_E \tilde{v}(e)\nu_s(de)ds - \delta_{[a_1,a_2]}(a) - \delta_{\tilde{\mathcal{N}}}(\nu) \right\}.$$

Since $[a_1,a_2]$ and $\tilde{\mathcal{N}}$ are closed convex spaces, $F_t(\omega,a,\nu)$ is the double Fenchel-Legendre transform in (a,ν) of the convex and lower semi-continuous function $(a,\nu)\mapsto \delta_{[a_1,a_2]}(a)+\delta_{\tilde{\mathcal{N}}}(\nu)$ and then

$$F_t(\omega,a,\nu) = \delta_{[a_1,a_2]}(a)+\delta_{\tilde{\mathcal{N}}}(\nu).$$

If the second order BSDEs with jumps are connected to a class of fully non linear partial integro-differential equations, with this particular function H and its transform F, the PIDE one would expect to find is the one given in Theorem 5.3.1. If moreover $H_t(\omega,\gamma,\tilde{v}) = H_t(\omega,\tilde{v})$ is independent of γ, and $\tilde{\mathcal{N}} = \{\lambda\delta_{\{1\}},\ \lambda_1\leq\lambda\leq\lambda_2\}$ (which is convex and closed), then F_t is independent of a and we would obtain a 2BSDEJ giving a representation of the G-Poisson process.

5.4 Uniqueness result

5.4.1 Representation of the solution

We have similarly as in Theorem 4.4 of [112]

Theorem 5.4.1 *Let Assumptions 5.3.1 and 5.3.2 hold. Assume $\xi\in\mathbb{L}_H^{2,\kappa}$ and that (Y,Z,U) is a solution to 2BSDE with jumps (5.3.3). Then, for any $\mathbb{P}\in\mathcal{P}_H^\kappa$ and $0\leq t_1<t_2\leq T$,*

$$Y_{t_1} = \underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t_1^+, \mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}} \; y_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2}), \; \mathbb{P}-a.s., \tag{5.4.1}$$

where, for any $\mathbb{P} \in \mathcal{P}_H^\kappa$, $\mathbb{F}^+$*-stopping time* τ, *and* $\mathcal{F}_\tau^+$*-measurable random variable* $\xi \in \mathbb{L}^2(\mathbb{P})$, $(y^{\mathbb{P}}(\tau,\xi), z^{\mathbb{P}}(\tau,\xi))$ *denotes the solution to the following standard BSDE on* $0 \leq t \leq \tau$

$$y_t^{\mathbb{P}} = \xi - \int_t^\tau \widehat{F}_s(y_s^{\mathbb{P}}, z_s^{\mathbb{P}}, u_s^{\mathbb{P}})ds + \int_t^\tau z_s^{\mathbb{P}} dB_s^c + \int_t^\tau \int_E u_s^{\mathbb{P}}(x)\tilde{\mu}_{B^d}(ds,\, dx), \; \mathbb{P}-a.s. \tag{5.4.2}$$

Remark 5.4.1 *We first emphasize that existence and uniqueness results for the standard BSDEs (5.4.2) are not given directly by the existing literature, since the compensator of the counting measure associated to the jumps of* B *is not deterministic. However, since all the probability measure we consider satisfy the martingale representation property and the Blumenthal* $0-1$ *law, it is clear that we can straightforwardly generalize the proof of existence and uniqueness of Li and Tang [116] (see also [9] and [33] for related results). Furthermore, the usual a priori estimates and comparison Theorems will also hold.*

Remark 5.4.2 *It is worth noticing that, unlike in the case of* 2*BSDEs (see [112] for example), this representation does not imply directly the uniquemess of the solution in* $\mathbb{D}_H^{2,\kappa} \times \mathbb{H}_H^{2,\kappa} \times \mathbb{J}_H^{2,\kappa}$*. Indeed, by taking* $t_2 = T$ *in this representation formula, we have*

$$Y_t = \underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+, \mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}} \; y_t^{\mathbb{P}'}(T, \xi), \; t \in [0,T], \; \mathbb{P}-a.s., \; \textit{for all } \mathbb{P} \in \mathcal{P}_H^\kappa,$$

and thus Y *is unique.*

Then, since we have that $d\left\langle Y^c, B^c\right\rangle_t = Z_t d\left\langle B^c\right\rangle_t$, $\mathcal{P}_H^\kappa-q.s.$, Z *is unique. However, here we are not able to obtain that* U *and* $K^{\mathbb{P}}$ *are uniquely determined. Nonetheless, this representation is necessary to prove some a priori estimates in Theorem 5.4.4 which, as for the standard BSDEJs, insure the uniqueness of the solution.*

Before giving the proof of the above theorem, we first state the following Lemma which is a generalization of the usual comparison theorem proved by Royer (see Theorem 2.5 in [106]). Its proof is a straightforward generalization so we omit it.

Lemma 5.4.1 *Let* $\mathbb{P} \in \mathcal{P}_H^\kappa$. *We consider two generators* f^1 *and* f^2 *satisfying Assumption* H_{comp} *in [106] (which is a consequence of our more restrictive assumptions). Given two non-decreasing processes* k^1 *and* k^2, *let* ξ^1 *and* ξ^2 *be two terminal conditions for the following BSDEJs for* $i = 1, 2$,

$$y_t^i = \xi^i + \int_t^T f_s^i(y_s^i, z_s^i, u_s^i)ds - \int_t^T z_s^i dW_s - \int_t^T \int_E u_s^i(x)\tilde{\mu}(ds,\, dx) + k_T^i - k_t^i, \;\; \mathbb{P}-a.s.$$

Denote by (y^1, z^1, u^1) *and* (y^2, z^2, u^2) *the respective solutions. If* $\xi^1 \leq \xi^2$, $k^1 - k^2$ *is non-increasing and* $f^1(t, y_t^1, z_t^1, u_t^1) \leq f^2(t, y_t^1, z_t^1, u_t^1)$, *then* $\forall t \in [0,T], Y_t^1 \leq Y_t^2$.

Proof of Theorem 5.4.1 The proof follows the lines of the proof of Theorem 4.4 in [112].

(i) Fix $0 \leq t_1 < t_2 \leq T$ and $\mathbb{P} \in \mathcal{P}_H^\kappa$. For any $\mathbb{P}' \in \mathcal{P}_H^\kappa(t_1^+, \mathbb{P})$ and $t_1 \leq t \leq t_2$, we have,

$$Y_t = Y_{t_2} + \int_t^{t_2} \widehat{F}_s(Y_s, Z_s, U_s)ds - \int_t^{t_2} Z_s dB_s^c - \int_t^{t_2}\int_E U_s(x)\tilde{\mu}_{B^d}(ds, dx)$$
$$+ K_{t_2} - K_t,\ \mathbb{P}' - a.s.,$$

where K is the aggregator for the family of non-decreasing processes.

With Assumption 5.3.1, we can apply the above Lemma 5.4.1 under $\mathbb{P}'$ to obtain $Y_{t_1} \geq y_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2})$, $\mathbb{P}'-a.s.$. Since $\mathbb{P}' = \mathbb{P}$ on $\mathcal{F}_{t_1}^+$, we get $Y_{t_1} \geq y_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2})$, $\mathbb{P} - a.s.$ and thus

$$Y_{t_1} \geq \underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t_1^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ y_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2}),\ \mathbb{P} - a.s.$$

(ii) We now prove the reverse inequality. Fix $\mathbb{P} \in \mathcal{P}_H^\kappa$. We will show in (iii) below that

$$C_{t_1}^{\mathbb{P}} := \underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t_1^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_{t_1}^{\mathbb{P}'}\left[(K_{t_2} - K_{t_1})^2\right] < +\infty,\ \mathbb{P} - a.s.$$

For every $\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+, \mathbb{P})$, denote

$$\delta Y := Y - y^{\mathbb{P}'}(t_2, Y_{t_2}),\ \delta Z := Z - z^{\mathbb{P}'}(t_2, Y_{t_2}) \text{ and } \delta U := U - u^{\mathbb{P}'}(t_2, Y_{t_2}).$$

By the Lipschitz Assumption 5.3.1(iii), there exist two bounded processes λ and η such that for all $t_1 \leq t \leq t_2$,

$$\delta Y_t = \int_t^{t_2}\left(\lambda_s \delta Y_s + \eta_s \widehat{a}_s^{1/2}\delta Z_s\right)ds + \int_t^{t_2}\left(\widehat{F}_s(y_s^{\mathbb{P}'}, z_s^{\mathbb{P}'}, U_s) - \widehat{F}_s(y_s^{\mathbb{P}'}, z_s^{\mathbb{P}'}, u_s^{\mathbb{P}'})\right)ds$$
$$- \int_t^{t_2}\delta Z_s dB_s^c - \int_t^{t_2}\int_E \delta U_s(x)\tilde{\mu}_{B^d}(ds, dx) + K_{t_2} - K_t,\ \mathbb{P}' - a.s.$$

Define for $t_1 \leq t \leq t_2$ the following processes

$$N_t := \int_{t_1}^t \eta_s \widehat{a}_s^{-1/2} dB_s^c + \int_{t_1}^t\int_E \gamma_s(x)\tilde{\mu}_{B^d}(ds,\ dx),$$

and

$$M_t := \exp\left(\int_{t_1}^t \lambda_s ds\right)\mathcal{E}(N)_t,$$

where $\mathcal{E}(N)_t$ denotes the Doléans-Dade exponential martingale of N_t.

By the boundedness of λ and η and the assumption on γ in Assumption 5.3.1(iv), we know that M has moments (positive or negative) of any order (see [80] for the positive moments and Lemma B.2.4 in the Appendix for the negative ones). Thus we have for $p \geq 1$

$$\mathbb{E}_{t_1}^{\mathbb{P}'}\left[\sup_{t_1 \leq t \leq t_2} M_t^p + \sup_{t_1 \leq t \leq t_2} M_t^{-p}\right] \leq C_p,\ \mathbb{P}' - a.s. \tag{5.4.3}$$

Then, by Itô's formula, we obtain

$$\begin{aligned}
d(M_t\delta Y_t) =& M_{t-}d(\delta Y_t) + \delta Y_{t-}dM_t + d\left[M,\delta Y\right]_t \\
=& M_{t-}\left[\left(-\lambda_t\delta Y_t - \eta_t\widehat{a}_t^{1/2}\delta Z_t - \widehat{F}_t(y_t^{\mathbb{P}'},z_t^{\mathbb{P}'},U_t) + \widehat{F}_t(y_t^{\mathbb{P}'},z_t^{\mathbb{P}'},u_t^{\mathbb{P}'})\right)dt\right. \\
&\left. +\delta Z_t dB_t^c + \int_E\left(\delta U_t(x) + \gamma_t(x)\delta U_t(x)\right)\tilde{\mu}_{B^d}(dt,dx)\right] \\
&+ \delta Y_{t-}M_{t-}\left(\lambda_t dt + \eta_t\widehat{a}_t^{-1/2}dB_t^c + \int_E\gamma_t(x)\tilde{\mu}_{B^d}(dt,dx)\right) \\
&+ M_t\left(\eta_t\widehat{a}_t^{1/2}\delta Z_t dt + \int_E\gamma_t(x)\delta U_t(x)\widehat{\nu}_t(dx)dt\right) - M_{t-}dK_t.
\end{aligned}$$

Thus, by Assumption 5.3.1(iv), we have

$$\begin{aligned}
\delta Y_{t_1} \leq& -\int_{t_1}^{t_2} M_s\left(\delta Z_s + \delta Y_s\eta_s\widehat{a}_s^{-1/2}\right)dB_s^c + \int_{t_1}^{t_2} M_{s-}dK_s \\
&- \int_{t_1}^{t_2} M_{s-}\int_E\left(\delta U_s(x) + \delta Y_s\gamma_s(x) + \gamma_s(x)\delta U_s(x)\right)\tilde{\mu}_{B^d}(ds,dx).
\end{aligned}$$

By taking conditional expectation, we obtain

$$\delta Y_{t_1} \leq \mathbb{E}_{t_1}^{\mathbb{P}'}\left[\int_{t_1}^{t_2} M_{t-}dK_t\right]. \tag{5.4.4}$$

Applying the Hölder inequality, we can now write

$$\begin{aligned}
\delta Y_{t_1} &\leq \mathbb{E}_{t_1}^{\mathbb{P}'}\left[\sup_{t_1\leq t\leq t_2}(M_t)\left(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'}\right)\right] \\
&\leq \left(\mathbb{E}_{t_1}^{\mathbb{P}'}\left[\sup_{t_1\leq t\leq t_2}(M_t)^3\right]\right)^{1/3}\left(\mathbb{E}_{t_1}^{\mathbb{P}'}\left[\left(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'}\right)^{3/2}\right]\right)^{2/3} \\
&\leq C(C_{t_1}^{\mathbb{P}})^{1/3}\left(\mathbb{E}_{t_1}^{\mathbb{P}'}\left[K_{t_2} - K_{t_1}\right]\right)^{1/3},\ \mathbb{P}-a.s.
\end{aligned}$$

Taking the essential infimum on both sides finishes the proof.

(iii) It remains to show that the estimate for $C_{t_1}^{\mathbb{P}}$ holds. But by definition, and the Lipschitz Assumption on F we clearly have

$$\sup_{\mathbb{P}\in\mathcal{P}_H^\kappa}\mathbb{E}^{\mathbb{P}}\left[(K_{t_2} - K_{t_1})^2\right] \leq C\left(\|Y\|^2_{\mathbb{D}_H^{2,\kappa}} + \|Z\|^2_{\mathbb{H}_H^{2,\kappa}} + \|U\|^2_{\mathbb{J}_H^{2,\kappa}} + \phi_H^{2,\kappa}\right) < +\infty, \tag{5.4.5}$$

since the last term on the right-hand side is finite thanks to the integrability assumed on ξ and $\widehat{F}^0$. We then use the definition of the essential supremum (see Neveu [93] for example) to have the following equality

$$\operatorname*{ess\,sup}_{\mathbb{P}'\in\mathcal{P}_H^\kappa(t_1^+,\mathbb{P})}{}^{\mathbb{P}}\ \mathbb{E}_{t_1}^{\mathbb{P}'}\left[(K_{t_2} - K_{t_1})^2\right] = \sup_{n\geq 1}\mathbb{E}_{t_1}^{\mathbb{P}_n}\left[(K_{t_2} - K_{t_1})^2\right],\ \mathbb{P}-a.s. \tag{5.4.6}$$

for some sequence $(\mathbb{P}_n)_{n\geq 1} \subset \mathcal{P}_H^\kappa(t_1^+, \mathbb{P})$. Morevoer, in Lemma B.2.1 of the Appendix, it is proved that the set $\mathcal{P}_H^\kappa(t_1^+, \mathbb{P})$ is upward directed which means that for any $\mathbb{P}_1^{'}$, $\mathbb{P}_2^{'} \in \mathcal{P}_H^\kappa(t_1^+, \mathbb{P})$, there exists $\mathbb{P}^{'} \in \mathcal{P}_H^\kappa(t_1^+, \mathbb{P})$ such that

$$\mathbb{E}_{t_1}^{\mathbb{P}^{'}}\left[(K_{t_2}-K_{t_1})^2\right] = \max\left\{\mathbb{E}_{\mathrm{t}_1}^{\mathbb{P}_1^{'}}\left[(\mathrm{K}_{\mathrm{t}_2}-\mathrm{K}_{\mathrm{t}_1})^2\right],\ \mathbb{E}_{\mathrm{t}_1}^{\mathbb{P}_2^{'}}\left[(\mathrm{K}_{\mathrm{t}_2}-\mathrm{K}_{\mathrm{t}_1})^2\right]\right\}.$$

Hence, by using a subsequence if necessary, we can rewrite (5.4.6) as

$$\underset{\mathbb{P}^{'}\in\mathcal{P}_H^\kappa(t_1^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_{t_1}^{\mathbb{P}^{'}}\left[(K_{t_2}-K_{t_1})^2\right] = \lim_{n\to\infty}\uparrow \mathbb{E}_{t_1}^{\mathbb{P}_n}\left[(K_{t_2}-K_{t_1})^2\right],\ \mathbb{P}-a.s.$$

With (5.4.5), we can then finish the proof exactly as in the proof of Theorem 4.4 in [112]. □

Finally, the comparison Theorem below follows easily from the classical one for BSDEJs (see for instance Theorem 2.5 in [106]) and the representation (5.4.1).

Theorem 5.4.2 *Let (Y, Z, U) and (Y', Z', U') be the solutions of 2BSDEJs with terminal conditions ξ and $\xi^{'}$, generators $\widehat{F}$ and $\widehat{F}^{'}$ respectively (with the corresponding function H and $H^{'}$), and let $(y^{\mathbb{P}}, z^{\mathbb{P}}, u^{\mathbb{P}})$ and $(y'^{\mathbb{P}}, z'^{\mathbb{P}}, u'^{\mathbb{P}})$ the solutions of the associated BSDEJs. Assume that they both verify our Assumptions 5.3.1 and 5.3.2 and that we have*

- $\mathcal{P}_H^\kappa \subset \mathcal{P}_{H'}^\kappa$.
- $\xi \leq \xi^{'}$, $\mathcal{P}_H^\kappa - q.s.$
- $\widehat{F}_t(y_t'^{\mathbb{P}}, z_t'^{\mathbb{P}}, u_t'^{\mathbb{P}}) \leq \widehat{F}_t^{'}(y_t'^{\mathbb{P}}, z_t'^{\mathbb{P}}, u_t'^{\mathbb{P}})$, $\mathbb{P}-a.s.$, *for all* $\mathbb{P} \in \mathcal{P}_H^\kappa$.

Then $Y \leq Y'$, $\mathcal{P}_H^\kappa - q.s.$

5.4.2 A priori estimates and uniqueness of the solution

We conclude this section by showing some a priori estimates which not only will imply uniqueness of the solution of the 2BSDEJ (5.3.3), but will be useful to obtain the existence of a solution in the next chapter.

Theorem 5.4.3 *Let Assumptions 5.3.1 and 5.3.2 hold. Assume $\xi \in \mathbb{L}_H^{2,\kappa}$ and $(Y, Z, U) \in \mathbb{D}_H^{2,\kappa} \times \mathbb{H}_H^{2,\kappa} \times \mathbb{J}_H^{2,\kappa}$ is a solution to the 2BSDEJ (5.3.3). Let $\left\{(y^{\mathbb{P}}, z^{\mathbb{P}}, u^{\mathbb{P}})\right\}_{\mathbb{P}\in\mathcal{P}_H^\kappa}$ be the solutions of the corresponding BSDEJs (5.4.2). Then, there exists a constant C_κ such that*

$$\|Y\|^2_{\mathbb{D}_H^{2,\kappa}} + \|Z\|^2_{\mathbb{H}_H^{2,\kappa}} + \|U\|^2_{\mathbb{J}_H^{2,\kappa}} + \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa}\mathbb{E}^{\mathbb{P}}\left[|K_T|^2\right] \leq C_\kappa\left(\|\xi\|^2_{\mathbb{L}_H^{2,\kappa}} + \phi_H^{2,\kappa}\right)$$

$$\sup_{\mathbb{P}\in\mathcal{P}_H^\kappa}\left\{\left\|y^{\mathbb{P}}\right\|^2_{\mathbb{D}^2(\mathbb{P})} + \left\|z^{\mathbb{P}}\right\|^2_{\mathbb{H}^2(\mathbb{P})} + \left\|u^{\mathbb{P}}\right\|^2_{\mathbb{J}^2(\mathbb{P})}\right\} \leq C_\kappa\left(\|\xi\|^2_{\mathbb{L}_H^{2,\kappa}} + \phi_H^{2,\kappa}\right).$$

Proof. As in the proof of the representation formula in Theorem 5.4.1, the Lipschitz assumption 5.3.1(iii) of F implies that there exist two bounded processes λ and η such that for all t, and $\mathbb{P}-a.s.$

$$y_t^{\mathbb{P}} = \xi + \int_t^T \left(\lambda_s y_s^{\mathbb{P}} + \eta_s \widehat{a}_s^{1/2} z_s^{\mathbb{P}} + \widehat{F}_s(0, 0, u_s^{\mathbb{P}})\right) ds - \int_t^T z_s^{\mathbb{P}} dB_s^c - \int_t^T \int_E u_s^{\mathbb{P}}(x) \tilde{\mu}_{B^d}(ds, dx).$$

Define the following processes

$$N_t := \int_t^T \eta_s \widehat{a}_s^{-1/2} dB_s^c + \int_t^T \int_E \gamma_s(x) \tilde{\mu}_{B^d}(ds, dx), \text{ and } M_t := \exp\left(\int_t^T \lambda_s ds\right) \mathcal{E}(N)_t,$$

where $\mathcal{E}(N)_t$ denotes the Doléans-Dade exponential martingale of N_t. Then by applying Itô's formula to $M_t y_t^{\mathbb{P}}$, we obtain

$$y_t^{\mathbb{P}} = \mathbb{E}_t^{\mathbb{P}}\left[M_T \xi + \int_t^T M_s \widehat{F}_s(0, 0, u_s^{\mathbb{P}}) ds - \int_t^T \int_E M_s \gamma_s(x) u_s^{\mathbb{P}}(x) \widehat{\nu}_s(dx) ds\right].$$

Finally with Assumption (5.3.1)(iv), the Hölder inequality and the inequality (5.4.3), we conclude that there exists a constant C_κ depending only on κ, T and the Lipschitz constant of F, such that for all $\mathbb{P}$

$$\left|y_t^{\mathbb{P}}\right| \leq C_\kappa \mathbb{E}_t^{\mathbb{P}}\left[|\xi|^\kappa + \int_t^T \left|\widehat{F}_s^0\right|^\kappa ds\right]^{1/\kappa}. \tag{5.4.7}$$

This immediately provides the estimate for $y^{\mathbb{P}}$. Now by definition of our norms, we get from (5.4.7) and the representation formula (5.4.1) that

$$\|Y\|_{\mathbb{D}_H^{2,\kappa}}^2 \leq C_\kappa \left(\|\xi\|_{\mathbb{L}_H^{2,\kappa}}^2 + \phi_H^{2,\kappa}\right). \tag{5.4.8}$$

Now apply Itô's formula to $|Y|^2$ under each $\mathbb{P} \in \mathcal{P}_H^\kappa$. We get as usual for every $\varepsilon > 0$

$$\begin{aligned}
&|Y_0|^2 + \int_0^T \left|\widehat{a}_t^{1/2} Z_t\right|^2 dt + \int_0^T \int_E |U_t(x)|^2 \widehat{\nu}_t(dx) dt \\
&= |\xi|^2 + 2\int_0^T Y_t \widehat{F}_t(Y_t, Z_t, U_t) dt + 2\int_0^T Y_{t-} dK_t \\
&\quad - 2\int_0^T Y_t Z_t dB_t^c - \int_0^T \int_E \left(|U_t(x)|^2 + 2Y_{t-} U_t(x)\right) \tilde{\mu}_{B^d}(dt, dx) \\
&\leq 2\int_0^T |Y_t| \, |\widehat{F}_t(Y_t, Z_t, U_t)| dt + 2 \sup_{0 \leq t \leq T} |Y_t| \, K_T \\
&\quad - 2\int_0^T Y_t Z_t dB_t^c - \int_0^T \int_E \left(|U_t(x)|^2 + 2Y_{t-} U_t(x)\right) \tilde{\mu}_{B^d}(dt, dx).
\end{aligned}$$

By our assumptions on F, we have

$$\left|\widehat{F}_t(Y_t, Z_t, U_t)\right| \leq C\left(|Y_t| + \left|\widehat{a}_t^{1/2} Z_t\right| + \left|\widehat{F}_t^0\right| + \left(\int_E |U_t(x)|^2 \widehat{\nu}_t(dx)\right)^{1/2}\right).$$

With the usual inequality $2ab \leq \frac{1}{\varepsilon}a^2 + \varepsilon b^2, \forall \varepsilon > 0$, we obtain

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|\widehat{a}_t^{1/2} Z_t\right|^2 dt + \int_0^T \int_E |U_t(x)|^2 \,\widehat{\nu}_t(dx)dt\right] \\
&\leq C\mathbb{E}^{\mathbb{P}}\left[|\xi|^2 + \int_0^T |Y_t| \left(\left|\widehat{F}_t^0\right| + |Y_t| + \left|\widehat{a}_t^{1/2} Z_t\right| + \left(\int_E |U_t(x)|^2 \,\widehat{\nu}_t(dx)\right)^{1/2}\right) dt\right] \\
&\quad + \mathbb{E}^{\mathbb{P}}\left[\int_0^T |Y_{t-}|\, dK_t\right] \\
&\leq C\left(\|\xi\|_{\mathbb{L}_H^{2,\kappa}} + \mathbb{E}^{\mathbb{P}}\left[\left(1 + \frac{C}{\varepsilon}\right) \sup_{0\leq t\leq T} |Y_t|^2 + \left(\int_0^T \left|\widehat{F}_t^0\right| dt\right)^2\right]\right) \\
&\quad + \varepsilon\mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|\widehat{a}_t^{1/2} Z_t\right|^2 dt + \int_0^T \int_E |U_t(x)|^2 \,\widehat{\nu}_t(dx)dt + |K_T|^2\right]. \qquad (5.4.9)
\end{aligned}
$$

Then by definition of our 2BSDEJ, we easily have

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[|K_T|^2\right] \leq & C_0\mathbb{E}^{\mathbb{P}}\left[|\xi|^2 + \sup_{0\leq t\leq T} |Y_t|^2 + \int_0^T \left|\widehat{a}_t^{1/2} Z_t\right|^2 dt\right. \\
&\left. + \int_0^T \int_E |U_t(x)|^2 \,\widehat{\nu}_t(dx)dt + \left(\int_0^T \left|\widehat{F}_t^0\right| dt\right)^2\right], \qquad (5.4.10)
\end{aligned}
$$

for some constant C_0, independent of ε. Now set $\varepsilon := (2(1 + C_0))^{-1}$ and plug (5.4.10) in (5.4.9). One then gets

$$
\mathbb{E}^{\mathbb{P}}\left[\int_0^T |\widehat{a}_t^{\frac{1}{2}} Z_t|^2 dt + \int_0^T \int_E |U_t(x)|^2 \,\widehat{\nu}_t(dx)dt\right] \leq C\mathbb{E}^{\mathbb{P}}\left[|\xi|^2 + \sup_{0\leq t\leq T} |Y_t|^2 + \left(\int_0^T \left|\widehat{F}_t^0\right| dt\right)^2\right].
$$

From this and the estimate for Y, we immediately obtain

$$
\|Z\|_{\mathbb{H}_H^{2,\kappa}} + \|U\|_{\mathbb{J}_H^{2,\kappa}} \leq C\left(\|\xi\|^2_{\mathbb{L}_H^{2,\kappa}} + \phi_H^{2,\kappa}\right).
$$

Then the estimate for K follows from (5.4.10). The estimates for $z^{\mathbb{P}}$ and $u^{\mathbb{P}}$ can be proved similarly. □

Theorem 5.4.4 *Let Assumptions 5.3.1 and 5.3.2 hold. For $i = 1, 2$, let us consider the solutions (Y^i, Z^i, U^i, K^i) of the 2BSDEJs (5.3.3) with terminal condition ξ^i. Then, there exists a constant C_κ depending only on κ, T and the Lipschitz constant of F such that*

$$
\begin{aligned}
&\left\|Y^1 - Y^2\right\|_{\mathbb{D}_H^{2,\kappa}} \leq C_\kappa \left\|\xi^1 - \xi^2\right\|_{\mathbb{L}_H^{2,\kappa}} \\
&\left\|Z^1 - Z^2\right\|^2_{\mathbb{H}_H^{2,\kappa}} + \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\sup_{0\leq t\leq T} \left|K_t^1 - K_t^2\right|^2\right] + \left\|U^1 - U^2\right\|^2_{\mathbb{J}_H^{2,\kappa}} \\
&\leq C_\kappa \left\|\xi^1 - \xi^2\right\|_{\mathbb{L}_H^{2,\kappa}} \left(\left\|\xi^1\right\|_{\mathbb{L}_H^{2,\kappa}} + \left\|\xi^2\right\|_{\mathbb{L}_H^{2,\kappa}} + (\phi_H^{2,\kappa})^{1/2}\right).
\end{aligned}
$$

Consequently, the 2BSDEJ (5.3.3) has at most one solution in $\mathbb{D}_H^{2,\kappa} \times \mathbb{H}_H^{2,\kappa} \times \mathbb{J}_H^{2,\kappa}$.

