Université de Paris

École Doctorale de Sciences Mathématiques de Paris-Centre (386)

Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR 8001)

Study of numerical methods for partial hedging and switching problems with costs uncertainty

Par Cyril Bénézet

Thèse de doctorat de **Mathématiques Appliquées**

Dirigée par Jean-François Chassagneux

Présentée et soutenue publiquement le 5 Novembre 2019

Devant un jury composé de :

Jean-François Chassagneux, Professeur, Univ. Paris Diderot, Directeur de thèse
Romuald Elie, Professeur, Univ. Paris-Est, Rapporteur
Emmanuel Gobet, Professeur, Ecole Polytechnique, Examinateur
Idris Kharroubi, Professeur, Sorbonne Université, Examinateur
Huyên Pham, Professeur, Univ. Paris Diderot, Examinateur
Adrien Richou, Maître de Conférences, Univ. de Bordeaux 1, Examinateur
Agnès Sulem, Directeur de Recherche, INRIA, Examinatrice
Jianfeng Zhang, Professeur, University of Southern California, Rapporteur

À ma tante Pascale...

Je souhaite tout d'abord remercier chaleureusement mon directeur de thèse, Jean-François, sans qui ce travail n'aurait pas vu le jour. Ces trois années de thèse passées dans la bonne humeur sont et resteront une étape importante dans ma vie, étape durant laquelle tu m'auras appris la recherche en mathématiques avec sérieux et rigueur, sans oublier les blagues auxquelles nul ne peut échapper, et qui me manqueront !

Je tiens également à sincèrement remercier Jérémie pour m'avoir encadré lors d'un stage au GIE AXA avant le début de ma thèse, puis pour avoir proposé tout au long de ma thèse des problématiques de recherche ayant des applications concrètes pour la gestion des risques dans le monde de l'Assurance. Les deux premières parties de ma thèse apportent quelques réponses à ces questions.

Cette thèse n'aurait pas pu être écrite sans le financement de l'initiative de recherche "Méthodes non-linéaires pour la gestion des risques financiers", sous l'égide de l'Institut Europlace de Finance et avec le support du fond pour la recherche d'AXA. Je remercie donc les membres de cette initiative de recherche : Bruno Bouchard, Xiaolu Tan, Jean-François Chassagneux, Jérémie Bonnefoy et Aymeric Kalife. J'ai une pensée particulière pour Bruno Bouchard qui m'a dirigé vers Jean-François et Jérémie lorsque, durant mon master, j'étais à la recherche d'une opportunité pour effectuer mon doctorat. Je remercie également Idris Kharroubi pour m'avoir encadré pour mes premiers pas dans la recherche en Contrôle Stochastique lors de mon master.

Après ces trois années de travail, le manuscrit étant terminé, je tiens à remercier Romuald Elie et Jianfeng Zhang qui me font l'honneur de rapporter ma thèse. Je remercie également Emmanuel Gobet, Idris Kharroubi, Huyên Pham, Adrien Richou et Agnès Sulem pour avoir accepté de faire partie de mon jury de thèse.

Je remercie tout le Laboratoire de Probabilités, Statistique et Modélisation, en particulier les doctorants et l'équipe de Mathématiques Financières et Actuarielles, Probabilités Numériques pour m'avoir accueilli durant ces trois années, toujours dans la bonne humeur. Merci à Nathalie et Valérie qui nous accompagnent dans les méandres administratifs que nous rencontrons au quotidien.

Je remercie également toute l'équipe du GRM Life du GIE AXA, les membres actuels comme les anciens, qui m'a accueilli, toujours les bras ouverts, durant mon stage et mon doctorat.

Je souhaite remercier les organisateurs du CEMRACS 2017, école d'été durant laquelle la deuxième partie de cette thèse a été rédigée, en collaboration avec Jérémie Bonnefoy, Jean-François Chassagneux, Camilo Garcia-Trillos, Shuoqing Deng et Lionel Lenôtre, que je remercie tous, en particulier Camilo avec qui le travail fut particulièrement agréable sous le soleil de Luminy. Un grand merci au CIRM pour l'accueil durant cet été 2017.

Je remercie également Christoph Reisinger, avec qui les discussions ont toujours été fructueuses, sans qui les méthodes numériques développées lors de la première partie de mon travail n'auraient pas été si poussées. Merci au Mathematical Institute de l'Université d'Oxford pour m'avoir accueilli durant une semaine et en particulier à Yufei pour sa présence, sa sympathie et nos discussions.

Je remercie Adrien Richou pour m'avoir introduit aux problèmes de switching "randomisés" et leurs liens avec les équations rétrogrades réfléchies. Cette collaboration a permis au dernier chapitre de ma thèse de voir le jour.

Enfin, je remercie ma famille et mes amis pour leur présence et leur soutien, et m'excuse auprès d'eux si les derniers mois de thèse m'ont rendu moins présent que d'habitude !

Un dernier mot pour Arij sur qui je peux toujours compter, à n'importe quel moment, et qui est toujours là pour me soutenir, m'encourager et (surtout !) me supporter. Merci.

Résumé

Nous apportons dans cette thèse quelques contributions à l'étude théorique et numérique de certains problèmes de contrôle stochastique, ainsi que leurs applications aux mathématiques financières et à la gestion des risques financiers. Ces applications portent sur des problématiques de valorisation et de couverture faibles de produits financiers, ainsi que sur des problématiques réglementaires. Nous proposons des méthodes numériques afin de calculer efficacement ces quantités pour lesquelles il n'existe pas de formule explicite.

Dans la première partie, nous étudions le problème de valorisation partielle d'options européennes, et en particulier le problème de sur-réplication avec probabilité p.
Dans le chapitre 3, lorsque le marché est linéaire, nous démontrons qu'il est possible de résoudre explicitement le problème de contrôle optimal, et que la valeur de celui-ci s'écrit essentiellement en fonction de la loi du prix du sous-jacent à la date terminale. Cette formulation permet une nouvelle démonstration de formules explicites dans le modèle de Black & Scholes pour des options vanille. De plus, nous calculons une vitesse de convergence lors de l'approximation par un problème avec contrôles bornés.
Dans le chapitre 4, nous considérons un marché non-linéaire, et nous approchons le problème par celui où l'on n'autorise que des contrôles constants par morceaux. Nous démontrons la convergence lorsque le pas de discrétisation temporelle tend vers 0. Lorsque la diffusion du log-prix du sous-jacent est brownienne, nous couplons cette discrétisation temporelle à une discrétisation spatiale par différences finies et démontrons la convergence vers la fonction valeur. Des applications numériques sont données, montrant que le schéma converge et montrant la nécessité pratique des conditions de compatibilité théoriques sur les paramètres de discrétisation.

La seconde partie de cette thèse est consacrée à l'approximation de mesures de risques sur la distribution du bilan futur d'une banque ou d'une société d'assurance. Ce problème est issu de la réglementation des compagnies d'assurance Solvabilité II. Dans un modèle Gaussien à taux stochastiques, l'entreprise vend un produit financier et se couvre discrètement dans le temps. Nous développons une méthode numérique permettant le calcul efficace, en dimension modérée, de la distribution recherchée. Nous démontrons que les mesures de risque spectrales, calculées en la distribution approchée, convergent vers les mesures de risque sur la distribution inconnue. Nous donnons des applications numériques montrant l'efficacité de la méthode, en comparaison avec la méthodes des

simulations imbriquées.

La troisième partie est consacrée à l'étude de problèmes de *switching* avec incertitude sur les coûts. Nous introduisons une nouvelle famille de problèmes de contrôle stochastique qui généralisent les problèmes de *switching* usuels. Nous montrons, sous des hypothèses classiques pour le générateur de l'équation rétrograde, et sous l'hypothèse de positivité des coûts de changement d'état, que la valeur de l'un de ces problèmes de contrôle est donné par la composante Y de la solution d'une équation rétrograde avec réflexions obliques. Nous en déduisons l'unicité des solutions à ces équations, sous réserve d'existence. Nous montrons ensuite que les équations susdites admettent effectivement des solutions, dans le cas que nous appelons "non-contrôlé" et "irréductible", sans hypothèse de positivité des coûts. Une description géométrique du domaine de réflexion est donnée, en fonctions des coûts.

Mots clés : Quantile Hedging, EDSRs, EDPs non-linéaires, solutions de viscosité, schémas de différences finies monotones, grilles sparses, mesures de risque, contrôle optimal stochastique, switching optimal, grossissement de filtration, EDSRs obliquement réfléchies.

Abstract

In this thesis, we give some contributions to the theoretical and numerical study to some stochastic optimal control problems, and their applications to financial mathematics and risk management. These applications are related to weak pricing and hedging of financial products and to regulation issues. We develop numerical methods in order to compute efficiently these quantities, when no closed formulae are available.

In the first part, we study the partial hedging of European claims, and in particular the quantile hedging problem.
In Chapter 3, when the market is linear, we show that it is possible to solve explicitly the optimal control problem, and that its value is a function of the law of the asset price at maturity. This formulation gives a new and simple proof for explicit formulae for the quantile hedging price of vanilla options in the Black & Scholes model. In addition, we compute a convergence rate for the approximation of the control problem by one with bounded controls.
In Chapter 4, we consider a non-linear market, and we approximate the control problem by one where a control is admissible only if it is piecewise constant over time. We show the convergence when the time discretization parameter goes to 0. When the log-price follows a Brownian diffusion, we couple this time discretization with a spatial discretization using finite differences, and we show the convergence towards the value function. Some numerical applications are given, showing that the scheme converges, and that the theoretical conditions on the discretization parameters are also necessary in practice.

The second part of this thesis is dedicated to the approximation of risk measures on the distribution of the future balance sheet of a bank or an insurance company. This problem comes from Solvency II, the insurance companies regulation. In a Gaussian model with stochastic rates, the company sells a financial product and hedges itself discretely in time. We develop a numerical method allowing efficient approximation, in moderate dimension, of the unknown distribution. We show that spectral risk measures, computed on the approximated distribution, converge to the risk measures computed on the unknown distribution. We give numerical applications showing the efficiency of the method, in comparison with the Nested Simulations approach.

The third part is dedicated to the study of some switching problems with costs uncertainty. We introduce a new family of stochastic optimal control problems, generalising usual switching problems. We show, under classical hypotheses on the driver of the backward equation and under positive costs, that the value of these control problems is given by the Y component of the solution of some backward equations with oblique reflections. We deduce the uniqueness of the solutions to these equations, assuming that there exists a solution. We then show that the precedent equations actually admits solutions, in the case we call "uncontrolled" and "irreducible", without assuming that the costs are positive. We give a geometric characterisation of the domain, with respect to the costs.

Key words: Quantile Hedging, BSDEs, Non-Linear PDEs, Viscosity solutions, Monotone Finite Difference schemes, Sparse Grids, Risk Measures, Stochastic Optimal Control, Optimal Switching, Enlargement of Filtration, Obliquely Reflected BSDEs.

Contents

Chapitre 1

Introduction - Français

Ce manuscrit étudie trois problèmes issus des Mathématiques Financières et du contrôle optimal de processus stochastiques. Nous introduisons et motivons les questions abordées, et nous résumons les résultats principaux obtenus.

1.1 Couverture partielle et Fonds propres réglementaires

Les deux premières parties de cette thèse sont dédiées à l'étude mathématique de la valorisation et la couverture des produits dérivés issus de la finance de marché, et à la gestion des risques issus de la couverture partielle.
Nous nous intéressons à une forme faible de valorisation, appelée valorisation sous contrainte de risque, et en particulier au prix de sur-réplication avec probabilité p. Nous étudions d'abord le cas d'un marché linéaire, où nous calculons des contrôles optimaux, et utilisons leur forme explicite afin d'obtenir une borne supérieure sur l'erreur commise lorsque l'on approxime le problème de contrôle avec des contrôles bornés. Dans un marché avec certaines imperfections, nous proposons un schéma numérique afin d'approcher le prix de sur-réplication avec probabilité p d'une option européenne. Nous prouvons sa convergence et montrons numériquement son efficacité.
Dans la deuxième partie, après modélisation d'un marché financier à taux stochastiques, nous proposons une méthode d'estimation de la distribution des pertes et profits (PnL) d'une institution financière, dans le but de calculer les fonds propres réglementaires dans le cadre de Solvabilité II. Nous obtenons, pour des mesures spectrales, une borne supérieure sur l'erreur commise par notre méthode d'approximation, et nous montrons sur des exemples numériques que la méthode proposée est compétitive, en dimension modérée, avec la méthode des simulations imbriquées.

1.1.1 Valorisation et couverture dans les marchés complets

Nous considérons un marché financier en temps continu, modélisé par un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, dans lequel sont définis un actif sans risque avec rendement constant $r \geq 0$ et un actif risqué d-dimensionnel de prix S, qui est une semi-martingale.

Supposons qu'un agent désire vendre une option européenne de maturité $T > 0$ et de *pay-off* $g(S_T)$, où g est une fonction mesurable de $\mathbb{R}$ dans $\mathbb{R}_+$, satisfaisant à des conditions d'intégrabilité adéquates. Une question importante en mathématiques financières est de déterminer un prix de vente. On suppose que le marché n'admet pas d'opportunité d'arbitrage et qu'il est possible de construire un portefeuille de réplication, c'est-à-dire que partant d'une richesse initiale p, il est possible de trouver une stratégie d'investissement telle que la valeur de liquidation du portefeuille à la date T soit exactement $g(S_T)$. Dans ce cas, le prix doit être p.

Mathématiquement, ces arguments ont été formalisés par les travaux fondateurs de Black et Scholes [BS73], Merton [Mer73 ; MS90], et les études de Harrison et Kreps [HK79], Harrison et Pliska [HP81], Duffie [Duf88], Karatzas [Kar89] et Delbaen et Schachermayer [DS94] entre autres : étant donnée une stratégie d'investissement (y, ν) où y est la richesse initiale et ν_t^i représente le nombre de parts de l'actif i détenu à la date t, la valeur du portefeuille évolue selon la dynamique suivante :

$$V_t = y + \int_0^t \nu_s \mathrm{d}S_s + \int_0^t r(V_s - \nu_s S_s)\mathrm{d}s, \quad t \in [0, T].$$

Dans ce cas, il est possible de trouver une mesure de probabilité *risque-neutre* $\mathbb{Q}$ équivalente à $\mathbb{P}$, sous laquelle $(e^{-rt}S_t)_{t\geq 0}$ est une martingale, et alors on obtient

$$p = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}g(S_T)\right].$$

Bien sûr, lorsque l'on veut modéliser les marchés financiers de façon plus précise, on réalise qu'à cause d'imperfections de marché (telles que les coûts de transaction, la présence de différents taux d'intérêt, des contraintes de *trading*), la valorisation et la couverture ne sont pas si directes : le marché peut devenir “non-linéaire” et incomplet. Dans ce cas, la dynamique des processus de richesse associés aux stratégies auto-financées devient potentiellement non-linéaire.
On peut écrire, par exemple,

$$V_t = y - \int_0^t f(s, S_s, V_s, \sigma(s, S_s)\nu_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\nu_s \mathrm{d}W_s, \quad t \in [0, T],$$

où $f : [0, T] \times (0, \infty)^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ est une fonction mesurable soumise à des hypothèses, et le processus de prix S est décrit par la dynamique

$$S_t = S_0 + \int_0^t \mu(s, S_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\mathrm{d}W_s, \quad t \in [0, T],$$

où W est un mouvement Brownien et μ, σ sont des coefficients à valeurs dans des espaces de dimension appropriée et satisfaisant à des hypothèses de sorte à ce que S et V soient uniquement définis.

Remarque 1.1.1. *D'autres dynamiques pour S et V peuvent et ont été considérées, avec par exemple des sauts [TL94 ; Par97].*

Exemple 1.1.1. *Si $\mu(t, x) \equiv \mu \in \mathbb{R}, \sigma(t, x) \equiv \sigma > 0$ et il existe un taux d'emprunt R et un taux de prêt $r \leq R$, la fonction f est donnée, suivant [EKPQ97], par*

$$f(t, x, y, z) = -ry - \sigma^{-1}\mu z + (R - r)(y - \sigma^{-1}z)^-.$$

Dans ce cadre non-linéaire, si le marché est complet, alors l'argument financier précédent s'applique. Le prix de tout produit dérivé est maintenant obtenu par la résolution d'une Équation Différentielle Stochastique Rétrograde (EDSR) :

$$Y_t = g(S_T) + \int_t^T f(t, S_t, Y_t, Z_t)\mathrm{d}t - \int_t^T Z_t \mathrm{d}W_t, \quad t \in [0, T],$$

où f, le générateur de l'EDSR, satisfait à des hypothèses appropriées.
Plus précisément, le prix à $t = 0$ est donné par Y_0 et la stratégie de couverture est donnée par le processus Z.

Dans le cas où $f(y, z) = -ry - \sigma^{-1}\mu z$ est linéaire, nous retrouvons le cas précédent, et on peut montrer, suivant [EKPQ97], que

$$Y_0 = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT} g(S_T)\right],$$

où $\mathbb{Q}$ est la mesure risque-neutre.

Pour chaque spécification des dynamiques, l'étude de ces EDSRs donne des informations sur le marché financier et la possibilité de valoriser et de couvrir des options Européennes, et leur résolution (exacte ou numérique) permet de calculer les prix et stratégies de couverture.

1.1.2 Sur-réplication de produits dérivés

Lorsque le marché est complet, les EDSRs introduites dans la sous-section précédente admettent une unique solution pour tout produit dérivé $g(S_T)$. Le prix du produit est alors donné par la valeur en $t = 0$ de Y et la stratégie de couverture est donnée par Z.
Cependant, il est possible de trouver des modèles de marché incomplet. Ceci peut arriver lorsque, par exemple, l'on autorise les processus de prix à admettre des sauts, ou alors si l'on restreint les stratégies admissibles à vivre dans des domaines fermés $D \subset \mathbb{R}^d$.

Même si le vendeur ne peut pas trouver de stratégie de réplication, il peut trouver une stratégie (y, ν) vérifiant $V_T \geq g(S_T)$ presque-sûrement. Le produit peut alors être vendu au *prix de sur-réplication*, suivant [EKQ95], défini par

$$p^s = \inf\left\{y \geq 0 : \exists \nu, V_T \geq g(S_T)\right\},$$

qui est le plus grand prix n'introduisant pas d'arbitrage sur le marché. Si l'infimum est atteint, ce prix induit une stratégie de couverture.

Cependant, celle-ci peut être difficile à implémenter en pratique, à cause de positions très longues ou très courtes, et de positions qui changent très rapidement au cours du temps.
De ce fait, une généralisation de ce problème a été considérée, où une stratégie est admissible seulement si elle satisfait à certaines contraintes de portefeuille. Entre autres, Cvitanic et Karatzas [CK93] et Föllmer et Kramkov [FK97] ont étudié ce problème respectivement dans le cas d'une diffusion et dans le cas semi-martingale.
Ce problème généralisé a alors été étudié systématiquement comme un nouveau problème de contrôle optimal stochastique, appelé "problème de cible stochastique", par Soner et Touzi [ST02a ; ST02b] pour des contrôles à valeurs dans un compact, et par Bouchard, Elie et Touzi [BET09] pour des contrôles à valeurs dans un fermé. En considérant la version dynamique du problème de contrôle, c'est-à-dire démarrant à la date

t avec un prix initial $S_t = x$, un principe de la programmation dynamique est prouvé [ST02a ; ST02b ; BET09] pour la fonction valeur $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$. Ce principe prend la forme suivante :

- **(GDP1)** Pour tout $y > v(t,x)$, il existe une stratégie ν telle que, pour tout temps d'arrêt $t \leq \tau \leq T$, on a
$$V_\tau^{y,\nu} \geq v(\tau, S_\tau).$$
- **(GDP2)** Pour tout $y < v(t,x)$, toute stratégie ν et tout temps d'arrêt $t \leq \tau \leq T$, on a
$$\mathbb{P}\left(V_\tau^{y,\nu} > v(\tau, S_\tau)\right) < 1.$$

De plus, la fonction valeur est une solution de viscosité d'une Équation aux Dérivées Partielle (EDP) de type Hamilton-Jacobi-Bellman, comme c'est le cas pour les problèmes de contrôle optimal stochastique classiques.

Cependant, en pratique, les prix obtenus en résolvant ce problème peuvent être trop élevés pour être intéressants. Ceci a mené les praticiens et les chercheurs à réfléchir à de nouvelles méthodes de valorisation. Lorsque le prix de vente p vérifie $p < p^s$, au vu du problème de contrôle précédent, il n'existe pas de stratégie (p, ν) telle que $V_T \geq g(S_T)$. Le vendeur du produit doit alors conserver du risque dans son portefeuille : quelle que soit la stratégie de couverture qu'il applique, il existe $\hat{\Omega} \subset \Omega$ tel que $\mathbb{P}(\hat{\Omega}) > 0$ et $V_T < g(S_T)$ sur $\hat{\Omega}$.

1.1.3 Valorisation et couverture avec perte contrôlée

Nous venons de voir que, pour obtenir un prix de vente raisonnable pour certains produits, il est nécessaire de choisir un prix inférieur au prix de sur-réplication. N'importe quelle stratégie de couverture échoue alors avec une probabilité strictement positive. On peut alors, par exemple, considérer un prix de sorte qu'il existe une stratégie minimisant une mesure de risque sur la distribution des pertes. D'autres notions de prix, comme la valorisation par indifférence d'utilité [Car08], ont beaucoup été étudiées.
Une autre possibilité est, étant donnée une probabilité de succès $p \in [0,1]$, de résoudre le problème suivant, introduit par Föllmer et Leukert dans [FL99],
$$p^{qh}(p) = \inf\left\{y \geq 0 : \exists \nu, \mathbb{P}\left(Y_T \geq g(S_T)\right) \geq p\right\}.$$
Ce problème est nommé sur-réplication avec probabilité p, et il est l'objet d'étude central de la première partie de cette thèse.
Nous remarquons que $p^{qh}(1) = p^s$, donc ce problème est une généralisation du problème de sur-réplication décrit précédemment.
Dans [FL99], une application élégante du lemme de Neyman-Pearson issu des mathématiques statistiques permet aux auteurs de résoudre le problème dans un marché linéaire, lorsque les prix sont dirigés par des semi-martingales. De plus, ils obtiennent des formules fermées pour le prix de sur-réplication avec probabilité p pour les options vanilles dans le modèle de Black et Scholes. Cependant, leur approche ne permet pas d'étudier le cas non-linéaire.

Bouchard, Elie et Touzi dans [BET09] ont généralisé le problème de [FL99] en considérant le problème de contrôle stochastique suivant, appelé “cible stochastique avec perte contrôlée” :

$$p^{cl}(p) = \inf \{y \geq 0 : \exists \nu, \mathbb{E}[\ell(V_T - g(S_T))] \geq p\},$$

où ℓ est une fonction croissante. En prenant $\ell = 1_{\mathbb{R}_+}$, on voit que le problème de [FL99] est un cas particulier du problème de cible avec perte contrôlée.
Dans leur travail [BET09], les auteurs imposent des contraintes de portefeuille, et le prix S est contrôlé par la stratégie de couverture, pour prendre en compte un impact de marché. Cependant, les prix et les richesses de portefeuille sont des processus de Markov, et plus particulièrement des diffusions Browniennes.
Le théorème de représentation des martingales dans la filtration Brownienne permet de montrer que

$$p^{cl}(p) = \inf \{y \geq 0 : \exists (\nu, \alpha), \ell(V_T - g(S_T)) - P^{\alpha}_T \geq 0\}, \tag{1.1.1}$$

où

$$P^{\alpha}_t = p + \int_0^t \alpha_s \mathrm{d}W_s, \quad t \in [0, T],$$

pour des processus α de carré intégrable vérifiant $P^{\alpha} \in [0, 1]$.
Cet élargissement des espaces d'état et des contrôles permet de reformuler le problème comme un problème classique de cible stochastique avec des contrôles à valeurs dans un fermé non-borné.
En conséquence du principe de programmation dynamique de [ST02a ; ST02b ; BET09], la fonction valeur de ce problème, définie sur $[0, T] \times (0, \infty)^d \times [0, 1]$, est une solution de viscosité d'une EDP non-linéaire. Cependant, l'unicité de la solution à cette EDP n'est pas obtenue dans [BET09], l'opérateur de l'EDP étant discontinu.
Ce travail a été théoriquement adapté à d'autres situations. Entre autres, dans un cadre Markovien, Moreau [Mor11] a étudié le problème en présence de sauts, Bouchard, Bouveret et Chassagneux [BBC16] ont considéré des options bermudéennes et Dumitrescu, Elie, Sabbagh et Zhou [Dum+17] ont considéré des options américaines. Dans un cadre non-Markovien, Bouchard, Elie et Reveillac ont introduit les EDSRs avec condition terminal faible [BER15], et Dumitrescu a étendu leur travail [Dum16]. Finalement, une extension à un nombre fini de contraintes a été étudiée par Bouchard et Vu [BNV12].
Dans un cadre Markovien et sans contrainte de portefeuille, Bouveret et Chassagneux [BC17] ont prouvé un théorème de comparaison pour l'EDP obtenue dans [BET09].

Dans le chapitre 3, nous étudions le problème de sur-réplication avec probabilité p dans le cas d'un marché linéaire. Plus précisément, nous montrons que ce prix est donné par sur-réplication d'un *pay-off* modifié. Nous montrons de plus que l'on peut approcher le problème de cibles stochastiques (1.1.1) par un problème où les contrôles α sont à valeurs dans $[-n, n]$, pour un $n > 0$, et nous obtenons, dans le cas linéaire toujours, une borne supérieure sur la différence entre les deux prix.
Dans le chapitre 4, nous développons un schéma numérique pour approcher le prix de sur-réplication avec probabilité p dans un modèle Markovien et non-linéaire, sans contrainte de portefeuille. Nous prouvons sa convergence et montrons son efficacité sur des exemples numériques.

1.1.4 Estimation de la distribution du PnL

Une autre approche pour contrôler les risques dans le portefeuille du vendeur est d'immobiliser de l'argent à la date $t = 0$, à utiliser dans le cas où la couverture a échoué. En fait, même si l'on vend un produit dérivé au prix de sur-réplication, en pratique il n'est pas facile de calculer et d'implémenter la stratégie de réplication associée. La stratégie effectivement implémentée est en temps discret il y a des contraintes de portefeuille. De plus, le prix de l'actif peut être impacté par la stratégie de couverture, et des coûts de transaction peuvent être appliqués. Certains de ces facteurs peuvent ne pas être pris en compte par le modèle théorique, et la couverture peut ne pas être parfaite en pratique.
Sachant cela, les compagnies d'assurance sont soumises à des contraintes de régulation dans le contexte de Solvabilité II, directive entrée en vigueur le 1er janvier 2016 qui harmonise la régulation des sociétés d'assurance européennes et qui est le cadre prudentiel permettant d'évaluer et de garantir le capital nécessaire pour la solvabilité de ces entreprises. Dans ce cadre, celles-ci doivent calculer les fonds propres réglementaires à immobiliser afin de payer les clients lorsque leur couverture n'est pas parfaite.

Le calcul explicite des fonds propres réglementaires est assez difficile et sophistiqué car il est demandé de calculer le quantile à 99.5% sur la distribution du PnL de la compagnie à horizon 1 an :

$$V = \inf\{v \geq 0 : \mathbb{P}(L_1 \geq v) \leq 0.05\},$$

où L_1 est la variable aléatoire représentant les pertes au temps $t = 1$ an. En d'autres termes, V est la *Value-at-Risk* (V@R) au niveau 99.5% de la distribution des pertes.

Pour calculer V, il est donc nécessaire de calculer, ou au moins d'approximer numériquement, la distribution de L_1.

Pour approximer la loi de L_1, l'approche principale est connue sous le nom de "simulations imbriquées" [GJ10]. Tout d'abord, un jeu de "simulations extérieures" est tiré, qui représente l'évolution des facteurs de risque sous la probabilité historique $\mathbb{P}$. Pour chaque simulation extérieure, à chaque pas de temps, un jeu de "simulations intérieures" est également tiré, afin de calculer les prix et les grecques des passifs pour rebalancer le portefeuille de couverture. Les simulations intérieures sont tirées sous la mesure risque-neutre $\mathbb{Q}$.
Bien que cette approche soit facile à comprendre et à implémenter, elle est très gourmande en temps de calcul et ressources. De plus, aucune information n'est stockée pour un travail futur, s'il est nécessaire de recalculer la distribution des pertes, par exemple à cause d'un changement de modèle.

Dans le chapitre 5, nous développons une nouvelle méthode pour approximer la distribution des pertes, et ainsi calculer les fonds propres réglementaires. Nous obtenons une borne supérieure sur l'erreur commise sur le calcul de mesures de risque spectrales en remplaçant la distribution inconnue par celle estimée par notre méthode. Nous montrons, sur des exemples numériques, que notre méthode est compétitive avec la méthode des simulations imbriquées.

1.1.5 Nos contributions

1.1.5.1 Chapitre 3

Ce chapitre est tout d'abord un chapitre d'introduction, où l'on redonne les définitions et les résultats concernant le problème de sur-réplication avec probabilité p introduit par Föllmer et Leukert [FL99]. Cependant, nous donnons des preuves originales et abordons également, dans la seconde section du chapitre, un nouveau problème.
Nous proposons une nouvelle approche pour résoudre le problème dans un marché linéaire avec un *pay-off* non-Markovien. La preuve est élémentaire et évite d'utiliser le lemme de Neyman-Pearson des mathématiques statistiques.
Soit ξ une variable de carré intégrable connue à la maturité T, qui représente le produit financier européen.
Le marché est linéaire, si (y, ν) est une stratégie, le processus de richesse associé est donné par

$$V_t = y - \int_0^t \left(a_s V_s + b_s^\top Z_s\right) \mathrm{d}s + \int_0^t Z_s \mathrm{d}W_s, \quad t \in [0, T].$$

Le prix de sur-réplication de tout produit dérivé européen est alors donné par

$$V^1 = \mathbb{E}[\Gamma_T \xi]\,,$$

où Γ est le processus vérifiant

$$\Gamma_t = 1 + \int_0^t a_s \Gamma_s \mathrm{d}s + \int_0^t b_s \Gamma_s \mathrm{d}W_s, \quad t \in [0, T].$$

On considère le prix de sur-réplication avec probabilité p de ξ, pour $p \in [0, 1]$,

$$V^p := \inf\left\{y \geq 0 : \exists \nu, \mathbb{P}\left[V_T \geq \xi\right] \geq p\right\}.$$

On montre que ce prix est obtenu comme infimum de prix de sur-réplication de produits dérivés avec *pay-off* $\xi 1_A$, où A varie dans $\mathcal{F}_T^p$ l'ensemble des ensembles $\mathcal{F}_T$-mesurables vérifiant $\mathbb{P}(A) \geq p$:

$$V^p = \inf_{A \in \mathcal{F}_T^p} \mathbb{E}[\Gamma_T \xi 1_A]\,.$$

Soit $q(p) := \inf\left\{q \geq 0 : \mathbb{P}\left[\Gamma_T \xi \leq q\right] \geq p\right\}$ le quantile de la loi de $\Gamma_T \xi$. Nous introduisons alors l'hypothèse suivante, qui permet de résoudre le problème explicitement.

Hypothèse 1.1.1. *Soit $p \in [0, 1]$. Il existe un ensemble $A \in \mathcal{F}_T$ vérifiant :*

i) $\mathbb{P}\left[A\right] = p$,

ii) $\{\Gamma_T \xi < q(p)\} \subset A \subset \{\Gamma_T \xi \leq q(p)\}$.

Sous cette hypothèse, nous prouvons le théorème suivant

Théorème 1.1.1. *Soit $p \in [0, 1]$ tel que l'hypothèse 1.1.1 est satisfaite. Alors n'importe quel ensemble $A^\star$ vérifiant les conditions i)-ii) est optimal pour le problème de contrôle V^p, ce qui signifie que*

$$V^p = \mathbb{E}[\Gamma_T \xi 1_{A^\star}]\,.$$

En application de ce théorème, nous donnons une nouvelle preuve des formules fermées obtenues dans [FL99] dans le modèle de Black et Scholes pour les options vanille.

Dans un second temps, nous considérons la version Markovienne du problème : ξ est de la forme $\xi = g(X_T)$ pour une fonction L-Lipschitzienne $g : \mathbb{R} \to \mathbb{R}_+$, où X est le processus de prix de l'actif sous-jacent. On suppose que X est un mouvement Brownien géométrique. Le marché est linéaire : si (y, ν) est une stratégie où y est la richesse initiale à la date t et $\frac{\nu}{\sigma}$ est la richesse investie dans l'actif, la dynamique du portefeuille est donnée par

$$Y_s^{t,y,\nu} = y + \int_t^s \lambda \nu_u \mathrm{d}u + \int_t^s \nu_u \mathrm{d}W_u, \quad s \in [t, T],$$

où $\lambda = \frac{\mu}{\sigma}$ est la prime de risque.

Dans ce cadre, pour $t \in [0, T]$, $x > 0$ et $p \in [0, 1]$, le problème s'écrit alors

$$v(t, x, p) = \inf \left\{ y \geq 0 \ \mid \ \exists \nu \in \mathcal{A}, \mathbb{P}(Y_T^{t,y,\nu} \geq g(X_T^{t,x})) \geq p \right\}.$$

En posant $P_s^{t,p,\alpha} = p + \int_t^s \alpha_u \mathrm{d}W_u$ pour $\alpha \in \mathbb{H}^2$, on peut montrer que

$$v(t, x, p) = \inf_{\alpha \in \mathcal{A}_p} \mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x}) P_T^{t,p,\alpha} \right],$$

où $\mathcal{A}_p$ est l'ensemble des processus $\alpha \in \mathbb{H}^2$ satisfaisant à $P^{t,p,\alpha} \in [0, 1]$.

Il est connu que la fonction v est convexe par rapport à sa troisième variable. On introduit $w(t, x, q) := \max_{p \in [0,1]} \{pq - v(t, x, p)\}$ la transformée de Fenchel-Legendre de v par rapport à p. Alors on a $v(t, x, p) = \max_{q > 0} \{pq - w(t, x, q)\}$ et $w(t, x, q) = \mathbb{E}^{\mathbb{Q}} \left[g^\sharp(X_T^{t,x}, qQ_T^t) \right]$ avec $g^\sharp(x, q) = (q - g(x))_+$.

On pose alors, pour $\epsilon > 0$, $g_\epsilon^\sharp = \rho_\epsilon * g^\sharp$, où ρ_ϵ est un noyau de régularisation et $*$ est l'opérateur de convolution. Enfin, on pose $w_\epsilon(t, x, q) = \mathbb{E}^{\mathbb{Q}} \left[g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t) \right]$ et $P_s = u_\epsilon(s, X_s^{t,x}, qQ_s^t)$, où $u_\epsilon := \partial_q w_\epsilon$. On a alors le résultat suivant :

Proposition 1.1.1. *P est une martingale, et $\mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x}) P_T \right] \leq v(t, x, p) + 2(1 + L)\epsilon$, où $p = u_\epsilon(t, x, q)$.*

Cette construction, par la formule d'Itô, permet alors d'obtenir un contrôle presque optimal, qui est de la forme

$$\alpha_s^\epsilon = \sigma X_s^{t,x} \partial_x u_\epsilon(s, X_s^{t,x}, qQ_s^t) + \lambda q Q_s^t \partial_q u_\epsilon(s, X_s^{t,x}, qQ_s^t).$$

Soit $N \geq 0$. En introduisant les temps d'arrêt suivants :

$$\tau_1^N = \inf \left\{ s \geq t \ \mid \ X_s^{t,x} \geq \frac{N\epsilon^2}{2K_1\sigma} \right\},$$
$$\tau_2^N = \inf \left\{ s \geq t \ \mid \ Q_s^t \geq \frac{N\epsilon^2}{2K_2|\lambda|q} \right\},$$

où $|\partial_x u_\epsilon| \leq \frac{K_1}{\epsilon^2}$ et $|\partial_q u_\epsilon| \leq \frac{K_2}{\epsilon^2}$, on montre que le contrôle $\alpha_s^{\epsilon,N} := \alpha^\epsilon 1_{\{s \leq \tau_1^N \wedge \tau_2^N\}}$ vérifie $\alpha^{\epsilon,N} \in [-N, N]$ et $P^{t,p,\alpha^{\epsilon,N}} \in [0, 1]$.

On a alors le théorème suivant :

Théorème 1.1.2. *Supposons qu'il existe $\iota > 0$ tel que pour tout $t \in [0,T]$ et $x > 0$, le processus β tel que*

$$\Gamma_T g(X_T^{t,x}) = \mathbb{E}\Big[\Gamma_T g(X_T^{t,x})\Big] + \int_t^T \beta_s \mathrm{d}W_s$$

vérifie $\mathbb{E}\Big[\int_t^T |\beta_s|^{2+\iota}\mathrm{d}s\Big] < +\infty$.
Soit $N > 0$ fixé. Pour $t \in [0,T]$, $x > 0$, $q > 0$, il existe $C > 0$ et $\epsilon > 0$ tels que, pour $p = u_\epsilon(t,x,q)$, on a

$$0 \leq v_N(t,x,p) - v(t,x,p) \leq C\frac{1}{N^{\frac{\iota}{12+8\iota}}}.$$

Ce théorème donne une borne supérieure sur l'erreur commise lorsque l'on remplace le problème initial de sur-réplication avec probabilité p par le problème où l'on demande de plus aux contrôles α d'être à valeurs dans l'intervalle $[-N,N]$.

L'ingrédient essentiel est la convexité de la fonction valeur dans le paramètre p. Il se trouve que ceci reste vrai, même lorsque le marché est non-linéaire, et un travail ultérieur serait de généraliser les résultats obtenus dans cette section dans le cas plus général que l'on étudie dans le chapitre suivant. C'est l'objet d'un travail en cours, en collaboration avec Jean-François Chassagneux.

1.1.5.2 Chapitre 4

Dans le chapitre 4, nous considérons le problème de sur-réplication avec probabilité p dans un marché complet avec un certain type d'imperfections. En particulier, il n'y a pas de contrainte de portefeuille. En s'appuyant d'une part sur le travail de Bouchard, Elie et Touzi [BET09], qui prouve une représentation de la fonction valeur comme solution (au sens de la viscosité) d'une EDP, et d'autre part sur le travail de Bouveret et Chassagneux [BC17] qui prouve dans ce contexte un principe de comparaison, nous construisons un schéma numérique pour approximer le prix de sur-réplication avec probabilité p.

Plus précisément, nous considérons un marché sur lequel un actif d-dimensionnel est coté, et dont le log-prix X évolue suivant

$$X_t = x + \int_0^t \mu(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s, \quad t \in [0,T].$$

Étant donnée une richesse initiale y et une stratégie $\nu \in \mathbb{H}^2$, le processus de richesse associé est donné par

$$Y_t = y - \int_0^t f(s,X_s,Y_s,\nu_s)\mathrm{d}s + \int_0^t \nu_s \mathrm{d}W_s, \quad t \in [0,T],$$

où f est Lipschitzienne par rapport à ses trois dernières variables. En particulier, le marché est complet : tout produit dérivé $\xi \in L^2(\mathcal{F}_T)$ est réplicable, puisque l'EDSR

$$Y_t = \xi + \int_t^T f(s,X_s,Y_s,Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, \quad t \in [0,T],$$

admet une unique solution (Y, Z).

Le prix de sur-réplication avec probabilité p, dans ce contexte, s'écrit

$$v(t,x,p) = \inf \left\{ y \geq 0 : \exists \nu \in \mathbb{H}^2, \mathbb{P}\left[Y_T^{t,x,y,\nu} \geq g(X_T^{t,x}) \right] \geq p \right\},$$

où $(X^{t,x}, Y^{t,x,y,\nu})$ sont les processus qui suivent les dynamiques décrites précédemment sur $[t,T]$, avec les conditions initiales $X_t^{t,x} = x$ et $Y_t^{t,x,y,\nu} = y$.

Dans ce contexte, il est connu que v est l'unique solution de viscosité continue de l'EDP suivante, sur $[0,T) \times \mathbb{R}^d \times (0,1)$,

$$\mathcal{H}(t,x,\varphi,\partial_t\varphi, D\varphi, D^2\varphi) = 0,$$

avec la condition terminale $v(T,x,p) = g(x)p$ et les conditions au bord $v(t,x,0) = 0$ et $v(t,x,1) = V(t,x)$, où V est le prix de sur-réplication du produit.
Ici, $\mathcal{H}$ est l'opérateur continu

$$\mathcal{H}(\Theta) = \sup_{\eta \in \mathcal{S}} H^\eta(\Theta),$$

où pour $\Theta = (t,x,y,b,q,A) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^{d+1}$ et $\eta \in \mathcal{S} \backslash \mathcal{D}$,

$$H^\eta(\Theta) = (\eta^1)^2 \left(-b - f(t,x,y,\mathfrak{z}(x,q,\eta^\flat)) - \mathcal{L}(x,q,A,\eta^\flat) \right)$$

avec $\mathcal{S}$ la sphère unité dans $\mathbb{R}^{d+1}$, $\mathcal{D}$ l'ensemble des vecteurs $\eta \in \mathcal{S}$ tels que leur première composante η^1 est égale à 0, $\eta^\flat := \frac{1}{\eta^1}(\eta^2, \ldots, \eta^{d+1}) \in \mathbb{R}^d$, et

$$\begin{aligned}
\mathfrak{z}(x,q,a) &= q^x \sigma(x) + q^p a, \\
\mathcal{L}(x,q,A,a) &= \mu(x)^\top q^x + \frac{1}{2} \operatorname{Tr}\left[\sigma(x)\sigma(x)^\top A^{xx} \right] + \frac{|a|^2}{2} A^{pp} + a^\top \sigma(x)^\top A^{xp}, \text{ et} \\
q &= (q^x, q^p)^\top, A = \begin{pmatrix} A^{xx} & A^{xp} \\ (A^{xp})^\top & A^{pp} \end{pmatrix}.
\end{aligned}$$

Pour approcher la fonction valeur, nous construisons un schéma numérique de la façon suivante :

1. Nous discrétisons l'espace des contrôles,
2. Nous considérons un schéma *Piecewise Constant Policy Timestepping* (PCPT) pour discrétiser le problème de contrôle en temps,
3. Nous utilisons un schéma de différences finies monotone pour discrétiser en espace le problème de contrôle.

Premièrement, soit $(\mathcal{R}_n)_{n\geq 1}$ une suite croissante d'ensembles fermés de $\mathcal{S}\backslash\mathcal{D}$ tels que

$$\overline{\bigcup_{n\geq 1} \mathcal{R}_n} = \mathcal{S},$$

et pour chaque $n \geq 1$, soit v_n l'unique solution de viscosité continue de l'EDP

$$\sup_{\eta \in \mathcal{R}_n} H^\eta(t,x,\varphi,\partial_t\varphi, D\varphi, D^2\varphi) = 0,$$

avec des conditions au bord appropriées.

Nous prouvons alors la proposition suivante

Proposition 1.1.2. *La fonction v_n converge vers v localement uniformément lorsque n tend vers l'infini. De plus, v_n est l'unique solution de viscosité continue de*

$$-\partial_t \varphi + \mathcal{F}_n(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

avec des conditions au bord appropriées et où, pour $\Xi = (t, x, y, q, A)$, nous posons $\mathcal{F}_n(\Xi) = \sup_{a \in \mathcal{R}_n^\flat} F^a(\Xi)$ et $F^a(\Xi) = -f(t, x, y, \mathfrak{z}(x, q, a)) - \mathcal{L}(x, q, A, a)$.

Nous introduisons ensuite le schéma PCPT [Kry00 ; BJ07 ; RF16 ; DRZ18] utilisé pour la discrétisation en temps.

On fixe $n \geq 1$. Pour $0 \leq t < s \leq T$, $a \in \mathcal{R}_n$ et une fonction continue $\phi : \mathbb{R}^d \times [0, 1] \to \mathbb{R}$, nous posons $S^a(s, t, \phi) : [t, s] \times \mathbb{R}^d \times [0, 1] \to \mathbb{R}$ l'unique solution de

$$-\partial_t \varphi + \mathcal{F}_n(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

avec condition terminale $S^a(s, t, \phi)(s, x, p) = \phi(x, p)$ et conditions au bord $S^a(s, t, \phi)(r, x, p) = B^p(s, t, \phi)(r, x)$ pour $p \in \{0, 1\}$, où $B^p(s, t, \phi)$ pour $p \in \{0, 1\}$ est la solution de

$$-\partial_t \varphi + F^0(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

avec condition terminale $B^p(s, t, \phi)(s, x) = \phi(x, p)$.

Étant donnée une grille discrète $\pi = \{0 = t_0 < t_1 < \cdots < t_\kappa = T\}$, nous définissons une approximation $v_{n,\pi}$ grâce à l'algorithme rétrograde suivant :

1. Initialisation : on pose $v_{n,\pi}(T, x, p) = g(x)p$.
2. Étape rétrograde : pour $k = \kappa - 1, \ldots, 0$, on calcule $w^{k,a} := S^a(t_k, t_{k+1}, v_{n,\pi}(t_{k+1}, \cdot))$ pour chaque $a \in \mathcal{R}_n^\flat$ et on pose

$$v_{n,\pi}(t_k, \cdot) = \min_{a \in \mathcal{R}_n^\flat} w^{k,a}.$$

Nous prouvons le théorème suivant :

Théorème 1.1.3. *La fonction $v_{n,\pi}$ converge vers v_n localement uniformément lorsque $|\pi|$ tend vers 0.*

Nous prouvons le théorème en utilisant le cadre de Barles et Souganidis [BS91], avec des estimations fines sur les EDSRs et des arguments de Barles et Jakobsen [BJ07] pour obtenir la consistance. Les difficultés principales apparaissent lorsque l'on veut vérifier la consistance du schéma au bord du domaine. La monotonie est obtenue en utilisant un théorème de comparaison et une récurrence descendante. La stabilité est montrée en remarquant que la solution du schéma semi-discret est croissante par rapport à la variable p, ce qui donne le prix de sur-réplication comme borne supérieure.

Dans la dernière partie de ce chapitre, lorsque le log-prix est un Brownien avec dérive en dimension 1, nous construisons un schéma entièrement discret en approximant les fonctions $w^{k,a}$ définies précédemment par un schéma de différences finies implicite. Le schéma utilise seulement des opérateurs uni-dimensionnels grâce à la dégénérescence des opérateurs différentiels. En effet, l'EDP est de dimension 2 en espace, mais il n'y a qu'un mouvement Brownien de dimension 1 dirigeant les deux processus contrôlés. Nous approchons $w^{k,a}$ par une fonction définie sur une grille spatiale Γ^a qui dépend du contrôle a. Pour effectuer l'opération de minimisation $v_{n,\pi}(t_k, \cdot) = \min_{a \in \mathcal{R}_n^\flat} w^{k,a}$, il est nécessaire de considérer une étape supplémentaire, qui consiste en une interpolation linéaire, pour étendre les solutions du schéma de différences finies au domaine entier. Nous renvoyons le lecteur à la section 4.3 pour plus de détails sur la définition exacte du schéma, qui produit une fonction $v_{n,\pi,\delta}$ pour un paramètre de discrétisation spatiale $\delta > 0$ et une grille temporelle uniforme π.

Enfin, nous prouvons le théorème suivant

Théorème 1.1.4. *Supposons que* $\pi = \{0 = t_0 < t_1 = h < \cdots < t_\kappa = T = \kappa h\}$ *et que* h, δ *satisfont aux conditions suivantes :*

$$\begin{aligned} \delta &\leq 1, \\ \frac{hL}{2\delta} \leq \theta &< \frac{1}{4}, \\ \mu h \leq \delta &\leq Mh, \end{aligned}$$

où $M > 0$ *est une constante indépendante de* h, δ *et* L *est la constante de Lipschitz du générateur* f*.*
Alors $v_{n,h,\delta}$ *converge vers* v_n *lorsque* $h \to 0$*, localement uniformément.*

Comme pour le théorème précédent, nous prouvons que le schéma numérique est monotone, stable et consistant. Pour obtenir la monotonie, il nous est nécessaire de prouver un théorème de comparaison pour les schémas de différences finies implicites à contrôle fixé entre deux dates de discrétisation, ainsi que d'utiliser la monotonie de l'interpolateur linéaire. Contrairement au cas semi-discret, nous ne prouvons pas que la solution du schéma est croissante en p. En revanche, nous obtenons tout de même des bornes uniformes en les paramètres de discrétisation en utilisant le principe de comparaison et une récurrence descendante. La consistance est obtenue par une analyse fine des propriétés locales du schéma numérique, en démontrant des estimées de type développement limité dans la direction de diffusion. Une analyse délicate de la consistance au bord est une nouvelle fois nécessaire.

Nous concluons ce chapitre par des applications numériques qui mettent en évidence que le schéma construit converge bien, lorsque le log-prix est un mouvement Brownien avec dérive et que la dynamique des portefeuilles de richesse est non-linéaire. Nous montrons également que les conditions de compatibilité sur les paramètres de discrétisations nécessaires théoriquement sont également nécessaires en pratique pour que le schéma converge. Enfin, dans le cas linéaire où il existe des formules fermées pour le prix des options vanille, nous montrons qu'il est également possible de converger vers ce prix théorique en choisissant correctement les paramètres de discrétisation.

Le contenu de ce chapitre est tiré d'un article écrit en collaboration avec Jean-François Chassagneux et Christoph Reisinger [BCR19], soumis au *SIAM Journal on Financial Mathematics.*

1.1.5.3 Chapitre 5

Dans le chapitre 5, nous construisons un schéma numérique pour estimer la distribution des pertes d'une compagnie d'assurance, et les fonds propres réglementaires. Cette procédure est une alternative à la méthode des "simulations imbriquées" [GJ10] et nous donnons des exemples numériques qui montrent que notre approche est compétitive en dimension modérée. De plus, la méthode développée ici permet un recalcul rapide lorsque le modèle et ses paramètres sous la mesure risque-neutre sont inchangés.

Nous décrivons tout d'abord l'approche mise en œuvre pour approcher la distribution des pertes :

1. Nous fixons une grille *sparse*, que nous décrivons ci-après, sur l'espace des facteurs de risque. Pour chaque point de cette grille et pour chaque date de rebalancement, nous calibrons la mesure risque-neutre $\mathbb{Q}$ et calculons le prix et les sensibilités du produit par rapport aux facteurs de risque, par des simulations de Monte Carlo sous $\mathbb{Q}$. Nous obtenons et stockons des interpolateurs qui seront utilisés à la place de la calibration et des "simulations intérieures" dans l'algorithme des simulations imbriquées.

2. Nous tirons un échantillon des facteurs de risque entre $t = 0$ et $t = 1$ (sur la grille des dates de rebalancement), et nous calculons le portefeuille de couverture à chaque date de rebalancement en utilisant les interpolateurs précédemment obtenus.

3. Nous déduisons un échantillon de la distribution des pertes, et nous approximons n'importe quelle mesure de risque à l'aide de cet échantillon.

L'avantage de cette méthode est que les calculs de l'étape 1. sont faits "hors ligne", alors que ceux des deux étapes suivantes sont faits "en ligne". En particulier, si le modèle sous $\mathbb{Q}$ n'est pas changé, les interpolateurs obtenus lors de l'étape 1. peuvent être réutilisés dans le futur, s'il est nécessaire de calculer de nouvelles mesures de risque.

Décrivons tout d'abord les grilles sparse considérées, suivant [BG04], qui permettent de réduire le nombre de points utilisés pour stocker l'approximation numérique des fonctions.
Pour chaque multi-indices $\mathbf{k} \leq \mathbf{l}$, nous définissons un pas via $h_{\mathbf{k}} = 2^{-\mathbf{k}}$ et les points sur la grille par

$$\breve{y}_{\mathbf{k},\mathbf{i}} = (m_1 + i_1(M_1 - m_1)h_{k_1}, \ldots, m_d + i_d(M_d - m_d)h_{k_d}), \ \mathbf{0} \leq \mathbf{i} \leq 2^{\mathbf{k}}.$$

En utilisant les fonctions *tente*,

$$y \in \mathbb{R} \mapsto \phi(y) := \begin{cases} 1 - |y| & \text{si } y \in [-1, 1] \\ 0 & \text{sinon} \end{cases}$$

on peut associer à la grille précédente un ensemble de fonctions de base nodale :

$$y \in \mathbb{R}^d \mapsto \phi_{\mathbf{k},\mathbf{i}}(y; A) = \prod_{l=1}^{d} \phi(\frac{y_l - \breve{y}^l_{\mathbf{k},\mathbf{i}}}{2^{-i_l}}) \ .$$

Lorsque l'on utilise une grille régulière, la fonction est approximée en utilisant toutes les fonctions de base nodale de l'ensemble précédent, au niveau $\mathbf{l}$ le plus fin. Ici, on considère l'espace nodal d'ordre κ défini par

$$\mathcal{V}_\kappa := \text{span}\{\phi_{\mathbf{l},\mathbf{j}}; (\mathbf{l}, \mathbf{j}) \in \mathcal{I}_\kappa(A)\},$$

avec

$$\mathcal{I}_\kappa := \Big\{(\mathbf{l},\mathbf{j}) : \quad 0 \le \sum_{i=1}^{d} l_i \le \kappa; \quad \mathbf{0} \le \mathbf{j} \le \mathbf{2^l}; \\ (l_i > 0 \text{ et } j_i \text{ est impair}) \text{ ou } (l_i = 0), \text{ pour } i = 1, \dots, d\Big\}.$$

Si $\psi : A \to \mathbb{R}$ est supporté dans A, nous définissons son $\mathcal{V}_\kappa$-interpolateur associé par

$$\pi^A_{\mathcal{V}_\kappa}[\psi](y) := \sum_{(\mathbf{l},\mathbf{j}) \in \mathcal{I}_\kappa(A)} \theta_{\mathbf{l},\mathbf{j}}(\psi; A)\phi_{\mathbf{l},\mathbf{j}}(y; A) \tag{1.1.2}$$

où l'opérateur $\theta_{\mathbf{l},\mathbf{j}}$ peut être défini récursivement en termes de r, la dimension de $\mathbf{l}$, par :

$$\theta_{\mathbf{l},\mathbf{j}}(\psi; A) = \begin{cases} \psi(\breve{y}_{\mathbf{l},\mathbf{j}}); & r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r}); A-); & l_r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r}); A-) - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r-1}); A-) & \\ \qquad -\frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \breve{y}^r_{l_r,j_r+1}); A-); & l_r > 0 \end{cases}$$

où, pour un hypercube $A = \prod_{l=1}^{d} [m_l, M_l]$, $A- := \prod_{l=1}^{d-1} [m_l, M_l]$ et pour un multi-indice $\mathbf{k}$ de dimension $r \ge 1$, $\mathbf{k}- = (k_1, \dots, k_{r-1})$.
Dans nos exemples numériques, la grille *sparse* est calculée avec la librairie C++ StOpt [Gev+18].

Nous décrivons maintenant le modèle financier considéré dans ce chapitre.

Dans notre modèle, le bilan de l'entreprise est vide avant la date $t = 0$. A la date $t = 0$, l'entreprise vend un put lookback de maturité $T = 30$ ans.

Le modèle que nous considérons sous la mesure de probabilité risque-neutre $\mathbb{Q}$ est une dynamique de Black et Scholes pour le processus de prix de l'actif risqué de dimension 1, et une dynamique de Hull et White pour le taux court. Plus précisément, les processus sont solutions des Équations Différentielles Stochastiques (EDS) suivantes :

$$\begin{aligned} \mathrm{d}r_t &= a(\mu_t - r_t)\mathrm{d}t + b\mathrm{d}B_t, \\ \mathrm{d}S_t &= r_t S_t \mathrm{d}t + \sigma \mathrm{d}W_t. \end{aligned}$$

Ici, W et B sont deux $\mathbb{Q}$-mouvements Browniens avec corrélation ρ, et les coefficients a, b, σ sont des constantes. La fonction déterministe μ est calibrée grâce à l'observation de la courbe des taux d'intérêt, qui est modélisée comme combinaison linéaire de 3 fonctions élémentaires, qui représentent le court/moyen/long terme.

Dans ce modèle, il y a donc 4 facteurs de risque à observer : le log-prix de l'actif X et les 3 nombres décrivant la courbe des taux $\theta = (\theta_1, \theta_2, \theta_3)$.

Sous la mesure physique $\mathbb{P}$, nous supposons qu'un échantillon des facteurs de risque à la date $t = 1$ nous est fourni. Nous supposons que c'est un vecteur Gaussien et nous estimons les paramètres de la loi grâce à l'échantillon. Nous simulons un échantillon des facteurs de risque entre $t = 0$ et $t = 1$ en supposant que c'est un processus Gaussien de la forme

$$\mathrm{d}\begin{pmatrix} X_t \\ \theta_t \end{pmatrix} = b\mathrm{d}t + c\mathrm{d}(W^i_t)^3_{i=0},$$

où $(W^i)_{i=0}^3$ est un $\mathbb{P}$-mouvement Brownien de dimension 4, $b \in \mathbb{R}^4$ et $c \in \mathbb{R}^{4\times 4}$ est une matrice triangulaire. Les coefficients b et c sont uniquement déterminés de sorte à ce que la loi de (X_1, θ_1) coincide avec la loi précédemment estimée.

Finalement, nous supposons que le portefeuille de couverture est construit avec l'actif sous-jacent et trois *swaps*, pour que l'entreprise soit couverte en Δ et en ρ. Pour ce faire, à chaque date de rebalancement, l'entreprise calcule la dérivée de son passif par rapport à chaque facteur de risque afin de déterminer la nouvelle composition du portefeuille de couverture.

Soit ρ une mesure de risque spectrale. Le résultat principal de ce chapitre donne une borne supérieure pour l'erreur commise lorsque l'on approxime $\rho(\mu)$ par $\rho(\mu^{\mathcal{S}})$, où μ est la vraie distribution des pertes, et $\mu^{\mathcal{S}}$ est la distribution approchée.

Théorème 1.1.5. *Il existe $C > 0$ et $\alpha > 0$ tels que*

$$\mathbb{E}\big[|\rho(\mu) - \rho(\mu^{\mathcal{S}})|^2\big]^{\frac{1}{2}} \leq C\left(\frac{1}{N^\alpha} + n\left\{\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}(\kappa - d + 1)^{(d-1)}\right\}\right),$$

où n est le nombre de dates de rebalancement, κ est le niveau maximum de la grille sparse*, d est le nombre de facteurs de risque, N (resp. M) est la taille de l'échantillon sous $\mathbb{P}$ (resp. sous $\mathbb{Q}$).*

Nous obtenons également une erreur pour l'approximation avec les simulations imbriquées, qui est de la forme

$$\mathbb{E}\big[|\rho(\mu) - \rho(\mu^{\mathcal{N}})|^2\big]^{\frac{1}{2}} \leq C\left(\frac{1}{N^\alpha} + n\sqrt{\frac{\log(N)}{M}}\right),$$

qui ressemble à l'erreur obtenue dans le théorème, où un terme supplémentaire apparaît à cause de l'interpolation dans la grille *sparse*. De plus, notre approche est plus gourmande en terme de mémoire, puisqu'il faut stocker les approximations des fonctions sur les grilles, ce qui est de l'ordre $O\left(n2^{\kappa-d+1}\frac{(\kappa-d+1)^{d-1}}{(d-1)!}\right)$. Cependant, comme nous l'avons déjà mentionné, le gain en temps d'exécution est très important une fois que les calculs hors-ligne ont été effectués. De plus, cette étape est facilement parallélisable puisque les calculs à effectuer sur chaque point de la grille sont indépendants.

Enfin, nous donnons des exemples numériques qui montrent que l'approche par grilles *sparse* est extrêmement compétitive avec l'approche par simulations imbriquées : avec un niveau maximal de 3 seulement, nous obtenons des distributions empiriques similaires par les deux approches, avec le même temps de calcul, alors que le nombre de simulations risque-neutres est seulement de 2000 pour les simulations imbriquées, contre 20000 pour la méthode des grilles. Nous donnons aussi une application au risque de modèle (sous $\mathbb{P}$) : si l'on utilise les simulations imbriquées, pour effectuer n'importe quel recalcul, il faut recommencer du début à chaque fois. En utilisant l'approche par grilles, si seul le modèle sous $\mathbb{P}$ est changé, il est possible de réutiliser les calculs hors-ligne afin d'obtenir la nouvelle distribution empirique quasi-instantanément.

Le contenu de ce chapitre est tiré d'un article écrit en collaboration avec Jérémie Bonnefoy, Jean-François Chassagneux, Shuoqing Deng, Camilo Garcia Trillos et Lionel Lenôtre [Bén+19], publié dans *ESAIM : Proceedings and Surveys*.

1.2 Problèmes de *switching* optimal “randomisés”

Dans la dernière partie de cette thèse, nous introduisons et étudions de nouveaux problèmes de *switching* optimal et leurs connexions avec les EDSRs à réflexions obliques.

1.2.1 Problèmes de *switching* optimal

Nous avons vu dans la première partie de cette introduction que le problème de valorisation d'actifs en finance peut être vu, d'un point de vue mathématique, comme un problème de contrôle stochastique tel qu'un problème de cibles stochastiques. Ces problèmes sont eux-même liés aux solutions d'équations différentielles stochastiques, plus précisément les EDSRs.

En fait, des problèmes économiques tels que la maximisation des profits d'une firme, et donc sa valorisation, peuvent également être vus comme des problèmes de *switching* que nous décrivons maintenant.

Supposons qu'une entreprise peut produire dans différents modes de production. Par exemple [HZ10], une centrale électrique peut être active ou non, ou encore différents niveaux de production peuvent être considérés. Une stratégie de gestion est alors définie par une suite croissante de temps d'arrêt $\mathcal{T} = (\tau_n)_{n\geq 0}$ qui décrit les temps où le gestionnaire décide de mettre en marche (resp. d'arrêter) la centrale. Alors, mathématiquement, les profits espérés de la firme peuvent être modélisés comme

$$J(\mathcal{T}) = \mathbb{E}\left[\int_0^T f(t, X_t, \mathcal{I}_t)\mathrm{d}t + g(X_T, \mathcal{I}_T) - \sum_{i=1}^{N_T} c(\mathcal{I}_t)\right],$$

où X est le processus du prix de l'électricité, f est le profit courant et g est le profit terminal, $\mathcal{I}_t = 1$ (resp. $\mathcal{I}_t = -1$) si la centrale est en marche (resp. à l'arrêt), et c est le coût de changement de mode. On a par exemple

$$\mathcal{I}_t = (-1)^{N_t} \text{ et } N_t = \sum_{n\geq 0} 1_{\{\tau_n \leq t\}}.$$

Une stratégie optimale et la valeur de la centrale sont obtenues en résolvant le problème suivant :

$$V = \sup_{\mathcal{T}} J(\mathcal{T}),$$

où le suprémum est pris sur l'ensemble des stratégies admissibles, par exemple telles que $N_T \in L^2(\mathcal{F}_T)$.

Plus généralement, supposons que l'entreprise peut fonctionner sous n modes de production, et qu'un processus stochastique contrôlé X^a dirige les profits, de dynamique

$$\mathrm{d}X_t = b(t, X_t, a_t)\mathrm{d}t + \sigma(t, X_t, a_t)\mathrm{d}W_t\,,$$

où $a_t = \sum_{i=1}^{N_t} \zeta_i 1_{\{\tau_i \leq t < \tau_{i+1}\}}$ représente le mode de production $\zeta \in \{1, \ldots, n\}$ utilisé sur l'intervalle $[\tau_i, \tau_{i+1})$. Ici, $(\tau_n)_{n\geq 1}$ est une suite croissante de temps d'arrêt, et ζ_i est mesurable par rapport à la tribu au temps τ_i. Le problème de *switching* optimal est alors

$$V = \sup_a \mathbb{E}\left[\int_0^T f(t, X_t^a, a_t)\mathrm{d}t + g(X_T^a, a_T) - \sum_{i\geq 0} c_{\zeta_i,\zeta_{i+1}} 1_{\{\tau_i \leq T\}}\right],$$

où a parcourt l'ensemble des stratégies admissibles, f est le gain courant, g le gain terminal et $c_{i,j}$ est le coût de passage de l'état i vers l'état j.

1.2.2 Lien avec les EDSRs

Le cas où il n'y a que deux modes a été traité par Hamadène et Jeanblanc [HJ07], où ils montrent que résoudre le problème de *switching* optimal revient à résoudre une EDSR à réflexions obliques de dimension deux. Dans ce contexte, cela revient à résoudre une EDSR doublement réfléchie [CK96]

$$\begin{aligned} Y_t = (g(X_T,1) - g(X_T,2)) + \int_t^T (f(s,X_s,1) - f(s,X_s,2))\,\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s \\ + \int_t^T \mathrm{d}K_s^+ - \int_t^T \mathrm{d}K_s^-, \\ -c_1 \leq Y_t \leq c_2 \text{ et } \int_0^T (Y_t + c_1)\mathrm{d}K_t^+ = \int_0^T (c_2 - Y_t)\mathrm{d}K_t^- = 0, \end{aligned}$$

dont la solution (Y, Z, K^+, K^-) satisfait à des conditions d'intégrabilité appropriées, et K^+, K^- sont des processus continus croissants.

Lorsque le nombre de modes est $n \geq 2$ et lorsque la diffusion X n'est pas contrôlée, on peut montrer de façon similaire qu'une EDSR à réflexions obliques donne la solution du problème, voir par exemple [DHP09]. C'est un système de n équations interconnectées (par le terme de réflexion) qui prend la forme suivante :

$$Y_t^i = g(X_T, i) + \int_t^T f(s, X_s, i) - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i, \tag{1.2.1}$$

$$Y_t^i \geq \max_{j\neq i}\left\{Y_t^j - c_{i,j}\right\}, \tag{1.2.2}$$

$$\int_0^T \left(Y_t^i - \max_{j\neq i}\left\{Y_t^j - c_{i,j}\right\}\right)\mathrm{d}K_t^i = 0. \tag{1.2.3}$$

Ici, Y_0^i donne la valeur de la firme, en supposant que l'on démarre du mode i. De plus, on obtient une stratégie optimale en utilisant l'EDSR à réflexions obliques : en partant de $\zeta_0^\star = i$ au temps $\tau_0^\star = 0$, une stratégie optimale est donnée par

$$\begin{aligned} \tau_{i+1}^\star &= \inf\left\{t \geq \tau_i^\star : Y_t^{\zeta_i^\star} = \max_{j\neq i}\left\{Y_t^j - c_{\zeta_i^\star,j}\right\}\right\}, \\ \zeta_{i+1}^\star &= \min\left\{j \in \arg\max\left\{Y_{\tau_{i+1}^\star}^j - c_{\zeta_i^\star,j}\right\}\right\}. \end{aligned}$$

Lorsque seule la dérive de la diffusion X est contrôlée, Hu et Tang [HT10] obtiennent une représentation de la valeur du problème de *switching* comme une EDSR réfléchie. Enfin, lorsque la diffusion X est entièrement contrôlée, le travail de Elie et Kharroubi [EK14] permet d'obtenir une représentation du même type.

1.2.3 EDSRs réfléchies

Nous avons vu que les problèmes de *switching* sont étroitement liés à l'étude des EDSRs réfléchies.
Tout d'abord, les efforts se concentrèrent sur les réflexions normales, avec les travaux de Gegout-Petit et Pardoux [GPP96] en dimension quelconque, et ceux de El Karoui, Kapoudjian, Pardoux, Peng et Quenez [EK+97] en dimension 1 pour une barrière, et de Cvitanic et Karatzas [CK96] pour deux barrières. Des applications respectives à la valorisation d'options américaines (et aux problèmes d'arrêt optimal) et d'options de jeu sont données dans ces deux travaux.

L'émergence des problèmes de *switching* motiva ensuite l'étude des EDSRs à réflexions obliques associées (1.2.1)-(1.2.2)-(1.2.3), mais également d'équations plus générales de la forme

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s, Z_s) - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i, \tag{1.2.4}$$

$$Y_t^i \geq \max_{j \neq i} \left\{ Y_t^j - c_{i,j} \right\}, \tag{1.2.5}$$

$$\int_0^T \left(Y_t^i - \max_{j \neq i} \left\{ Y_t^j - c_{i,j} \right\} \right) \mathrm{d}K_t^i = 0, \tag{1.2.6}$$

où $\xi = (\xi^1, \ldots, \xi^n)$ est une variable aléatoire de carré intégrable connue à la date T, et les processus K^i sont continus et croissants.

Dans le cas où $f^i(t, y, z) = f^i(t, y^i, z^i)$ pour chaque i, l'existence et l'unicité sont obtenus par Hu et Tang [HT10]. L'existence est obtenue par un argument de pénalisation, et l'unicité est obtenue par vérification, en identifiant la solution avec un problème de *switching* formel défini par des *EDSRs switchées*. Malheureusement, cette approche ne s'étend pas au cas d'un générateur plus général.
Hamadène et Zhang [HZ10] ont considéré des réflexions obliques plus générales et ont autorisé le générateur f à dépendre de y tout entier, mais ils imposent que f soit croissant par rapport à chaque y^i.
Dans le contexte de réflexions du type *switching*, Chassagneux, Elie et Kharroubi [CEK12] démontrent l'existence et l'unicité de solutions à (1.2.4)-(1.2.5)-(1.2.6) dans le cas où $f^i(t, y, z) = f^i(t, y, z^i)$.

Enfin, Chassagneux et Richou [CR18] considèrent un cadre général pour les EDSRs à réflexions obliques en étudiant les équations de la forme

$$\begin{aligned}
&Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s - \int_t^T H(s, Y_s, Z_s)\Phi_s \mathrm{d}s, \quad t \in [0, T], \\
&Y_t \in \mathcal{D}, \Phi_t \in \mathcal{C}(Y_t), \quad t \in [0, T], \\
&\int_0^T \Phi_t 1_{\{Y_t \notin \partial\mathcal{D}\}} \mathrm{d}t = 0,
\end{aligned}$$

où $\mathcal{D}$ est un domaine convexe de $\mathbb{R}^d$, et si $y \in \partial\mathcal{D}$, $\mathcal{C}(y)$ est le cône normal extérieur à $\mathcal{D}$ au point y. La fonction $H : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^{d \times k} \to \mathbb{R}^{d \times d}$ est fixée *a priori* et modélise la direction de réflexion, qui peut donc potentiellement être aléatoire et dépendre du processus Z.
Dans un cadre non-Markovien, sous des hypothèses de régularité sur le bord du domaine,

de symétrie et de non-dégénérescence sur H, les auteurs prouvent l'existence et l'unicité de la solution. Dans un cadre Markovien, ils obtiennent l'existence de solutions sous des hypothèses plus faibles. La non-dégénérescence de la direction de réflexion H reste demandée.

1.2.4 Nos contributions

Dans le chapitre 6, nous considérons une nouvelle classe de problèmes de *switching*. Dans le cas où le générateur f vérifie $f^i(t,y,z) = f^i(t,y^i,z^i)$ pour tout $i \in \{1,\dots,n\}$, $t \in [0,T]$, $y \in \mathbb{R}^d$ et $z \in \mathbb{R}^{d\times k}$, nous prouvons que, sous l'hypothèse d'existence d'une solution pour une EDSRs à réflexions obliques, il y a unicité, et cette solution coïncide avec la fonction valeur d'un problème de *switching*. Utilisant les résultats de [CR18], nous prouvons l'existence de solutions pour ces EDSRs dans ce que nous appelons le cas *irréductible et non-contrôlé*, que nous décrivons en détails plus bas.

Le problème que nous considérons est un problème de contrôle optimal de type *switching*. La nouveauté est que l'agent ne peut pas choisir directement le nouveau mode. À la place, il peut choisir parmi des distributions de probabilité, et le nouveau mode sera tiré selon la loi qu'il choisit.

Remarque 1.2.1. *Cette spécification donne lieu à des coûts incertains : si la distribution du nouveau mode est $\sum_{i=1}^n p_i\delta_i$, alors le coût à payer est $c = \sum_{i=1}^n c_i 1_{\{\zeta=i\}}$, où ζ est le nouveau mode. Cependant, dans l'analyse mathématique du problème, ce point de vue n'est pas mis en évidence. En effet, l'agent cherchant à optimiser ses gains espérés, seul le gain moyen $\frac{1}{n}\sum_{i=1}^n p_i c_i$ intervient dans l'analyse.*

Remarque 1.2.2. *En utilisant des mesures de Dirac sur $\{1,\dots,n\}$, on peut écrire le problème de* switching *classique dans ce cadre.*

Nous travaillons dans un espace de probabilités $(\Omega, \mathcal{G}, \mathbb{P})$ où $\mathcal{G}$ est engendrée par un mouvement Brownien W et une famille de variables aléatoires indépendantes $(X_n)_{n\geq 1}$, distribuées uniformément sur $[0,1]$ et indépendantes de W.

Nous considérons un espace métrique compact ordonné $\mathscr{C}$, qui est l'espace des contrôles.

A chaque $u \in \mathscr{C}$ est associé une probabilité de transition sur l'espace d'états $\{1,\dots,n\}$, définie par $p^u_{i,j} := \mathbb{P}(F(u,i,X) = j)$, où X est distribuée uniformément sur $[0,1]$ et $F : \mathscr{C} \times \{1,\dots,n\} \times [0,1] \to \{1,\dots,n\}$ est mesurable. En termes du jeu, cela signifie que si le mode actuel est i et que l'agent décide de changer de mode en utilisant le contrôle u, le nouvel état sera choisi, indépendamment de tout ce qui précède, selon la distribution de probabilités $\sum_{j=1}^n p^u_{i,j}\delta_j$ sur $\{1,\dots,n\}$.

On se donne également une fonction de coût $c : \{1,\dots,n\} \times \mathscr{C} \to \mathbb{R}, (i,u) \mapsto c_{i,u}$ qui représente le coût de changement de mode, partant de l'état i et utilisant le contrôle u.

Nous introduisons $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ la filtration Brownienne.

Une stratégie $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1})$ pour le jeu est alors donnée par une suite croissante de temps aléatoires $(\tau_n)_{n\geq 0}$ et une suite de variables aléatoires à valeurs dans $\mathscr{C}$, vérifiant :

- $\tau_0 \in [0,T]$ et $\zeta_0 \in \{1,\dots,n\}$ sont déterministes.

- Pour tout $n \geq 0$, τ_{n+1} est un $\mathbb{F}^n$-temps d'arrêt et α_{n+1} est $\mathcal{F}^n_{\tau_{n+1}}$-mesurable. Le nouveau mode est alors $\zeta_{n+1} := F(\alpha_{n+1}, \zeta_n, X_{n+1})$, et nous définissons une nouvelle filtration $\mathbb{F}^{n+1} = (\mathcal{F}^{n+1}_t)_{t\geq 0}$ avec $\mathcal{F}^{n+1}_t = \mathcal{F}^n_t \vee \sigma(X_{n+1}1_{\{\tau_{n+1}\leq t\}})$ qui incorpore l'information du nouvel état.

Pour une stratégie ϕ, nous définissons $\mathbb{F}^\infty = (\mathcal{F}^\infty_t)_{t\geq 0}$ avec $\mathcal{F}^\infty_t = \bigvee_{n\geq 0} \mathcal{F}^n_t$.

Une stratégie est *admissible* si $A^\phi_T - A^\phi_{\tau_0} \in L^2(\mathcal{F}^\infty_T)$ et $\mathbb{E}\Big[\Big(A^\phi_{\tau_0}\Big)^2 |\mathcal{F}^0_{\tau_0}\Big]$ est fini presque-sûrement, où $A^\phi_t = \sum_{n\geq 0} c_{\zeta_n,\alpha_{n+1}} 1_{\{\tau_{n+1}\leq t\}}$ est le processus de coûts cumulés.

Suivant Hu et Tang [HT10], l'agent cherche à maximiser $\mathbb{E}\Big[U^\phi_{\tau_0} - A^\phi_{\tau_0}|\mathcal{F}^0_{\tau_0}\Big]$, où (U^ϕ, V^ϕ, M^ϕ) est la solution unique sur $(\Omega, \mathcal{G}, \mathbb{F}^\infty, \mathbb{P})$ de l'EDSR suivante sur $[0, T]$,

$$U_t = \xi^{a_T} + \int_t^T f^{a_s}(s, U_s, V_s)\mathrm{d}s - \int_t^T V_s \mathrm{d}W_s - \int_t^T \mathrm{d}M_s - \int_t^T \mathrm{d}A^\phi_s.$$

Une part importante de ce travail est d'étudier la filtration $\mathbb{F}^\infty$. Plus précisément, nous détaillons la structure des martingales dans cette filtration, et nous montrons que les EDSRs avec générateur Lipschitzien et condition terminale L^2 admettent une solution unique. Par exemple, nous obtenons un théorème de représentation pour les variables aléatoires $\xi \in L^2(\mathcal{F}^\infty_T)$ qui permet de décomposer ξ comme la somme de son espérance, d'une intégrale stochastique contre le mouvement Brownien et d'un terme de sauts. On montre que les sauts ne peuvent se produire qu'aux instants $\tau_i, i \geq 0$, et sont nécessairement de la forme

$$\mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^k_{\tau_{k+1}}\Big].$$

Nous renvoyons le lecteur à la section 6.4.1 pour plus de détails sur la filtration $\mathbb{F}^\infty$. Les résultats de cette partie nécessitent une analyse fine de la filtration $\mathbb{F}^\infty$ et des filtrations $\mathbb{F}^k, k \geq 0$.

Cette étude montre que le problème d'optimisation est bien défini. De plus, dans le cas de coûts positifs et en supposant l'existence d'une solution à l'EDSR à réflexions obliques suivante :

$$Y^i_t = \xi^i + \int_t^T f^i(s, Y^i_s, Z^i_s)\mathrm{d}s - \int_t^T Z^i_s \mathrm{d}W_s + \int_t^T \mathrm{d}K^i_s, \quad (1.2.7)$$

$$Y_t \in \mathcal{D}, \quad (1.2.8)$$

$$\int_0^T \left(Y^i_t - \sup_{u\in\mathscr{C}} \left\{\sum_{j=1}^d p^u_{i,j} Y^j_t - c_{i,u}\right\}\right) \mathrm{d}K^i_t = 0, \quad (1.2.9)$$

où

$$\mathcal{D} = \left\{y \in \mathbb{R}^d : y_i \geq \sup_{u\in\mathscr{C}} \left\{\sum_{j=1}^d p^u_{i,j} y_j - c_{i,u}\right\}, i = 1, \dots, d\right\}$$

est d'intérieur non-vide et $\xi \in L^2(\mathcal{F}^\infty_T)$ prend ses valeurs dans $\mathcal{D}$, nous identifions Y^i_t avec la valeur du jeu démarrant dans le mode i à la date t. De plus, une stratégie optimale est donnée par $\zeta_0^\star = i, \tau_0^\star = t$ et

$$\tau_{k+1}^\star = \inf\left\{\tau_k^\star \leq s \leq T : Y_s^{\zeta_k^\star} = \sup_{u\in\mathscr{C}}\left\{\sum_{j=1}^d p^u_{\zeta_k^\star,j} Y_s^j - c_{\zeta_k^\star,u}\right\}\right\} \wedge (T+1),$$
$$\alpha_{k+1}^\star = \inf\left\{v \in \mathscr{C}, v \in \arg\max_{u\in\mathscr{C}}\left\{\sum_{j=1}^d p^u_{\zeta_k^\star,j} Y^j_{\tau_k^\star} - c_{\zeta_k^\star,u}\right\}\right\}.$$

Les difficultés de l'analyse proviennent du fait que nous ne travaillons pas dans la filtration Brownienne, comme dans le cas de problèmes de *switching* usuels. Le grossissement de filtration induit des termes de martingales orthogonales au mouvement Brownien qu'il faut traiter dans l'analyse des EDSRs. Il faut alors adapter soigneusement des arguments classiques des EDSRs pour tenir compte de ces termes supplémentaires.
De plus, puisque l'agent ne choisit pas directement le nouveau mode, il est possible, même pour la stratégie optimale, d'observer des changements de mode simultanés. Il nous est alors nécessaire de montrer que pour la stratégie optimale, le coût cumulé engendré par les sauts simultanés au temps initial t reste dans $L^2(\mathcal{F}_t^0)$. Il est possible de démontrer cela en étudiant soigneusement la chaîne de Markov non-homogène associée à la stratégie optimale dans un "coin" du bord du domaine, en utilisant le fait que celui-ci n'est pas d'intérieur vide.
Ces deux ingrédients permettent de montrer que la stratégie optimale est admissible, et qu'elle induit effectivement un profit espéré maximal.

Dans la suite du chapitre, nous nous intéressons à l'existence de la solution à (1.2.7)-(1.2.8)-(1.2.9). Nous considérons le cas "irréductible" et "non-contrôlé", c'est-à-dire lorsque l'ensemble des contrôles est un singleton $\mathscr{C} = \{u\}$. Il y a donc une seule probabilité de transition représentée par une matrice P, ainsi qu'un vecteur de coûts c. L'agent décide seulement quand changer d'état. S'il est actuellement dans l'état i, il paye c_i pour changer d'état, et son nouvel état est déterminé sous la loi $\sum_{j=1}^n p_{i,j}\delta_j$. Nous supposons que la matrice P est irréductible.
Nous considérons des coûts arbitraires, c'est-à-dire que nous ne supposons pas $c_i > 0$ pour tout i. Si c'est le cas, il est clair que le domaine est d'intérieur non-vide.
Nous démontrons des conditions nécessaires et suffisantes sur les coûts c pour obtenir un domaine d'intérieur non-vide. Cette caractérisation fait intervenir la matrice $C = (C^{i,j})_{i,j=1,\dots,n}$, où $C^{i,j}$ est le coût pour passer "directement" de l'état i à l'état j. En pratique, $C^{i,j}$ est obtenu comme le coût "moyen" à payer pour passer de l'état i à l'état j, ce qui se traduit mathématiquement par

$$C^{i,j} = \left((I_{n-1} - P^{(i)})^{-1}c^{(i)}\right)_j \text{ si } i \neq j,$$
$$= 0 \text{ si } i = j,$$

où $P^{(i)}$ (resp. $c^{(i)}$) est la matrice (resp. le vecteur) où l'on a retiré la ligne et la colonne (resp. la coordonnée) i, et où I_{n-1} est la matrice identité de taille $n-1$.
Nous démontrons, lorsque le domaine est non-vide, que chaque colonne de la matrice $-C$ est un point de $\mathcal{D}$.
De plus, nous démontrons que l'ensemble $\mathcal{D} \cap \{y_n = 0\}$ est un simplexe dont les points extrémaux sont donnés par les colonnes de la matrice $(C^{n,j} - C^{i,j})_{i,j=1,\dots,n}$. On obtient

alors des conditions nécessaires et suffisantes sur la matrice C pour que le domaine soit d'intérieur non-vide.

Lorsque le domaine est d'intérieur non-vide, sous une hypothèse technique de copositivité sur la matrice P, nous montrons qu'il existe une solution à (1.2.7)-(1.2.8)-(1.2.9) dans un cadre Markovien. Pour ce faire, nous appliquons le théorème de Chassagneux et Richou [CR18] en construisant une application H qui, pour chaque point sur la frontière du domaine, envoie son cône normal sur le cône des directions de réflexion imposées par le jeu. Cette application est tout d'abord construite sur les points extrémaux de $\mathcal{D} \cap \{y_n = 0\}$, qui est ici un simplexe. Après avoir déterminé le cône normal en chacun de ces points, on obtient naturellement une construction de la fonction H en ces points. On étend ensuite la fonction H à $\mathcal{D} \cap \{y_n = 0\}$ par combinaisons convexes, et à $\mathcal{D}$ tout entier en utilisant l'invariance du domaine par translation par le vecteur $(1, \ldots, 1)$. Finalement, on obtient une fonction définie sur $\mathbb{R}^d$ par projection sur $\mathcal{D}$.

Le contenu de ce chapitre est tiré d'un travail en cours de préparation, en collaboration avec Jean-François Chassagneux et Adrien Richou.

Chapter 2

Introduction - English

This manuscript investigates three problems coming from Mathematical Finance and the optimal control of stochastic processes. The aim of this chapter is to introduce and motivate the questions we studied and to summarize the main results obtained.

2.1 Quantile hedging and Capital requirement

The first two parts of this thesis are dedicated to the pricing and hedging of contingent claims in mathematical finance, and to the risk management associated to partial hedging.
We are, in the first part of this thesis, interested in a weak pricing procedure called pricing with controlled loss, and in particular quantile hedging. We first study the case of a linear market, where we compute optimal controls and take advantage of their explicit form to derive an upper bound for the error made when approximating the control problem with bounded controls. In a market with some imperfections, we design a numerical scheme to approximate the quantile hedging price of an European contingent claim. We prove its convergence and we numerically show its pertinence.
In the second part, after we model a financial market with stochastic interest rates, we suggest a method to estimate the distribution of the Profit and Loss (PnL for short) of a financial institution, in order to compute the capital requirement in the Solvency II framework. We obtain, for spectral risk measures, an upper bound for the error made by our approximation procedure, and we show on numerical examples that the suggested method is competitive with nested simulations.

2.1.1 Pricing and hedging in complete markets

We consider a continuous time financial market modelled by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, in which there is a risk free asset with constant return $r \geq 0$, and a d-dimensional risky asset S, which is a semi-martingale.

Suppose that one agent wants to sell an European claim with maturity $T > 0$ and pay-off $g(S_T)$, where g is a measurable function from $\mathbb{R}$ to $\mathbb{R}_+$, satisfying appropriate integrability conditions. An important question in financial mathematics is to find a price. We assume that the market is arbitrage-free and that it is possible to construct a replicating portfolio: starting from a wealth p, one can find an investment strategy such that the portfolio liquidation value at time T equals exactly $g(S_T)$. In that case, the price must be p.

Mathematically, these arguments were formalised by the pioneering work of Black and Scholes [BS73], Merton [Mer73; MS90] and the studies of Harrison and Kreps [HK79], Harrison and Pliska [HP81], Duffie [Duf88], Karatzas [Kar89] and Delbaen and Schachermayer [DS94] among others: given an investment strategy (y, ν) where y is the initial wealth and ν_t^i represents the number of shares in asset i owned at time t, the portfolio value evolves according to the following dynamics:

$$V_t = y + \int_0^t \nu_s \mathrm{d}S_s + \int_0^t r(V_s - \nu_s S_s)\mathrm{d}s, \quad t \in [0, T].$$

In that case, it is possible to find a *risk-neutral* probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$, under which $(e^{-rt}S_t)_{t\geq 0}$ is a martingale, and we then obtain

$$p = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}g(S_T)\right].$$

Of course, when we want to model financial markets more accurately, we realize that, due to imperfections in the market (for example transaction costs, non-uniqueness of interest rate, trading constraints), pricing and hedging are not so straightforward: the market can now be "non-linear" and incomplete. In that case, the portfolio's dynamics may become non-linear.
One can write, for example,

$$V_t = y - \int_0^t f(s, S_s, V_s, \sigma(s, S_s)\nu_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\nu_s \mathrm{d}W_s, \quad t \in [0, T],$$

where $f : [0, T] \times (0, \infty)^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a measurable function subject to assumptions, and the price process S is described by the dynamics

$$S_t = S_0 + \int_0^t \mu(s, S_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\mathrm{d}W_s, \quad t \in [0, T],$$

with W being a Brownian motion and μ, σ being coefficients valued in space of appropriate dimension and satisfying hypotheses so that S and V are uniquely defined.

Remark 2.1.1. *Other dynamics for S and V can be and have been considered, for example with jumps [TL94; Par97].*

Example 2.1.1. *If $\mu(t, x) \equiv \mu \in \mathbb{R}, \sigma(t, x) \equiv \sigma > 0$ and there is a borrowing rate R and a lending rate $r \leq R$, the function f is given, see [EKPQ97], by*

$$f(y, z) = -ry - \sigma^{-1}\mu z + (R - r)(y - \sigma^{-1}z)^-.$$

In this non-linear setting, if the market is complete, then the financial argument above applies. The price of any contingent claim is now obtained through solving a Backward Stochastic Differential Equation (BSDE for short):

$$Y_t = g(S_T) + \int_t^T f(t, S_t, Y_t, Z_t)\mathrm{d}t - \int_t^T Z_t \mathrm{d}W_t, \quad t \in [0, T],$$

where f, the driver of the BSDE, satisfies appropriate conditions.
More precisely, the price at $t = 0$ is given by Y_0, and the hedging strategy is given by the process Z.

In the case where $f(y, z) = -ry - \sigma^{-1}\mu z$ is linear, we recover the previous case, and one can show, see [EKPQ97], that

$$Y_0 = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}g(S_T)\right],$$

where $\mathbb{Q}$ is the risk-neutral probability measure.

For each specification of the dynamics, studying these BSDEs gives informations about a financial market and the possibility to price and hedge European claims, and the (exact or numerical) resolution allows to compute prices and hedging strategies.

2.1.2 Super-replication of contingent claims

When the market is complete, the BSDEs introduced in the previous subsection admit a unique solution, for every contingent claim $g(S_T)$. The price of the claim is then given by the value at $t = 0$ of Y and the associated hedge is given by Z.
However, one can find models where the market is incomplete. This can happen, for example, if we allow the price process S to have jumps, or else if we restrict the admissible self-financing strategies ν to live in a closed subset $D \subset \mathbb{R}^d$.

Even if the seller cannot find a replicating strategy, he may be able to find a strategy (y, ν) satisfying $V_T \geq g(S_T)$ almost surely. The claim can thus be sold at the *super-replication price*, see [EKQ95], defined by

$$p^s = \inf\left\{y \geq 0 : \exists \nu, V_T \geq g(S_T)\right\},$$

the largest price which doesn't introduce arbitrage in the market. If the infimum is attained, this price induces a hedging strategy for the claim.

However, this strategy can be difficult to implement in practice, due to large and fast varying positions.
Thus a generalization of this problem was considered, where a strategy is admissible only if it satisfies to some portfolio constraints. Among others, Cvitanic and Karatzas [CK93], and Föllmer and Kramkov [FK97] studied this problem respectively in the diffusion case and in the semi-martingale case.
This generalized problem was then studied systematically as a new stochastic optimal control problem, called "stochastic target problem", by Soner and Touzi [ST02a; ST02b] for controls valued in a compact set, and by Bouchard, Elie, Touzi [BET09] for controls valued in a closed set. Using the dynamic version of the stochastic control problem, that is the problem starting at time t with initial prices $S_t = x$, a dynamic programming principle is proved [ST02a; ST02b; BET09] for the value function $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}$. This principle takes the following form:

- **(GDP1)** For all $y > v(t, x)$, there exists a strategy ν such that, for all stopping time $t \leq \tau \leq T$, we have

$$V_\tau^{y,\nu} \geq v(\tau, S_\tau).$$

- **(GDP2)** For all $y < v(t,x)$, all strategy ν and all stopping time $t \leq \tau \leq T$, we have

$$\mathbb{P}\left(V_\tau^{y,\nu} > v(\tau, S_\tau)\right) < 1.$$

Moreover, the value function is a viscosity solution of a Partial Differential Equation (PDE for short) of Hamilton-Jacobi-Bellman type, as it is the case for classical stochastic optimal control problems.

However, in practice, the prices obtained by solving this problem can be too high to be interesting. This led practitioners and academics to think about new pricing methods. When the actual price p satisfies $p < p^{sr}$, in view of the previous control problem, there does not exist any strategy (p,ν) satisfying $V_T \geq g(S_T)$. The seller of the claim has some risk in its portfolio to manage: whatever hedging strategy he applies, there exists $\hat{\Omega} \subset \Omega$ with $\mathbb{P}(\hat{\Omega}) > 0$ and $V_T < g(S_T)$ on $\hat{\Omega}$.

2.1.3 Pricing and hedging with controlled loss

We have just seen that, to have a reasonable price for some contingent claims, it is needed to chose a price below the super-replication one. Any hedging strategy is thus unsuccessful with positive probability. Thus one can consider a price such that there exists a strategy which allows to minimise some risk measure on the loss distribution. There are different notions of price, as pricing by indifference utility [Car08], which have also been studied.
One first possibility is, given a probability of success $p \in [0,1]$, to solve the following problem, introduced by Föllmer and Leukert in [FL99],

$$p^{qh}(p) = \inf\left\{y \geq 0 : \exists \nu, \mathbb{P}\left(Y_T \geq g(S_T)\right) \geq p\right\}.$$

This problem is the so-called "quantile hedging problem", and it is the main concern of the first part of this thesis.
We notice that $p^{qh}(1) = p^s$, so this problem is a generalization of the super-replication problem described above.
In [FL99], a subtle use of the Neyman-Pearson from mathematical statistics allows the authors to solve the problem in a linear market where the prices are driven by general semi-martingales. Moreover, they derive closed form formulae for the quantile hedging price of vanilla options in the Black and Scholes framework. However, their approach does not extend to the non-linear case.

Bouchard, Elie and Touzi in [BET09] generalised [FL99] by considering the following stochastic problem, called *stochastic target problem with controlled loss*:

$$p^{cl}(p) = \inf\left\{y \geq 0 : \exists \nu, \mathbb{E}\left[\ell(V_T - g(S_T))\right] \geq p\right\},$$

where ℓ is a non-decreasing function. Taking $\ell = 1_{\mathbb{R}_+}$ shows that the quantile hedging problem of [FL99] is a special case of the previous one.
In their work [BET09], the authors allow for portfolio constraints and the price S to be controlled by the strategy, to take into account market impact. However the stock prices and the portfolio wealth processes are necessarily Markov processes, and more

particularly Brownian diffusions.
The martingale representation theorem in the Brownian filtration allows to show that

$$p^{cl}(p) = \inf\left\{y \geq 0 : \exists(\nu, \alpha), \ell(V_T - g(S_T)) - P_T^{\alpha} \geq 0\right\}, \tag{2.1.1}$$

where

$$P_t^{\alpha} = p + \int_0^t \alpha_s \mathrm{d}W_s, \quad t \in [0, T],$$

for square-integrable processes α such that $P^{\alpha} \in [0, 1]$.
This enlargement of the state and control spaces allows to recast the problem as a standard stochastic target problem with controls living in an unbounded closed set.
A consequence of the dynamic programming principle of [ST02a; ST02b; BET09] is that the value function of this problem, defined on $[0, T] \times (0, \infty)^d \times [0, 1]$, is a viscosity solution of a non-linear PDE. However, uniqueness for the solution of the PDE is not proved in [BET09], as the operator is not continuous.
This work was theoretically extended in various settings. Among others, in a Markovian setting, Moreau [Mor11] studied the problem in presence of jumps, Bouchard, Bouveret and Chassagneux [BBC16] considered Bermudan claims and Dumitrescu, Elie, Sabbagh and Zhou [Dum+17] considered American ones. In a non-Markovian setting, Bouchard, Elie and Reveillac introduced BSDEs with weak terminal conditions [BER15], and Dumitrescu extended their work [Dum16]. Finally, an extension to a finite number of constraints was studied by Bouchard and Vu [BNV12].
In a Markovian framework and without portfolio constraints, Bouveret and Chassagneux [BC17] prove that a comparison theorem holds for the PDE obtained in [BET09].

In Chapter 3, we study the quantile hedging problem in a linear market. More precisely, we show that this price is given by super-replication of a modified pay-off. We moreover show that we can approach the stochastic target problem (2.1.1) with the problem with controls α valued in $[-n, n]$, pour $n > 0$, and we obtain, in the linear case, an upper bound for the difference between the two prices.
In Chapter 4, we design a numerical scheme to approximate the quantile hedging problem in a non-linear Markovian framework without portfolio constraints. We prove its convergence and show its efficiency on numerical examples.

2.1.4 Estimation of the PnL distribution

An other approach to manage the risks in the portfolio of the seller is to lock in money at time $t = 0$ to use in case of unsuccessful hedge to pay the client.
In fact, even if one sells the claim with the theoretical super-replication price, in practice it is not easy to compute and implement the associated hedging strategy. The strategy actually implemented is discretised in time and there are portfolio constraints. In addition, the asset price can be impacted by the strategy, and transaction costs may apply. Some of these factors can be ignored by the theoretical model, hence the hedge can be not perfect.
This is why nowadays insurance companies have to met regulation constraints in the Solvency II framework, which is a directive harmonising European insurance firms regulation since January 1st 2016 and which is the prudential framework assessing the

required solvency capital of these companies. Since then, they have to compute the capital to immobilize to pay clients when their hedge is not perfect.

Explicit computation of the capital requirement is quite difficult and involved as it requires to compute the 99.5% quantile of the distribution of the PnL of the company at horizon 1 year:

$$V = \inf \{v \geq 0 : \mathbb{P}(L_1 \geq v) \leq 0.05\},$$

where L_1 is the random variable representing the loss at horizon 1 year. In other terms, V is the so-called Value-at-Risk (V@R for short) at level 99.5% of the loss distribution.

To compute V, it is thus needed to compute, or at least to approximate numerically, the distribution of L_1.

To approximate the law of L_1, the main approach is known as the *nested simulations* approach [GJ10]. First, a set of "outer simulations" is drawn, which represents the evolution of the risk factors under the physical probability $\mathbb{P}$. For each outer simulation, at each time step, a sample of "inner simulations" is drawn, to compute the prices and the Greeks of the liabilities, in order to rebalance the hedging portfolio. The inner sample is drawn using the risk-neutral probability measure $\mathbb{Q}$.
While this approach is easy to understand and to implement, it is quite greedy. In addition, no information is stored for future work or for recalculation due to a model change, for example.

In Chapter 5, we develop a new method to approximate the loss distribution, and hence the capital requirement. We obtain an upper bound on the error made by computing spectral risk measures with the estimated distribution instead of the unknown one. We show, on numerical examples, that our procedure is competitive with nested simulations.

2.1.5 Our contributions

2.1.5.1 Chapter 3

First, this chapter is an introductory chapter, where we give definitions and first results about the quantile hedging problem introduced by Föllmer and Leukert [FL99]. However, we give new proofs and we also study, in the second section, a new problem. We suggest a new approach to solve the problem in a linear market with a non-Markovian pay-off. The proof is elementary and avoids to invoke the Neyman-Pearson lemma from mathematical statistics.
Let the non-Markovian European claim be represented by a random variable ξ which is square-integrable and known at the maturity T.
The market is linear, if (y, ν) is a strategy, the associated wealth process is given by

$$V_t = y - \int_0^t \left(a_s V_s + b_s^\top Z_s\right) \mathrm{d}s + \int_0^t Z_s \mathrm{d}W_s, \quad t \in [0, T].$$

The super-replication price of any European claim is thus given by

$$V^1 = \mathbb{E}[\Gamma_T \xi],$$

where Γ is the process satisfying to

$$\Gamma_t = 1 + \int_0^t a_s \Gamma_s \mathrm{d}s + \int_0^t b_s \Gamma_s \mathrm{d}W_s, \quad t \in [0,T].$$

We consider the quantile hedging price of ξ, for $p \in [0,1]$,

$$V^p := \inf \left\{ y \geq 0 : \exists \nu, \mathbb{P}\left[V_T \geq \xi\right] \geq p \right\}.$$

We show that the quantile hedging price of ξ is obtained as the infimum of the super-replication price of contingent claims with pay-off $\xi 1_A$, where A varies inside $\mathcal{F}_T^p$, the set of $\mathcal{F}_T$-measurable sets A satisfying to $\mathbb{P}(A) \geq p$:

$$V^p = \inf_{A \in \mathcal{F}_T^p} \mathbb{E}[\Gamma_T \xi 1_A].$$

Let $q(p) := \inf \left\{ q \geq 0 : \mathbb{P}\left[\Gamma_T \xi \leq q\right] \geq p \right\}$ the quantile of the law of $\Gamma_T \xi$. We then introduce the following assumption, which allows to solve the problem explicitly.

Assumption 2.1.1. *Let $p \in [0,1]$. There exists a set $A \in \mathcal{F}_T$ satisfying:*

i) $\mathbb{P}\left[A\right] = p$,

ii) $\{\Gamma_T \xi < q(p)\} \subset A \subset \{\Gamma_T \xi \leq q(p)\}$.

Under this assumption, we prove the following theorem

Theorem 2.1.1. *Let $p \in [0,1]$ such that Assumption 2.1.1 is satisfied. Then any set $A^\star$ satisfying to the conditions i)-ii) is optimal for the control problem V^p, meaning that*

$$V^p = \mathbb{E}[\Gamma_T \xi 1_{A^\star}].$$

As an application to this theorem, we provide a new derivation of the closed formulae obtained in [FL99] in the Black and Scholes model for vanilla options.

In the second part, we consider a Markovian setting: ξ is of the form $\xi = g(X_T)$ for a L-Lipschitz function $g : \mathbb{R} \to \mathbb{R}_+$ and X is the asset price process. We assume that X is a geometric Brownian motion. The market is linear, if (y, ν) is a strategy where y is the initial (at time t) endowment and $\frac{\nu}{\sigma}$ is the wealth invested in the asset, the portfolio's dynamics is given by

$$Y_s^{t,y,\nu} = y + \int_t^s \lambda \nu_u \mathrm{d}u + \int_t^s \nu_u \mathrm{d}W_u, \quad s \in [t,T],$$

where $\lambda = \frac{\mu}{\sigma}$ is the risk-premium.

In this setting, for $t \in [0,T]$, $x > 0$ and $p \in [0,1]$, the problem writes

$$v(t,x,p) = \inf \left\{ y \geq 0 \;\mid\; \exists \nu \in \mathcal{A}, \mathbb{P}(Y_T^{t,y,\nu} \geq g(X_T^{t,x})) \geq p \right\}.$$

Setting $P_s^{t,p,\alpha} = p + \int_t^s \alpha_u \mathrm{d}W_u$ for $\alpha \in \mathbb{H}^2$, one can show that

$$v(t,x,p) = \inf_{\alpha \in \mathcal{A}_p} \mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x}) P_T^{t,p,\alpha} \right],$$

where $\mathcal{A}_p$ is the set of processes $\alpha \in \mathbb{H}^2$ satisfying to $P^{t,p,\alpha} \in [0,1]$.

It is known that v is convex with respect to its third argument. We introduce $w(t,x,q) := \max_{p\in[0,1]} \{pq - v(t,x,p)\}$ the Fenchel-Legendre transform of v with respect to p. Then we have $v(t,x,p) = \max_{q>0} \{pq - w(t,x,q)\}$ and $w(t,x,q) = \mathbb{E}^{\mathbb{Q}} \left[g^\sharp(X_T^{t,x}, qQ_T^t) \right]$ with $g^\sharp(x,q) = (q - g(x))_+$.

We then set, for $\epsilon > 0$, $g_\epsilon^\sharp = \rho_\epsilon * g^\sharp$, where ρ_ϵ is a mollifier and $*$ the convolution operator. Last, we set $w_\epsilon(t,x,q) = \mathbb{E}^{\mathbb{Q}} \left[g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t) \right]$ and $P_s = u_\epsilon(s, X_s^{t,x}, qQ_s^t)$, where $u_\epsilon := \partial_q w_\epsilon$. We then obtain the following result:

Proposition 2.1.1. *P is a martingale and $\mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x} P_T) \right] \leq v(t,x,p) + 2(1+L)\epsilon$, where $p = u_\epsilon(t,x,q)$.*

This construction, by Itô's formula, shows that an almost optimal control is given by

$$\alpha_s^\epsilon = \sigma X_s^{t,x} \partial_x u_\epsilon(s, X_s^{t,x}, qQ_s^t) + \lambda q Q_s^t \partial_q u_\epsilon(s, X_s^{t,x}, qQ_s^t).$$

Let $N \geq 0$. Introducing the following stopping times:

$$\tau_1^N = \inf \left\{ s \geq t \mid X_s^{t,x} \geq \frac{N\epsilon^2}{2K_1\sigma} \right\},$$
$$\tau_2^N = \inf \left\{ s \geq t \mid Q_s^t \geq \frac{N\epsilon^2}{2K_2|\lambda|q} \right\},$$

where $|\partial_x u_\epsilon| \leq \frac{K_1}{\epsilon^2}$ and $|\partial_q u_\epsilon| \leq \frac{K_2}{\epsilon^2}$, we show that the control $\alpha_s^{\epsilon,N} := \alpha^\epsilon 1_{\{s \leq \tau_1^N \wedge \tau_2^N\}}$ satisfies $\alpha^{\epsilon,N} \in [-N,N]$ and $P^{t,p,\alpha^{\epsilon,N}} \in [0,1]$.

We then have the following theorem:

Theorem 2.1.2. *Assume that there exists $\iota > 0$ such that for all $t \in [0,T]$ and $x > 0$, the process β such that*

$$\Gamma_T g(X_T^{t,x}) = \mathbb{E}\Big[\Gamma_T g(X_T^{t,x})\Big] + \int_t^T \beta_s \mathrm{d}W_s$$

satisfies $\mathbb{E}\Big[\int_t^T |\beta_s|^{2+\iota} \mathrm{d}s\Big] < +\infty$.
Let $N > 0$. For $t \in [0,T]$, $x > 0$, $q > 0$, there exists $C > 0$ and $\epsilon > 0$ such that, for $p = u_\epsilon(t,x,q)$, we have

$$0 \leq v_N(t,x,p) - v(t,x,p) \leq C \frac{1}{N^{\frac{\iota}{12+8\iota}}}.$$

This theorem gives an upper bound on the error made when we replace the quantile hedging problem by a similar problem with controls in $[-N,N]$.

The main tool is the convexity of v in p. One can notice that it remains true, even in a non-linear market, and a future work would be to generalise the results obtained here in the more general case studied in the next chapter. This is a work in progress with Jean-François Chassagneux.

2.1.5.2 Chapter 4

In Chapter 4, we consider the quantile hedging problem in a complete market with some imperfections. In particular, there is no portfolio constraints. Using the work of Bouchard, Elie and Touzi [BET09], which provides a PDE representation (in the viscosity sense) for the value function of the problem, and the work of Bouveret and Chassagneux [BC17] which proves in this context a comparison theorem, we construct a numerical scheme to approximate the quantile hedging price.

More precisely, we consider a market in which trades a d-dimensional risky asset, whose log-price X evolves as

$$X_t = x + \int_0^t \mu(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s, \quad t \in [0,T].$$

Given an initial wealth y and a strategy $\nu \in \mathbb{H}^2$, the associated wealth process satisfies

$$Y_t = y - \int_0^t f(s, X_s, Y_s, \nu_s)\mathrm{d}s + \int_0^t \nu_s \mathrm{d}W_s, \quad t \in [0,T],$$

where f is Lipschitz-continuous with respect to its three last variables. In particular, the market is complete as every contingent claim $\xi \in L^2(\mathcal{F}_T)$ is replicable, since the BSDE

$$Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, \quad t \in [0,T],$$

admits a unique solution (Y, Z).

The quantile hedging price, in this context, writes

$$v(t,x,p) = \inf\left\{y \geq 0 : \exists \nu \in \mathbb{H}^2, \mathbb{P}\left[Y_T^{t,x,y,\nu} \geq g(X_T^{t,x})\right] \geq p\right\},$$

where $(X^{t,x}, Y^{t,x,y,\nu})$ are the processes following the dynamics described above on $[t,T]$, with initial conditions $X_t^{t,x} = x$ and $Y_t^{t,x,y,\nu} = y$.

In this context, it is known that v is the unique continuous viscosity solution of the following PDE in $[0,T) \times \mathbb{R}^d \times (0,1)$,

$$\mathcal{H}(t, x, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0,$$

with the terminal condition $v(T,x,p) = g(x)p$ and the boundary conditions $v(t,x,0) = 0$ and $v(t,x,1) = V(t,x)$, where V is the super-replication price of the claim.
Here, $\mathcal{H}$ is the continuous operator

$$\mathcal{H}(\Theta) = \sup_{\eta \in \mathcal{S}} H^\eta(\Theta),$$

where for $\Theta = (t,x,y,b,q,A) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^{d+1}$ and $\eta \in \mathcal{S}\backslash\mathcal{D}$,

$$H^\eta(\Theta) = (\eta^1)^2\left(-b - f(t,x,y,\mathfrak{z}(x,q,\eta^\flat)) - \mathcal{L}(x,q,A,\eta^\flat)\right)$$

with $\mathcal{S}$ the unit sphere in $\mathbb{R}^{d+1}$, $\mathcal{D}$ the set of vectors $\eta \in \mathcal{S}$ such that their first component η^1 equals 0, $\eta^\flat := \frac{1}{\eta^1}(\eta^2, \ldots, \eta^{d+1}) \in \mathbb{R}^d$, and

$$\mathfrak{z}(x, q, a) = q^x \sigma(x) + q^p a,$$
$$\mathcal{L}(x, q, A, a) = \mu(x)^\top q^x + \frac{1}{2} \operatorname{Tr}\left[\sigma(x)\sigma(x)^\top A^{xx}\right] + \frac{|a|^2}{2} A^{pp} + a^\top \sigma(x)^\top A^{xp}, \text{ and}$$
$$q = (q^x, q^p)^\top, A = \begin{pmatrix} A^{xx} & A^{xp} \\ (A^{xp})^\top & A^{pp} \end{pmatrix}.$$

To approximate the value function, we design a numerical scheme as follows:

1. We discretise the control set,
2. We consider an associated Piecewise Constant Policy Timestepping (PCPT) scheme for the control process to discretise in time the problem.
3. We use a monotone finite difference scheme to discretise in space the control problem.

First, let $(\mathcal{R}_n)_{n\geq 1}$ be an increasing sequence of closed subsets of $\mathcal{S}\backslash\mathcal{D}$ such that

$$\overline{\bigcup_{n\geq 1} \mathcal{R}_n} = \mathcal{S},$$

and for each $n \geq 1$, let v_n be the unique continuous viscosity solution of the PDE

$$\sup_{\eta \in \mathcal{R}_n} H^\eta(t, x, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0,$$

with appropriate boundary conditions.

We then prove the following proposition:

Proposition 2.1.2. *The functions v_n converge to v locally uniformly as $n \to \infty$. Moreover, v_n is the unique viscosity solution to*

$$-\partial_t \varphi + \mathcal{F}_n(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

with appropriate boundary conditions and where, for $\Xi = (t, x, y, q, A)$, we set $\mathcal{F}_n(\Xi) = \sup_{a \in \mathcal{R}_n^\flat} F^a(\Xi)$ and $F^a(\Xi) = -f(t, x, y, \mathfrak{z}(x, q, a)) - \mathcal{L}(x, q, A, a)$.

We next introduce the PCPT scheme [Kry00; BJ07; RF16; DRZ18] used for the time discretization.

We fix $n \geq 1$. For $0 \leq t < s \leq T$, $a \in \mathcal{R}_n$ and a continuous $\phi : \mathbb{R}^d \times [0, 1] \to \mathbb{R}$, we denote by $S^a(s, t, \phi) : [t, s] \times \mathbb{R}^d \times [0, 1] \to \mathbb{R}$ the unique solution of

$$-\partial_t \varphi + \mathcal{F}_n(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

with terminal condition $S^a(s, t, \phi)(s, x, p) = \phi(x, p)$ and boundary conditions $S^a(s, t, \phi)(r, x, p) = B^p(s, t, \phi)(r, x)$ for $p \in \{0, 1\}$, where $B^p(s, t, \phi)$ for $p \in \{0, 1\}$ is the solution to

$$-\partial_t \varphi + F^0(t, x, \varphi, D\varphi, D^2\varphi) = 0,$$

with terminal condition $B^p(s, t, \phi)(s, x) = \phi(x, p)$.

Given a discrete grid $\pi = \{0 = t_0 < t_1 < \cdots < t_\kappa = T\}$, we define our approximation $v_{n,\pi}$ by the following backward algorithm:

1. Initialisation: set $v_{n,\pi}(T,x,p) = g(x)p$.
2. Backward step: For $k = \kappa - 1, \dots, 0$, compute $w^{k,a} := S^a(t_k, t_{k+1}, v_{n,\pi}(t_{k+1}, \cdot))$ for each $a \in \mathcal{R}_n^\flat$ and set

$$v_{n,\pi}(t_k, \cdot) = \min_{a \in \mathcal{R}_n^\flat} w^{k,a}.$$

We prove the following theorem:

Theorem 2.1.3. *The function $v_{n,\pi}$ converges to v_n locally uniformly as $|\pi| \to 0$.*

We prove the theorem using Barles and Souganidis [BS91] framework, together with fine BSDEs estimates and arguments from Barles and Jakobsen [BJ07] to obtain consistency. The main difficulties arise while checking consistency of the scheme at the boundary of the domain. Monotonicity is obtained by a comparison theorem and a backward induction. Stability is shown by noticing that the solution to the semi-discrete scheme is non-decreasing with respect to p, which gives an upper bound, namely the super-replication price.

In the last part of this chapter, when the log-price follows is a drifted Brownian motion in dimension 1, we construct a fully discrete numerical scheme by approximating the functions $w^{k,a}$ defined above with implicit finite difference schemes. The scheme uses only one-dimensional derivatives, exploiting the degeneracy of the differential operators involved. Indeed, the PDE is of spatial dimension 2, but there is only one Brownian motion of dimension 1 driving the two controlled processes.
We approximate $w^{k,a}$ by a function defined on a spatial grid Γ^a, which depends on the control a. To perform the minimisation operation $v_{n,\pi}(t_k, \cdot) = \min_{a \in \mathcal{R}_n^\flat} w^{k,a}$, it is needed to consider a supplementary step which consists in a linear interpolation to extend the solution to the finite difference scheme to the whole domain.
We refer to Section 4.3 for more details about the exact definition of the scheme, which produces a function $v_{n,\pi,\delta}$ for a spatial discretization parameter $\delta > 0$ and an uniform time grid π.

Lastly, we prove the following theorem

Theorem 2.1.4. *Assume that $\pi = \{0 = t_0 < t_1 = h < \dots < t_\kappa = T = \kappa h\}$ and that h, δ satisfy to the following conditions:*

$$\begin{aligned} \delta &\le 1, \\ \frac{hL}{2\delta} \le \theta &< \frac{1}{4}, \\ \mu h \le \delta &\le Mh, \end{aligned}$$

where $M > 0$ is a constant independent of h, δ and L is the Lipschitz constant of the driver f.
Then $v_{n,\pi,\delta} \to v_n$ as $h \to 0$, locally uniformly.

As for the precedent theorem, we prove that the scheme is monotone, stable and consistent. To obtain the monotonicity, we need to prove a comparison theorem for the finite differences schemes for fixed control between two discretization dates, and we use the monotonicity of the linear interpolator. In contrast with the semi-discrete case, we do not prove that the solution is non-decreasing with p. However, we still get uniform bounds with respect to the discretization parameters using the comparison principle and a backward induction. Consistency is obtained by a fine analysis of local properties of the scheme. More precisely, we obtain Taylor expansions in the direction of diffusion. A delicate analysis of the consistency at the boundary is once again necessary.

We conclude this chapter with numerical applications showing that the scheme is converging, when the log-price is a drifted Brownian motion and the dynamics of the wealth portfolios is non-linear. We also show that the compatibility conditions on the discretization parameters, theoretically necessary, are also necessary in practice to obtain convergence. In a linear market, where there exists closed formulae for the quantile hedging price of vanilla options, we show that it is possible to make our scheme converge toward this theoretical price by tuning carefully the discretization parameters.

The content of this chapter is from an article in collaboration with Jean-François Chassagneux and Christoph Reisinger [BCR19], submitted to SIAM Journal on Financial Mathematics.

2.1.5.3 Chapter 5

In Chapter 5, we design an numerical procedure to estimate the loss distribution of an insurance company and the capital requirement. This procedure is an alternative to the nested simulations approach [GJ10] and we provide numerical examples which show that is is competitive in moderate dimensions. In addition, the method developed here allows fast recalculation when the model and the parameters under the risk-neutral probability remain the same.

We describe first the approach we consider to approximate the loss distribution:

1. We fix a sparse grid, described below, on the space of risk factors. For each point of the sparse grid and for each rebalancing time, we calibrate the risk-neutral measure $\mathbb{Q}$ and we compute the claim's price and its derivatives with respect to each risk factors, by Monte Carlo simulations under $\mathbb{Q}$. We obtain and store interpolators to be used in place of the calibration and the "inner simulations" in the nested simulations algorithm.

2. We draw a sample of the risk factors evolution between $t = 0$ and $t = 1$ (on the grid of rebalancing dates), and we compute the hedging portfolio at each rebalancing time using the previously computed interpolators.

3. We obtain a sample of the loss distribution and we approximate any risk measure using this sample.

The advantage of this method is that the computations at step 1. are done "offline", while the two next steps are done "online". In particular, if the model under $\mathbb{Q}$ remains unchanged, the interpolators computed in 1. can be used again in the future in case of a recomputation.

Let us describe first the sparse grids we consider, following [BG04], which allow to reduce the number of points used to store the numerical approximation of functions. For each multi-index $\mathbf{k} \leq \mathbf{l}$, we define a grid mesh $h_{\mathbf{k}} = 2^{-\mathbf{k}}$ and grid points

$$\check{y}_{\mathbf{k},\mathbf{i}} = (m_1 + i_1(M_1 - m_1)h_{k_1}, \ldots, m_d + i_d(M_d - m_d)h_{k_d}), \ \mathbf{0} \leq \mathbf{i} \leq 2^{\mathbf{k}}.$$

Using the *hat* function,

$$y \in \mathbb{R} \mapsto \phi(y) := \begin{cases} 1 - |y| & \text{if } y \in [-1, 1] \\ 0 & \text{otherwise} \end{cases}$$

we can associate to the previous grid a set of nodal basis function:

$$y \in \mathbb{R}^d \mapsto \phi_{\mathbf{k},\mathbf{i}}(y; A) = \prod_{l=1}^{d} \phi(\frac{y_l - \check{y}^l_{\mathbf{k},\mathbf{i}}}{2^{-i_l}}) \; .$$

When using a "full" linear interpolation, the function is approximated using the whole set of nodal basis function at the finest level $\mathbf{l}$. Instead, we consider the sparse grid nodal space of order κ defined by

$$\mathcal{V}_\kappa := \text{span}\{\phi_{\mathbf{l},\mathbf{j}}; (\mathbf{l}, \mathbf{j}) \in \mathcal{I}_\kappa(A)\},$$

where

$$\mathcal{I}_\kappa := \Big\{(\mathbf{l}, \mathbf{j}) : \quad 0 \le \sum_{i=1}^{d} l_i \le \kappa; \quad \mathbf{0} \le \mathbf{j} \le \mathbf{2^l}; \\ (l_i > 0 \text{ and } j_i \text{ is odd}) \text{ or } (l_i = 0), \text{ for } i = 1, \ldots, d\Big\}.$$

For a function $\psi : A \to \mathbb{R}$ with support in A, we define its associated $\mathcal{V}_\kappa$-interpolator by

$$\pi^A_{\mathcal{V}_\kappa}[\psi](y) := \sum_{(\mathbf{l},\mathbf{j}) \in \mathcal{I}_\kappa(A)} \theta_{\mathbf{l},\mathbf{j}}(\psi; A)\phi_{\mathbf{l},\mathbf{j}}(y; A) \qquad (2.1.2)$$

where the operator $\theta_{\mathbf{l},\mathbf{j}}$ can be defined recursively in terms of r, the dimension of $\mathbf{l}$, by:

$$\theta_{\mathbf{l},\mathbf{j}}(\psi; A) = \begin{cases} \psi(\check{y}_{\mathbf{l},\mathbf{j}}); & r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \check{y}^r_{l_r,j_r}); A-); & l_r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \check{y}^r_{l_r,j_r}); A-) - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \check{y}^r_{l_r,j_r-1}); A-) & \\ \qquad - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot, \check{y}^r_{l_r,j_r+1}); A-); & l_r > 0 \end{cases}$$

where, for a hypercube $A = \prod_{l=1}^{d}[m_l, M_l]$, $A- := \prod_{l=1}^{d-1}[m_l, M_l]$ and for a multi-index $\mathbf{k}$ with dimension $r \ge 1$, $\mathbf{k}- = (k_1, \ldots, k_{r-1})$.
In our numerical examples, the sparse grid will be computed using the StOpt C++ library [Gev+18].