Proof. As in the previous Theorem, we can obtain that there exists a constant C_κ depending only on κ, T and the Lipschitz constant of $\widehat{F}$, such that for all $\mathbb{P}$

$$\left|y_t^{\mathbb{P},1} - y_t^{\mathbb{P},2}\right| \leq C_\kappa \mathbb{E}_t^{\mathbb{P}}\left[\left|\xi^1 - \xi^2\right|^\kappa\right]^{1/\kappa}. \tag{5.4.11}$$

Now by definition, we get from (5.4.11) and the representation formula (5.4.1) that

$$\left\|Y^1 - Y^2\right\|^2_{\mathbb{D}^{2,\kappa}_H} \leq C_\kappa \left\|\xi^1 - \xi^2\right\|^2_{\mathbb{L}^{2,\kappa}_H}. \tag{5.4.12}$$

Applying Itô's formula to $\left|Y^1 - Y^2\right|^2$, under each $\mathbb{P} \in \mathcal{P}^\kappa_H$, leads to

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|\widehat{a}_t^{\frac{1}{2}}(Z_t^1 - Z_t^2)\right|^2 dt + \int_0^T \int_E \left|U_t^1(x) - U_t^2(x)\right|^2 \widehat{\nu}_t(dx)dt\right] \\
&\leq C\mathbb{E}^{\mathbb{P}}\left[\left|\xi^1 - \xi^2\right|^2\right] + \mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|Y_t^1 - Y_t^2\right| d(K_t^1 - K_t^2)\right] \\
&\quad + C\mathbb{E}^{\mathbb{P}}\Big[\int_0^T \left|Y_t^1 - Y_t^2\right| \Big(\left|Y_t^1 - Y_t^2\right| + |\widehat{a}_t^{\frac{1}{2}}(Z_t^1 - Z_t^2)| \\
&\quad + \left(\int_E \left|U_t^1(x) - U_t^2(x)\right|^2 \widehat{\nu}_t(dx)dt\right)^{1/2}\Big) dt\Big] \\
&\leq C\left(\left\|\xi^1 - \xi^2\right\|^2_{\mathbb{L}^{2,\kappa}_H} + \left\|Y^1 - Y^2\right\|^2_{\mathbb{D}^{2,\kappa}_H}\right) \\
&\quad + \frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|\widehat{a}_t^{1/2}(Z_t^1 - Z_t^2)\right|^2 dt + \int_0^T \int_E \left|U_t^1(x) - U_t^2(x)\right|^2 \widehat{\nu}_t(dx)dt\right] \\
&\quad + C\left\|Y^1 - Y^2\right\|_{\mathbb{D}^{2,\kappa}_H} \left(\mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^2 \left(K_T^i\right)^2\right]\right)^{\frac{1}{2}}.
\end{aligned}$$

The estimates for $(Z^1 - Z^2)$ and $(U^1 - U^2)$ are now obvious from the above inequality and the estimates of Theorem 5.4.3. Finally the estimate for the difference of the increasing processes is obvious by definition. $\square$

Chapter 6

Existence of 2BSDEs with Jumps and Applications

Contents

6.1 Introduction **159**

6.2 A direct existence argument **160**

6.2.1 Notations 160

6.2.2 Existence when ξ is in $\mathrm{UC_b}(\Omega)$ 163

6.2.3 Main result 172

6.2.4 An extension of the representation formula 172

6.3 Application to a robust utility maximization problem **175**

6.3.1 The Market 175

6.3.2 Solving the optimization problem with a Lipschitz 2BSDEJ 177

6.3.3 A link with a particular quadratic 2BSDEJ 180

6.1 Introduction

The results of the previous chapter allowed us to find an aggregated version $\widehat{\nu}$ of the compensator of the jump measure of the canonical process. Furthermore, it is always possible to define an aggregated version $\widehat{a}$ of the density with respect to the Lebesgue measure on $\mathbb{R}^+$ of the quadratic variation of the continuous part of the canonical process. Equipped with the pair $(\widehat{a}, \widehat{\nu})$, we were able to define a generator $\widehat{F}$ depending on both $\widehat{a}$ and $\widehat{\nu}$, and to give a good formulation of second order BSDEs with jumps (see equation (5.3.3)). With this definition, we proved some a priori estimates and a uniqueness result.

Our aim in this chapter is to pursue the study undertaken in the previous one. More precisely, we prove existence of a solution to equation (5.3.3) by a direct approach. We construct a solution inspired by the representation obtained in Theorem 5.4.1, and using the tool of regular conditional probability distributions. This gives complete wellposedness for second order BSDEs with jumps. As an application of these results, in the spirit of [59] and [87], we treat a robust exponential utility maximization problem in a market with jumps, under model uncertainty. The uncertainty affects both the volatility process and the jump measure compensator. We prove existence of an optimal strategy, and that the value function of the problem is the unique solution of a particular second order BSDE with jumps.

The last part of this study would be to establish a connection with partial integro-differential equations (PIDEs for short). Indeed, Soner, Touzi and Zhang proved in [112] that Markovian second order BSDEs, are connected in the continuous case to a class of parabolic fully non-linear PDEs. On the other hand, we know that solutions to standard Markovian BSDEJs provide viscosity solutions to some parabolic partial integro-differential equations whose non local operator is given by a quantity similar to $<\tilde{v},\nu>$ defined in (5.3.1) (see [4] for more details). Then in the Markovian case, second order BSDEs are the natural candidate for the probabilistic interpretation of fully non-linear PIDEs.

6.2 A direct existence argument

In the article [112], the main tool to prove existence of a solution is the so called regular conditional probability distributions of Stroock and Varadhan [113]. Indeed, these tools allow to give a pathwise construction for conditional expectations. Since, at least when the generator is null, the y component of the solution of a BSDE can be written as a conditional expectation, the r.c.p.d. allows us to construct solutions of BSDEs pathwise. In the previous chapter, we have identified a candidate solution to the 2BSDEJ as an essential supremum of solutions of classical BSDEs with jumps (see (5.4.1)). However those BSDEs with jumps are written under mutually singular probability measures. Hence, being able to construct them pathwise allows us to avoid the problems related to negligible-sets. In this section we will generalize the approach of [112] to the jump case.

6.2.1 Notations

For the convenience of the reader, we recall below some of the notations introduced in [112]. Remember that we are working in the Skorohod space $\Omega=\mathbb{D}\left([0,T],\mathbb{R}^d\right)$ endowed with the Skorohod metric which makes it a separable space.

- For $0\leq t\leq T$, we denote by $\Omega^t:=\left\{\omega\in\mathbb{D}\left([t,T],\mathbb{R}^d\right)\right\}$ the shifted canonical space of càdlàg paths on $[t,T]$ which are null at t, B^t the shifted canonical process. Let $\mathcal{N}^t$ be the set of measures ν on $\mathcal{B}(E)$ satisfying

$$\int_t^T\int_E(1\wedge|x|^2)\nu_s(dx)ds<+\infty \text{ and } \int_t^T\int_{|x|>1}x\nu_s(dx)ds<+\infty,\ \forall\tilde{\omega}\in\Omega^t, \tag{6.2.1}$$

and let $\mathcal{D}^t$ be the set of $\mathbb{F}^t$-progressively measurable processes α taking values in $\mathbb{S}_d^{>0}$ with $\int_t^T|\alpha_s|ds<+\infty$, for every $\tilde{\omega}\in\Omega^t$.

$\mathbb{F}^t$ is the filtration generated by B^t. We define similarly the continuous part of B^t, denoted $B^{t,c}$, its discontinuous part denoted $B^{t,d}$, the density of the quadratic variation of $B^{t,c}$, denoted $\widehat{a}^t$, and $\mu_{B^{t,d}}$ the counting measure associated to the jumps of B^t.

Exactly as in Section 5.2, we can define semimartingale problems and the corresponding probability measures. We then restrict ourselves to deterministic (α,ν) and we let $\tilde{\mathcal{A}}^t$ be the corresponding separable class of coefficients and $\mathcal{P}_{\tilde{\mathcal{A}}^t}$ the corresponding family of probability measures, which will be noted $\mathbb{P}^{t,\alpha,\nu}$. Then, this family also satisfies the aggregation property of Theorem 5.2.1 in Chapter 5, and we can define $\widehat{v}^t$, the aggregator of the predictable compensators of B^t.

- For $0 \leq s \leq t \leq T$ and $\omega \in \Omega^s$, we define the shifted path $\omega^t \in \Omega^t$ by

$$\omega_r^t := \omega_r - \omega_t, \ \forall r \in [t, T].$$

- For $0 \leq s \leq t \leq T$ and $\omega \in \Omega^s$, $\widetilde{\omega} \in \Omega^t$ we define the concatenation path $\omega \otimes_t \widetilde{\omega} \in \Omega^s$ by

$$(\omega \otimes_t \widetilde{\omega})(r) := \omega_r 1_{[s,t)}(r) + (\omega_t + \widetilde{\omega}_r) 1_{[t,T]}(r), \ \forall r \in [s, T].$$

- For $0 \leq s \leq t \leq T$ and a $\mathcal{F}_T^s$-measurable random variable ξ on Ω^s, for each $\omega \in \Omega^s$, we define the shifted $\mathcal{F}_T^t$-measurable random variable $\xi^{t,\omega}$ on Ω^t by

$$\xi^{t,\omega}(\widetilde{\omega}) := \xi(\omega \otimes_t \widetilde{\omega}), \ \forall \ \widetilde{\omega} \in \Omega^t.$$

Similarly, for an $\mathbb{F}^s$-progressively measurable process X on $[s, T]$ and $(t, \omega) \in [s, T] \times \Omega^s$, we can define the shifted process $\left\{X_r^{t,\omega}, r \in [t, T]\right\}$, which is $\mathbb{F}^t$-progressively measurable.

- For a $\mathbb{F}$-stopping time τ, we use the same simplification as [112]

$$\omega \otimes_\tau \widetilde{\omega} := \omega \otimes_{\tau(\omega)} \widetilde{\omega}, \ \xi^{\tau,\omega} := \xi^{\tau(\omega),\omega}, \ X^{\tau,\omega} := X^{\tau(\omega),\omega}.$$

- We define our "shifted" generator

$$\widehat{F}_s^{t,\omega}(\widetilde{\omega}, y, z, u) := F_s(\omega \otimes_t \widetilde{\omega}, y, z, u, \widehat{a}_s^t(\widetilde{\omega}), \widehat{\nu}_s^t(\widetilde{\omega})), \ \forall (s, \widetilde{\omega}) \in [t, T] \times \Omega^t.$$

Then note that since F is assumed to be uniformly continuous in ω for the Skorohod topology, then so is $\widehat{F}^{t,\omega}$. Notice that this implies that for any $\mathbb{P} \in \mathcal{P}_{\tilde{\mathcal{A}}^t}$

$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^T \left|\widehat{F}_s^{t,\omega}(0,0,0)\right|^\kappa ds\right)^{\frac{2}{\kappa}}\right] < +\infty,$$

for some ω if and only if it holds for all $\omega \in \Omega$.

- Finally, we extend Definition 5.3.1 in the shifted spaces

Definition 6.2.1 *$\mathcal{P}_H^{t,\kappa}$ consists of all $\mathbb{P} := \mathbb{P}^{t,\alpha,\nu} \in \mathcal{P}_{\tilde{\mathcal{A}}^t}$ such that*

$$\underline{a}^{\mathbb{P}} \leq \widehat{a}^t \leq \overline{a}^{\mathbb{P}}, \ dt \times d\mathbb{P} - a.s. \text{ for some } \underline{a}^{\mathbb{P}}, \overline{a}^{\mathbb{P}} \in \mathbb{S}_d^{>0}$$
$$\mathbb{E}^{\mathbb{P}}\left[\left(\int_t^T \left|\widehat{F}_s^{t,\omega}(0,0,0)\right|^\kappa ds\right)^{\frac{2}{\kappa}}\right] < +\infty, \text{ for all } \omega \in \Omega.$$

For given $\omega \in \Omega$, $\mathbb{F}$-stopping time τ and $\mathbb{P} \in \mathcal{P}_H^\kappa$, the r.c.p.d. of $\mathbb{P}$ is a probability measure $\mathbb{P}_\tau^\omega$ on $\mathcal{F}_T$ such that for every bounded $\mathcal{F}_T$-measurable random variable ξ

$$\mathbb{E}_\tau^{\mathbb{P}}[\xi](\omega) = \mathbb{E}^{\mathbb{P}_\tau^\omega}[\xi], \text{ for } \mathbb{P}\text{-a.e. } \omega.$$

Furthermore, $\mathbb{P}_\tau^\omega$ naturally induces a probability measure $\mathbb{P}^{\tau,\omega}$ on $\mathcal{F}_T^{\tau(\omega)}$ such that the $\mathbb{P}^{\tau,\omega}$-distribution of $B^{\tau(\omega)}$ is equal to the $\mathbb{P}_\tau^\omega$-distribution of $\left\{B_t - B_{\tau(\omega)}, \ t \in [\tau(\omega), t]\right\}$. Besides, we have

$$\mathbb{E}^{\mathbb{P}_\tau^\omega}[\xi] = \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\xi^{\tau,\omega}].$$

Remark 6.2.1 *We emphasize that the above notations correspond to the ones used in [112] when we consider the subset of Ω consisting of all continuous paths from $[0,T]$ to $\mathbb{R}^d$ whose value at time 0 is 0.*

We now prove the following Proposition which gives a relation between $(\widehat{a}^{t,\omega},\widehat{\nu}^{t,\omega})$ and $(\widehat{a}^t,\widehat{\nu}^t)$.

Proposition 6.2.1 *Let $\mathbb{P}\in\mathcal{P}_H^\kappa$ and τ be an $\mathbb{F}$-stopping time. Then, for $\mathbb{P}$-a.e. $\omega\in\Omega$, we have for $ds\times d\mathbb{P}^{\tau,\omega}$-a.e. $(s,\widetilde{\omega})\in[\tau(\omega),T]\times\Omega^{\tau(\omega)}$*

$$\widehat{a}_s^{\tau,\omega}(\widetilde{\omega})=\widehat{a}_s^{\tau(\omega)}(\widetilde{\omega}),\ \textit{and}\ \widehat{\nu}_s^{\tau,\omega}(\widetilde{\omega},A)=\widehat{\nu}_s^{\tau(\omega)}(\widetilde{\omega},A)\ \textit{for every}\ A\in\mathcal{B}(E).$$

This result is important for us, because it implies that for $\mathbb{P}$-a.e. $\omega\in\Omega$ and for $ds\times d\mathbb{P}^{t,\omega}-$ a.e. $(s,\widetilde{\omega})\in[t,T]\times\Omega^t$

$$F_s\left(\omega\otimes_t\widetilde{\omega},y,z,u,\widehat{a}_s(\omega\otimes_t\widetilde{\omega}),\widehat{\nu}_s(\omega\otimes_t\widetilde{\omega})\right)=F_s\left(\omega\otimes_t\widetilde{\omega},y,z,u,\widehat{a}_s^t(\widetilde{\omega}),\widehat{\nu}_s^t(\widetilde{\omega})\right).$$

Whereas the left-hand side has in general no regularity in ω, the right-hand side, that we choose as our shifted generator, is uniformly continuous in ω.

Proof. The proof of the equality for $\widehat{a}$ is the same as in Lemma 4.1 of [110], so we omit it.

Now, for $s\geq\tau$ and for any $A\in\mathcal{B}(E)$, we know by the Doob-Meyer decomposition that there exist a $\mathbb{P}$-local martingale M and a $\mathbb{P}^{\tau,\omega}$-martingale N such that

$$\mu_{B^d}([0,s],A)=M_s+\int_0^s\widehat{\nu}_r(A)dr,\ \mathbb{P}-a.s.$$
$$\mu_{B^{\tau(\omega),d}}([\tau(\omega),s],A)=N_s+\int_\tau^s\widehat{\nu}_r^{\tau(\omega)}(A)dr,\ \mathbb{P}^{\tau,\omega}-a.s.$$

Then, we can rewrite the first equation above for $\mathbb{P}$-a.e. $\omega\in\Omega$ and for $\mathbb{P}^{\tau,\omega}$-a.e. $\widetilde{\omega}\in\Omega^{\tau(\omega)}$

$$\mu_{B^d}(\omega\otimes_\tau\widetilde{\omega},[0,s],A)=M_s^{\tau,\omega}(\widetilde{\omega})+\int_0^s\widehat{\nu}_r^{\tau,\omega}(\widetilde{\omega},A)dr. \tag{6.2.2}$$

Now, by definition of the measures μ_{B^d} and $\mu_{B^{\tau(\omega),d}}$, we have

$$\mu_{B^d}(\omega\otimes_\tau\widetilde{\omega},[0,s],A)=\mu_{B^d}(\omega,[0,\tau],A)+\mu_{B^{\tau(\omega),d}}(\widetilde{\omega},[\tau,s],A).$$

Hence, we obtain from (6.2.2) that for $\mathbb{P}$-a.e. $\omega\in\Omega$ and for $\mathbb{P}^{\tau,\omega}$-a.e. $\widetilde{\omega}\in\Omega^{\tau(\omega)}$

$$\mu_{B^d}(\omega,[0,\tau],A)-\int_0^\tau\widehat{\nu}_r(\omega,A)dr+N_s(\widetilde{\omega})-M_s^{\tau,\omega}(\widetilde{\omega})=\int_\tau^s\left(\widehat{\nu}_r^{\tau,\omega}(\widetilde{\omega},A)-\widehat{\nu}_r^{\tau(\omega)}(\widetilde{\omega},A)\right)dr$$

In the left-hand side above, the terms which are $\mathcal{F}_\tau$-measurable are constants in $\Omega^{\tau(\omega)}$ and using the same arguments as in Step 1 of the proof of Lemma B.1.2, we can show that $M^{\tau,\omega}$ is a $\mathbb{P}^{\tau,\omega}$-local martingale for $\mathbb{P}$-a.e. $\omega\in\Omega$. This means that the left-hand side is a $\mathbb{P}^{\tau,\omega}$-local martingale while the right-hand side is a predictable finite variation process. By the martingale representation property which still holds in the shifted canonical spaces, we deduce that for $\mathbb{P}$-a.e. $\omega\in\Omega$ and for $ds\times d\mathbb{P}^{\tau,\omega}$-a.e. $(s,\widetilde{\omega})\in[\tau(\omega),T]\times\Omega^{\tau(\omega)}$

$$\int_\tau^s\left(\widehat{\nu}_r^{\tau,\omega}(\widetilde{\omega},A)-\widehat{\nu}_r^{\tau(\omega)}(\widetilde{\omega},A)\right)dr=0,$$

which is the desired result. □

6.2.2 Existence when ξ is in $\mathrm{UC_b}(\Omega)$

When ξ is in $\mathrm{UC_b}(\Omega)$, we know that there exists a modulus of continuity function ρ for ξ and F in ω. Then, for any $0 \leq t \leq s \leq T$, $(y, z, \nu) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{V}$ and $\omega, \omega' \in \Omega$, $\widetilde{\omega} \in \Omega^t$,

$$\left|\xi^{t,\omega}(\widetilde{\omega}) - \xi^{t,\omega'}(\widetilde{\omega})\right| \leq \rho\left(d_{S,t}(\omega, \omega')\right), \ \left|\widehat{F}_s^{t,\omega}(\widetilde{\omega}, y, z, u) - \widehat{F}_s^{t,\omega'}(\widetilde{\omega}, y, z, u)\right| \leq \rho\left(d_{S,t}(\omega, \omega')\right),$$

where $d_{S,t}(\omega, \omega') := d_S(\omega_{\cdot\wedge t}, \omega'_{\cdot\wedge t})$, d_S being the Skorohod distance.

We then define for all $\omega \in \Omega$

$$\Lambda(\omega) := \sup_{0 \leq t \leq T} \Lambda_t(\omega) := \sup_{0 \leq t \leq T} \sup_{\mathbb{P} \in \mathcal{P}_H^t} \left(\mathbb{E}^{\mathbb{P}}\left[\left|\xi^{t,\omega}\right|^2 + \int_t^T |\widehat{F}_s^{t,\omega}(0,0,0)|^2 ds\right]\right)^{1/2}. \tag{6.2.3}$$

Now since $\widehat{F}^{t,\omega}$ is also uniformly continuous in ω, it is easily verified that

$$\Lambda(\omega) < \infty \text{ for some } \omega \in \Omega \text{ iff it holds for all } \omega \in \Omega. \tag{6.2.4}$$

Moreover, when Λ is finite, it is uniformly continuous in ω for the Skorohod topology and is therefore $\mathcal{F}_T$-measurable. Now, by Assumption 5.3.2, we have

$$\Lambda_t(\omega) < \infty \text{ for all } (t, \omega) \in [0, T] \times \Omega. \tag{6.2.5}$$

To prove existence, we define the following value process V_t pathwise

$$V_t(\omega) := \sup_{\mathbb{P} \in \mathcal{P}_H^{t,\kappa}} \mathcal{Y}_t^{\mathbb{P},t,\omega}(T, \xi), \text{ for all } (t, \omega) \in [0, T] \times \Omega, \tag{6.2.6}$$

where, for any $(t_1, \omega) \in [0, T] \times \Omega$, $\mathbb{P} \in \mathcal{P}_H^{t_1,\kappa}, t_2 \in [t_1, T]$, and any $\mathcal{F}_{t_2}$-measurable $\eta \in \mathbb{L}^2(\mathbb{P})$, we denote $\mathcal{Y}_{t_1}^{\mathbb{P},t_1,\omega}(t_2, \eta) := y_{t_1}^{\mathbb{P},t_1,\omega}$, where $\left(y^{\mathbb{P},t_1,\omega}, z^{\mathbb{P},t_1,\omega}, u^{\mathbb{P},t_1,\omega}\right)$ is the solution of the following BSDE with jumps on the shifted space Ω^{t_1} under $\mathbb{P}$

$$\begin{aligned} y_s^{\mathbb{P},t_1,\omega} &= \eta^{t_1,\omega} + \int_s^{t_2} \widehat{F}_r^{t_1,\omega}\left(y_r^{\mathbb{P},t_1,\omega}, z_r^{\mathbb{P},t_1,\omega}, \nu\right) dr - \int_s^{t_2} z_r^{\mathbb{P},t_1,\omega} dB_r^{t_1,c} \\ &\quad - \int_s^{t_2} \int_{\mathbb{R}^d} u_s^{\mathbb{P},t_1,\omega}(x) \widetilde{\mu}_{B^{t_1,d}}(ds, dx), \ \mathbb{P}-a.s., \ s \in [t, T], \end{aligned} \tag{6.2.7}$$

where as usual $\widetilde{\mu}_{B^{t_1,d}}(ds, dx) := \mu_{B^{t_1,d}}(ds, dx) - \widehat{\nu}_s^t(dx)ds$. In view of the Blumenthal zero-one law, $y_t^{\mathbb{P},t,\omega}$ is constant for any given (t, ω) and $\mathbb{P} \in \mathcal{P}_H^{t,\kappa}$, and therefore the value process V is well defined. Let us now show that V inherits some properties from ξ and F.