We now describe the financial model we considered.

In our model, the balance sheet of the company is empty before time $t = 0$. At time $t = 0$, it sells a put lookback with maturity $T = 30$ years.

The model we consider under the risk-neutral probability $\mathbb{Q}$ is a Black and Scholes dynamics of dimension 1 for the stock process S, together with a Hull and White dynamics for the short rate. More precisely, the processes are solution to the following Stochastic Differential Equations:

$$\mathrm{d}r_t = a(\mu_t - r_t)\mathrm{d}t + b\mathrm{d}B_t, \\ \mathrm{d}S_t = r_t S_t \mathrm{d}t + \sigma \mathrm{d}W_t.$$

Here, W and B are two $\mathbb{Q}$-Brownian motions with correlation parameter ρ, and the coefficients a, b, σ are constants. The deterministic function μ is calibrated with the observation of the interest rates curve, which is modelled as a linear combination of 3 elementary functions, which represent short/middle/long term.

In this model, there are then 4 risk factors to observe: the stock log-price X, and the 3 numbers describing the interest rates curve $\theta = (\theta_1, \theta_2, \theta_3)$.

Under the physical probability $\mathbb{P}$, we assume that a sample of the risk factors' law at $t = 1$ is given to us. We assume that it is a Gaussian vector and we estimate the parameters of the law thanks to the sample. Then samples of the risk factors between 0 and 1 are simulated assuming that it is a Gaussian process of the form

$$\mathrm{d}\begin{pmatrix} X_t \\ \theta_t \end{pmatrix} = b\mathrm{d}t + c\mathrm{d}(W_t^i)_{i=0}^3,$$

where $(W^i)_{i=0}^3$ is a four-dimensional $\mathbb{P}$-Brownian motion, $b \in \mathbb{R}^4$ and $c \in \mathbb{R}^{4\times 4}$ is a triangular matrix. The coefficients b and c are determined uniquely so that the law of (X_1, θ_1) coincides with the previously estimated law.

Finally, we assume that the hedging portfolio is constructed using the stock and three swaps, in order to be Δ-hedged and ρ-hedged. To this effect, at each rebalancing time, the company differentiates the price of the claim with respect to each risk factors to compute the new portfolio's composition.

Let ρ be a spectral risk measure.
The main result of the chapter gives an upper bound for the error made when we approximate $\rho(\mu)$ by $\rho(\mu^{\mathcal{S}})$, where μ is the true loss distribution and $\mu^{\mathcal{S}}$ is the approximate distribution.

Theorem 2.1.5. *There exists $C > 0$ and $\alpha > 0$ such that*

$$\mathbb{E}\big[|\rho(\mu) - \rho(\mu^{\mathcal{S}})|^2\big]^{\frac{1}{2}} \leq C\left(\frac{1}{N^\alpha} + n\left\{\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}(\kappa - d + 1)^{(d-1)}\right\}\right),$$

where n is the number of rebalancing dates, κ is the maximum level of the sparse grid, d is the number of risk factors and N (resp. M) is the size of the $\mathbb{P}$-sample (resp. $\mathbb{Q}$-sample).

We also obtain a bound for the approximation with the nested simulations approach, which is of the form

$$\mathbb{E}\big[|\rho(\mu) - \rho(\mu^{\mathcal{N}})|^2\big]^{\frac{1}{2}} \leq C\left(\frac{1}{N^\alpha} + n\sqrt{\frac{\log(N)}{M}}\right),$$

which looks like the error obtained in the theorem, where an extra term comes from the interpolation due to the sparse grid approach. The sparse grid approach is also greedier in term of memory requirement as it is needed to store the sparse grid approximations, of order $O\left(n2^{\kappa-d+1}\frac{(\kappa-d+1)^{d-1}}{(d-1)!}\right)$. However, as already mentioned, the gain in running time is important once the "offline" computations are done. Moreover, since the computations on each point of the grid is independent, our algorithm is easily parallelizable.

Lastly, we exhibit numerical examples showing that the sparse grid approach is highly competitive with the nested simulations approach: with a maximal level of 3 only, we obtain a similar distribution with the two approaches in the same amount of time, while the number of risk-neutral simulations is 2000 for the nested simulations, while it is 20000 for the sparse grid method. We also provide an application to risk of model (under $\mathbb{P}$): if one uses the nested simulations approach, to perform any recomputation, one has to start from the beginning at each time. Using the grid approach, when we change the model under $\mathbb{P}$, we can use again the "offline" computations to compute the new empirical distribution almost instantly.

The content of this chapter is from an article in collaboration with Jérémie Bonnefoy, Jean-François Chassagneux, Shuoqing Deng, Camilo Garcia Trillos and Lionel Lenôtre [Bén+19], published in ESAIM: Proceedings and Surveys.

2.2 Randomised switching control problems

In the last part of this thesis, we introduce and study new optimal switching problems, and their connections with obliquely reflected BSDEs.

2.2.1 Optimal switching problems

We have seen in the first half of the introduction that pricing problems from finance can be viewed, from a mathematical point of view, as stochastic control problems, like stochastic target problems for example. These problems in turn are closely related to solutions of stochastic differential equations, namely BSDEs.

In fact, economic problems like maximisation of a firm's profit, and thus its valuation, can be viewed as stochastic control problems of switching type, as we describe now.

Suppose that a firm can run its production in different modes. For example [HZ10], a power plant can be activated or not, or different levels of production can be considered. A management strategy is then defined as a non-decreasing sequence of stopping times $\mathcal{T} = (\tau_n)_{n\geq 0}$ which describes the times when the manager decides to turn on (resp. off) the production. Then, mathematically, the firm's expected profits can be modelled as

$$J(\mathcal{T}) = \mathbb{E}\left[\int_0^T f(t, X_t, \mathcal{I}_t)\mathrm{d}t + g(X_T, \mathcal{I}_T) - \sum_{i=1}^{N_T} c(\mathcal{I}_t)\right],$$

where X is the price process of electricity, f is the running profit over time, g is the terminal profit, $\mathcal{I}_t = 1$ (resp. $\mathcal{I}_t = -1$) if the power plant is running (resp is shut down), and c is the cost of switching. We have for example

$$\mathcal{I}_t = (-1)^{N_t} \text{ and } N_t = \sum_{n\geq 0} 1_{\{\tau_n \leq t\}}.$$

An optimal management strategy and the value of the power plant are obtained by solving the following problem:

$$V = \sup_{\mathcal{T}} J(\mathcal{T}),$$

where the sup is taken over the set of admissible strategies, for example such that $N_T \in L^2(\mathcal{F}_T)$.

More generally, assume that the firm can be run with n different modes of production, and that a controlled stochastic process X^a drives the profit, with dynamics

$$\mathrm{d}X_t = b(t, X_t, a_t)\mathrm{d}t + \sigma(t, X_t, \nu_t)\mathrm{d}W_t$$

where $a_t = \sum_{i=1}^{N_t} \zeta_i 1_{\{\tau_i \le t < \tau_{i+1}\}}$ represents the mode of production $\zeta \in \{1, \dots, n\}$ on the time interval $[\tau_i, \tau_{i+1})$. Here, $(\tau_n)_{n \ge 1}$ is a sequence of non-decreasing stopping times and ζ_i is measurable with respect to the sigma-algebra stopped at τ_i. The optimal switching problem is then

$$V = \sup_a \mathbb{E}\left[\int_0^T f(t, X_t^a, a_t)\mathrm{d}t + g(X_T^a, a_T) - \sum_{i \ge 0} c_{\zeta_i, \zeta_{i+1}} 1_{\{\tau_i \le T\}}\right],$$

where a runs through the set of admissible strategies, f is the running profit, g is the terminal profit and $c_{i,j}$ is the cost for switching from regime i to regime j.

2.2.2 Connection with BSDEs

The case where there are only two modes was dealt with by Hamadène and Jeanblanc [HJ07], where they showed that solving the optimal switching problem is equivalent to solving a two-dimensional obliquely reflected BSDE, which, in their context, can be transformed into a doubly reflected BSDE [CK96]

$$\begin{aligned} Y_t = (g(X_T, 1) - g(X_T, 2)) &+ \int_t^T (f(s, X_s, 1) - f(s, X_s, 2))\,\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s \\ &+ \int_t^T \mathrm{d}K_s^+ - \int_t^T \mathrm{d}K_s^-, \\ -c_1 \le Y_t \le c_2 \text{ and } &\int_0^T (Y_t + c_1)\mathrm{d}K_t^+ = \int_0^T (c_2 - Y_t)\mathrm{d}K_t^- = 0, \end{aligned}$$

whose solution (Y, Z, K^+, K^-) satisfies appropriate integrability conditions, and K^+, K^- are continuous increasing processes.

When the number of modes is $n \ge 2$, one can show similarly that an obliquely reflected BSDE gives the solution of the problem, see for example [DHP09]. It is a system of n interconnected equations (through the reflection term) which takes the following form:

$$Y_t^i = g(X_T, i) + \int_t^T f(s, X_s, i) - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i, \tag{2.2.1}$$

$$Y_t^i \ge \max_{j \ne i} \left\{Y_t^j - c_{i,j}\right\}, \tag{2.2.2}$$

$$\int_0^T \left(Y_t^i - \max_{j \ne i}\left\{Y_t^j - c_{i,j}\right\}\right) \mathrm{d}K_t^i = 0. \tag{2.2.3}$$

Here, Y_0^i gives the firm's value, assuming that one is starting from mode i. In addition, we also obtain an optimal strategy using the obliquely reflected BSDE: starting from $\zeta_0^\star = i$ at time $\tau_0^\star = 0$, an optimal strategy is given by

$$\tau_{i+1}^\star = \inf\left\{t \geq \tau_i^\star : Y_t^{\zeta_i^\star} = \max_{j\neq i}\left\{Y_t^j - c_{\zeta_i^\star,j}\right\}\right\},$$
$$\zeta_{i+1}^\star = \min\left\{j \in \arg\max\left\{Y_{\tau_{i+1}^\star}^j - c_{\zeta_i^\star,j}\right\}\right\}.$$

When only the drift of X is controlled, Hu and Tang [HT10] obtain a representation of the value as a reflected BSDE. Lastly, when the diffusion of X is also controlled, Elie and Kharroubi [EK14] obtain a similar representation.

2.2.3 Reflected BSDEs

We have seen that switching problems are closely related to the study of obliquely reflected BSDEs.
First, focus was given on normal reflections, with the works of Gegout-Petit and Pardoux [GPP96] in arbitrary dimension, and these of El Karoui, Kapoudjian, Pardoux, Peng and Quenez [EK+97] in dimension 1 for one barrier, and of Cvitanic and Karatzas [CK96] for two barriers. In these two last works, some applications are given respectively to American option pricing (and optimal stopping problems) and game options.

The growing interest in switching problems was then a motivation to study obliquely reflected BSDEs of the form (2.2.1)-(2.2.2)-(2.2.3), and more generally of the form

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s, Z_s) - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i, \tag{2.2.4}$$

$$Y_t^i \geq \max_{j\neq i}\left\{Y_t^j - c_{i,j}\right\}, \tag{2.2.5}$$

$$\int_0^T \left(Y_t^i - \max_{j\neq i}\left\{Y_t^j - c_{i,j}\right\}\right) \mathrm{d}K_t^i = 0, \tag{2.2.6}$$

where $\xi = (\xi^1, \ldots, \xi^n)$ is a square-integrable random-variable known at time T and the processes K^i are continuous and non-decreasing.

In the case where $f^i(t, y, z) = f^i(t, y^i, z^i)$ for each i, existence and uniqueness was obtained by Hu and Tang [HT10]. While existence is obtained by a penalization argument, uniqueness is obtained via a verification argument, identifying the solution with a formal switching problem involving *switched BSDEs*. Unfortunately, this approach does not extend to the case of a more general driver.
Hamadène and Zhang [HZ10] considered more general oblique constraints and allowed the driver f to depend on the whole y, but they require f to be increasing with respect to each y^i.
In the context of switching-type reflections, Chassagneux, Elie and Kharroubi [CEK12] provide existence and uniqueness of solutions of (2.2.4)-(2.2.5)-(2.2.6) in the case where $f^i(t, y, z) = f^i(t, y, z^i)$.

Lastly, Chassagneux and Richou [CR18] consider a general framework for obliquely reflected BSDEs, studying equations of the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s - \int_t^T H(s, Y_s, Z_s)\Phi_s \mathrm{d}s, \quad t \in [0,T],$$
$$Y_t \in \mathcal{D}, \Phi_t \in \mathcal{C}(Y_t), \quad t \in [0,T],$$
$$\int_0^T \Phi_t 1_{\{Y_t \notin \partial\mathcal{D}\}} \mathrm{d}t = 0,$$

where $\mathcal{D}$ is a convex domain of $\mathbb{R}^d$, and if $y \in \partial\mathcal{D}$, $\mathcal{C}(y)$ is the outward normal cone of $\mathcal{D}$ at y. The function $H : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^{d\times d}$ is *a priori* fixed and model the reflection direction, which may be random and may depend on Z.
In a non-Markovian framework, under some regularity hypothesis on the boundary of the domain, under symmetry and non-degeneracy of H, the authors prove that these equations admit a unique solution. In a Markovian framework, they obtain existence of solutions under weaker hypotheses. The non-degeneracy of H remains a key hypothesis.

2.2.4 Our contributions

In Chapter 6, we consider a new class of switching problems. When the driver f satisfies $f^i(t,y,z) = f^i(t,y^i,z^i)$ for all $i \in \{1,\dots,n\}$, $t \in [0,T]$, $y \in \mathbb{R}^d$ and $z \in \mathbb{R}^{d\times k}$, we prove that, under existence of solutions for an obliquely reflected BSDE, there is uniqueness, and it coincides with the value function of the switching problem. Using the results of [CR18], we provide existence of solutions for these BSDEs in what we call the *uncontrolled irreducible* case, that we describe in details below.

The problem we consider is an optimal control problem of switching type. The novelty is that the agent cannot chose directly the new mode. Instead, he is given probability distributions, and he gets to decide under which the probability measure the new mode will be drawn.

Remark 2.2.1. *This specification leads to uncertainty on costs: if the new mode is drawn according to the distribution $\sum_{i=1}^n p_i\delta_i$, the cost is $c = \sum_{i=1}^n c_i 1_{\{\zeta=i\}}$, where ζ is the new mode. However, in the mathematical analysis, this viewpoint is not emphasized. Indeed, since the agent optimizes his expected reward, only the mean cost $\frac{1}{n}\sum_{i=1}^n p_i c_i$ is relevant for the analysis.*

Remark 2.2.2. *Using Dirac measures on $\{1,\dots,n\}$, one can write the classical switching problem in this setting.*

We work on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$, where $\mathcal{G}$ is generated by a Brownian motion W and an independent family $(X_n)_{n\geq 1}$ of uniformly distributed on $[0,1]$ random variables, independent of W.

We consider an ordered compact metric space $\mathscr{C}$, which is the control space.
To each $u \in \mathscr{C}$ is associated a transition probability function on the state space $\{1,\dots,n\}$ by $p^u_{i,j} := \mathbb{P}(F(u,i,X) = j)$ where X is uniformly distributed on $[0,1]$ and $F : \mathscr{C} \times \{1,\dots,n\} \times [0,1] \to \{1,\dots,n\}$ is measurable. In terms of the game, that means that if the current mode is i and the agent decides to switch using the control u, the new mode will be chosen, independently of everything up to now, according to the probability distribution $\sum_{j=1}^n p^u_{i,j}\delta_j$ on $\{1,\dots,n\}$.

We are also given a cost function $c : \{1, \dots, n\} \times \mathscr{C} \to \mathbb{R}, (i, u) \mapsto c_{i,u}$ which represents the cost to switch from mode i using the control u.

We introduce $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t\geq 0}$ the Brownian filtration.

A strategy $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1})$ for the game is then given by a non-decreasing sequence of random times $(\tau_n)_{n\geq 0}$ and a sequence of $\mathscr{C}$-valued random variables satisfying:

- $\tau_0 \in [0, T]$ and $\zeta_0 \in \{1, \dots, n\}$ are deterministic.
- For all $n \geq 0$, τ_{n+1} is a $\mathbb{F}^n$-stopping time and α_{n+1} is $\mathcal{F}^n_{\tau_{n+1}}$-measurable. The new mode is then $\zeta_{n+1} := F(\alpha_{n+1}, \zeta_n, X_{n+1})$, and we define a new filtration $\mathbb{F}^{n+1} = (\mathcal{F}_t^{n+1})_{t\geq 0}$, with $\mathcal{F}_t^{n+1} = \mathcal{F}_t^n \vee \sigma(X_{n+1} 1_{\{\tau_{n+1}\leq t\}})$, which incorporates the information of the new state.

For a strategy ϕ, we define $\mathbb{F}^\infty = (\mathcal{F}_t^\infty)_{t\geq 0}$ with $\mathcal{F}_t^\infty = \bigvee_{n\geq 0} \mathcal{F}_t^n$.
A strategy is *admissible* if $A_T^\phi - A_{\tau_0}^\phi \in L^2(\mathcal{F}_T^\infty)$ and $\mathbb{E}\Big[\Big(A_{\tau_0}^\phi\Big)^2 | \mathcal{F}_{\tau_0}^0\Big]$ is almost surely finite, where $A_t^\phi = \sum_{n\geq 0} c_{\zeta_n, \alpha_{n+1}} 1_{\{\tau_{n+1}\leq t\}}$ is the cumulative cost process.

Following Hu and Tang [HT10], the agent aims to maximise $\mathbb{E}\Big[U_{\tau_0}^\phi - A_{\tau_0}^\phi | \mathcal{F}_{\tau_0}^0\Big]$, where (U^ϕ, V^ϕ, M^ϕ) is the unique solution on $(\Omega, \mathcal{G}, \mathbb{F}^\infty, \mathbb{P})$, to the following BSDE on $[0, T]$,

$$U_t = \xi^{a_T} + \int_t^T f^{a_s}(s, U_s, V_s)\mathrm{d}s - \int_t^T V_s \mathrm{d}W_s - \int_t^T \mathrm{d}M_s - \int_t^T \mathrm{d}A_s^\phi.$$

An important part is to study the filtration $\mathbb{F}^\infty$. More precisely, we exhibit the structure of the martingales in this filtration, and we show that BSDEs with Lipschitz driver and L^2 terminal condition admit a unique solution. For example, we obtain a representation theorem for L^2 random variables in $\mathcal{F}_T^\infty$, which allows to decompose ξ as the sum of its expectation, a stochastic integral against the Brownian motion, and a jump term. We show that the jumps can only occur at the random times $\tau_i, i \geq 0$, and are necessarily of the form

$$\mathbb{E}\Big[\xi | \mathcal{F}^{k+1}_{\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi | \mathcal{F}^{k}_{\tau_{k+1}}\Big].$$

We refer the reader to Section 6.4.1 for more details on the filtration $\mathbb{F}^\infty$. A careful analysis of the filtrations $\mathbb{F}^k, k \geq 0$ and $\mathbb{F}^\infty$ is needed to obtain the results.

This study shows that the optimisation problem is well-defined. Moreover, in the case of positive costs and assuming the existence of a solution to the following obliquely reflected BSDE:

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s^i, Z_s^i)\mathrm{d}s - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i, \tag{2.2.7}$$

$$Y_t \in \mathcal{D}, \tag{2.2.8}$$

$$\int_0^T \left(Y_t^i - \sup_{u\in\mathscr{C}} \left\{ \sum_{j=1}^d p_{i,j}^u Y_t^j - c_{i,u} \right\} \right) \mathrm{d}K_t^i = 0, \tag{2.2.9}$$

where

$$\mathcal{D} = \left\{ y \in \mathbb{R}^d : y_i \geq \sup_{u \in \mathscr{C}} \left\{ \sum_{j=1}^{d} p^u_{i,j} y_j - c_{i,u} \right\}, i = 1, \ldots, d \right\}$$

has non-empty interior and $\xi \in L^2(\mathcal{F}^\infty_T)$ takes values in $\mathcal{D}$, we are able to identify Y^i_t with the value of the game starting in mode i at time t. In addition, an optimal strategy is given by $\zeta^\star_0 = i, \tau^\star_0 = t$ and

$$\tau^\star_{k+1} = \inf \left\{ \tau^\star_k \leq s \leq T : Y^{\zeta^\star_k}_s = \sup_{u \in \mathscr{C}} \left\{ \sum_{j=1}^{d} p^u_{\zeta^\star_k, j} Y^j_s - c_{\zeta^\star_k, u} \right\} \right\} \wedge (T+1),$$

$$\alpha^\star_{k+1} = \inf \left\{ v \in \mathscr{C}, v \in \arg\max_{u \in \mathscr{C}} \left\{ \sum_{j=1}^{d} p^u_{\zeta^\star_k, j} Y^j_{\tau^\star_k} - c_{\zeta^\star_k, u} \right\} \right\}.$$

The difficulties of the analysis come from the fact that we do not work in the Brownian filtration, as in usual switching problems. Due to the enlargement of filtration, there are new terms (coming from martingales which are orthogonal to the Brownian motion) when dealing with BSDEs. It is then needed to carefully adapt classical BSDEs arguments to take into account these new terms.
In addition, since the agent does not chose directly the new mode, it is possible, even for the optimal strategy, to observe simultaneous switches. It is thus necessary to show that for the optimal strategy, the cumulative cost generated by these simultaneous switches at a the initial time t stays in $L^2(\mathcal{F}^0_t)$. It is possible to show this by studying carefully the time-inhomogeneous Markov chain associated to the optimal strategy in a "corner" of the domain, using the fact that it has non-empty interior.
This two tools allow to show that the optimal strategy is admissible, and that it indeed induces a maximal expected profit.

In the rest of the chapter, we are interested at the existence of the solution to (2.2.7)-(2.2.8)-(2.2.9). We consider the "irreducible" and "uncontrolled" case, that is when the control space is a singleton $\mathscr{C} = \{u\}$. Thus there is only one probability transition represented by a matrix P, and only one cost vector c. The agent only decides when to switch. If the mode is i, he pays c_i to switch, and the new state is determined under $\sum_{j=1}^{n} p_{i,j} \delta_j$. We assume that the matrix P is irreducible.
We consider arbitrary costs, that is we do not suppose that $c_i > 0$ for all i. If that is the case, it is clear that the domain $\mathcal{D}$ has non-empty interior.
We give sufficient and necessary conditions on the costs c to obtain a domain with non-empty interior. This characterization involves a matrix $C = (C^{i,j})_{i,j=1,\ldots,n}$, where $C^{i,j}$ represents the cost to switch "directly" from state i to state j. In practice, $C^{i,j}$ is computed as the mean cost to pay to switch from i to j, which is expressed mathematically as

$$\begin{aligned} C^{i,j} &= \left((I_{n-1} - P^{(i)})^{-1} c^{(i)} \right)_j \text{ if } i \neq j, \\ &= 0 \text{ if } i = j, \end{aligned}$$

where $P^{(i)}$ (resp. $c^{(i)}$) is the matrix (resp. the vector) where we have removed line i and column i (resp. coordinate i), and I_{n-1} is the identity matrix of size $n-1$.

We show, when the domain is non-empty, that each column of the matrix $-C$ is a point of $\mathcal{D}$.
In addition, we prove that the set $\mathcal{D} \cap \{y_n = 0\}$ is a simplex whose extremal points are given by the columns of the matrix $(C^{n,j} - C^{i,j})_{i,j=1,\dots,n}$. We then obtain necessary and sufficient conditions on C for the domain to have non-empty interior.

When the domain has non-empty interior, under a technical assumption of copositivity on the matrix P, we show that there is a solution to (2.2.7)-(2.2.8)-(2.2.9) in a Markovian setting. To do so, we apply the theorem of Chassagneux and Richou [CR18] by constructing a map H sending, at each boundary point, its normal cone into the cone of directions of reflection imposed by the game. This map is first constructed on the extremal points of $\mathcal{D} \cap \{y_n = 0\}$, which is a simplex. After having determined the outward normal cone at each of these points, we naturally obtain the function H at these points. We next extend H to $\mathcal{D} \cap \{y_n = 0\}$ by convex combinations, and to $\mathcal{D}$ using the invariance of the domain by translations by $(1, \dots, 1)$. Lastly, we obtain H on $\mathbb{R}^d$ by projection onto $\mathcal{D}$.

The content of this chapter is from a work in progress, in collaboration with Jean-François Chassagneux and Adrien Richou.

Acknowledgements

This thesis was funded in the scope of the research project “Advanced techniques for non-linear pricing and risk management of derivatives” under the aegis of the Europlace Institute of Finance, with the support of AXA Research Fund.

The author would also like to thank the organizers of the 2017 CEMRACS for the opportunity to study at CIRM, and to thank the all funding sources that supported us during CEMRACS, in particular University Paris Diderot. The chapter 5 of this thesis was written in this context.

Part I

Pricing with controlled loss

Chapter 3

Quantile hedging in a linear market

3.1 Introduction

In this chapter, we consider the quantile hedging problem [FL99] in a linear market, and we derive a characterisation of the price as the super-replication price of a modified pay-off. In the second part, we consider the control formulation of [BET09] and we obtain, in a Black & Scholes market, a convergence rate for the control problem with bounded controls.

The quantile hedging problem for an European contingent claim was popularised by [FL99], who studied some special cases and provided a closed form solution in some settings. A key ingredient in their analysis is the Neyman-Pearson lemma from mathematical statistics. In the first part of this chapter, we show that it is not needed to use this lemma as we provide new proofs for their results.

In [BET09], the quantile hedging was then viewed as a special case of a broader class of approximate hedging problems, which are formulated as stochastic target control problem. This point of view allows to obtain a PDE characterisation in an incomplete market with portfolio constraints. This work was extended in various settings, Moreau [Mor11] introduced jump dynamics, Bouchard, Bouveret and Chassagneux [BBC16] studied the Bermudan case and Dumitrescu, Elie, Sabbagh and Zhou [Dum+17] the American case.

These works are of a theoretical nature, except for [BBC16]. In fact, the control space of the stochastic target problem from [BET09] being unbounded, the study of the numerical approximation is difficult. Similarly, the PDE characterizing the quantile hedging price has a discontinuous operator, making it difficult to discretise. A comparison theorem for the PDE was obtained by Bouveret and Chassagneux [BC17], and we refer to Chapter 4 where we introduce a numerical scheme to approximate the quantile hedging price. However, we do not obtain convergence rate for the PDE discretization. In the Black & Scholes model, we show in this chapter that it is possible to obtain some rate of convergence when we truncate the control space, using duality.

The rest of this chapter is organised as follows. First, we characterise the quantile hedging price as a super-replication price of a modified pay-off, to be determined. If moreover the market is linear, we construct explicitly this pay-off, using elementary probabilistic arguments. This allows, as an application, to derive closed formulae for

the quantile hedging price of vanilla options in the Black & Scholes model, which are known [FL99].
In the second part of this chapter, we study the problem from a control point of view, using the formulation of [BET09], when the stock price is driven by a SDE. It is well-known that the control space associated to this stochastic target problem is the whole real line $\mathbb{R}$, hence its numerical approximation is difficult. The goal here is to consider the control problem where the admissible controls are valued in a compact subset $[-n, n]$, and to obtain a convergence rate when we let n go to infinity. The value function of this control problem is convex in the p variable, hence we consider its Fenchel-Legendre transform. In a linear market, this allows to derive a ϵ-optimal control, hence we obtain an approximation of the value function. The form of the ϵ-optimal control allows to obtain an upper bound for the difference between the original problem and the one with bounded controls.

Notations If $n \geq 1$, we let $\mathcal{B}^n$ be the Borelian sigma-algebra on $\mathbb{R}^n$.
For any filtered probability space $(\Omega, \mathcal{G}, \mathbb{F}, \mathbb{P})$ and constants $T > 0$ and $p \geq 1$, we define the following spaces:

- $L^p(\mathcal{G})$ is the set of real $\mathcal{G}$-measurable random variables X satisfying $\mathbb{E}[|X|^p] < +\infty$,
- $\mathcal{P}(\mathbb{F})$ is the predictable sigma-algebra on $\Omega \times [0, T]$,
- $\mathbb{H}^p_n(\mathbb{F})$ is the set of predictable processes ϕ valued in $\mathbb{R}^n$ such that

$$\mathbb{E}\left[\int_0^T |\phi_t|^p \mathrm{d}t\right] < +\infty,$$

- $\mathbb{S}^p_n(\mathbb{F})$ is the set of predictable processes ϕ valued in $\mathbb{R}^n$ such that

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\phi_t|^p\right] < +\infty,$$

If $n = 1$, we omit the subscript n in previous notations.

3.2 Resolution of the quantile hedging problem in a linear market

In this section, we show that the quantile hedging price is the minimum price over some modified pay-off's. When the market is linear, we show that there is a modified pay-off which attains this minimum. This allows to obtain the formulae of [FL99] in the Black & Scholes model for vanilla options.

3.2.1 The quantile hedging price

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We assume that it is endowed with a d-dimensional Brownian motion $W = (W_t)_{t \geq 0}$, that $\mathbb{F}$ is its natural augmented filtration and that $\mathcal{F} = \mathcal{F}_T$ for some finite terminal time $T > 0$.

We consider a $\mathcal{F}_T$-random variable $0 \leq \xi \in L^2(\mathcal{F}_T)$.
Let $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{P}(\mathbb{F}) \otimes \mathcal{B} \otimes \mathcal{B}^d$-measurable function such that $f(\cdot,\cdot,0,0) = 0, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.s., and there exists $C > 0$ such that, for all $(y,y') \in \mathbb{R}^2$ and $(z,z') \in (\mathbb{R}^d)^2$,

$$|f(\omega,t,y,z) - f(\omega,t,y',z')| \leq C\left(|y-y'| + |z-z'|\right), \mathrm{d}\mathbb{P} \otimes \mathrm{d}t\text{-a.s..}$$

An admissible strategy is a couple (y,ν), where $y \geq 0$ and $\nu \in \mathbb{H}^2(\mathbb{F})$ is a $\mathbb{R}^d$-valued process such that the real-valued process $Y^{y,\nu}$, solution to the following SDE:

$$Y_t = y - \int_0^t f(s,Y_s,\nu_s)\mathrm{d}s + \int_0^t \nu_s \mathrm{d}W_s, t \in [0,T],$$

stays non-negative $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t-$a.s.. We denote by $\mathcal{A}$ the set of admissible strategies.
For $p \in [0,1]$, we consider the following stochastic target problem:

$$V^p = \inf\{y \geq 0 : \exists \nu, (y,\nu) \in \mathcal{A} \text{ and } \mathbb{P}(Y_T^{y,\nu} \geq \xi) \geq p\}. \tag{3.2.1}$$

The case $p = 1$ is well-known and was extendedly studied before. The next proposition shows that V^1 is obtained by solving a BSDE:

Proposition 3.2.1. *Let $(Y,Z) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{H}^2(\mathbb{F})$ be the solution to the following BSDE:*

$$Y_t = \xi + \int_t^T f(s,Y_s,Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s,\ t \in [0,T].$$

Then $(Y_0,Z) \in \mathcal{A}$ and $V^1 = Y_0$.

Proof. Since $f(\cdot,\cdot,0,0) = 0, \mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.s., it is straightforward to see that $(Y_t^0 = 0, Z_t^0 = 0)$ is the unique solution to the BSDE with terminal condition 0 and driver f. By the comparison theorem [EKPQ97], since $\xi \geq 0$, we obtain $Y_t \geq 0$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.s., which amounts to say that $(Y_0,Z) \in \mathcal{A}$.
Since $Y_T = \xi$, we get $Y_0 \geq V^1$ by definition.
Let $(y,\nu) \in \mathcal{A}$ such that $Y_T^{y,\nu} \geq \xi = Y_T, \mathbb{P}$-a.s.. The comparison theorem ensures that $Y_0 \leq Y_0^{y,\nu} = y$. Thus $Y_0 \leq V^1$ and the proof is complete. □

The previous proposition generalizes easily when $p < 1$. However, we do not get a BSDE representation for V^p, but only a characterisation as an infimum of BSDEs.

Proposition 3.2.2. *For $A \in \mathcal{F}_T$, we define $(Y^A,Z^A) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{H}^2(\mathbb{F})$ as the solution of the following BSDE:*

$$Y_t = \xi 1_A + \int_t^T f(s,Y_s,Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s,\ t \in [0,T].$$

Then $(Y_0^A, Z^A) \in \mathcal{A}$ and $V^p = \inf_{A\in\mathcal{F}_T^p} Y_0^A$, where $\mathcal{F}_T^p = \{A \in \mathcal{F}_T : \mathbb{P}(A) \geq p\}$.

Proof. Let $A \in \mathcal{F}_T$. By the comparison theorem, since $\xi 1_A \geq 0$ we get that $Y_t^A \geq 0, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.s., and thus $(Y_0^A, Z^A) \in \mathcal{A}$.
If moreover $A \in \mathcal{F}_T^p$, since $Y_T^A = \xi 1_A$, we have $A \subset \{Y_T^A \geq \xi\}$. This implies $\mathbb{P}(Y_T^A \geq \xi) \geq \mathbb{P}(A) \geq p$ and $V^p \leq Y_0^A$. Since A is arbitrary in $\mathcal{F}_T^p$, we get $V^p \leq \inf_{a\in\mathcal{F}_T^p} Y^a A_0$.
Conversely, let $(y,\nu) \in \mathcal{A}$ such that $\mathbb{P}(Y_T^{y,\nu} \geq \xi) \geq p$. Then set $A := \{Y_T^{y,\nu} \geq \xi\} \in \mathcal{F}_T^p$ and consider the associated processes $(Y^A,Z^A) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{H}^2(\mathbb{F})$. We have on A: $Y_T^A = \xi \leq Y_T^{y,\nu}$ by definition of A, and on A^c: $Y_T^A = 0 \leq Y_T^{y,\nu}$ since $(y,\nu) \in \mathcal{A}$. Thus $Y_T^A \leq Y_T^{y,\nu} \mathbb{P}$-a.s., and the comparison theorem gives $Y_0^A \leq Y_0^{y,\nu} = y$. Since (y,ν) is arbitrary in $\mathcal{A}$, we get $\inf_{A\in\mathcal{F}_T^p} Y_0^A \leq V^p$. □

3.2.2 An explicit solution

In this section, we are going to give a BSDE representation for $V^p, p < 1$, when the driver f is a linear function of (y, z).
More precisely, we consider the following hypothesis:

Assumption 3.2.1. *There exists a bounded* $(\mathbb{R} \times \mathbb{R}^d)$*-valued predictable process* (a, b) *such that:*

$$f(\omega, t, y, z) = a_t(\omega)y + b_t(\omega)^\top z.$$

In this linear setting, for every terminal condition $Y_T \in L^2(\mathcal{F}_T)$, consider the solution $(Y_t, Z_t) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{H}^2(\mathbb{F})$ of the following BSDE:

$$Y_t = Y_T + \int_t^T \left(a_s Y_s + b_s^\top Z_s\right) \mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, t \in [0, T].$$

Then, we have the following representation for Y [EKPQ97]:

$$Y_t = \Gamma_t^{-1} \mathbb{E}[\Gamma_T Y_T] \, ,$$

where Γ is the unique solution of the following SDE:

$$\Gamma_t = 1 + \int_0^t \Gamma_s a_s \mathrm{d}s + \int_0^t \Gamma_s b_s^\top \mathrm{d}W_s, t \in [0, T]. \tag{3.2.2}$$

In particular, in this setting, Proposition 3.2.2 rewrites:

Proposition 3.2.3. *For* $p \in [0, 1]$*, we have:*

$$V^p = \inf_{A \in \mathcal{F}_T^p} \mathbb{E}[\Gamma_T \xi 1_A]$$

Lastly, we introduce the technical assumption used in the proof of the theorem. For any $p \in [0, 1]$, let $q(p) = \inf\{q \geq 0 : \mathbb{P}(\Gamma_T \xi \leq q) \geq p\}$ the p-quantile of the law of $\Gamma_T \xi$.

Assumption 3.2.2. *For any* $p \in [0, 1]$*, there exists a set* $A \in \mathcal{F}_T$ *such that:*

i. $\mathbb{P}(A) = p$,

ii. $\{\Gamma_T \xi < q(p)\} \subset A \subset \{\Gamma_T \xi \leq q(p)\}$.

This Assumption is satisfied in numerous practical examples, in particular when the random variable $\Gamma_T \xi$ is absolutely continuous with respect to the Lebesgue measure.
The next results gives a BSDE characterisation for V^p, for all $p \in [0, 1]$.

Theorem 3.2.1. *For* $p \in [0, 1]$*, let* $A^\star$ *be a set satisfying the hypotheses of Assumption 3.2.2. Then we have:*

$$V^p = Y_0^{A^\star} = \mathbb{E}[\Gamma_T \xi 1_{A^\star}] \, .$$

Proof. Let $p \in [0,1]$. In view of Proposition 3.2.2 and Proposition 3.2.3, we need to prove that $\mathbb{E}[\Gamma_T \xi 1_{A^\star}] = \inf_{A \in \mathcal{F}_T^p} \mathbb{E}[\Gamma_T \xi 1_A]$.
Let $A \in \mathcal{F}_T^p$. To prove $\mathbb{E}[\Gamma_T \xi 1_{A^\star}] \leq \mathbb{E}[\Gamma_T \xi 1_A]$, since we have

$$A = (A \cap A^\star) \cup (A \cap (A^\star)^c), \text{ and} \tag{3.2.3}$$
$$A^\star = (A^\star \cap A) \cup (A^\star \cap A^c), \tag{3.2.4}$$

it is enough to prove

$$\mathbb{E}[\Gamma_T \xi 1_{A^\star \cap A^c}] \leq \mathbb{E}\big[\Gamma_T \xi 1_{A \cap (A^\star)^c}\big] .$$

We have:

$$A^\star \cap A^c \subset A^\star \subset \{\Gamma_T \xi \leq q(p)\}, \text{ and}$$
$$A \cap (A^\star)^c \subset (A^\star)^c \subset \{\Gamma_T \xi < q(p)\}^c = \{\Gamma_T \xi \geq q(p)\}.$$

Thus:

$$\mathbb{E}[\Gamma_T \xi 1_{A^\star \cap A^c}] \leq q(p)\mathbb{E}[1_{A^\star \cap A^c}] = q(p)\mathbb{P}(A^\star \cap A^c), \text{ and} \tag{3.2.5}$$
$$\mathbb{E}\big[\Gamma_T \xi 1_{A \cap (A^\star)^c}\big] \geq q(p)\mathbb{E}\big[1_{A \cap (A^\star)^c}\big] = q(p)\mathbb{P}(A \cap (A^\star)^c). \tag{3.2.6}$$

By Assumption 3.2.2, (3.2.3)-(3.2.4) and since $A \in \mathcal{F}_T^p$, we have:

$$p = \mathbb{P}(A^\star) = \mathbb{P}(A^\star \cap A) + \mathbb{P}(A^\star \cap A^c), \text{ and}$$
$$p \leq \mathbb{P}(A) = \mathbb{P}(A \cap A^\star) + \mathbb{P}(A \cap (A^\star)^c).$$

From these two equations, we get:

$$\mathbb{P}(A \cap (A^\star)^c) \geq p - \mathbb{P}(A^\star \cap A) = \mathbb{P}(A^\star \cap A). \tag{3.2.7}$$

Using the fact that $q(p) \geq 0$ and (3.2.5)-(3.2.6) with (3.2.7), we get:

$$\mathbb{E}[\Gamma_T \xi 1_{A^\star \cap A^c}] \leq q(p)\mathbb{P}(A^\star \cap A^c) \leq q(p)\mathbb{P}(A \cap (A^\star)^c) \leq \mathbb{E}\big[\Gamma_T \xi 1_{A \cap (A^\star)^c}\big] ,$$

which ends the proof of the theorem. □

3.2.3 Application: quantile hedging in the Black & Scholes model for a put option

As an application for the previous theorem, we show that the closed formulae for $V^p, p \in [0,1]$, in the Black & Scholes model, derived in [FL99], are consequences of Theorem 3.2.1, and can be obtained without invoking the Neyman-Pearson Lemma.
We assume here that the random variable ξ is given by

$$\xi = (K - X_T)^+,$$

where $K > 0$ and X_T is the value at time T of the geometric Brownian motion:

$$X_t = x \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right), \quad t \in [0,T],$$

with $\mu \in \mathbb{R}$, $\sigma > 0$ and $x > 0$.
In this linear setting, the function f is given by:

$$f(t, y, z) = -ry - \frac{\mu - r}{\sigma} z =: -ry - \lambda z,$$

where $r \geq 0$ is the interest rate, and $\lambda := \frac{\mu - r}{\sigma}$ is the *risk premium*.
The process Γ defined in (3.2.2) is thus given by:

$$\Gamma_t = \exp\left(-\left(r + \frac{\lambda^2}{2}\right)t - \lambda W_t\right), \quad t \in [0, T].$$

By Theorem 3.2.1, for $p \in [0, 1]$, we have:

$$V^p = \inf_{A \in \mathcal{F}_T^p} \mathbb{E}[\Gamma_T g(X_T) 1_A] = \mathbb{E}[\Gamma_T g(X_T) 1_{A^\star}],$$

where $A^\star$ is a set satisfying Assumption 3.2.2.
To find V^p, following 3.2.2, it is enough to study the law of

$$\begin{aligned}
H &:= \Gamma_T g(X_T) \\
&= \exp\left(-\left(r + \frac{\lambda^2}{2}\right)T - \lambda W_T\right) \cdot \left(K - x\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma W_T\right)\right)^+ \\
&=^d \exp\left(-\left(r + \frac{\lambda^2}{2}\right)T - \lambda\sqrt{T}N\right) \cdot \left(K - x\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}N\right)\right)^+ \\
&=: \tilde{H}(N),
\end{aligned}$$

where N is normally distributed and $\tilde{H}$ is defined by:

$$\begin{aligned}
&\tilde{H} : \mathbb{R} \to [0, \infty), \\
&\quad y \mapsto \exp\left(-\left(r + \frac{\lambda^2}{2}\right)T - \lambda\sqrt{T}y\right) \cdot \left(K - x\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}y\right)\right)^+.
\end{aligned}$$

First, observe that:

$$\begin{aligned}
p^0 := \mathbb{P}(H = 0) &= \mathbb{P}\left(x\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}N\right) \geq K\right) \\
&= \mathbb{P}\left(N \geq \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{K}{x}\right) - \left(\mu - \frac{\sigma^2}{2}\right)T\right)\right) \\
&= \Phi\left(N_{max}\right),
\end{aligned}$$

with $N_{max} := \frac{1}{\sigma\sqrt{T}}\left(\left(\mu - \frac{\sigma^2}{2}\right)T - \ln\left(\frac{K}{x}\right)\right)$ and Φ the cumulative distribution function of the standard normal law.
Thus, if $p \leq p^0$, taking the set $A = \{H = 0\}$ which has probability greater or equal than p gives:

$$0 \leq V^p \leq \mathbb{E}[H 1_A] = 0,$$

and $V^p = 0$.

3.2.3.1 Case 1: $\lambda \geq 0$.

The function $\tilde{H}$ is non-increasing on $\mathbb{R}$ as the product of two non-negative non-increasing functions, and it is a bijection from $E := (-\infty, -N_{max}]$ to $[0, +\infty)$. Thus there exists a non-increasing mapping $l : [0, +\infty) \to E$ such that $\tilde{H}(l(q)) = q$ and $l(\tilde{H}(y)) = y$ for $q \geq 0$ and $y \in E$. In particular, it satisfies:

$$\begin{aligned} \{H \leq q\} &=^d \{N \geq l(q)\}, \\ \{N \geq l\} &=^d \{H \leq \tilde{H}(l)\}, \\ l(0) &= \frac{1}{\sigma\sqrt{T}}\left(\left(\mu - \frac{\sigma^2}{2}\right)T - \frac{K}{x}\right). \end{aligned}$$

Thus, for any $p \in [p^0, 1]$:

$$\begin{aligned} q(p) &= \inf\{q \geq 0 : \mathbb{P}(H \leq q) \geq p\} \\ &= \inf\{q \geq 0 : \mathbb{P}(N \geq l(q)) \geq p\} \\ &= \tilde{H}(\sup\{l : \mathbb{P}(N \geq l) \geq p\}) \\ &= \tilde{H}(-\inf\{q : \mathbb{P}(N \leq q) \geq p\}) \\ &= \tilde{H}(-\Phi^{-1}(p)), \end{aligned}$$

where $\Phi^{-1}(p)$ is the p-quantile for the standard normal law.
Since $\mathbb{P}(H \leq q(p)) = \mathbb{P}(H \leq \tilde{H}(\Phi^{-1}(p))) = \mathbb{P}(N \geq l(\tilde{H}(-\Phi^{-1}(p))) = \mathbb{P}(N \geq -\Phi^{-1}(p)) = \mathbb{P}(N \leq \Phi^{-1}(p)) = p$, $\{H \leq q(p)\}$ is optimal for V^p by Assumption 3.2.2 and Theorem 3.2.1.

In particular, one can compute V^p for $p \in [p^0, 1]$:

$$\begin{aligned} V^p &= \mathbb{E}\big[H 1_{H \leq q(p)}\big] \\ &= \mathbb{E}\Big[\tilde{H}(N) 1_{N \geq -\Phi^{-1}(p)}\Big] \\ &= \int_{-\Phi^{-1}(p)}^{+\infty} \frac{\exp\left(-\frac{1}{2}y^2\right)}{\sqrt{2\pi}} \tilde{H}(y) \\ &= \int_{-\Phi^{-1}(p)}^{-N_{max}} \frac{\exp\left(-\frac{1}{2}y^2\right)}{\sqrt{2\pi}} \tilde{H}(y), \end{aligned}$$

since $H = 0$ on $[-N_{max}, +\infty)$. Thus:

$$\begin{aligned}
V^p &= \frac{K \exp(-rT)}{\sqrt{2\pi}} \int_{-\Phi^{-1}(p)}^{-N_{max}} \exp\left(-\frac{1}{2}y^2 - \lambda\sqrt{T}y - \frac{\lambda^2}{2}T\right) \mathrm{d}y \\
&- \frac{x}{\sqrt{2\pi}} \int_{-\Phi^{-1}(p)}^{-N_{max}} \exp\left(-\frac{1}{2}y^2 - \left(r + \frac{\lambda^2}{2}\right)T - \lambda\sqrt{T}y + \left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}y\right) \mathrm{d}y \\
&= \frac{K \exp(-rT)}{\sqrt{2\pi}} \int_{-\Phi^{-1}(p)}^{-N_{max}} \exp\left(-\frac{1}{2}\left(y + \lambda\sqrt{T}\right)^2\right) \mathrm{d}y \\
&- \frac{x}{\sqrt{2\pi}} \int_{-\Phi^{-1}(p)}^{-N_{max}} \exp\left(-\frac{1}{2}\left(y - \sqrt{T}(\sigma - \lambda)\right)^2\right) \mathrm{d}y \\
&= K \exp(-rT) \left[\Phi\left(\lambda\sqrt{T} - N_{max}\right) - \Phi\left(\lambda\sqrt{T} - \Phi^{-1}(p)\right)\right] \\
&- x\left[\Phi\left(-N_{max} - \sqrt{T}(\sigma - \lambda)\right) - \Phi\left(-\Phi^{-1}(p) - \sqrt{T}(\sigma - \lambda)\right)\right].
\end{aligned}$$

3.2.3.2 Case 2: $\lambda < 0$

Studying the function $\tilde{H}$ on $(-\infty, -N_{max})$ shows that it increases from $(-\infty, y^\star]$ to $(0, \tilde{H}(y^\star)]$, and then decreases from $(y^\star, -N_{max})$ to $(0, \tilde{H}(y^\star))$, with

$$y^\star = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{\lambda K}{(\lambda - \sigma)x}\right) - \left(\mu - \frac{\sigma^2}{2}\right)T\right).$$

Therefore, there exists an increasing function $l_1 : (0, \tilde{H}(y^\star)] \to (-\infty, y^\star]$ and a decreasing function $l_2 : (0, \tilde{H}(y^\star)) \to (y^\star, -N_{max})$ such that, for all $q \in (0, y^\star)$ and $l_1, l_2 \in \mathbb{R}$:

$$\begin{aligned}
l_1(\tilde{H}(q)) &= q \text{ if } q \leq y^\star, \\
l_2(\tilde{H}(q)) &= q \text{ if } q \geq y^\star, \\
l_1(\tilde{H}(q)) &< l_2(\tilde{H}(q)) \text{ iif } q \neq y^\star.
\end{aligned}$$

Moreover, since $\tilde{H}$ is monotonic and differentiable on each interval, l_1 and l_2 are differentiable.
In particular, $\mathbb{P}(H \leq \cdot)$ is a differentiable function on $(0, \tilde{H}(y^\star))$:

$$\begin{aligned}
\mathbb{P}(H \leq q) &= \mathbb{P}(N \leq l_1(q)) + \mathbb{P}(N \geq l_2(q)) \\
&= 1 - \Phi(l_2(q)) + \Phi(l_1(q))
\end{aligned}$$

This means that for any $p \in [p_0, 1]$, the p-quantile $q(p)$ of H satisfies $\mathbb{P}(H \leq q(p)) = p$. In particular, by Theorem 3.2.1, an optimal set for $V^p, p \in [p_0, 1]$ is $\{H \leq q(p)\} = \{N \leq l_1(q(p))\} \cup \{N \geq l_2(q(p))\}$.

3.3 A control point of view: truncation of the control space

In this section, we consider that the pay-off ξ is of the form $\xi = g(X_T)$, where X is the price process for the underlying asset and $g : \mathbb{R} \to \mathbb{R}_+$ is a L-Lipschitz function. We

suppose here that X is a geometric Brownian motion, and we consider the dynamic version for the quantile hedging problem.
More precisely, for all $t \in [0,T]$ and $x \in \mathbb{R}$, $X^{t,x}$ is the solution to the following SDE:

$$X_s = y + \int_t^s \mu X_u \mathrm{d}u + \int_t^s \sigma X_u \mathrm{d}W_u, \quad s \in [t,T],$$

where W is a 1-dimensional Brownian motion.
In a linear market with zero interest-rate, given a strategy (y,ν) where y is the initial wealth and $\frac{\nu}{\sigma}$ is the wealth invested in the asset, the portfolio's value is

$$Y_s^{t,y,\nu} = y + \int_t^s \lambda \nu_u \mathrm{d}u + \int_t^s \nu_u \mathrm{d}W_u, \quad s \in [t,T],$$

where $\lambda := \frac{\mu}{\sigma}$ is the *risk-premium.*

Let $p \in [0,1]$. The problem (3.2.1) writes, in this dynamic setting,

$$v(t,x,p) = \inf \left\{ y \geq 0 | \exists \nu \in \mathcal{A}, \mathbb{P}(Y^{t,y,\nu} \geq g(X_T^{t,x})) \geq p \right\}.$$

For $t \in [0,T]$, $p \in [0,1]$ and $\alpha \in \mathbb{H}^2(\mathbb{F})$, we let

$$P_s^{t,p,\alpha} = p + \int_t^s \alpha_u \mathrm{d}W_u.$$

We also set

$$\mathcal{A}_p := \left\{ \alpha \in \mathbb{H}^2(\mathbb{F}) | P^{t,p,\alpha} \in [0,1] \mathrm{d}t \otimes \mathrm{d}\mathbb{P} - \mathrm{a.s.} \right\}.$$

We then have the following problem reformulation

Proposition 3.3.1 ([BET09; BBC16]). *For $t \in [0,T]$, $x > 0$, $p \in [0,1]$, we have*

$$\begin{aligned} v(t,x,p) &= \inf \left\{ y \geq 0 | \exists \nu \in \mathcal{A}, \exists \alpha \in \mathcal{A}_p, 1_{\{Y_T^{t,y,\alpha} \geq g(X_T^{t,x})\}} \geq P_T^{t,p,\alpha} \mathbb{P} - a.s. \right\} \\ &= \inf_{\alpha \in \mathcal{A}_p} \mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x}) 1_{\{P_T^{t,p,\alpha} \geq 0\}} \right] = \inf_{\alpha \in \mathcal{A}_p} \mathbb{E}^{\mathbb{Q}} \left[g(X_T^{t,x}) P_T^{t,p,\alpha} \right], \end{aligned}$$

where $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \Gamma_T^t$ and

$$\Gamma_s^t = 1 - \int_t^s \lambda \mathrm{d}W_u, \quad s \in [t,T].$$

3.3.1 An (almost) optimal control using duality

It is well-known (see Remark 2.3 in [BBC16]) that the function v is convex in its third variable. Moreover, we have the following result.

Proposition 3.3.2 ([BER15; BBC16]). *For $t \in [0,T]$, $x > 0$ and $q \in \mathbb{R}$, let $w(t,x,q) = \max_{p \in [0,1]} \{pq - v(t,x,p)\}$ the Fenchel-Legendre transform of v with respect to p. Then*

$$w(t,x,q) = \mathbb{E}^{\mathbb{Q}} \left[(qQ_T^t - g(X_T^{t,x}))_+ \right] = \mathbb{E}^{\mathbb{Q}} \left[g^\sharp(X_T^{t,x}, qQ_T^t) \right],$$

with, for $(x,q) \in \mathbb{R}^2$,

$$g^\sharp(x,q) = (q - g(x))_+ = \max(q - g(x), 0).$$

To construct an (almost) optimal control, we introduce a mollified version of w, as follows.
We fix $\epsilon > 0$. For all $t \in [0,T]$, $x > 0$, $q > 0$, we set

$$g^\sharp_\epsilon(t,x,q) := \left(\rho_\epsilon * g^\sharp\right)(t,x,q) = \int_{[-\epsilon,\epsilon]^2} \rho_\epsilon(y,r) g^\sharp(t, x-y, q-r)\mathrm{d}y\mathrm{d}r,$$

where $\rho_\epsilon(x,q) = \epsilon^{-2}\rho(x/\epsilon, q/\epsilon)$ with $\rho : \mathbb{R}^2 \to \mathbb{R}$ is a positive mollifier, that is a smooth function supported on $[-1,1]^2$ satisfying $\int_{\mathbb{R}^2}\rho = 1$, and $*$ is the convolution operator. We define a smooth function by $w_\epsilon(t,x,q) = \mathbb{E}^{\mathbb{Q}}\left[g^\sharp_\epsilon(X^{t,x}_T, qQ^t_T)\right]$ for $t \in [0,T]$, $x \in (0,\infty)$ and $q > 0$.

First, we give some easy but useful properties of w_ϵ.

Lemma 3.3.1. *We have $|w_\epsilon - w| \leq (1+L)\epsilon$. In addition, w_ϵ is convex in its third variable.*

Proof. We have

$$\begin{aligned} |w_\epsilon(t,x,q) - w(t,x,q)| &= |\mathbb{E}^{\mathbb{Q}}\left[g^\sharp_\epsilon(X^{t,x}_T, qQ^t_T) - g^\sharp(X^{t,x}_T, qQ^t_T)\right]| \\ &\leq \mathbb{E}^{\mathbb{Q}}\left[|g^\sharp_\epsilon(X^{t,x}_T, qQ^t_T) - g^\sharp(X^{t,x}_T, qQ^t_T)|\right], \end{aligned}$$

so it is enough to obtain the inequality for $g^\sharp_\epsilon$ and $g^\sharp$. For $x \in (0,\infty)$ and $q > 0$, we have,

$$\begin{aligned} |g^\sharp_\epsilon(x,q) - g^\sharp(x,q)| &= |\int \rho_\epsilon(y,r) g^\sharp(x-y, q-r)\mathrm{d}y\mathrm{d}r - g^\sharp(x,q)| \\ &= |\int \rho_\epsilon(y,r)\left(g^\sharp(x-y,q-r) - g^\sharp(x,q)\right)\mathrm{d}y\mathrm{d}r| \\ &\leq \int \rho_\epsilon(y,r)|g^\sharp(x-y,q-r) - g^\sharp(x,q)|\mathrm{d}y\mathrm{d}r \\ &\leq (1+L)\epsilon. \end{aligned}$$

To show that w_ϵ is convex in q, notice first that $g^\sharp$ is convex in q as $g^\sharp(x,q) = (q - g(x))_+$ for all $(x,q) \in (0,\infty) \times \mathbb{R}_+$.
Now, observe that, for all $x \in (0,\infty), \lambda \in [0,1], (q_1,q_2) \in \mathbb{R}^2_+$, since $\rho_\epsilon \geq 0$,

$$\begin{aligned} g^\sharp_\epsilon(x, \lambda q_1 + (1-\lambda)q_2) &= \int \rho_\epsilon(y,r) g^\sharp(x-y, \lambda q_1 + (1-\lambda)q_2 - r)\mathrm{d}y\mathrm{d}r \\ &= \int \rho_\epsilon(y,r) g^\sharp(x-y, \lambda(q_1 - r) + (1-\lambda)(q_2 - r))\mathrm{d}y\mathrm{d}r \\ &\leq \lambda \int \rho_\epsilon(y,r) g^\sharp(x-y, q_1 - r)\mathrm{d}y\mathrm{d}r \\ &\qquad + (1-\lambda)\int \rho_\epsilon(y,r) g^\sharp(x-y, q_2 - r)\mathrm{d}y\mathrm{d}r \\ &\leq \lambda g^\sharp_\epsilon(x,q_1) + (1-\lambda) g^\sharp_\epsilon(x,q_2). \end{aligned}$$

Last, if $t \in [0,T], x \in (0,\infty), \lambda \in [0,1], (q_1,q_2) \in (0,\infty)^2$, we obtain

$$\begin{aligned} w_\epsilon(t,x,\lambda q_1 + (1-\lambda)q_2) &= \mathbb{E}^{\mathbb{Q}}\left[g_\epsilon^\sharp(X_T^{t,x}, (\lambda q_1 + (1-\lambda)q_2)Q_T^t)\right] \\ &\leq \lambda \mathbb{E}^{\mathbb{Q}}\left[g_\epsilon^\sharp(X_T^{t,x}, q_1 Q_T^t)\right] + (1-\lambda)\mathbb{E}^{\mathbb{Q}}\left[g_\epsilon^\sharp(X_T^{t,x}, q_2 Q_T^t)\right] \\ &= \lambda w_\epsilon(t,x,q_1) + (1-\lambda) w_\epsilon(t,x,q_2). \end{aligned}$$

□

We now use the function w_ϵ to construct a martingale P, which will induce an (almost) optimal control.