Lemma 6.2.1 *Let Assumptions 5.3.1 and 5.3.2 hold and consider some ξ in $\mathrm{UC_b}(\Omega)$. Then for all $(t, \omega) \in [0, T] \times \Omega$ we have $|V_t(\omega)| \leq C\Lambda_t(\omega)$. Moreover, for all $(t, \omega, \omega') \in [0, T] \times \Omega^2$, $|V_t(\omega) - V_t(\omega')| \leq C\rho(d_{S,t}(\omega, \omega'))$. Consequently, V_t is $\mathcal{F}_t$-measurable for every $t \in [0, T]$.*

Proof. (i) For each $(t, \omega) \in [0, T] \times \Omega$ and $\mathbb{P} \in \mathcal{P}_H^{t,\kappa}$, let α be some positive constant which will be fixed later and let $\eta \in (0, 1)$. Since F is uniformly Lipschitz in (y, z) and satisfies Assumption 5.3.1(iv), we have

$$\left|\widehat{F}_s^{t,\omega}(y, z, u)\right| \leq \left|\widehat{F}_s^{t,\omega}(0, 0, 0)\right| + C\left(|y| + \left|\left(\widehat{a}_s^t\right)^{1/2} z\right| + \left(\int_E |u(x)|^2 \widehat{\nu}_s^t(dx)\right)^{1/2}\right).$$

Now apply Itô's formula. We obtain

$$
\begin{aligned}
& e^{\alpha t}\left|y_t^{\mathbb{P},t,\omega}\right|^2+\int_t^T e^{\alpha s}\left|(\widehat{a}_s^t)^{1/2} z_s^{\mathbb{P},t,\omega}\right|^2 ds+\int_t^T\int_E e^{\alpha s}\left|u_s^{\mathbb{P},t,\omega}(x)\right|^2\widehat{\nu}_s^t(dx)ds \\
&= e^{\alpha T}\left|\xi^{t,\omega}\right|^2+2\int_t^T e^{\alpha s} y_s^{\mathbb{P},t,\omega}\widehat{F}_s^{t,\omega}(y_s^{\mathbb{P},t,\omega},z_s^{\mathbb{P},t,\omega},u_s^{\mathbb{P},t,\omega})ds \\
&\quad -\alpha\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}\right|^2 ds-2\int_t^T e^{\alpha s}y_{s^-}^{\mathbb{P},t,\omega}z_s^{\mathbb{P},t,\omega}dB_s^{t,c} \\
&\quad -\int_t^T\int_E e^{\alpha s}\left(2y_{s^-}^{\mathbb{P},t,\omega}u_s^{\mathbb{P},t,\omega}(x)+\left|u_s^{\mathbb{P},t,\omega}(x)\right|^2\right)\widetilde{\mu}_{B^{t,d}}(ds,dx) \\
&\leq e^{\alpha T}\left|\xi^{t,\omega}\right|^2+\int_t^T e^{\alpha s}\left|\widehat{F}_s^{t,\omega}(0,0,0)\right|^2 ds+\left(1+2C+\frac{2C^2}{\eta}-\alpha\right)\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}\right|^2 ds \\
&\quad +\eta\int_t^T e^{\alpha s}\left|(\widehat{a}_s^t)^{1/2}z_s^{\mathbb{P},t,\omega}\right|^2 ds+\eta\int_t^T\int_E e^{\alpha s}\left|u_s^{\mathbb{P},t,\omega}(x)\right|^2\widehat{\nu}_s^t(dx)ds \\
&\quad -2\int_t^T e^{\alpha s}y_{s^-}^{\mathbb{P},t,\omega}z_s^{\mathbb{P},t,\omega}dB_s^{t,c}-\int_t^T\int_E e^{\alpha s}\left(2y_{s^-}^{\mathbb{P},t,\omega}u_s^{\mathbb{P},t,\omega}(x)+\left|u_s^{\mathbb{P},t,\omega}(x)\right|^2\right)\widetilde{\mu}_{B^{t,d}}(ds,dx).
\end{aligned}
$$

Now choose $\eta=1/2$ for instance and α large enough. By taking expectation we obtain easily

$$
\left|y_t^{\mathbb{P},t,\omega}\right|^2\leq C\left|\Lambda_t(\omega)\right|^2.
$$

The result then follows from the arbitrariness of $\mathbb{P}$.

(ii) The proof is exactly the same as above, except that one has to use uniform continuity in ω of $\xi^{t,\omega}$ and $\widehat{F}^{t,\omega}$. Indeed, for each $(t,\omega)\in[0,T]\times\Omega$ and $\mathbb{P}\in\mathcal{P}_H^t$, let α be some positive constant which will be fixed later and let $\eta\in(0,1)$. By Itô's formula we have, since $\widehat{F}$ is uniformly Lipschitz

$$
\begin{aligned}
& e^{\alpha t}\left|y_t^{\mathbb{P},t,\omega}-y_t^{\mathbb{P},t,\omega'}\right|^2+\int_t^T e^{\alpha s}\left(\left|(\widehat{a}_s^t)^{\frac{1}{2}}(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})\right|^2+\int_E e^{\alpha s}(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})^2(x)\widehat{\nu}_s^t(dx)\right)ds \\
&\leq e^{\alpha T}\left|\xi^{t,\omega}-\xi^{t,\omega'}\right|^2+2C\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|^2 ds \\
&\quad +2C\int_t^T\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|\left|(\widehat{a}_s^t)^{\frac{1}{2}}(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})\right|ds \\
&\quad +2C\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|\left(\int_{\mathbb{R}^d}\left|u_s^{\mathbb{P},t,\omega}(x)-u_s^{\mathbb{P},t,\omega'}(x)\right|^2\widehat{\nu}_s^t(dx)\right)^{1/2}ds \\
&\quad +2C\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|\left|\widehat{F}_s^{t,\omega}(y_s^{\mathbb{P},t,\omega},z_s^{\mathbb{P},t,\omega},u_s^{\mathbb{P},t,\omega})-\widehat{F}_s^{t,\omega'}(y_s^{\mathbb{P},t,\omega},z_s^{\mathbb{P},t,\omega},u_s^{\mathbb{P},t,\omega})\right|ds \\
&\quad -\alpha\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|^2 ds-2\int_t^T e^{\alpha s}(y_{s^-}^{\mathbb{P},t,\omega}-y_{s^-}^{\mathbb{P},t,\omega'})(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})dB_s^{t,c} \\
&\quad -\int_t^T\int_E e^{\alpha s}\left(2(y_{s^-}^{\mathbb{P},t,\omega}-y_{s^-}^{\mathbb{P},t,\omega'})(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})+(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})^2\right)(x)\widetilde{\mu}_{B^{t,d}}(ds,dx).
\end{aligned}
$$

We then deduce

$$
\begin{aligned}
&e^{\alpha t}\left|y_t^{\mathbb{P},t,\omega}-y_t^{\mathbb{P},t,\omega'}\right|^2+\int_t^T e^{\alpha s}\left(\left|(\widehat{a}_s^t)^{\frac{1}{2}}(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})\right|^2+\int_E e^{\alpha s}(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})^2(x)\widehat{\nu}_s^t(dx)\right)ds\\
&\leq e^{\alpha T}\left|\xi^{t,\omega}-\xi^{t,\omega'}\right|^2+\int_t^T e^{\alpha s}\left|\widehat{F}_s^{t,\omega}(y_s^{\mathbb{P},t,\omega},z_s^{\mathbb{P},t,\omega},u_s^{\mathbb{P},t,\omega})-\widehat{F}_s^{t,\omega'}(y_s^{\mathbb{P},t,\omega},z_s^{\mathbb{P},t,\omega},u_s^{\mathbb{P},t,\omega})\right|^2 ds\\
&+\eta\int_t^T e^{\alpha s}\left|(\widehat{a}_s^t)^{1/2}(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})\right|^2 ds+\eta\int_t^T\int_E e^{\alpha s}\left|u_s^{\mathbb{P},t,\omega}(x)-u_s^{\mathbb{P},t,\omega'}(x)\right|^2\widehat{\nu}_s^t(dx)ds\\
&+\left(2C+C^2+\frac{2C^2}{\eta}-\alpha\right)\int_t^T e^{\alpha s}\left|y_s^{\mathbb{P},t,\omega}-y_s^{\mathbb{P},t,\omega'}\right|^2 ds\\
&-2\int_t^T e^{\alpha s}(y_{s^-}^{\mathbb{P},t,\omega}-y_{s^-}^{\mathbb{P},t,\omega'})(z_s^{\mathbb{P},t,\omega}-z_s^{\mathbb{P},t,\omega'})dB_s^{t,c}\\
&-\int_t^T\int_E e^{\alpha s}\left(2(y_{s^-}^{\mathbb{P},t,\omega}-y_{s^-}^{\mathbb{P},t,\omega'})(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})+(u_s^{\mathbb{P},t,\omega}-u_s^{\mathbb{P},t,\omega'})^2\right)(x)\widetilde{\mu}_{B^{t,d}}(ds,dx).
\end{aligned}
$$

Now choose $\eta=1/2$ and α such that $\nu:=\alpha-2C-C^2-\frac{2C^2}{\eta}\geq 0$. We obtain the desired result by taking expectation and using the uniform continuity in ω of ξ and F. □

The next proposition is a dynamic programming property verified by the value process, which will prove useful when proving that V provides a solution to the 2BSDEJ with generator F and terminal condition ξ. The result and its proof are intimately connected and use the same arguments as the proof of Proposition 4.7 in [110].

Proposition 6.2.2 *Under Assumptions 5.3.1, 5.3.2 and for $\xi\in \mathrm{UC_b}(\Omega)$, we have for all $0\leq t_1<t_2\leq T$ and for all $\omega\in\Omega$*

$$V_{t_1}(\omega)=\sup_{\mathbb{P}\in\mathcal{P}_H^{t_1,\kappa}}\mathcal{Y}_{t_1}^{\mathbb{P},t_1,\omega}(t_2,V_{t_2}^{t_1,\omega}).$$

The proof is almost the same as the proof in [110], with minor modifications due to the introduction of jumps.

Proof. Without loss of generality, we assume that $t_1=0$ and $t_2=t$. Thus, we have to prove

$$V_0(\omega)=\sup_{\mathbb{P}\in\mathcal{P}_H^{\kappa}}\mathcal{Y}_0^{\mathbb{P}}(t,V_t).$$

Denote $(y^{\mathbb{P}},z^{\mathbb{P}},u^{\mathbb{P}}):=(\mathcal{Y}^{\mathbb{P}}(T,\xi),\mathcal{Z}^{\mathbb{P}}(T,\xi),\mathcal{U}^{\mathbb{P}}(T,\xi))$

(i) For any $\mathbb{P}\in\mathcal{P}_H^{\kappa}$, we know by Lemma B.1.2 in the Appendix, that for $\mathbb{P}-a.e.\ \omega\in\Omega$, the r.c.p.d. $\mathbb{P}^{t,\omega}\in\mathcal{P}_H^{t,\kappa}$. Now thanks to the paper of Li and Tang [116], we know that the solution of BSDEs on the Wiener-Poisson space with Lipschitz generator can be constructed via Picard iteration. Thus, it means that at each step of the iteration, the solution can be formulated as a conditional expectation under $\mathbb{P}$. By the properties of the r.p.c.d., this entails that

$$y_t^{\mathbb{P}}(\omega)=\mathcal{Y}_t^{\mathbb{P}^{t,\omega},t,\omega}(T,\xi),\ \text{ for }\mathbb{P}-a.e.\ \omega\in\Omega. \tag{6.2.8}$$

Hence, by definition of V_t and the comparison principle for BSDEs with jumps, we get that $y_0^{\mathbb{P}} \leq \mathcal{Y}_0^{\mathbb{P}}(t, V_t)$. By arbitrariness of $\mathbb{P}$, this leads to

$$V_0(\omega) \leq \sup_{\mathbb{P} \in \mathcal{P}_H^\kappa} \mathcal{Y}_0^{\mathbb{P}}(t, V_t).$$

(ii) For the other inequality, we proceed as in [110]. Let $\mathbb{P} \in \mathcal{P}_H^\kappa$ and $\varepsilon > 0$. By separability of Ω, there exists a partition $(E_t^i)_{i\geq 1} \subset \mathcal{F}_t$ such that $d_S(\omega, \omega')_t \leq \varepsilon/2$ for any i and any $\omega, \omega' \in E_t^i$.

Now for each i, fix a $\widehat{\omega}_i \in E_t^i$ and let $\mathbb{P}_t^i$ be an $\varepsilon-$optimizer of $V_t(\widehat{\omega}_i)$. If we define for each $n \geq 1$, $\mathbb{P}^n := \mathbb{P}^{n,\varepsilon}$ by

$$\mathbb{P}^n(E) := \mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^n \mathbb{E}^{\mathbb{P}_t^i}\left[1_E^{t,\omega}\right] 1_{E_t^i}\right] + \mathbb{P}(E \cap \widehat{E}_t^n), \text{ where } \widehat{E}_t^n := \cup_{i>n} E_t^i, \tag{6.2.9}$$

then, by Lemma B.1.3, we know that $\mathbb{P}^n \in \mathcal{P}_H^\kappa$. Besides, by Lemma 6.2.1 and its proof, we have for any i and any $\omega \in E_t^i$

$$\begin{aligned} V_t(\omega) &\leq V_t(\widehat{\omega}_i) + C\rho(\varepsilon) \leq \mathcal{Y}_t^{\mathbb{P}_t^i, t, \widehat{\omega}_i}(T, \xi) + \varepsilon + C\rho(\varepsilon) \\ &\leq \mathcal{Y}_t^{\mathbb{P}_t^i, t, \omega}(T, \xi) + \varepsilon + C\rho(\varepsilon) = \mathcal{Y}_t^{(\mathbb{P}^n)^{t,\omega}, t, \omega}(T, \xi) + \varepsilon + C\rho(\varepsilon), \end{aligned}$$

where we used successively the uniform continuity of V in ω, the definition of $\mathbb{P}_t^i$, the uniform continuity of $\mathcal{Y}_t^{\mathbb{P}, t, \omega}$ in ω and finally the definition of $\mathbb{P}^n$.

Then, it follows from (6.2.8) that

$$V_t \leq y_t^{\mathbb{P}^n} + \varepsilon + C\rho(\varepsilon), \ \mathbb{P}^n - a.s. \text{ on } \cup_{i=1}^n E_t^i. \tag{6.2.10}$$

Let now $(y^n, z^n, u^n) := (y^{n,\varepsilon}, z^{n,\varepsilon}, u^{n,\varepsilon})$ be the solution of the following BSDE on $[0, t]$

$$\begin{aligned} y_s^n = &\left[y_t^{\mathbb{P}^n} + \varepsilon + C\rho(\varepsilon)\right] 1_{\cup_{i=1}^n E_t^i} + V_t 1_{\widehat{E}_t^n} + \int_s^t \widehat{F}_r(y_r^n, z_r^n, u_r^n) dr - \int_s^t z_r^n dB_r^c \\ &- \int_s^t \int_E u_r^n(x) \widetilde{\mu}_{B^d}(dr, dx), \ \mathbb{P}^n - a.s. \end{aligned} \tag{6.2.11}$$

By the comparison principle for BSDEs with jumps, we know that $\mathcal{Y}_0^{\mathbb{P}}(t, V_t) \leq y_0^n$. Then since $\mathbb{P}^n = \mathbb{P}$ on $\mathcal{F}_t$, the equality (6.2.11) also holds $\mathbb{P} - a.s.$ Using the same arguments and notations as in the proof of Lemma 6.2.1, we obtain

$$\left|y_0^n - y_0^{\mathbb{P}^n}\right|^2 \leq C\mathbb{E}^{\mathbb{P}}\left[\varepsilon^2 + \rho(\varepsilon)^2 + \left|V_t - y_t^{\mathbb{P}^n}\right|^2 1_{\widehat{E}_t^n}\right].$$

Then, by Lemma 6.2.1, we have

$$\mathcal{Y}_0^{\mathbb{P}}(t, V_t) \leq y_0^n \leq y_0^{\mathbb{P}^n} + C\left(\varepsilon + \rho(\varepsilon) + \left(\mathbb{E}^{\mathbb{P}}\left[\Lambda_t^2 1_{\widehat{E}_t^n}\right]\right)^{\frac{1}{2}}\right) \leq V_0 + C\left(\varepsilon + \rho(\varepsilon) + \left(\mathbb{E}^{\mathbb{P}}\left[\Lambda_t^2 1_{\widehat{E}_t^n}\right]\right)^{\frac{1}{2}}\right).$$

Then it suffices to let n go to $+\infty$, use the dominated convergence theorem, and let ε go to 0. $\square$

Now we are facing the problem of the regularity in t of V. Indeed, if we want to obtain a solution of the 2BSDE, then it has to be at least càdlàg, $\mathcal{P}_H^\kappa - q.s.$ To this end, we define now for all (t,ω), the $\mathbb{F}^+$-progressively measurable process

$$V_t^+ := \overline{\lim_{r\in\mathbb{Q}\cap(t,T],r\downarrow t}} V_r.$$

Lemma 6.2.2 *Under the conditions of the previous Proposition, we have*

$$V_t^+ = \lim_{r\in\mathbb{Q}\cap(t,T],r\downarrow t} V_r,\ \mathcal{P}_H^\kappa - q.s.$$

and thus V^+ is càdlàg $\mathcal{P}_H^\kappa - q.s.$

Proof. For each $\mathbb{P}$, we define $\widetilde{V}^\mathbb{P} := V - \mathcal{Y}^\mathbb{P}(T,\xi)$. Then, we recall that we have

$$\widetilde{V}^\mathbb{P} \geq 0,\ \mathbb{P} - a.s.$$

Now for any $0 \leq t_1 < t_2 \leq T$, let $(y^{\mathbb{P},t_2}, z^{\mathbb{P},t_2}, u^{\mathbb{P},t_2}) := (\mathcal{Y}^\mathbb{P}(t_2,V_{t_2}), \mathcal{Z}^\mathbb{P}(t_2,V_{t_2}), \mathcal{U}^\mathbb{P}(t_2,V_{t_2}))$. Once more, we remind that since solutions of BSDEs can be defined by Picard iterations, we have by the properties of the r.p.c.d. that

$$\mathcal{Y}_{t_1}^\mathbb{P}(t_2,V_{t_2})(\omega) = \mathcal{Y}_{t_1}^{\mathbb{P}^{t_1,\omega},t_1,\omega}(t_2,V_{t_2}^{t_1,\omega}),\ \text{for } \mathbb{P} - a.e.\ \omega.$$

Hence, we conclude from Proposition 6.2.2 $V_{t_1} \geq y_{t_1}^{\mathbb{P},t_2},\ \mathbb{P} - a.s.$ Denote

$$\widetilde{y}_t^{\mathbb{P},t_2} := y_t^{\mathbb{P},t_2} - \mathcal{Y}_t^\mathbb{P}(T,\xi),\ \widetilde{z}_t^{\mathbb{P},t_2} := \widehat{a}_t^{-1/2}(z_t^{\mathbb{P},t_2} - \mathcal{Z}_t^\mathbb{P}(T,\xi)),\ \widetilde{u}_t^{\mathbb{P},t_2} := u_t^{\mathbb{P},t_2} - \mathcal{U}_t^\mathbb{P}(T,\xi).$$

Then $\widetilde{V}_{t_1}^\mathbb{P} \geq \widetilde{y}_{t_1}^{\mathbb{P},t_2}$ and $(\widetilde{y}^{\mathbb{P},t_2}, \widetilde{z}^{\mathbb{P},t_2}, \widetilde{u}^{\mathbb{P},t_2})$ satisfies the following BSDE on $[0,t_2]$

$$\widetilde{y}_t^{\mathbb{P},t_2} = \widetilde{V}_{t_2}^\mathbb{P} + \int_t^{t_2} f_s^\mathbb{P}(\widetilde{y}_s^{\mathbb{P},t_2}, \widetilde{z}_s^{\mathbb{P},t_2}, \widetilde{u}_s^{\mathbb{P},t_2})ds - \int_t^{t_2} \widetilde{z}_s^{\mathbb{P},t_2} dW_s^\mathbb{P} - \int_t^{t_2}\int_{\mathbb{R}^d} \widetilde{u}_s^{\mathbb{P},t_2}(x)\widetilde{\mu}_{B^d}(ds,dx),$$

where

$$\begin{aligned} f_t^\mathbb{P}(\omega,y,z,u) :&= \widehat{F}_t(\omega, y + \mathcal{Y}_t^\mathbb{P}(\omega), \widehat{a}_t^{-\frac{1}{2}}(\omega)(z + \mathcal{Z}_t^\mathbb{P}(\omega)), u + \bar{\mathcal{U}}_t^\mathbb{P}(\omega)) \\ &\quad - \widehat{F}_t(\omega, \mathcal{Y}_t^\mathbb{P}(\omega), \mathcal{Z}_t^\mathbb{P}(\omega), \mathcal{U}_t^\mathbb{P}(\omega)). \end{aligned}$$

By the definition given in Royer [106], we conclude from the above that $\widetilde{V}^\mathbb{P}$ is a positive $f^\mathbb{P}$-supermartingale under $\mathbb{P}$. Since $f^\mathbb{P}(0,0,0) = 0$, we can apply the downcrossing inequality proved in [106] to obtain classically that for $\mathbb{P} - a.e.\ \omega$, the limit

$$\lim_{r\in\mathbb{Q}\cup(t,T],r\downarrow t} \widetilde{V}_r^\mathbb{P}(\omega)$$

exists for all t. Finally, since $\bar{\mathcal{Y}}^\mathbb{P}$ is càdlàg, we obtain the desired result. □

We follow now Remark 4.9 in [110], and for a fixed $\mathbb{P} \in \mathcal{P}_H^\kappa$, we introduce the following RBSDE with jumps and with lower obstacle V^+ under $\mathbb{P}$

$$\begin{aligned} &\widetilde{Y}_t^\mathbb{P} = \xi + \int_t^T \widehat{F}_s(\widetilde{Y}_s^\mathbb{P}, \widetilde{Z}_s^\mathbb{P}, \widetilde{U}_s^\mathbb{P}, \nu)ds - \int_t^T \widetilde{Z}_s^\mathbb{P} dB_s^c - \int_t^T\int_E \widetilde{U}_s^\mathbb{P}(x)\widetilde{\mu}_{B^d}(ds,dx) + \widetilde{K}_T^\mathbb{P} - \widetilde{K}_t^\mathbb{P} \\ &\widetilde{Y}_t^\mathbb{P} \geq V_t^+,\ 0 \leq t \leq T,\ \mathbb{P} - a.s. \\ &\int_0^T \left(\widetilde{Y}_{s^-}^\mathbb{P} - V_{s^-}^+\right) d\widetilde{K}_s^\mathbb{P} = 0,\ \mathbb{P} - a.s., \end{aligned}$$

where we emphasize that the process $\widetilde{K}^\mathbb{P}$ is predictable.

Remark 6.2.2 *Existence and uniqueness of the above RBSDE under our Assumptions, with the restrictions that the compensator is not random, have been proved by Hamadène and Ouknine [58] or Essaky [52]. However, their proofs can be easily generalized to our context.*

Let us now argue by contradiction and suppose that $\widetilde{Y}^{\mathbb{P}}$ is not equal $\mathbb{P}-a.s.$ to V^+. Then we can assume without loss of generality that $\widetilde{Y}_0^{\mathbb{P}} > V_0^+$, $\mathbb{P}-a.s.$ fix now some $\varepsilon > 0$ and define the following stopping-time

$$\tau^\varepsilon := \inf\left\{t \geq 0,\ \widetilde{Y}_t^{\mathbb{P}} \leq V_t^+ + \varepsilon\right\}.$$

Then $\widetilde{Y}^{\mathbb{P}}$ is strictly above the obstacle before τ^ε, and therefore $\widetilde{K}^{\mathbb{P}}$ is identically equal to 0 in $[0,\tau^\varepsilon]$. Hence, we have

$$\widetilde{Y}_t^{\mathbb{P}} = \widetilde{Y}_{\tau^\varepsilon}^{\mathbb{P}} + \int_t^{\tau^\varepsilon} \widehat{F}_s(\widetilde{Y}_s^{\mathbb{P}}, \widetilde{Z}_s^{\mathbb{P}}, \widetilde{U}_s^{\mathbb{P}})ds - \int_t^{\tau^\varepsilon} \widetilde{Z}_s^{\mathbb{P}} dB_s^c - \int_t^{\tau^\varepsilon}\int_E \widetilde{U}_s^{\mathbb{P}}(x)\widetilde{\mu}_{B^d}(ds,dx),\ \mathbb{P}-a.s.$$

Let us now define the following BSDE on $[0,\tau^\varepsilon]$

$$y_t^{+,\mathbb{P}} = V_{\tau^\varepsilon}^+ + \int_t^{\tau^\varepsilon} \widehat{F}_s(y_s^{+,\mathbb{P}}, z_s^{+,\mathbb{P}}, u_s^{+,\mathbb{P}})ds - \int_t^{\tau^\varepsilon} z_s^{+,\mathbb{P}} dB_s^c - \int_t^{\tau^\varepsilon}\int_E u_s^{+,\mathbb{P}}(x)\widetilde{\mu}_{B^d}(ds,dx),\ \mathbb{P}-a.s.$$

By the standard a priori estimates already used in this chapter, we obtain that

$$\widetilde{Y}_0^{\mathbb{P}} \leq y_0^{+,\mathbb{P}} + C\left|V_{\tau^\varepsilon}^+ - \widetilde{Y}_{\tau^\varepsilon}^{\mathbb{P}}\right| \leq y_0^{+,\mathbb{P}} + C\varepsilon,$$

by definition of τ^ε. Following the arguments in Step 1 of the proof of Theorem 4.5 in [110], we can show that $y_0^{+,\mathbb{P}} \leq V_0^+$ which in turn implies

$$\widetilde{Y}_0^{\mathbb{P}} \leq V_0^+ + C\varepsilon,$$

hence a contradiction by arbitrariness of ε. Therefore, we have obtained the following decomposition

$$V_t^+ = \xi + \int_t^T \widehat{F}_s(V_s^+, \widetilde{Z}_s^{\mathbb{P}}, \widetilde{U}_s^{\mathbb{P}})ds - \int_t^T \widetilde{Z}_s^{\mathbb{P}} dB_s^c - \int_t^T\int_E \widetilde{U}_s^{\mathbb{P}}(x)\widetilde{\mu}_{B^d}(ds,dx) + \widetilde{K}_T^{\mathbb{P}} - \widetilde{K}_t^{\mathbb{P}},\ \mathbb{P}-a.s.$$

Finally, we can use the result of Nutz [95] to aggregate the families

$$\left\{\widetilde{Z}^{\mathbb{P}},\ \mathbb{P}\in\mathcal{P}_H^\kappa\right\} \text{ and } \left\{\widetilde{U}^{\mathbb{P}},\ \mathbb{P}\in\mathcal{P}_H^\kappa\right\},$$

into universal processes $\widetilde{Z}$ and $\widetilde{U}$.

We next prove the representation (5.4.1) for V and V^+, and that, as shown in Proposition 4.11 of [110], we actually have $V = V^+$, $\mathcal{P}_H - q.s.$, which shows that in the case of a terminal condition in $UC_b(\Omega)$, the solution of the 2RBSDE is actually $\mathbb{F}$-progressively measurable.

Proposition 6.2.3 *Assume that $\xi \in UC_b(\Omega)$ and that Assumptions 5.3.1 and 5.3.2 hold. Then we have*

$$V_t = \underset{\mathbb{P}'\in\mathcal{P}_H(t,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathcal{Y}_t^{\mathbb{P}'}(T,\xi) \text{ and } V_t^+ = \underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathcal{Y}_t^{\mathbb{P}'}(T,\xi),\ \mathbb{P}-a.s.,\ \forall \mathbb{P}\in\mathcal{P}_H^\kappa.$$

Besides, we also have for all t, $V_t = V_t^+$, $\mathcal{P}_H^\kappa - q.s.$

Proof. The proof for the representations is the same as the proof of proposition 4.10 in [110], since we also have a stability result for BSDEs under our assumptions. For the equality between V and V^+, we also refer to the proof of Proposition 4.11 in [110]. $\square$

Therefore, in the sequel we will use V instead of V^+. Finally, we have to check that the minimum condition (5.3.5) holds. Fix $\mathbb{P}$ in $\mathcal{P}_H^\kappa$ and $\mathbb{P}^{'} \in \mathcal{P}_H^\kappa(t^+, \mathbb{P})$. Then, proceeding exactly as in Step 2 of the proof of Theorem 5.4.1 in Chapter 5, but introducing the process γ' of Assumption 5.3.1(iv) instead of γ, we can similarly obtain

$$V_t - y_t^{\mathbb{P}^{'}} \geq \mathbb{E}_t^{\mathbb{P}^{'}}\left[\int_t^T M_s' d\widetilde{K}_s^{\mathbb{P}^{'}}\right] \geq \mathbb{E}_t^{\mathbb{P}^{'}}\left[\inf_{t\leq s\leq T} M_s'\left(\widetilde{K}_T^{\mathbb{P}^{'}} - \widetilde{K}_t^{\mathbb{P}^{'}}\right)\right],$$

where M' is defined as M but with γ' instead of γ. Now let us prove that for any $n > 1$

$$\mathbb{E}_t^{\mathbb{P}^{'}}\left[\left(\inf_{t\leq s\leq T} M_s'\right)^{-n}\right] < +\infty,\ \mathbb{P}^{'} - a.s. \tag{6.2.12}$$

First we have

$$M_s' = \exp\left(\int_t^s \lambda_r dr + \int_t^s \eta_r \widehat{a}_r^{-1/2} dB_s^c - \frac{1}{2}\int_t^s |\eta_r|^2 dr + \int_t^s \int_E \gamma_r'(x)\widetilde{\mu}_{B^d}(dr, dx)\right)$$
$$\times \prod_{t\leq r\leq s}(1 + \gamma_r'(\Delta B_r))e^{-\gamma_r'(\Delta B_r)}.$$

Define, then $A_s = \mathcal{E}\left(\int_t^s \eta_r \widehat{a}_r^{-1/2} dB_s^c\right)$ and $C_s = \mathcal{E}\left(\int_t^s \int_E \gamma_r'(x)\widetilde{\mu}_{B^d}(dr, dx)\right)$. Notice that both this processes are strictly positive martingales, since η and γ' are bounded and we have assumed that γ' is strictly greater than -1. We have

$$M_s' = \exp\left(\int_t^s \lambda_r dr\right) A_s C_s.$$

Since the process λ is bounded, we have

$$\left(\inf_{t\leq s\leq T} M_s'\right)^{-n} \leq C\left(\inf_{t\leq s\leq T} A_s C_s\right)^{-n} = C\left(\sup_{t\leq s\leq T}(A_s C_s)^{-1}\right)^n.$$

Using the Doob inequality for the submartingale $(A_s C_s)^{-1}$, we obtain

$$\mathbb{E}_t^{\mathbb{P}^{'}}\left[\left(\inf_{t\leq s\leq T} M_s'\right)^{-n}\right] \leq C\mathbb{E}_t^{\mathbb{P}^{'}}\left[(C_T A_T)^{-n}\right] \leq C\left(\mathbb{E}_t^{\mathbb{P}^{'}}\left[(C_T)^{-2n}\right]\mathbb{E}_t^{\mathbb{P}^{'}}\left[(A_T)^{-2n}\right]\right)^{1/2} < +\infty,$$

where we used the fact that since η is bounded, the continuous stochastic exponential A has negative moments of any order, and where the same result holds for the purely discontinuous stochastic exponential C by Lemma B.2.4 in the Appendix B.

Then, we have for any $p>1$

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{P}'}\left[\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right]\\
&=\mathbb{E}_t^{\mathbb{P}'}\left[\left(\inf_{t\leq s\leq T}M_s'\right)^{1/p}\left(\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right)\left(\inf_{t\leq s\leq T}M_s'\right)^{-1/p}\right]\\
&\leq\left(\mathbb{E}_t^{\mathbb{P}'}\left[\inf_{t\leq s\leq T}M_s'\left(\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right)\right]\right)^{1/p}\left(\mathbb{E}_t^{\mathbb{P}'}\left[\inf_{t\leq s\leq T}M_s'^{-\frac{2}{p-1}}\right]\mathbb{E}_t^{\mathbb{P}'}\left[\left(\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right)^2\right]\right)^{\frac{p-1}{2p}}\\
&\leq C\left(\underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}^{\mathbb{P}'}\left[\left(\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right)^2\right]\right)^{\frac{p-1}{2p}}\left(V_t-y_t^{\mathbb{P}'}\right)^{1/p},
\end{aligned}$$

where we used (6.2.12). Arguing as in Step (iii) of the proof of Theorem 5.4.1, the above inequality along with Proposition 6.2.3 shows that we have

$$\underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,inf}^{\mathbb{P}}}\ \mathbb{E}^{\mathbb{P}'}\left[\widetilde{K}_T^{\mathbb{P}'}-\widetilde{K}_t^{\mathbb{P}'}\right]=0,$$

that is to say that the minimum condition 5.3.5 is satisfied. This implies that the family $\left\{\widetilde{K}^{\mathbb{P}}\right\}_{\mathbb{P}\in\mathcal{P}_H^\kappa}$ satisfies the consistency condition (i) of Theorem 5.2.1 in Chapter 5 and therefore can be aggregated by this Theorem.

6.2.3 Main result

We are now in position to state the main result of this section

Theorem 6.2.1 *Let $\xi\in\mathcal{L}_H^{2,\kappa}$. Under Assumptions 5.3.1 and 5.3.2, there exists a unique solution $(Y,Z,U)\in\mathbb{D}_H^\infty\times\mathbb{H}_H^2\times\mathbb{J}_H^{2,\kappa}$ of the* 2BSDEJ (5.3.3).

Proof. The proof follow the lines of the proof of Theorem 4.7 in [112], using the a priori estimates of Theorem 5.4.4 in Chapter 5, therefore we omit it. □

6.2.4 An extension of the representation formula

So far, we managed to provide wellposedness results for 2BSDEs with jumps, by working under a set a probability measures which, if restricted to the ones for which the canonical process is a continuous local martingale is strictly smaller than the one considered in [112] or [104]. This is due mainly to the fact that we had to restrict ourselves to processes α satisfying the separability condition defined in the previous chapter in order to retrieve the aggregation result of Theorem 5.2.1, which was crucial to our analysis since it allowed us to define an aggregator for the family of predictable compensators.

This is clearly not very satisfying, not only from the theoretical point of view, but also from the practical one. Indeed, the set from which the processes α are allowed to be chosen corresponds in financial applications to the set of possible volatility processes for the market considered. It is therefore desirable to have the greatest possible generality. However, we emphasize that the restrictions we put on the predictable compensators ν are clearly not a problem from the point of view of the applications. Indeed, our set of compensators is strictly greater than the one associated to pure jump additive processes. Those processes, and more precisely the Lévy processes, being the most widely used in applications, our set is not really restrictive.

The aim of this section is to show that under additional assumptions, we can show that the representation formula (5.4.1) also holds for a larger set of probability measures for which there is no longer any restrictions on the processes α. In this regard, we recall the set of probability measures $\overline{\mathcal{P}}_{\widetilde{\mathcal{A}}}$ defined in Remark 5.2.6. We recall that every probability measure in this set satisfy the Blumenthal $0-1$ law and the martingale representation property. Moreover, exactly as in Definition 5.3.1, we define and restrict ourselves to the subset $\overline{\mathcal{P}}_H^\kappa$ of $\overline{\mathcal{P}}_{\widetilde{\mathcal{A}}}$. We define the following space for each $p \geq \kappa$,

$$\overline{\mathbb{L}}_H^{p,\kappa} := \left\{\xi,\ \|\xi\|_{\overline{\mathbb{L}}_H^{p,\kappa}} < +\infty\right\} \text{ where } \|\xi\|^p_{\overline{\mathbb{L}}_H^{p,\kappa}} := \sup_{\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\operatorname*{ess\,sup}_{0\leq t\leq T}{}^{\mathbb{P}} \left(\mathbb{E}_t^{H,\mathbb{P}}[|\xi|^\kappa]\right)^{\frac{p}{\kappa}}\right],$$

and we let

$$\overline{\mathcal{L}}_H^{p,\kappa} := \text{the closure of } \mathrm{UC}_b(\Omega) \text{ under the norm } \|\cdot\|_{\overline{\mathbb{L}}_H^{p,\kappa}}, \text{ for every } 1<\kappa\leq p.$$

We then have the following result, which is similar to Theorem 5.3 in [110]

Theorem 6.2.2 *Let $\xi \in \overline{\mathcal{L}}_H^{2\kappa}$ and in addition to Assumptions 5.3.1 and 5.3.2, assume that*

- *F is uniformly continuous in a for $a \in D^1_{F_t}$, and for all (t,ω,y,z,u,a,ν)*

$$|F_t(\omega,y,z,u,a,\nu)| \leq C\left(1 + d_{S,t}(\omega,0) + |y| + |z| + \left(\int_E u_s^2(x)\nu(dx)\right)^{1/2} + \left|a^{1/2}\right|\right). \quad (6.2.13)$$

- *$\mathcal{P}_H^\kappa$ is dense in $\overline{\mathcal{P}}_H^\kappa$ in the sense that for any $\mathbb{P}^{\alpha,\nu} \in \overline{\mathcal{P}}_H^\kappa$ and for any $\varepsilon > 0$, there exists $\mathbb{P}^{\alpha^\varepsilon,\nu} \in \mathcal{P}_H^\kappa$ such that*

$$\mathbb{E}^{\mathbb{P}_\nu}\left[\int_0^T |(\alpha_t^\varepsilon)^{1/2} - \alpha_t^{1/2}|^2 dt\right] \leq \varepsilon. \quad (6.2.14)$$

Then, we have

$$Y_0 = \sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} y_0^{\mathbb{P}} = \sup_{\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa} y_0^{\mathbb{P}},$$

where under any $\mathbb{P} := \mathbb{P}^{\alpha,\nu} \in \overline{\mathcal{P}}_H^\kappa$, $(y^{\mathbb{P}}, z^{\mathbb{P}}, u^{\mathbb{P}})$ is the unique solution of the BSDEJ

$$y_t^{\mathbb{P}} = \xi + \int_t^T F_s(y_s^{\mathbb{P}}, z_s^{\mathbb{P}}, u_s^{\mathbb{P}}, \widehat{a}_s, \nu_s)ds - \int_t^T Z_s dB_s^c - \int_t^T\int_E u_s^{\mathbb{P}}(x)\widetilde{\mu}^{\mathbb{P}}_{B^d}(ds,dx),\ \mathbb{P}-a.s.$$

Remark 6.2.3 *A sufficient condition for (6.2.13) to hold is that $F_t(0,0,0,0,0,\nu)$ is bounded uniformly, since F is already uniformly Lipschitz in (y,z,u) and is assumed to be uniformly continuous in (ω,a).*

Proof. First, we remind that Remark 5.4.1 ensures existence and uniqueness of the solutions of our BSDEs under any $\mathbb{P} \in \overline{\mathcal{P}}_H^\kappa$. We will proceed in two steps.

(i) $\xi \in \mathrm{UC_b}(\Omega)$

For any $\mathbb{P} := \mathbb{P}^{\alpha,\nu} \in \overline{\mathcal{P}}_H^\kappa$ and any $\varepsilon > 0$, let $\mathbb{P}^\varepsilon := \mathbb{P}^{\alpha^\varepsilon,\nu} \in \mathcal{P}_H^\kappa$ be given by (6.2.14). Using the process $L^\mathbb{P}$ defined in (5.2.7), we have $\mathbb{P}-a.s.$

$$\begin{aligned} y_t^\mathbb{P} &= \xi(B_\cdot) + \int_t^T F_s(B_\cdot, y_s^\mathbb{P}, z_s^\mathbb{P}, u_s^\mathbb{P}, \widehat{a}_s, \nu_s)ds - \int_t^T \widehat{a}_s^{1/2} z_s^\mathbb{P} dL_s^{\mathbb{P},c} \\ &\quad - \int_t^T \int_E u_s^\mathbb{P}(x)\left(\mu_{L^{\mathbb{P},d}}(ds,dx) - \nu_s(dx)ds\right). \end{aligned}$$

Let now $(\overline{y}^\mathbb{P}, \overline{z}^\mathbb{P}, \overline{u}^\mathbb{P})$ denote the unique solution of the following BSDEJ under $\mathbb{P}_\nu$

$$\begin{aligned} \overline{y}_t^\mathbb{P} &= \xi(X_\cdot^{\alpha,\nu}) + \int_t^T F_s(X_\cdot^{\alpha,\nu}, \overline{y}_s^\mathbb{P}, \overline{z}_s^\mathbb{P}, \overline{u}_s^\mathbb{P}, \alpha_s, \nu_s)ds - \int_t^T \alpha_s^{1/2} \overline{z}_s^\mathbb{P} dB_s^c \\ &\quad - \int_t^T \int_E \overline{u}_s^\mathbb{P}(x)\left(\mu_{B^d}(ds,dx) - \nu_s(dx)ds\right). \end{aligned}$$

By definition of $\mathbb{P}^{\alpha,\nu}$, we know that the distribution of $y^\mathbb{P}$ under $\mathbb{P}$ is equal to the distribution of $\overline{y}^\mathbb{P}$ under $\mathbb{P}_\nu$. Since the Blumenthal $0-1$ law also holds, this implies clearly that we have

$$y_0^\mathbb{P} = \overline{y}_0^\mathbb{P}.$$

Similarly, we define $y^{\mathbb{P}^\varepsilon}$ and $\overline{y}^{\mathbb{P}^\varepsilon}$. Then, using classical estimates from the BSDEJ theory (see [4] for instance) we have

$$\begin{aligned} &|y_0^\mathbb{P} - y_0^{\mathbb{P}^\varepsilon}|^2 = |\overline{y}_0^\mathbb{P} - \overline{y}_0^{\mathbb{P}^\varepsilon}|^2 \\ &\leq C\mathbb{E}^{\mathbb{P}_\nu}\left[|\xi(X_\cdot^{\alpha,\nu}) - \xi(X_\cdot^{\alpha^\varepsilon,\nu})|^2 + \int_0^T \left|F_t(X_\cdot^{\alpha}, \overline{y}_t^\mathbb{P}, \overline{z}_t^\mathbb{P}, \alpha_t, \nu_t) - F_t(X_\cdot^{\alpha^\varepsilon}, \overline{y}_t^\mathbb{P}, \overline{z}_t^\mathbb{P}, \alpha_t^\varepsilon, \nu_t)\right|^2 dt\right]. \end{aligned}$$

Then, we have by (6.2.13)

$$\begin{aligned} \left|F_t(X_\cdot^{\alpha^\varepsilon}, \overline{y}_t^\mathbb{P}, \overline{z}_t^\mathbb{P}, \alpha_t^\varepsilon, \nu_t)\right| &\leq C\left(1 + d_{S,t}(X^{\alpha^\varepsilon,\nu}, 0) + |\overline{y}_t^\mathbb{P}| + |\overline{z}_t^\mathbb{P}| + |\alpha_t^\varepsilon|^{1/2}\right) \\ &\leq C\left(1 + d_{S,t}(X^{\alpha,\nu}, 0) + |\overline{y}_t^\mathbb{P}| + |\overline{z}_t^\mathbb{P}| + |\alpha_t|^{1/2}\right) \\ &\quad + C\left(d_{S_t}(X^{\alpha^\varepsilon,\nu}, X^{\alpha,\nu}) + |\alpha_t^\varepsilon - \alpha_t|^{\mathbf{1}/2}\right) \\ &\leq C\left(1 + \sup_{0\leq s\leq T}|X_s^{\alpha,\nu}| + |\overline{y}_t^\mathbb{P}| + |\overline{z}_t^\mathbb{P}| + |\alpha_t|^{1/2}\right) \\ &\quad + C\left(\sup_{0\leq s\leq T}\left|X_s^{\alpha^\varepsilon,\nu} - X_s^{\alpha,\nu}\right|) + |\alpha_t^\varepsilon - \alpha_t|^{\mathbf{1}/2}\right), \end{aligned} \tag{6.2.15}$$

where we used in the last inequality the fact that the Skorohod distance is dominated by the uniform norm (see for instance Billingsley [12]).

Using Doob's inequality and Itô's isometry, it is easy to see that (6.2.14) implies that

$$\mathbb{E}^{\mathbb{P}_\nu}\left[\sup_{0\leq t\leq T}\left|X_t^{\alpha^\varepsilon,\nu} - X_t^{\alpha,\nu}\right|^2\right] \leq \varepsilon.$$

Since ξ is also uniformly continuous and bounded in ω, we can apply the dominated convergence Theorem in (6.2.15) to obtain

$$\lim_{\varepsilon\to 0}|y_0^{\mathbb{P}} - y_0^{\mathbb{P}^\varepsilon}| = 0.$$

This clearly implies the result in that case.

(ii) $\xi \in \overline{\mathcal{L}}_H^{2,\kappa}$

In that case, with the same ntations as above, let $\xi_n \in \mathrm{UC_b}(\Omega)$ such that $\|\xi - \xi_n\|_{\overline{\mathbb{L}}_H^{2,\kappa}} \underset{n\to+\infty}{\longrightarrow} 0$. Then, we define $y^{\mathbb{P},n}$ the solution of the BSDEJ with terminal condition ξ_n and generator $F_t(\cdot, \widehat{a}_s, \nu_s)$ under $\mathbb{P}$. Then, we have

$$\sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} y_0^{\mathbb{P},n} = \sup_{\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa} y_0^{\mathbb{P},n}. \tag{6.2.16}$$

Moreover, using classical estimates for BSDEs, we can show that

$$\left|y_0^{\mathbb{P},n} - y_0^{\mathbb{P}}\right|^2 \leq C\,\|\xi_n - \xi\|^2_{\overline{\mathbb{L}}_H^{2,\kappa}}.$$

This shows that the convergence of $y_0^{\mathbb{P},n}$ to $y_0^{\mathbb{P}}$ is uniform with respect to $\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa$. Hence we can pass to the limit in (6.2.16) and exchange the limit and the suprema to obtain the desired result. □

We finish this Section by recalling a result from [110] (see Proposition 5.4) which gives a sufficient condition for the density condition (6.2.14)

Lemma 6.2.3 *Assume that the domain of F does not depend on t and that D_F^1 contains a countable dense subset. Then* (6.2.14) *holds.*

Proof. It suffices to notice that in our framework, all the constant mappings belong to $\widetilde{A}_0$. Then Proposition 5.4 in [110] applies. □

6.3 Application to a robust utility maximization problem

We study in this part a robust exponential utility maximization problem in the spirit of [59] and [87]. We consider the case of an agent maximizing the utility of his terminal wealth, over all possible trading strategies. The agent also have at the terminal date T a claim ξ, which is a $\mathcal{F}_T$-measurable random variable. We work in a setting with model ambiguity, due to the presence of the set $\mathcal{P}_H^\kappa$, composed of non dominated probability measures, accounting for the volatility uncertainty and jump measure uncertainty (see also remark 6.3.2). In [51] and [59], the utility maximization problems are solved by means of stochastic control techniques, and the value function is linked with some particular BSDEs. In our case, since we work with a whole set of non dominated probability measures, our value function will be linked to some second order BSDE with jumps, and the general martingale optimality principle described in [51] is treated via the use of non-linear martingales.

6.3.1 The Market

In this section, we will always assume that the matrices $\underline{a} := \underline{a}^{\mathbb{P}}$ and $\overline{a} := \overline{a}^{\mathbb{P}}$ are uniformly bounded in $\mathbb{P}$. In particular, this implies that we can restrict ourselves to the case where the parameter a in the definition of a generator F is bounded. We consider a financial market consisting of one riskless asset, whose price is assumed to be equal to 1 for simplicity, and one risky asset whose price process $(S_t)_{0\leq t\leq T}$ is assumed to follow a mixed-diffusion

$$\frac{dS_t}{S_{t^-}} = b_t dt + dB_t^c + \int_E \beta_t(x)\mu_{B^d}(dt,dx),\ \overline{\mathcal{P}}_H^\kappa - q.s., \tag{6.3.1}$$

where we assume that

Assumption 6.3.1 (i) (b_t) *is a bounded* $\mathbb{F}$*-predictable process uniformly continuous in* ω *for the Skorohod topology.*

(ii) (β_t) *is a bounded* $\mathbb{F}$*-predictable process which is uniformly continuous in* ω *for the Skorohod topology, verifies*

$$\sup_{\nu\in\mathcal{N}}\int_0^T\int_E |\beta_t(x)|\,\nu_t(dx)dt < +\infty,\ \overline{\mathcal{P}}_H^\kappa - q.s.,$$

and satisfies

$$C_1(1\wedge|x|) \leq \beta_t(x) \leq C_2(1\wedge|x|),\ \overline{\mathcal{P}}_H^\kappa - q.s., \text{ for all } (t,x)\in[0,T]\times E,$$

where $C_2 \geq 0 \geq C_1 > -1+\delta$*, where* $\delta > 0$ *is fixed.*

Remark 6.3.1 *The uniform continuity assumption on* ω *is here to ensure that the 2BSDEs we will encounter in the sequel indeed have solutions. The assumption on* β *is classical [82] and implies that the price process* S *is positive.*

Remark 6.3.2 *The volatility is implicitly embedded in the model. Indeed, under each* $\mathbb{P}\in\mathcal{P}_H^\kappa$*, we have* $dB_s^c \equiv \widehat{a}_t^{1/2}dW_t^{\mathbb{P}}$ *where* $W^{\mathbb{P}}$ *is a Brownian motion under* $\mathbb{P}$*. Therefore,* $\widehat{a}^{1/2}$ *plays the role of volatility under each* $\mathbb{P}$ *and thus allows us to model the volatility uncertainty. Similarly, we have incertitude on the jumps of our price process, since the predictable compensator associated to the jumps of the discontinuous part of the canonical process changes with the probability considered. This allows us to have incertitude not only about the size of the jumps but also about their laws.*

We then denote $\pi = (\pi_t)_{0\leq t\leq T}$ a trading strategy, which is a 1-dimensional $\mathbb{F}$-predictable process, supposed to take its value in some compact set C. The process π_t describes the amount of money invested in the stock at time t. The number of shares is $\frac{\pi_t}{S_{t^-}}$. So the liquidation value of a trading strategy π with positive initial capital x is given by the following wealth process:

$$X_t^\pi = x + \int_0^t \pi_s\left(dB_s^c + b_s ds + \int_E \beta_s(x)\mu_{B^d}(ds,dx)\right),\ 0\leq t\leq T,\ \overline{\mathcal{P}}_H^\kappa - q.s.$$

The problem of the investor in this financial market is to maximize his expected exponential utility under model uncertainty from his total wealth $X_T^\pi - \xi$ where ξ is a liability at time T which is a random variable assumed to be $\mathcal{F}_T$-measurable. Then the value function V of the maximization problem can be written as

$$V^\xi(x) := \sup_{\pi\in\mathcal{C}}\inf_{\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa}\mathbb{E}^{\mathbb{P}}\left[-\exp\left(-\eta\left(X_T^\pi - \xi\right)\right)\right] = -\inf_{\pi\in\mathcal{C}}\sup_{\mathbb{P}\in\overline{\mathcal{P}}_H^\kappa}\mathbb{E}^{\mathbb{P}}\left[\exp\left(-\eta\left(X_T^\pi - \xi\right)\right)\right]. \tag{6.3.2}$$

where

$$\mathcal{C} := \{(\pi_t) \text{ which are predictable and take values in } C\},$$

is our set of admissible strategies.