Proposition 3.3.3. *If $u_\epsilon := \partial_q w_\epsilon$, then $u_\epsilon \in [0,1]$. In addition, the process P defined by $P_s = u_\epsilon(s, X_s^{t,x}, qQ_s^t)$, $s \in [t,T]$, is a martingale.*

Proof. Using the dominated convergence theorem, we obtain

$$u_\epsilon(t,x,q) := \mathbb{E}\Big[\partial_q g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t)\Big] .$$

To check that u_ϵ is valued in $[0,1]$, it is thus enough to check that $\partial_q g_\epsilon^\sharp$ is. We have, for $t \in [0,T]$, $x \in (0,\infty)$, $q_1, q_2 > 0$, since $g^\sharp(t,x,\cdot)$ is 1-Lipschitz,

$$\begin{aligned} |g_\epsilon^\sharp(t,x,q_1) - g_\epsilon^\sharp(t,x,q_2)| &\leq \int \rho_\epsilon(y,r)|g^\sharp(t,x-y,q_1-r) - g^\sharp(t,x-y,q_2-r)|\mathrm{d}y\mathrm{d}r \\ &\leq |q_1 - q_2|. \end{aligned}$$

In addition, we have for $t \in [0,T]$, $x \in (0,\infty)$ and $0 < q_1 \leq q_2$, since $g^\sharp(t,x,\cdot)$ is non-decreasing,

$$g_\epsilon^\sharp(t,x,q_2) - g_\epsilon^\sharp(t,x,q_1) = \int \rho_\epsilon(y,r)\left(g^\sharp(t,x-y,q_2-r) - g^\sharp(t,x-y,q_1-r)\right)\mathrm{d}y\mathrm{d}r \geq 0.$$

Since $g_\epsilon^\sharp(t,x,\cdot)$ is 1-Lipschitz and non-decreasing, we obtain $\partial_q g_\epsilon^\sharp(t,x,q) \in [0,1]$. □

Theorem 3.3.1. *The process $(P_s)_{s\in[t,T]}$ induces a $2(1+L)\epsilon-$optimal control.*

Proof. We need to prove that $\mathbb{E}^{\mathbb{Q}}\left[P_T g(X_T^{t,x})\right] \leq v(t,x,p) + 2(1+L)\epsilon$.
Set $g_\epsilon(x,p) = \sup_{q>0}\{pq - g_\epsilon^\sharp(x,q)\}$.
First, notice that if $p = \partial_q g_\epsilon^\sharp(x,q)$, we have, for all $r > 0$, since $g_\epsilon^\sharp$ is convex in its last variable,

$$g_\epsilon^\sharp(x,r) \geq g_\epsilon^\sharp(x,q) + (r-q)p,$$

which writes

$$pq - g_\epsilon^\sharp(x,q) \geq pr - g_\epsilon^\sharp(x,r).$$

Since r is arbitrary, this gives $g_\epsilon(x,p) = pq - g_\epsilon^\sharp(x,q)$.
We have, for $x \in (0,\infty)$ and $p \in [0,1]$, using Lemma 3.3.1,

$$\begin{aligned} g_\epsilon(x,p) &= \sup_{q>0} pq - g_\epsilon^\sharp(x,q) \\ &\geq \sup_{q>0} pq - g^\sharp(x,q) - (1+L)\epsilon \\ &\geq pg(x) - (1+L)\epsilon. \end{aligned}$$

Thus:

$$\mathbb{E}^{\mathbb{Q}}\left[P_T g(X_T^{t,x})\right] \leq \mathbb{E}^{\mathbb{Q}}\left[g_\epsilon(X_T^{t,x}, P_T)\right] + (1+L)\epsilon. \tag{3.3.1}$$

In addition, we have, for all $t \in [0,T]$, $x \in (0,\infty)$ and $p \in [0,1]$, Since $P_T = u_\epsilon(T, X_T^{t,x}, qQ_T^t) = \partial_q w_\epsilon(T, X_T^{t,x}, qQ_T^t) = \partial_q g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t)$, we have

$$\mathbb{E}^{\mathbb{Q}}\left[g_\epsilon(X_T^{t,x}, P_T)\right] = \mathbb{E}^{\mathbb{Q}}\left[qQ_T^t P_T - g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t)\right].$$

Since $\Gamma_T^t Q_T^t = 1$ and P is a $\mathbb{P}$-martingale, we obtain

$$\mathbb{E}^{\mathbb{Q}}\left[g_\epsilon(X_T^{t,x}, P_T)\right] = qp - \mathbb{E}^{\mathbb{Q}}\left[g_\epsilon^\sharp(X_T^{t,x}, qQ_T^t)\right] = qp - w_\epsilon(t,x,q).$$

Since $p = u_\epsilon(t,x,q) = \partial_q w_\epsilon(t,x,q)$, we obtain, by Lemma 3.3.1,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[g_\epsilon(X_T^{t,x}, qQ_T^t)\right] &= qp - w_\epsilon(t,x,q) \\ &= \max_{r>0} pr - w_\epsilon(t,x,r) \\ &\leq \max_{r>0} pr - w(t,x,r) + (1+L)\epsilon \\ &\leq v(t,x,p) + (1+L)\epsilon. \end{aligned} \tag{3.3.2}$$

We finally obtain, combining (3.3.1) and (3.3.2),

$$\mathbb{E}^{\mathbb{Q}}\left[P_T g(X_T^{t,x})\right] \leq v(t,x,p) + 2(1+L)\epsilon.$$

□

3.3.2 Control space truncation

We fix $N > 0$, and we define

$$\mathcal{A}_p(N) := \{\alpha \in \mathcal{A}_p | \alpha \in [-N,N] \mathrm{d}t \otimes \mathrm{d}\mathbb{P} - \text{a.s.}\}.$$

The goal is to obtain a convergence rate for the approximate control problem

$$v_N(t,x,p) = \inf_{\alpha \in \mathcal{A}_p(N)} \mathbb{E}^{\mathbb{Q}}\left[g(X_T^{t,x}) P_T^{t,p,\alpha}\right].$$

Using Itô's formula and Theorem 3.3.1, we obtain that

$$\alpha_s^\epsilon := \sigma X_s^{t,x} \partial_x u_\epsilon(s, X_s^{t,x}, qQ_s^t) + \lambda q Q_s^t \partial_q u_\epsilon(s, X_s^{t,x}, qQ_s^t)$$

is a $2(L+1)\epsilon$-optimal control.
Using α^ϵ, we construct a control $\alpha^{\epsilon,N}$ valued in $[-N,N]$. This will allow us to upper bound the difference between $v_N(t,x,p)$ and $v(t,x,p)$.

First, we observe that there exists $K_1 > 0$ such that, for $t \in [0,T]$, $x > 0$ and $q > 0$,

$$\begin{aligned}|\partial_x u_\epsilon(t,x,q)| = |\partial_{xq} w_\epsilon(t,x,q)| &= \left|\partial_{xq}\mathbb{E}^{\mathbb{Q}}\left[g^\sharp_\epsilon(X^{t,x}_T, qQ^t_T)\right]\right| \\ &\leq \mathbb{E}^{\mathbb{Q}}\left[\left|\left((\partial_{xq}\rho_\epsilon) * g^\sharp\right)(t,x,q)\right|\right] \leq \frac{K_1}{\epsilon^2},\end{aligned}$$

and similarly $|\partial_q u_\epsilon(t,x,q)| = |\partial_{qq} w_\epsilon(t,x,q)| \leq \frac{K_2}{\epsilon^2}$ for some constant $K_2 > 0$.

Consider the following stopping times:

$$\begin{aligned}\tau_1^N &:= \inf\left\{s \geq t \mid X^{t,x}_s \geq \frac{N\epsilon^2}{2K_1\sigma}\right\}, \\ \tau_2^N &:= \inf\left\{s \geq t \mid Q^t_T \geq \frac{N\epsilon^2}{2K_2|\lambda|q}\right\}, \\ \tau^N &:= \tau_1^N \wedge \tau_2^N.\end{aligned}$$

we have, if $t \leq s \leq \tau^N$,

$$\begin{aligned}|\alpha^\epsilon_s| &\leq \sigma X^{t,x}_s |\partial_x u_\epsilon(s, X^{t,x}_s, qQ^t_s)| + |\lambda| q Q^t_s |\partial_q u_\epsilon(s, X^{t,x}_s, qQ^t_s)| \\ &\leq \sigma \frac{N\epsilon^2}{2K_1\sigma}\frac{K_1}{\epsilon^2} + |\lambda| q \frac{N\epsilon^2}{2K_2|\lambda|q}\frac{K_2}{\epsilon^2} \\ &\leq \frac{N}{2} + \frac{N}{2} = N.\end{aligned}$$

We consider $\alpha^{\epsilon,n}_s := \alpha^\epsilon_s 1_{\{s \leq \tau^N\}}$.

Let $M^{t,x} := \Gamma^t_T g(X^{t,x}_T)$. Using the martingale representation theorem, there exists $m \in \mathbb{R}_+$ and $\beta \in \mathbb{H}^2(\mathbb{F})$ such that:

$$M^{t,x} = m + \int_t^T \beta_s \mathrm{d}W_s.$$

Assumption 3.3.1. *There exists $\iota > 0$ such that*

$$\mathbb{E}\left[\int_t^T |\beta_s|^{2+\iota}\mathrm{d}s\right] < +\infty.$$

Theorem 3.3.2. *For $t \in [0,T]$, $x \in (0,\infty)$, $q > 0$, there exists $C_\iota > 0$ and $\epsilon > 0$ such that, for $p = u_\epsilon(t,x,q)$,*

$$0 \leq v_N(t,x,p) - v(t,x,p) \leq C_\iota \frac{1}{N^{\frac{\iota}{12+8\iota}}}.$$

Proof. By Theorem 3.3.1, we have

$$\begin{aligned}v_N(t,x,p) - v(t,x,p) &\leq \mathbb{E}^{\mathbb{Q}}\left[P^{t,p,\alpha^{\epsilon,N}}_T g(X^{t,x}_T)\right] - \mathbb{E}^{\mathbb{Q}}\left[P^{t,p,\alpha^\epsilon}_T g(X^{t,x}_T)\right] + 2(L+1)\epsilon \\ &\leq \mathbb{E}\Big[M^{t,x}(P^{t,p,\alpha^{\epsilon,N}}_T - P^{t,p,\alpha^\epsilon}_T)\Big] + 2(L+1)\epsilon.\end{aligned}$$

Using Itô's formula and Hölder inequality twice, we obtain, by Assumption 3.3.1,

$$\begin{aligned}
\mathbb{E}\Big[M^{t,x}(P_T^{t,p,\alpha^{\epsilon,N}} - P_T^{t,p,\alpha^{\epsilon}})\Big] &= \mathbb{E}\Big[\int_t^T \beta_s \alpha_s^{\epsilon} 1_{\{s\geq\tau^N\}} \mathrm{d}s\Big] \\
&\leq \mathbb{E}\Big[\int_t^T |\beta_s \alpha_s^{\epsilon}| 1_{\{s\geq\tau^N\}} \mathrm{d}s\Big] \\
&\leq \mathbb{E}\Big[\int_t^T |\beta_s|^{2+\iota} \mathrm{d}s\Big]^{\frac{1}{2+\iota}} \mathbb{E}\Big[\int_t^T |\alpha_s^{\epsilon}|^{\frac{2+\iota}{1+\iota}} 1_{\{s\geq\tau^N\}} \mathrm{d}s\Big]^{\frac{1+\iota}{2+\iota}} \\
&\leq \mathbb{E}\Big[\int_t^T |\beta_s|^{2+\iota} \mathrm{d}s\Big]^{\frac{1}{2+\iota}} \mathbb{E}\Big[\int_t^T |\alpha_s^{\epsilon}|^2 \mathrm{d}s\Big]^{\frac{1}{2}} \mathbb{E}\Big[\int_t^T 1_{s\geq\tau^N} \mathrm{d}s\Big]^{\frac{\iota}{2(2+\iota)}}.
\end{aligned}$$

We have, using Markov's inequality,

$$\begin{aligned}
\mathbb{E}\Big[\int_t^T 1_{\{s\geq\tau^N\}} \mathrm{d}s\Big] &= \int_t^T \mathbb{P}(s \geq \tau^N) \mathrm{d}s \\
&\leq T\mathbb{P}(T \geq \tau^N) \\
&\leq T\left(\mathbb{P}(T \geq \tau_1^N) + \mathbb{P}(T \geq \tau_2^N)\right) \\
&\leq T\mathbb{P}(\sup_{s\in[t,T]} X_s^{t,x} \geq \frac{N\epsilon^2}{2K_1\sigma}) + T\mathbb{P}(\sup_{s\in[t,T]} Q_s^t \geq \frac{N\epsilon^2}{2K_2|\lambda|q}) \\
&\leq T\mathbb{E}\Big[\sup_{s\in[t,T]} X_s^{t,x}\Big] \frac{2K_1\sigma}{N\epsilon^2} + T\mathbb{E}\Big[\sup_{s\in[t,T]} Q_s^t\Big] \frac{2K_2|\lambda|q}{N\epsilon^2} \\
&\leq \frac{C_1}{N\epsilon^2},
\end{aligned} \tag{3.3.3}$$

for some $C_1 > 0$. Moreover:

$$\begin{aligned}
&\mathbb{E}\Big[\int_t^T |\alpha_s^{\epsilon}|^2 \mathrm{d}s\Big] \\
&\leq \frac{\sigma^2}{\epsilon^4}\mathbb{E}\Big[\int_t^T |X_s^{t,x}|^2 \mathrm{d}s\Big] + \frac{\lambda^2 q^2}{\epsilon^4}\mathbb{E}\Big[\int_t^T |Q_s^t|^2 \mathrm{d}s\Big] + \frac{2\lambda\sigma q}{\epsilon^4}\mathbb{E}\Big[\int_t^T X_s^{t,x} Q_s^t \mathrm{d}s\Big] \\
&\leq \frac{C_2}{\epsilon^4},
\end{aligned} \tag{3.3.4}$$

for some $C_2 > 0$. Using (3.3.3) and (3.3.4), we obtain,

$$\begin{aligned}
v_N(t,x,p) - v(t,x,p) &\leq \|\beta\|_{2+\iota} \frac{C_2^{\frac{1}{2}}}{\epsilon^2} \frac{C_1^{\frac{\iota}{2(2+\iota)}}}{(N\epsilon^2)^{\frac{\iota}{2(2+\iota)}}} + 2(L+1)\epsilon \\
&\leq \frac{\|\beta\|_{2+\iota} C_2^{\frac{1}{2}} C_1^{\frac{\iota}{2(2+\iota)}}}{N^{\frac{\iota}{2(2+\iota)}} \epsilon^{2+\frac{\iota}{2+\iota}}} + 2(L+1)\epsilon.
\end{aligned}$$

One can minimize this quantity taking

$$\epsilon = \left(\frac{4+3\iota}{2+\iota} \frac{\|\beta\|_{2+\iota} C_2^{\frac{1}{2}} C_1^{\frac{\iota}{2(2+\iota)}}}{2(L+1)N^{\frac{\iota}{2(2+\iota)}}}\right)^{\frac{2+\iota}{6+4\iota}} =: \tilde{C}_\iota \frac{1}{N^{\frac{\iota}{12+8\iota}}}, \tag{3.3.5}$$

and one obtains the result. □

Remark 3.3.1. *If g is bounded, we can take $\iota \to \infty$ and we obtain, for all $\eta > 0$ and $t \in [0,T]$, $x > 0$, $q > 0$ and $p = u_\epsilon(t,x,q)$ (with ϵ defined by (3.3.5)), that there exists $C_\eta > 0$ such that*

$$0 \leq v_N(t,x,p) - v(t,x,p) \leq C_\eta N^{\eta - \frac{1}{8}}.$$

Chapter 4

A numerical scheme for the quantile hedging problem

The content of this chapter is from an article in collaboration with Jean-François Chassagneux and Christoph Reisinger [BCR19], submitted to SIAM Journal on Financial Mathematics.

4.1 Introduction

In this work, we study the numerical approximation of the *quantile hedging price* of a European contingent claim in a market with possibly some imperfections. The quantile hedging problem is a specific case of a broader class of approximate hedging problems. It consists in finding the minimal initial endowment of a portfolio that will allow hedging a European claim with a given probability p of success, the case $p = 1$ corresponding to the classical problem of (super)replication. This approach has been popularised by Föllmer and Leukert in [FL99] who provided a closed form solution in a special setting.

The first PDE characterisation was introduced by [BET09] in a possibly incomplete market setting with portfolio constraints. Various extensions have been considered since this work: to jump dynamics [Mor11]; to the Bermudan case [BBC16] and American case [Dum+17]; to a non-Markovian setting [BER15; Dum16]; and to a finite number of quantile constraints [BNV12].

Except for [BBC16; BNV12], all the aforementioned works are of a theoretical nature. The lack of established numerical methods for these problems is a clear motivation for our study. In the following, we present in more detail the quantile hedging problem and the new numerical method we introduce and study in this paper.

On a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a d-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ and denote by $(\mathcal{F}_t)_{t\in[0,T]}$ its natural filtration. We suppose that all the randomness comes from the Brownian motion and assume that $\mathcal{F} = \mathcal{F}_T$.

Let $\mu : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathcal{M}_d(\mathbb{R})$, where $\mathcal{M}_d(\mathbb{R})$ is the set of $d \times d$ matrices with real entries, and $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be Lipschitz continuous functions, with Lipschitz constant L.

For $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$ and $\nu \in \mathbb{H}^2$, which denotes the set of predictable square-integrable processes, we consider the solution $(X^{t,x}, Y^{t,x,y,\nu})$ to the following stochastic differential equations:

$$\begin{aligned} X_s &= x + \int_t^s \mu(X_u)\mathrm{d}u + \int_t^s \sigma(X_u)\mathrm{d}W_u, \\ Y_s &= y - \int_t^s f(u, X_u, Y_u, \nu_u)\mathrm{d}u + \int_t^s \nu_u \mathrm{d}W_u, \qquad s \in [t,T]. \end{aligned}$$

In the financial applications we are considering, X will typically represent the log-price of risky assets, the control process ν is the amount invested in the risky assets, and the function f is non-linear to allow the inclusion of market frictions in the model. In a typical financial example, which will be investigated in the numerical section, the underlying diffusion X is a one-dimensional Brownian motion with constant drift $\mu \in \mathbb{R}$ and volatility $\sigma > 0$, there is a constant borrowing rate R and a lending rate r with $R \geq r$, and in this situation, the function f is given by (see [EKPQ97], Example 1.1),

$$f(t,x,y,z) = -ry - \sigma^{-1}\mu z + (R-r)(y - \sigma^{-1}z)^-.$$

The quantile hedging problem corresponds to the following stochastic control problem: for $(t,x,p) \in [0,T] \times \mathbb{R}^d \times [0,1]$ find

$$v(t,x,p) := \inf\left\{ y \geq 0 : \exists \nu \in \mathbb{H}_2,\ \mathbb{P}\left(Y_T^{t,x,y,\nu} \geq g(X_T^{t,x})\right) \geq p \right\}. \tag{4.1.1}$$

The main objective of this paper is to design a numerical procedure to approximate the function v by discretizing an associated non-linear PDE first derived in [BET09]. A key point in the derivation of this PDE is to reformulate the problem as a classical stochastic target problem by introducing a new control process representing the conditional probability of success. To this end, for $\alpha \in \mathbb{H}^2$, we denote

$$P_s^{t,p,\alpha} := p + \int_t^s \alpha_s \mathrm{d}W_s\ ,\ t \leq s \leq T\ ,$$

and by $\mathcal{A}^{t,p}$ the set of α such that $P_\cdot^{t,p,\alpha} \in [0,1]$. The problem (4.1.1) can be rewritten as

$$v(t,x,p) := \inf\left\{ y \geq 0 : \exists(\nu,\alpha) \in (\mathbb{H}_2)^2,\ Y_T^{t,x,y,\nu} \geq g(X_T^{t,x})\mathbf{1}_{\{P_T^{t,p,\alpha}>0\}} \right\}$$

(see Proposition 3.1 in [BET09] for details). In our framework, the above singular stochastic control problem admits a representation in terms of a non-linear expectation, generated by a Backward Stochastic Differential Equation (BSDE),

$$v(t,x,p) = \inf_{\alpha \in \mathcal{A}^{t,p}} \mathcal{Y}_t^\alpha, \tag{4.1.2}$$

where $(\mathcal{Y}^\alpha, \mathcal{Z}^\alpha)$ is the solution to

$$\mathcal{Y}_s^\alpha = g(X_T^{t,x})\mathbf{1}_{\{P_T^{t,p,\alpha}>0\}} + \int_s^T f(s, X_s, \mathcal{Y}_s^\alpha, \mathcal{Z}_s^\alpha)\mathrm{d}s - \int_s^T \mathcal{Z}_s^\alpha \mathrm{d}W_s\ ,\ t \leq s \leq T\ .$$

The results in [BER15] justify the previous representation and give a dynamic programming principle for the control problem in a general setting. In the Markovian setting, this would lead naturally to the following PDE for v in $[0,T) \times \mathbb{R}^d \times (0,1)$:

$$\text{“} -\partial_t\varphi + \sup_{a\in\mathbb{R}^d} F^a(t,x,\varphi,D\varphi,D^2\varphi) = 0\text{”}, \tag{4.1.3}$$

where, for $(t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^+$, $q := \begin{pmatrix} q^x \\ q^p \end{pmatrix} \in \mathbb{R}^{d+1}$ and $A := \begin{pmatrix} A^{xx} & A^{xp} \\ A^{xp\top} & A^{pp} \end{pmatrix} \in \mathbb{S}^{d+1}$, $A^{xx} \in \mathbb{S}^d$, we define

$$F^a(\Xi) := -f(t,x,y,\mathfrak{z}(x,q,a)) - \mathcal{L}(x,q,A,a) \ , \tag{4.1.4}$$

denoting $\Xi := (t,x,y,q,A)$, with

$$\mathfrak{z}(x,q,a) := q^x\sigma(x) + q^p a \ , \tag{4.1.5}$$
$$\mathcal{L}(x,q,A,a) := \mu(x)^\top q^x + \frac{1}{2}\operatorname{Tr}\left[\sigma(x)\sigma(x)^\top A^{xx}\right] + \frac{|a|^2}{2}A^{pp} + a^\top\sigma(x)^\top A^{xp} \tag{4.1.6}$$

The PDE formulation in (4.1.3) is not entirely correct as the supremum part may degenerate and requires semi-limit relaxation to be mathematically rigourous. We refer to [BET09], where it has been obtained in a more general context. We shall use an alternative PDE formulation to this "natural" one (4.1.3), which we give at the start of Section 4.2.

Moreover, the value function v continuously satisfies the following boundary conditions in the p-variable:

$$v(t,x,0) = 0 \ \text{ and } \ v(t,x,1) = V(t,x) \text{ on } [0,T] \times (0,\infty)^d, \tag{4.1.7}$$

where V is the super-replication price of the contingent claim with pay-off $g(\cdot)$.

It is also known that v has a discontinuity as $t \to T$. By definition, the terminal condition is

$$\mathbb{R}^d \times [0,1] \ni (x,p) \mapsto g(x)1_{p>0} \in \mathbb{R}^+ \ ,$$

but the values which are continuously attained are obtained by convexification [BET09]:

$$v(T-,x,p) = pg(x) \text{ on } \mathbb{R}^d \times [0,1], \tag{4.1.8}$$

and we shall work with this terminal condition at $t = T$ from now on.

To design the numerical scheme to approximate v, we use the following strategy:

1. Bound and discretise the set where the controls α take their values.

2. Consider an associated Piecewise Constant Policy Timestepping (PCPT) scheme for the control processes.

3. Use a monotone finite difference scheme to approximate in time and space the PCPT solution resulting from 1. & 2.

The approximation of controlled diffusion processes by ones where policies are piecewise constant in time was first analysed by [Kry99]; in [Kry00], this procedure is used in conjunction with Markov chain approximations to diffusion processes to construct fully discrete approximation schemes to the associated Bellman equations and to derive their convergence order. An improvement to the order of convergence from [Kry99] was shown recently in [JPR19] using a refinement of Krylov's original, probabilistic techniques.

Using purely viscosity solution arguments for PDEs, error bounds for such approximations are derived in [BJ07], which are weaker than those in [Kry99] for the control approximation scheme, but improve the bounds in [Kry00] for the fully discrete scheme. In [RF16], using a switching system approximation introduced in [BJ07], convergence is proven for a generalised scheme where linear PDEs are solved piecewise in time on different meshes, and the control optimisation is carried out at the end of time intervals using possibly non-monotone, higher order interpolations. An extension of the analysis in [RF16] to jump-processes and non-linear expectations is given in [DRZ18].

Our first contribution is to prove that the approximations built in step 1. and 2. above are convergent for the quantile hedging problem, which has substantial new difficulties compared to the settings considered in the aforementioned works. For this we rely heavily on the comparison theorem for the formulation in (4.2.1) and we take advantage of the monotonicity property of the approximating sequences. The main new difficulties come from the non-linear form of the PDE including unbounded controls, and in particular the boundaries in the p-variable. To deal with the latter especially, we rely on some fine estimates for BSDEs to prove the consistency of the scheme including the strong boundary conditions (see Lemma 4.2.2 and Lemma 4.2.3).

Our second contribution is to design the monotone scheme in step 3. and to prove its convergence. The main difficulties here come from the non-linearity of the new term – the driver of the BSDE – in the gradient; the degeneracy of the diffusion operator given in (4.1.6) due to the two state processes being driven by a single Brownian motion; and again the boundedness of the domain in the p direction. In particular, a careful analysis of the consistency of the boundary condition is needed (see Proposition 4.3.4).

To the best of our knowledge, this is the first numerical method for the quantile hedging problem in this non-linear market specification. In the linear market setting, using a dual approach, [BBC16] combines the solution of a linear PDE with Fenchel-Legendre transforms to tackle the problem of Bermudan quantile hedging. Their approach cannot be directly adapted here due to the presence of the non-linearity. The dual approach in the non-linear setting would impose some convexity assumption on f and would require us to solve fully non-linear PDEs. Note that here f is only required to be Lipschitz continuous in (y, z). We believe that an interesting alternative to our method would be to extend the work of [Bok+09] to the non-linear market setting considered here.

The rest of the paper is organised as follows. In Section 4.2, we derive the control approximation and PCPT scheme associated with items 1. & 2. above and prove their convergence. In Section 4.3, we give a monotone finite difference approximation which is shown to converge to the semi-discrete PCPT scheme. In Section 4.4, we present numerical results for a specific application and analyse the observed convergence. Finally, the appendix contains some of the longer, more technical proofs and collects useful background results used in the paper.

Notations Throughout, $\operatorname{diag}(x)$ is the diagonal matrix of size d, whose diagonal is given by x. We denote by $\mathcal{S}$ the sphere in $\mathbb{R}^{d+1}$ of radius 1 and by $\mathcal{D}$ the set of vectors $\eta \in \mathcal{S}$ such that their first component $\eta^1 = 0$. For any $\eta \in \mathcal{S} \setminus \mathcal{D}$, we denote $\eta^\flat := \frac{1}{\eta^1}(\eta^2, \dots, \eta^{d+1})^\top \in \mathbb{R}^d$. By extension, for $\mathcal{Z} \subset \mathcal{S} \setminus \mathcal{D}$, $\mathcal{Z}^\flat := \{\eta^\flat \in \mathbb{R}^d \,|\, \eta \in \mathcal{Z}\}$.

We further denote by $\mathcal{BC} := L^\infty([0,T], \mathcal{C}^0(\mathbb{R}^d \times [0,1]))$, the space of functions u that are essentially bounded in time and continuous with respect to their space variable. The convergence in $\mathcal{C}^0([0,T] \times \mathbb{R}^d \times [0,1])$ considered here is local uniform convergence.

4.2 Convergence of a discrete-time scheme

In this section, we design a Piecewise Constant Policy Timestepping (PCPT) scheme which is convergent to the value function v defined in (4.1.2).

Following [Bok+09], it has been shown in [BC17] that the function v is equivalently a viscosity solution of the following PDE (see Theorems 3.1 and 3.2 in [BC17]):

$$\mathcal{H}(t, x, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0 \tag{4.2.1}$$

in $(0,T) \times \mathbb{R}^d \times (0,1)$, where $\mathcal{H}$ is the continuous operator

$$\mathcal{H}(\Theta) = \sup_{\eta \in \mathcal{S}} H^\eta(\Theta) \ ,$$

where for $(t, x, y, b) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}$, $q \in \mathbb{R}^{d+1}$ and $A \in \mathbb{S}^{d+1}$ as given above (4.1.4), and $\Theta := (t, x, y, b, q, A)$, we define

$$H^\eta(\Theta) = (\eta^1)^2 \left(-b - f(t, x, y, \mathfrak{z}(x, q, \eta^\flat)) - \mathcal{L}(x, q, A, \eta^\flat)\right) \ , \quad \text{for } \eta \in \mathcal{S} \setminus \mathcal{D} \ .$$

Recall also the definition of $\mathcal{L}$ and $\mathfrak{z}$ in (4.1.5) and (4.1.6).

This representation and its properties are key in the proof of convergence. Loosely speaking, it is obtained by "compactifying" the set $\{1\} \times \mathbb{R}^d$ to the unit sphere $\mathcal{S}$. A comparison theorem is shown in Theorem 3.2 in [BC17].

As partially stated in the introduction, we will work under the following assumption:

(H)(i) The functions μ, σ are L-Lipschitz continuous and g is bounded and L_g-Lipschitz.

(ii) The function f is measurable and for all $t \in [0,T]$, $f(t, \cdot, \cdot, \cdot)$ is L-Lipschitz. For all $(t, x, z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d$, the function $y \mapsto f(t, x, y, z)$ is decreasing. Moreover,

$$f(t, x, 0, 0) = 0. \tag{4.2.2}$$

Under this Lipschitz continuity assumption, the mapping $\mathcal{S} \setminus \mathcal{D} \ni \eta \mapsto H^\eta(\Theta) \in \mathbb{R}$ extends continuously to $\mathcal{S}$ by setting (see Remark 3.1 in [BC17]), for all $\eta \in \mathcal{D}$,

$$H^\eta(\Theta) = -\frac{1}{2} A^{pp} \ .$$

Remark 4.2.1. *(i) In (H)(ii), the monotonicity assumption is not a restriction, as in a Lipschitz framework, the classical transformation $\tilde{v}(t, x, p) := e^{\lambda t} v(t, x, p)$ for λ large enough allows to reach this setting; see Remark 3.3 in [BC17] for details.*
(ii) The condition (4.2.2) is a reasonable financial modelling assumption: It says that starting out in the market with zero initial wealth and making no investments will lead to a zero value of the wealth process.
(iii) Since f is decreasing and g is bounded, it is easy to see that $|V|_\infty \leq |g|_\infty$, where V is the super-replication price.

4.2.1 Discrete set of controls

In order to introduce a discrete-time scheme which approximates the solution v of (4.2.1), (4.1.7) and (4.1.8), we start by discretizing the set of controls $\mathcal{S}$.

Let $(\mathcal{R}_n)_{n\geq 1}$ be an increasing sequence of closed subsets of $\mathcal{S} \setminus \mathcal{D}$ such that

$$\overline{\bigcup_{n\geq 1} \mathcal{R}_n} = \mathcal{S} \,. \tag{4.2.3}$$

For $n \geq 1$, let $v_n : [0,T] \times \mathbb{R}^d \times [0,1] \to \mathbb{R}$ be the unique continuous viscosity solution of the following PDE:

$$\mathcal{H}_n(t, x, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0, \tag{4.2.4}$$

satisfying the boundary conditions (4.1.7)-(4.1.8), see Corollary 4.6.1. Above, the operator $\mathcal{H}_n$ is naturally given by

$$\mathcal{H}_n(\Theta) := \sup_{\eta \in \mathcal{R}_n} H^\eta(\Theta) \,.$$

Proposition 4.2.1. *The functions v_n converge to v in $\mathcal{C}^0([0,T] \times \mathbb{R}^d \times [0,1])$.*

Proof. 1. For $n' < n$, we observe that $v_{n'}$ is a super-solution of (4.2.4) as $\mathcal{R}_{n'} \subset \mathcal{R}_n$. Using the comparison result of Proposition 4.6.1 in the appendix, we obtain that $v_{n'} \geq v_n$. Similarly, using the comparison principle ([BC17], Theorem 3.2), we obtain that $v_n \geq v$, for all $n \geq 1$.

For all $(t, x, p) \in [0,T] \times (0,\infty)^d \times [0,1]$, let:

$$\overline{v}(t,x,p) = \lim_{j\to\infty} \sup \left\{ v_n(s,y,q) : n \geq j \text{ and } \|(s,y,q) - (t,x,p)\| \leq \frac{1}{j} \right\},$$
$$\underline{v}(t,x,p) = \lim_{j\to\infty} \inf \left\{ v_n(s,y,q) : n \geq j \text{ and } \|(s,y,q) - (t,x,p)\| \leq \frac{1}{j} \right\}.$$

From the above discussion, recalling that v_1 and v are continuous, we have

$$v_1 \geq \overline{v} \geq \underline{v} \geq v \,,$$

which shows that $\overline{v}$ and $\underline{v}$ satisfy the boundary conditions (4.1.7)-(4.1.8).

In order to prove the theorem, it is enough to show that $\overline{v}$ is a viscosity subsolution of (4.2.1) and $\underline{v}$ is a viscosity supersolution (which follows similarly and is therefore omitted). The comparison principle ([BC17], Theorem 3.2) then implies that $v = \overline{v} = \underline{v}$, and it follows from [CIL92], Remark 6.4 that the convergence $v_n \to v$ as $n \to \infty$ is uniform on every compact set. Using Theorem 6.2 in [Ach+13], we obtain that $\overline{v}$ is a subsolution to

$$\underline{\mathcal{H}}(t, x, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0 \text{ on } (0,T) \times (0,\infty)^d \times (0,1) \,,$$

where

$$\underline{\mathcal{H}}(\Theta) = \lim_{j\to\infty} \inf \left\{ \mathcal{H}_n(\Theta') : n \geq j \text{ and } \|\Theta - \Theta'\| \leq \frac{1}{j} \right\}.$$

In the next step, we prove that $\underline{\mathcal{H}} = \mathcal{H}$, which concludes the proof of the proposition.

2. Let us denote by $\mathfrak{P}_n : \mathcal{S} \to \tilde{\mathcal{S}}_n$ the closest neighbour projection on the closed set $\tilde{\mathcal{S}}_n$. From (4.2.3), we have that $\lim_{n\to\infty} \mathfrak{P}_n(\eta) = \eta$, for all $\eta \in \mathcal{S}$. We also have that

$$\mathcal{H}(\Theta) = H^{\eta^*}(\Theta)$$

for some $\eta^* \in \operatorname{argmax}_{\eta\in\mathcal{S}} H^\eta(\Theta)$ as $\mathcal{S}$ is compact. Let us now introduce $\eta_n := \mathfrak{P}_n(\eta^*)$ and by continuity of H, we have

$$H^{\eta_n}(\Theta) \to \mathcal{H}(\Theta) \quad \text{as } n \to \infty\ .$$

We also observe that

$$H^{\eta_n}(\Theta) \le \mathcal{H}_n(\Theta) \le \mathcal{H}(\Theta)\ .$$

This proves the convergence $\mathcal{H}_n(\Theta) \uparrow \mathcal{H}(\Theta)$, for all Θ. As $\mathcal{H}$ is continuous, we conclude by using Dini's Theorem that the convergence is uniform on compact subsets, leading to $\underline{\mathcal{H}} = \mathcal{H}$. □

4.2.2 The PCPT scheme

From now on, we fix $n \ge 1$ and $\mathcal{R}_n$ the associated discrete set of control. For $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^+$, $q \in \mathbb{R}^{d+1}$ and $A \in \mathbb{S}^{d+1}$, denoting $\Xi := (t, x, y, q, A)$, we define

$$\mathcal{F}_n(\Xi) = \sup_{a\in\mathcal{R}_n^\flat} F^a(\Xi) \text{ with } F^a(\Xi) := -f(t, x, y, \mathfrak{z}(x, q, a)) - \mathcal{L}(x, q, A, a)\ .$$

Following the proof of Corollary 4.6.1 in the appendix, we easily observe that v_n is also the unique viscosity solution to

$$-\partial_t\varphi + \mathcal{F}_n(t, x, \varphi, D\varphi, D^2\varphi) = 0 \text{ on } [0, T) \times \mathbb{R}^d \times (0, 1) \tag{4.2.5}$$

with the same boundary conditions (4.1.7) and (4.1.8). The above PDE is written in a more classical way and we will mainly consider this form in the sequel. Let us observe in particular that $K := \mathcal{R}_n^\flat$ is a discrete subset of $\mathbb{R}^d$, such that (4.2.5) appears as a natural discretization of (4.1.3) and will be simpler to study.

To approximate v_n, we consider an adaptation of the PCPT scheme in [Kry00; BJ07; RF16], and especially [DRZ18], to our setting, as described below.

For $\kappa \in \mathbb{N}^*$, we consider a grid on the time interval $[0, T]$:

$$\pi = \{0 =: t_0 < \cdots < t_k < \cdots < t_\kappa := T\},$$

and denote $|\pi| := \max_{0\le k<\kappa}(t_{k+1} - t_k)$.

For $0 \le t < s \le T$, $a \in K$ and a continuous $\phi : \mathbb{R}^d \times [0, 1] \to \mathbb{R}$, we denote by $S^a(s, t, \phi) : [t, s] \times \mathbb{R}^d \times [0, 1] \to \mathbb{R}$ the unique solution of

$$\begin{aligned} -\partial_t\varphi + F^a(t, x, \varphi, D\varphi, D^2\varphi) &= 0 && \text{on } [t, s) \times \mathbb{R}^d \times (0, 1), && (4.2.6)\\ \varphi(s, x, p) &= \phi(x, p) && \text{on } \mathbb{R}^d \times [0, 1], \\ \varphi(r, x, p) &= B^p(t, s, \phi)(r, x), && \text{on } [t, s) \times \mathbb{R}^d \times \{0, 1\}. && (4.2.7)\end{aligned}$$

Here, let $B^p(t,s,\phi)$, for $p \in \{0,1\}$, be the solution to

$$-\partial_t\varphi + F^0(r,x,\varphi,D\varphi,D^2\varphi) = 0 \text{ on } [t,s)\times\mathbb{R}^d\ , \tag{4.2.8}$$

with terminal condition $B^p(t,s,\phi)(s,x) = \phi(x,p)$.

The solution to the PCPT scheme associated with the grid π is then the function $v_{n,\pi} : [0,T]\times\mathbb{R}^d\times[0,1]\to\mathbb{R}$ such that

$$\mathfrak{S}(\pi,t,x,p,v_{n,\pi}(t,x,p),v_{n,\pi}) = 0,$$

where, for a grid π, $(t,x,p,y)\in[0,T]\times\mathbb{R}^d\times[0,1]\times\mathbb{R}^+$ and a function $u\in\mathcal{BC}$,

$$\mathfrak{S}(\pi,t,x,p,y,u) = \begin{cases} y - \min_{a\in K} S^a\left(t_\pi^+, t_\pi^-, u(t_\pi^+,\cdot)\right)(t,x,p) & \text{if } t<T, \\ y - \hat{g}(x)p & \text{otherwise,}\end{cases} \tag{4.2.9}$$

with

$$t_\pi^+ := \inf\{r\in\pi \,|\, r>t\} \ \text{ and } \ t_\pi^- := \sup\{r\in\pi\,|\,r\le t\}. \tag{4.2.10}$$

We will drop the subscript π for brevity whenever we consider a fixed mesh.

Let us observe that the function $v_{n,\pi}$ can alternatively, and perhaps more intuitively, be described by the following backward algorithm:

1. Initialisation: set $v_{n,\pi}(T,x,p) := g(x)p$, $x\in\mathbb{R}^d\times[0,1]$.
2. Backward step: For $k=\kappa-1,\dots,0$, compute $w^{k,a} := S^a(t_k,t_{k+1},v_{n,\pi}(t_{k+1},\cdot))$ and set

$$v_{n,\pi}(\cdot) := \min_{a\in K} w^{k,a}\ . \tag{4.2.11}$$

Remark 4.2.2. *In our setting, we can easily identify the boundary values (of the scheme):*

(i) At $p=0$, the terminal condition is $\phi(T,x)=0$ (recall that $v(T,x,p)=g(x)\mathbf{1}_{\{p>0\}}$), and this propagates through the backward iteration, so that $v_{n,\pi}(t,x,0)=0$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$.

(ii) At $p=1$, the terminal condition is $\phi(T,x)=g(x)$ and the boundary condition is thus given by $v_{n,\pi}(t,x,1)=V(t,x)$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$, where V is the super-replication price.

The main result of this section is the following.

Theorem 4.2.1. *The function $v_{n,\pi}$ converges to v_n in $\mathcal{C}^0([0,T]\times\mathbb{R}^d\times[0,1])$ as $|\pi|\to 0$.*

Proof. 1. We first check the consistency with the boundary condition. Let $\hat{a}\in K$ and $\hat{w}$ be the (continuous) solution of

$$-\partial_t\varphi + F^{\hat{a}}(t,x,\varphi,D\varphi,D^2\varphi) = 0 \text{ on } [0,T)\times\mathbb{R}^d\times(0,1) \tag{4.2.12}$$

with boundary condition $v(t,x,p) = pV(t,x)$ on $[0,T] \times \mathbb{R}^d \times \{0,1\} \bigcup \{T\} \times \mathbb{R}^d \times [0,1]$. By backward induction on π, one gets that

$$v_{n,\pi} \le \hat{w} \,.$$

Indeed, we have $v_{n,\pi}(T,\cdot) = \hat{w}(T,\cdot)$. Now if the inequality is true at time t_k, $k \ge 1$, we have, using the comparison result for (4.2.12), recalling Proposition 4.6.1, that

$$w^{k,\hat{a}}(t,\cdot) \le \hat{w}(t,\cdot) \text{ for } t \in [t_{k-1}, t_k] \,,$$

and thus *a fortiori* $\hat{w}(t,\cdot) \ge v_{n,\pi}(t,\cdot)$, for $t \in [t_{k-1}, t_k]$.

We also obtain that

$$v_{n,\pi}(\cdot) \ge v_n(\cdot) \tag{4.2.13}$$

by backward induction. Indeed, we have $v_{n,\pi}(T,\cdot) = v_n(T,\cdot)$. Assume that the inequality is true at time t_k, $k \ge 1$. We observe that $w^{k,a}$ is a supersolution of (4.2.4), namely the PDE satisfied by v_n. By the comparison result, this implies that $w^{k,a}(t,\cdot) \ge v_n(t,\cdot)$, for $t \in [t_{k-1}, t_k]$. Taking the infimum over $a \in K$ yields then (4.2.13).

Since

$$v_n \le \underline{w} \le \overline{w} \le \hat{w} \,,$$

where

$$\overline{w}(t,x,p) = \limsup_{(t',x',p',|\pi|) \to (t,x,p,0)} v_{n,\pi}(t',x',p') \text{ and } \underline{w} = \liminf_{(t',x',p',|\pi|) \to (t,x,p,0)} v_{n,\pi}(t',x',p'),$$

we obtain that $\underline{w}$ and $\overline{w}$ satisfy the boundary conditions (4.1.7)–(4.1.8).

2. We prove below that the scheme is *monotone*, *stable* and *consistent*, see Proposition 4.2.2, Proposition 4.2.3, and Proposition 4.2.4, respectively. Combining this with step 1. and Theorem 2.1 in [BS91] then ensures the convergence in $\mathcal{C}^0$ of $v_{n,\pi}$ to v_n as $|\pi| \to 0$. □

Remark 4.2.3. *We prove the following properties by a combination of viscosity solution arguments and – mostly – BSDE arguments where they appear more natural. It should be possible to use only PDE arguments with similar main steps as in [BJ07] .*

Proposition 4.2.2 (Monotonicity)**.** *Let $u \ge v$ for $u, v \in \mathcal{BC}$, $(t,x,p) \in [0,T] \times \mathbb{R}^d \times [0,1]$, $y \in \mathbb{R}$. We have:*

$$\mathfrak{S}(\pi,t,x,p,y,u) \le \mathfrak{S}(\pi,t,x,p,y,v).$$

Proof. Let $t < T, x \in \mathbb{R}^d, p \in [0,1]$. By definition of $v_{n,\pi}$, recalling (4.2.11), it is sufficient to prove that, for any $a \in K$, we have:

$$S^a(t^+,t^-,u(t^+,\cdot))(t,x,p) \ge S^a(t^+,t^-,v(t^+,\cdot))(t,x,p),$$

with t^+, t_- defined in (4.2.10) . But this follows directly from the comparison result given in Proposition 4.6.1. □

We now study the stability of the scheme. We first show that the solution of the scheme $v_{n,\pi}$ is increasing in its third variable. This is not only an interesting property in its own right which the piecewise constant policy solution inherits from the solution to the original problem (4.1.1), but it also allows us to obtain easily a uniform bound for $v_{n,\pi}$, namely the boundary condition at $p = 1$.

Lemma 4.2.1. *The scheme (4.2.9) has the property, for all $t \in [0, T]$ and $x \in \mathbb{R}^d$:*

$$v_{n,\pi}(t, x, q) \leq v_{n,\pi}(t, x, p) \text{ if } 0 \leq q \leq p \leq 1.$$

Proof. We will prove the assertion by induction on $k \in \{0, \dots, \kappa\}$. For $t = T = t_\kappa$ and every $x \in \mathbb{R}^d$, we have $(x, p) \mapsto v_{n,\pi}(T, x, p) := g(x)p$, which increasing in p.

Let $1 \leq k < \kappa - 1$. Assume now that $v_{n,\pi}(t, x, \cdot)$ is increasing for all $t \geq t_{k+1}$ and $x \in \mathbb{R}^d$. We show that $v_{n,\pi}(t, x, \cdot)$ is also increasing for $t \in [t_k, t_{k+1})$ and $x \in \mathbb{R}^d$.

Let $0 \leq q \leq p \leq 1$. By the definition of $v_{n,\pi}$ in (4.2.11), it is sufficient to show that for each $a \in K$, we have, for $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$w^{k,a}(t, x, q) \leq w^{k,a}(t, x, p) \; .$$

From Lemma 4.6.1(i) in the appendix, these two quantities admit a probabilistic representation with two different random terminal times

$$\tau^q = \inf\{s \geq t : P_s^{t,q,a} \in \{0, 1\}\} \wedge t_{k+1},$$
$$\tau^p = \inf\{s \geq t : P_s^{t,p,a} \in \{0, 1\}\} \wedge t_{k+1}.$$

However, using Lemma 4.6.1(ii), we can write probabilistic representations with BSDEs with terminal time t_{k+1}: we have that $S^a(t_{k+1}, t_k, w_\pi(t_{k+1}, \cdot))(t, x, p) = \tilde{Y}_t^{t,x,p,a}$, where $\tilde{Y}_t^{t,x,p,a}$ is the first component of the solution of the following BSDE:

$$Y_s = v_{n,\pi}(t_{k+1}, X_{t_{k+1}}^{t,x}, \tilde{P}_{t_{k+1}}^{t,p,a}) + \int_s^{t_{k+1}} f(u, X_u^{t,x}, Y_u, Z_u) \mathrm{d}u - \int_s^{t_{k+1}} Z_u \mathrm{d}W_u,$$

where $\tilde{P}^{t,p,a}$ is the process defined by:

$$\tilde{P}_s^{t,p,a} = p + \int_t^s a 1_{\{u \leq \tau^p\}} \mathrm{d}W_u,$$

and a similar representation holds for $S^a(t_{k+1}, t_k, w_\pi(t_{k+1}, \cdot))(t, x, q)$.

It remains to show that

$$v_{n,\pi}(t_{k+1}, X_{t_{k+1}}^{t,x}, \tilde{P}_{t_{k+1}}^{t,p,a}) \geq v_{n,\pi}(t_{k+1}, X_{t_{k+1}}^{t,x}, \tilde{P}_{t_{k+1}}^{t,q,a}). \tag{4.2.14}$$

If this is true, the classical comparison theorem for BSDEs (see e.g. Theorem 2.2 in [EKPQ97]), concludes the proof.

First, we observe that $P_{\tau_p}^{t,p,a} \geq P_{\tau_p}^{t,q,a}$. On $\{\tau_p = T\}$, (4.2.14) holds straightforwardly by the induction hypothesis. On $\{\tau_p < T\}$, if $P_{\tau_p}^{t,p,a} = 1$ then $P_T^{t,p,a} = 1$ and (4.2.14) holds by induction hypothesis, as $P_T^{t,q,a} \leq 1$; if $P_{\tau_p}^{t,p,a} = 0$ then *a fortiori* $P_{\tau_p}^{t,q,a} = 0$ and $P_T^{t,p,a} = P_T^{t,q,a} = 0$, which concludes the proof. □

Proposition 4.2.3 (Stability)**.** *The solution to scheme (4.2.9) is bounded.*

Proof. For any π and any $(t,x,p)\in[0,T]\times\mathbb{R}^d\times[0,1]$, we have $v_{n,\pi}(t,x,p)\le v_{n,\pi}(t,x,1)=V(t,x)$. □

To prove the consistency of the scheme, we will need the two following lemmata.

Lemma 4.2.2. *For $0\le\tau\le t\le\theta\le T$, $\xi\in\mathbb{R}$, and $\phi\in\mathcal{C}^\infty([0,T]\times\mathbb{R}^d\times[0,1])$, the following holds*

$$|S^a(\tau,\theta,\phi(\theta,\cdot)+\xi)(t,\cdot)-S^a(\tau,\theta,\phi(\theta,\cdot))(t,\cdot)-\xi|_\infty\le C|\theta-t||\xi|\,.$$

Proof. We denote $w=S^a(\tau,\theta,\phi(\cdot))$ and $\tilde{w}=S^a(\tau,\theta,\phi(\cdot)+\xi)$. Using Lemma 4.6.1, we have that, for $(t,x,p)\in[\tau,\theta]\times\mathbb{R}^d\times[0,1]$,

$$w(t,x,p)=Y_t\ \text{ and }\ \tilde{w}(t,x,p)=\hat{Y}_t$$

where (Y,Z) and $(\hat{Y},\hat{Z})$ are solutions to, respectively,

$$\begin{aligned}Y_r&=\phi(X^{t,x}_\theta,\tilde{P}^{t,p,a}_\theta)+\int_r^T f(s,X^{t,x}_s,Y_s,Z_s)\mathrm{d}s-\int_r^\theta Z_s\mathrm{d}W_s\ ,\ t\le r\le\theta\,,\\ \hat{Y}_r&=\phi(X^{t,x}_\theta,\tilde{P}^{t,p,a}_\theta)+\xi+\int_r^T f(s,X^{t,x}_s,\hat{Y}_s,\hat{Z}_s)\mathrm{d}s-\int_r^\theta \hat{Z}_s\mathrm{d}W_s\ ,\ t\le r\le\theta\,.\end{aligned}$$

Denoting $\Gamma:=Y+\xi$ and $f_\xi(t,x,y,z)=f(t,x,y-\xi,z)$, one observes then that (Γ,Z) is the solution to

$$\Gamma_r=\phi(X^{t,x}_\theta,\tilde{P}^{t,p,a}_\theta)+\xi+\int_r^T f_\xi(s,X^{t,x}_s,\Gamma_s,Z_s)\mathrm{d}s-\int_r^\theta Z_s\mathrm{d}W_s\ ,\ t\le r\le\theta\,.$$

Let $\Delta:=\Gamma-\hat{Y}$, $\delta Z=Z-\hat{Z}$ and $\delta f_s=f_\xi(s,X^{t,x}_s,\Gamma_s,Z_s)-f(s,X^{t,x}_s,\Gamma_s,Z_s)$, for $s\in[t,\theta]$. We then get

$$\Delta_r:=\int_r^\theta\left(f(s,X^{t,x}_s,\Gamma_s,Z_s)-f(s,X^{t,x}_s,Y_s,Z_s)+\delta f_s\right)\mathrm{d}s-\int_r^\theta\delta Z_s\mathrm{d}W_s\,.$$

Classical energy estimates for BSDEs [EKPQ97; CR18] lead to

$$\mathbb{E}\left[\sup_{r\in[t,\theta]}|\Delta_r|^2\right]\le C\mathbb{E}\left[\int_t^\theta|\Delta_s\delta f_s|\mathrm{d}s\right]\,. \tag{4.2.15}$$

Next, we compute

$$\int_t^\theta|\Delta_s\delta f_s|\mathrm{d}s\le\frac{1}{2C}\sup_{s\in[t,\theta]}|\Delta_s|^2+2C\left(\int_t^\theta|\delta f_s|\mathrm{d}s\right)^2\,.$$

Combining the previous inequality with (4.2.15), we obtain

$$\mathbb{E}\left[\sup_{r\in[t,\theta]}|\Delta_r|^2\right]\le 4C^2\mathbb{E}\left[\left(\int_t^\theta|\delta f_s|\right)^2\right]\,.$$

Using the Lipschitz property of f, we get from the definition of f_ξ,

$$|\delta f_s| \leq L\xi ,$$

which eventually leads to

$$\mathbb{E}\left[\sup_{r\in[t,\theta]} |\Delta_r|^2\right] \leq C|\theta - t|^2\xi^2$$

and concludes the proof. □

Lemma 4.2.3. *Let $0 \leq \tau < \theta \leq T$ and $\phi \in \mathcal{C}^\infty([0,T] \times \mathbb{R}^d \times [0,1])$. For $(t,x,p) \in [\tau,\theta) \times \mathbb{R}^d \times (0,1)$,*

$$\phi(t,x,p) - S^a(\tau,\theta,\phi(\theta,\cdot))(t,x,p) - (\theta - t)G^a\phi(t,x,p) = o(\theta - t) ,$$

where $G^a\phi(t,x,p) := -\partial_t\phi(t,x,p) + F^a(t,x,p,\phi,D\phi,D^2\phi)$.

Proof. We first observe that $S^a(\tau,\theta,\phi(\cdot))(t,x,p) = Y_t$, with (Y^a, Z^a) the solution to

$$Y_r = \Phi_\theta + \int_r^\theta f(s, X_s^{t,x}, Y_s, Z_s)\mathrm{d}s - \int_r^\theta Z_s \mathrm{d}W_s$$

with, for $t \leq s \leq \theta$,

$$\Phi_s = \phi(s, X_s^{t,x}, P_s^{t,p,\alpha}) \text{ and } \alpha := a\mathbf{1}_{[0,\tau]} .$$

By a direct application of Itô's formula, we observe that

$$\Phi_r = \Phi_\theta - \int_r^\theta \{\partial_t\phi + \mathcal{L}^\alpha\phi\}(s, X_s^{t,x}, P_s^{t,p,\alpha})\mathrm{d}s - \int_r^\theta \mathfrak{Z}_s \mathrm{d}W_s , \quad t \leq r \leq \theta ,$$

where $\mathfrak{Z}_s := \mathfrak{z}(X_s^{t,x}, D\phi(s, X_s^{t,x}, P_s^{t,p,a}), \alpha_s)$, $t \leq s \leq \theta$.

For ease of exposition, we also introduce an "intermediary" process $(\hat{Y}, \hat{Z})$ as the solution to

$$\hat{Y}_r = \Phi_\theta + \int_r^\theta f(s, X_s^{t,x}, \Phi_s, \mathfrak{Z}_s)\mathrm{d}s - \int_r^\theta \hat{Z}_s \mathrm{d}W_s , \quad t \leq r \leq \theta .$$

Now, we compute

$$\hat{Y}_t - \Phi_t + (\theta - t)G^a\phi(t,x,p) =$$
$$\mathbb{E}\left[\int_t^\theta \left(\{\partial_t\phi(s, X_s^{t,x}, P_s^{t,p,a}) - \partial_t\phi(t,x,p)\} + \{F^a\phi(s, X_s^{t,x}, P_s^{t,p,a}) - F^a\phi(t,x,p)\}\right) \mathrm{d}s\right] .$$

Using the smoothness of ϕ, the Lipschitz property of f and the following control

$$\mathbb{E}\left[|X_s^{t,x} - x| + |P_s^{t,p,\alpha} - p|\right] \leq C_a|\theta - t|^{\frac{1}{2}} ,$$

we obtain

$$|\hat{Y}_t - \Phi_t + (\theta - t)G^a\phi(t,x,p)| \leq C_{a,\phi}(\theta - t)^{\frac{3}{2}} . \tag{4.2.16}$$

We also have

$$\hat{Y}_r - \Phi_r = \int_r^\theta G^a\phi(s, X_s^{t,x}, P_s^{t,p,\alpha})\mathrm{d}s - \int_r^\theta (\hat{Z}_s - \mathfrak{Z}_s)\mathrm{d}W_s\,.$$

Applying classical energy estimates for BSDEs, we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[t,\theta]} |\hat{Y}_r - \Phi_r|^2 + \int_t^\theta |\hat{Z}_s - \mathfrak{Z}_s|^2\mathrm{d}s\right] &\le C\mathbb{E}\left[\left(\int_t^\theta |G^a\phi(s, X_s^{t,x}, P_s^{t,p,\alpha}|\mathrm{d}s\right)^2\right]\\ &\le C_{a,\phi}(\theta - t)^2\,,\end{aligned} \tag{4.2.17}$$

where in the last line we used the smoothness of ϕ and the linear growth of f and σ. We also observe that

$$\hat{Y}_r - Y_r = \int_r^\theta \{\delta f_s + f(s, X_s^{t,x}, \hat{Y}_s, \hat{Z}_s) - f(s, X_s^{t,x}, Y_s, Z_s)\}\mathrm{d}s - \int_r^\theta \{\hat{Z}_s - Z_s\}\mathrm{d}W_s\,,$$

where $\delta f_s := f(s, X_s^{t,x}, \Phi_s, \mathfrak{Z}_s) - f(s, X_s, \hat{Y}_s, \hat{Z}_s)$, for $t \le s \le \theta$. Once again, from classical energy estimates [EKPQ97; CR18], we obtain

$$|\hat{Y}_t - Y_t|^2 \le C\mathbb{E}\left[\left(\int_t^\theta \delta f_s\mathrm{d}s\right)^2\right]\,.$$

Using the Cauchy-Schwarz inequality and the Lipschitz property of f,

$$|\hat{Y}_t - Y_t|^2 \le C(\theta - t)\mathbb{E}\left[\sup_{r\in[t,\theta]} |\hat{Y}_r - \Phi_r|^2 + \int_t^\theta |\hat{Z}_s - \mathfrak{Z}_s|^2\mathrm{d}s\right]\,.$$

This last inequality, combined with (4.2.17), leads to

$$|\hat{Y}_t - Y_t| \le C(\theta - t)^{\frac{3}{2}}\,.$$

The proof is concluded by combining the above inequality with (4.2.16). □

Finally, we can prove the following consistency property.

Proposition 4.2.4 (Consistency)**.** *Let $\phi \in \mathcal{C}^\infty([0,T] \times \mathbb{R}^d \times [0,1])$. For $(t,x,p) \in [0,T) \times \mathbb{R}^d \times (0,1)$,*

$$\left|\frac{1}{t_\pi^+ - t}\mathfrak{S}\big(\pi, t, x, p, \phi(t,x,p) + \xi, \phi(\cdot) + \xi\big) + \partial_t\phi - \mathcal{F}_n(t,x,p,\phi,D\phi,D^2\phi)\right| \to 0$$

as $(|\pi|, \xi) \to 0$.

Proof. We first observe that by Lemma 4.2.2, it is sufficient to prove

$$\left|\frac{1}{t_\pi^+ - t}\mathfrak{S}\big(\pi, t, x, p, \phi(t,x,p), \phi(\cdot)\big) + \partial_t\phi - \mathcal{F}_n(t,x,p,\phi,D\phi,D^2\phi)\right| \underset{|\pi|\downarrow 0}{\longrightarrow} 0$$

We have that

$$
\begin{aligned}
&\left| \frac{1}{t_\pi^+ - t} \mathfrak{S}\big(\pi, t, x, p, \phi(t,x,p), \phi(\cdot)\big) + \partial_t \phi - \mathcal{F}_n(t, x, p, \phi, D\phi, D^2\phi) \right| \\
&= \left| \frac{1}{t_\pi^+ - t} \{\phi(t,x,p) - \min_{a\in K} S^a\left(t_\pi^-, t_\pi^+, \phi(t_\pi^+, \cdot)\right)(t,x,p)\} - \max_{a\in K} G^a(t,x,p)\phi \right| \\
&\leq \max_{a\in K} \left| \frac{1}{t_\pi^+ - t} \{\phi(t,x,p) - S^a\left(t_\pi^-, t_\pi^+, \phi(t_\pi^+, \cdot)\right)(t,x,p)\} - G^a \phi \right| .
\end{aligned}
$$

The proof is then concluded by applying Lemma 4.2.3. □

To conclude this section, let us observe that we obtain the following result, combining Proposition 4.2.1 and Theorem 4.2.1.