Before going on, we emphasize immediately, that in the sequel we will limit ourselves to probability measures in $\mathcal{P}_H^\kappa$. We will recover the supremum over all probability measures in $\overline{\mathcal{P}}_H^\kappa$ at the end by showing that Theorem 6.2.2 applies.

To find the value function V^ξ and an optimal trading strategy π^*, we follow the ideas of the general *martingale optimality principle* approach as in [51] and [59] but adapt it here to a nonlinear framework. This is the same approach as in [87].

Let $\{R^\pi\}$ be a family of processes which satisfy the following properties

Properties II.1 (i) $R_T^\pi = \exp(-\eta(X_T^\pi - \xi))$ *for all* $\pi \in \mathcal{C}$.

(ii) $R_0^\pi = R_0$ *is constant for all* $\pi \in \mathcal{C}$.

(iii) *We have*

$$R_t^\pi \leq \operatorname*{ess\,sup^{\mathbb{P}}}_{\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+,\mathbb{P})} \mathbb{E}_t^{\mathbb{P}'}[R_T^\pi], \ \forall \pi \in \mathcal{C}$$

$$R_t^{\pi^*} = \operatorname*{ess\,sup^{\mathbb{P}}}_{\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+,\mathbb{P})} \mathbb{E}_t^{\mathbb{P}'}[R_T^{\pi^*}] \text{ for some } \pi^* \in \mathcal{C}, \ \mathbb{P}-a.s. \text{ for all } \mathbb{P} \in \mathcal{P}_H^\kappa.$$

Then it follows

$$\sup_{\mathbb{P} \in \mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}[U(X_T^\pi - \xi)] \geq R_0 = \sup_{\mathbb{P} \in \mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}[U(X_T^{\pi^*} - \xi)] = -V^\xi(x). \tag{6.3.3}$$

6.3.2 Solving the optimization problem with a Lipschitz 2BSDEJ

To construct R^π, we set

$$R_t^\pi = \exp(-\eta X_t^\pi) Y_t, \ t \in [0,T], \ \pi \in \mathcal{C},$$

where $(Y, Z, U) \in \mathbb{D}_H^{2,\kappa} \times \mathbb{H}_H^{2,\kappa} \times \mathbb{J}_H^{2,\kappa}$ is the unique solution of the following 2BSDEJ

$$Y_t = e^{\eta\xi} + \int_t^T \widehat{F}_s(Y_s, Z_s, U_s)ds - \int_t^T Z_s dB_s^c - \int_t^T \int_E U_s(x)\widetilde{\mu}_{B^d}(ds,dx) + K_T - K_t. \tag{6.3.4}$$

The generator $\widehat{F}$ is chosen so that R^π satisfies the Properties II.1. Let us apply Itô's formula to $\exp(-\eta X_t^\pi)Y_t$ under some $\mathbb{P} \in \mathcal{P}_H^\kappa$. We obtain after some calculations

$$\begin{aligned} d\left(e^{-\eta X_t^\pi} Y_t\right) = e^{-\eta X_{t^-}^\pi} \Big[& -\eta\pi_s b_s Y_s ds + \frac{\eta^2}{2}\pi_s^2 \widehat{a}_s Y_s ds - \eta\pi_s \widehat{a}_s Z_s ds - \widehat{F}_s(Y_s, Z_s, U_s)ds \\ & + \int_E \left(e^{-\eta\pi_s\beta_s(x)} - 1\right)(Y_s + U_s(x))\widehat{\nu}_s(dx)ds + (Z_s - \eta\pi_s Y_s)\, dB_s^c \\ & + \int_E \left(e^{-\eta\pi_s\beta_s(x)} - 1\right)(Y_{s^-} + U_s(x)) + U_s(x)\widetilde{\mu}_{B^d}(ds,dx) - dK_t \Big]. \end{aligned} \tag{6.3.5}$$

Hence the appropriate choice for F

$$F_s(y,z,u,a,\nu) := \inf_{\pi \in C} \left\{ (-\eta b_s + \frac{\eta^2}{2}\pi a)\pi y - \eta\pi a z + \int_E \left(e^{-\eta\pi\beta_s(x)} - 1 \right)(y + u(x))\nu(dx) \right\}.$$

First, because of Assumption 6.3.1, F is uniformly Lipschitz in (y,z), uniformly continuous in ω. It is also continuous in a and since $D_F^1 = [\underline{a}, \overline{a}]$, it is even uniformly continuous in a. Besides, it is convex in a and ν (since it is the infimum of a family of linear functions) and hence can be written as a Fenchel-Legendre transform. Moreover, its domain clearly does not depend on (ω,t,y,z,u) by our boundedness assumptions. Besides, D_F^1 clearly contains a countable dense subset. This in particular shows that Theorem 6.2.2 applies here. Finally,

$$\inf_{\pi \in C} \int_E \left(e^{-\eta\pi\beta_s(\omega,x)} - 1 \right)(u(x) - u'(x))\nu(dx) \leq F_s(\omega,y,z,u,a,\nu) - F_s(\omega,y,z,u',a,\nu)$$
$$F_s(\omega,y,z,u,a,\nu) - F_s(\omega,y,z,u',a,\nu) \leq \sup_{\pi \in C} \int_E \left(e^{-\eta\pi\beta_s(\omega,x)} - 1 \right)(u(x) - u'(x))\nu(dx).$$

Since C is compact and β is bounded, it is therefore clear from the above inequalities that Assumption 5.3.1(iv) is satisfied. Therefore, if we assume that $e^{\eta\xi} \in \mathcal{L}_H^{2,\kappa}$ (for instance if $\xi \in \mathcal{L}_H^{\infty,\kappa}$), the 2BSDEJ (6.3.4) indeed has a unique solution and R^π is well defined. Let us now prove that it satisfies the properties II.1. The property (i) is clear by definition and (ii) holds because of Proposition 6.2.3 together with the Bluementhal $0-1$ law. Now for any $0 \leq t \leq T$, any $\pi \in \mathcal{C}$, any $\mathbb{P} \in \mathcal{P}_H^\kappa$ and any $\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+,\mathbb{P})$, we have from (6.3.5)

$$\mathbb{E}_t^{\mathbb{P}'}\left[R_T^\pi\right] - R_t^\pi \geq -\mathbb{E}_t^{\mathbb{P}'}\left[\int_t^T e^{-\eta X_{s^-}^\pi} dK_s\right]. \tag{6.3.6}$$

Let us now prove that for any $\pi \in \mathcal{C}$ and for any $\mathbb{P}$, we have

$$\underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,inf}^{\mathbb{P}}} \mathbb{E}_t^{\mathbb{P}'}\left[\int_t^T e^{-\eta X_{s^-}^\pi} dK_s\right] = 0. \tag{6.3.7}$$

This is similar to what we did in the proof of Theorem 5.4.1, and therefore we know that it is sufficient to prove that for any $p > 1$

$$\mathbb{E}_t^{\mathbb{P}'}\left[\sup_{t \leq s \leq T} e^{-pX_s^\pi}\right] \leq C_p, \tag{6.3.8}$$

for some positive constant C_p depending only on p and the bounds for π, b and β. Let M be such that $-M \leq \pi \leq M$, we have

$$\begin{aligned}
e^{-X_s^\pi} &\leq e^{TM\|b\|_\infty - \int_0^t \pi_s dB_s^c + M\int_0^t \int_E |\beta_s(x)| \mu_{B^d}(ds,dx)} \\
&= e^{TM\|b\|_\infty + \frac{1}{2}\int_0^t \pi_s^2 \widehat{a}_s ds + M\int_0^T \int_E |\beta_s(x)| \widehat{\nu}_s(dx)ds + M\int_0^T \int_E |\beta_s(x)| - \ln(1+|\beta_s(x)|)\widehat{\nu}_s(dx)ds} \\
&\quad \times \mathcal{E}\left(\int_0^t \pi_s dB_s^c\right) \mathcal{E}\left(\int_0^t |\beta_s(x)| \widetilde{\mu}_{B^d}(ds,dx)\right)^M \\
&\leq C\mathcal{E}\left(-\int_0^t \pi_s dB_s^c\right) \mathcal{E}\left(\int_0^t |\beta_s(x)| \widetilde{\mu}_{B^d}(ds,dx)\right)^M,
\end{aligned}$$

where we used in the last inequality the fact that π, $\widehat{a}$ and β are uniformly bounded and that

$$\sup_{\nu\in\mathcal{N}}\int_0^T\int_E |\beta_t(x)|\,\nu_t(dx)dt<+\infty \text{ and } |\beta_s(x)-\ln(1+\beta_s(x))|\leq C|\beta_s(x)|^2.$$

Then (6.3.8) comes from the fact that above Doléans-Dade exponentials have moments of any order (it is clear for the continuous one, and we refer again to Lemma B.2.4 for the discontinuous one). Using (6.3.7) in (6.3.6), we obtain

$$R_t^\pi \leq \underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_t^{\mathbb{P}'}[R_T^\pi].$$

Now, using a classical mesurable selection argument (see [36] (chapitre III) or [45] or Lemma 3.1 in [49]) we can define a predictable process $\pi^*\in\mathcal{C}$ such that

$$\widehat{F}_s(Y_s,Z_s,U_s)=(-\eta b_s+\frac{\eta^2}{2}\pi_s^*\widehat{a}_s)\pi_s^*Y_s-\eta\pi_s^*\widehat{a}_s+\int_E\left(e^{-\eta\pi_s^*\beta_s(x)}-1\right)(Y_s+U_s(x))\widehat{\nu}_s(dx).$$

Using the same arguments as above, we obtain

$$R_t^{\pi^*} = \underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ \mathbb{E}_t^{\mathbb{P}'}[R_T^{\pi^*}],$$

which proves (iii) of Property II.1 holds.

We summarize everything in the following proposition

Proposition 6.3.1 *Assume that* $\exp(\eta\xi)\in\overline{\mathcal{L}}_H^{2,\kappa}$. *Then, under Assumption 6.3.1, the value function of the optimization problem (6.3.2) is given by*

$$V^\xi(x)=-e^{-\eta x}Y_0,$$

where Y_0 *is defined as the initial value of the unique solution* $(Y,Z,U)\in\mathbb{D}_H^{2,\kappa}\times\mathbb{H}_H^{2,\kappa}\times\mathbb{J}_H^{2,\kappa}$ *of the following 2BSDEJ*

$$Y_t=e^{\eta\xi}+\int_t^T\widehat{F}_s(Y_s,Z_s,U_s)ds-\int_t^T Z_s dB_s^c-\int_t^T\int_E U_s(x)\widetilde{\mu}_{B^d}(ds,dx)+K_T-K_t, \tag{6.3.9}$$

where the generator is defined as follows

$$\widehat{F}_t(\omega,y,z,u):=F_t(\omega,y,z,u,\widehat{a}_t,\widehat{\nu}_t),$$

where

$$F_t(y,z,u,a,\nu):=\inf_{\pi\in C}\left\{(-\eta b_t+\frac{\eta^2}{2}\pi a)\pi y-\eta\pi a z+\int_E\left(e^{-\eta\pi\beta_t(x)}-1\right)(y+u(x))\nu(dx)\right\}.$$

Moreover, there exists an optimal trading strategy π^* *realizing the infimum above.*

6.3.3 A link with a particular quadratic 2BSDEJ

In the case without uncertainty on the parameters, coming back to the paper of El Karoui and Rouge [51], utility maximization problems in a continuous framework with constrained strategies have usually been linked to BSDEs with a quadratic growth generator. The same type of results were later proved by Morlais [92] in a discontinuous setting, that is to say when the assets in the market are assumed to have jumps. More recently, Lim and Quenez [82] showed that when the strategies are constrained in a compact set, then one only needed to consider a BSDE with a Lipschitz generator, thus simplifying the approach of Morlais. Since we provided in this chapter a wellposedness theory for Lipscitz 2BSDEs with jumps, we used the ideas of [82] in the previous section. Our aim in this section is to show that with the result of Proposition 6.3.1 and by making a change of variables, we can prove the existence of a solution to a particular 2BSDEJ, similar to the one in [92], whose generator satisfies a quadratic growth condition.

Before proceeding with the proof, we recall that as for quadratic BSDEs and 2BSDEs, we always have a deep link between the martingale part of the solution and the BMO spaces. So we need to introduce the following spaces.

$\mathbb{J}^{2,\kappa,H}_{BMO}$ denotes the space of predictable and $\mathcal{E}$-mesurable applications $U : \Omega \times [0,T] \times E$ such that

$$\|U\|^2_{\mathbb{J}^{2,\kappa,H}_{BMO}} := \sup_{\mathbb{P}\in\mathcal{P}^\kappa_H} \left\| \int_0^\cdot \int_E U_s(x)\widetilde{\mu}_{B^d}(ds,dx) \right\|_{\mathrm{BMO}(\mathbb{P})} < +\infty,$$

where $\|\cdot\|_{\mathrm{BMO}(\mathbb{P})}$ is the usual BMO($\mathbb{P}$) norm under $\mathbb{P}$.

$\mathbb{H}^{2,\kappa,H}_{BMO}$ denotes the space of all $\mathbb{F}^+$-progressively measurable $\mathbb{R}^d$-valued processes Z with

$$\|Z\|_{H^{2,\kappa,H}_{BMO}} := \sup_{\mathbb{P}\in\mathcal{P}^\kappa_H} \left\| \int_0^\cdot Z_s dB^c_s \right\|_{\mathrm{BMO}(\mathbb{P})} < +\infty.$$

We refer the reader to the book by Kazamaki [72] and the references therein for more information on the BMO spaces.

Proposition 6.3.2 *Assume that* $\exp(\eta\xi) \in \mathcal{L}^{2,\kappa}_H$ *and that* ξ *is bounded quasi-surely. Then, there exists a solution* $(Y', Z', U') \in \mathbb{D}^{2,\kappa}_H \times \mathbb{H}^{2,\kappa}_H \times \mathbb{J}^{2,\kappa}_H$ *to the quadratic 2BSDEJ*

$$Y'_t = \xi + \int_t^T \widehat{F}'_s(Z'_s, U'_s)ds - \int_t^T Z'_s dB^c_s - \int_t^T \int_E U'_s(x)\widetilde{\mu}_{B^d}(ds,dx) + K'_T - K'_t, \tag{6.3.10}$$

where the generator is defined as follows

$$\widehat{F}'_t(\omega, z, u) := F'_t(\omega, z, u, \widehat{a}_t, \widehat{\nu}_t), \tag{6.3.11}$$

where

$$F'_t(z,u,a,\nu) := \inf_{\pi\in C}\left\{ \frac{\eta}{2}\left|\pi a^{1/2} - \left(a^{1/2}z + \frac{b_t + \int_E \beta_t(x)\nu(dx)}{a^{1/2}\eta}\right)\right|^2 + \frac{1}{\eta} j\left(\eta(u - \pi\beta_t)\right)\right\}$$
$$- \left(b_t + \int_E \beta_t(x)\nu(dx)\right) a^{1/2}z - \frac{|b_t + \int_E \beta_t(x)\nu(dx)|^2}{2a\eta},$$

where $j(u,\nu) := \int_E \left(e^{u(x)} - 1 - u(x)\right)\nu(dx)$. *Moreover,*

$$Y'_t = \underset{\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}} \ y_t'^{\mathbb{P}},$$

where $y'^{\mathbb{P}}$ *is the solution to the quadratic BSDE with the same terminal condition* ξ *and generator* $\widehat{F}'$. *Besides,* Y' *is also bounded quasi-surely.*

Proof.
Step 1: As in Morlais [92], we can verify that the generator $\widehat{F}'$ satisfies the following conditions.

(i) $\widehat{F}'$ has the quadratic growth property. There exists $(\alpha,\delta) \in \mathbb{R}_+ \times \mathbb{R}_+^*$ such that for all (t,z,u), $\mathcal{P}_H^\kappa - q.s.$ $\left|\widehat{F}'_t(0,0)\right| \leq \alpha$ and

$$-\left|\widehat{F}'_t(0,0)\right| - \frac{\delta}{2}\left|\widehat{a}_t^{1/2}z\right|^2 - \frac{1}{\delta}\widehat{j}_t(-\eta u) \leq \widehat{F}'_t(z,u) \leq \left|\widehat{F}'_t(0,0)\right| + \frac{\delta}{2}\left|\widehat{a}_t^{1/2}z\right|^2 + \frac{1}{\delta}\widehat{j}_t(\gamma u),$$

where $\widehat{j}_t(u) := \int_E \left(e^{u(x)} - 1 - u(x)\right)\widehat{\nu}_t(dx)$.

(ii) We have a "local Lipschitz" condition in z, $\exists \mu > 0$ and a progressively measurable process $\phi \in \mathbb{H}^{2,\kappa,H}_{BMO}$ such that for all (t,z,z',u), $\mathcal{P}_H^\kappa - q.s.$

$$\left|\widehat{F}'_t(z,u) - \widehat{F}'_t(z',u) - \phi_t.(\widehat{a}_t^{1/2}z - \widehat{a}_t^{1/2}z')\right| \leq \mu\left|\widehat{a}_t^{1/2}z - \widehat{a}_t^{1/2}z'\right|\left(\left|\widehat{a}_t^{1/2}z\right| + \left|\widehat{a}_t^{1/2}z'\right|\right).$$

(iii) For every (z,u,u') there exists a predictable and $\mathcal{E}$-mesurable process (γ_t) such that $\mathcal{P}_H^\kappa - q.s.$,

$$\widehat{F}'_t(z,u) - \widehat{F}'_t(z,u') \leq \int_0^t \int_E \gamma_s(x)\left(u(x) - u'(x)\right)\widehat{\nu}_t(dx)ds$$

$$\int_0^t \int_E \gamma'_s(x)\left(u(x) - u'(x)\right)\widehat{\nu}_t(dx)ds \leq \widehat{F}'_t(z,u) - \widehat{F}'_t(z,u'),$$

where there exists constants $C_1, C'_1 > -1$ and $C_2, C'_2 > 0$, independent of (z,u,u') such that

$$C_1(1 \wedge |x|) \leq \gamma_t(x) \leq C_2(1 \wedge |x|), \text{ and } C'_1(1 \wedge |x|) \leq \gamma'_t(x) \leq C'_2(1 \wedge |x|).$$

In particular, γ and γ' are in $\mathbb{J}^{2,\kappa,H}_{BMO}$.

We know from [92], that under each $\mathbb{P}$, the BSDEJ with terminal condition ξ and generator $\widehat{F}'$ has a solution, that we denote $(y'^{\mathbb{P}}, z'^{\mathbb{P}}, u'^{\mathbb{P}})$. Moreover, $y'^{\mathbb{P}}$ is bounded and there exists a bounded version of $u'^{\mathbb{P}}$ (that we still denote $u'^{\mathbb{P}}$) , with the following estimates

$$\left\|y'^{\mathbb{P}}\right\|_\infty \leq C(1 + \|\xi\|_\infty) \text{ and } |u'^{\mathbb{P}}| \leq 2\left\|y'^{\mathbb{P}}\right\|_\infty,$$

where $C > 0$. Next by applying Itô's formula, we have $y_t^{\mathbb{P}} = e^{\eta y'^{\mathbb{P}}}$, $t \in \mathbb{R}^+$, where $(y^{\mathbb{P}}, z^{\mathbb{P}}, u^{\mathbb{P}})$ is the solution under $\mathbb{P}$ of equation (6.3.9). This gives that

$$e^{-C'} \leq y_t^{\mathbb{P}} \leq e^{C'}, \ t \in \mathbb{R}^+, \ C' \in \mathbb{R}.$$

In particular $y^{\mathbb{P}}$ is strictly positive and bounded away from 0, and by representation (5.4.1), so is Y solving (6.3.9). Thus, we can make the following change of variables

$$Y_t' := \frac{1}{\eta}\log\left(Y_t\right).$$

Then by Itô's formula and the fact that K has only predictable jumps, we can verify that the triple (Y', Z', U') satisties (6.3.10) where

$$Z_t' := \frac{1}{\eta}\frac{Z_t}{Y_t},\ U_t' := \frac{1}{\eta}\log\left(1+\frac{U_t}{Y_{t-}}\right),\ K_t' := \int_0^t \frac{1}{\eta Y_s}dK_s^c - \sum_{0<s\leq t}\frac{1}{\eta}\log\left(1-\frac{\Delta K_s^d}{Y_{s-}}\right).$$

In particular, K' is predictable and nondecreasing with ${K'}_0 = 0$. Finally, due to the monotonicity of the function $\frac{1}{\eta}\log(x)$, we have the following representation for Y'

$$Y_t' = \underset{\mathbb{P}'\in\mathcal{P}_H^{\kappa}(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}}\ y_t'^{\mathbb{P}}.$$

Step 2: Next, we will prove the minimum condition for K'. As in [104] for 2BSDE with quadratic growth generator, we use the above representation of Y' and the properties verified by $\widehat{F}'$ in z and u.