Corollary 4.2.1. *In the setting of this section, assuming (H), we have in $\mathcal{C}^0([0,T] \times \mathbb{R}^d)$,*

$$
\lim_{n\to\infty} \lim_{|\pi|\downarrow 0} v_{n,\pi} = v \, .
$$

Remark 4.2.4. *An important question, from a numerical perspective, is to understand how to fix the parameters n and π in relation to each other. The theoretical difficulty here is to obtain a precise rate of convergence for the approximations given in Proposition 4.2.1 and Theorem 4.2.1, along the lines of the continuous dependence estimates with respect to the control discretization in [JK05; DRZ18], and estimates of the approximation by piecewise constant controls as in [Kry99; JPR19]. To answer this question in our general setting is a challenging task, extending also to error estimates for the full discretization in the next section, which is left for further research.*

4.3 Application to the Black-Scholes model: a fully discrete monotone scheme

The goal of this section is to introduce a fully implementable scheme and to prove its convergence. The scheme is obtained by adding a finite difference approximation to the PCPT procedure described in Section 4.2.2. Then in Section 4.4, we present numerical tests that demonstrate the practical feasibility of our numerical method. From now on, we will assume that the log-price process X is a one-dimensional Brownian motion with drift, for $(t,x) \in [0,T] \times \mathbb{R}$:

$$
X_s^{t,x} = x + \mu(s-t) + \sigma(W_s - W_t), \qquad s \in [t,T],
$$

with $\mu \in \mathbb{R}$ and $\sigma > 0$.

This restriction to Black-Scholes is not essential, as the main difficulty and non-linearities are already present in this case and the analysis technique can be extended straightforwardly to more general monotone schemes in the setting of more complex SDEs for X. We take advantage of the specific dynamics to design a simple to implement numerical scheme, which also simplifies the notation.

We shall moreover work under the following hypothesis.

Assumption 4.3.1. *The coefficient μ is non-negative.*

Remark 4.3.1. *This assumption is introduced without loss of generality in order to alleviate the notation in the definition of the scheme. We detail in Remark 4.3.2(ii) how to modify the scheme for non-positive drift μ. The convergence properties are the same.*

We now fix $n \geq 1$, $\mathcal{R}_n$ the associated discrete set of controls (see Section 4.2.1). We denote $K = \mathcal{R}_n^\flat$ assuming that $0 \notin K$ and recall that v_n is the solution to (4.2.5). We consider the grid $\pi = \{0 =: t_0 < \cdots < t_k < \cdots < t_\kappa := T\}$ on $[0, T]$ and approximate v_n by a PCPT scheme, extending Section 4.2.2.

The main point here is that we introduce a finite difference approximation for the solution $S^a(\cdot)$, $a \in K$ to (4.2.6)–(4.2.7). This approximation, denoted by $S^a_\delta(\cdot)$ for a parameter $\delta > 0$, will be specified in Section 4.3.1 below. For $\delta > 0$ and $a \in K$, each approximation $S^a_\delta(\cdot)$ is defined on a spatial grid

$$\mathcal{G}^a_\delta := \delta\mathbb{Z} \times \Gamma^a_\delta \subset \mathbb{R} \times [0, 1].$$

where Γ^a_δ is a uniform grid on $[0, 1]$, with $N^a_\delta + 1$ points and mesh size $1/N^a_\delta$. An element of $\mathcal{G}^a_\delta$ is denoted by $(x_k, p_l) := (k\delta, l/N^a_\delta)$, and an element of $\ell^\infty(\mathcal{G}^a_\delta)$ by $u_{k,l} := u(x_k, p_l)$, for all $k \in \mathbb{Z}$ and $0 \leq l \leq N^a_\delta$. For $0 \leq t < s \leq T$, and $\varphi : \delta\mathbb{Z} \times [0, 1] \to \mathbb{R}$ a bounded function, we have that $S^a_\delta(s, t, \varphi) \in \ell^\infty(\mathcal{G}^a_\delta)$.

In order to define our approximation of v_n, it is not enough to replace $S^a(\cdot)$ in the minimisation (4.2.9), or similarly (4.2.11), by $S^a_\delta(\cdot)$, as the approximations are normally not defined on the same grid for the p-variable. This flexibility of having different grids will be important for the construction of the schemes later on. We thus have to consider a supplementary step which consists in a linear interpolation in the p-variable, namely, any mapping $u \in \ell^\infty(\mathcal{G}^a_\delta)$ is extended into $\mathcal{I}^a_\delta(u) : \delta\mathbb{Z} \times [0, 1] \to \mathbb{R}$ by linear interpolation in the second variable: if $u \in \ell^\infty(\mathcal{G}^a_\delta)$, $k \in \mathbb{Z}$ and $p \in [p_l, p_{l+1})$ with $0 \leq l < N^a_\delta$,

$$\mathcal{I}^a_\delta(x_k, p) = \frac{p_{l+1} - p}{p_{l+1} - p_l} u_{k,l} + \frac{p - p_l}{p_{l+1} - p_l} u_{k,l+1},$$

and obviously $\mathcal{I}^a_\delta(x_k, 1) = u_{k,N^a_\delta}$.

The solution to the numerical scheme associated with π, δ is then $v_{n,\pi,\delta} : \pi \times \delta\mathbb{Z} \times [0, 1] \to \mathbb{R}$ satisfying

$$\widehat{\mathfrak{S}}(\pi, \delta, t, x, p, v_{n,\pi,\delta}(t, x, p), v_{n,\pi,\delta}) = 0,$$

where, for any $0 \leq t \in \pi, x \in \delta\mathbb{Z}, p \in [0, 1], y \in \mathbb{R}^+$ and any bounded function $u : \pi \times \delta\mathbb{Z} \times [0, 1] \to \mathbb{R}$:

$$\widehat{\mathfrak{S}}(\pi, \delta, t, x, p, y, u) = \begin{cases} y - \min_{a \in K} \mathcal{I}^a_\delta \left(S^a_\delta \left(t^+_\pi, t_k, u(t^+_\pi, \cdot) \right) \right) (t_k, x, p) & \text{if } k < \kappa, \\ y - g(x)p & \text{otherwise,} \end{cases} \tag{4.3.1}$$

where $t^+_\pi = \inf\{s \in \pi : s \geq t\}$.

Alternatively, $v_{n,\pi,\delta}$ is defined by the following backward induction:

1. Initialisation: set $v_{n,\pi,\delta}(T, x, p) := g(x)p$, $x \in \mathbb{R}^d \times [0, 1]$.

2. Backward step: For $k = \kappa - 1, \ldots, 0$, compute $w_\delta^{k,a} := S_\delta^a(t_k, t_{k+1}, v_{n,\pi,\delta}(t_{k+1}, \cdot))$ and set, for $(x,p) \in \delta\mathbb{Z} \times [0,1]$,

$$v_{n,\pi,\delta}(t_k, x, p) := \inf_{a \in K} \mathcal{I}_\delta^a(w_\delta^{k,a})(t_k, x, p)\ .$$

Before stating the main convergence result of this section, see Theorem 4.3.1 below, we give the precise definition of $S_\delta^a(\cdot)$ using finite difference operators.

4.3.1 Finite difference scheme and convergence result

Let $0 \leq t < s \leq T, \delta > 0, \varphi : \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$. We set $h := s - t > 0$. For $a \in K$, we will describe the grid $\mathcal{G}_\delta^a = \delta\mathbb{Z} \times \Gamma_\delta^a \subset \delta\mathbb{Z} \times [0,1]$ and the finite difference scheme used to define S_δ^a.

First, we observe that for the model of this section, (4.2.5) can be rewritten as

$$\sup_{a \in K} \left(-D^a\varphi - \mu\nabla^a\varphi - \frac{\sigma^2}{2}\Delta^a\varphi - f(t, x, \varphi, \sigma\nabla^a\varphi) \right) = 0, \tag{4.3.2}$$

with:

$$\begin{aligned} \nabla^a\varphi &:= \partial_y\varphi + \frac{a}{\sigma}\partial_p\varphi, \\ \Delta^a\varphi &:= \partial^2_{yy}\varphi + 2\frac{a}{\sigma}\partial^2_{yp}\varphi + \frac{a^2}{\sigma^2}\partial^2_{pp}\varphi, \\ D^a\varphi &:= \partial_t\varphi - \frac{a}{\sigma}\mu\partial_p\varphi. \end{aligned} \tag{4.3.3}$$

Exploiting the degeneracy of the operators ∇^a and Δ^a in the direction $(a, -\sigma)$, we construct Γ_δ^a so that the solution to (4.3.2) is approximated by the solution of an implicit finite difference scheme with only one-directional derivatives.

To take into account the boundaries $p = 0, p = 1$, we set

$$N_\delta^a := \min\left\{ j \geq 1 : j\frac{|a|}{\sigma}\delta \geq 1 \right\} = \left\lceil \frac{\sigma}{|a|\delta} \right\rceil \tag{4.3.4}$$

and

$$\mathfrak{a}(a, \delta) := \operatorname{sgn}(a)\frac{\sigma}{\delta N_\delta^a}, \tag{4.3.5}$$

where $a \neq 0$. We have $N_\delta^a = \sigma/\delta|\mathfrak{a}(a,\delta)|$. We finally set:

$$\Gamma_\delta^a := \left\{ 0, \frac{|\mathfrak{a}(a,\delta)|}{\sigma}\delta, \ldots, N_\delta^a\frac{|\mathfrak{a}(a,\delta)|}{\sigma}\delta = 1 \right\} = \left\{ \frac{j}{N_\delta^a}\ :\ j = 0, \ldots, N_\delta^a \right\}.$$

Using the degeneracy of $\nabla^{\mathfrak{a}(a,\delta)}$ and $\Delta^{\mathfrak{a}(a,\delta)}$ in the direction $(\mathfrak{a}(a,\delta), -\sigma)$, we define the following finite difference operators, for $v = (v_{k,l})_{k \in \mathbb{Z}, 0 \leq l \leq N_\delta^a} = (v(x_k, p_l))_{k \in \mathbb{Z}, 0 \leq l \leq N_\delta^a} \in \ell^\infty(\mathcal{G}_\delta^a)$ and $w = (w_k)_{k \in \mathbb{Z}} \in \ell^\infty(k\mathbb{Z})$:

$$\begin{aligned} \nabla_\delta^a v_{k,l} &:= \frac{1}{2\delta}\left(v_{k+1,l+\operatorname{sgn}(a)} - v_{k-1,l-\operatorname{sgn}(a)}\right), & \nabla_\delta w_k &:= \frac{1}{2\delta}\left(w_{k+1} - w_{k-1}\right), \\ \nabla_{+,\delta}^a v_{k,l} &:= \frac{1}{\delta}\left(v_{k+1,l+\operatorname{sgn}(a)} - v_{k,l}\right), & \nabla_{+,\delta} w_k &:= \frac{1}{\delta}\left(w_{k+1} - w_k\right), \\ \Delta_\delta^a v_{k,l} &:= \frac{1}{\delta^2}\left(v_{k+1,l+\operatorname{sgn}(a)} + v_{k-1,l-\operatorname{sgn}(a)} - 2v_{k,l}\right), & \Delta_\delta w_k &:= \frac{1}{\delta^2}\left(w_{k+1} + w_{k-1} - 2w_k\right). \end{aligned}$$

Let $\theta > 0$ a parameter to be fixed later. We define, for $(t, x, y, q, q_+, A) \in [0, T] \times \mathbb{R}^5$,

$$F(t, x, y, q, A) := -\mu q - \frac{\sigma^2}{2} A - f(t, x, y, \sigma q), \text{ and} \tag{4.3.6}$$

$$\widehat{F}(t, x, y, q, q_+, A) := -\mu q_+ - \left(\frac{\sigma^2}{2} + \theta \frac{\delta^2}{h}\right) A - f(t, x, y, \sigma q)\,. \tag{4.3.7}$$

Now, $S^a_\delta(s, t, \varphi) \in \ell^\infty(\mathcal{G}^a_\delta)$ is defined as the unique solution to (see Proposition 4.3.1 below for the well-posedness of this definition)

$$S(k, l, v_{k,l}, \nabla^a_\delta v_{k,l}, \nabla^a_{+,\delta} v_{k,l}, \Delta^a_\delta v_{k,l}, \varphi) = 0, \tag{4.3.8}$$

$$v_{k,0} = \underline{v}_k, v_{k,N^a_\delta} = \overline{v}_k, \tag{4.3.9}$$

where

$$S(k, l, v, q, q_+, A, u) = v - u\left(x_k, \mathfrak{p}^a(p_l)\right) + h\widehat{F}(t, k\delta, v, q, q_+, A), \tag{4.3.10}$$

for $k \in \mathbb{Z}, 0 < l < N^a_\delta, (v, v_+, v_-) \in \mathbb{R}^3$, and any bounded $u : \delta\mathbb{Z} \times [0, 1] \to \mathbb{R}$, with

$$\mathfrak{p}^a(p) := p - \mu \frac{\mathfrak{a}(a, \delta)}{\sigma} h \tag{4.3.11}$$

for $p \in [0, 1]$, and where $(\underline{v}_k)_{k\in\mathbb{Z}}$ (resp. $(\overline{v}_k)_{k\in\mathbb{Z}}$) is the solution to

$$S_b(k, \underline{v}_k, \nabla_\delta \underline{v}_k, \nabla_{+,\delta} \underline{v}_k, \Delta_\delta \underline{v}_k, (\underline{\varphi}_k)_{k\in\mathbb{Z}}) = 0, \tag{4.3.12}$$

$$(\text{resp. } S_b(k, \overline{v}_k, \nabla_\delta \overline{v}_k, \nabla_{+,\delta} \overline{v}_k, \Delta_\delta \overline{v}_k, (\overline{\varphi}_k)_{k\in\mathbb{Z}}) = 0) \tag{4.3.13}$$

with $\underline{\varphi}_k = \varphi(k\delta, 0)$ (resp. $\overline{\varphi}_k = \varphi(k\delta, N^a_\delta)$) and, for $k \in \mathbb{Z}, (v, v_+, v_-) \in \mathbb{R}^3, u \in \ell^\infty(\mathbb{Z})$:

$$S_b(k, v, q, q_+, A, u) = v - u_k + h\widehat{F}(t, k\delta, v, q, q_+, A). \tag{4.3.14}$$

Remark 4.3.2. *(i) Here, as stated before, we have assumed $\mu \geq 0$. If the opposite is true, one has to consider $\nabla^a_-(\delta) v_{k,l} := \frac{1}{\delta}\left(v_{k,l} - v_{k-1,l-\mathrm{sgn}(a)}\right)$ (resp. $\nabla_-(\delta) w_k := \frac{1}{\delta}(w_k - w_{k-1})$) instead of $\nabla^a_{+,\delta} v_{k,l}$ (resp. $\nabla_{+,\delta} w_k$), in the definition of $S^a_\delta(s, t, \varphi)$ (resp. $\underline{v}_k, \overline{v}_k$), to obtain a monotone scheme.*
(ii) For the non-linearity f, we used the Lax-Friedrichs scheme [CL84; DRZ18], adding the term $\theta(v_+ + v_- - 2v)$ term in the definition of $\widehat{F}$ to enforce monotonicity.

We now assume that the following conditions on the parameters are satisfied:

$$\delta \leq 1, \tag{4.3.15}$$

$$\frac{hL}{2\delta} \leq \theta < \frac{1}{4},$$

$$\mu h \leq \delta \leq Mh, \tag{4.3.16}$$

for a constant $M > 0$ independent of h and δ. Under these conditions, we prove that $S^a_\delta(s, t, \varphi)$ is uniquely defined, and that it can be obtained by Picard iteration.

Remark 4.3.3. *Since $\mu h \leq \delta$, we have $|\mu \frac{\mathfrak{a}(a,\delta)}{\sigma} h| \leq \frac{|\mathfrak{a}(a,\delta)|}{\sigma} \delta$, which ensures that from (4.3.11), $\mathfrak{p}^a(p_l) \in [0, 1]$ for all $0 < l < N^a_\delta$.*

Proposition 4.3.1. *For every bounded function $\varphi : \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$, there exists a unique solution to (4.3.8)–(4.3.9).*

Proof. First, $\underline{v} \in \ell(\delta\mathbb{Z})$ (resp. $\overline{v} \in \ell(\delta\mathbb{Z})$) is uniquely defined by (4.3.12) (resp. (4.3.13)), see Proposition 4.6.2.
We consider the following map:

$$\ell^\infty(\mathcal{G}^a_\delta) \to \ell^\infty(\mathcal{G}^a_\delta),$$
$$v \mapsto \psi(v),$$

where $\psi(v)$ is defined by, for $k \in \mathbb{Z}$ and $l \in \{1, \dots, N_a - 1\}$:

$$\begin{aligned}\psi(v)_{k,l} \ = & \frac{1}{1 + \frac{h}{\delta}\mu + \sigma^2\frac{h}{\delta^2} + 2\theta}\left(\varphi\left(k\delta, \mathfrak{p}^a(p_l)\right) + \right. \\ & \frac{h}{\delta}\mu v_{k+1,l+\mathrm{sgn}(a)} + \frac{\sigma^2}{2}\frac{h}{\delta^2}(v_{k+1,l+\mathrm{sgn}(a)} + v_{k-1,l-\mathrm{sgn}(a)}) + \\ & \left. hf\left(t^-, k\delta, v_{k,l}, \frac{\sigma}{2\delta}(v_{k+1,l+\mathrm{sgn}(a)} - v_{k-1,l-\mathrm{sgn}(a)})\right) + \theta(v_{k+1,l+\mathrm{sgn}(a)} + v_{k-1,l-\mathrm{sgn}(a)})\right),\end{aligned}$$

$$\psi(v)_{k,0} = \underline{v}_k, \psi(v)_{k,N_a} = \overline{v}_k.$$

Notice that v is a solution to (4.3.8)-(4.3.9) if and only if v is a fixed point of ψ. It is now enough to show that ψ maps $\ell^\infty(\mathcal{G}^a_\delta)$ into $\ell^\infty(\mathcal{G}^a_\delta)$ and is contracting. If $v \in \ell^\infty(\mathcal{G}^a_\delta)$, by boundedness of $\varphi, \underline{v}$ and $\overline{v}$, it is clear that $\psi(v)$ is bounded. If $v^1, v^2 \in \ell^\infty(\mathcal{G}^a_\delta)^2$, we have, for all $k \in \mathbb{Z}$ and $1 \le l \le N_a - 1$:

$$|\psi(v^1)_{k,l} - \psi(v^2)_{k,l}| \le \frac{\frac{h}{\delta}\mu + \sigma^2\frac{h}{\delta^2} + 2\theta + hL + \frac{hL}{\delta}}{1 + \frac{h}{\delta}\mu + \sigma^2\frac{h}{\delta^2} + 2\theta}|v^1 - v^2|_\infty.$$

Since $\delta \le 1$ by assumption (4.3.15), one has $hL + \frac{hL}{\delta} \le 2\frac{hL}{\delta} \le 4\theta$, thus:

$$|\psi(v^1) - \psi(v^2)|_\infty \le \frac{4\theta + \frac{h}{\delta}\mu + \sigma^2\frac{h}{\delta^2} + 2\theta}{1 + \frac{h}{\delta}\mu + \sigma^2\frac{h}{\delta^2} + 2\theta}|v^1 - v^2|_\infty.$$

Since $4\theta < 1$ by assumption (4.3.1) and the function $x \mapsto \frac{4\theta + x}{1+x}$ is increasing on $[0, \infty)$ with limit 1 when $x \to +\infty$, this proves that ψ is a contracting mapping. $\square$

For this scheme, we have the following strong uniqueness result:

Proposition 4.3.2. *Let $\varphi^1, \varphi^2 : \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$ two bounded functions satisfying $\varphi^1 \le \varphi^2$ on $\delta\mathbb{Z} \times [0,1]$.*

1. *(Monotonicity) For all $k \in \mathbb{Z}$, $1 \le l \le N_a$, $(v, q, q_+, A) \in \mathbb{R}^4$, we have:*
$$S(k, l, v, q, q_+, A, \varphi^2) \le S(k, l, v, q, q_+, A, \varphi^1).$$

2. *(Comparison theorem) Let $(v^1, v^2) \in \ell^\infty(\mathcal{G}^a_\delta)^2$ satisfy:*
$$\begin{aligned} S(k, l, v^1_{k,l}, \nabla^a_\delta v^1_{k,l}, \nabla^a_{+,\delta} v^1_{k,l}, \Delta^a_\delta v^1_{k,l}, \varphi^2) & \\ \le S(k, l, v^2_{k,l}, \nabla^a_\delta v^2_{k,l}, \nabla^a_{+,\delta} v^2_{k,l}, \Delta^a_\delta v^2_{k,l}, \varphi^2) & \quad (4.3.17) \\ v^1_{k,0} \le v^2_{k,0}, & \\ v^1_{k,N^a_\delta} \le v^2_{k,N^a_\delta} & \end{aligned}$$
for all $k \in \mathbb{Z}$ and $1 \le l \le N^a_\delta - 1$. Then $v^1 \le v^2$.

3. *We have* $S^a_\delta(s,t,\varphi^1)_{k,l} \le S^a_\delta(s,t,\varphi^2)_{k,l}$ *for all* $k \in \mathbb{Z}$ *and* $0 \le l \le N^a_\delta$.

Proof. Let φ^1, φ^2 as stated in the proposition.

1. We have, for $k \in \mathbb{Z}$ and $0 < l < N^a_\delta$:

$$\begin{aligned} &S(k,l,v,q,q_+,A,\varphi^2) - S(k,l,v,q,q_+,A,\varphi^1) \\ &= (\varphi^1 - \varphi^2)\,(x_k, \mathfrak{p}^a(p_l)) \le 0. \end{aligned}$$

2. We assume here that $a > 0$. For $k \in \mathbb{Z}$, let $M_k = \max_{0 \le l \le N^a_\delta}(v^1_{k+l,l} - v^2_{k+l,l}) < \infty$ (if $a < 0$, we have to consider $\max_{0\le l\le N^a_\delta}(v^1_{k-l,l} - v^2_{k-l,l})$). We want to prove that $M_k \le 0$ for all k. Assume to the contrary that there exists $k \in \mathbb{Z}$ such that $M_k > 0$. Then there exists $0 \le l \le N^a_\delta$ such that

$$v^1_{k+l,l} - v^2_{k+l,l} = M_k > 0. \tag{4.3.18}$$

First, we have $v^1_{k,0} \le v^2_{k,0}$ and $v^1_{k+N^a_\delta, N^a_\delta} \le v^2_{k+N^a_\delta, N^a_\delta}$. Thus $0 < l < N^a_\delta$. Moreover, using (4.3.17), re-arranging the terms, using the fact that f is non-increasing with respect to its third variable and Lipschitz-continuous, by (4.3.18),

$$\begin{aligned} (1+2\theta)M_k \le \frac{hL}{2\delta}\left|v^2_{k+l+1,l+1} - v^1_{k+l+1,l+1}\right| - \theta(v^2_{k+l+1,l+1} - v^1_{k+l+1,l+1}) + \\ \frac{hL}{2\delta}\left|v^2_{k+l-1,l-1} - v^1_{k+l-1,l-1}\right| - \theta(v^2_{k+l-1,l-1} - v^1_{k+l-1,l-1}). \end{aligned} \tag{4.3.19}$$

For $j \in \{l-1, l+1\}$, we observe that

$$\frac{hL}{2\delta}|v^2_{k+j,j} - v^1_{k+j,j}| - \theta(v^2_{k+j,j} - v^1_{k+j,j}) \le \left(\frac{hL}{2\delta} + \theta\right) M_k. \tag{4.3.20}$$

Indeed, if $v^2_{k+j,j} \ge v^1_{k+j,j}$ then

$$\frac{hL}{2\delta}|v^2_{k+j,j} - v^1_{k+j,j}| - \theta(v^2_{k+j,j} - v^1_{k+j,j}) = \left(\frac{hL}{2\delta} - \theta\right)(v^2_{k+j,j} - v^1_{k+j,j}) \le 0,$$

since $\frac{hL}{2\delta} \le \theta$. Otherwise, if $v^2_{k+j,j} < v^1_{k+j,j}$

$$\begin{aligned} \frac{hL}{2\delta}|v^2_{k+j,j} - v^1_{k+j,j}| - \theta(v^2_{k+j,j} - v^1_{k+j,j}) &= \left(\frac{hL}{2\delta} + \theta\right)(v^1_{k+j,j} - v^2_{k+j,j}), \\ &\le \left(\frac{hL}{2\delta} + \theta\right) M_k. \end{aligned}$$

Inserting (4.3.20) into (4.3.19), we get

$$(1+2\theta)M_k \le 2\left(\frac{hL}{2\delta} + \theta\right) M_k.$$

Thus,

$$\left(1 - \frac{hL}{\delta}\right) M_k \le 0,$$

which is a contradiction to $M_k > 0$ since $\frac{hL}{\delta} \le 2\theta < \frac{1}{2}$.

3. Let $v^i = S^a_\delta(s,t,\varphi^i)$ for $i = 1,2$. Since $\underline{\varphi}^1 \le \underline{\varphi}^2$ and $\overline{\varphi}^1 \le \overline{\varphi}^2$, we get by Proposition 4.6.2 that $v^1_{k,0} \le v^2_{k,0}$ and $v^1_{k,N^a_\delta} \le v^2_{k,N^a_\delta}$ for all $k \in \mathbb{Z}$.

   By monotonicity, we get, for all $k \in \mathbb{Z}$ and $0 < l < N^a_\delta$,
$$S(k,l,v^1_{k,l},\nabla^a_\delta v^1_{k,l},\nabla^a_{+,\delta} v^1_{k,l},\Delta^a_\delta v^1_{k,l},\varphi^2) \\ \le S(k,l,v^1_{k,l},\nabla^a_\delta v^1_{k,l},\nabla^a_{+,\delta} v^1_{k,l},\Delta^a_\delta v^1_{k,l},\varphi^1)$$
   Moreover,
$$S(k,l,v^1_{k,l},\nabla^a_\delta v^1_{k,l},\nabla^a_{+,\delta} v^1_{k,l},\Delta^a_\delta v^1_{k,l},\varphi^1) \\ = S(k,l,v^2_{k,l},\nabla^a_\delta v^2_{k,l},\nabla^a_{+,\delta} v^2_{k,l},\Delta^a_\delta v^2_{k,l},\varphi^2) = 0$$
   So that,
$$S(k,l,v^1_{k,l},\nabla^a_\delta v^1_{k,l},\nabla^a_{+,\delta} v^1_{k,l},\Delta^a_\delta v^1_{k,l},\varphi^2) \\ \le S(k,l,v^2_{k,l},\nabla^a_\delta v^2_{k,l},\nabla^a_{+,\delta} v^2_{k,l},\Delta^a_\delta v^2_{k,l},\varphi^2)$$
   and the proof is concluded applying the previous point.

□

We last give a refinement of the comparison theorem, which we use in the sequel.

Proposition 4.3.3. *Let $u : \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$ be a bounded function, and let $v^1, v^2 \in \ell^\infty(\mathcal{G}^a_\delta)$. Assume that, for all $k \in \mathbb{Z}$ and $0 < l < N^a_\delta$, we have*
$$S(k,l,v^1_{k,l},\nabla^a_\delta v^1_{k,l},\nabla^a_{+,\delta} v^1_{k,l},\Delta^a_\delta v^1_{k,l},u) \\ \le 0 \le S(k,l,v^2_{k,l},\nabla^a_\delta v^2_{k,l},\nabla^a_{+,\delta} v^2_{k,l},\Delta^a_\delta v^2_{k,l},u).$$
Then:
$$v^1_{k,l} - v^2_{k,l} \le e^{-4\frac{\mathfrak{a}(a,\delta)^2}{\sigma^2}C(h,\delta)l(N^a_\delta - l)}\left(|(v^1_{\cdot,0} - v^2_{\cdot,0})^+|_\infty + |(v^1_{\cdot,N^a_\delta} - v^2_{\cdot,N^a_\delta})^+|_\infty\right),$$
where
$$C(h,\delta) := \frac{\ln\left(\frac{1+\frac{h}{\delta}\mu+\sigma^2\frac{h}{\delta^2}+2\theta+\frac{hL}{2\delta}}{\frac{h}{\delta}\mu+\sigma^2\frac{h}{\delta^2}+2\theta+\frac{hL}{2\delta}}\right)}{\delta^2}. \tag{4.3.21}$$
Moreover,
$$C(h,\delta) \ge \frac{1}{\left((\mu+\frac{L}{2})M + 2\theta M^2\right)h^2 + \sigma^2 h} - \frac{M^2}{2\sigma^4}. \tag{4.3.22}$$

Remark 4.3.4. *(i) To prove the consistency of the scheme, we define in Lemma 4.3.2 smooth functions $w^\pm$ so that $(w^\pm(x_k,p_l)) \in l^\infty(\mathcal{G}^a_\delta)$ satisfy $S \ge 0$ or $S \le 0$, but we cannot use the comparison theorem directly, as the values at the boundary cannot be controlled. The previous proposition will be used in Lemma 4.3.3 to show that the difference between $w^\pm$ and the linear interpolant of a solution of $S = 0$ is small.*

(ii) The coefficient $\exp\Big(-4\frac{\mathfrak{a}(a,\delta)}{\sigma^2}C(h,\delta)l(N^a_\delta - l\Big)$ *that appears in the first equation of the previous proposition shows that the dependance on the boundary values decays exponentially with the distance to the boundary. This was to be expected and was already observed in similar situations, see for example Lemma 3.2 in [BJ07] for Hamilton-Jacobi-Bellman equations.*

We now can state the main result of this section.

Theorem 4.3.1. *The function* $v_{n,\pi,\delta}$ *converges to* v_n *uniformly on compact sets, as* $|\pi|, \delta \to 0$ *satisfying conditions* (4.3.15)–(4.3.16)*, where* $\pi = \{0 = t_0 < t_1 < \cdots < t_\kappa = T\}$.

We prove below that the scheme is *monotone* (see Proposition 4.3.5), *stable* (see Proposition 4.3.6), *consistent* with (4.2.5) in $[0,T) \times \mathbb{R} \times (0,1)$ (see Proposition 4.3.7) and with the boundary conditions (see Proposition 4.3.4). The theorem then follows by identical arguments to [BS91].

4.3.2 Proof of Theorem 4.3.1

We first show that the numerical scheme is consistent with the boundary conditions. For any discretization parameters π, δ, we define $V_{\pi,\delta} : \pi \times \delta\mathbb{Z} \to \mathbb{R}$ as the solution to the following system, with S_b from (4.3.14):

$$\begin{aligned} S_b(k, v^j_k, \nabla_\delta v^j_k, \nabla_{+,\delta} v^j_k, \Delta_\delta v^j_k, v^{j+1}_k) &= 0, k \in \mathbb{Z}, 0 \le j < \kappa, \\ v^\kappa_k &= g(x_k), k \in \mathbb{Z}, \end{aligned} \tag{4.3.23}$$

where $v^j_k := v(t_j, x_k)$ for $0 \le j \le \kappa$ and $k \in \mathbb{Z}$. We set $(U_{\pi,\delta})^j_k := \nabla_\delta (V_{\pi,\delta})^j_k = \frac{1}{2\delta}((V_{\pi,\delta})^j_{k+1} - (V_{\pi,\delta})^j_{k-1})$. By Proposition 4.6.3 in the appendix, $V_{\pi,\delta}$ and $U_{\pi,\delta}$ are bounded, uniformly in π, δ, and, by [BS91], $V_{\pi,\delta}$ converges to V (the super-replication price) uniformly on compact sets as $|\pi| \to 0$ and $\delta \to 0$.

Proposition 4.3.4. *There exist constants* $K_1, K_2, K_3 > 0$ *such that, for all discretization parameters* π, δ *with* $|\pi|$ *small enough, we have, for* $(t_j, x_k, p) \in \pi \times \delta\mathbb{Z} \times [0,1]$*:*

$$\begin{aligned} pV_{\pi,\delta}(t_j, x_k) - K_1(T - t_j) \le v_{n,\pi,\delta}(t_j, x_k, p) &\le pV_{\pi,\delta}(t_j, x_k) + K_1(T - t_j), \\ pV_{\pi,\delta}(t_j, x_k) - (1 - e^{-K_2 p})(1 - e^{-K_2(1-p)}) &\le v_{n,\pi,\delta}(t_j, x_k, p) \\ &\le pV_{\pi,\delta}(t_j, x_k) + (1 - e^{-K_2 p})(1 - e^{-K_2(1-p)}). \end{aligned}$$

Proof. We only prove, by backward induction, the lower bounds, while the proof of the upper bounds is similar. For $0 \le j \le \kappa$, $k \in \delta\mathbb{Z}$ and $0 \le l \le N^a_\delta$, we set $V^j_k := V_{\pi,\delta}(t_j, x_k)$ and $U^j_k := U_{\pi,\delta}(t_j, x_k)$. For $\epsilon \in \{0,1\}$, we define:

$${}^\epsilon w(t_j, x_k, p) := pV^j_k - {}^\epsilon c(t_j, p),$$

with

$${}^\epsilon c(t_j, p) := \epsilon K_1(T - t_j) + (1 - \epsilon)(1 - e^{-K_2 p})(1 - e^{-K_2(1-p)}),$$

and ${}^{\epsilon}w^j_{k,l} = {}^{\epsilon}w(t_j, x_k, p_l), {}^{\epsilon}c^j_l = {}^{\epsilon}c(t_j, p_l)$, $p_l \in \Gamma^a_\delta$. The proof now proceeds in two steps.
1. First, we have ${}^{\epsilon}w(T, x_k, p) \leq pV_{\pi,\delta}(T, x_k) = pg(x_k) = v_{n,\pi,\delta}(T, x_k, p)$ on $\delta\mathbb{Z} \times [0, 1]$. Suppose that, for $0 \leq j < \kappa$, on $\delta\mathbb{Z} \times [0, 1]$, we have

$$ {}^{\epsilon}w(t_{j+1}, x_k, p) \leq v_{n,\pi,\delta}(t_{j+1}, x_k, p). $$

We want to prove on $\delta\mathbb{Z} \times [0, 1]$

$$ {}^{\epsilon}w(t_j, x_k, p) \leq v_{n,\pi,\delta}(t_j, x_k, p). $$

Since ${}^{\epsilon}w$ is convex in p, ${}^{\epsilon}w(t_j, x_k, \cdot) \leq \mathcal{I}^a_\delta({}^{\epsilon}w^j_{k,\cdot})$ on $[0, 1]$. By definition, we have $v_{n,\pi,\delta}(t_j, x_k, p) = \min_{a\in K} \mathcal{I}^a_\delta(S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot)))(x_k, p)$, we are thus going to prove

$$ {}^{\epsilon}w^j_{k,l} \leq S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot))(t_j, x_k, p_l) \tag{4.3.24} $$

for all $a \in K$ and all $k \in \mathbb{Z}, 0 \leq l \leq N^a_\delta$. For $a \in K$, by the induction hypothesis, ${}^{\epsilon}w(t_{j+1}, \cdot) \leq v_{n,\pi,\delta}(t_{j+1}, \cdot)$, so if we are able to get

$$ S_b(k, {}^{\epsilon}\underline{w}_k, \nabla_\delta {}^{\epsilon}\underline{w}_k, \nabla_{+,\delta} {}^{\epsilon}\underline{w}_k, \Delta_\delta {}^{\epsilon}\underline{w}_k, {}^{\epsilon}w^{j+1}_{k,0}) \leq 0, k \in \mathbb{Z}, \tag{4.3.25} $$
$$ S_b(k, {}^{\epsilon}\overline{w}_k, \nabla_\delta {}^{\epsilon}\overline{w}_k, \nabla_{+,\delta} {}^{\epsilon}\overline{w}_k, \Delta_\delta {}^{\epsilon}\overline{w}_k, {}^{\epsilon}w^{j+1}_{k,N^a_\delta}) \leq 0, k \in \mathbb{Z}, \tag{4.3.26} $$
$$ S(k, l, {}^{\epsilon}w_{k,l}, \nabla^a_\delta {}^{\epsilon}w_{k,l}, \nabla^a_{+,\delta} {}^{\epsilon}w_{k,l}, \Delta^a_\delta {}^{\epsilon}w_{k,l}, {}^{\epsilon}w(t_{j+1}, \cdot)) \leq 0, k \in \mathbb{Z}, 0 < l < N^a_\delta \tag{4.3.27} $$

where ${}^{\epsilon}\underline{w}^j_k = {}^{\epsilon}w(t_j, x_k, 0), {}^{\epsilon}\overline{w}^j_k = {}^{\epsilon}w(t_j, x_k, 1)$, we obtain that (4.3.24) holds true by the comparison result in Proposition 4.3.2, which concludes the proof. We now proceed with the proof of (4.3.25), (4.3.26) and (4.3.27).
2.a Now, observe that ${}^{\epsilon}\underline{w}^j_k = -\epsilon K_1(T - t_j)$, for $k \in \mathbb{Z}$. We have, since $f(t_j, x_k, 0, 0) = 0$ and f is non-increasing in its third variable,

$$ S_b(k, {}^{\epsilon}\underline{w}_k, \nabla_\delta {}^{\epsilon}\underline{w}_k, \nabla_{+,\delta} {}^{\epsilon}\underline{w}_k, \Delta_\delta {}^{\epsilon}\underline{w}_k, {}^{\epsilon}w^{j+1}_{k,0}) = -\epsilon K h - h f(t_j, x_k, -\epsilon K(T - t_j), 0) \leq 0. $$

2.b We have that ${}^{\epsilon}\overline{w}^j_k = V^j_k - \epsilon K_1(T - t_j)$, for $k \in \mathbb{Z}$. Since

$$ f(t_j, x_k, V^j_k - \epsilon K_1(T - t_j), U^j_k) \geq f(t_j, x_k, V^j_k, U^j_k)\,, $$

and by definition of $V_{\pi,\delta}$:

$$ \begin{aligned} S_b(k, {}^{\epsilon}\overline{w}_k, \nabla_\delta {}^{\epsilon}\overline{w}_k, \nabla_{+,\delta} {}^{\epsilon}\overline{w}_k, \Delta_\delta {}^{\epsilon}\overline{w}_k, {}^{\epsilon}w^{j+1}_{k,N^a_\delta}) &= -\epsilon K h + S_b(k, V^j_k, \nabla_\delta V^j_k, \nabla_{+,\delta} V^j_k, \Delta_\delta V^j_k) \\ &\leq -\epsilon K h \leq 0. \end{aligned} $$

2.c We now prove (4.3.27). Let $k \in \mathbb{Z}$, $0 < l < N^a_\delta$. We have, by definition (4.3.10) of S:

$$ \begin{aligned} S(k, l, {}^{\epsilon}w^j_{k,l}, &\nabla^a_\delta {}^{\epsilon}w^j_{k,l}, \nabla^a_{+,\delta} {}^{\epsilon}w^j_{k,l}, \Delta^a_\delta {}^{\epsilon}w^j_{k,l}, {}^{\epsilon}w(t_{j+1}, \cdot)) \\ &= {}^{\epsilon}w^j_{k,l} - {}^{\epsilon}w(t_{j+1}, x_k, \mathfrak{p}^a(p_l)) \\ &\quad + h\widehat{F}(t, k\delta, {}^{\epsilon}w^j_{k,l}, \nabla^a_\delta {}^{\epsilon}w^j_{k,l}, \nabla^a_{+,\delta} {}^{\epsilon}w^j_{k,l}, \Delta^a_\delta {}^{\epsilon}w^j_{k,l}) \\ &\leq -{}^{\epsilon}c^j_{k,l} + \mu \frac{\mathfrak{a}(a, \delta)}{\sigma} h V^{j+1}_k + {}^{\epsilon}c(t_{j+1}, x_k, \mathfrak{p}^a(p_l)) \\ &\quad - p_l h \widehat{F}(t, x_k, V^j_k, U^j_k, \nabla_{+,\delta} V^j_k, \Delta_\delta V^j_k) \\ &\quad + h\widehat{F}(t, x_k, p_l V^j_k, \nabla^a_\delta {}^{\epsilon}w^j_{k,l}, \nabla^a_{+,\delta} {}^{\epsilon}w^j_{k,l}, \Delta^a_\delta {}^{\epsilon}w^j_{k,l}), \end{aligned} $$

where we have used (4.3.23) and $f(t, x_k, {}^\epsilon w^j_{k,l}, \sigma\nabla^{a\epsilon}_\delta w^j_{k,l}) \geq f(t, x_k, p_l V^j_k, \sigma\nabla^{a\epsilon}_\delta w^j_{k,l})$. By adding $\pm p_l h f(t_j, x_k, p_l V^j_k, \sigma\nabla^{a\epsilon}_\delta w^j_{k,l})$, using the Lipschitz continuity of f and

$$\nabla^{a\epsilon}_\delta w^j_{k,l} = p_l U^j_k + \frac{\mathfrak{a}(a,\delta)}{2\sigma}(V^j_{k+1} + V^j_{k-1}) + \frac{1}{2\delta}\left({}^\epsilon c^j_{l-\mathrm{sgn}(a)} - {}^\epsilon c^j_{l+\mathrm{sgn}(a)}\right),$$

we get, by definition (4.3.7) of $\widehat{F}$,

$$\begin{aligned}
S(k, l, {}^\epsilon w_{k,l}, &\nabla^{a\epsilon}_\delta w^j_{k,l}, \nabla^a_{+,\delta}{}^\epsilon w^j_{k,l}, \Delta^{a\epsilon}_\delta w^j_{k,l}, {}^\epsilon w^j(t_{j+1},\cdot)) \\
\leq & h\frac{\mathfrak{a}(a,\delta)}{\sigma}\mu(V^{j+1}_k - V^j_{k+1}) - h\sigma\mathfrak{a}(a,\delta)U^j_k - 2\theta\frac{\mathfrak{a}(a,\delta)}{\sigma}\delta^2 U^j_k \\
& + 2hLp_l(1-p_l)(V^j_k + |U^j_k|) + hL\frac{|\mathfrak{a}(a,\delta)|}{2\sigma}(V^j_{k+1}+V^j_{k-1}) \\
& - \left({}^\epsilon c^j_l - {}^\epsilon c(t_{j+1}, \mathfrak{p}^a(p_l)) - \mu h\nabla^a_{+,\delta}{}^\epsilon c^j_l - \left(\frac{\sigma^2}{2}h + \theta\delta^2\right)\Delta^{a\epsilon}_\delta c^j_l\right) \\
& + hL|\nabla^{a\epsilon}_\delta c^j_l|.
\end{aligned}$$

Since $|\mathfrak{a}(a,\delta)| \leq \max\{|a|, a \in K\} \leq n$ and V and U are bounded uniformly in h, δ (see Proposition 4.6.3 in the appendix), there exists a constant $K_{n,\theta,M,L} > 0$ such that

$$\begin{aligned}
& h\frac{\mathfrak{a}(a,\delta)}{\sigma}\mu(V^{j+1}_k - V^j_{k+1}) - h\sigma\mathfrak{a}(a,\delta)U^j_k - 2\theta\frac{\mathfrak{a}(a,\delta)}{\sigma}\delta^2 U^j_k \\
& + 2hLp_l(1-p_l)(V^j_k + |U^j_k|) + hL\frac{|\mathfrak{a}(a,\delta)|}{2\sigma}(V^j_{k+1}+V^j_{k-1}) \leq hK_{n,\theta,M,L}.
\end{aligned}$$

When $\epsilon = 1$, the terms of the last three lines all vanish except the first one, and $c^j_l - c(t_{j+1}, p_l - \mu\frac{\mathfrak{a}(a,\delta)}{\sigma}h) = K_1 h$. Thus we get:

$$S(k, l, {}^\epsilon w_{k,l}, \nabla^{a\epsilon}_\delta w_{k,l}, \nabla^a_{+,\delta}{}^\epsilon w_{k,l}, \Delta^{a\epsilon}_\delta w_{k,l}, {}^\epsilon w(t_{j+1},\cdot)) \leq h(-K_1 + K_{n,\theta,M,L}).$$

Hence, choosing K_1 large enough gives the result.

We now deal with the case $\epsilon = 0$. By Taylor expansions of ${}^\epsilon c$ around (t_j, p_l), we get:

$$\begin{aligned}
& S(k, l, {}^\epsilon w_{k,l}, \nabla^{a\epsilon}_\delta w^j_{k,l}, \nabla^a_{+,\delta}{}^\epsilon w^j_{k,l}, \Delta^{a\epsilon}_\delta w^j_{k,l}, {}^\epsilon w^j(t_{j+1},\cdot)) \\
& \leq hK_{n,\theta,M,L} + hL\frac{|\mathfrak{a}(a,\delta)|}{\sigma}|\partial_p{}^\epsilon c(t_j,p_l)| + h\partial_t{}^\epsilon c(t_j,p_l) + h\frac{\mathfrak{a}(a,\delta)^2}{2}\partial^2_{pp}{}^\epsilon c(t_j,p_l) + h\varepsilon(h;K_2),
\end{aligned}$$

with $\lim_{h\to 0}\varepsilon(h;K_2) = 0$. By definition of ${}^\epsilon c$, we get, for $h_0 > 0$ to be fixed later on and $h \in [0, h_0]$:

$$\begin{aligned}
& S(k, l, {}^\epsilon w_{k,l}, \nabla^{a\epsilon}_\delta w^j_{k,l}, \nabla^a_{+,\delta}{}^\epsilon w^j_{k,l}, \Delta^{a\epsilon}_\delta w^j_{k,l}, {}^\epsilon w^j(t_{j+1},\cdot)) \\
& \qquad \leq h\Bigg[K_{n,\theta,M,L} + K_2 L\frac{|\mathfrak{a}(a,\delta)|}{\sigma}e^{-K_2 p_l} + K_2 L\frac{\mathfrak{a}(a,\delta)}{\sigma}e^{-K_2(1-p_l)} \\
& \qquad\qquad - K_2^2\frac{\mathfrak{a}(a,\delta)^2}{2}e^{-K_2 p_l} - K_2^2\frac{\mathfrak{a}(a,\delta)^2}{2}e^{-K_2(1-p_l)} + |\varepsilon(h;K_2)|\Bigg] \\
& \leq h\left[\max_{h\in[0,h_0]}|\varepsilon(h;K_2)| + K_{n,\theta,M,L} + K_2|\mathfrak{a}(a,\delta)|(e^{-K_2 p_l} + e^{-K_2(1-p_l)})\left(\frac{L}{\sigma} - \frac{|\mathfrak{a}(a,\delta)|}{2}K_2\right)\right].
\end{aligned}$$

To conclude, one can choose K_2 large enough so that $K_{n,\theta,M,L} + K_2|\mathfrak{a}(a,\delta)|(e^{-K_2 p_l} + e^{-K_2(1-p_l)})(\frac{L}{\sigma} - \frac{|\mathfrak{a}(a,\delta)|}{2}K_2) \leq -\eta < 0$, and then consider $h_0 > 0$ small enough so that $|\varepsilon(h;K_2)| \leq \eta$ for $h \in [0, h_0]$. □

Proposition 4.3.5 (Monotonicity). *Let π be a grid of $[0,T]$ and $\delta > 0$ satisfying (4.3.15)–(4.3.16). Let $y \in \mathbb{R}, 0 \le k \le \kappa, j \in \mathbb{Z}$, $p \in [0,1]$, and let $\mathcal{U}, \mathcal{V} : \pi \times \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$ be two bounded functions such that $\mathcal{U} \le \mathcal{V}$. Then:*

$$\widehat{\mathfrak{S}}(\pi, \delta, k, j, p, y, \mathcal{U}) \ge \widehat{\mathfrak{S}}(\pi, \delta, k, j, p, y, \mathcal{V}).$$

Proof. The result is clear for $k = \kappa$. If $k < \kappa$, it is sufficient to show that:

$$\mathcal{I}^a_\delta(S^a_\delta(t_{k+1}, t_k, \mathcal{U}(t_{k+1}, \cdot))) \le \mathcal{I}^a_\delta(S^a_\delta(t_{k+1}, t_k, \mathcal{V}(t_{k+1}, \cdot))),$$

for all $a \in K$, recalling (4.3.1). This is a consequence of the comparison result in Proposition 4.3.2 and the monotonicity of the linear interpolator. □

We now prove the stability of the scheme. Here, in contrast to Lemma 4.2.1, we are not able to prove that the solution of the scheme is increasing in p. However, due to the boundedness of the terminal condition, we obtain uniform bounds for $v_{n,\pi,\delta}$.

Proposition 4.3.6 (Stability). *For all π and $\delta > 0$, there exists a unique solution $v_{n,\pi,\delta}$ to (4.3.1), which satisfies:*

$$0 \le v_{n,\pi,\delta} \le |g|_\infty \text{ on } \pi \times \delta\mathbb{Z} \times [0,1].$$

Proof. We prove the proposition by backward induction. First, since $v_{n,\pi,\delta}$ is a solution to (4.3.1), $v_{n,\pi,\delta}(T,x,p) = pg(x)$ on $\delta\mathbb{Z} \times [0,1]$, and we have $0 \le v_{n,\pi,\delta}(T,x,p) \le |g|_\infty$ for all $(x,p) \in \delta\mathbb{Z} \times [0,1]$.

Let $0 \le j \le \kappa - 1$ and assume that $v_{n,\pi,\delta}(t_k, \cdot)$ is uniquely determined for $k > j$, and that $0 \le v_{n,\pi,\delta}(t_{j+1}, \cdot) \le |g|_\infty$. Since $v_{n,\pi,\delta}$ is a solution to (4.3.1), we have

$$v_{n,\pi,\delta}(t_j, x, p) = \min_{a \in K} \mathcal{I}^a_\delta(S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot))),$$

and for each $a \in K$, $S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot))$ is uniquely determined by Proposition 4.3.1, so $v_{n,\pi,\delta}(t_j, \cdot)$ is uniquely determined. Next, we show that, for all $k \in \mathbb{Z}$ and $0 \le l \le N_a$:

$$0 \le S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot)) \le |g|_\infty.$$

Then it is easy to conclude that $0 \le v_{n,\pi,\delta}(t_j, \cdot) \le e^{LT}|g|_\infty$ on $\mathbb{R} \times [0,1]$, by properties of the linear interpolation and the minimisation.

First, it is straightforward that $\check{u}$ defined by $\check{u}_{k,l} = 0$ for all $k \in \mathbb{Z}$ and $0 \le l \le N_a$ satisfies $\check{u} = S^a_\delta(t_{j+1}, t_j, 0)$. The comparison theorem gives $0 \le S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot))$, since $0 \le v_{n,\pi,\delta}(t_{j+1}, \cdot)$. To obtain the upper bound, we notice that $\hat{u}$ defined by $\hat{u}_{k,l} := |g|_\infty$ for all $k \in \mathbb{Z}$ and $0 \le l \le N_a$ satisfies

$$S(k, l, \hat{u}_{k,l}, \nabla^a_\delta \hat{u}_{k,l}, \nabla^a_{+,\delta} \hat{u}_{k,l}, \Delta^a_\delta \hat{u}_{k,l}, \hat{u}) = -hf(t_j, x_k, \hat{u}, 0) \ge -hf(t_j, x_k, 0, 0) \ge 0.$$

Hence the comparison result in Proposition 4.3.2 yields $S^a_\delta(t_{j+1}, t_j, v_{n,\pi,\delta}(t_{j+1}, \cdot)) \le |g|_\infty$. □

We now prove the consistency. The proof requires several lemmata. First, we show that the perturbation induced by the change of controls vanishes as $\delta \to 0$.

Lemma 4.3.1. *For all $a \in K$, a and $\mathfrak{a}(a,\delta)$ have the same sign, and:*

$$0 \leq |a| - |\mathfrak{a}(a,\delta)| \leq \frac{n^2}{\sigma}\delta. \tag{4.3.28}$$

Moreover, there exists $c > 0$ such that for all $a \in K$ and $\delta > 0$, $|\mathfrak{a}(a,\delta)| \geq c > 0$.

Proof. By definition of N_δ^a,

$$(N_\delta^a - 1)\frac{|a|}{\sigma}\delta < 1 = N_\delta^a \frac{|\mathfrak{a}(a,\delta)|}{\sigma}\delta \leq N_\delta^a \frac{|a|}{\sigma}\delta,$$

thus

$$|a| - \frac{|a|}{N_\delta^a} < |\mathfrak{a}(a,\delta)| \leq |a|.$$

Also, we observe

$$\frac{|a|}{N_\delta^a} = \frac{|a|}{\lceil \frac{\sigma}{|a|\delta} \rceil} \leq \frac{|a|}{\frac{\sigma}{|a|\delta}} = \frac{a^2\delta}{\sigma} \leq \frac{n^2}{\sigma}\delta,$$

which concludes the proof of (4.3.28).

By (4.3.4), we have:

$$N_\delta^a \leq \frac{\sigma}{|a|\delta} + 1 \leq \frac{\sigma}{a_m\delta} + 1 \leq \frac{\sigma}{c\delta}$$

where $a_m = \min\{|a| : a \in K\}$ and $c > 0$ is independent of a, δ. Now, by (4.3.5), we get:

$$|\mathfrak{a}(a,\delta)| = \frac{\sigma}{\delta N_\delta^a} \geq c.$$

□

Last, we give explicit supersolutions and subsolutions satisfying appropriate conditions. Let $0 \leq t < s \leq T$, $\delta > 0$ and $a \in K$ be fixed. For $\epsilon > 0$, we set

$$\begin{aligned} f_\epsilon(t,x,y,\nu) &:= (f(t,\cdot,\cdot,\cdot) * \rho_\epsilon)(t,y,\nu) \\ &:= \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}} f(t,x-u,y-z,\nu-\eta)\rho_\epsilon(u,z,\eta)\mathrm{d}u\mathrm{d}z\mathrm{d}\eta, \end{aligned}$$

where $*$ is the convolution operator and, for $\epsilon > 0$, $\rho_\epsilon(x) := \epsilon^{-3}\rho(x/\epsilon)$ with $\rho : \mathbb{R}^3 \to \mathbb{R}$ is a mollifier, i.e. a smooth function supported on $[-1,1]^3$ satisfying $\int_{\mathbb{R}^3} \rho = 1$. We set

$$F_\epsilon(t,x,y,q,A) = \left(\frac{1}{2}\sigma^2 - \mu\right) q - \frac{\sigma^2}{2}A - f_\epsilon(t,x,y,\sigma q).$$

Remark 4.3.5. *Since f is L-Lipschitz continuous with respect to its three last variables, we have $|f_\epsilon - f|_\infty \leq L\epsilon$.*

The lengthy proof of the following lemma by insertion is given in the appendix.

Lemma 4.3.2. *Let $0 \le t < s \le T, \varphi \in \mathcal{C}_b^\infty(\mathbb{R}\times\mathbb{R},\mathbb{R}), a \in K$. We set $h := s-t$. Let $\epsilon > 0$ such that $\epsilon \to 0$ and $\frac{\delta}{\epsilon^2} \to 0$ as $h \to 0$, observing (4.3.16).*
Then there exist bounded functions $S_\delta^{a,\pm}(s,t,\varphi) : \delta\mathbb{Z}\times[0,1] \to \mathbb{R}$ of the form

$$\begin{aligned} S_{\delta,\epsilon}^{a,\pm}(s,t,\varphi)(x,p) =& \varphi(x,\mathfrak{p}^a(p)) \\ & - hF_\epsilon(t,x,\varphi(x,\mathfrak{p}^a(p)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x,\mathfrak{p}^a(p)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x,\mathfrak{p}^a(p))) \\ & \pm C_{\varphi,n}(h,\epsilon), \end{aligned} \tag{4.3.29}$$

where $\mathfrak{p}^a$ is defined in (4.3.11), and where $C_{\varphi,n}(h,\epsilon) > 0$ satisfies $\frac{C_{\varphi,n}(h,\epsilon)}{h} \to 0$ as $h \to 0$, such that $w^\pm := (S_{\delta,\epsilon}^{a,\pm}(s,t,\varphi)(x_k,p_l))_{k\in\mathbb{Z},0\le l\le N_\delta^a} \in \ell^\infty(\mathcal{G}_\delta^a))$ satisfy

$$S(k,l,w_{k,l}^+,\nabla_\delta^a w_{k,l}^+,\nabla_{+,\delta}^a w_{k,l}^+,\Delta_\delta^a w_{k,l}^+) \ge 0, \tag{4.3.30}$$
$$S(k,l,w_{k,l}^-,\nabla_\delta^a w_{k,l}^-,\nabla_{+,\delta}^a w_{k,l}^-,\Delta_\delta^a w_{k,l}^-) \le 0, \tag{4.3.31}$$

for all $k \in \mathbb{Z}$ and $0 < l < N_\delta^a$.
Furthermore, for all $x \in \delta\mathbb{Z}$, $S_{\delta,\epsilon}^{a,\pm}(s,t,\varphi)(x,\cdot) \in \mathcal{C}^2([0,1],\mathbb{R})$, and $|\partial_{pp}^2 S_{\delta,\epsilon}^{a,\pm}(s,t,\varphi)|_\infty \le \frac{C_\varphi(h)}{\epsilon^2}$ for some constant $C_\varphi(h) > 0$ independent of ϵ.

Lemma 4.3.3. *Let $0 \le t < s \le T, \delta > 0, a \in K, \varphi \in \mathcal{C}_b^\infty(\mathbb{R}\times\mathbb{R})$ be fixed. Let $h = s-t$, $k \in \mathbb{Z}, x_k \in \delta\mathbb{Z}, p \in (0,1)$, and assume that h is sufficiently small so that $p \in [p_1, p_{N_\delta^a-1}]$, observing (4.3.16). Let $\epsilon > 0$ such that $\epsilon \to 0$ and $\frac{\delta}{\epsilon^2} \to 0$ as $h \to 0$. Then we have:*

$$\begin{aligned} S_{\delta,\epsilon}^{a,-}(s,t,\varphi)(x_k,p) - \mathcal{I}_\delta^a(S_\delta^a(s,t,\varphi))(x_k,p) \le C'_{\varphi,n}(h,\epsilon), \\ \mathcal{I}_\delta^a(S_\delta^a(s,t,\varphi))(x_k,p) - S_{\delta,\epsilon}^{a,+}(s,t,\varphi)(x_k,p) \le C'_{\varphi,n}(h,\epsilon), \end{aligned}$$

where $C'_{\varphi,n}(h,\epsilon) > 0$ satisfies $\frac{C'_{\varphi,n}(h,\epsilon)}{h} \to 0$, as $h \to 0$ and where the functions $S_{\delta,\epsilon}^{a,\pm}(s,t,\varphi)$ are introduced in Lemma 4.3.2.