Fix $\mathbb{P}$ in $\mathcal{P}_H^{\kappa}$ and $\mathbb{P}^{'} \in \mathcal{P}_H^{\kappa}(t^+,\mathbb{P})$, denote

$$\delta Y' := Y' - y'^{\mathbb{P}^{'}},\ \delta Z' := Z' - z'^{\mathbb{P}^{'}} \text{ and } \delta U' := U' - u'^{\mathbb{P}^{'}}.$$

By the "local Lipschitz" condition (ii) of $\widehat{F}'$ in z, there exist a process η with

$$|\eta_t| \leq \mu\left(\left|\widehat{a}_t^{1/2}Z_t\right| + \left|\widehat{a}_t^{1/2}z_t^{\mathbb{P}^{'}}\right|\right),\ \mathbb{P}^{'} - a.s.$$

such that we have for all $0\leq t\leq T$, $\mathbb{P}^{'} - a.s.$

$$\begin{aligned}\delta Y_t' &= \int_t^T (\eta_s+\phi_s)\widehat{a}_s^{\frac{1}{2}}\delta Z_s' ds - \int_t^T \delta Z_s' dB_s^c - \int_t^T\int_E \delta U_s'(x)\left[\widetilde{\mu}_{B^d}(ds,dx) - \gamma_s'(x)\widehat{\nu}(dx)ds\right]\\ &+ \int_t^T\left[\widehat{F}_s'(z_s',U_s') - \widehat{F}_s'(z_s',u_s')\right]ds - \int_t^T\int_E \gamma_s'(x)\delta U_s'(x)\widehat{\nu}(dx)ds + K_T' - K_t'. \end{aligned} \tag{6.3.12}$$

Let us now prove that we have $Z' \in \mathbb{H}^{2,\kappa,H}_{BMO}$, $U' \in \mathbb{J}^{2,\kappa,H}_{BMO}$ and that U' (or more precisely a version of U') is uniformly bounded. Actually, this can be proved using exactly the same arguments as in the proof of Lemma 3.3.1 in [104] on the one hand, that is to say applying Itô's formula to $e^{-\nu Y_t'}$ for a well chosen $\nu>0$. This allows then to get rid of the non-decreasing process K' in the estimates. Then on the other hand, one can use the same calculations as in the proof of Lemma 4.2.1 to obtain that

$$Z' \in \mathbb{H}^{2,\kappa,H}_{BMO},\ \left(e^{-\delta U'}-1\right) \in \mathbb{J}^{2,\kappa,H}_{BMO} \text{ and that } U' \text{ is uniformly bounded.}$$

Then, using the mean value theorem, it is easy to show that for every $t\in[0,T]$ and every $x\in E$, there exists some $c\in[-(U_t(x))^-, (U_t(x))^+]$ such that

$$U_t'(x) = \frac{1-e^{-\delta U_t'(x)}}{\delta}e^{\delta c}.$$

Since c is bounded, this clearly implies that $U' \in \mathbb{J}^{2,\kappa,H}_{BMO}$. Then the process η defined above is also in $\mathbb{H}^{2,\kappa,H}_{BMO}$. Since the process γ' is bounded and has jumps greater than $-1+\delta$, Kazamaki criterion (see [70]) implies that we can define an equivalent probability measure $\mathbb{Q}'$ such that

$$\frac{d\mathbb{Q}'}{d\mathbb{P}'} = \mathcal{E}\left(\int_0^{\cdot} (\eta_s + \phi_s)\,\widehat{a}_s^{-1/2} dB_s^c + \int_0^{\cdot}\int_E \gamma_s'(x)\widetilde{\mu}_{B^d}(ds,dx)\right).$$

By taking conditional expectations in (6.3.12) and using property (iii) satisfied by F', we obtain

$$Y_t' - y_t'^{\mathbb{P}'} \geq \mathbb{E}_t^{\mathbb{Q}'}\left[K_T' - K_t'\right].$$

For notational convenience, denote

$$\mathcal{E}_t^1 := \mathcal{E}\left(\int_0^t (\phi_s + \eta_s)\widehat{a}_s^{-1/2} dB_s^c\right) \text{ and } \mathcal{E}_t^2 := \mathcal{E}\left(\int_0^t\int_E \gamma_s'(x)\widetilde{\mu}_{B^d}(ds,dx)\right).$$

Let $r>1$ be the number given by Lemma 3.2.2 in [104] applied to $\mathcal{E}^1$. Then we estimate

$$\begin{aligned}
&\mathbb{E}_t^{\mathbb{P}'}\left[K'_T - K_t'\right]\\
&\leq \mathbb{E}_t^{\mathbb{P}'}\left[\frac{\mathcal{E}_T}{\mathcal{E}_t}\left(K_T' - K_t'\right)\right]^{\frac{1}{2r-1}} \mathbb{E}_t^{\mathbb{P}'}\left[\left(\frac{\mathcal{E}_t}{\mathcal{E}_T}\right)^{\frac{1}{2(r-1)}}\left(K_T' - K_t'\right)\right]^{\frac{2(r-1)}{2r-1}}\\
&\leq \left(\delta Y_t'\right)^{\frac{1}{2r-1}}\left(\mathbb{E}_t^{\mathbb{P}'}\left[\left(\frac{\mathcal{E}_t^1}{\mathcal{E}_T^1}\right)^{\frac{1}{r-1}}\right]\right)^{\frac{r-1}{2r-1}}\left(\mathbb{E}_t^{\mathbb{P}'}\left[\left(\frac{\mathcal{E}_t^2}{\mathcal{E}_T^2}\right)^{\frac{2}{r-1}}\right]\mathbb{E}_t^{\mathbb{P}'}\left[\left(K_T' - K_t'\right)^4\right]\right)^{\frac{r-1}{2(2r-1)}}\\
&\leq C\left(\mathbb{E}_t^{\mathbb{P}'}\left[\left(K_T'\right)^4\right]\right)^{\frac{r-1}{2(2r-1)}}\left(\delta Y_t'\right)^{\frac{1}{2r-1}},
\end{aligned}$$

where we used in the last inequality Lemma 3.2.2 in [104] for the term involving $\mathcal{E}^1$ and where we used Lemma B.2.4 for the term involving $\mathcal{E}^2$ (this is possible because of the properties verified by γ').

With the same argument as in Step (iii) of the proof of Theorem 5.4.1, the above inequality along with the representation for Y' shows that we have

$$\underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,inf}^{\mathbb{P}}} \mathbb{E}^{\mathbb{P}'}\left[K_T' - K_t'\right] = 0,$$

that is to say that the minimum condition 5.3.5 is verified. Indeed, the only thing to verify is that

$$\underset{\mathbb{P}'\in\mathcal{P}_H^\kappa(t^+,\mathbb{P})}{\operatorname{ess\,sup}^{\mathbb{P}}} \mathbb{E}_t^{\mathbb{P}'}\left[\left(K_T'\right)^4\right] < +\infty. \tag{6.3.13}$$

With the BMO properties satisfied by Z' and U' we can follow the proof of Theorem 3.2 in [104] to show that

$$\sup_{\mathbb{P}\in\mathcal{P}_H^\kappa} \mathbb{E}^{\mathbb{P}}\left[\left(K_T'\right)^4\right] < +\infty.$$

Then (6.3.13) can be obtained exactly as in the proof of Step (iii) of Theorem 5.4.1. □

Appendix A

Technical Results for Quadratic BSDEs with Jumps

Proposition A.0.3 *Let g be a function satisfying Assumption 4.2.1(i), which is uniformly Lipschitz in (y,z,u). Let Y be a càdlàg g-submartingale (resp. g-supermartingale) in $\mathcal{S}^\infty$. Assume further that g is concave (resp. convex) in (z,u) and that*

$$g_t(y,z,u) \geq -M - \beta\,|y| - \frac{\gamma}{2}\,|z|^2 - \frac{1}{\gamma}j_t(-\gamma u).$$
$$\left(\textit{resp. } g_t(y,z,u) \leq M + \beta\,|y| + \frac{\gamma}{2}\,|z|^2 + \frac{1}{\gamma}j_t(\gamma u)\right).$$

Then there exists a predictable non-decreasing (resp. non-increasing) process A null at 0 and processes $(Z,U)\in\mathbb{H}^2\times\mathbb{J}^2$ such that

$$Y_t = Y_T + \int_t^T g_s(Y_s,Z_s,U_s)ds - \int_t^T Z_s dB_s - \int_t^T\int_E U_s(x)\widetilde{\mu}(dx,ds) - A_T + A_t,\ t\in[0,T].$$

Proof. As usual, we can limit ourselves to the g-supermartingale case. Let us consider the following reflected BSDEJ with lower obstacle Y, generator g and terminal condition Y_T

$$\begin{aligned}
&\widetilde{Y}_t = Y_T + \int_t^T g_s(\widetilde{Y}_s,\widetilde{Z}_s,\widetilde{U}_s)ds - \int_t^T \widetilde{Z}_s dB_s - \int_t^T\int_E \widetilde{U}_s(x)\widetilde{\mu}(dx,ds) + A_T - A_t,\ t\in[0,T],\ \mathbb{P}-a.s.\\
&\widetilde{Y}_t \geq Y_t,\ t\in[0,T],\ \mathbb{P}-a.s.\\
&\int_0^T \left(\widetilde{Y}_{s^-} - Y_{s^-}\right) dA_s = 0,\ \mathbb{P}-a.s.
\end{aligned}\tag{A.0.1}$$

The existence and uniqueness of a solution consisting of a predictable non-decreasing process A and a triplet $(\widetilde{Y},\widetilde{Z},\widetilde{U})\in\mathcal{S}^2\times\mathbb{H}^2\times\mathbb{J}^2$ follow from the results of [58] for instance, since the generator is Lipschitz and the obstacle is càdlàg and bounded.

We will now show that the process $\widetilde{Y}$ must always be equal to the lower obstacle Y, which will provide us the desired decomposition. We proceed by contradiction and assume without loss of generality that $\widetilde{Y}_0 > Y_0$. For any $\varepsilon>0$, we now define the following bounded stopping time

$$\tau_\varepsilon := \inf\left\{t>0,\ \widetilde{Y}_t \leq Y_t + \varepsilon,\ \mathbb{P}-a.s.\right\}\wedge T.$$

By the Skorokhod condition, it is a classical result that the non-decreasing process A never acts before τ_ε. Therefore, we have for any $t\in[0,\tau_\varepsilon]$

$$\widetilde{Y}_t = \widetilde{Y}_{\tau_\varepsilon} + \int_t^{\tau_\varepsilon} g_s(\widetilde{Y}_s,\widetilde{Z}_s,\widetilde{U}_s)ds - \int_t^{\tau_\varepsilon} \widetilde{Z}_s dB_s - \int_t^{\tau_\varepsilon}\int_E \widetilde{U}_s(x)\widetilde{\mu}(dx,ds),\ \mathbb{P}-a.s. \tag{A.0.2}$$

Consider now the BSDEJ on $[0, \tau_\varepsilon]$ with terminal condition Y_{τ_ε} and generator g (existence and uniqueness of the solution are consequences of the result of [4] or [81])

$$\widehat{Y}_t = Y_{\tau_\varepsilon} + \int_t^{\tau_\varepsilon} g_s(\widehat{Y}_s, \widehat{Z}_s, \widehat{U}_s)ds - \int_t^{\tau_\varepsilon} \widehat{Z}_s dB_s - \int_t^{\tau_\varepsilon} \int_E \widehat{U}_s(x)\widehat{\mu}(dx, ds),\ \mathbb{P} - a.s. \tag{A.0.3}$$

Notice also that since Y and $g(0,0,0)$ are bounded, $\widehat{Y}$ and $\widetilde{Y}$ are also bounded, as a consequence of classical *a priori* estimates for Lipschitz BSDEJs and RBSDEJs. Then, using the fact that g is convex in (z, u) and that

$$g_t(y, z, u) \leq M + \beta |y| + \frac{\gamma}{2} |z|^2 + \frac{1}{\gamma} j_t(\gamma u),$$

we can proceed exactly as in the Step 2 of the proof of Proposition 4.2.7 to obtain that for any $\theta \in (0, 1)$

$$\begin{aligned}
\widetilde{Y}_0 - \theta \widehat{Y}_0 &\leq \frac{1-\theta}{\gamma} \ln \left(\mathbb{E}^{\mathbb{P}} \left[\exp \left(\gamma \int_0^{\tau_\varepsilon} \left(M + C \left| \widehat{Y}_t \right| + \frac{C}{1-\theta} \left| \widetilde{Y}_s - \theta \widehat{Y}_s \right| \right) ds + \frac{\gamma}{1-\theta} (\widetilde{Y}_{\tau_\varepsilon} - \theta Y_{\tau_\varepsilon}) \right) \right] \right) \\
&\leq (1-\theta) \ln(C_0) + C \left(\int_0^{\tau_\varepsilon} \left\| \widetilde{Y}_s - \theta \widehat{Y}_s \right\|_\infty ds + \left\| \widetilde{Y}_{\tau_\varepsilon} - \theta Y_{\tau_\varepsilon} \right\|_\infty \right),
\end{aligned}$$

where C_0 is some constant which does not depend on θ.

Letting θ go to 1, and using the fact that by definition $\widetilde{Y}_{\tau_\varepsilon} - \theta Y_{\tau_\varepsilon} \leq \varepsilon$, we obtain

$$\widetilde{Y}_0 - \widehat{Y}_0 \leq C\varepsilon + C \int_0^{\tau_\varepsilon} \left\| \widetilde{Y}_s - \widehat{Y}_s \right\|_\infty ds. \tag{A.0.4}$$

But since Y is a g-supermartingale, we have by definition that $\widehat{Y}_0 \leq Y_0$. Since we assumed that $\widetilde{Y}_0 > Y_0$, this implies that $\widetilde{Y}_0 > \widehat{Y}_0$. Therefore, we can use Gronwall's lemma in (A.0.4) to obtain

$$0 < \widetilde{Y}_0 - \widehat{Y}_0 \leq C_1 \varepsilon,$$

for some constant $C_1 > 0$, independent of ε.

By arbitrariness of ε, this gives us the desired contradiction and ends the proof. □

APPENDIX B

Technical Results for 2BSDEs with Jumps

B.1 The measures $\mathbb{P}^{\alpha,\nu}$

Lemma B.1.1 *Let τ be an $\mathbb{F}$-stopping time. Let $\omega \in \Omega$ and S be a bounded $\mathbb{F}^{\tau(\omega)}$-stopping time. There exists an $\mathbb{F}$ stopping time $\widetilde{S}$, such that*

$$S = \widetilde{S}^{\tau,\omega}. \tag{B.1.1}$$

Proof.
(B.1.1) means that

$$\forall \widetilde{\omega} \in \Omega^{\tau(\omega)}, \;\; S(\widetilde{\omega}) = \widetilde{S}^{\tau,\omega}(\widetilde{\omega}) = \widetilde{S}(\omega \otimes_{\tau(\omega)} \widetilde{\omega}).$$

We set

$$\forall \omega_1 \in \Omega, \; \widetilde{S}(\omega_1) := S(\omega_1^{\tau(\omega)}).$$

Using the fact that

$$\text{for } s \geq \tau(\omega), \; (\omega \otimes_{\tau(\omega)} \widetilde{\omega})^{\tau(\omega)}(s) = \omega(\tau(\omega)) + \widetilde{\omega}(s) - \omega(\tau(\omega)) - \widetilde{\omega}(\tau(\omega)) = \widetilde{\omega}(s),$$

we have that $\widetilde{S}$ satisfies (B.1.1) by construction, indeed

$$\widetilde{S}^{\tau,\omega}(\widetilde{\omega}) = S\left[\left(\omega \otimes_{\tau(\omega)} \widetilde{\omega}\right)^{\tau(\omega)}\right] = S(\widetilde{\omega}).$$

Notice now that for any $\omega \in \Omega$, $\widetilde{S} : \Omega \to [\tau(\omega), T]$ is measurable as a composition of measurable mappings.

Finally, let us prove that $\widetilde{S}$ is an $\mathbb{F}$-stopping time.
For $t < \tau(\omega)$, $\{\widetilde{S} \leq t\} = \{\emptyset\} \in \mathcal{F}_t$.
For $t \geq \tau(\omega)$,

$$\begin{aligned}
\{\widetilde{S} \leq t\} &= \{\omega_1 \in \Omega \,|\, S(\omega_1^{\tau(\omega)}) \leq t\} \\
&= \left\{\omega_1 \otimes_{\tau(\omega)} \omega_1^{\tau(\omega)} \in \Omega \,|\, S(\omega_1^{\tau(\omega)}) \leq t\right\} \\
&= \left\{\omega_1 \in \Omega \,|\, \omega_1^{\tau(\omega)} \in A^{\omega}\right\} \\
&= \Pi^{-1}_{\tau(\omega)}\left(A^{\omega}\right).
\end{aligned}$$

where $A^\omega := \{\widetilde{\omega} \in \Omega^{\tau(\omega)} \,|\, S(\widetilde{\omega}) \leq t\} \in \mathcal{F}_t^{\tau(\omega)}$ and

$\Pi_{\tau(\omega)} : (\Omega, \mathcal{F}_t) \to \left(\Omega^{\tau(\omega)}, \mathcal{F}_t^{\tau(\omega)}\right)$ is defined by $\Pi_{\tau(\omega)}(\omega_1) := w_1^{\tau(\omega)}$. By definition of the measurability of $\Pi_{\tau(\omega)}$, we have $\{\widetilde{S} \leq t\} = \Pi^{-1}_{\tau(\omega)}(A^\omega) \in \mathcal{F}_t$. □

Lemma B.1.2 *Let $\mathbb{P} \in \mathcal{P}_{\tilde{\mathcal{A}}}$ and τ be an $\mathbb{F}^B$-stopping time. Then*

$$\mathbb{P}^{\tau,\omega} \in \mathcal{P}^{\tau(\omega)}_{\tilde{\mathcal{A}}} \textit{ for } \mathbb{P}-a.e.\ \omega \in \Omega.$$

Proof.

<u>**Step** 1</u>: Let us first prove that

$$(\mathbb{P}_\nu)^{\tau,\omega} = \mathbb{P}^{\tau(\omega)}_{\nu^{\tau,\omega}},\ \ \mathbb{P}_\nu\text{-a.s. on } \Omega, \tag{B.1.2}$$

where $(\mathbb{P}_\nu)^{\tau,\omega}$ denotes a probability measure on Ω^τ, constructed from a regular conditional probability distribution (r.c.p.d.) of $\mathbb{P}_\nu$ for the stopping time τ, evaluated at ω, and $\mathbb{P}^{\tau(\omega)}_{\nu^{\tau,\omega}}$ is the unique solution of the martingale problem $(\mathbb{P}^1, \tau(\omega), T, Id, \nu^{\tau,\omega})$, where $\mathbb{P}^1$ is such that $\mathbb{P}^1(B^\tau_\tau = 0) = 1$.

It is enough to show that outside a $\mathbb{P}_\nu$-negligible set, the shifted processes M^τ, J^τ, Q^τ are $(\mathbb{P}_\nu)^{\tau,\omega}$-local martingales, where M, J and Q are defined in Remark 5.2.1. For this, for any ω in Ω, take a bounded $\mathbb{F}^{\tau(\omega)}$-stopping time S. By Lemma B.1.1, there exists a bounded $\mathbb{F}$-stopping time $\tilde{S}$ such that $S = \tilde{S}^{\tau,\omega}$. Then, following the definitions in Subsection 6.2.1,

$$\Delta B^{\tau,\omega}_S(\widetilde{\omega}) = \Delta B_S(\omega \otimes_\tau \widetilde{\omega}) = \Delta(\omega \otimes_\tau \widetilde{\omega})(S) = \Delta\omega_S \mathbf{1}_{\{S\leq\tau\}} + \Delta\widetilde{\omega}_S\mathbf{1}_{\{S>\tau\}},$$

and that for $S \geq \tau$

$$B_S(\omega \otimes_\tau \widetilde{\omega}) = (\omega \otimes_\tau \widetilde{\omega})(S) = \omega_\tau + \widetilde{\omega}_S = B_\tau(\omega) + B^\tau_S(\widetilde{\omega}).$$

From this we get

$$\begin{aligned}
M^{\tau,\omega}_S(\widetilde{\omega}) = M_S(\omega \otimes_\tau \widetilde{\omega}) =& B_S(\omega \otimes_\tau \widetilde{\omega}) - \sum_{u\leq S} \mathbf{1}_{|\Delta B_u(\omega\otimes_\tau\widetilde{\omega})|>1} \Delta B_u(\omega \otimes_\tau \widetilde{\omega})\\
&+ \int_0^S \int_E x\mathbf{1}_{|x|>1}\nu_u(\omega \otimes_\tau \widetilde{\omega})(dx)\, du\\
=& B^\tau_S(\widetilde{\omega}) + B_t(\omega) - \sum_{u\leq\tau} \mathbf{1}_{|\Delta\omega_u|>1}\Delta\omega_u - \sum_{\tau<u\leq S} \mathbf{1}_{|\Delta B^\tau_u(\widetilde{\omega})|>1}\Delta B^\tau_u(\widetilde{\omega})\\
&+ \int_0^\tau \int_E x\mathbf{1}_{|x|>1}\nu_u(\omega)(dx)\, du + \int_\tau^S \int_E x\mathbf{1}_{|x|>1}\nu^{\tau,\omega}_u(\widetilde{\omega})(dx)\, du\\
=& M^\tau_S(\widetilde{\omega}) + M_\tau(\omega),\ \ \forall\omega \in \Omega.
\end{aligned}$$

And we can now compute

$$\begin{aligned}
\mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}[M^\tau_S] &= \mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[M^{\tau,\omega}_S - M_\tau(\omega)\right]\\
&= \mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[M^{\tau,\omega}_{\tilde{S}^{\tau,\omega}}\right] - M_\tau(\omega)\\
&= \mathbb{E}^{\mathbb{P}_\nu}_\tau[M_{\tilde{S}}](\omega) - M_\tau(\omega) = 0, \text{for } \mathbb{P}_\nu\text{-a.e. } \omega.
\end{aligned} \tag{B.1.3}$$

Since S is an arbitrary bounded stopping time, we have that M^τ is a $(\mathbb{P}_\nu)^{\tau,\omega}$-local martingale for $\mathbb{P}_\nu$-a.e. ω.

Notice that since M is a càdlàg local martingale under $\mathbb{P}_\nu$, we can find a $\mathbb{P}_\nu$-negligible set N, such that for any bounded $\mathbb{F}$-stopping times $\sigma_1 \leq \sigma_2$, for any ω outside of N, we have $\mathbb{E}^{\mathbb{P}_\nu}_{\sigma_1}[M_{\sigma_2}](\omega) = M_{\sigma_1}(\omega)$. In particular, the negligible set appearing in the equality (B.1.3) is independent of the choice of $\widetilde{S}$.

We treat the case of the process J^τ analogously and write

$$\begin{aligned}
J_S^{\tau,\omega}(\widetilde{\omega}) &= \big(M_S^{\tau,\omega}(\widetilde{\omega})\big)^2 - S - \int_0^\tau \int_E x^2 \nu_u(\omega)(dx)du - \int_\tau^S \int_E x^2 \nu_u^{\tau,\omega}(\widetilde{\omega})(dx)du \\
&= (M_S^\tau(\widetilde{\omega}))^2 + (M_\tau(\omega))^2 + 2M_S^\tau(\widetilde{\omega})M_\tau(\omega) - (S-\tau) - \int_\tau^S \int_E x^2 \nu_u^{\tau,\omega}(\widetilde{\omega})(dx)du \\
&\quad - \int_0^\tau \int_E x^2 \nu_u(\omega)(dx)du - \tau \\
&= J_S^\tau(\widetilde{\omega}) + J_\tau(\omega) + 2M_S^\tau(\widetilde{\omega})M_\tau(\omega).
\end{aligned}$$

Then we can compute the expectation

$$\mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[J_S^\tau\right] = \mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[J_S^{\tau,\omega} - 2M_S^\tau M_\tau(\omega)\right] - J_\tau(\omega) = 0, \text{for } \mathbb{P}_\nu\text{-a.e. } \omega.$$

J^τ is then a $(\mathbb{P}_\nu)^{\tau,\omega}$-local martingale for $\mathbb{P}_\nu$-a.e. ω.. Finally, we do the same kind of calculation for Q^τ, and we obtain

$$\begin{aligned}
Q_S^{\tau,\omega}(\widetilde{\omega}) &= \int_0^S \int_E g(x)\mu_B(\omega \otimes_\tau \widetilde{\omega}, dx, du) - \int_0^S \int_E g(x)\nu_u^{\tau,\omega}(\widetilde{\omega})(dx)\,du \\
&= \int_0^\tau \int_E g(x)\mu_B(\omega, dx, du) + \int_\tau^S \int_E g(x)\mu_{B^\tau}(\widetilde{\omega}, dx, du) \\
&\quad - \int_0^\tau \int_E g(x)\nu_u(\omega)(dx)du - \int_\tau^S \int_E g(x)\nu_u^{\tau,\omega}(\widetilde{\omega})(dx)\,du \\
&= Q_S^\tau(\widetilde{\omega}) + Q_\tau(\omega).
\end{aligned}$$

And again we compute the expectation over the $\widetilde{\omega} \in \Omega^\tau$, under the measure $(\mathbb{P}_\nu)^{\tau,\omega}$

$$\begin{aligned}
\mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[Q_S^\tau\right] &= \mathbb{E}^{(\mathbb{P}_\nu)^{\tau,\omega}}\left[Q_S^{\tau,\omega} - Q_\tau(\omega)\right] \\
&= \mathbb{E}_\tau^{\mathbb{P}_\nu}[Q_S](\omega) - Q_\tau(\omega) = 0, \text{for } \mathbb{P}_\nu\text{-a.e. } \omega.
\end{aligned}$$

We have the desired result, and conclude that (B.1.2) holds true. We can now deduce that for any $(\alpha,\nu) \in \tilde{\mathcal{A}}$

$$\mathbb{P}^{\alpha^{\tau,\omega},\nu^{\tau,\omega}} \in \mathcal{P}_{\tilde{\mathcal{A}}}^{\tau(\omega)}\ \mathbb{P}_\nu\text{-a.s. on } \Omega. \tag{B.1.4}$$

Indeed, if $(\alpha,\nu) \in \mathcal{D}\times\mathcal{N}$, then $(\alpha^{\tau,\omega},\nu^{\tau,\omega}) \in \mathcal{D}^{\tau(\omega)}\times\mathcal{N}^{\tau(\omega)}$, because

$$\int_{\tau(\omega)}^T \int_E (1\wedge|x|^2)\nu_s^{\tau,\omega}(\widetilde{\omega})(dx)ds + \int_{\tau(\omega)}^T \int_{|x|>1} x\nu_s^{\tau,\omega}(\widetilde{\omega})(dx)ds + \int_{\tau(\omega)}^T \left|\alpha_s^{\tau,\omega}(\widetilde{\omega})\right| ds < \infty.$$

Step 2: We define $\tilde{\tau} := \tau \circ X^\alpha$, $\tilde{\alpha}^{\tau,\omega} := \alpha^{\tilde{\tau},\beta_\alpha(\omega)}$ and $\tilde{\nu}^{\tau,\omega} := \nu^{\tilde{\tau},\beta_\alpha(\omega)}$ where β_α is a measurable map such that $B = \beta_\alpha(X^\alpha)$, $\mathbb{P}_\nu$-a.s. Moreover, $\tilde{\tau}$ is a stopping time, we have $\tau = \tilde{\tau} \circ \beta_\alpha$ and using (B.1.4),

$$\mathbb{P}^{\tilde{\alpha}^{\tau,\omega},\tilde{\nu}^{\tau,\omega}} \in \mathcal{P}^{\tau(\omega)}_{\tilde{\mathcal{A}}} \ \ \mathbb{P}^{\alpha,\nu}\text{-a.s. on } \Omega.$$

Step 3: We show that

$$\mathbb{E}^{\mathbb{P}^{\alpha,\nu}}\left[\phi\left(B_{t_1\wedge\tau},\dots,B_{t_n\wedge\tau}\right)\psi\left(B_{t_1},\dots,B_{t_n}\right)\right] = \mathbb{E}^{\mathbb{P}_{\alpha,\nu}}\left[\phi\left(B_{t_1\wedge\tau},\dots,B_{t_n\wedge\tau}\right)\psi_\tau\right]$$

for every $0 < t_1 < \dots < t_n \leq T$, every continuous and bounded functions ϕ and ψ and

$$\psi_\tau(\omega) = \mathbb{E}^{\mathbb{P}^{\tilde{\alpha}^{\tau,\omega},\tilde{\nu}^{\tau,\omega}}}\Big[\psi(\omega(t_1),\dots,\omega(t_k),\omega(t)+B^t_{t_{k+1}},\dots,\omega(t)+B^t_{t_n})\Big],$$

for $t := \tau(\omega) \in [t_k, t_{k+1})$.

Recall that $\mathbb{P}^{\tilde{\alpha}^{\tau,\omega},\tilde{\nu}^{\tau,\omega}}$ is defined by $\mathbb{P}^{\tilde{\alpha}^{\tau,\omega},\tilde{\nu}^{\tau,\omega}} = \mathbb{P}_{\tilde{\nu}^{\tau,\omega}} \circ \left(X^{\tilde{\alpha}^{\tau,\omega}}\right)^{-1}$, then

$$\begin{aligned}\psi_\tau(\omega) = \mathbb{E}^{\mathbb{P}^{\tau(\omega)}_{\tilde{\nu}^{\tau,\omega}}}\Big[\psi\Big(&\omega(t_1),\dots,\omega(t_k),\omega(t) + \int_t^{t_{k+1}}\left(\alpha_s^{\tilde{\tau},\beta_\alpha(\omega)}\right)^{1/2} d(B^c_s)^{\tau(\omega)}\\ &+ \int_t^{t_{k+1}}\int_E x(\mu_{B^{\tau(\omega)}}(ds,dx) - \nu_s^{\tilde{\tau},\beta_\alpha(\omega)}(dx)ds),\dots,\omega(t) + \int_t^{t_n}\left(\alpha_s^{\tilde{\tau},\beta_\alpha(\omega)}\right)^{1/2} d(B^c_s)^{\tau(\omega)}\\ &+ \int_t^{t_n}\int_E x(\mu_{B^{\tau(\omega)}}(ds,dx) - \nu_s^{\tilde{\tau},\beta_\alpha(\omega)}(dx)ds)\Big)\Big].\end{aligned}$$

Then, $\forall\omega\in\Omega$, if $t := \tilde{\tau}(\omega) = \tau\left(X^\alpha(\omega)\right) \in [t_k, t_{k+1}[$,

$$\begin{aligned}\psi_\tau\left(X^\alpha(\omega)\right) = \mathbb{E}^{\mathbb{P}^{\tilde{\tau}(\omega)}_{\nu^{\tilde{\tau},\omega}}}\Big[\psi\Big(&X^\alpha_{t_1}(\omega),\dots,X^\alpha_{t_k}(\omega),X^\alpha_t(\omega) + \int_t^{t_{k+1}}\left(\alpha_s^{\tilde{\tau},\omega}\right)^{1/2} d(B_s^{\tilde{\tau}(\omega)})^c\\ &+ \int_t^{t_{k+1}}\int_E x(\mu_{B^{\tilde{\tau}(\omega)}}(ds,dx) - \nu_s^{\tilde{\tau},\omega}(dx)ds),\dots,X^\alpha_t(\omega) + \int_t^{t_n}\left(\alpha_s^{\tilde{\tau},\omega}\right)^{1/2} d(B_s^{\tilde{\tau}(\omega)})^c\\ &+ \int_t^{t_n}\int_E x(\mu_{B^{\tilde{\tau}(\omega)}}(ds,dx) - \nu_s^{\tilde{\tau},\omega}(dx)ds)\Big)\Big].\end{aligned} \tag{B.1.5}$$

We remark that for every $\omega\in\Omega$,

$$\begin{aligned}&\alpha_s(\omega) = \alpha_s\left(\omega\otimes_{\tilde{\tau}(\omega)}\omega^{\tilde{\tau}(\omega)}\right) = \alpha_s^{\tilde{\tau},\omega}\left(\omega^{\tilde{\tau}(\omega)}\right)\\ &\text{and } \ \nu_s(\omega)(dx) = \nu_s\left(\omega\otimes_{\tilde{\tau}(\omega)}\omega^{\tilde{\tau}(\omega)}\right)(dx) = \nu_s^{\tilde{\tau},\omega}\left(\omega^{\tilde{\tau}(\omega)}\right)(dx)\end{aligned}$$

By definition, the $(\mathbb{P}_\nu)^{\tilde{\tau},\omega}$-distribution of $B^{\tilde{\tau}(\omega)}$ is equal to the $(\mathbb{P}_\nu)^\omega_{\tilde{\tau}}$-distribution of $(B_\cdot - B_{\tilde{\tau}(\omega)})$. (B.1.5) then becomes

$$\begin{aligned}\psi_\tau\left(X^\alpha(\omega)\right) =& \mathbb{E}^{(\mathbb{P}_\nu)^\omega_{\tilde{\tau}}}\Big[\psi\Big(X^\alpha_{t_1}(\omega),\dots,X^\alpha_{t_k}(\omega),X^\alpha_t(\omega) + \int_t^{t_{k+1}}\alpha_s^{1/2}(B^c)d(B^c_s)\\ &+ \int_t^{t_{k+1}}\int_E x(\mu_B(ds,dx) - \nu_s(dx)ds),\dots,X^\alpha_t(\omega) + \int_t^{t_n}\alpha_s^{1/2}(B^c)d(B^c_s)\\ &+ \int_t^{t_n}\int_E x(\mu_B(ds,dx) - \nu_s(dx)ds)\Big)\Big]\\ =& \mathbb{E}^{(\mathbb{P}_\nu)^\omega_{\tilde{\tau}}}\Big[\psi\Big(X^\alpha_{t_1},\dots,X^\alpha_{t_k},X^\alpha_{t_{k+1}},\dots,X^\alpha_{t_n}\Big)\Big]\\ =& \mathbb{E}^{\mathbb{P}_\nu}\Big[\psi\Big(X^\alpha_{t_1},\dots,X^\alpha_{t_k},X^\alpha_{t_{k+1}},\dots,X^\alpha_{t_n}\Big)|\mathcal{F}_{\tilde{\tau}}\Big](\omega), \ \ \mathbb{P}_\nu\text{-a.s. on } \Omega.\end{aligned}$$

Then we have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^{\alpha,\nu}}\left[\phi\left(B_{t_1\wedge\tau},\ldots,B_{t_n\wedge\tau}\right)\psi_\tau\right] = \mathbb{E}^{\mathbb{P}_\nu}\left[\phi\Big(X^\alpha_{t_1\wedge\tilde\tau},\ldots,X^\alpha_{t_n\wedge\tilde\tau}\Big)\psi_{\tilde\tau}(X^\alpha)\right]\\
&= \mathbb{E}^{\mathbb{P}_\nu}\left[\phi\Big(X^\alpha_{t_1\wedge\tilde\tau},\ldots,X^\alpha_{t_n\wedge\tilde\tau}\Big)\mathbb{E}^{\mathbb{P}_\nu}\left[\psi\Big(X^\alpha_{t_1},\ldots,X^\alpha_{t_k},X^\alpha_{t_{k+1}},\ldots,X^\alpha_{t_n}\Big)|\mathcal{F}_{\tilde\tau}\right]\right]\\
&= \mathbb{E}^{\mathbb{P}_\nu}\left[\phi\Big(X^\alpha_{t_1\wedge\tilde\tau},\ldots,X^\alpha_{t_n\wedge\tilde\tau}\Big)\psi\Big(X^\alpha_{t_1},\ldots,X^\alpha_{t_k},X^\alpha_{t_{k+1}},\ldots,X^\alpha_{t_n}\Big)\right]\\
&= \mathbb{E}^{\mathbb{P}_{\alpha,\nu}}\left[\phi\left(B_{t_1\wedge\tau},\ldots,B_{t_n\wedge\tau}\right)\psi\left(B_{t_1},\ldots,B_{t_n}\right)\right].
\end{aligned}$$

<u>**Step** 4</u>: Now we prove that $\mathbb{P}^{\tau,\omega} = \mathbb{P}^{\tilde\alpha^{\tau,\omega},\tilde\nu^{\tau,\omega}}$, $\mathbb{P}$-a.s. on Ω.

By definition of the conditional expectation,

$$\psi_\tau(\omega) = \mathbb{E}^{\mathbb{P}^{\tau,\omega}}\left[\psi(\omega(t_1),\ldots,\omega(t_k),\omega(t)+B^t_{t_{k+1}},\ldots,\omega(t)+B^t_{t_n})\right],\ \mathbb{P}^{\alpha,\nu}\text{-a.s.},$$

where $t := \tau(\omega) \in [t_k, t_{k+1}[$, and where the $\mathbb{P}^{\alpha,\nu}$-null set can depend on $(t_1,\ldots,t_n)$ and ψ, but we can choose a common null set by standard approximation arguments.

Then by a density argument we obtain

$$\mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\eta] = \mathbb{E}^{\mathbb{P}^{\tilde\alpha^{\tau,\omega},\tilde\nu^{\tau,\omega}}}[\eta],\ \text{for } \mathbb{P}^{\alpha,\nu}\text{-a.e. } \omega,$$

for every bounded and $\mathcal{F}^{\tau(\omega)}_T$-mesurable random variable η. This implies $\mathbb{P}^{\tau,\omega} = \mathbb{P}^{\tilde\alpha^{\tau,\omega},\tilde\nu^{\tau,\omega}}$, $\mathbb{P}$-a.s. on Ω. And from the Step 1 we deduce that $\mathbb{P}^{\tau,\omega} \in \overline{\mathcal{P}}^{\tau(\omega)}_{\tilde{\mathcal{A}}}$. □

Lemma B.1.3 *We have $\mathbb{P}^n \in \mathcal{P}^\kappa_H$, where $\mathbb{P}^n$ is defined by (6.2.9).*

Proof. Since by definition, $\mathbb{P}^i_t \in \mathcal{P}^t_H$ and $\mathbb{P} \in \mathcal{P}_H$, we have $\mathbb{P}^i_t = \mathbb{P}^{\alpha^i,\nu^i}$ and $\mathbb{P} = \mathbb{P}^{\alpha,\nu}$, for $(\alpha^i,\nu^i) \in \widetilde{\mathcal{A}}^t$ and $(\alpha,\nu) \in \widetilde{\mathcal{A}}$, $i = 1,\ldots,n$. Next we define

$$\begin{aligned}
\overline{\alpha}_s &:= \alpha_s \mathbf{1}_{[0,t)}(s) + \left[\sum_{i=1}^n \alpha^i_s \mathbf{1}_{E^i_t}(X^\alpha) + \alpha_s \mathbf{1}_{\hat E^n_t}(X^\alpha)\right]\mathbf{1}_{[t,T]}(s),\ \text{and}\\
\overline{\nu}_s &:= \nu_s \mathbf{1}_{[0,t)}(s) + \left[\sum_{i=1}^n \nu^i_s \mathbf{1}_{E^i_t}(X^\alpha) + \nu_s \mathbf{1}_{\hat E^n_t}(X^\alpha)\right]\mathbf{1}_{[t,T]}(s).
\end{aligned}$$

Now following the arguments in the proof of step 3 of Lemma B.1.2, we prove that for any $0 < t_1 < \cdots < t_k = t < t_{k+1} < t_n$ and any continuous and bounded functions ϕ and ψ,

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^{\alpha,\nu}}&\left[\phi(B_{t_1},\ldots,B_{t_k})\sum_{i=1}^n \mathbb{E}^{\mathbb{P}^{\alpha^i_t,\nu^i_t}}\left[\psi(B_{t_1},\ldots,B_{t_k},B_t+B^t_{t_{k+1}},\ldots,B_t+B^t_{t_n})\right]\mathbf{1}_{E^i_t}\right]\\
&= \mathbb{E}^{\mathbb{P}^{\overline\alpha,\overline\nu}}\left[\phi(B_{t_1},\ldots,B_{t_k})\psi(B_{t_1},\ldots,B_{t_n})\right].
\end{aligned}$$

This implies that $\mathbb{P}^n = \mathbb{P}^{\overline\alpha,\overline\nu} \in \mathcal{P}_{\widetilde{\mathcal{A}}}$. And since all the probability measures $\mathbb{P}^i$ satisfy the requirements of Definition 6.2.1, we have $\mathbb{P}^n = \mathbb{P}^{\overline\alpha,\overline\nu} \in \mathcal{P}^\kappa_H$. □

B.2 Other properties

B.2.1 Upward directed set

Lemma B.2.1 *Fix an arbitrary measure $\mathbb{P} = \mathbb{P}^{\alpha,\nu}$ in $\mathcal{P}_H^\kappa$. The set $\mathcal{P}_H^\kappa(t^+, \mathbb{P})$ is upward directed, i.e. for each $\mathbb{P}_1 := \mathbb{P}^{\alpha_1,\nu_1}$ and $\mathbb{P}_2 := \mathbb{P}^{\alpha_2,\nu_2}$ in $\mathcal{P}_H^\kappa(t^+, \mathbb{P})$, there exists $\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+, \mathbb{P})$ such that $\forall u > t$,*

$$\mathbb{E}_t^{\mathbb{P}'}\left[(K_u - K_t)^2\right] = \max\left\{\mathbb{E}_t^{\mathbb{P}_1}\left[(K_u - K_t)^2\right], \mathbb{E}_t^{\mathbb{P}_2}\left[(K_u - K_t)^2\right]\right\}. \tag{B.2.1}$$

Proof.

We define the following $\mathcal{F}_t$-measurable sets

$$E_1 := \left\{\omega \in \Omega \,:\, \mathbb{E}_t^{\mathbb{P}_2}\left[(K_u - K_t)^2\right](\omega) \leq \mathbb{E}_t^{\mathbb{P}_1}\left[(K_u - K_t)^2\right](\omega)\right\},$$

and $E_2 := \Omega \backslash E_1$. Then for all $A \in \mathcal{F}_T$, we define the probability measure $\mathbb{P}'$ by,

$$\mathbb{P}'(A) := \mathbb{P}_1(A \cap E_1) + \mathbb{P}_2(A \cap E_2).$$

By definition, $\mathbb{P}'$ satisfies (B.2.1). Let us prove now that $\mathbb{P}' \in \mathcal{P}_H^\kappa(t^+, \mathbb{P})$. As in the proof of claim (4.17) in [112], for $s \in [0, T]$, we define the processes α^* and ν^* as follows

$$\begin{aligned}
\alpha_s^*(\omega) &:= \alpha_s(\omega)\mathbf{1}_{[0,t)}(s) + \left(\alpha_s^1(\omega)\mathbf{1}_{\{X^\alpha \in E_1\}}(\omega) + \alpha_s^2(\omega)\mathbf{1}_{\{X^\alpha \in E_2\}}(\omega)\right)\mathbf{1}_{[t,T]}(s),\\
\nu_s^*(\omega) &:= \nu_s(\omega)\mathbf{1}_{[0,t)}(s) + \left(\nu_s^1(\omega)\mathbf{1}_{\{X^\alpha \in E_1\}}(\omega) + \nu_s^2(\omega)\mathbf{1}_{\{X^\alpha \in E_2\}}(\omega)\right)\mathbf{1}_{[t,T]}(s),
\end{aligned}$$

where X^α is defined in (5.2.6). We have

$$0 < \underline{\alpha} \wedge \underline{\alpha}^1 \wedge \underline{\alpha}^2 \leq \alpha^* \leq \overline{\alpha} \vee \overline{\alpha}^1 \vee \overline{\alpha}^2,$$

where $\underline{\alpha}$, $\overline{\alpha}$, $\underline{\alpha}^i$, $\overline{\alpha}^i$ are the lower and upper bounds of the processes α, α^1 and α^2. Next, we have

$$\begin{aligned}
\int_0^T \int_E (1 \wedge |x|^2)\nu_s^*(ds, dx) \leq &\int_0^T \int_E (1 \wedge |x|^2)\nu_s(ds, dx) + \int_0^T \int_E (1 \wedge |x|^2)\nu_s^1(ds, dx)\\
&+ \int_0^T \int_E (1 \wedge |x|^2)\nu_s^2(ds, dx) < +\infty,
\end{aligned}$$

and the same way we see that $\int_0^T \int_{\{|x|>1\}} x\nu_s^*(ds, dx) < +\infty$. Then, we have therefore clearly $(\alpha^*, \nu^*) \in \tilde{\mathcal{A}}$, and we can define the element $\mathbb{P}^{\alpha^*,\nu^*}$ of $P_{\tilde{\mathcal{A}}}$.

Now using the same arguments as in the Step 3 of the proof of Lemma B.1.2, we obtain that for any $t_1 < \cdots < t_k = t < t_{k+1} < \cdots < t_n$ and any bounded and continuous functions ϕ and ψ,

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}^{\alpha^*,\nu^*}}\left[\phi(B_{t_1}, \ldots, B_{t_k})\psi(B_{t_1}, \ldots, B_{t_n})\right]\\
&= \mathbb{E}^{\mathbb{P}^{\alpha,\nu}}\Big[\phi(B_{t_1}, \ldots, B_{t_k})\Big(\mathbb{E}^{\mathbb{P}^{\alpha_1,\nu_1}}[\psi(B_{t_1}, \ldots, B_{t_k}, B_t + B_{t_{k+1}}^t, \ldots, B_t + B_{t_n}^t)]\mathbf{1}_{E_1}\\
&\quad + \mathbb{E}^{\mathbb{P}^{\alpha_2,\nu_2}}[\psi(B_{t_1}, \ldots, B_{t_k}, B_t + B_{t_{k+1}}^t, \ldots, B_t + B_{t_n}^t)]\mathbf{1}_{E_2}\Big)\Big].
\end{aligned}$$

This shows that

$$(\mathbb{P}^{\alpha^*,\nu^*})_t^\omega = (\mathbb{P}')_t^\omega \text{ for } \mathbb{P}^{\alpha^*,\nu^*}\text{-almost every } \omega \text{ in } \Omega \text{ and every } t > 0.$$

Thus $\mathbb{P}^{'} = \mathbb{P}^{\alpha^*,\nu^*} \in \mathcal{P}_{\tilde{\mathcal{A}}}$. To prove that $\mathbb{P}^{'} \in \mathcal{P}_H^{\kappa}$, we compute

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^{'}}\left[\int_0^T \left|\hat{F}_s^0\right|^2 ds\right] &= \mathbb{E}^{\mathbb{P}}\left[\int_0^t \left|\hat{F}_s^0\right|^2 ds\right] + \mathbb{E}^{\mathbb{P}_1}\left[\int_t^T \left|\hat{F}_s^0\right|^2 ds\,\mathbf{1}_{E_1}\right] + \mathbb{E}^{\mathbb{P}_2}\left[\int_t^T \left|\hat{F}_s^0\right|^2 ds\,\mathbf{1}_{E_2}\right]\\ &\leq \mathbb{E}^{\mathbb{P}}\left[\int_0^T \left|\hat{F}_s^0\right|^2 ds\right] + \mathbb{E}^{\mathbb{P}_1}\left[\int_0^T \left|\hat{F}_s^0\right|^2 ds\,\mathbf{1}_{E_1}\right] + \mathbb{E}^{\mathbb{P}_2}\left[\int_0^T \left|\hat{F}_s^0\right|^2 ds\,\mathbf{1}_{E_2}\right] < +\infty.\end{aligned}$$

Since by construction $\mathbb{P}^{'}$ coincides with $\mathbb{P}$ on $\mathcal{F}_t$, the proof is complete.

□

B.2.2 L^r-Integrability of exponential martingales

Lemma B.2.2 *Let $\delta > 0$ and $n \in \mathbb{N}^*$. Then there exists a constant $C_{n,\delta}$ depending only on δ and n such that*

$$(1+x)^{-n} - 1 + nx \leq C_{n,\delta}x^2, \text{ for all } x \in [-1+\delta, +\infty).$$

Proof. The inequality is clear for x large enough, let's say $x \geq M$ for some $M > 0$. Then, a simple Taylor expansion shows that this also holds in a neighborhood of 0, that is to say for $x \in [-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$. Finally, for $x \in (-1+\delta, -\varepsilon) \cup (\varepsilon, M)$, it is clear that we can choose C large enough such that the inequality also holds. □

Mémin [89] and then Lépingle and Mémin [80] proved some useful multiplicative decompositions of exponential semimartingales. We give here one of these representations that we will use in the proof of Lemma B.2.3.

Proposition B.2.1 (Proposition II.1 of [79]) *Let N be a local martingale and let A be a predictable process with finite variation such that $\Delta A \neq -1$. We assume $N_0 = A_0 = 0$. Then there exists a local martingale $\tilde{N}$ with $\tilde{N}_0 = 0$ and such that*

$$\mathcal{E}(N + A) = \mathcal{E}(\tilde{N})\mathcal{E}(A).$$

Lemma B.2.3 *Let $\lambda > 0$ and M be a local martingale with bounded jumps, such that $\Delta M \geq -1+\delta$, for a fixed $\delta > 0$. Let $V^{-\lambda}$ be the predictable compensator of*

$$\left\{W_t^{-\lambda} = \sum_{s\leq t}\left[(1+\Delta M_s)^{-\lambda} - 1 + \lambda\Delta M_s\right],\ t \geq 0\right\}.$$

We have

(i) $\mathcal{E}^{-\lambda}(M) = \mathcal{E}(N^{-\lambda} + A^{-\lambda})$ *where*

$$A^{-\lambda} = \frac{\lambda(\lambda+1)}{2}\langle M^c, M^c\rangle^T + V^{-\lambda},\ N^{-\lambda} = -\lambda M^T + W^{-\lambda} - V^{-\lambda}.$$

(ii) There exist a local martingale $\tilde{N}^{-\lambda}$ such that

$$\mathcal{E}^{-\lambda}(M) = \mathcal{E}(\tilde{N}^{-\lambda})\mathcal{E}(A^{-\lambda}).$$

Proof. First note that thanks to lemma B.2.2, for $\lambda > 0$, $(1+x)^{-\lambda} - 1 + \lambda x \le Cx^2$, and thus $W^{-\lambda}$ is integrable. We set

$$T_n = \inf\{t \ge 0 \ : \ \mathcal{E}(M)_t \le \frac{1}{n}\} \text{ and } M_t^n = M_{t\wedge T_n}.$$

Then M^n and $\mathcal{E}(M^n)$ are local martingales, $\mathcal{E}(M^n) \ge \frac{1}{n}$ and $\mathcal{E}(M^n)_t = \mathcal{E}(M)_t$ if $t < T_n$. The assumption $\Delta M > -1$ shows that T_n tends to infinity when n tends to infinity. For each $n \ge 1$, we apply Itï¿½'s formula to a $\mathcal{C}^2$ function f_n that coincides with $x^{-\lambda}$ on $[\frac{1}{n}, +\infty)$:

$$\begin{aligned}\mathcal{E}^{-\lambda}(M^n)_t =& 1 - \lambda \int_0^t \mathcal{E}^{-\lambda-1}(M^n)_{s^-} d\mathcal{E}(M^n)_s + \frac{\lambda(\lambda+1)}{2} \int_0^t \mathcal{E}^{-\lambda-2}(M^n)_{s^-} d\left\langle (\mathcal{E}(M^n))^c \right\rangle_s \\ &+ \sum_{s\le t} \left[\mathcal{E}^{-\lambda}(M^n)_s - \mathcal{E}^{-\lambda}(M^n)_{s^-} + \lambda \mathcal{E}^{-\lambda-1}(M^n)_{s^-} \Delta \mathcal{E}^{-\lambda}(M^n)_s\right] \\ =& 1 + \int_0^t \mathcal{E}(M^n)_{s^-} dX_s^n,\end{aligned}$$

where

$$X_t^n := -\lambda M_t^n + \frac{\lambda(\lambda+1)}{2} \left\langle (M^n)^c, (M^n)^c \right\rangle_t + \sum_{s\le t} \left[(1+\Delta M_s)^{-\lambda} - 1 + \lambda \Delta M_s\right],$$

and then $\mathcal{E}^{-\lambda}(M^n) = \mathcal{E}(X^n)$. Let us define the non-truncated counterpart X of X^n:

$$X = -\lambda M + \frac{\lambda(\lambda+1)}{2} \langle M^c, M^c \rangle + W^{-\lambda}.$$

On the interval $[0, T_n[$, we have $X^n = X$ and $\mathcal{E}^{-\lambda}(M) = \mathcal{E}(X)$, now letting n tends to infinity, we obtain that $\mathcal{E}^{-\lambda}(M)$ and $\mathcal{E}(X)$ coincide on $[0, +\infty[$, which is the point (i) of the Lemma.

We want to use the proposition B.2.1 to prove the point (ii), so we need to show that $\Delta A > -1$. We set

$$S = \inf\{t \ge 0 \ : \ \Delta A_t^{-\lambda} \le -1\}.$$

It is a predictable time. Using this, and the fact that M and $(W^{-\lambda} - V^{-\lambda})$ are local martingales, we have

$$\Delta A_S^{-\lambda} = \mathbb{E}\left[\Delta A_S^{-\lambda} | \mathcal{F}_{S^-}\right] = \mathbb{E}\left[\Delta X_S | \mathcal{F}_{S^-}\right] = \mathbb{E}\left[(1+\Delta M_S)^{-\lambda} | \mathcal{F}_{S^-}\right],$$

and since $\{S < +\infty\} \in \mathcal{F}_{S^-}$,

$$0 \ge \mathbb{E}\left[\mathbf{1}_{\{S<+\infty\}}(1 + \Delta A_S^{-\lambda})\right] = \mathbb{E}\left[\mathbf{1}_{\{S<+\infty\}}(1+\Delta M_S)^{-\lambda}\right].$$

Then $\Delta M_S \le -1$ on $\{S < +\infty\}$, which means that $S = +\infty$ and $\Delta A > -1$ a.s. The proof is now complete. □

We are finally in a position to state the Lemma on L^r integrability of exponential martingales for a negative exponent r.