Proof. We prove the first identity, while the second one is similar.
Set $w := S_\delta^a(s,t,\varphi)$ and $w^- := S_{\delta,\epsilon}^{a,-}(s,t,\varphi)$. By definition of w and by (4.3.31), one can apply Proposition 4.3.3. For all $k \in \mathbb{Z}$ and $0 < l < N_\delta^a$:

$$w_{k,l}^- - w_{k,l} \le Be^{-4\frac{\mathfrak{a}(a,\delta)^2}{\sigma^2}C(h,\delta)l(N_\delta^a-l)} \le Be^{-4\frac{\mathfrak{a}(a,\delta)^2}{\sigma^2}C(h,\delta)(N_\delta^a-1)}, \tag{4.3.32}$$

with $B = |(w_{\cdot,0}^- - w_{\cdot,0})^+|_\infty + |(w_{\cdot,N_\delta^a}^- - w_{\cdot,N_\delta^a})^+|_\infty$ and $C(h,\delta)$ is defined in (4.3.21). By Lemma 4.3.1, there exists a constant $c > 0$ such that $|\mathfrak{a}(a,\delta)| \ge c$. In addition, using (4.3.22), we get:

$$\begin{aligned} \frac{B}{h}e^{-4C(h,\delta)\frac{\mathfrak{a}(a,\delta)^2}{\sigma}(N_\delta^a-1)} &\le \frac{B}{h}e^{-4\frac{c^2}{\sigma^2}\left(\frac{1}{\left((\mu+\frac{L}{2})M+2\theta M^2\right)h^2+\sigma^2 h}-\frac{M^2}{2\sigma^4}\right)} \\ &= Be^{4\frac{c^2M^2}{2\sigma^6}}\frac{e^{-4\frac{c^2}{\sigma^2}\frac{1}{\left((\mu+\frac{L}{2})M+2\theta M^2\right)h^2+\sigma^2 h}}}{h} \to 0 \quad \text{as } h \to 0. \end{aligned}$$

Now let $p \in [p_1, p_{N_\delta^a - 1})$ and $k \in \mathbb{Z}$. By definition of $\mathcal{I}_\delta^a$, one has:

$$\mathcal{I}_\delta^a(S_\delta^a(s,t,\varphi))(x_k,p) = \lambda w_{k,l} + (1-\lambda) w_{k,l+1},$$

where $p \in [p_l, p_{l+1})$ with $0 < l < N_\delta^a - 1$, and $\lambda = \frac{p_{l+1}-p}{p_{l+1}-p_l}$. Thus:

$$\begin{aligned}
& S_{\delta,\epsilon}^{a,-}(s,t,\varphi)(x_k,p) - \mathcal{I}_\delta^a(w)(x_k,p) \\
&= S_{\delta,\epsilon}^{a,-}(s,t,\varphi)(x_k,p) - \mathcal{I}_\delta^a(w^-)(x_k,p) + \mathcal{I}_\delta^a(w^-)(x_k,p) - \mathcal{I}_\delta^a(w)(x_k,p) \\
&= S_{\delta,\epsilon}^{a,-}(s,t,\varphi)(x_k,p) - \mathcal{I}_\delta^a(w^-)(x_k,p) + \lambda(w_{k,l}^- - w_{k,l}) + (1-\lambda)(w_{k,l+1}^- - w_{k,l+1}).
\end{aligned}$$

The two last terms are controlled using (4.3.32), and, by properties of linear interpolation of the function $p \mapsto S_{\delta,\epsilon}^{a,-}(t^+,t^-,\varphi)(x_k,p) \in \mathcal{C}^2([0,1],\mathbb{R})$ with $|\partial_{pp}^2 S_{\delta,\epsilon}^{a,-}(t^+,t^-,\varphi)|_\infty \leq \frac{C_\varphi(h)}{\epsilon^2}$ (recall the previous Lemma) the first term is of order $\frac{\delta^2}{\epsilon^2} = o(h)$ since (4.3.16) is in force and $\frac{\delta}{\epsilon^2} \to 0$. □

Lemma 4.3.4. *For $0 \leq t < s \leq T$ such that $L(s-t) \leq 1, \xi > 0, \varphi : \delta\mathbb{Z} \times [0,1] \to \mathbb{R}$ a bounded function, the following holds for all $a \in K$:*

$$S_\delta^a(s,t,\varphi) + \xi - L(s-t)\xi \leq S_\delta^a(s,t,\varphi+\xi) \leq S_\delta^a(s,t,\varphi) + \xi,$$

where L is the Lipschitz constant of f.

Proof. Let $v = S_\delta^a(s,t,\varphi), w = v + \xi - L(s-t)\xi$. Since v satisfies (4.3.8), we have, for $k \in \mathbb{Z}$ and $0 < l < N_\delta^a$,

$$\begin{aligned}
S(k,l,w_{k,l},\nabla_\delta^a w_{k,l},\nabla_{+,\delta}^a w_{k,l},\Delta_\delta^a w_{k,l},\varphi+\xi) &= -L(s-t)\xi \\
&+ (s-t)\left(f(t,x_k,v_{k,l},\nabla_\delta^a v_{k,l}) - f(t,x_k,w_{k,l},\nabla_\delta^a v_{k,l})\right).
\end{aligned}$$

Since f is non-increasing in its third variable and Lipschitz continuous, we get:

$$S(k,l,w_{k,l},\nabla_\delta^a w_{k,l},\nabla_{+,\delta}^a w_{k,l},\Delta_\delta^a w_{k,l},\varphi+\xi) \leq 0.$$

The same computation with $l = 0$ or $l = N_\delta^a$ and S_b instead of S gives

$$S_b(k,l,w_{k,l},\nabla_\delta w_{k,l},\nabla_{+,\delta} w_{k,l},\Delta_\delta w_{k,l},\varphi+\xi) \leq 0,$$

and the comparison theorem given in Proposition 4.6.2 gives $w_{k,l} \leq S_\delta^a(s,t,\varphi+\xi)_{k,l}$ for $k \in \mathbb{Z}$ and $l \in \{0, N_\delta^a\}$.

The comparison result from Proposition 4.3.2 gives the first inequality of the lemma. The second one is proved similarly. □

Proposition 4.3.7 (Consistency). *Let $\varphi \in \mathcal{C}_b^\infty([0,T] \times \mathbb{R} \times \mathbb{R}, \mathbb{R}), (t,x,p) \in [0,T) \times \mathbb{R} \times (0,1)$. We have, with the notation in (4.3.2):*

$$\begin{aligned}
\Bigg| & \frac{1}{t_{j+1}-t_j} \widehat{\mathfrak{S}}(\pi,\delta,t_j,x_k,q,\varphi(t_j,x_k,q)+\xi,\varphi+\xi) \\
& - \sup_{a \in K}\left[-D^a\varphi(t,x,p) + F(t,x,\varphi(t,x,p),\nabla^a\varphi(t,x,p),\Delta^a\varphi(t,x,p))\right]\Bigg| \to 0,
\end{aligned}$$

as $\delta, |\pi| \to 0$ satisfying (4.3.15)–(4.3.16), $\pi \times \delta\mathbb{Z} \times [0,1] \ni (t_j,x_k,q) \to (t,x,p), \xi \to 0$.

Proof. Let φ, j, k, p, l as in the statement of the Proposition. Without loss of generality, we can consider π, δ, t_j, x_k, q such that, for all $a \in K$:

$$0 \leq \mathfrak{p}^a(q) \leq 1.$$

Since φ is smooth and $(t_k, x_j, p_l) \to (t, x, p)$, we have

$$\begin{aligned}|\sup_{a\in K}&[-D^a\varphi(t,x,p) + F(t,x,p,\varphi(t,x,p),\nabla^a\varphi(t,x,p),\Delta^a\varphi(t,x,p))] \\ &- \sup_{a\in K}[-D^a\varphi(t_j,x_k,p_l) + F(t_j,x_k,p_l,\varphi(t_j,x_k,p_l),\nabla^a\varphi(t_j,x_k,p_l),\Delta^a\varphi(t_j,x_k,p_l)))]| \to 0.\end{aligned}$$

Thanks to Lemma 4.3.4, it suffices to prove:

$$\begin{aligned}&\Bigg|\frac{1}{t_{j+1}-t_j}\widehat{\mathfrak{S}}(\pi,\delta,t_j,x_k,q,\varphi(t_j,x_k,q),\varphi) \\ &\quad - \sup_{a\in K}(-D^a\varphi(t_j,x_k,q) + F(t_j,x_k,\varphi(t_j,x_k,q),\nabla^a\varphi(t_j,x_k,q),\Delta^a\varphi(t_j,x_k,q)))\Bigg| \to 0,\end{aligned}$$

as $|\pi| \to 0$ and $\pi \times \delta\mathbb{Z} \times (0,1) \ni (t_j, x_k, q) \to (t, x, p)$.

Let $\epsilon > 0$ such that $\epsilon \to 0$ and $\frac{\delta}{\epsilon^2} \to 0$. Using $|\inf - \sup| \leq \sup|\cdot - \cdot|$, adding

$$\begin{aligned}\pm\Bigg(&\frac{1}{t_{j+1}-t_j}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) \\ &+F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)))\Bigg)\end{aligned}$$

and using Lemma 4.3.1, it is enough to show that, for all $a \in K$,

$$\begin{aligned}&\Bigg|\frac{1}{t_{j+1}-t_j}\left(\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) - \mathcal{I}^a_\delta(S^a_\delta(t_{j+1},t_j,\varphi(t_{j+1},\cdot)))(x_k,q)\right) \\ &-F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)))\Bigg| \to 0.\end{aligned}$$

The proof is concluded using the equality $|\cdot| = \max(\cdot, -\cdot)$ and the two following inequalities, obtained by Lemma 4.3.3 and by definition (4.3.30) to (4.3.31) of $S^{a,\pm}_{\delta,\epsilon}$:

$$\begin{aligned}&\frac{1}{t_{j+1}-t_j}\left(\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) - \mathcal{I}^a_\delta(S^a_\delta(t_{j+1},t_j,\varphi(t_{j+1},\cdot)))(x_k,q)\right) \\ &- F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q))) \\ &\leq \frac{1}{t_{j+1}-t_j}(\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) - S^{a,-}_{\delta,\epsilon}(t_{j+1},t_j,\varphi(t_{j+1},\cdot)))(x_k,q) + o(t_{j+1}-t_j)) \\ &- F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q))),\end{aligned}$$

and

$$\begin{aligned}&F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q))) \\ &- \frac{1}{t_{j+1}-t_j}\left(\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) - \mathcal{I}^a_\delta(S^a_\delta(t_{j+1},t_j,\varphi(t_{j+1},\cdot)))(x_k,q)\right) \\ &\leq F_\epsilon(t_j,x_k,\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(t_{j+1},x_k,\mathfrak{p}^a(q))) \\ &- \frac{1}{t_{j+1}-t_j}(\varphi(t_{j+1},x_k,\mathfrak{p}^a(q)) - S^{a,+}_{\delta,\epsilon}(t_{j+1},t_j,\varphi(t_{j+1},\cdot)))(x_k,q) + o(t_{j+1}-t_j)).\end{aligned}$$

□

4.4 Numerical studies

We now present a numerical application of the algorithm from Section 4.3.

4.4.1 Model

We keep the notation of the previous section: the process X is a Brownian motion with drift. In this numerical example, the drift of Y is given by the following functions:

$$f_1(t,x,y,z) := -\sigma^{-1}\mu z, \text{ and}$$
$$f_2(t,x,y,z) := -\sigma^{-1}\mu z + R(y-\sigma^{-1}z)^-,$$

where, for $x \in \mathbb{R}$, $x^- = \max(-x,0)$ denotes the negative part of x. The function f_1 corresponds to a linear complete Black & Scholes market. It is well-known that there are explicit formulae for the quantile hedging price for a vanilla put or call, see [FL99]. In both cases, we compute the quantile hedging price of a put option with strike $K = 30$ and maturity $T = 1$, i.e. $g(x) = \max(K - \exp(x), 0)$. The parameters of X are $\sigma = 0.25$ and $\mu = 0.01875$ (this corresponds to a parameter $b = 0.05$ for the associated geometric Brownian motion, where $\mu = b - \sigma^2/2$).

In the rest of this section, we present the following numerical experiments. First, using the non-linear driver f_2, we observe the convergence of $v_{\pi,\delta}$ towards v_n for a fixed discrete control set. Second, we show that the conditions (4.3.15) to (4.3.16) are not only theoretically important, but also numerically. Last, we use the fact that the analytical solution to the quantile hedging problem with driver f_1 is known (see [FL99]) to assess the convergence (order) of the scheme more precisely. We observe that a judicious choice of control discretization, time and space discretization leads to convergence of $v_{\pi,\delta}$ to v. However, the unboundedness of the optimal control as $t \to T$ leads to expensive computations.

The scheme obtained in the previous section deals with an infinite domain in the x variable. In practice, one needs to consider a bounded interval $[B_1, B_2]$ and to add some boundary conditions. Here, we choose $B_1 = \log(10)$ and $B_2 = \log(45)$, and the approximate Dirichlet boundary values for $v(t, B_i, p)$ are the limits $\lim v_{th}(t,x,p)$ as $x \to 0$ or $x \to +\infty$, where v_{th} is the analytical solution obtained in [FL99] for the linear driver f_1. Since the non-linearity in f_2 is small for realistic parameters (we choose $R = 0.05$ in our tests), it is expected that the prices are close (see also [GP15]). Furthermore, we will consider values obtained for points (t,x,p) with x far from to the boundary. In this situation, the influence of our choice of boundary condition should be small, as noticed for example in Proposition 4.3.3. This was studied more systematically, for example, in [BDR95].

4.4.2 Convergence towards v_n with the non-linear driver

In this section, we consider the non-linear driver f_2 defined above, where there is no known analytical expression for the quantile hedging price. We now fix a discrete control set, and we compute the value function $v_{\pi,\delta}$ for various discretization parameters π, δ

δ	N_t	N_x	N_c	N_p	$a_{\max}$	$N_p(a_{\max})$	$a_{\min}$	$N_p(a_{\min})$	time (in seconds)
0.1	10	18	12	15	2.5	2	0.3125	9	0.04
0.05	20	33	15	38	2.5	3	0.3125	17	0.15
0.01	100	153	22	255	2.77	10	0.33	76	166
0.005	200	304	22	585	2.94	18	0.33	151	4608

Table 4.1 – Parameters for selected values of δ for non-linear driver test.

satisfying (4.3.15) to (4.3.16). We consider the following control set with 22 controls:

$$\left([-2,2]\cap\frac{\mathbb{Z}\backslash\{0\}}{2}\right)\bigcup\left([-3,3]\cap\frac{\mathbb{Z}\backslash\{0\}}{3}\right)$$
$$=\{-2,-1.5,\dots,1.5,2\}\cup\left\{-3,-3+\frac{1}{3},\dots,3-\frac{1}{3},3\right\},$$

and $\delta\in\{0.1,0.07,0.05,0.03,0.01,0.007,0.005\}$. For a given δ, we set $h=C\delta$ with $C:=\min(1,2\frac{\theta}{L},\frac{1}{|\sigma^2-\mu|})$, $\theta=\frac{1}{5}$ and $L=|\mu|+R$, so that (4.3.15) to (4.3.16) are satisfied.

In Table 4.1, we show some discretization parameters obtained by this construction with selected values of δ. Recall that different meshes are applied in each step of the PCPT algorithm for different a. Here, N_x is the number of points for the x-variable, N_c the number of controls, and N_p the *total* number of points for the p variable, i.e., for all meshes for different values of a combined. Moreover, $a_{\max}$ (resp. $a_{\min}$) is the greatest (resp. smallest) control obtained, using the modification of the control set (4.4.2) as described in Section 4.3.1. We also report $N_p(a_{\max})$ (resp. $N_p(a_{\min})$), the number of points for the p variable for the control $a_{\max}$ (resp. $a_{\min}$). With our choice of parameters, we have $h=\delta$, so that the number of time steps is always $\frac{1}{\delta}$.

For different choices of δ, we get the graphs shown in Figure 6.1 for the function $p\mapsto v_{\pi,\delta}(t,x,p)$, where $(t,x)=(0,30),(0,37)$.

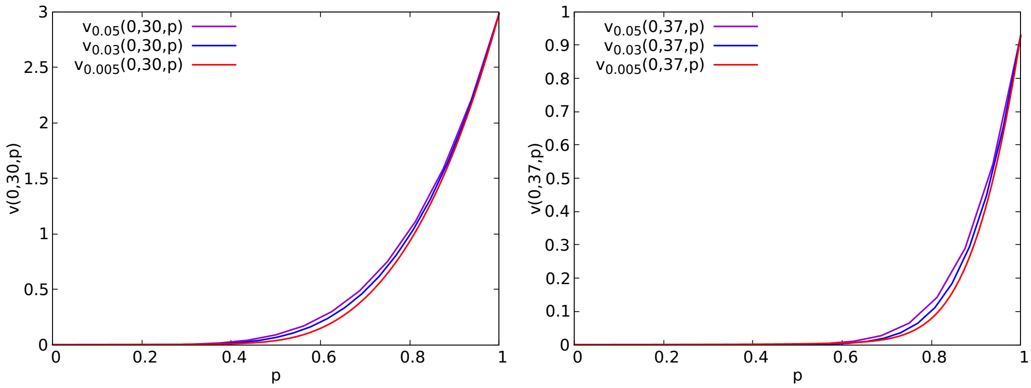

Figure 4.1 – The functions $v_{\pi,\delta}(t,x,\cdot)$, $t=0$ and $x\in\{30,37\}$. We drop the subscript π in the legend above as $h=\delta$.

We observe experimentally, while not proved, that the numerical approximation always gives an upper bound for v_n, which is itself greater than the quantile hedging

price v. This, if confirmed more systematically, could be a practically useful feature of this numerical method.
The scheme preserves a key feature of the exact solution, namely that the quantile hedging price is 0 exactly for p below a certain threshold, depending on t, x. This is a consequence of the diffusion stencil Δ_δ^a respecting the degeneracy of the diffusion operator Δ^a in (4.3.3), which acts only in direction $(1, p)$, and by the specific construction of the meshes.

4.4.3 CFL conditions

Using the same discrete control set as in Section 4.4.2, we now fix $h = 0.1$ and compute $v_{\pi,\delta}$ for δ chosen as above. The conditions (4.3.15) to (4.3.16) are then not satisfied and the results of Proposition 4.3.1 are not valid anymore.

First, while π is coarse, we observe that the computational time to obtain $v_{\pi,\delta}(t_j, \cdot)$ from $v_{\pi,\delta}(t_{j+1}, \cdot)$ is increased. While convergence of the Picard iteration to a fixed point is still observed, significantly more such iterations are needed. For example, for $\delta = 0.005$ and $h = 0.1$, we observe that 3000 Picard iterations are needed, while in the example where (4.3.15) to (4.3.16) were satisfied, 250 iterations sufficed to obtain convergence subject to a tolerance of 10^{-5}.

The second observation is that, while we seem to observe convergence of the sequence to some limit (judging from our chosen range of δ: it might start to diverge for smaller δ, as seen for the case δ fixed and varying h below), it is not the limit seen in the previous subsection. We show in the left panel of Figure 4.2 the difference between the solution obtained with $\delta = 0.005, h = 0.1$, and $\delta = 0.005, h = C\delta$. When the conditions (4.3.15) to (4.3.16) are not met, we generally have a non-monotone scheme, and convergence to the unique viscosity solution of the PDE, which equals the value function of the stochastic target problem, is not guaranteed.

Conversely, when δ is fixed and we vary h, the situation is different. There is no adverse effect on the Picard iterations, as the conditions needed for Proposition 4.3.1 are still satisfied. The issue here is that the consistency hypothesis is violated, and convergence to the true solution is again not observed: when h is too close to 0, the value $v_{\pi,\delta}$ grows, as seen in the right panel of Figure 4.2. Here, δ is fixed to 0.05 and h goes from 0.025 to 1.2×10^{-5}.

4.4.4 Convergence to the analytical solution with linear driver

We now consider the linear driver f_1. In that case, the quantile hedging price can be found explicitly (see [FL99]), and, for each (t, x, p), the optimal control is then:

$$\alpha(t, x, p) = \frac{1}{\sqrt{2\pi(T-t)}} \exp\left(-\frac{N^{-1}(p)^2}{2}\right),$$

where N is the cumulative distribution function of the standard normal distribution.

In particular, if the uniform grid $\pi = \{0, h, \ldots, \kappa h = T\}$ is fixed, one obtains that the optimal controls are contained in the interval $[0, \frac{1}{\sqrt{2\pi h}}]$ on $\pi \backslash \{T\}$. On the other hand, if δ is fixed, one sees from (4.3.5) that the largest control one can reach (with a non-trivial grid for the p variable) is $\frac{\sigma}{2\delta}$.

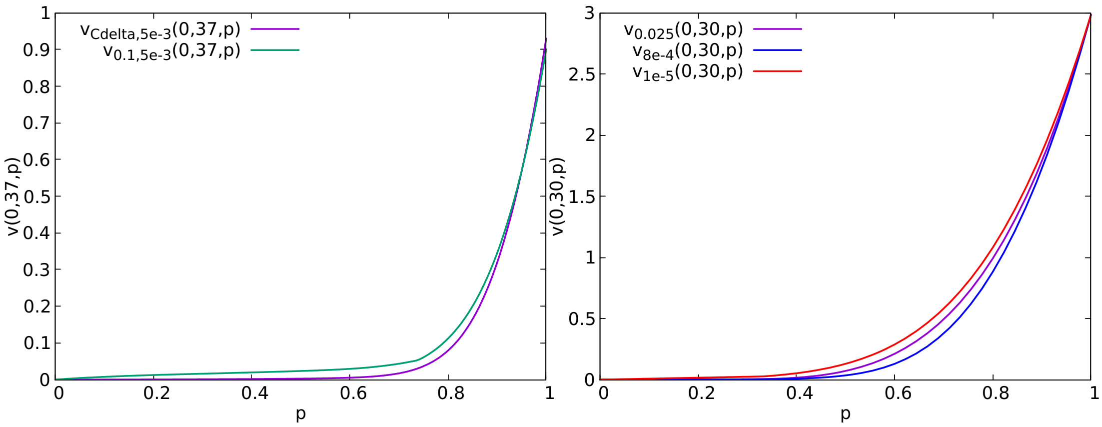

Figure 4.2 – Left: comparison of $v_{0.1,5e-3}(0, 37, \cdot)$ and $v_{C\delta,5e-3}(0, 37, \cdot)$. Right: comparison of $c_{h,0.05}(0, 30, \cdot)$ for different h. We drop the subscript δ in the legend above as $\delta = 0.05$ is fixed.

For the parameters, we first choose $n \geq 2$, then we pick δ such that $\frac{\sigma}{n\delta} \geq \frac{1}{\sqrt{2\pi C\delta}}$, and we set $h = C\delta$. It is easy to see that therefore δ is proportional to n^{-2}.

We now pick the controls as a subset of $\{\frac{\sigma}{m\delta}, m \geq n\}$ to obtain $K_n := \{a_i := \frac{\sigma}{m_i\delta}, i = 1, \ldots, M\}$ as follows: let $m_1 = n$ so that $a_1 = a^n_{\max} = \frac{\sigma}{n\delta}$. If $m_1, \ldots, m_i$ are constructed, we set $m_{i+1} = \inf\{m \geq m_i, \frac{\sigma}{m_i\delta} - \frac{\sigma}{m\delta} \geq \frac{1}{n}\}$ and $a_{i+1} = \frac{\sigma}{m_{i+1}\delta}$. If $m_{i+1} < n^{-1}$, then we set $M = i + 1$ and we are done. This eliminates clustering of control values and speeds up the computation.

In the left panel of Figure 4.3, we observe convergence towards the quantile hedging price. Moreover, the right panel of Figure 4.3 demonstrates that the pointwise error, here for $(t, x, p) = (0, 30, 0.8)$, has a convergence rate of about 1 with respect to n in the construction described previously. Last, in Table 4.2, we report the values of δ and $a_{\max}$ obtained for different choices of n.

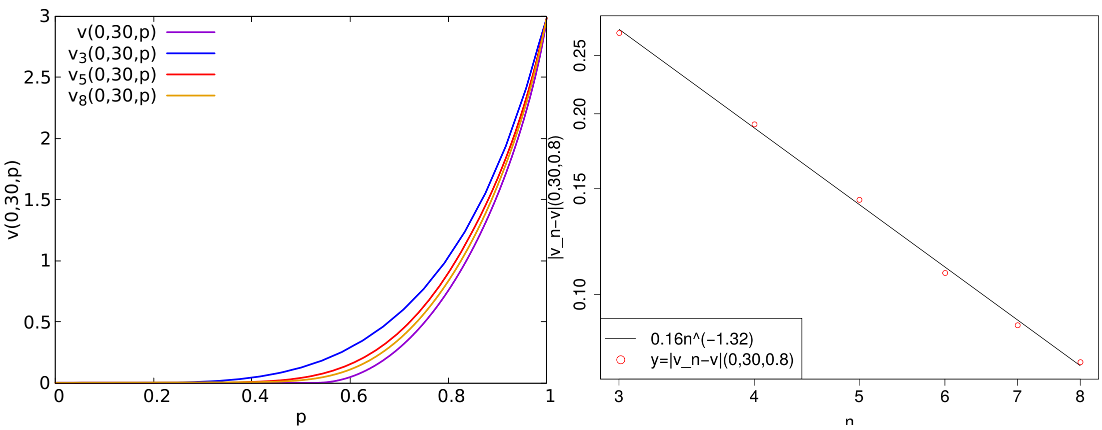

Figure 4.3 – Left: $v_n(0, 30, \cdot)$ and $v(0, 30, \cdot)$ for $n = 3, 5, 8$. Right: convergence rate estimation of $v_n(0, 30, 0.8)$ to $v(0, 30, 0.8)$ and log-log plot.

n	δ	$a_{\max}$	N_x	N_c	N_p
3	0.04	1.91	37	5	33
5	0.01	3.18	97	12	244
7	0.006	5.09	248	26	1138

Table 4.2 – Discretization parameters for selected values of n for linear driver.

4.5 Conclusions and extensions

We have introduced semi-discrete and discrete schemes for the quantile hedging problem, proven their convergence, and illustrated their behaviour in a numerical test.

The scheme, based on piecewise constant policy time-stepping, has the attractive feature that semi-linear PDEs for individual controls can be solved independently on adapted meshes. In the example of the Black-Scholes dynamics this had the effect that in spite of the degeneracy of the diffusion operator it was possible to construct on each mesh a local scheme, i.e. one where only neighbouring points are involved in the discretisation. This does not contradict known results on the necessity of non-local stencils for monotone consistent schemes in this degenerate situation (see e.g. [Rei18]), because of the superposition of different highly anisotropic meshes to arrive at a scheme which is consistent overall.

A more accurate scheme could be constructed by exploiting higher order, limited interpolation in the p-variable, such as in [RF16]. It should be possible to deduce convergence from the results of this paper and the properties of the interpolator using the techniques in [War16].

4.6 Appendix

4.6.1 Proofs

Proposition 4.3.3. For ease of notation, we set,

$$\begin{aligned} e(l) &:= e^{-4\frac{\mathfrak{a}(a,\delta)^2}{\sigma^2}C(h,\delta)l(N^a-l)}, \\ e^\star &:= \min_{x\in[0,N^a_\delta]} e(x) = e(N^a_\delta/2) = e^{-\frac{\mathfrak{a}(a,\delta)^2}{\sigma^2}C(h,\delta)(N^a_\delta)^2} = e^{-\frac{C(h,\delta)}{\delta^2}}, \\ B &:= |(v^1_{\cdot,0} - v^2_{\cdot,0})^+|_\infty + |(v^1_{\cdot,N^a_\delta} - v^2_{\cdot,N^a_\delta})^+|_\infty. \end{aligned}$$

By the comparison theorem, it is enough to show that $w \in \ell^\infty(\mathcal{G}^a_\delta)$ defined by

$$w_{k,l} := v^2_{k,l} + Be(l)$$

satisfies $w_{k,0} \geq v^1_{k,0}$, $w_{k,N^a_\delta} \geq v^1_{k,N^a_\delta}$ and $S(k,l,w_{k,l},\nabla^a_\delta w_{k,l},\nabla^a_{+,\delta} w_{k,l},\Delta^a_\delta w_{k,l},u) \geq 0$, for all $k \in \mathbb{Z}$ and $0 < l < N^a_\delta$.

The boundary conditions are easily checked: if $k \in \mathbb{Z}$ and $l \in \{0, N_a\}$, we have, since $e(0) = e(N_a) = 1$:

$$w_{k,l} = v^2_{k,l} + B \geq v^2_{k,l} + (v^1_{k,l} - v^2_{k,l})^+ \geq v^1_{k,l}.$$

For $k \in \mathbb{Z}$, $1 \leq l \leq N^a_\delta - 1$, we prove $S(k, l, w_{k,l}, \nabla^a_\delta w_{k,l}, \nabla^a_{+,\delta} w_{k,l}, \Delta^a_\delta w_{k,l}, u) \geq 0$. By definition (4.3.10), inserting $\pm h f(t^-, e^{k\delta}, v^2_{k,l}, \frac{1}{2\delta}(w_{k+1,l+\mathrm{sgn}(a)} - w_{k-1,l-\mathrm{sgn}(a)}))$, since $S(k, l, v^2_{k,l}, \nabla^a_\delta v^2_{k,l}, \nabla^a_{+,\delta} v^2_{k,l}, \Delta^a_\delta v^2_{k,l}, u) \geq 0$ and since f is non-increasing with respect to its third variable and Lipschitz continuous with respect to its fourth variable, we have:

$$
\begin{aligned}
&S(k, l, w_{k,l}, \nabla^a_\delta w_{k,l}, \nabla^a_{+,\delta} w_{k,l}, \Delta^a_\delta w_{k,l}, u) \\
&\geq B\left[\left(1 + \mu\frac{h}{\delta} + \sigma^2\frac{h}{\delta^2} + 2\theta\right) e(l) - \left(\mu\frac{h}{\delta} + \frac{\sigma^2}{2}\frac{h}{\delta^2} + \theta\right) e(l + \mathrm{sgn}(a))\right. \\
&\left. - \left(\frac{\sigma^2}{2}\frac{h}{\delta^2} + \theta\right) e(l - \mathrm{sgn}(a)) - \frac{hL}{2\delta}|e(l + \mathrm{sgn}(a)) - e(l - \mathrm{sgn}(a))|\right].
\end{aligned}
$$

We have $|e(l + \mathrm{sgn}(a)) - e(l - \mathrm{sgn}(a))| \leq 1 - e^\star$, thus:

$$
\begin{aligned}
&S(k, l, w_{k,l}, \nabla^a_\delta w_{k,l}, \nabla^a_{+,\delta} w_{k,l}, \Delta^a_\delta w_{k,l}, u) \\
&\geq B\left[\left(1 + \mu\frac{h}{\delta} + \sigma^2\frac{h}{\delta^2} + 2\theta + \frac{hL}{2\delta}\right) e^\star - \mu\frac{h}{\delta} - \sigma^2\frac{h}{\delta^2} - 2\theta - \frac{hL}{2\delta}\right].
\end{aligned}
$$

It is thus enough to have

$$
\left(1 + \mu\frac{h}{\delta} + \sigma^2\frac{h}{\delta^2} + 2\theta + \frac{hL}{2\delta}\right) e^\star - \mu\frac{h}{\delta} - \sigma^2\frac{h}{\delta^2} - 2\theta - \frac{hL}{2\delta} \geq 0,
$$

and one can easily check that this is the case with our choice of $C(h, \delta)$.
It remains to prove (4.3.22). Since $\ln(1+x) > x - \frac{x^2}{2}$ for all $x > 0$, we have, by (4.3.16):

$$
\begin{aligned}
C(h, \delta) &> \frac{1}{\delta^2}\left(\frac{1}{\mu\frac{h}{\delta} + \sigma^2\frac{h}{\delta^2} + 2\theta + \frac{hL}{2\delta}} - \frac{1}{2}\frac{1}{\left(\mu\frac{h}{\delta} + \sigma^2\frac{h}{\delta^2} + 2\theta + \frac{hL}{2\delta}\right)^2}\right) \\
&= \frac{1}{(\mu + \frac{L}{2})\delta h + \sigma^2 h + 2\theta\delta^2} - \frac{1}{2\left((\mu + \frac{L}{2})h + \sigma^2\frac{h}{\delta} + 2\delta\theta\right)^2} \\
&\geq \frac{1}{\left((\mu + \frac{L}{2})M + 2\theta M^2\right) h^2 + \sigma^2 h} - \frac{1}{2\left(((1 + 2\theta)\mu + \frac{L}{2})h + \frac{\sigma^2}{M}\right)^2} \\
&\geq \frac{1}{\left((\mu + \frac{L}{2})M + 2\theta M^2\right) h^2 + \sigma^2 h} - \frac{M^2}{2\sigma^4}.
\end{aligned}
$$

□

Lemma 4.3.2. We show the result for $S^{a,+}_{\delta,\epsilon}$, while the proof is similar for $S^{a,-}_{\delta,\epsilon}$. For $k \in \mathbb{Z}$ and $0 \leq l \leq N^a_\delta$, let

$$
z_{k,l} := \varphi(x_k, \mathfrak{p}^a(p_l)) - hF_\epsilon(t, x_k, \varphi(x, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k, \mathfrak{p}^a(p_l))). \quad (4.6.1)
$$

It is enough to show that, for all $k \in \mathbb{Z}$ and $0 < l < N^a_\delta$,

$$
S(k, l, z_{k,l}, \nabla^a_\delta z_{k,l}, \nabla^a_{+,\delta} z_{k,l}, \Delta^a_\delta z_{k,l}, \varphi) \geq -C_{\varphi,n}(h, \epsilon).
$$

Then, since f is non-increasing in its third variable, it is then easy to show that $w^+ = z + C_{\varphi,n}(h, \epsilon)$ satisfies (4.3.30).

Let $k \in \mathbb{Z}$ and $1 \leq l \leq N_a - 1$. We have, by definition (4.3.10):

$$\begin{aligned}
&S(k, l, z_{k,l}, \nabla^a_\delta z_{k,l}, \nabla^a_{+,\delta} z_{k,l}, \Delta^a_\delta z_{k,l}, \varphi) \\
&= h \Big(\widehat{F}(t, x_k, z_{k,l}, \nabla^a_\delta z_{k,l}, \nabla^a_{+,\delta} z_{k,l}, \Delta^a_\delta z_{k,l}) \\
&- F_\epsilon(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) \Big),
\end{aligned}$$

so it is enough to show

$$\begin{aligned}
&\widehat{F}(t, x_k, z_{k,l}, \nabla^a_\delta z_{k,l}, \nabla^a_{+,\delta} z_{k,l}, \Delta^a_\delta z_{k,l}) \\
&\quad - F_\epsilon(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) \geq -\frac{C_{\phi,n}(h, \epsilon)}{h}.
\end{aligned}$$

We split the sum into three terms:

$$\begin{aligned}
A &= \widehat{F}(t, x_k, z_{k,l}, \nabla^a_\delta z_{k,l}, \nabla^a_{+,\delta} z_{k,l}, \Delta^a_\delta z_{k,l}) \\
&\quad - \widehat{F}(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^a_{+,\delta} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))), \\
B &= \widehat{F}(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^a_{+,\delta} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))) \\
&\quad - F(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))), \\
C &= F(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) \\
&\quad - F_\epsilon(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)), \Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))).
\end{aligned}$$

First, we have

$$\begin{aligned}
C &= f_\epsilon(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) - f(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) \\
&\geq -|f_\epsilon - f|_\infty.
\end{aligned}$$

Secondly, by (4.3.6)-(4.3.7), we have,

$$\begin{aligned}
B =& -\theta \frac{\delta^2}{h} \Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l)) \\
&+ \mu(\nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)) - \nabla^a_{+,\delta} \varphi(x_k, \mathfrak{p}^a(p_l))) + \frac{\sigma^2}{2}(\Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)) - \Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))) \\
&+ (f(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \sigma \nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l))) - f(t, x_k, \varphi(x_k, \mathfrak{p}^a(p_l)), \sigma \nabla^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l)))) \\
\geq& -\theta \frac{\delta^2}{h} |\Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))| \\
&- \mu |\nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)) - \nabla^a_{+,\delta} \varphi(x_k, \mathfrak{p}^a(p_l))) - \frac{\sigma^2}{2} |\Delta^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)) - \Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))| \\
&- \sigma L |\nabla^{\mathfrak{a}(a,\delta)} \varphi(x_k, \mathfrak{p}^a(p_l)) - \nabla^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))|
\end{aligned}$$

The first term goes to 0 since $\frac{\delta^2}{h} \to 0$ as $h \to 0$ and $\Delta^a_\delta \varphi(x_k, \mathfrak{p}^a(p_l))$ is bounded. The last three terms go to 0 by Taylor expansion and Lemma 4.3.1, since φ is smooth.

Finally, by (4.3.6)-(4.3.7), using the linearity of the discrete differential operators and (4.6.1), and since f is Lipschitz-continuous, we have,

$$
\begin{aligned}
A \geq & - h\mu|\nabla^a_{+,\delta}F_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& - h(\frac{\sigma^2}{2}+\theta\frac{\delta^2}{h})|\Delta^a_\delta F_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& - Lh|F_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& - L\sigma h|\nabla^a_\delta F_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)))|.
\end{aligned}
$$

We can show that each term goes to 0 as $h \to 0$. For example,

$$
\begin{aligned}
& h\frac{\sigma^2}{2}|\Delta^a_\delta F_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)),\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& \quad \geq h\frac{\sigma^2}{2}\mu|\Delta^a_\delta\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l))| \\
& \quad - h\frac{\sigma^4}{4}|\Delta^a_\delta\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l))| \\
& \quad - h\frac{\sigma^2}{2}|\Delta^a_\delta f_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\sigma\nabla^a_\delta\varphi(x_k,\mathfrak{p}^a(p_l)))|.
\end{aligned}
$$

The first two terms go to 0 with h since $|\Delta^a_\delta\nabla^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l))|$ and $|\Delta^a_\delta\Delta^{\mathfrak{a}(a,\delta)}\varphi(x_k,\mathfrak{p}^a(p_l))|$ are bounded, by smoothness of φ and by Lemma 4.3.1.

We can control the derivatives of $\mathfrak{f}_\epsilon : (x,p) \mapsto f_\epsilon(t,x,\varphi(x,p),\sigma\varphi(x,p))$ with respect to ϵ: for any $\alpha = (\alpha_1,\alpha_2) \in \mathbb{N}^2$, we have

$$
|D^\alpha \mathfrak{f}_\epsilon|_\infty \leq \frac{C_{\varphi,\alpha}}{\epsilon^{\alpha_1+\alpha_2}}, \tag{4.6.2}
$$

for a constant $C_{\varphi,\alpha} > 0$. By the triangle inequality and Taylor expansion, we get:

$$
\begin{aligned}
& -h\frac{\sigma^2}{2}|\Delta^a_\delta f_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\sigma\nabla^a_\delta\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& \quad \geq -h\frac{\sigma^2}{2}|(\Delta^a_\delta - \Delta^{\mathfrak{a}(a,\delta)})f_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\sigma\nabla^a_\delta\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& \quad - h\frac{\sigma^2}{2}|\Delta^{\mathfrak{a}(a,\delta)}f_\epsilon(t,x_k,\varphi(x_k,\mathfrak{p}^a(p_l)),\sigma\nabla^a_\delta\varphi(x_k,\mathfrak{p}^a(p_l)))| \\
& \quad \geq -C_1 h\frac{\sigma^2}{2}\frac{\delta^2}{\epsilon^4} - C_2 h\frac{\sigma^2}{2}\frac{1}{\epsilon^2},
\end{aligned}
$$

where $C_1, C_2 > 0$, and this quantity goes to 0 by our choice of ϵ.

Last, the smoothness of $S^{a,\pm}_{\delta,\epsilon}$ is straightforward by (4.3.29) and the control on its second derivative with respect to p is obtained by (4.6.2). □

4.6.2 Representation and comparison results

For $(t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^+$, $q := \begin{pmatrix} q^x \\ q^p \end{pmatrix} \in \mathbb{R}^{d+1}$ and $A := \begin{pmatrix} A^{xx} & A^{xp} \\ A^{xp\top} & A^{pp} \end{pmatrix} \in \mathbb{S}^{d+1}$, $A^{xx} \in \mathbb{S}^d$, denoting $\Xi := (t,x,y,q,A)$, we define, recalling (4.1.4)-(4.1.5)-(4.1.6),

$$
\mathcal{F}(\Xi) = \sup_{a\in\mathcal{R}^\flat} F^a(\Xi) \text{ with } F^a(\Xi) := -f(t,x,y,\mathfrak{z}(x,q,a)) - \mathcal{L}(x,q,A,a)\ ,
$$

where $\mathcal{R} \subset \mathcal{S} \setminus \mathcal{D}$ with a *finite* number of elements.

Proposition 4.6.1. *Let $0 \leq \tau < \theta \leq T$ and u_1 (resp. u_2) be a lower semi-continuous super-solution (resp. upper semi-continuous sub-solution) with polynomial growth, of*

$$-\partial_t \varphi + \mathcal{F}(t, x, \varphi, D\varphi, D^2\varphi) = 0 \quad on \;\; [\tau, \theta) \times \mathbb{R}^d \times (0,1)\,, \tag{4.6.3}$$

with $u_1 \geq u_2$ on $[\tau,\theta] \times \mathbb{R}^d \times \{0,1\} \bigcup \{\theta\} \times \mathbb{R}^d \times [0,1]$, then $u_1 \geq u_2$ on $[\tau,\theta] \times \mathbb{R}^d \times [0,1]$.

Corollary 4.6.1. *There exists a unique continuous solution w to (4.6.3) or equivalently*

$$\sup_{\eta \in \mathcal{R}} H^\eta(t, x, \varphi, \partial_t\varphi, D\varphi, D^2\varphi) = 0 \quad on \;\; [\tau, \theta) \times \mathbb{R}^d \times (0,1)\,, \tag{4.6.4}$$

satisfying $w(\cdot) = \Psi(\cdot)$ on $[\tau,\theta] \times \mathbb{R}^d \times \{0,1\} \bigcup \{\theta\} \times \mathbb{R}^d \times [0,1]$, where $\Psi \in \mathcal{C}^0$.

Proof. This is a direct application of the comparison principles. The equivalence between (4.6.3) and (4.6.4) comes from the fact that $H^\eta(\Theta)$ and $-b - F^{\eta^\flat}(\Xi)$ have the same sign. □

Lemma 4.6.1. *(i) Let $a \in \mathbb{R}^d$ and w^a be the unique solution to*

$$-\partial_t \varphi + F^a(t, x, \varphi, D\varphi, D^2\varphi) = 0 \quad on \;\; [\tau, \theta) \times \mathbb{R}^d \times (0,1)\,,$$

satisfying $w^a(\cdot) = \Psi(\cdot)$ on $[\tau,\theta] \times \mathbb{R}^d \times \{0,1\} \bigcup \{\theta\} \times \mathbb{R}^d \times [0,1]$, where $\Psi \in \mathcal{C}^0$. Then it admits the following probabilistic representation:

$$w^a(t,x,p) = Y_t,$$

where Y is the first component of the solution (Y, Z) to the following BSDE with random terminal time

$$Y_\cdot = \Psi(\mathcal{T}, X^{t,x}_{\mathcal{T}}, P^{t,p,a}_{\mathcal{T}}) + \int_\cdot^{\mathcal{T}} f(s, X^{t,x}_s, Y_s, Z_s)\mathrm{d}s - \int_\cdot^{\mathcal{T}} Z_s \mathrm{d}W_s,$$

with $\mathcal{T} := \inf\{s \geq t : P^{t,p,a}_s \in \{0,1\}\} \wedge \theta$ and

$$\begin{cases} P^{t,p,a}_\cdot = p + a(W_\cdot - W_t) \\ X^{t,x}_\cdot = x + \int_t^\cdot \mu(X^{t,x}_s)\mathrm{d}s + \int_t^\cdot \sigma(X^{t,x}_s)\mathrm{d}W_s\,. \end{cases}$$

(ii) Assume moreover, that $\Psi(T,\cdot) = \phi(\cdot)$ and $\Psi(\cdot,1) = B^1(\cdot,\phi)$, $\Psi(\cdot,0) = B^0(\cdot,\phi)$, with the notation of (4.2.8). Then the solution $(\tilde{Y}, \tilde{Z})$ to

$$Y_\cdot = \phi(X^{t,x}_\theta, \tilde{P}^{t,p,a}_\theta) + \int_\cdot^\theta f(s, X^{t,x}_s, Y_s, Z_s)\mathrm{d}s - \int_\cdot^\theta Z_s \mathrm{d}W_s\,, \tag{4.6.5}$$

where $\tilde{P}^{t,p,a}_\cdot := P^{t,p,a}_{\cdot \wedge \mathcal{T}}$, satisfies

$$Y = \tilde{Y} \;\; on \; [t, \mathcal{T}].$$

Proof. (i) The probabilistic representation is proved in [DP97]. Note that uniqueness to the PDE comes from the previous corollary in the special case where $\mathcal{R}$ is a singleton.
(ii) Let $A := \{\mathcal{T} = \theta\}, B := \{\mathcal{T} < \theta, P^{p,a}_{\mathcal{T}} = 1\}$, and $C := \{\mathcal{T} < \theta, P^{p,a}_{\mathcal{T}} = 0\}$, so that $\Omega = A \cup B \cup C$. For $e \in \{0,1\}$, let $({}^eY^{t,x}, {}^eZ^{t,x})$ the solution to

$$Y_\cdot = \phi(X^{t,x}_\theta, e) + \int_\cdot^\theta f(s, X^{t,x}_s, Y_s, Z_s)\mathrm{d}s - \int_\cdot^\theta Z_s \mathrm{d}W_s \,.$$

By (4.2.8), we have $B^e(\tau, X^{t,x}_\tau, \phi) = {}^eY_\tau$ for $e \in \{0,1\}$.

We introduce the following auxiliary processes, for $s \in [t, \theta]$,

$$\begin{aligned}
\check{Y}_s &:= Y_s 1_{t \le s \le \mathcal{T}} + {}^1Y_s 1_{s > \mathcal{T}} 1_B + {}^0Y_s 1_{s>\mathcal{T}} 1_C, \\
\check{Z}_s &:= Z_t 1_{t \le s \le \mathcal{T}} + {}^1Z_t 1_{s > \mathcal{T}} 1_B + {}^0Z_t 1_{s>\mathcal{T}} 1_C.
\end{aligned}$$

First, by construction, we have $Y = \check{Y}$ on $[t, \mathcal{T}]$. To prove the proposition it is thus sufficient to show that $\tilde{Y} = \check{Y}$ on $[t, \theta]$. To this effect, we show that $(\check{Y}, \check{Z})$ is solution of (4.6.5). We have, for all $s \in [t, \theta]$,

$$\begin{aligned}
\check{Y}_s = &\left[\Psi(\mathcal{T}, X^{t,x}_{\mathcal{T}}, P^{t,p,a}_{\mathcal{T}}) + \int_s^{\mathcal{T}} f(u, X^{t,x}_u, Y_u, Z_u)\mathrm{d}u - \int_s^{\mathcal{T}} Z_u \mathrm{d}W_u\right] 1_{s \le \mathcal{T}} \\
&+ \left[\phi(X^{t,x}_\theta, 1) + \int_s^\theta f(u, X^{t,x}_u, {}^1Y_u, {}^1Z_u)\mathrm{d}u - \int_s^\theta {}^1Z_u \mathrm{d}W_u\right] 1_{\mathcal{T}<s} 1_B \\
&+ \left[\phi(X^{t,x}_\theta, 0) + \int_s^\theta f(u, X^{t,x}_u, {}^0Y_u, {}^0Z_u)\mathrm{d}u - \int_s^\theta {}^0Z_u \mathrm{d}W_u\right] 1_{\mathcal{T}<s} 1_C.
\end{aligned}$$

By our hypotheses and by the definition of ${}^pY_\cdot$, since $\tilde{P}^{t,p,a}_\theta = 1$ on B and $\tilde{P}^{t,p,a}_\theta = 0$ on C, we have

$$\begin{aligned}
\Psi(\mathcal{T}, X^{t,x}_{\mathcal{T}}, P^{t,p,a}_{\mathcal{T}}) &= \phi(X^{t,x}_\theta, P^{t,p,a}_\theta) 1_A + B^1(\mathcal{T}, X^{t,x}_{\mathcal{T}}, \phi) 1_B + B^0(\mathcal{T}, X^{t,x}_{\mathcal{T}}, \phi) 1_C \\
&= \phi(X^{t,x}_\theta, P^{t,p,a}_\theta) 1_A + {}^1Y_{\mathcal{T}} 1_B + {}^0Y_{\mathcal{T}} 1_C \\
&= \phi(X^{t,x}_\theta, \tilde{P}^{t,p,a}_\theta) \\
&\quad + \left[\int_{\mathcal{T}}^\theta f(u, X^{t,x}_u, {}^1Y_u, {}^1Z_u)\mathrm{d}u - \int_{\mathcal{T}}^\theta {}^1Z_u \mathrm{d}W_u\right] 1_B \\
&\quad + \left[\int_{\mathcal{T}}^\theta f(u, X^{t,x}_u, {}^0Y_u, {}^0Z_u)\mathrm{d}u - \int_{\mathcal{T}}^\theta {}^0Z_u \mathrm{d}W_u\right] 1_C.
\end{aligned}$$

Thus, since $A \cap \{\mathcal{T} < s\} = \emptyset$, we deduce

$$\begin{aligned}
\check{Y}_s =& \phi(X_\theta^{t,x}, \tilde{P}_\theta^{t,p,a}) \\
&+ \left[\int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, {}^1Y_u, {}^1Z_u) \mathrm{d}u 1_{s \le \mathcal{T}} 1_B + \int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, {}^0Y_u, {}^0Z_u) \mathrm{d}u 1_{s \le \mathcal{T}} 1_C \right] \\
&+ \left[\int_s^{\mathcal{T}} f(u, X_u^{t,x}, Y_u, Z_u) \mathrm{d}u 1_{s \le \mathcal{T}} + \int_s^{\mathcal{T}} f(u, X_u^{t,x}, {}^1Y_u, {}^1Z_u) \mathrm{d}u 1_{\mathcal{T} < s} 1_B \right. \\
&\qquad \left. + \int_s^{\mathcal{T}} f(u, X_u^{t,x}, {}^0Y_u, {}^0Z_u) \mathrm{d}u 1_{\mathcal{T} < s} 1_C \right] \\
&- \left[\int_{\mathcal{T}}^{\theta} {}^1Z_u \mathrm{d}W_u 1_{s \le \mathcal{T}} 1_B + \int_{\mathcal{T}}^{\theta} {}^0Z_u \mathrm{d}W_u 1_{s \le \mathcal{T}} 1_C \right] \\
&- \left[\int_s^{\mathcal{T}} Z_u \mathrm{d}W_u 1_{s \le \mathcal{T}} + \int_s^{\mathcal{T}} {}^1Z_u \mathrm{d}W_u 1_{\mathcal{T} < s} 1_B + \int_s^{\mathcal{T}} {}^0Z_u \mathrm{d}W_u 1_{\mathcal{T} < s} 1_C \right].
\end{aligned}$$

Now, since $({}^1Y_u, {}^1Z_u) = (\check{Y}_u, \check{Z}_u)$ on $(\mathcal{T}, \theta] \cap B$ and $({}^0Y_u, {}^0Z_u) = (\check{Y}_u, \check{Z}_u)$ on $(\mathcal{T}, \theta] \cap B$, and $\int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, \check{Y}_u, \check{Z}_u) \mathrm{d}u 1_A = 0$, we get

$$\begin{aligned}
\int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, {}^1Y_u, {}^1Z_u) \mathrm{d}u 1_{s \le \mathcal{T}} 1_B &+ \int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, {}^0Y_u, {}^0Z_u) \mathrm{d}u 1_{s \le \mathcal{T}} 1_C \\
&= \int_{\mathcal{T}}^{\theta} f(u, X_u^{t,x}, \check{Y}_u, \check{Z}_u) \mathrm{d}u 1_{s \le \mathcal{T}}.
\end{aligned}$$

A similar analysis for the other terms shows that $(\check{Y}, \check{Z})$ is a solution to (4.6.5) and concludes the proof. □

4.6.3 Finite differences operator S_b

We give relevant results about the finite difference approximation defined by the operator S_b, see (4.3.14), which follow from standard arguments.

Proposition 4.6.2 (Comparison theorem)**.** *Let $0 \le t < s \le T, \delta > 0, h = s - t$ such that (4.3.15) to (4.3.16) is satisfied. Let $(u^1, u^2, v^1, v^2) \in \ell^\infty(\delta\mathbb{Z})^4$ such that $u_k^1 \le u_k^2$ for all $k \in \mathbb{Z}$. Then:*

1. *For all $k \in \mathbb{Z}, (v, \nabla, \nabla_+, \Delta) \in \mathbb{R}^4$:*

$$S_b(k, v, \nabla, \nabla_+, \Delta, u_k^2) \le S_b(k, v, \nabla, \nabla_+, \Delta, u_k^1).$$

2. *Assume that, for all $k \in \mathbb{Z}$:*

$$S_b(k, v_k^1, \nabla_\delta v_k^1, \nabla_{+,\delta} v_k^1, \Delta_\delta v_k^1, u_k^1) \le 0, S_b(k, v_k^2, \nabla_\delta v_k^2, \nabla_{+,\delta} v_k^2, \Delta_\delta v_k^2, u_k^1) \ge 0.$$

   Then $v_k^1 \le v_k^2$ for all $k \in \mathbb{Z}$.

3. *Assume that, for all $k \in \mathbb{Z}$:*

$$S_b(k, v_k^1, \nabla_\delta v_k^1, \nabla_{+,\delta} v_k^1, \Delta_\delta v_k^1, u_k^1) = 0, S_b(k, v_k^2, \nabla_\delta v_k^2, \nabla_{+,\delta} v_k^2, \Delta_\delta v_k^2, u_k^2) = 0.$$

   Then $v_k^1 \le v_k^2$ for all $k \in \mathbb{Z}$.

Proof. The proof is similar to the proof of Proposition 4.3.2 and is omitted. □

Proposition 4.6.3. *Let $\pi, \delta > 0$ such that (4.3.15) to (4.3.16) is satisfied for all $h = t_{j+1} - t_j, j = 0, \dots, \kappa - 1$. Let $V_{\pi,\delta} : \pi \times \delta\mathbb{Z} \to \mathbb{R}$ the solution to:*

$$\begin{aligned} v_k^\kappa &= g(x_k), \quad k \in \mathbb{Z}, \\ S_b(k, V_k^j, \nabla_\delta v_k^j, \nabla_{+,\delta} v_k^j, \Delta_\delta v_k^j, v_k^{j+1}) &= 0, \quad k \in \mathbb{Z}, 0 \le j < \kappa. \end{aligned}$$

For all $k \in \mathbb{Z}$, let $U_{\pi,\delta})_k := \frac{(V_{\pi,\delta})_{k+1} - (V_{\pi,\delta})_{k-1}}{2\delta}$. Then:

1. *$(V_{\pi,\delta}, U_{\pi,\delta}) \in \ell^\infty(\delta\mathbb{Z})^2$ and their bound is independent of π, δ.*
2. *$V_{\pi,\delta}$ converges uniformly on compacts sets to V, the super-replication price of the contingent claim with pay-off g.*

Proof. We only show the first point, the second one is obtained by applying the arguments of [BS91], after proving monotonicity, stability and consistency following the steps of Subsection 4.3.2.

Since g is bounded, it is easy to show that $V_{\pi,\delta}$ is also bounded independently of π, δ, and the proof is similar to the proof of Proposition 4.3.6.

Since g is Lipschitz-continuous, we get that $U_{\pi,\delta}(T, \cdot)$ is bounded. Using the Lipschitz-continuity of f, one deduces easily that $U_{\pi,\delta}$ is a solution of

$$\begin{aligned} &u_k^j - u_k^{j+1} - h\left(-\mu\nabla_{+,\delta} u_k^j - (\frac{\sigma^2}{2} + \theta\frac{\delta^2}{h})\Delta_\delta u_k^j - L - L|u_k^j| - L|\nabla_\delta u_k^j|\right) \ge 0, k \in \mathbb{Z}, 0 \le j < \kappa, \\ &u_k^j - u_k^{j+1} - h\left(-\mu\nabla_{+,\delta} u_k^j - (\frac{\sigma^2}{2} + \theta\frac{\delta^2}{h})\Delta_\delta u_k^j + L + L|u_k^j| + L|\nabla_\delta u_k^j|\right) \le 0, k \in \mathbb{Z}, 0 \le j < \kappa, \\ &u_k^\kappa = \frac{g(x_{k+1}) - g(x_{k-1})}{2\delta} \in [-L, L], k \in \mathbb{Z}. \end{aligned}$$

Again, comparison theorems can be proved, and it is now enough to show that there exists $(\underline{u}, \overline{u}) \in \ell^\infty(\pi \times \delta\mathbb{Z})^2$ which are bounded uniformly in π, δ such that

$$\begin{aligned} &\underline{u}_k^j - \underline{u}_k^{j+1} - h\left(-\mu\nabla_{+,\delta} \underline{u}_k^j - (\frac{\sigma^2}{2} + \theta\frac{\delta^2}{h})\Delta_\delta \underline{u}_k^j - L - L|\underline{u}_k^j| - L|\nabla_\delta \underline{u}_k^j|\right) \le 0, k \in \mathbb{Z}, 0 \le j < \kappa, \\ &\overline{u}_k^j - \overline{u}_k^{j+1} - h\left(-\mu\nabla_{+,\delta} \overline{u}_k^j - (\frac{\sigma^2}{2} + \theta\frac{\delta^2}{h})\Delta_\delta \overline{u}_k^j + L + L|\overline{u}_k^j| + L|\nabla_\delta \overline{u}_k^j|\right) \ge 0, k \in \mathbb{Z}, 0 \le j < \kappa, \\ &\underline{u}_k^\kappa \le -L, \overline{u}_k^\kappa \ge L, k \in \mathbb{Z}. \end{aligned}$$

We deal with $\underline{u}$ only, we obtain similar results for $\overline{u}$. One can easily show that $\underline{u}^j := 1 - (L+1)\prod_{k=j+1}^{\kappa} \frac{1}{1 - h_k L}$, where $h_k := t_k - t_{k-1}$, satisfies the requirements. Furthermore, one gets $\underline{u}^j \ge \underline{u}^0 \ge 1 - (L+1)2^{\frac{T}{2L}}$, thus one gets that u is lower bounded by $1 - (L + 1)2^{\frac{T}{2L}}$. □

Part II

A sparse grid approach to balance sheet measureement

Chapter 5

A sparse grid approach to balance sheet measureement

The content of this chapter is from an article in collaboration with Jérémie Bonnefoy, Jean-François Chassagneux, Shuoqing Deng, Camilo Garcia Trillos and Lionel Lenôtre [Bén+19], published in ESAIM: Proceedings and Surveys.

5.1 Introduction

The goal of this paper is to present a robust and efficient method to numerically assess risks on the balance sheet distribution of, say, an insurance company, at a given horizon. In practice, it is chosen to be one year, consistently with the Solvency 2 regulation, the prudential framework for assessing the required solvency capital for an European insurance company.

On a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, the balance sheet of the company is a random process summarised, at any time $t \geq 0$, by the value of the assets of the company $(A_t)_{t\geq 0}$ and the value of the liabilities $(L_t)_{t\geq 0}$. The quantity of interest is the Profit and Loss (PnL in the sequel) associated to the balance sheet, which is given by

$$P_t = L_t - A_t \;, \quad t \geq 0 \,.$$

By convention, and adopting the point of view of risk management, we measure the loss as a positive quantity.

On the Liability side, the insurance company has sold a structured financial product which depends on the evolution of a one-dimensional stock price (S_t) and the risk-free interest rate (r_t). Several insurance products could be of this type, in particular Unit-Linked (with or without financial guarantees) and Variable Annuity contracts. For those contracts, client's money is invested in equity and bond markets while the insurance company might also provide with financial guarantees similar to long-term put options. The long maturity of those contracts requires the introduction of a model for interest rate as they are very sensitive to Interest Rate curve movements. The value L_1 is just the price of this product taking into account the value of some risk factors $\mathcal{X}_1$ (stock price, interest rate curve etc.) at time $t = 1$ used to calibrate the pricing model.

On the Asset side, the insurance company manages some assets to hedge the risk associated to the product sale. The pricing actually includes a margin which is secured through hedging. The hedging assets are the stock and swaps of several maturities, in practice mostly concentrated on the long term. In practice, bond futures are also included sometimes. The hedging portfolio is typically rebalanced on a weekly basis and the hedging quantities are determined by a financial model, taken to be the same as the liability pricing model, whose inputs are the risk factors $\mathcal{X}_t$ at the time t when the hedge is computed.

We describe precisely in Section 5.2, the pricing and hedging model, the dynamics of the risk factor $\mathcal{X}$ and the value of the asset and liability side of the balance sheet. Let us stress that the risk factor model is given under the so-called *real-world* probability measure $\mathbb{P}$, which might be objectively calibrated using time series of financial markets or represents the management view. This *real-world* model may be –and most of the time is– completely different from the pricing and hedging model which might be simplified for runtime/trackability purposes, prudent (pricing and hedging include a margin) or being constrained by regulation.

Our goal is then to compute various risk indicator for the loss distribution of the balance sheet at one year namely the distribution of P_1 under the *real-world* probability measure $\mathbb{P}$, that we denote hereafter η.

Precisely, we measure the risk associated to η using a (law invariant) risk measure defined over the class of square integrable measures $\varrho : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$. First, we consider for ϱ the so called Value-at-Risk ($V@R$), which is defined by the left-side quantile:

$$V@R_p(\eta) = \inf \{q \in \mathbb{R} \mid \eta((-\infty, q]) \geq p\} \ . \tag{5.1.1}$$

We will also work with the class of *spectral risk measures*: a spectral risk measure is defined as

$$\varrho_h(\eta) = \int_0^1 V@R_p(\eta) h(p) \mathrm{d}p \ , \tag{5.1.2}$$

where h is a non-decreasing probability density on $[0, 1]$. In the numerics, we will focus on the Average Value-at-Risk ($AV@R$) which is given by

$$AV@R_\alpha(\eta) = \frac{1}{1-\alpha} \int_\alpha^1 V@R_p(\eta) \mathrm{d}p \ , \tag{5.1.3}$$

and is a special case of a spectral risk measure.

For a law invariant risk measure ϱ, we denote by $\Re$ its "lift" on $L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}) =: L^2$, namely $\Re[X] = \varrho([X])$ for any $X \in L^2$, where $[X]$ denotes the law of X. The lift $\Re_h$ from a spectral risk measure ϱ_h satisfies the following properties:

1. *Monotonicity*: $\Re_h[X] \leq \Re_h[Y]$, for $X \leq Y \in L^2$;
2. *Cash invariance*: $\Re_h[X + c] = \Re_h[X] + c$ for $X \in L^2$ and $c \in \mathbb{R}$;
3. *Positive homogeneity*: $\Re[tX] = t\Re[X]$, $t \geq 0$ and $X \in L^2$.
4. *Convexity*: $\Re[tX + (1-t)Y] \leq t\Re[X] + (1-t)\Re[Y]$, whenever $0 \leq t \leq 1$, for $X, Y \in L^2$;

Let us stress the fact that $V@R$ only satisfies 1-3. We refer to [Pic13] and the references therein for more insights on risk measures and spectral risk measures.

In our setting, the loss distribution η of the balance sheet PnL is obtained through the following expression:

$$\eta = p_1 \sharp \nu \ ,$$

where $\sharp$ denotes the push-forward operator, $p_1 : \mathbb{R}^\theta \to \mathbb{R}$ is the *function* describing the PnL in terms of the risk factors, and ν stands for the distribution of the risk factors $\mathcal{X}$. In practice the estimation of $\varrho(\eta)$ requires to sample from η. In turn, this demands for a sample of the model parameter distribution ν and for a numerical approximation of p_1. In this note, we compare two main approaches to form the sample of η given one of ν.

The first one is known as the *nested simulation* approach: It is a two-step method. First, a set of "outer simulation", describing the random values of the risk factors, is drawn. Then, for each value of the risk factors, a sample of "inner simulation" is drawn to compute the various hedge and prices. In this approach, all computations are realised "online". The main advantage of this approach is its simplicity to implement in practice, described in the first paragraph of Subsection 5.3.1. However, it is well known that this approach is quite greedy, even if optimised as in [GJ10]. We also want to stress the fact that when computing the η-sample, no information about p_1 is stored for future work: for example if ν is modified, due to time or a model change, a full recalculation would be required.

The other approach we chose to adopt and would like to promote is a *grid approach* where the approximation of p_1 is made "offline", by a Monte Carlo approach, and then stored. The numerical computation is then done through a (multi-linear) interpolation on a grid. The main drawback of this approach is that the size of the grid, in high dimension, can become untractable, especially if one uses regular grid. To partially circumvent this difficulty, we introduce a *sparse grid* [BG04] which reduces drastically the number of point to be used (equivalently, values to be stored) with only small reduction of the accuracy of the method.

We prove that, for a spectral risk measure, the two approaches give an estimation of $\varrho(\eta)$ which converges to the true value, see Theorem 5.3.1.

Furthermore, we show in the numerical Section 5.4 that using the grid approach together with a sparse grid of low level allows to get a good approximation of the loss distributions η, and of some related risk measures, while reducing drastically the computational time and allowing to keep information about the balance sheet function p_1. Last, this permits to numerically quantify uncertainty. Indeed, since the computations on the grid are stored, the computation of the distribution of the PnL under other distributions for the parameters is almost instantaneous and can be compared with the results obtained with the initial one. An application to uncertainty estimation is given in the last numerical application.

The rest of the paper is organised as follows. In the Section 2, we first describe the mathematical models that are used to describe the evolution of the prices under the risk-neutral measure $\mathbb{Q}$. We then describe precisely how A and L are specified. In Section 3, we describe the two numerical methods used to compute L_t and A_t at any given time $t \geq 0$. In particular, we show how to efficiently compute, at time t, the quantities to hold in the hedging portfolio, which are expressed in term of the derivatives of the claim's price. We also explain how to compute the price of the product and of the assets used to construct the hedging portfolio, leading to the computation of L_t and A_t. We show how to obtain an approximation of the distribution of P_1 under the physical measure $\mathbb{P}$, and we prove an upper bound for the mean square error of the overall procedure. Finally, in Section 4, we present our numerical results, comparing the two methods.

5.2 Financial Model

In this section, we give the precise specification of the asset and liability sides of the balance sheet. We also present the *risk-neutral* model and the *real-world* model that are used.

5.2.1 Description of the sold product

Let us assume that a company sells a contingent claim at time $t = 0$ which is a (discretely) path-dependent option with a pay-off function G paid at the maturity $T > 0$, depending upon the evolution of a one-dimensional risky asset's price S. We focus here on:
A put lookback option, that is a discretely path-dependent option whose strike at maturity T is given by the maximum of the asset's price S over the times $t \in \{\tau_0 = 0, \tau_1, \cdots, \tau_\kappa = T\}$ where $\kappa \geq 1$:

$$G(S_{\tau_0}, \ldots, S_{\tau_\kappa}) = \left(\max_{0 \leq \ell \leq \kappa} S_{\tau_\ell} \right) - S_T. \tag{5.2.1}$$

Remark 5.2.1. *The proxy provided above is close to financial guarantees offered in Variable Annuity contracts. Those contracts are structured insurance products composed of a fund investment on top of which both insurance and financial protection are added. In our case, the contract is a* Guaranteed Minimum Accumulation Benefit including a ratchet mechanism. *At time* $t = 0$, *the customer invests his/her money in the underlying fund and will receive at a given maturity the maximum between the terminal fund value and its* terminal benefit base *in case she is still alive. The terminal benefit base is equal to the maximum of the underlying fund values observed at each anniversary date of the contract (ratchet mechanism). We do not consider the modelling of death/survival in this proceedings, neither the possibility that client can surrender at any time during the life of the contract.*

5.2.2 Market model under the risk-neutral measure

We assume that all pricing and hedging is done with a market risk-neutral measure $\mathbb{Q}$.

The derivative with a pay-off function G as above depends upon a one-dimensional stock's price $S = (S_t)_{t\in[0,T]}$. We assume here that the dynamics of the asset under $\mathbb{Q}$ are of the Black & Scholes type as described in Section 5.2.2.2 with a stochastic interest rate $r = (r_t)_{t\in[0,T]}$ which follows a Hull & White model.

As the pay-off G is a proxy of Variable Annuity guarantee which is a long term Savings product (in practice maturity ranges from 10 to 30 years depending on product type), the modelling of interest rate is essential as the product and therefore the overall balance sheet of the company is very sensitive to this risk.

5.2.2.1 The short rate model

Let $\Theta \in \mathbb{R}^d$ ($d := 3$ in the sequel) be a set of parameters representing some market observations. The short rate evolution is governed by the Hull & White dynamics

$$r_s^{t,\Theta} = r_t^{t,\Theta} + \int_t^s a\left(\mu_u^{t,\Theta} - r_u^{t,\Theta}\right) \mathrm{d}u + b\left(B_s - B_t\right), \; s \in [t,T], \tag{5.2.2}$$

where B is a $\mathbb{Q}$-Brownian motion, a and b are real constants and $\mu^{t,\Theta} : [t,T] \to \mathbb{R}$ is a function. We refer to [BM07] for a more complete analysis of the Hull & White short rate model.

The parameter $\mu^{t,\Theta}$ is calibrated using the market observations Θ, so that the model reproduces the interest rate curve observed on the market. It is given by

$$\mu_s^{t,\Theta} = f^\Theta(t,s) + \frac{1}{a}\frac{\partial f^\Theta(t,s)}{\partial s} + \frac{b^2}{2a^2}\left(1 - e^{-2a(s-t)}\right), s \in [t,T]. \tag{5.2.3}$$

We refer to the Appendix for a derivation of (5.2.3).

As a consequence, the Θ parameter must be chosen in order to represent adequately the forward rate curve observed on the market.

We suppose here that the forward rate curve $f^\Theta(t,\cdot)$ is directly observed and is a linear combination of three elementary functions $h^{t,1}, h^{t,2}, h^{t,3}$ from $[t,T]$ to $\mathbb{R}$, given by

$$h^{t,1}(s) := h^1(s-t), \; h^{t,2}(s) := h^2(s-t), \text{ and } h^{t,3}(s) := h^3(s-t), \; s \in [t,T],$$

where, for $u \in [0,T]$:

$$h^1(u) = \begin{cases} 1 & \text{if } u \le \frac{t_1+t_2}{2}, \\ 2\frac{\frac{t_2+t_3}{2} - u}{t_3 - t_1} & \text{if } u \in [\frac{t_1+t_2}{2}, \frac{t_2+t_3}{2}], \\ 0 & \text{otherwise,} \end{cases} \quad h^3(u) = \begin{cases} 0 & \text{if } u \le \frac{t_2+t_3}{2} \\ 2\frac{u - \frac{t_2+t_3}{2}}{t_4 - t_2} & \text{if } u \in [\frac{t_2+t_3}{2}, \frac{t_3+t_4}{2}], \\ 1 & \text{otherwise,} \end{cases}$$

$$\text{and } h^2(u) = 1 - h^1(u) - h^3(u),$$

where $0 \le t_1 < t_2 < t_3 < t_4 \le T$ are four fixed real numbers.

The function $h^{t,1}$ (resp. $h^{t,2}, h^{t,3}$) model the short (resp. middle, long) term structure of the interest rates curve.

In a nutshell, the short rate model is determined by:

1. the time of observation $t \in [0,T]$,
2. the three-dimensional parameter $\Theta := \{\theta_1, \theta_2, \theta_3\} \in \mathbb{R}^3$, where $\theta_1, \theta_2, \theta_3$ are such that
$$f^\Theta(t,\cdot) = \theta_1 h^{t,1} + \theta_2 h^{t,2} + \theta_3 h^{t,3}, \tag{5.2.4}$$
$f^\Theta(t,\cdot)$ being the observed forward rates curve.

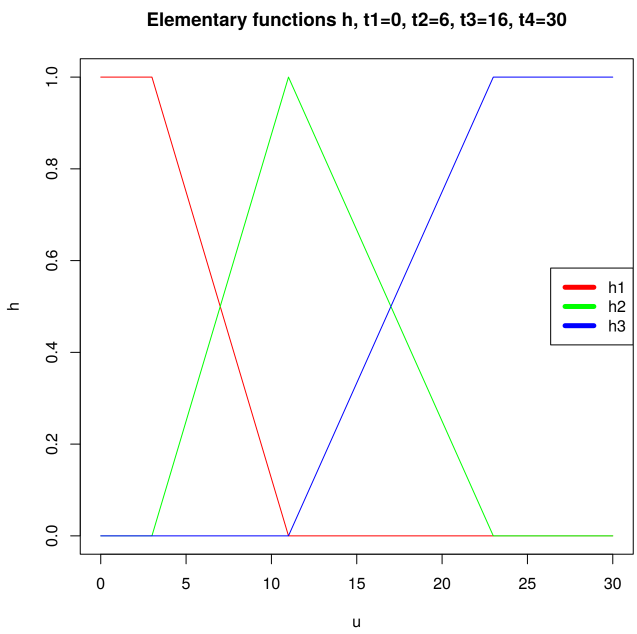

Figure 5.1 – Building blocks for the forward interest rate curve.

As a result, given an observation (t, Θ) as above, the short rate process $r^{t,\Theta}$ has the dynamics (5.2.2), where $\mu^{t,\Theta}$ is computed using (5.2.3).

Remark 5.2.2. *In practice, the parameters a, b appearing in (5.2.2) can be calibrated so that the model reproduces the prices, observed on the market, of some contracts such as swaps or swaptions. We could more generally allow the parameters a, b of the Hull & White model to depend upon the market observations Θ. The parameter Θ should live in a higher-dimensional space to take into account the observed swap(tion)s prices. Regular recalibration of parameters is largely performed by practitioners, in particular when they perform dynamic hedging.*

There are several reasons explaining the choice of this model. First of all, it is quite simple to calibrate using the data. In fact, the function μ is directly given as a function of the forward rate curve. We should note again that the choice of keeping a, b fixed through time simplifies the calibration. Secondly, we will see later in Proposition 5.3.1 that this short rate model, associated to the stock model described below, leads to an exact simulation under the risk-neutral measure. Lastly, closed and easily tractable formulas can be obtained for the prices of the zero-coupon bonds and swaps which are the products used to construct the hedging portfolio and then to compute the value of the company's assets A.

These prices are as follows.

Proposition 5.2.1. *Let $(t, \Theta) \in [0, 1] \times \mathbb{R}^3$ be a market observation, and consider the process $\left(r_s^{t,\Theta}\right)_{s \in [t,T]}$ given by (5.2.2), where the parameter $\mu^{t,\Theta}$ is defined with (5.2.3) and (5.2.4).*

1. *The price at time t of a zero-coupon maturing at time $u \in [t,T]$ is given by:*

$$P^{t,\Theta,u} = \exp\left(-\int_t^u f^{\Theta}(t,s)\mathrm{d}s\right),$$

and its derivatives with respect to $\Theta := (\theta_1,\theta_2,\theta_3)$ are given by:

$$\frac{\partial P^{t,\Theta,u}}{\partial \theta_i} = -P^{t,\Theta,u}\int_t^u h^{t,i}(s)\mathrm{d}s.$$

2. *Let $(0,\Theta_0)$ be the observation made at time 0. Consider a swap contract issued in $s=0$, with maturity $M>0$, rate $R>0$, with coupons versed at every time $i \in \{1,\dots,M\}$. Then, the price of this contract at time t is given by:*

$$SW^{t,\Theta,M,R} = \frac{P^{t,\Theta,1}}{P^{0,\Theta_0,1}} - P^{t,\Theta,M} - R\sum_{i=1}^{M} P^{t,\Theta,i},$$

and its derivatives with respect to Θ are given by:

$$\frac{\partial SW^{t,\Theta,M,R}}{\partial \theta_j} = -\frac{P^{t,\Theta,1}}{P^{0,\Theta_0,1}}\int_t^1 h^{t,j}(s)\mathrm{d}s + P^{t,\Theta,M}\int_t^M h^{t,j}(r)\mathrm{d}r + R\sum_{i=1}^{M}\left(P^{t,\Theta,i}\int_t^{t+i} h^{t,j}(s)\mathrm{d}s\right).$$

5.2.2.2 The stock model

Given the observations Θ of the interest rate factors and the risky asset's price $x \in (0,\infty)$, the evolution of the price under the neutral-risk measure $\mathbb{Q}$ is given by

$$S_s^{t,x,\Theta} = x + \int_t^s r_u^{t,\Theta} S_u^{t,x,\Theta}\mathrm{d}u + \int_t^s \sigma S_u^{t,x,\Theta}\mathrm{d}\tilde{W}_u, \quad s \in [t,T], \tag{5.2.5}$$

where $\sigma > 0$, $\tilde{W}$ is another $\mathbb{Q}$-Brownian motion, whose quadratic covariation with B is given by

$$\langle B,\tilde{W}\rangle_t := \rho\, t, \quad t \in [0,T],$$

where $\rho \in [-1,1]$. Equivalently, $S^{t,x,\Theta}$ can be written as:

$$\begin{aligned} S_s^{t,x,\Theta} = x + \int_t^s r_u^{t,\Theta} S_u^{t,x,\Theta}\mathrm{d}u &+ \int_t^s \rho\sigma S_u^{t,x,\Theta}\mathrm{d}B_u \\ &+ \int_t^s \sqrt{1-\rho^2}\sigma S_u^{t,x,\Theta}\mathrm{d}W_u, \quad s \in [t,T], \end{aligned} \tag{5.2.6}$$

where W is a $\mathbb{Q}$-Brownian motion, independent of B.

Remark 5.2.3. *In practice, the parameter σ appearing in (5.2.6) can be calibrated so that the model reproduces the prices of some derivatives over the risky asset. This can be taken into account by increasing the dimension of the space where Θ lives, and by adding this calibration procedure.*

Remark 5.2.4. *Naturally, the general sparse grid approach can be applied to different models and functional representations. We made the choice of using a Black & Scholes model for the stock value and a Hull & White model with the given functional representation in terms of h^1,h^2,h^3 for the short rate, since they are convenient to obtain explicit pricing and sensitivities formulae as we show in the following.*

5.2.3 Modelling the Balance Sheet

The key point for us is to approximate the distribution of the balance sheet of an insurance company at time $t = 1$ (here a year) given the market observations at $t = 0$. As mentioned in the introduction, the PnL is a process P which can be decomposed as

$$P_t = L_t - A_t,\ t \in [0, 1],$$

where L is the value of the liabilities of the company and A is the value of the assets. We assume that at time $t = 0$, the balance sheet is clear, meaning that the company has no asset nor liability, that is $L_0 = A_0 = 0$.

We describe precisely in the two subsequent sections how these quantities are defined. Importantly, we denote by $\bar{\mathcal{X}}_t := (\bar{S}_t, \bar{\Theta}_t)$, $0 \leq t \leq 1$, the stochastic process representing the random evolution of the market parameter under the *real-world* measure $\mathbb{P}$. Namely, $\bar{S}$ is the stock price and $\bar{\Theta}$ the interest rate curve parameters as described above in Section 5.2.2.1. It is important to have in mind that the model chosen for the stock price S (under $\mathbb{Q}$) and $\bar{S}$ (under $\mathbb{P}$) will be completely different as they do not serve the same purpose (pricing-hedging on one hand, risk management of the Balance Sheet or regulatory assessment of required capital on the other hand).

5.2.3.1 Liability side

For any market observation $\bar{\mathcal{X}}_t := (\bar{S}_t, \bar{\Theta}_t)$, the value $L_t =: \ell(t, \bar{S}_t, \bar{\Theta}_t)$ of the liabilities has to be computed, especially at time $t = 1$ in our application. The company's liabilities are reduced to one derivative product sold at $t = 0$. In our setting, the contingent claim's price is simply given by:

$$\ell(t, x, \Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta} \mathrm{d}s} G(S^{t,x,\Theta})\right], \tag{5.2.7}$$

where $(r^{t,\Theta}, S^{t,x,\Theta})_{t \leq s \leq T}$ are the risk neutral dynamics of the short rate and stock price, see Section 5.2.2.1 and Section 5.2.2.2, calibrated from the observed market parameter (x, Θ) at time t.

We recall that the pay-off G depends on $S^{t,x,\Theta}$ only through the values $S_\tau^{t,x,\Theta}$, $\tau \in \Gamma_G := \{\tau_0, \ldots, \tau_\kappa\}$ $(\kappa \geq 0)$, see (5.2.1).

As explained in more detail below, the computation of P_1 first requires to approximate $\ell(t, x, \Theta)$ for (t, x, Θ) on a (possibly stochastic) discrete grid of $[0, 1] \times (0, \infty) \times \mathbb{R}^3$. This approximation L at any point $(t, x, \Theta) \in [0, 1] \times (0, \infty) \times \mathbb{R}^3$ basically follows from the simulation of the processes $r^{t,\Theta}$ and $S^{t,x,\Theta}$ under the risk-neutral measure $\mathbb{Q}$ and a Monte Carlo procedure. We will see in Section 5.3 that the simulation can be done in an exact manner in our model.

Remark 5.2.5. *A classical approach to compute ℓ would be to use a dynamic programming principle. Step by step, it requires*

1. *To numerically obtain $\ell(1, x, \Theta)$ for all (x, Θ) on the grid with (5.2.7),*
2. *Then to iteratively compute $\ell(t_{k+1}, x, \Theta)$ for all x, Θ on the grid using*

$$\ell(t_k, x, \Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t_k}^{t_{k+1}} r_s^{t_k,\Theta} \mathrm{d}s} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t_{k+1}}^{T} r_s^{t_k,\Theta} \mathrm{d}s} G(S^{t_k,x,\Theta}) | \mathcal{F}_{t_{k+1}}\right]\right]. \tag{5.2.8}$$

However, the inner conditional expectation is not *of the form $\ell(t_{k+1}, x, \Theta)$ for $x > 0$ and $\Theta \in \mathbb{R}^3$. In fact, the time of the market observation Θ is still t_k, while the discount factor goes only from t_{k+1} to T, in contrast with (5.2.7). Therefore, using the dynamic programming equation (5.2.8) would require to introduce some additional and artificial parameters, namely the time of calibration and the value of $r^{t_k,\Theta}$ at time t_{k+1}. This would made the overall procedure heavier that is why we compute ℓ using simply (5.2.7).*

5.2.3.2 Asset side

The company wants to replicate the product with pay-off G[1]. The classical theory of mathematical finance ensures that it is equivalent, in theory, to possess a portfolio which negates the variations of the price of the product with respect to the evolution of the underlying parameters.
In our context, the insurer wants to be protected against the variations with respect to the stock price S_t and the interest rate curve, which is modelled through the parameter Θ.
The dynamic hedging portfolio is constructed and rebalanced in discrete time, on the time grid $\Gamma := \{t_0 = 0 < t_1 < \cdots < t_n = 1\}$ (in practice, the portfolio will be rebalanced up to the maturity of the product, but in our setting, we are only interested in the portfolio's value up to $t = 1$). At each time $t \in \Gamma$, the insurer computes the derivatives of the price with respect to S_t and Θ_t, and then buys some financial assets (the stock and swaps) in order to construct a portfolio whose derivatives match those computed.
To model this framework, we decompose the hedging portfolio's value A in two parts:

$$A_t = A_t^{\Delta} + A_t^{\rho}.$$

The process A^{Δ} is the value of the portfolio obtained to cancel the variations of the price with respect to S, while A^{ρ} is defined to deal with the variations with respect to Θ.

Remark 5.2.6. *Obviously, since in practice the hedging is done in discrete time and some underlying parameter are not considered, the pay-off G is not exactly replicated, nor super-replicated. Therefore the PnL of the company is not null, nor always positive, and the goal of this proceedings is precisely to propose a new numerical method to estimate the distribution of this quantity at time $t = 1$.*

To construct the hedging portfolio, the insurer can buy the underlying stock, together with three swap contracts, defined by some rates $R_1, R_2, R_3 > 0$ and maturity dates $T_1, T_2, T_3 \in [1, T]$. Interest rate hedging is performed so that the portfolio is insensitive to the variations of the main maturities of the interest rate curve. For long-term products, this means building an hedging portfolio containing several different maturities from 1 year to 30 year. Here, only 3 maturities representing short, medium and long-term part of the curve are considered for simplicity. The formula for their price $\mathrm{SW}^{t,\Theta,T_i,R_i}, i = 1, 2, 3$ is given in Proposition 5.2.1 above. We now describe how to compute the quantities of assets and swaps to buy at a time $t \in \Gamma$, to rebalance the

[1] In practice, the pricing embeds a margin and the objective of replicating the pay-off G is to secure it.

hedging portfolio. Denote by Δ (resp. $\rho_i, i=1,2,3$) the quantities of stock (resp. swap with rate R_i and maturity date T_i, $i=1,2,3$). Then the value of the portfolio of the company is given by:

$$\Pi_t = \Delta S_t + \sum_{i=1}^{3} \rho_i \mathrm{d}SW^{t,\Theta_t,T_i,R_i} - \ell(t,S_t,\Theta_t).$$

By Itô's formula, assuming a semi-martingale decomposition for the process Θ under $\mathbb{P}$, we get:

$$\begin{aligned}
\mathrm{d}\Pi_t &= (\Delta - \Delta(t,S_t,\Theta_t))\,\mathrm{d}S_t \\
&+ \left(\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_1} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_1} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_1} - \frac{\partial \ell}{\partial\theta_1}(t,x,\Theta)\right)\mathrm{d}\theta_1 \\
&+ \left(\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_2} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_2} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_2} - \frac{\partial \ell}{\partial\theta_2}(t,x,\Theta)\right)\mathrm{d}\theta_2 \\
&+ \left(\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_3} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_3} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_3} - \frac{\partial \ell}{\partial\theta_3}(t,x,\Theta)\right)\mathrm{d}\theta_3 \\
&+ \mathrm{d}t \text{ terms.}
\end{aligned}$$

To cancel the risks induced by the variations of the stock price and the interest rate curve, it is needed that the four first terms in the previous equation cancel.
Those considerations lead to the following construction for the hedging portfolio $A = A^\Delta + A^\rho$:

Δ-hedging: The value of A^Δ at time 1 is

$$A_1^\Delta = \sum_{i=0}^{n-1} \Delta(t_i,\bar{S}_{t_i},\bar{\Theta}_{t_i})\left(\bar{S}_{t_{i+1}} - \bar{S}_{t_i}\right) \quad \text{where} \quad \Delta(t,x,\Theta) := \frac{\partial L}{\partial x}(t,x,\Theta). \tag{5.2.9}$$

Note that the function Δ has to be computed, at each time $t_i, i=0,\dots,n$, and for any market situation (x,Θ) at this time. A method leading to a numerical estimation of Δ is proposed in Section 5.3.

ρ-hedging: The value of A_1^ρ is

$$A_1^\rho = \sum_{i=0}^{n-1}\sum_{j=1}^{3} \rho^j(t_i,\bar{S}_{t_i},\bar{\Theta}_{t_i})\left(\mathrm{SW}^{t_{i+1},\bar{\Theta}_{t_{i+1}},T_j,R_j} - \mathrm{SW}^{t_i,\bar{\Theta}_{t_i},T_j,R_j}\right). \tag{5.2.10}$$

At time $t \in [0,T]$, for a market at (x,Θ), the quantities $\rho^i(t,x,\Theta)$, $i=1,2,3$ of each swap contract required for the hedging are given by the solution of the linear system

$$\begin{cases}
\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_1} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_1} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_1} &= \frac{\partial \ell}{\partial\theta_1}(t,x,\Theta) \\
\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_2} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_2} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_2} &= \frac{\partial \ell}{\partial\theta_2}(t,x,\Theta) \\
\rho_1 \frac{\partial \mathrm{SW}^{t,\Theta,T_1,R_1}}{\partial\theta_3} + \rho_2 \frac{\partial \mathrm{SW}^{t,\Theta,T_2,R_2}}{\partial\theta_3} + \rho_3 \frac{\partial \mathrm{SW}^{t,\Theta,T_3,R_3}}{\partial\theta_3} &= \frac{\partial \ell}{\partial\theta_3}(t,x,\Theta)
\end{cases} \tag{5.2.11}$$

One key quantity to compute for us in this setting is thus the vector of sensitivities $(\frac{\partial \ell}{\partial\theta_i}(t,x,\Theta))$, $i=1,2,3$.

Remark 5.2.7. *(i) We choose to always use the same swap contracts issued at $t = 0$ as hedging instruments. We could have decided to enter for free in swaps (at the swap rate) at each rebalancing time. However, this strategy requires to keep the memory of the swap rate in order to compute the swap price at the next rebalancing date.*
(ii) The T_i, R_i should be chosen so that they represent some liquid contracts.

5.2.3.3 The global PnL function

From the previous two sections, we conclude that the PnL of the balance sheet at time 1 can be expressed as,

$$P_1 = p_1\left((t, \bar{S}_t, \bar{\Theta}_t)_{t\in\Gamma}\right)$$

where $(\bar{S}, \bar{\Theta})$ are the market parameters (risk factors) and the PnL function $p_1 : \mathbb{R}^\gamma \to \mathbb{R}$, with $\gamma = 4 \times (n+1)$, is given by

$$\ell(t_n, x_n, \Theta_n) - \sum_{i=0}^{n-1} \Delta(t_i, x_i, \Theta_i)(x_{i+1} - x_i) - \sum_{i=0}^{n-1}\sum_{j=1}^{3} \rho^j(t_i, x_i, \Theta_i)\left(\mathrm{SW}^{t_{i+1},\Theta_{i+1},T_j,R_j} - \mathrm{SW}^{t_i,\Theta_i,T_j,R_j}\right). \tag{5.2.12}$$

In the next section, we describe the model we will consider for the market parameter $(\bar{S}, \bar{\Theta})$.

5.2.4 Market parameters under the *real-world* measure

We describe here the model that will be used for the simulation of the market parameters in the numerical part. Let us insist that this *real-world* measure $\mathbb{P}$ might represent the view of the management on the evolution of the market parameter on the period $[0, 1]$. As already mentioned, it can be completely different from the model used for the *risk-neutral* pricing.

We assume that we know, or at least are able to estimate, the first two moments of the distribution of $(X_1 := \log(S_1), \Theta_1) = (X_1, (\theta_1)_1, (\theta_2)_1, (\theta_3)_1)$ under $\mathbb{P}$. More precisely, we assume that under $\mathbb{P}$, this random vector has mean and covariance matrix given by

$$\mu = (\mu_X, \mu_1, \mu_2, \mu_3)\,,$$
$$V = (V_{ij})_{i,j=0,1,2,3}.$$

To model the random process $(X_t, (\theta_1)_t, (\theta_2)_t, (\theta_3)_t)_{t\in[0,1]}$ under $\mathbb{P}$, we assume that its dynamics are given by

$$\begin{aligned} X_t &= X_0 + b_0 t + c_{00}W_t^0 + c_{01}W_t^1 + c_{02}W_t^2 + c_{03}W_t^3,\\ (\theta_1)_t &= (\theta_1)_0 + b_1 t + c_{11}W_t^1 + c_{12}W_t^2 + c_{13}W_t^3\\ (\theta_2)_t &= (\theta_2)_0 + b_2 t + c_{22}W_t^2 + c_{23}W_t^3\\ (\theta_3)_t &= (\theta_3)_0 + b_3 t + c_{33}W_t^3, \end{aligned}$$

where $W^i, i = 0, 1, 2, 3$ are independent $\mathbb{P}$-Brownian motions.

Proposition 5.2.2. *There exists at most one set of coefficients* $b_i, c_{ij}, i, j = 0, 1, 2, 3$, *such that the random vector* $(X_1, (\theta_1)_1, (\theta_2)_2, (\theta_3)_2)$ *has mean* μ *and covariance matrix* V.

Proof. We refer to the appendix for a proof, cf. Proposition 5.5.1. □

Remark 5.2.8. *It is well-known that it is difficult to estimate accurately the drift parameter in a Black-Scholes model. This makes our computation subject to model risk. We leave it to a future research work to find a robust way to approximate the law of* P_1 *under* $\mathbb{P}$. *Nevertheless, let us point out that the grid approach allows us to compute, with minimal re-computation, risk measures for different approximations of the law of* P_1. *This is one of the advantages of this method with respect to using "nested simulations", as illustrated in Section 5.4.*

5.3 Numerical methods

In this section, we describe the two numerical methods that we use to compute the risk indicator on the balance sheet PnL. The first one is known as *nested simulation* approach and is already used in the industry, see the seminal paper [GJ10]. The second one is a *sparse grid* approach and is designed to be more efficient that the *nested simulation* approach in the case of moderate dimensions. In the next section, we present numerical simulations that confirms this fact for the model with moderate dimension that we consider here.

5.3.1 Estimating the risk measure

Given a risk measure ϱ and the loss distribution η of the balance sheet at one year, we estimate the quantity of interest $\varrho(\eta)$ by simulating a sample of N i.i.d random variables $(\Psi_j)_{1\le j\le N}$ with law η and then computing simply $\varrho(\eta^N)$ using the formulae (5.1.1), (5.1.2) and (5.1.3) with η^N instead of η. Here, η^N stands for the empirical measure associated to the Ψ_j i.e.

$$\eta^N = \frac{1}{N}\sum_{j=1}^{N} \delta_{\Psi_j} ,$$

where δ_x is the Dirac mass at the point x.

Recall, that in our work, the loss distribution η is obtained through the following expression:

$$\eta = p_1 \sharp \nu$$

p_1 being described in (5.2.12) and ν stands for the distribution of the market parameters. Namely, ν is the law of the random variable

$$\bar{\mathcal{X}} := (\bar{S}_t, \bar{\Theta}_t)_{t\in\Gamma}$$

under the *real world* probability measure $\mathbb{P}$.

In order to estimate $\varrho(\eta)$ for a chosen risk measure, we need to be able to sample from η which implies two steps in our setting. First, we need to be able to sample $\bar{\mathcal{X}}$ and then we use an approximation p_1^{υ} of p_1, $\upsilon \in \{\mathcal{N}, \mathcal{S}\}$, namely:

- $p_1^{\mathcal{N}}$ if one chooses the *nested simulation* approach;
- $p_1^{\mathcal{S}}$ if one chooses the *sparse grid* approach.

Eventually, the estimator of $\varrho(\mu)$ is given by

$$R^{\upsilon} := \varrho(p_1^{\upsilon}\sharp\nu^N) \ , \quad \text{for } \upsilon \in \{\mathcal{N}, \mathcal{S}\} \, .$$

To summarise, the two numerical methods have the following steps.

Nested simulation approach

1. Outer step: Simulate the model parameters $(\mathcal{X}_j)_{j=1,\ldots,N}$.
2. Inner step: Simulate $\Psi_j = p^{\mathcal{N}}(\mathcal{X}_j)$ using MC simulations. This requires to compute the option prices with Monte Carlo estimates, the interest rate derivative prices, and the various sensitivities of the price, see subsection 5.3.2.2. All these computations are done using closed-form formulae that are derived below.
3. Estimate the risk measure.

All computations are made "online".

Sparse grid approach

1. Fix a (sparse) grid $\mathcal{V}$ and compute the approximation $p^{\mathcal{S}}$ at each required value on the grid by an MC simulation. This involves exactly the same computations as 2. above.
2. Simulate the N model parameter samples $(\mathcal{X}_j)$ and evaluate $\Psi_j = p^{\mathcal{S}}(\mathcal{X}_j)$.
3. Estimate the risk measure.

Computations at step 1. are done "offline". The next two steps are done "online".

We now describe precisely how to compute $p^{\mathcal{N}}$ and $p^{\mathcal{S}}$.

5.3.2 Nested Simulation approach

In the *nested simulation* approach, once the market parameters $\bar{\mathcal{X}}$ have been simulated, the function $p_1^{\mathcal{N}}$ has itself to be computed. Recalling (5.2.12), this requires to evaluate the function $\ell^{\mathcal{N}}, \Delta^{\mathcal{N}}, \rho^{\mathcal{N}}$ (approximations of ℓ, Δ, ρ) at the value of the market parameters. This computation is made using again a Monte Carlo approach.

In order to compute the risk-neutral expectations in the above formulae, one has to sample the short rate process and compute its integral over $[t, T]$, and also to simulate the stock price S at least at the times $\tau_\ell \in \Gamma_G \cap [t, T]$. Under the model described in section 5.2.2, the simulation is exact and is described in the two following statements whose proofs are postponed to the appendix.

Let $(t, x, \Theta) \in [0, T] \times (0, \infty) \times \mathbb{R}^3$ be a market observation, and consider the processes $\left(r_s^{t,\Theta}\right)_{s\in[t,T]}$ and $\left(S_s^{t,x,\Theta}\right)_{s\in[t,T]}$ defined by (5.2.2), (5.2.6). We first start with an easy and well known observation.

Lemma 5.3.1. *We have the following decomposition for the short rate process:*

$$r^{t,\Theta} = \xi^t + \alpha^{t,\Theta},$$

where $(\alpha_s^{t,\Theta})_{s\in[t,T]}$ *is the deterministic function of time*

$$\alpha_s^{t,\Theta} = \frac{1}{a} f^{\Theta}(t,s) + \frac{b^2}{2a^3}\left[1 - e^{-a(s-t)}\right]^2, s \in [t,T], \tag{5.3.1}$$

and $\left(\xi_s^t\right)_{s\in[t,T]}$ *is the solution of the SDE*

$$\begin{aligned} d\xi_s^t &= -a\xi_s^t \mathrm{d}s + b\mathrm{d}B_s, s \in [t,T] \\ \xi_t^t &= 0. \end{aligned} \tag{5.3.2}$$

The following proposition provides a recursive way to produce sample paths of the triplet $(\xi^t, A^t, X^{t,x,\Theta})$ on the discrete grid $\{t\} \cup (\Gamma_G \cap [t,1])$. We are thus in a position to simulate the evolution of $(r^{t,\Theta}, S^{t,x,\Theta})$ and the discount factor $\beta_T^{t,\theta} := e^{-\int_t^T r_s^{t,\Theta} \mathrm{d}s}$ under the risk-neutral measure $\mathbb{Q}$.

Proposition 5.3.1. *Fix* $t \le s \le u \le T$*. Conditionally upon* $\mathcal{F}_s$*, the triplet*

$$\left(\xi_u^t, A_u^{t,s} := \int_s^u \xi_r^t \mathrm{d}r, X_u^{t,x,\Theta} := \log\left(S_u^{t,x,\Theta}\right)\right)$$

is a Gaussian vector of dimension 3*, with mean vector and covariance matrix respectively given by*

$$\begin{aligned} \mathbb{E}_s\big[\xi_u^t\big] &= \xi_s^t e^{-a(u-s)}, \\ \mathbb{E}_s\big[A_u^{t,s}\big] &= \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}), \\ \mathbb{E}_s\big[X_u^{t,x,\Theta}\big] &= X_s^{t,x,\Theta} + \int_s^u \alpha_r \mathrm{d}r - \frac{\sigma^2}{2}(u-s) + \frac{\xi_s^t}{a}(1 - e^{-a(u-s)}), \end{aligned} \tag{5.3.3}$$

and

$$\begin{aligned} \mathrm{Var}_s(\xi_u^t) &= \frac{b^2}{2a}(1 - e^{-2a(u-s)}), \\ \mathrm{Cov}_s(\xi_u^t, A_u^{t,s}) &= \frac{b^2}{a^2}(1 - e^{-a(u-s)}) - \frac{b^2}{2a^2}(1 - e^{-2a(u-s)}), \\ \mathrm{Var}_s(A_u^{t,s}) &= \frac{b^2}{a^2}(u-s) - \frac{2b^2}{a^3}(1 - e^{-a(u-s)}) + \frac{b^2}{2a^3}(1 - e^{-2a(u-s)}), \\ \mathrm{Cov}_s(\xi_u^t, X_u^{t,x,\Theta}) &= \frac{b}{a}(\frac{b}{a} + \rho\sigma)(1 - e^{-a(u-s)}) - \frac{b^2}{2a^2}(1 - e^{-2a(u-s)}), \\ \mathrm{Cov}_s(A_u^{t,s}, X_u^{t,x,\Theta}) &= -\frac{1}{a}\mathrm{Cov}(\xi_u^t, X_u^{t,x,\Theta}) + \frac{b}{a}(\frac{b}{a} + \rho\sigma)(u-s) - \frac{b^2}{a^3}(1 - e^{-a(u-s)}), \\ \mathrm{Var}_s(X^{t,x,\Theta}) &= (\rho\sigma + \frac{b}{a})^2(u-s) + (1-\rho^2)\sigma^2(u-s) - 2\frac{b}{a^2}(\rho\sigma + \frac{b}{a})(1 - e^{-a(u-s)}) + \frac{b^2}{2a^3}(1 - e^{-2a(u-s)}). \end{aligned} \tag{5.3.4}$$

5.3.2.1 Approximation of the Liability side

With the above results, the approximation at any time t of the liability function ℓ, denoted $\ell^{\mathcal{N}}$, is straightforward. It is given by

$$\ell^{\mathcal{N}}(t,x,\Theta) = \frac{1}{M}\sum_{k=1}^{M} \beta_T^{t,\Theta,k} G(S^{t,x,\Theta,k}), \tag{5.3.5}$$

with $((\beta_\tau^{t,\Theta,k}, S_\tau^{t,x,\Theta,k})_{\tau\in\Gamma_G})_{1\le k\le M}$ are i.i.d realisations of $(\beta_\tau^{t,\Theta}, S_\tau^{t,x,\Theta})_{\tau\in\Gamma_G}$, recall (5.2.1).

5.3.2.2 Approximation of the Asset side

The approximation of the asset side is slightly more involved as it requires the computation of sensitivities with respect to the model parameters: $\frac{\partial\ell}{\partial x}$ and $\frac{\partial\ell}{\partial\theta_i}$, $i=1,2,3$, see (5.2.9), (5.2.10) and (5.2.11). We choose to compute the sensitivities using a "weight" approach namely we express them as expectation of the discounted pay-off times a well chosen random weight. Note that other techniques are available to compute these sensitivities e.g. *Automatic differentiation*. In our context, we have compared the two methods and observed that the weight approach is more efficient, see Section 5.5.4 in the Appendix.

We now describe how to obtain the sensitivities in a convenient form for Monte Carlo simulation. Recall that we consider a liability function ℓ of the form:

$$\ell(t,x,\Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s} G(S^{t,x,\Theta})\right],$$

where G depends upon $S^{t,x,\Theta}$ through its values on a finite set $\Gamma_G = \{\tau_0 = 0, \dots, \tau_\kappa = T\} \subset [0,T]$. In the following, we assume for simplicity that $\tau_1 > t$. Otherwise, there are deterministic terms that are to be added, but the method remains the same.

The Delta: We want to compute:

$$\frac{\partial\ell}{\partial x}(t,x,\Theta) = \frac{\partial}{\partial x}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s} G(S^{t,x,\Theta})\right]$$

and we have the following result.

Proposition 5.3.2. *For all $(t,x,\Theta)\in[0,1]\times\mathbb{R}\times\mathbb{R}^3$, the following holds:*

$$\frac{\partial\ell}{\partial x}(t,x,\Theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s} G(S^{t,x,\Theta}) H^{x,\Theta}\left(e^{-\int_t^{\tau_1}\xi_s^t\mathrm{d}s}, S_{\tau_1}^{t,x,\Theta}\right)\right], \tag{5.3.6}$$

with

$$H^{x,\Theta}(a,y) = \frac{\Sigma_{1,2}^{-1}a + \Sigma_{2,2}^{-1}\times\left(\log(y/x) - \int_t^{\tau_1}\alpha_r^{t,\Theta}\mathrm{d}r + \frac{\sigma^2}{2}(\tau_1 - t)\right)}{x}, \tag{5.3.7}$$

where Σ is the covariance matrix of $(A_{\tau_1}^t, X_{\tau_1}^{t,x,\Theta})$, see (5.3.9).

Proof. We write the expectation as an integral, remembering that we know the law of the couple $(\xi_u^t, A_u^t, X_u^{t,x,\Theta})$ conditionally upon $\mathcal{F}_s$:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s}G(S^{t,x,\Theta})\right]\\
&\quad= e^{-\int_t^T \alpha_s^{t,\Theta}\mathrm{d}s}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{\tau_1}\xi_s^t\mathrm{d}s}\prod_{\ell=1}^{\kappa-1}e^{-\int_{\tau_\ell}^{\tau_{\ell+1}}\xi_s^t\mathrm{d}s}G(S^{t,x,\Theta})\right]\\
&\quad= e^{-\int_t^T \alpha_s^{t,\Theta}\mathrm{d}s}\int_{\mathbb{R}^{2\kappa}}e^{-a_1}\prod_{\ell=1}^{\kappa-1}e^{-a_{\ell+1}}G(e^{x_1},\dots,e^{x_\kappa})\mathrm{d}\mathbb{Q}_{(A_u^t,X_u^{t,x,\Theta})_{u\in\Gamma_G}}(a_1,\cdots,a_\kappa,x_1,\cdots,x_\kappa)\\
&\quad= e^{-\int_t^T \alpha_s^{t,\Theta}\mathrm{d}s}\int_{\mathbb{R}^{2\kappa}}e^{-a_1}\prod_{\ell=1}^{\kappa-1}e^{-a_{\ell+1}}G(e^{x_1},\dots,e^{x_\kappa})p^{\Theta}(t,0,\log(x),1,a_1,x_1)\times\dots\\
&\qquad\qquad\times p^{\Theta}(t_{\kappa-1},0,x_{\kappa-1},t_\kappa,a_\kappa,x_\kappa)\mathrm{d}a_1\cdots\mathrm{d}a_\kappa\mathrm{d}x_1\cdots\mathrm{d}x_\kappa,
\end{aligned}$$

where $p^{\Theta}(s,a,x,u,.,.)$ is the density of the couple $(A_u^{t,a}, X_u^{t,x,\Theta})$, conditionally upon $\mathcal{F}_s$. We have previously seen that it is a Gaussian vector with explicit mean vector and covariance matrix. Thus, using Fubini's theorem, we get, since there is no dependence on x except in the first density:

$$\begin{aligned}
\frac{\partial\ell(t,x,\Theta)}{\partial x} &= e^{-\int_t^T \alpha_s^{t,\Theta}\mathrm{d}s}\int_{\mathbb{R}^{2\kappa}}e^{-a_1}\prod_{\ell=1}^{\kappa-1}e^{-a_{\ell+1}}G(e^{x_1},\dots,e^{x_\kappa})\frac{\partial p^{\Theta}(t,0,\log(x),1,a_1,x_1)}{\partial x}\times\dots\\
&\qquad\times p^{\Theta}(t_{\kappa-1},0,x_{\kappa-1},t_\kappa,a_\kappa,x_\kappa)\mathrm{d}a_1\cdots\mathrm{d}a_\kappa\mathrm{d}x_1\cdots\mathrm{d}x_\kappa.
\end{aligned} \tag{5.3.8}$$

Consequently, the sensibility of the discounted price with respect to the initial stock price is computed only by calculating the derivative of the density with respect to x.

We have

$$p^{\Theta}(t,0,\log(x),s,a,y) = \frac{1}{\det(2\pi\Sigma)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(((a,y)-\mu)\,\Sigma^{-1}\,((a,y)-\mu)^{\top}\right)\right),$$

where, thanks to (5.3.3)-(5.3.4),

$$\begin{aligned}
\mu &= \left(\mathbb{E}_t\big[A_{\tau_1}^t\big],\mathbb{E}_t\big[X_{\tau_1}^{t,x,\Theta}\big]\right)\\
\Sigma &= \begin{pmatrix}\mathrm{Var}(A_{\tau_1}^t)_t & \mathrm{Cov}(A_{\tau_1}^t,X_{\tau_1}^{t,x,\Theta})_t\\ \mathrm{Cov}(A_{\tau_1}^t,X_{\tau_1}^{t,x,\Theta})_t & \mathrm{Var}(X_{\tau_1}^{t,x,\Theta})_t\end{pmatrix}.
\end{aligned} \tag{5.3.9}$$

Still by (5.3.3)-(5.3.4), we see that only $\mathbb{E}_t\Big[X_{\tau_1}^{t,x,\Theta}\Big]$ depends upon x. Thus we get:

$$\frac{\partial f(t,0,\log(x),s,a,y)}{\partial x} = \frac{\Sigma_{1,2}^{-1}a+\Sigma_{2,2}^{-1}\times\left(\log(y/x)-\int_t^{\tau_1}\alpha_r^{t,\Theta}\mathrm{d}r+\frac{\sigma^2}{2}(\tau_1-t)\right)}{x}f(t,0,\log(x),s,a,y).$$

Reinjecting this equality into (5.3.8) and rewriting the result in term of expectations, we finally get the result. □

The function $\Delta(t_i,\cdot)$ is computed using the Monte Carlo estimator given in (5.3.6) as in (5.3.5) where we simulate in addition the weight H.

Sensitivities with respect to the interest rates curve. We consider now the derivatives with respect to the interest rates curve. For $i=1,2,3$, we want to compute:

$$\frac{\partial \ell}{\partial \theta_i}(t,x,\Theta)=\frac{\partial}{\partial \theta_i}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s}G(S^{t,x,\Theta})\right],\ i=1,2,3.$$

Proposition 5.3.3. *For all $(t,x,\Theta)\in[0,1]\times(0,\infty)\times\mathbb{R}^3$ and all $i=1,2,3$, we have the following identity (where we have set $\tau_0=t$):*

$$\frac{\partial \ell}{\partial \theta_i}(t,x,\Theta)=\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s^{t,\Theta}\mathrm{d}s}G(S^{t,x,\Theta})H_i^{x,\Theta}\left((\xi_{\tau_l}^t,e^{-\int_{\tau_{l-1}}^{\tau_l}\xi_u^t\mathrm{d}u},S_{\tau_l}^{t,x,\Theta})_{l=1,\dots,\kappa}\right)\right] \tag{5.3.10}$$

with

$$\begin{aligned}H_i^{t,x,\Theta}\left((r_\ell,a_\ell,s_\ell)_{\ell=1,\dots,\kappa}\right)=-\int_t^{\tau_\kappa}h_s^{t,i}\mathrm{d}s+\sum_{\ell=1}^{\kappa}\left(\int_{\tau_{\ell-1}}^{\tau_\ell}h_s^{t,i}\mathrm{d}s\right)\Big((\Sigma^{\tau_{\ell-1},\tau_\ell})^{-1}_{1,3}(r_\ell-\mu_1^{\tau_{\ell-1},\tau_\ell})\\ +(\Sigma^{\tau_{\ell-1},\tau_\ell})^{-1}_{2,3}(a_\ell-\mu_2^{\tau_{\ell-1},\tau_\ell})\\ +(\Sigma^{\tau_{\ell-1},\tau_\ell})^{-1}_{3,3}(\log(s_\ell)-\mu_3^{\tau_{\ell-1},\tau_\ell})\Big),\end{aligned} \tag{5.3.11}$$

where $\mu^{s,u}$ and $\Sigma^{s,u}$ are the mean and the covariance matrix of the Gaussian vector $(\xi_u^t,A_u^t,X^{t,x,\Theta})$ conditionally upon $\mathcal{F}_s$.

Proof. Performing a similar analysis as the one above,

$$\frac{\partial \ell}{\partial \theta_i}(t,x,\Theta)=\frac{\partial e^{-\int_t^T\alpha_s^{t,\Theta}\mathrm{d}s}}{\partial\theta_i}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T\xi_s^t\mathrm{d}s}G(S^{t,x,\Theta})\right]+e^{-\int_t^T\alpha_s^{t,\Theta}\mathrm{d}s}\int_{\mathbb{R}^{2\kappa}}e^{-a_1}\prod_{\ell=1}^{\kappa-1}e^{-a_{\ell+1}}G(e^{x_1},\dots,e^{x_\kappa})$$
$$\frac{\partial\left(p^\Theta(t,0,\log(x),1,a_1,x_1)\dots p^\Theta(t_{\kappa-1},a_{\kappa-1},x_{\kappa-1},t_\kappa,a_\kappa,x_\kappa)\right)}{\partial\theta_i}\mathrm{d}a_1\cdots\mathrm{d}a_\kappa\mathrm{d}x_1\cdots\mathrm{d}x_\kappa.$$

Here, the computations are more involved since every density depends upon the $a_i, i=1,2,3$, but the idea is the same as before.

The only difference is that, when we use (5.3.3)-(5.3.4) to differentiate, we see that the short term itself appears in the formulae. This is not a problem as we can rewrite the previous integral as an integral over $\mathbb{R}^{3\kappa}$, with the short rate process taken at times $\tau_l, l=1,\dots,\kappa$, as new variables to integrate against. □

The quantity $\frac{\partial \ell}{\partial\theta_i}$ $(i=1,2,3)$ is computed using the Monte Carlo estimator of the formula (5.3.10). Then solving the system (5.2.11) allows to obtain the coefficients ρ_1,ρ_2,ρ_3 .

This method to calculate derivatives allows us to compute the function $\ell(t,x,\Theta)$ and its four derivatives with only one Monte Carlo simulation. Furthermore, given a risk-neutral scenario, each quantity involved in formulae (5.3.7) and (5.3.11) can be exactly computed by integrating the elementary functions $h^{t,i}$ and by inverting real symmetric matrices of size 3×3. Thus, the weight functions $H^{t,x,\Theta}, H_i^{t,x,\Theta}$ are easily and accurately computed.

5.3.3 Sparse Grid Approach

The *nested simulation* approach requires the approximation of the function ℓ and its derivative $\frac{\partial \ell}{\partial x}$, $\frac{\partial \ell}{\partial \theta_i}$, $i = 1, 2, 3$ for each path $(\bar{S}^j_t, \bar{\Theta}^j_t)_{t\in\Gamma}$ of the market parameters. These values are computed on the fly which is quite time consuming.

We suggest here an alternative method which will pre-compute the quantities of interest (ℓ and its derivatives) and store them. The requested value for a given market parameter will then be obtained by an interpolation procedure.

A first simple approach is to consider an equidistant grid of the domain $A := \prod_{l=1}^{d}[m_p, M_p]$ which is a truncation of the support of $\mathcal{X}$ ($\mathbb{R}^d$, $d = 4$ in our setting). Then one can use a multi-linear interpolation to reconstruct the function in the whole space. If one sets 2^p points in one dimension, the total number of points will be 2^{dp} for one function at one given time and overall $(n + 1)2^{dp}$ to store (here $d = 4$). This will become rapidly too big, especially if one allows the number of market parameters to grow. This is a typical example of the "curse of dimensionality" encountered in numerical analysis when dealing with problem of high dimension.

Instead of considering a regular grid, we shall rely on the use of *sparse grid*, which allows to lower the number of points required to store the numerical approximation of the function. We now present rapidly the main concepts linked to sparse grids, see [BG04] for a comprehensive survey. In our numerical examples, see Section 5.4, the sparse grid will be computed using the StOpt C++ library [Gev+18].

For each multi-index $\mathbf{k} \le \mathbf{l}$, we define a grid mesh $h_{\mathbf{k}} = 2^{-\mathbf{k}}$ and grid points

$$\breve{y}_{\mathbf{k},\mathbf{i}} = (m_1 + i_1(M_1 - m_1)h_{k_1}, \ldots, m_d + i_d(M_d - m_d)h_{k_d}),\ \mathbf{0} \le \mathbf{i} \le 2^{\mathbf{k}}.$$

Using the *hat* function,

$$y \in \mathbb{R} \mapsto \phi(y) := \begin{cases} 1 - |y| & \text{if } y \in [-1, 1] \\ 0 & \text{otherwise} \end{cases}$$

we can associate to the previous grid a set of nodal basis function:

$$y \in \mathbb{R}^d \mapsto \phi_{\mathbf{k},\mathbf{i}}(y; A) = \prod_{l=1}^{d} \phi(\frac{y_l - \breve{y}^l_{\mathbf{k},\mathbf{i}}}{2^{-i_l}})\ .$$

When using a "full" linear interpolation, the function is approximated using the whole set of nodal basis function at the finest level $\mathbf{l}$. Instead, we consider the sparse grid nodal space of order p defined by

$$\mathcal{V}_\kappa := \text{span}\{\phi_{\mathbf{l},\mathbf{j}}; (\mathbf{l}, \mathbf{j}) \in \mathcal{I}_p(A)\},$$

where

$$\mathcal{I}_\kappa := \Big\{(\mathbf{l}, \mathbf{j}) : \quad 0 \le \sum_{i=1}^{d} l_i \le p; \quad \mathbf{0} \le \mathbf{j} \le \mathbf{2}^{\mathbf{l}};$$
$$(l_i > 0 \text{ and } j_i \text{ is odd}) \text{ or } (l_i = 0),\ \text{for } i = 1, \ldots, d\Big\}.$$

For a function $\psi : A \to \mathbb{R}$ with support in A, we define its associated $\mathcal{V}_\kappa$-interpolator by

$$\pi^A_{\mathcal{V}_\kappa}[\psi](y) := \sum_{(\mathbf{l},\mathbf{j})\in\mathcal{I}_\kappa(A)} \theta_{\mathbf{l},\mathbf{j}}(\psi;A)\phi_{\mathbf{l},\mathbf{j}}(y;A) \tag{5.3.12}$$

where the operator $\theta_{\mathbf{l},\mathbf{j}}$ can be defined recursively in terms of r, the dimension of $\mathbf{l}$, by:

$$\theta_{\mathbf{l},\mathbf{j}}(\psi;A) = \begin{cases} \psi(\breve{y}_{\mathbf{l},\mathbf{j}}); & r=0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot,\breve{y}^r_{l_r,j_r});A-); & l_r = 0 \\ \theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot,\breve{y}^r_{l_r,j_r});A-) - \frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot,\breve{y}^r_{l_r,j_r-1});A-) & \\ \qquad -\frac{1}{2}\theta_{\mathbf{l}-,\mathbf{j}-}(\psi(\cdot,\breve{y}^r_{l_r,j_r+1});A-); & l_r > 0 \end{cases}$$

where, for a hypercube $A = \prod_{l=1}^d [m_l, M_l]$, $A- := \prod_{l=1}^{d-1}[m_l, M_l]$ and for a multi-index $\mathbf{k}$ with dimension $r \geq 1$, $\mathbf{k}- = (k_1, \ldots, k_{r-1})$.

Let us now introduce the approximation that we use to compute the loss distribution, namely

$$\ell^{\mathcal{S}}(\cdot) := \pi_{\mathcal{V}_\kappa}[\ell^{\mathcal{N}}](\cdot), \left(\frac{\partial \ell}{\partial x}(t,\cdot)\right)^{\mathcal{S}} := \pi_{\mathcal{V}_\kappa}[\left(\frac{\partial \ell}{\partial x}(t,\cdot)\right)^{\mathcal{N}}](\cdot)$$
$$\text{and } \left(\frac{\partial \ell}{\partial \theta_i}(t,\cdot)\right)^{\mathcal{S}} := \pi_{\mathcal{V}_\kappa}[\left(\frac{\partial \ell}{\partial \theta_i}(t,\cdot)\right)^{\mathcal{N}}](\cdot)\ i \in \{1,2,3\}\ ,\ t \in \Gamma\ .$$

These functions are built by computing the coefficients appearing in (5.3.12), which are then stored in memory. For $\ell^{\mathcal{S}}$ say, this amounts to compute the function $\ell^{\mathcal{N}}$ on the sparse grid $\mathcal{V}_p$ and this is done by Monte Carlo simulation as in the previous approach, recall the definition of $\ell^{\mathcal{N}}$ in (5.3.5).

Complexity The main limitation of this method is the memory usage and the time to pre-compute the functions. This is proportional to the number of point in the grid. This number can be estimated to be of $O(2^{\kappa-d+1}\frac{(\kappa-d+1)^{d-1}}{(d-1)!})$, see Proposition 4.1 in [BG04]. Let us insist on the fact that this is done "offline" compared to the *nested simulation* approach. For the "online" computations, the main effort is put in the evaluation of the function which is slightly more evolved than a linear interpolation and is of $O(\kappa)$ where κ is the maximum level that is chosen.

Remark 5.3.1. *The computations at each point being independent, this Sparse Grid approach can be easily parallelised, hence improving further the gain of time observed in Subsection 5.4.1.*

5.3.4 Convergence study

The goal is to obtain a reasonable approximation of the risk associated to the loss distribution of the balance sheet in an efficient way. In this section, we explain why the methods introduced above are indeed good approximations of the risk indicators. We also study theoretically the numerical complexity of both methods in terms of memory and time consumption.

5.3.4.1 Error analysis

For the risk estimation, we will investigate a root Mean Square Error (rMSE) of the following form

$$\epsilon^{\upsilon} := \mathbb{E}\big[|\varrho(p_1\sharp\eta) - \varrho(p_1^{\upsilon}\sharp\eta^N)|^2\big]^{\frac{1}{2}}\ ,\ \text{for } \upsilon \in \{\mathcal{N}, \mathcal{S}\}.$$

The expectation operator $\mathbb{E}[\cdot]$ acts under $\mathbb{P}^{\otimes N} \otimes \mathbb{Q}^{\otimes M}$, namely it averages both on the simulation of the market parameters under the *real-world* measure used for calibration and the risk neutral evolution of the market model under the pricing measure.