Lemma B.2.4 *Let $\lambda > 0$ and let M be a local martingale with bounded jumps, such that $\Delta M \ge -1 + \delta$, for a fixed $\delta > 0$, and $\langle M, M \rangle_t$ is bounded $dt \times \mathbb{P}$-a.s. Then*

$$\mathbb{E}^{\mathbb{P}}\left[\mathcal{E}(M)_t^{-\lambda}\right] < +\infty \ \ dt \times \mathbb{P} - a.s.$$

Proof. Let $n \geq 1$ be an integer. We will denote $\tilde{\mu}_M = \mu_M - \nu_M$ the compensated jump measure of M. Thanks to lemma B.2.3, we write the decompostion

$$\mathcal{E}(M)^{-n} = \mathcal{E}(\tilde{N}^{-n})\mathcal{E}\left(\frac{1}{2}n(n+1) < M^c, M^c > +V^{-n}\right),$$

where $\tilde{N}^{-n}$ is a local martingale and V^{-n} is defined as $V^{-\lambda}$. Using Lemma B.2.2, we have the inequality

$$V_t^{-n} \leq \int_0^t \int_E Cx^2 \nu_M(ds, dx)$$

and using the previous representation we obtain

$$\begin{aligned}
\mathcal{E}(M)_t^{-n} &\leq \mathcal{E}(\tilde{N}^{-n})_t \mathcal{E}\left(\frac{1}{2}n(n+1)\langle M^c, M^c\rangle + \int_0^{\cdot} \int_E Cx^2 \nu_M(ds, dx)\right)_t \\
&\leq \mathcal{E}(\tilde{N}^{-n})_t \exp\left((\frac{1}{2}n(n+1) + C)\langle M, M\rangle_t\right) \\
&\leq C\mathcal{E}(\tilde{N}^{-n})_t \text{ since } \langle M, M\rangle_t \text{ is bounded.}
\end{aligned}$$

Let us prove now that the jumps of $\tilde{N}^{-n}$ are strictly bigger than -1. We compute

$$\begin{aligned}
\Delta\tilde{N}^{-n} &= \frac{\Delta N^{-n}}{1 + \Delta A^{-n}} \text{ where } A^{-n} \text{ is defined as in lemma B.2.4} \\
&= \frac{(1 + \Delta M)^{-n}}{1 + \Delta V^{-n}} - 1 > -1 \text{ since } -1 < \Delta M \leq B \text{ and } \Delta V^{-n} > -1.
\end{aligned}$$

This implies that $\mathcal{E}(\tilde{N}^{-n})$ is a positive supermartingale which equals 1 at $t = 0$. We deduce

$$\mathbb{E}\left[\mathcal{E}(M)_t^{-n}\right] \leq C\mathbb{E}\left[\mathcal{E}(\tilde{N}^{-n})_t\right] \leq C.$$

We have the desired integrability for negative integers. We extend the property to any negative real number by Hölder's inequality. □

Bibliography

Bibliography

[1] Acciaio, B. (2007). Optimal risk sharing with non-monotone monetary functionals, *Finance Stoch* 11:267–289. (Cited on page 5.)

[2] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999). Coherent measures of risk, *Mathematical Finance*, 9:203-228. (Cited on pages 2, 3, 32 and 33.)

[3] Avellaneda, M., Lévy, A., Paras, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities, *Applied Math. Finance*, 2:73–88. (Cited on page 128.)

[4] Barles, G., Buckdahn, R., Pardoux, E. (1997). Backward stochastic differential equations and integral-partial differential equations, *Stochastics and stochastic reports*, 60:57-83. (Cited on pages 17, 18, 82, 85, 143, 144, 160, 174 and 186.)

[5] Barrieu, P., El Karoui, N. (2005). Inf-convolution of risk measures and optimal risk transfer. *Finance and Stochastics*,9:269-298. (Cited on pages 3, 36, 39 and 44.)

[6] Barrieu, P., El Karoui, N. (2009). Pricing, Hedging and Optimally Designing Derivatives via Minimization of Risk Measures. Volume on Indifference Pricing (ed: Rene Carmona), Princeton University Press. (Cited on pages 4, 30, 32, 71, 101, 102, 113, 115, 117 and 118.)

[7] Barrieu, P., El Karoui, N. (2011). Monotone stability of quadratic semimartingales with applications to general quadratic BSDEs and unbounded existence result, preprint. (Cited on pages 19, 70, 81 and 104.)

[8] Barrieu, P., Cazanave, N., El Karoui, N. (2008). Closedness results for BMO semi-martingales and application to quadratic BSDEs, *C. R. Acad. SCI. Paris, Ser. I*, 346:881-886. (Cited on pages 103 and 104.)

[9] Becherer, D. (2006). Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging, *Ann. of App. Prob*, 16(4):2027-2054. (Cited on pages 70 and 150.)

[10] Biagini, S. and Frittelli, M. (2009). On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. *In Optimality and risk-modern trends in mathematical finance*, 1-28. (Cited on pages 6, 33, 34 and 40.)

[11] Bichteler, K. (1981). Stochastic integration and L^p theory of semimartingales, *Ann. of Prob.*, 9:49-89. (Cited on pages 22 and 131.)

[12] Billingsley, P. (1999). Convergence of Probability Measures, 2nd Edition, *Wiley Series in Probability and Statistics*. (Cited on pages 129 and 174.)

[13] Bion-Nadal, J. (2012), A new probabilistic approach to non local and fully non linear second order partial differential equations, preprint arXiv:1210.1955. (Cited on page 129.)

[14] Bion-Nadal, J., Kervarec, M. (2011). Risk measuring under model uncertainty. *Ann. Appl. Proba.*, to appear. (Cited on page 128.)

[15] Bismut, J.M. (1973). Conjugate convex functions in optimal stochastic control, *J. Math. Anal. Appl.*, 44:384-404. (Cited on pages 12 and 69.)

[16] Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica* 30. (Cited on page 30.)

[17] Bouchard, B. and Touzi, N. (2011). Weak Dynamic Programming Principle for Viscosity Solutions, *SIAM Journal on Control and Optimization*, 49:948-962. (Cited on page 134.)

[18] Briand, Ph., Hu, Y. (2006). BSDE with quadratic growth and unbounded terminal value, *Probab. Theory Relat. Fields*, 136:604-618. (Cited on pages 70 and 81.)

[19] Briand, Ph., Hu, Y. (2008). Quadratic BSDEs with convex generators and unbounded terminal conditions, *Probab. Theory Relat. Fields*, 141:543-567. (Cited on pages 92, 94 and 95.)

[20] Burgert, C. and Ruschendorf, L. (2006). On the optimal risk allocation problem. *Statistics & Decisions* 24:153–171. (Cited on page 5.)

[21] Carlier, G. and Dana, R.-A. (2006). Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints. *Statist. Decisions*, 24:127-152. (Cited on page 30.)

[22] Carlier, G. and Dana, R.-A. (2007). Are generalized call-spreads efficient ? *J. Math. Econom.*, 43:581-596. (Cited on page 31.)

[23] Carlier, G., Dana, R.-A. and Shahidi, N. (2003). Efficient insurance contracts under epsilon-contaminated utilities *The Geneva Papers on Risk and Insurance Theory*, 28: 59–71. (Cited on page 52.)

[24] Chateauneuf, A., Dana R.-A., and Tallon, J.-M. (2000). Optimal risk-sharing rules and equilibria with Choquet-expected-utility. *J. Math. Econom.*, 34:191-214. (Cited on page 30.)

[25] Chateauneuf, A., and Tallon, J.-M. (2002). Diversification, convex preferences and nonempty core in the Choquet expected utility model. *Econom. Theory*, 19:509-523. (Cited on page 30.)

[26] Chateauneuf, A., Maccheroni, F., Marinacci M., and Tallon, J.-M. (2005). Monotone continuous multiple priors. *Econom. Theory*, 26:973-982. (Cited on page 30.)

[27] Chen, Z. (1998). A property of backward stochastic differential equations, *C. R. Acad. Sci. Paris*, 326:483-488. (Cited on page 109.)

[28] Cheridito, P., Soner, H.M., Touzi, N., and Victoir, N. (2007). Second Order Backward Stochastic Differential Equations and Fully Non-Linear Parabolic PDEs, *Communications on Pure and Applied Mathematics*, 60:1081-1110. (Not cited.)

[29] Choquet, G. (1955). Theory of capacities. *Ann. Inst. Fourier, Grenoble*, 5:131-295. (Cited on pages 29 and 30.)

[30] Cohen, S.N. (2011) Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces , *Elec. Journal of Prob.*, 17:62. (Cited on pages 139 and 140.)

[31] Constantinides, G.M., Magill, M.J.P. (1976). Portfolio selection with transaction costs, *J. Econ. Theory*, 49:33-83. (Cited on page 25.)

[32] Cramér, H. (1930). On the mathematical theory of risk, *Skandia Jubilee Volume*, Stockholm. (Cited on page 61.)

[33] Crépey, S., Matoussi, A. (2008). Reflected and doubly reflected BSDEs with jumps, *Ann. of App. Prob.*, 18:2041-2069. (Cited on page 150.)

[34] Cvitanić, J., Karatzas, I. (1992). Convex duality in constrained portfolio optimization, *Ann. Appl. Proba.*, 2:767-818. (Cited on page 25.)

[35] De Finetti, B. (1957). Su un'impostazione alternativa della teoria collettiva del rischio. *Transactions XVth International Congress of Actuaries.* (Cited on page 31.)

[36] Dellacherie, C., and Meyer, P.-A. (1975). Probabilités et potentiel. Chapitre I à IV,*Hermann, Paris.* (Cited on page 179.)

[37] Denis, L., Hu, M., Peng, S. (2011). Function spaces and capacity related to a Sublinear Expectation: application to G-Brownian Motion Paths, *Potential Analysis*, 34:139-161. (Cited on pages 21, 55 and 128.)

[38] Denis, L., and Kervarec, M. (2007). Utility functions and optimal investment in nondominated models, preprint. (Cited on page 25.)

[39] Denis, L., Martini, C. (2006). A theoretical framework for the pricing of contingent claims in the presence of model uncertainty, *Annals of Applied Probability*, 16:827-852. (Cited on page 128.)

[40] Denneberg, D. (1994). Non-additive measure and integral, *Theory and Decision Library. Series B: Mathematical and Statistical Methods*, 27. Kluwer Academic Publishers Group, Dordrecht. (Cited on pages 5, 30, 40 and 48.)

[41] Denuit, M. and Charpentier, A. (2004). Mathématiques de l'assurance non-vie. Tome 1 : principes fondamentaux de théorie du risque. *Economica.* (Cited on page 61.)

[42] Doléans-Dade, C., Meyer, P.A. (1977). Une caractérisation de BMO, *Séminaire de Probabilités XI, Univ. de Strasbourg*, Lecture Notes in Math., 581:383-389. (Cited on pages 75 and 76.)

[43] Dow, J. and Werlang, S.R.C. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica*, 60:197-204. (Cited on pages 30 and 52.)

[44] Eichberger, J. and Kelsey, D. (1999). E-capacities and the Ellsberg paradox.*Theory and Decision*, 46:107-140. (Cited on page 30.)

[45] El Karoui, N. (1981). Les aspects probalilistes du contrôle stochastique, *Ecole d'Eté de Probabilités de Saint-Flour IX-1979, Lecture Notes in Mathematics*, 876:73-238. (Cited on pages 23, 134 and 179.)

[46] El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S. and Quenez, M. C. (1997). Reflected solutions of backward SDE's, and related obstacle problems for PDE's,*Annals of Applied Probability*, 25(2):702-737. (Cited on pages 24 and 128.)

[47] El Karoui, N., Matoussi, A., Ngoupeyou, A. (2012). Quadratic BSDE with jumps and unbounded terminal condition, preprint (Cited on pages 70 and 77.)

[48] El Karoui, N., Nguyen, D.H., Jeanblanc-Picqué, M. (1987). Compactification methods in the control of degenerate diffusions: existence of an optimal control, *Stochastics*, 20:169-219. (Cited on page 134.)

[49] El Karoui, N., Peng, S., Quenez, M.C. (1994). Backward stochastic differential equations in finance, *Mathematical Finance*, 7(1): 1-71. (Cited on page 179.)

[50] El Karoui, N. and Ravanelli, C. (2009). Cash subadditive risk measures and interest rate ambiguity. *Math. Finance*, 19:561-590. (Cited on pages 49 and 50.)

[51] El Karoui, N., Rouge, R. (2000). Pricing via utility maximization and entropy, *Mathematical Finance*, 10:259-276. (Cited on pages 25, 175, 177 and 180.)

[52] Essaky, E.H. (2006). Reflected backward stochastic differential equations with jumps and RCLL obstacle, preprint. (Cited on page 169.)

[53] Filipović D. and Svindland G. (2008), Optimal capital and risk allocations for law- and cash-invariant convex functions, *Finance Stoch* 12:423–439 (Cited on page 5.)

[54] Föllmer, H. and Schied, A. (2004). Stochastic Finance: An introduction in discrete time, second edition, *de Gruyter studies in mathematics 27.* (Cited on pages 3, 6, 32, 33, 36, 37, 38, 40, 41, 43, 44 and 48.)

[55] Frittelli, M. and Rosazza Gianin, E. (2002). Putting Order in Risk Measures. *Journal of Banking Finance*, 26:1473-1486. (Cited on pages 3, 32 and 33.)

[56] Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with nonunique prior. *J. Math. Econom.*, 18:141-153. (Cited on page 30.)

[57] Greco, G. (1977). Integrale monotono. *Rend. Sem. Mat. Padova*, 57:149-166. (Cited on page 30.)

[58] Hamadène, S. and Ouknine, Y. (2011). Reflected backward SDEs with general jumps, preprint. (Cited on pages 169 and 185.)

[59] Hu, Y., Imkeller, P., and Müller, M. (2005). Utility maximization in incomplete markets, *Ann. Appl. Proba.*, 15(3):1691–1712. (Cited on pages 159, 175 and 177.)

[60] Hipp, C. and Taksar, M. (2010). Optimal non-proportional reinsurance control. *Insurance Math. Econom.*, 47:246-254. (Cited on page 31.)

[61] Hipp, C. and Vogt, M. (2003). Optimal dynamic XL reinsurance.*ASTIN Bulletin*, 33:193-207. 2003. (Cited on page 31.)

[62] Hu, M. and Peng, S. (2009). G-Lèvy Processes under Sublinear Expectations, preprint. (Cited on pages 148 and 149.)

[63] W. Hürlimann, Excess of loss reinsurance with reinstatements revisited, *Astin Bulletin* 35 (2005) 211-238. (Cited on page 11.)

[64] Izumisawa, M., Sekiguchi, T., Shiota, Y. (1979). Remark on a characterization of BMO-martingales, *Tôhoku Math. Journ.*, 31:281–284. (Cited on page 76.)

[65] Jacod, J., Shiryaev, A.N. (1987). Limit theorems for stochastic processes, *Springer-Verlag*. (Cited on pages 129, 131, 132, 137 and 141.)

[66] Jacod, J. (1979). Calcul stochastique et problèmes de martingales, *Springer*. (Cited on page 109.)

[67] Jouini, E., Schachermayer, W. and Touzi, N. (2006). Law invariant risk measures have the Fatou property.*Advances in mathematical economics*, 9:49–71. (Cited on page 43.)

[68] Jouini, E., Schachermayer, W. and Touzi, N. (2008). Optimal risk sharing for law invariant monetary utility functions.*Math. Finance*, 18:269-292. (Cited on pages 3, 4, 5, 30, 35, 39, 56 and 58.)

[69] Karandikar, R. (1995). On pathwise stochastic integration. *Stochastic Processes and their Applications*, 57:11-18. (Cited on pages 22 and 131.)

[70] Kazamaki, N. (1979). A sufficient condition for the uniform integrability of exponential martingales, *Math. Rep. Toyama University*, 2:1-11. (Cited on pages 75, 94 and 183.)

[71] Kazamaki, N. (1979). On transforming the class of BMO-martingales by a change of law, *Tôhoku Math. J.*, 31:117-125. (Cited on page 76.)

[72] Kazamaki, N. (1994). Continuous exponential martingales and BMO. *Springer-Verlag*. (Cited on pages 74, 75, 76 and 180.)

[73] Kervarec, M. (2008). Etude des modèles non dominés en mathématiques financières. PhD Dissertation. (Cited on page 40.)

[74] Kobylanski, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth, *Ann. Prob.* 28:259-276. (Cited on pages 70, 103 and 116.)

[75] Kramkov, D. and Schachermayer, W. (1999). The asymptotic elasticity of utility functions and optimal investement in incomplete markets. *The Annals of Applied Probability*, 9(3):904-950. (Cited on page 25.)

[76] Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in mathematical economics*, 3:83-95. (Cited on pages 7, 38 and 42.)

[77] Laeven, R.J.A., Stadje, M.A. (2012). Robust portfolio choice and indifference valuation, preprint. (Cited on page 70.)

[78] Lepeltier, J. P. and San Martin, J. (1997). Backward stochastic differential equations with continuous coefficient, *Statistics & Probability Letters*, 32(5):42-430. (Cited on pages 104 and 105.)

[79] Lépingle, D. and Mémin, J. (1978). Sur l'intégrabilité uniforme des martingales exponentielles. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 42:175-203. (Cited on pages 76 and 194.)

[80] Lépingle, D. and Mémin, J. (1978). Intégrabilité uniforme et dans L^r des martingales exponentielles. *Sém. Prob. de Rennes.* (Cited on pages 76, 152 and 194.)

[81] Li, X., Tang, S. (1994). Necessary conditions for optimal control of stochastic systems with random jumps, *SIAM J. Control Optim.*, 32(5):1447-1475. (Cited on pages 70, 82, 85 and 186.)

[82] Lim, T., Quenez, M.-C. (2010). Exponential utility maximization in an incomplete market with defaults, *Electronic Journal of Probability*, 16:53. (Cited on pages 26, 176 and 180.)

[83] Lin, Q.Q. (2003). Nonlinear Doob-Meyer decomposition with jumps, *Acta Mathematica Sinica*, 19(1):69-78. (Cited on pages 18, 71, 100 and 106.)

[84] Ma, J., Yao, S. (2010). On quadratic g-evaluations / expectations and related analysis, *Stochas. Ana. and App.*, 28(4):711-734. (Cited on pages 18, 71, 100, 103, 109 and 113.)

[85] Maccheroni F., Marinacci M., Rustichini A. and Taboga M., Portfolio Selection With Monotone Mean-Variance Preferences,*Mathematical Finance*, 19(3):487-521. (Cited on page 5.)

[86] Mata, A.J. (2000). Pricing excess of loss reinsurance with reinstatements, *Astin Bulletin*, 30:349-368. (Cited on pages 11 and 62.)

[87] Matoussi, A., Possamai, D., Zhou, C. (2012). Robust utility maximization in non-dominated models with 2BSDEs, *Mathematical Finance*, to appear. (Cited on pages 26, 159, 175 and 177.)

[88] Merton, R. (1969). Lifetime portfolio selection under uncertainty: the continuous time case,*Rev. Econ. Stat.*, 51:239-265. (Cited on page 25.)

[89] Mémin, J. (1978). Décompositions multiplicatives de semimartingales exponentielles et applications, *Séminaire de probabilités de Strasbourg*, 12:35-46. (Cited on page 194.)

[90] Meyer, P.-A. (1974). Une représentation de surmartingales, *Séminaire de probabilités (Strasbourg)*, 8:310-315. (Cited on page 108.)

[91] Morlais, M.A. (2007). Equations différentielles stochastiques rétrogrades à croissance quadratique et applications , *Université de Rennes*, PhD Thesis. (Cited on pages 70, 79 and 114.)

[92] Morlais, M.-A. (2009). Utility maximization in a jump market model, *Stochastics and Stochastics Reports*, 81:1-27. (Cited on pages 180, 181 and 182.)

[93] Neveu, J. (1975). Discrete parameter martingale, *North Holland Publishing Company*. (Cited on page 153.)

[94] Ngoupeyou, A.B. (2010). Optimisation des portefeuilles d'actifs soumis au risque de défaut, *Université d'Evry*, PhD Thesis. (Cited on pages 70 and 104.)

[95] Nutz, M. (2012). Pathwise construction of stochastic integrals, *Electronic Communications in Probability*, 17:1-7. (Cited on pages 22, 131 and 170.)

[96] Panjer, H. (1981). Recursive evaluation of a family of compound distributions.*ASTIN Bulletin*, 12: 22–26. (Cited on page 62.)

[97] Pardoux, E. and Peng, S (1990). Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, 14:55-61. (Cited on pages 12, 13, 69, 70, 128 and 144.)

[98] Partrat, C., Besson, J.-L. (2004). Assurance non-vie. Modélisation, simulation. *Economica*. (Cited on page 61.)

[99] Peng, S. (1997). Backward SDE and related g-expectation, *in Backward stochastic differential equations*, Pitman Research notes Math. Series, 364:141-159. (Cited on pages 18, 71, 100 and 101.)

[100] Peng, S. (1999). Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type, *Prob. Theo. and Rel. Fields*, 113(4):23-30. (Cited on pages 19, 21 and 103.)

[101] Peng, S. (2007). G-Expectation, G-Brownian motion, and related stochastic calculus of Itô type, arXiv:math/0601035v2. (Cited on pages 55 and 128.)

[102] Peng, S. (2010) Nonlinear Expectations and Stochastic Calculus under Uncertainty, arXiv:1002.4546v1 (Cited on pages 128 and 148.)

[103] Pliska, S.R. (1986). A stochastic calculus model of continuous trading: optimal portfolios, *Math. Operations Research*, 11:371-382. (Cited on page 25.)

[104] Possamai, D., Zhou, C. (2010). Second order backward stochastic differential equations with quadratic growth, preprint. (Cited on pages 172, 182, 183 and 184.)

[105] Protter, P.E. (2004). Stochastic integration and differential equations, *Springer*. (Cited on page 75.)

[106] Royer, M. (2006). Backward stochastic differential equations with jumps and related nonlinear expectations, *Stochastic Processes and their Applications*, 116:1358-1376. (Cited on pages 17, 18, 71, 92, 97, 100, 103, 106, 144, 150, 151, 153 and 169.)

[107] Sato, K.-I. (1999). Lévy processes and infinitely divisible distributions, *Cambridge University Press*. (Cited on pages 82 and 137.)

[108] Schmeidler, D. (1986). Integral representation without additivity. *Proc. Amer. Math. Soc.*, 97:255-261. (Cited on page 30.)

[109] Schmidli, H. (2002). On minimizing the ruin probability by investment and reinsurance. *Ann. Appl. Probab.*, 12:890-907. (Cited on page 31.)

[110] Soner, H.M., Touzi, N., Zhang J. (2010). Dual formulation of second order target problems, *Ann. of App. Prob*, to appear. (Cited on pages 162, 166, 167, 169, 170, 173 and 175.)

[111] Soner, H.M., Touzi, N., Zhang J. (2011). Quasi-sure stochastic analysis through aggregation, *Elec. Journal of Prob.*, 16:1844-1879. (Cited on pages iii, v, 22, 130, 131, 133, 134, 135, 136, 139, 140 and 141.)

[112] Soner, H.M., Touzi, N., Zhang J. (2012). Wellposedness of second order BSDE's, *Prob. Th. and Related Fields*, 153:149-190. (Cited on pages iii, v, 21, 24, 55, 128, 129, 130, 143, 144, 145, 147, 150, 151, 153, 160, 161, 162, 172 and 192.)

[113] Stroock, D.W., Varadhan, S.R.S. (1979). Multidimensional diffusion processes, *Springer-Verlag, Berlin, Heidelberg, New-York.* (Cited on pages 24 and 160.)

[114] Sundt, B. (1991). On excess of loss reinsurance with reinstatements. *Bulletin of the Swiss Association of Actuaries*, no. 1. (Cited on pages 11 and 62.)

[115] Taksar, M. (2000). Optimal risk and dividend distribution control models for an insurance company. *Math. Methods Oper. Res.*, 51:1-42. (Cited on page 31.)

[116] Tang S., Li X.(1994). Necessary condition for optimal control of stochastic systems with random jumps, *SIAM JCO*, 332:1447-1475. (Cited on pages 13, 14, 150 and 166.)

[117] Tevzadze, R. (2008). Solvability of backward stochastic differential equations with quadratic growth, *Stoch. Proc. and their App.*, 118:503-515. (Cited on pages v, 15, 17, 70, 71, 88, 89 and 92.)

[118] Vogt, M. (2004). Optimal dynamic reinsurance. *Blätter der DGVFM*, 26:407-420. (Cited on page 31.)

[119] Walhin, J.F. and Paris, J. (2001). Excess of loss reinsurance with reinstatements: Premium calculation and ruin probability for the cedent, *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* 25:257-270. (Cited on pages 11 and 62.)

[120] Wang, S. (2002). A universal framework for pricing financial and insurance risks, *ASTIN Bulletin*, 32:213-234. (Cited on page 37.)

[121] Zariphopoulou, T. (1994). Consumption-investment models with constraints, SIAM J. Control and Optimization, 32(1):59-85. (Cited on page 25.)