The first observation is that under reasonable assumptions on the risk measure used in the risk indicator, the error performed in the numerical simulation can be separated in two main contribution: the error due to the sampling of the loss distribution coming from the sampling of the market parameters and the error made when approximating the different pricing and hedging functions.

Lemma 5.3.2. *Assume that ϱ has a* Monotonic *and* Cash Invariant *lift $\Re$, then*

$$\epsilon^{\upsilon} \le \mathbb{E}\big[|\varrho(p_1\sharp\eta) - \varrho(p_1\sharp\eta^N)|^2\big]^{\frac{1}{2}} + \mathbb{E}\left[\sup_{1\le j\le N}|p_1(\mathcal{X}^j) - p_1^{\upsilon}(\mathcal{X}^j)|^2\right]^{\frac{1}{2}}\ .$$

Proof. We denote by $\widehat{\mathcal{X}}^N(\omega)$ the random variable with distribution $\eta^N(\omega)$. Note that

$$p_1(\widehat{\mathcal{X}}^N(\omega)) \le p_1^{\upsilon}(\widehat{\mathcal{X}}^N(\omega)) + \sup_{1\le j\le N}|p_1(\mathcal{X}^j(\omega)) - p_1^{\upsilon}(\mathcal{X}^j(\omega))|$$

which leads to

$$\Re(p_1(\widehat{\mathcal{X}}^N(\omega))) \le \Re(p_1^{\upsilon}(\widehat{\mathcal{X}}^N(\omega))) + \sup_{1\le j\le N}|p_1(\mathcal{X}^j(\omega)) - p_1^{\upsilon}(\mathcal{X}^j(\omega))|\ .$$

By symmetry, we easily get

$$|\varrho(p_1\sharp\eta) - \varrho(p_1^{\upsilon}\sharp\eta^N(\omega))| \le |\varrho(p_1\sharp\eta) - \varrho(p_1\sharp\eta^N(\omega))| + \sup_{1\le j\le N}|p_1(\mathcal{X}^j(\omega)) - p_1^{\upsilon}(\mathcal{X}^j(\omega))|$$

and then the proof is concluded using Minkowski inequality. □

The error due to the approximation of the function p_1 is well understood when the function is smooth enough. Note that the asset side of the function is quite involved and we will not attempt to obtain the condition for smoothness of the overall function p_1. We will now simply review the error done on the liability part $\ell(1,\cdot)$ assuming that the mapping G is bounded and

$$(x,\Theta) \to \beta_T^{1,\Theta}G(S^{1,x,\Theta}) \in \mathcal{C}_b^2. \tag{5.3.13}$$

Even though, this cannot be almost surely true in the model presented above, we will assume that $\beta_T^{1,\Theta}$ is bounded in the discussion below. A more precise analysis should take care of these extreme events arising with small probability. Another possibility would be to force the interest rate to be non-negative, by truncation or by considering a CIR type of model for (5.2.2).

Lemma 5.3.3. *Assume that* (5.3.13) *holds true. Recall the definition of* $\ell^{\mathcal{N}}$ *in* (5.3.5)*, then*

$$\mathbb{E}\Big[\max_{1\le j\le N}|\ell_1(\mathcal{X}^j)-\ell_1^{\mathcal{N}}(\mathcal{X}^j)|^2\Big]^{\frac{1}{2}} \le C\sqrt{\frac{\log(N)}{M}}\;.$$

Proof. We denote by c the bound on the mapping $(x,\Theta)\to\beta_T^{1,\Theta}G(S^{1,x,\Theta})$ (recall the discussion after equation (5.3.13)) and thus the bound on ℓ_1. For the reader's convenience, we introduce

$$\Sigma_M^j := \sum_{k=1}^{M}\ell_1(\mathcal{X}^j)-\beta_T^{t,\mathcal{X}^j,k}G(S_T^{t,\mathcal{X}^j,k}),$$

and observe that $\mathbb{E}\Big[\Sigma_M^j\Big]=0$ and recall that the (Σ_M^j) are i.i.d. We have, using Hoeffding Inequality,

$$\mathbb{E}\Big[\mathbf{1}_{\{|\Sigma_M^j|^2>z\}}\Big]\le 2\exp\left(-\frac{z}{cM}\right)\;.$$

Using the independence property, we obtain

$$\mathbb{E}\Big[\mathbf{1}_{\{\max_j|\Sigma_M^j|^2\le z\}}\Big]\ge\left(1-2\exp\left(-\frac{z}{cM}\right)\right)^N\;.$$

Now set $\xi:=cM\log(N)$ and compute

$$\begin{aligned}\mathbb{E}\Big[\max_{1\le j\le N}|\Sigma_M^j|^2\Big] &= \int_0^\infty\mathbb{E}\Big[\mathbf{1}_{\{\max_j|\Sigma_M^j|^2>z\}}\Big]\,\mathrm{d}z\\ &\le \xi+\int_\xi^\infty\mathbb{E}\Big[\mathbf{1}_{\{\max_j|\Sigma_M^j|^2>z\}}\Big]\,\mathrm{d}z\\ &\le \xi+\int_\xi^\infty\left\{1-\left(1-2\exp\left(-\frac{z}{cM}\right)\right)^N\right\}\mathrm{d}z.\end{aligned}$$

Now, we observe that for $N\ge 2$, $2\exp\left(-\frac{z}{cM}\right)\le 1$ for $z\ge\xi$. Using the fact that $1-(1-u)^N\le Nu$, for $u\in[0,1]$, we get

$$\mathbb{E}\Big[\max_{1\le j\le N}|\Sigma_M^j|^2\Big]\le\xi+2N\int_\xi^\infty\exp\left(-\frac{z}{cM}\right)\mathrm{d}z$$

which leads to

$$\mathbb{E}\Big[\max_{1\le j\le N}|\Sigma_M^j|^2\Big]\le\xi+2cM,$$

and concludes the proof.

□

We conclude this section by giving the overall estimation error induced by the numerical procedure above. We will admit that the upper bound for the error given for $\ell(1,\cdot)$ in Lemma 5.3.3 holds true for the PnL function $p(1,\cdot)$ with a scaling by n coming from the number of rebalancing date.

Theorem 5.3.1. *Assume that ϱ_h is a spectral risk measure. Then, the following holds, for some $\alpha > 0$,*

1. *for the* Nested Simulation *approach*

$$\epsilon^{\mathcal{N}} \le C\left(\frac{1}{N^\alpha} + n\sqrt{\frac{\log(N)}{M}}\right) ; \tag{5.3.14}$$

2. *for the* Sparse Grid *approach with maximum level κ*

$$\epsilon^{\mathcal{S}} \le C\left(\frac{1}{N^\alpha} + n\left\{\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}(\kappa - d + 1)^{(d-1)}\right\}\right) . \tag{5.3.15}$$

Proof. 1. We first show that

$$\mathbb{E}\left[\max_{1\le j\le N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{S}}(\mathcal{X}^j)|^2\right]^{\frac{1}{2}} \le C\left(\sqrt{\frac{\log(N)}{M}} + 2^{-2\kappa}\kappa^{(d-1)}\right) . \tag{5.3.16}$$

Indeed, we have that

$$\begin{aligned} \ell_1^{\mathcal{S}} &= \pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] \\ &= \ell_1^{\mathcal{N}} + \pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}} . \end{aligned}$$

And we observe that

$$\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}} = \frac{1}{M}\sum_{j=1}^{M} \pi_{\mathcal{V}_\kappa}[e^{-\int_1^T r_s^{1,\cdot,k}\mathrm{d}s} G(S^{1,\cdot,k})] - e^{-\int_1^T r_s^{1,\cdot,k}\mathrm{d}s} G(S^{1,\cdot,k})$$

Let us denote by $(x,\Theta) \mapsto \phi^k(x,\Theta) = e^{-\int_1^T r_s^{1,\Theta,k}\mathrm{d}s} G(S^{1,x,\Theta,k})$ which is a random function as it depends on the random realisation of the (r,S) process. Under (5.3.13), ϕ^k is smooth enough to apply the results in Proposition 4.1 in [BG04] and we obtain

$$|\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}}|_\infty \le \frac{1}{M}\sum_{j=1}^{M} |\pi_{\mathcal{V}_\kappa}[\phi^k] - \phi^k|_\infty \le C 2^{-2\kappa}\kappa^{d-1} . \tag{5.3.17}$$

We then observe that

$$\mathbb{E}\left[\max_{1\le j\le N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{S}}(\mathcal{X}^j)|^2\right]^{\frac{1}{2}} \le C\left(\mathbb{E}\left[\max_{1\le j\le N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{N}}(\mathcal{X}^j)|^2\right]^{\frac{1}{2}} + |\pi_{\mathcal{V}_\kappa}[\ell_1^{\mathcal{N}}] - \ell_1^{\mathcal{N}}|_\infty\right) .$$

The proof of (5.3.16) is concluded by combining the above inequality with (5.3.17) and Lemma 5.3.3.
2. We now prove (5.3.14). Applying Lemma 5.3.2, we obtain

$$\epsilon^{\mathcal{N}} \le \mathbb{E}\left[|\varrho(p_1\sharp\eta) - \varrho(p_1\sharp\eta^N)|^2\right]^{\frac{1}{2}} + \mathbb{E}\left[\max_{1\le j\le N} |\ell_1(\mathcal{X}^j) - \ell_1^{\mathcal{N}}(\mathcal{X}^j)|^2\right]^{\frac{1}{2}} .$$

The second term in the right-hand side of the above inequality is controlled by using Lemma 5.3.3. We now study the first term in the right-hand side, which is the error introduced by the sampling of the loss distribution. Applying Corollary 11 in [Pic13] to the spectral risk measure, we first get

$$\mathbb{E}\big[|\varrho(p_1\sharp\eta) - \varrho(p_1\sharp\eta^N)|^2\big]^{\frac{1}{2}} \leq C\mathbb{E}\big[\mathcal{W}_2(\eta,\eta^N)^2\big]^{\frac{1}{2}} .$$

We then use Theorem 1 in [FG15] to bound the Wasserstein distance, which concludes the proof for this step.
3. To prove (5.3.15), we follow similar arguments as in step 2. but using (5.3.16) instead of invoking Lemma 5.3.3. □

Remark 5.3.2. *We can compare the bound obtained for the nested simulation with the ones in [GJ10]. Using a different approach, the authors prove a very nice bound on the overall error given by*

$$C\left(\frac{1}{\sqrt{N}} + \frac{1}{M}\right) ,$$

for the $V@R$ (which is not a spectral risk measure) and $AV@R$. Note that the term $\frac{1}{M}$ is obtained by cancellation of the first order term through an error expansion. It would be interesting to understand under which assumptions their bound can be retrieved in our setting of general spectral risk measure. This topic is left for further research.

We conclude this Section by a short account on the numerical complexity of the two methods.

The *Nested Simulation* approach is a pure "online" method which is very simple to implement but has a huge drawback in term of running time. Each time an estimation is requested the numerical complexity is overall of nNM, where recall n is the number of rebalancing date, M the number of sample for the risk neutral simulation and N the number of sample for the real-world simulation. The memory requirements comes only from the estimation of the loss distribution and are of order N.

As already mentioned, the *Sparse Grid* approach is both an "online" and "offline" method. In terms of memory requirement, it is thus greedier than the *nested simulation* approach. On top of the memory needed to store the sample distribution (of order N), memory is also needed to store the sparse grid approximation $p^{\mathcal{S}}$, the requirement are of order $O(n2^{\kappa-d+1}\frac{(\kappa-d+1)^{d-1}}{(d-1)!})$. In term of running time, the gain is important as the complexity of evaluating $p^{\mathcal{S}}$ is of $O(\kappa)$ only, where κ is the maximum level used.

5.4 Numerics

In the numerical applications below, we will compare the loss distribution obtained via our two numerical procedures by computing the Wasserstein distance between the two empirical distributions. Since the loss distribution is one-dimensional, we use the following formula [Pro56]: for two probability distribution on $\mathbb{R}$, η and $\tilde{\eta}$,

$$W_2(\eta,\tilde{\eta}) = (\int_0^1 |F_\eta^{-1}(u) - F_{\tilde{\eta}}^{-1}(u)|^2 \mathrm{d}u)^{\frac{1}{2}} .$$

In the setting of empirical distributions, the above distance is easily computed. Suppose $\eta = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$ and $\tilde{\eta} = \frac{1}{N}\sum_{i=1}^{N}\delta_{y_i}$.
We straightforwardly compute

$$W_2(\eta,\tilde{\eta})^2 = \sum_{i=1}^{N}\int_{\frac{i-1}{N}}^{\frac{i}{N}} |F_\eta^{-1}(u) - F_{\tilde{\eta}}^{-1}(u)|^2 \mathrm{d}u = \frac{1}{N}\sum_{i=1}^{N}|x_{(i)} - y_{(i)}|^2, \qquad (5.4.1)$$

where the subscript (i) refers to the i-th order statistic of the distribution, since $x_{(i)}$ (resp. $y_{(i)}$) is simply the $\frac{i}{N}$-th quantile of η (resp. $\tilde{\eta}$).

Besides the Wasserstein distance between the two empirical distributions, we will also compare the estimated V@R and AV@R, which are computed in a similar way. Indeed, for $\alpha \in (0,1]$, we have:

$$V@R_\alpha(\eta) = F_\eta^{-1}(\alpha) = x_{(i_\alpha)}, \qquad (5.4.2)$$

where $\frac{i_\alpha - 1}{N} < \alpha \leq \frac{i_\alpha}{N}$, $i_\alpha \in \{1,\dots,N\}$.
For a given $\alpha \in (0,1]$, we observe that

$$AV@R_\alpha(\eta) = \frac{1}{1-\alpha}\int_\alpha^1 V@R_p(\eta)\mathrm{d}p = \frac{1}{1-\alpha}\left(\int_\alpha^{\frac{i_\alpha}{N}} V@R_p(\eta)\mathrm{d}p + \sum_{i=i_\alpha}^{N-1}\int_{\frac{i}{N}}^{\frac{i+1}{N}} V@R_p(\eta)\mathrm{d}p\right)$$

which leads to

$$AV@R_{\bar{\alpha}}(\eta) = \frac{1}{1-\alpha}\left(\{\frac{i_\alpha}{N} - \alpha\}x_{(i_\alpha)} + \frac{1}{N}\sum_{i=i_\alpha}^{N-1} x_{(i+1)}\right). \qquad (5.4.3)$$

Using the formulae (5.4.1), (5.4.2) and (5.4.3), we will now present numerical results showing the efficiency and usefulness of the *sparse grid* approach. We first start with a comparison with the classically used *nested simulation* approach.

5.4.1 Sparse grid approach versus nested simulations approach

We computed the empirical distribution of the PnL at horizon 1 year using the *nested simulations* approach, recall Section 5.3.2, and the *sparse grid* approach, recall Section 5.3.3.
For both methods, we used a sample of size $N = 11000$ describing the *real-world* evolution of S and Θ, recall Section 5.2.4.
For the nested simulations approach, using the overall error bound given in [GJ10] we want to approximate the risk-neutral expectations with Monte Carlo simulations with samples of size $M \simeq \sqrt{N}$, that is $M = 100$. In practice however, we observe that convergence has not occurred yet and we observe non-neglectable changes in the risk measures taken into account, see Table 5.1. In this table, we compute the Wasserstein distance with respect to the distribution obtained for $M = 10000$. Operational constraints do not allow us to change the size of the sample used for the Monte Carlo simulations under $\mathbb{P}$, We thus consider a sample of size $M = 2000$ for the *risk-neutral* Monte Carlo simulations. Following Proposition 5.2.2, we calibrated a Gaussian model

M	100	500	1000	2000	10000
Wasserstein distance	0.73	0.20	0.09	0.05	0
V@R/AV@R $\alpha = 0.005$	7.03 / 8.75	4.26 / 5.08	3.95 / 4.33	3.72 / 4.01	3.56 / 3.85
V@R/AV@R $\alpha = 0.01$	5.82 / 7.52	3.88 / 4.57	3.68 / 4.07	3.48 / 3.80	3.37 / 3.65
V@R/AV@R $\alpha = 0.05$	4.01 / 5.22	3.08 / 3.62	2.93 / 3.36	2.84 / 3.23	2.77 / 3.13
V@R/AV@R $\alpha = 0.1$	3.31 / 4.42	2.72 / 3.24	2.61 / 3.06	2.54 / 2.96	2.49 / 2.88

Table 5.1 – Comparison of the metrics obtained with Nested simulations for varying risk-neutral simulation sample size.

such that $(X_1, (\theta_1)_1, (\theta_2)_1, (\theta_3)_1)$ has mean and covariance matrix given by:

$$\mu = (4.1 \times 10^{-5}, 0.01, 0.03, 0.01),$$
$$V = \begin{pmatrix} 0.004 & 3.2 \times 10^{-5} & 6.76 \times 10^{-6} & 0.000008 \\ 3.2 \times 10^{-5} & 3.1 \times 10^{-5} & 1.82 \times 10^{-5} & 1.5 \times 10^{-5} \\ 6.76 \times 10^{-6} & 1.82 \times 10^{-5} & 7.5 \times 10^{-5} & 8.1 \times 10^{-6} \\ 0.000008 & 1.5 \times 10^{-5} & 8.1 \times 10^{-6} & 2.7 \times 10^{-5} \end{pmatrix}$$

The risk-neutral simulations were computed by a Monte Carlo procedure, computed with the exact formulae in the Hull & White and Black & Scholes setting we used, recall Proposition 5.3.1. The volatility parameter used in the Black & Scholes model is set to $\sigma = 0.3$ while the parameters defining the Hull & White model are set to $a = 0.05$ and $b = 0.01$. Last, the covariation parameter between the two Brownian motions is set to $\rho = 0$.
In this setting, the nested simulations method was tested with the Put Lookback option described in 5.2.1, with maturity $T = 30$ years. Figure 5.2 shows the outcome PnL's distribution.

We next looked at the grid method. Figure 5.3 shows the outcome PnL's distributions for sparse grids of level $1, 2, 3$, which respectively have cardinal 81, 297, 945. For each level, we chose the number of risk-neutral simulations M so that the error induced by the Monte Carlo estimation is small compared with the sparse interpolation error, meaning that we can run the program multiple times without changing significantly the outcome. Furthermore, as in the *nested simulations* case, we choose M so that increasing M has no effect on the distribution. Empirically, we chose $M = 20000$ for the sparse grid of level 1, 2 or 3. Figure 5.4 compares the distribution obtained with nested simulations with the distribution obtained with the sparse grid of level 3. Table 5.2 shows computational times comparison, and Table 5.3 shows V@R and AV@R comparison for the empirical distributions obtained in each case.

We observe that the computational time on the Sparse grid of level 3 with $M = 20000$ is similar to the one for the Nested simulations with only $M = 2000$. Moreover, a significant gain in time is obtained by the use of the sparse grid of level 2 only, which already gives good results, see Table 5.2. As already observed in Remark 5.3.1, this gain in time can be further improved by parallelisation of the computations. In addition, we observe that, once the computations on the grid are done, then the PnL distribution is almost straightforwardly obtained. This is a key feature of the method since the computations on the grid are to be kept. Indeed, if one needs to change the distribution of (S, Θ) under $\mathbb{P}$, say because the view of the risk management on the

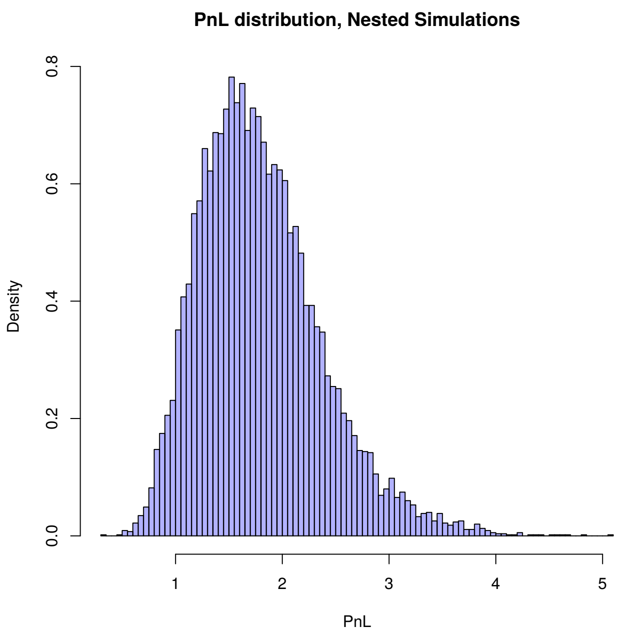

Figure 5.2 – PnL distribution, nested simulations

evolution of the market parameters has changed, then they can be re-used easily. In the next section, we give an application in this direction.

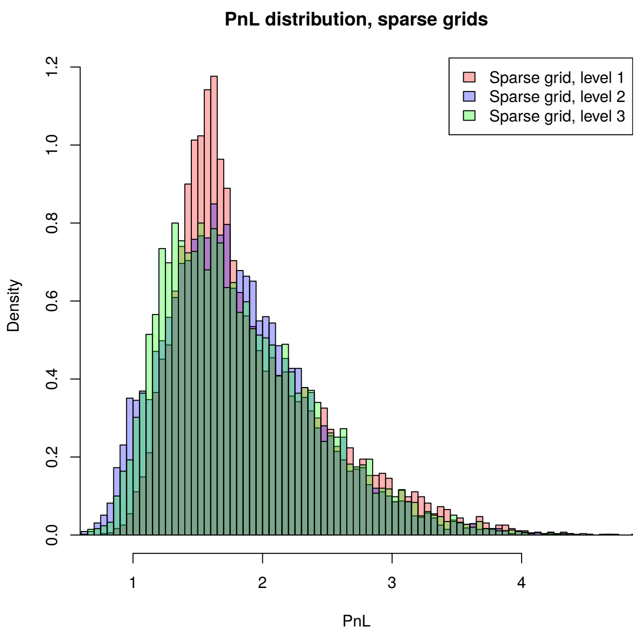

Figure 5.3 – PnL distribution, sparse grids

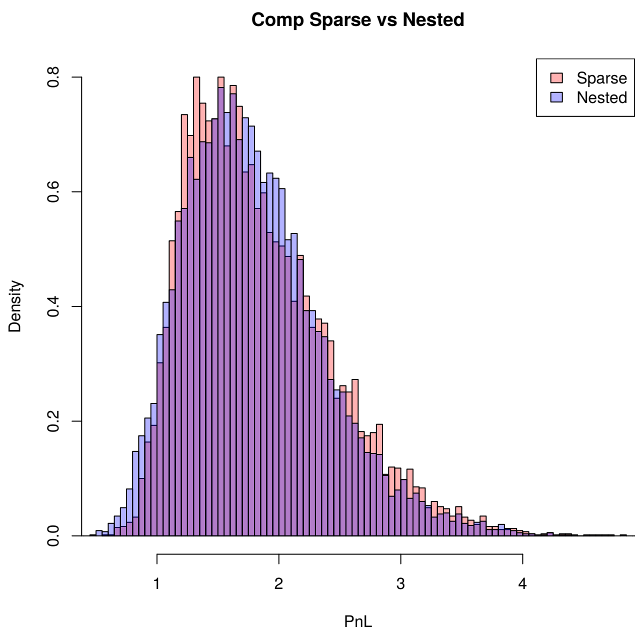

Figure 5.4 – PnL distribution, nested simulations versus sparse grid method

Level of the sparse grid	$l = 1$	$l = 2$	$l = 3$	Nested sim.
Computations on the grid	9 min 30 sec	35 min 10 sec	1h 58 min	
Computation of the PnL distribution	2 sec	4 sec	8 sec	2h 15min

Table 5.2 – Computational times

5.4.2 Model risk

As mentioned above, an interesting feature of the *sparse grid* approach developed in this paper is the ability to change the distribution of the processes $X = log(S)$ and Θ under the *real-world* probability $\mathbb{P}$. In this Section, we use the model described in Section 5.2.4 to simulate a first sample. Then, we consider some uncertainty over the estimated moments μ, V of (X_1, Θ_1) used to calibrate the gaussian model: we only assume that the true moments lie in centered intervals around the estimations. In practice, we consider intervals of the form $[m \times 0.95, m \times 1.05]$, where m is the estimated moment under consideration.

Level of the sparse grid	$l = 1$	$l = 2$	$l = 3$	Nested simulations
Wasserstein distance	0.13	0.06	0.05	0
V@R/AV@R $\alpha = 0.005$	3.87 / 4.09	3.64 / 3.94	3.72 / 3.96	3.72 / 4.01
V@R/AV@R $\alpha = 0.01$	3.68 / 3.93	3.43 / 3.73	3.54 / 3.79	3.48 / 3.80
V@R/AV@R $\alpha = 0.05$	3.08 / 3.43	2.81 / 3.17	2.95 / 3.30	2.84 / 3.23
V@R/AV@R $\alpha = 0.1$	2.73 / 3.16	2.54 / 2.92	2.64 / 3.04	2.54 / 2.96

Table 5.3 – Comparison of the empirical distributions

To better understand the risk associated with this uncertainty under $\mathbb{P}$, we simulate two "extreme" new samples of (X, Θ), where every moment taken into account to calibrate the model are multiplied by 0.95 (resp. 1.05), and, thanks to the grid computations done before with the initial model, we are in position to compute almost instantaneously the empirical PnL distributions associated with these two new samples.

Table 5.4 shows the Wasserstein distance between the initial distributions and the two obtained for the shifted parameters, and the V@R and AV@R obtained at different quantile levels. We observe that with these small change the distribution are quite close to each other. The main discrepancy are obtained for the diminished moments.

Model	Initial model	Diminished moments	Augmented moments
Wasserstein distance	0	0.030	0.0070
V@R/AV@R $\alpha = 0.005$	3.37 / 3.63	3.43 / 3.81	3.33 / 3.67
V@R/AV@R $\alpha = 0.01$	3.18 / 3.45	3.17 / 3.54	3.16 / 3.45
V@R/AV@R $\alpha = 0.05$	2.61 / 2.95	2.63 / 2.97	2.60 / 2.93
V@R/AV@R $\alpha = 0.1$	2.34 / 2.71	2.37 / 2.73	2.32 / 2.69

Table 5.4 – Comparison of the empirical distributions

5.5 Appendix

5.5.1 Proof of (5.2.3)

In this subsection, we shall give the proof of Proposition 2.1 for completeness. We remind that in the Hull-White model, the dynamics of the short rate is given by the following:

$$dr_s^{t,\Theta} = a(\mu_s^{t,\Theta} - r_s^{t,\Theta})\mathrm{d}s + b\mathrm{d}B_s \tag{5.5.1}$$

with $a, b \in \mathbb{R}$. We will prove that the mean-reverting θ_s can be calibrated by forward interest rate curve $f^\Theta(t,s)$ by:

$$\mu_s^{t,\Theta} = f^\Theta(t,s) + \frac{1}{a}\frac{\partial f^\Theta(t,s)}{\partial s} + \frac{b^2}{2a^2}(1 - e^{-2a(s-t)}) \tag{5.5.2}$$

The method is to express the price of the zero-coupon bond $P(t,s)$ in the following way:

$$\mathbb{E}[\exp(-\int_t^s r_u^{t,\Theta}\mathrm{d}u)] = P(t,s) = \exp(-\int_t^s f^\Theta(t,u)\mathrm{d}u) \tag{5.5.3}$$

Then by comparing both sides, we can determine $f^\Theta(t,s)$. First, it is easy to find out that the solution to (5.5.1) is

$$r_s^{t,\Theta} = r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}\mathrm{d}u + b\int_t^s e^{-a(s-u)}\mathrm{d}B_u$$

Then by straightforward calculation we have that

$$\int_t^s r_u^{t,\Theta}\mathrm{d}u = \frac{r_t^{t,\Theta}}{a}(1 - e^{-a(s-t)}) + \int_t^s \mu_u^{t,\Theta}(1 - e^{-a(s-u)})\mathrm{d}u + \frac{b}{a}\int_t^s (1 - e^{-a(s-u)})\mathrm{d}B_u$$

So $\int_t^s r_u^{t,\Theta}\mathrm{d}u$ follows a normal distribution with mean

$$\mathbb{E}[\int_t^s r_u^{t,\Theta}\mathrm{d}u] = \frac{r_t^{t,\Theta}}{a}(1 - e^{-a(s-t)}) + \int_t^s \mu_u^{t,\Theta}(1 - e^{-a(s-u)})\mathrm{d}u$$

and variance

$$\mathbb{V}[\int_t^s r_u^{t,\Theta}\mathrm{d}u] = \frac{b^2}{a^2}\int_t^s (1 - e^{-a(s-u)})^2\mathrm{d}u$$

Now comparing the both sides of (5.5.3), we have that

$$\int_t^s f^\Theta(t,u)\mathrm{d}u = \mathbb{E}[\int_t^s r_u^{t,\Theta}\mathrm{d}u] - \frac{1}{2}\mathbb{V}[\int_t^s r_u^{t,\Theta}\mathrm{d}u]$$

Thus

$$\begin{aligned} f^\Theta(t,s) &= \frac{\partial}{\partial s}\mathbb{E}[\int_t^s r_u^{t,\Theta}\mathrm{d}u] - \frac{1}{2}\frac{\partial}{\partial s}\mathbb{V}[\int_t^s r_u^{t,\Theta}\mathrm{d}u] \\ &= r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}\mathrm{d}u - \frac{b^2}{2a^2}2a\int_t^s (e^{-a(s-u)} - e^{-2a(s-u)})\mathrm{d}u \\ &= r_t^{t,\Theta}e^{-a(s-t)} + a\int_t^s \mu_u^{t,\Theta}e^{-a(s-u)}du - \frac{b^2}{2a^2}(1 - e^{-a(s-t)})^2 \end{aligned}$$

By straightforward differentiation, we have

$$\frac{\partial}{\partial s} f^{\Theta}(t,s) = -ar_t^{t,\Theta} e^{-a(s-t)} - a^2 \int_t^s \mu_u^{t,\Theta} e^{-a(s-u)} \mathrm{d}u + a\mu_s^{t,\Theta} - \frac{b^2}{a}(e^{-a(s-t)} - e^{-2a(s-t)}) \tag{5.5.5}$$

Now by (5.5.4) and (5.5.5), we can easily verify that (5.5.2) is valid.

5.5.2 Proof of Proposition 5.2.2

We provide a recursive proof of Proposition 5.2.2, which allows to compute effectively the coefficients defining the processes.

Suppose more generally that a vector $\mu \in \mathbb{R}^n$ and a covariance matrix $V \in \mathbb{R}^{n\times n}$ is given. We look for n processes $X^i (i = 0, \dots, n)$ defined by:

$$X_t^i = X_0^i + b_i t + \sum_{j=1}^n c_{ij} W_t^j,$$

where $W_t^j (j = 1, \dots, n)$ are n independent Brownian motions, and $b \in \mathbb{R}^n, C = (c_{ij}) \in \mathbb{R}^{n\times n}$.

Proposition 5.5.1. *There is at most one $(b, C) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$ such that:*

- $c_{ij} = 0$ *whenever* $i > j$,
- $\mathbb{E}[X_1^i] = \mu_i (i = 1, \dots, n)$,
- $Cov(X_1^i, X_1^j) = V_{ij} (i, j = 1, \dots, n)$.

Proof. We have $\mathbb{E}[X_1^i] = X_0^i + b_i$, so $b_i := \mu_i - X_0^i$ ensures $\mathbb{E}[X_1^i] = \mu_i$ for all i.

We next determine the matrix C thanks to a recursive algorithm:

Ascending step: Let $i, l \in \{1, \dots, n\}$, and assume $c_{ik}, k > l$ and $c_{lk}, k \geq l$ are determined. Then we can determine c_{il}.

Indeed, if $i > l$, we set $c_{il} = 0$. If $i < l$, we have:

$$\begin{aligned} V_{il} &= Cov(X_1^i, X_1^l) \\ &= \sum_{j=1}^n c_{ij} c_{lj} \\ &= \sum_{j=l}^n c_{ij} c_{lj} \\ &= c_{il} c_{ll} + \sum_{j>l} c_{ij} c_{lj}. \end{aligned}$$

Thus we set:

$$c_{il} = \frac{1}{c_{ll}} \left(V_{il} - \sum_{j>l} c_{ij} c_{lj} \right).$$

Back step: Let $l \in \{1, \dots, n\}$ and assume c_{lj} is determined, for $k > l$. Then we can determine c_{ll}.

Indeed:

$$V_{ll} = \mathbb{V}(X_1^l) = \sum_{j=1}^{n} c_{lj}^2 = \sum_{j=l}^{n} c_{lj}^2 = c_{ll}^2 + \sum_{j>l} c_{lj}^2.$$

Thus we set:

$$c_{ll} = \sqrt{V_{ll} - \sum_{j>l} c_{lj}^2}.$$

□

5.5.3 Proofs of Lemma 5.3.1 and Proposition 5.3.1

We prove here the Lemma 5.3.1 and the Proposition 5.3.1, which give a recursive procedure to simulate exactly under $\mathbb{Q}$.

Proof of Lemma 5.3.1. Let $(t, \Theta) \in [0, T] \times \mathbb{R}^3$ and consider the process $r^{t,\Theta} = (r_s^{t,\Theta})_{s \in [t,T]}$ defined by (5.2.2)-(5.2.3).

Let $s \in [t, T]$. An application of Itô's formula gives:

$$e^{as} r_s^{t,\Theta} = e^{at} r_t^{t,\Theta} + a \int_t^s e^{au} \mu_u^{t,\Theta} \mathrm{d}u + b \int_t^s e^{au} \mathrm{d}B_u,$$

and an easy computation using equality (5.2.3) shows that:

$$a \int_t^s e^{-a(s-u)} \mu_u^{t,\Theta} \mathrm{d}u = \alpha_s^{t,\Theta} - \alpha_t^{t,\Theta} e^{-a(s-t)},$$

where $\alpha^{t,\Theta}$ is the defined by (5.3.1).

In addition, if ξ^t is defined by (5.3.2), applying Itô's formula again gives:

$$\xi_s^t = b \int_t^s e^{-a(s-u)} \mathrm{d}B_u.$$

Thus:

$$r_s^{t,\Theta} = e^{-a(s-t)} r_t^{t,\Theta} + \alpha_s^{t,\Theta} - e^{-a(s-t)} \alpha_t^{t,\Theta} + \xi_s^t,$$

which ends the proof as $r_t^{t,\Theta} = \alpha_t^{t,\Theta}$, by (5.5.4). □

We now turn to the proof of Proposition 5.3.1.

Proof of Proposition 5.3.1. Let $t \le s \le u \le T$. Itô's formula implies that the triplet $(\xi_r^t, A_r^{t,s}, X_r^{t,x,\Theta})_{r \in [s,u]}$ is the solution of the following linear stochastic differential equation:

$$d \begin{pmatrix} \xi_r \\ A_r \\ X_r \end{pmatrix} = \left[\begin{pmatrix} -a & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_r \\ A_r \\ X_r \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \alpha_r^{t,\Theta} - \frac{\sigma^2}{2} \end{pmatrix} \right] \mathrm{d}t + \begin{pmatrix} b & 0 \\ 0 & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \mathrm{d}B_r \\ \mathrm{d}W_r \end{pmatrix}, r \in [s, u],$$

with the initial conditions $\xi_s = \xi_s^t, A_s = 0, X_s = X_s^{t,x,\Theta}$. This linear equation has a closed form solution, and we find:

$$\begin{aligned}
\xi_u^t &= e^{-a(u-s)}\xi_s^t + b\int_s^u e^{-a(u-r)}\mathrm{d}B_r, \\
A_u^{t,s} &= \frac{\xi_s^t}{a}(1-e^{-a(u-s)}) + \frac{b}{a}\int_s^u (1-e^{-a(u-r)})\mathrm{d}B_r, \\
X_u^{t,x,\Theta} &= X_s^{t,x,\Theta} + \int_s^u \alpha_r^{t,\Theta}\mathrm{d}r - \frac{\sigma^2}{2}(u-s) + \frac{\xi_s^t}{a}(1-e^{-a(u-s)}) + \int_s^u \sqrt{1-\rho^2}\sigma\mathrm{d}W_r \\
&\quad + \frac{b}{a}\int_s^u (1-e^{-a(u-r)})\mathrm{d}B_r + \int_s^u \rho\sigma\mathrm{d}B_r.
\end{aligned}$$

Conditionally upon $\mathcal{F}_s$, the vector $(\xi_u^t, A_u^{t,s}, X_u^{t,x,\Theta})$ is Gaussian and the expectations and covariations given in the Proposition are easily computed thanks to the above formulae. □

5.5.4 Comparison with Automatic Differentiation.

We use the stan math C++ library [Car+15] which allows to easily implement a (Reverse Mode) Automatic Differentiation procedure in order to deduce the derivatives directly from the Monte-Carlo computation of the function L. We compare the results obtained with the weights method developed here with the results obtained by Automatic Differentiation. We also provide a comparison about the computational times.

Precisely, we computed the derivatives of ℓ with respect to the 4 variables $(x, \theta_1, \theta_2, \theta_3)$ at 256 points $(x^i, \theta_1^i, \theta_2^i, \theta_3^i)_{i=1,\ldots,256}$. In the case of the automatic differentiation, we only take 1000 risk-neutral simulations to compute ℓ, while for the approach involving the computation of weights, we took 10000 simulations to compute ℓ and its four derivatives.

Table 5.5 sums up the time taken for the computations. Clearly, the gain in time resulting by using the weights algorithm is really significant. Additionally, Figure 5.5 shows the accuracy in the computation using the weights derivatives in comparison with the Automatic Differentiation.

Algorithm - Option	Put Lookback
Automatic Differentiation	179 sec
Weights	97 sec

Table 5.5 – Computational times

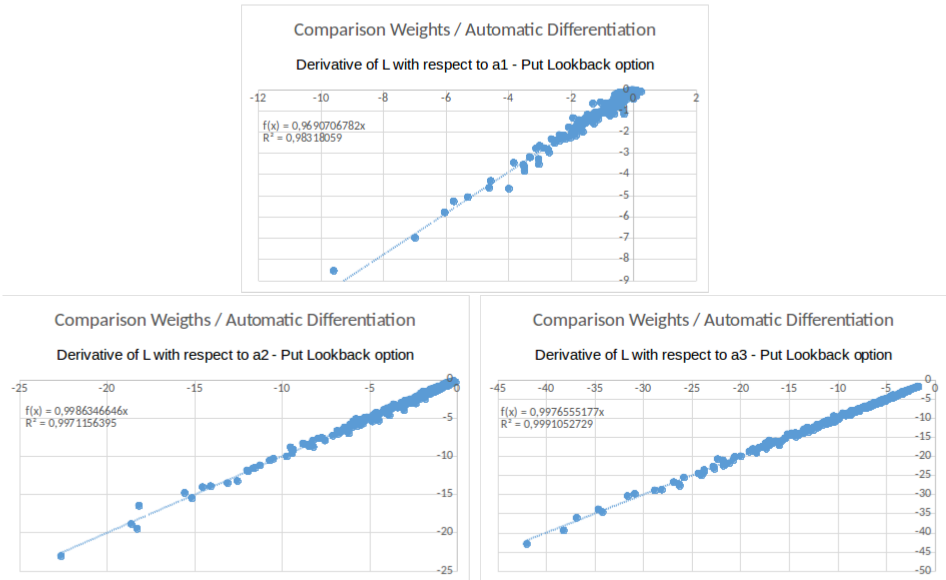

Figure 5.5 – Results from the grid method versus results from Automatic Differentiation

Part III

Obliquely Reflected BSDEs and Randomised switching

Chapter 6

Obliquely Reflected BSDEs and Randomised switching

The content of this chapter is from a work in progress, in collaboration with Jean-François Chassagneux and Adrien Richou.

6.1 Introduction

In this work, we introduce and study a new class of optimal switching problems in stochastic control. We show, in the spirit of [HT10], when the costs of switching are positive, that the value of these problems and an optimal policy are obtained by solving obliquely reflected Backward Stochastic Differential Equations (BSDEs for short). We then study, with arbitrary costs and in what we call the uncontrolled irreducible case, the question of existence of these obliquely reflected BSDEs in a Markovian framework, using an existence theorem from [CR18].

The interest in switching problems comes from their connections with financial and economics problems, like the pricing of *real options* or the valuation of a company which has to choose over time between multiple modes of production.
In [HJ07], Hamadène and Jeanblanc are interested in the case of two modes of production, and they notice that the solution is obtained by solving an obliquely reflected BSDE, which in this setting writes as a doubly reflected BSDE [CK96]. The general case of d modes was tackled in [DHP09].

In these works where the obliquely reflected BSDE is obtained by solving a switching problem, the driver does not depend on the components Y, Z of the solution. In [HT10], Hu and Tang consider equations in which the driver f can depend on Y, Z, with the assumption that $f^i(t, y, z) = f^i(t, y^i, z^i)$ for $1 \leq i \leq d$, $t \in [0, T]$, $y \in \mathbb{R}^d$, $z \in \mathbb{R}^{d\times k}$ (where k is the dimension of the Brownian motion). They obtain existence of the solutions by a classical penalization argument, and uniqueness is obtained through a verification argument, introducing a formal control problem with *switched BSDEs*. The hypothesis was relaxed in [CEK12] where the authors obtain existence and uniqueness by assuming only that $f^i(t, y, z) = f^i(t, y, z^i)$ for $1 \leq i \leq d$, $t \in [0, T]$, $y \in \mathbb{R}^d$,

$z \in \mathbb{R}^{d\times k}$. In the case of a general driver, existence is obtained in [CR18] in a Markovian framework. Existence in the non-Markovian case and uniqueness are open problems.

In all these works, the same underlying switching problem is considered, hence the domain in which Y lives, and the reflections, are of the same nature. More general oblique constraints and a driver depending on the global solution of the BSDE were considered in [HZ10], but they require the driver to be increasing with the components of Y.

Here, we consider a modification of the usual switching problem: when it is decided to change the mode, the agent cannot directly switch to the desired one. Instead, she is proposed probability distributions on the state space, among which she can choose. The new state is then drawn, independently of everything up to now, according to this distribution.

One first remark is that using Dirac measures, one can easily notice that the classical switching problem writes in this setup, so this new problem is a generalization of the classical switching problem.

In a framework similar to [HT10], we show the connection between the randomised switching problems and obliquely reflected BSDEs. More precisely, assuming existence of the solution of a well-chosen obliquely reflected BSDE, we show that its Y component is the value of the randomised switching problem, and an immediate corollary is the uniqueness of solutions for the BSDE.

We observe that the domain and the directions of reflections of the obliquely reflected BSDEs can be entirely different from one problem specification to an other. To study the existence of these BSDEs, our starting point is to apply results about obliquely reflected BSDEs from a more geometric point of view.

Reflected BSDEs were first considered by Gegout-Petit and Pardoux [GPP96], in the setting of normal reflections. In dimension 1, they have also been studied in [EK+97] in the simply reflected case, and in [CK96] in the doubly reflected case. In dimension greater than 1, the case of oblique reflections in an orthant was tackled in [Ram02]. The first attempt to study the general case of oblique reflections is [GRR15] in a restrictive setting for the driver and the directions of reflections. In [CR18], Chassagneux and Richou obtain an existence and uniqueness results in a non-Markovian setting, and an existence result in a Markovian framework.

In the setting of randomised switching problems studied here, a careful study of the geometry of the domain allows to invoke the results of [CR18] to obtain existence of the obliquely reflected BSDE in the so-called uncontrolled irreducible case.

The rest of the paper is organised as follows. In Section 6.2, we introduce the randomised switching problem by defining the admissible strategies and the associated reward, through the solution of a switched BSDE. We prove, if there exists a solution to some BSDEs with oblique reflections, that its first component coincides with the value of the switching problem, following ideas of Hu and Tang [HT10]. In particular, a classical verification argument allows to deduce uniqueness of the solution of the BSDE with oblique reflections. In Section 6.3, we show that there exists a solution of the obliquely reflected BSDE under some conditions on the parameters of the switching problem. In particular, we provide, in the uncontrolled irreducible case, necessary and sufficient conditions so that the domain in which the first component of the BSDE lives has non-empty interior, and that a solution actually exists by showing that the existence Theorem in [CR18] can be invoked. Last, we gather in Section 6.4.1 some technical results (martingale representation theorems and existence and uniqueness of BSDEs) about an enlarged filtration associated to an admissible strategy.

Notations If $n \geq 1$, we let $\mathcal{B}^n$ be the Borelian sigma-algebra on $\mathbb{R}^n$.
For any filtered probability space $(\Omega, \mathcal{G}, \mathbb{F}, \mathbb{P})$ and constants $T > 0$ and $p \geq 1$, we define the following spaces:

- $L^p_n(\mathcal{G})$ is the set of $\mathcal{G}$-measurable random variables X valued in $\mathbb{R}^n$ satisfying $\mathbb{E}[|X|^p] < +\infty$,
- $\mathcal{P}(\mathbb{F})$ is the predictable sigma-algebra on $\Omega \times [0, T]$,
- $\mathbb{H}^p_n(\mathbb{F})$ is the set of predictable processes ϕ valued in $\mathbb{R}^n$ such that
$$\mathbb{E}\left[\int_0^T |\phi_t|^p \mathrm{d}t\right] < +\infty,$$
- $\mathbb{S}^p_n(\mathbb{F})$ is the set of predictable processes ϕ valued in $\mathbb{R}^n$ such that
$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\phi_t|^p\right] < +\infty,$$
- $\mathbb{A}^p_n(\mathbb{F})$ is the set of continuous processes ϕ valued in $\mathbb{R}^n$ such that $\phi_T \in L^p_n(\mathcal{F}_T)$ and ϕ^i is non-decreasing for all $i = 1, \dots, n$.

If $n = 1$, we omit the subscript n in previous notations.

For $d \geq 1$, we denote by $(e_i)_{i=1}^d$ the canonical basis of $\mathbb{R}^d$, and $S_d(\mathbb{R})$ is the set of symmetric matrices of size $d \times d$ with real coefficients.

If $\mathcal{D}$ is a convex subset of $\mathbb{R}^d$ ($d \geq 1$) and $y \in \mathcal{D}$, we denote by $\mathcal{C}(y)$ the outward normal cone at y, defined by

$$\mathcal{C}(y) := \{v \in \mathbb{R}^d : v^\top (z - y) \leq 0 \text{ for all } z \in \mathcal{D}\}.$$

We also set $\mathfrak{n}(y) := \mathcal{C}(y) \cap \{v \in \mathbb{R}^d : |v| = 1\}$.

If X, Y are two matrices of the same size, we say that $X \geq Y$ if it is true coefficientwise.
If X is a matrix of size $n \times m$, $\mathcal{I} \subset \{1, \dots, n\}$ and $\mathcal{J} \subset \{1, \dots, m\}$, we set $X^{(\mathcal{I},\mathcal{J})}$ the matrix of size $(n - |\mathcal{I}|) \times (m \times |\mathcal{J}|)$ obtained from X by deleting rows with index $i \in \mathcal{I}$ and columns with index $j \in \mathcal{J}$. If $\mathcal{I} = \{i\}$ we set $X^{(i,\mathcal{J})} := X^{(\mathcal{I},\mathcal{J})}$, and similarly if $\mathcal{J} = \{j\}$.
If v is a vector of size n and $1 \leq i \leq n$, we set $v^{(i)}$ the vector of size $n - 1$ obtained from v by deleting coefficient i.

6.2 Randomised switching control problems

We consider an optimal switching problem where, in contrast with the usual switching problems [HT10; HZ10; CEK12], the agent does not choose directly the new state, but chooses a probability distribution under which the new state will be determined. Assuming the existence of a solution to some Obliquely Reflected BSDE, we characterise the value function and a family of optimal strategies for the switching problem.

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space. We fix a finite time horizon $T > 0$ and $k, d \geq 1$ two integers.

We assume that there exists a k-dimensional Brownian motion W and a sequence $(X_n)_{n\geq 1}$ of independent random variables, independent of W, uniformly distributed on $[0, 1]$. We also assume that $\mathcal{G}$ is generated by the Brownian motion W and the family $(X_n)_{n\geq 1}$.

We define $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t\geq 0}$ as the augmented Brownian filtration, which satisfies the usual conditions.

Let $\mathscr{C}$ be an ordered compact metric space and $F : \mathscr{C} \times \{1, \dots, d\} \times [0, 1] \to \{1, \dots, d\}$ a measurable map.

To each $u \in \mathscr{C}$ is associated a transition probability function on the state space $\{1, \dots, d\}$, given by $p_{i,j}^u := \mathbb{P}(F(u, i, X) = j)$ for X uniformly distributed on $[0, 1]$.

Let $c : \{1, \dots, d\} \times \mathscr{C} \to \mathbb{R}_+, (i, u) \mapsto c_{i,u}$ a map such that $c_{i,\cdot}$ is continuous for all $i = 1, \dots, d$. We assume that $\inf_{i\in\{1,\dots,d\},u\in\mathscr{C}} c_{i,u} = \underline{c} > 0$, and we set $\sup_{i\in\{1,\dots,d\},u\in\mathscr{C}} c_{i,u} = \bar{c}$.

Let $\xi = (\xi^1, \dots, \xi^d) \in L^2(\mathcal{F}_T^0)$ and $f : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^{k\times d} \to \mathbb{R}^d$ a map satisfying

- For all $(t, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^{k\times d}$, we have almost surely,

$$f(t, y, z) = (f^i(t, y_i, z_i))_{i=1}^d,$$

  where, for each $i = 1, \dots, d$, $f^i : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$ is $\mathcal{P}(\mathbb{F}^0) \otimes \mathcal{B} \otimes \mathcal{B}^k$-measurable and $f^i(\cdot, 0, 0) \in \mathbb{H}^2(\mathbb{F}^0)$.

- There exists $L \geq 0$ such that, for all $(t, y_1, y_2, z_1, z_2) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{k\times d} \times \mathbb{R}^{k\times d}$,

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq L(|y_1 - y_2| + |z_1 - z_2|),$$

6.2.1 The control problem

We define in this subsection the stochastic optimal control problem. We first introduce the strategies available to the agent, and then the associated reward she tries to maximise.

A strategy is given by $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1})$ where $\zeta_0 \in \{1, \dots, d\}$, $(\tau_n)_{n\geq 0}$ is a non-decreasing sequence of random times and $(\alpha_n)_{n\geq 1}$ is a sequence of $\mathscr{C}$-valued random variables, which satisfy:

- $\tau_0 \in [0, T]$ and $\zeta_0 \in \{1, \dots, d\}$ are deterministic.
- For all $n \geq 0$, τ_{n+1} is a $\mathbb{F}^n$-stopping time and α_{n+1} is $\mathcal{F}_{\tau_{n+1}}^n$-measurable (recall that $\mathbb{F}^0$ is the augmented Brownian filtration). We then set $\mathbb{F}^{n+1} = (\mathcal{F}_t^{n+1})_{t\geq 0}$ with $\mathcal{F}_t^{n+1} := \mathcal{F}_t^n \vee \sigma(X_{n+1} 1_{\{\tau_{n+1}\leq t\}})$.

If $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1})$ is a strategy, we set, for all $n \geq 0$, $\zeta_{n+1} = F(\alpha_{n+1}, \zeta_n, X_{n+1})$, and we define the state process as $a_t = \sum_{k=0}^{+\infty} \zeta_k 1_{[\tau_k, \tau_{k+1})}(t)$. The cumulative cost process is defined by $A_t^\phi = \sum_{k=0}^{+\infty} c_{\zeta_k, \alpha_{k+1}} 1_{\{\tau_{k+1} \leq t\}}$, $t \geq 0$, and the number of switches before t is denoted by $N_t^\phi := \sum_{k\geq 0} 1_{\{\tau_{k+1} \leq t\}}$.
Lastly, we define $\mathbb{F}^\infty = (\mathcal{F}_t^\infty)_{t\geq 0}$ with $\mathcal{F}_t^\infty := \bigvee_{n\geq 0} \mathcal{F}_t^n, t \geq 0$. Notice that the state process a is adapted to $\mathbb{F}^\infty$.

Remark 6.2.1. *For technical reasons involving simultaneous jumps, we do not consider the generated filtration associated to a, and one can easily see that $\mathbb{F}^\infty$ is larger.*

We say that $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1})$ is an *admissible strategy* if the cumulative cost process satisfies $A_T^\phi - A_{\tau_0}^\phi \in L^2(\mathbb{F}_T^\infty)$ and $\mathbb{E}\Big[\Big(A_{\tau_0}^\phi\Big)^2 | \mathcal{F}_{\tau_0}^0\Big] < +\infty$ a.s.. We define $\mathscr{A}$ as the set of admissible strategies, and for $t \in [0, T]$ and $i \in \{1, \dots, d\}$, we define $\mathscr{A}_t^i$ as the subset of admissible strategies satisfying $\zeta_0 = i$ and $\tau_0 = t$.

Remark 6.2.2. *The definition of $\mathscr{A}$ is slightly weaker than usual [HT10], which writes $A_T^\phi \in L^2(\mathcal{F}_T^\infty)$. One notices that our definition is enough to define the switched BSDE associated to an admissible control, see below. Moreover, we observe in the next subsection that the optimal strategies is admissible with respect to our definition, and not necessarily with the usual one, due to the simultaneous jumps at the initial time.*

We are now in position to introduce the reward associated to an admissible strategy. If $\phi = (\zeta_0, (\tau_n)_{n\geq 0}, (\alpha_n)_{n\geq 1}) \in \mathscr{A}$, the reward is defined as the value $\mathbb{E}\Big[U_{\tau_0}^\phi - A_{\tau_0}^\phi | \mathcal{F}_{\tau_0}^0\Big]$, where $(U^\phi, V^\phi, M^\phi) \in \mathbb{S}^2(\mathbb{F}^\infty) \times \mathbb{H}_k^2(\mathbb{F}^\infty) \times \mathbb{H}^2(\mathbb{F}^\infty)$ is the solution of the following switched BSDE [HT10] on the filtered probability space $(\Omega, \mathcal{G}, \mathbb{F}^\infty, \mathbb{P})$:

$$U_t = \xi^{a_T} + \int_t^T f^{a_s}(s, U_s, V_s)\mathrm{d}s - \int_t^T V_s \mathrm{d}W_s - \int_t^T \mathrm{d}M_s - \int_t^T \mathrm{d}A_s^\phi,\ t \in [\tau_0, T]. \quad (6.2.1)$$

Remark 6.2.3. *This switched BSDE rewrites as a classical BSDE in $\mathbb{F}^\infty$, and since $A^\phi - A_t^\phi \in \mathbb{S}^2(\mathbb{F}^\infty)$, the terminal condition and the driver are standard parameters, hence there exists a unique solution to (6.2.1) for all $\phi \in \mathscr{A}$. We refer to Subsection 6.4.1.2 for more details.*

For $t \in [0, T]$ and $i \in \{1, \dots, d\}$, the agent aims to solve the problem

$$\mathcal{V}_t^i = \operatorname*{ess\,sup}_{\phi \in \mathscr{A}_t^i} \mathbb{E}\Big[U_t^\phi - A_t^\phi | \mathcal{F}_t^0\Big].$$

6.2.2 An optimal strategy

The standing assumption in this subsection is:

Assumption 6.2.1.
- *There exists a solution $(Y, Z, K) \in \mathbb{S}_d^2(\mathbb{F}^0) \times \mathbb{H}_{k\times d}^2(\mathbb{F}^0) \times$*

$\mathbb{A}^2_d(\mathbb{F}^0)$ *to the following Obliquely Reflected BSDE:*

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s^i, Z_s^i)\mathrm{d}s - \int_t^T Z_s^i \mathrm{d}W_s + \int_t^T \mathrm{d}K_s^i,\ t \in [0,T],\ i \in \mathcal{I}, \quad (6.2.2)$$

$$Y_t \in \mathcal{D}, \quad t \in [0,T], \quad (6.2.3)$$

$$\int_0^T \left(Y_t^i - \sup_{u \in \mathscr{C}} \left\{ \sum_{j=1}^d p_{i,j}^u Y_t^j - c_{i,u} \right\} \right) \mathrm{d}K_t^i = 0, \quad i \in \mathcal{I}, \quad (6.2.4)$$

where $\mathcal{I} := \{1, \dots, d\}$ *and* $\mathcal{D}$ *is the following convex subset of* $\mathbb{R}^d$*:*

$$\mathcal{D} := \left\{ y \in \mathbb{R}^d : y_i \geq \sup_{u \in \mathscr{C}} \left\{ \sum_{j=1}^d p_{i,j}^u y_j - c_{i,u} \right\}, i \in \mathcal{I} \right\}. \quad (6.2.5)$$

- $\mathcal{D}$ *has non-empty interior,*
- *For all* $u \in \mathscr{C}$ *and* $i \in \{1, \dots, d\}$*, we have* $p_{i,i}^u \neq 1$*.*

Let $i \in \{1, \dots, d\}$ and $t \in [0,T]$. We set, for $\phi \in \mathscr{A}_t^i$ and $t \leq s \leq T$,

$$\mathcal{Y}_s^\phi := \sum_{k \geq 0} Y_s^{\zeta_k} 1_{[\tau_k, \tau_{k+1})}(s),$$

$$\mathcal{Z}_s^\phi := \sum_{k \geq 0} Z_s^{\zeta_k} 1_{[\tau_k, \tau_{k+1})}(s),$$

$$\mathcal{K}_s^\phi := \sum_{k \geq 0} K_s^{\zeta_k} 1_{[\tau_k, \tau_{k+1})}(s),$$

$$\mathcal{M}_s^\phi := \sum_{k \geq 0} \left(Y_{\tau_{k+1}}^{\zeta_{k+1}} - \mathbb{E}\Big[Y_{\tau_{k+1}}^{\zeta_{k+1}} | \mathcal{F}_{\tau_{k+1}}^k \Big] \right) 1_{\{t < \tau_{k+1} \leq s\}},$$

$$\mathcal{A}_s^\phi = \sum_{k \geq 0} \left(Y_{\tau_{k+1}}^{\zeta_k} - \mathbb{E}\Big[Y_{\tau_{k+1}}^{\zeta_{k+1}} | \mathcal{F}_{\tau_{k+1}}^k \Big] + c_{\zeta_k, \alpha_{k+1}} \right) 1_{\{t < \tau_{k+1} \leq s\}}. \quad (6.2.6)$$

Remark 6.2.4. *For all* $k \geq 0$*, since* $\alpha_{k+1} \in \mathcal{F}_{\tau_{k+1}}^k$*, we have*

$$\mathbb{E}\Big[Y_{\tau_{k+1}}^{\zeta_{k+1}} | \mathcal{F}_{\tau_{k+1}}^k \Big] = \sum_{j=1}^d \mathbb{P}\left[\zeta_{k+1} = j | \mathcal{F}_{\tau_{k+1}}^k \right] Y_{\tau_{k+1}}^j = \sum_{j=1}^d p_{\zeta_k, j}^{\alpha_k} Y_{\tau_{k+1}}^j. \quad (6.2.7)$$

We now introduce a strategy, and the aim of this subsection is to prove that it is admissible and optimal for the control problem. Let $\phi^\star = (\zeta_0^\star, (\tau_n^\star)_{n \geq 0}, (\alpha_n^\star)_{n \geq 1})$ defined by $\tau_0^\star = t$ and $\zeta_0^\star = i$ and inductively by:

$$\tau_{k+1}^\star = \inf \left\{ \tau_k^\star \leq s \leq T : Y_s^{\zeta_k^\star} = \sup_{u \in \mathscr{C}} \left\{ \sum_{j=1}^d p_{\zeta_k^\star, j}^u Y_s^j - c_{\zeta_k^\star, u} \right\} \right\} \wedge (T+1), \quad (6.2.8)$$

$$\alpha_{k+1}^\star = \inf \left\{ \alpha \in \arg\max_{u \in \mathscr{C}} \left\{ \sum_{j=1}^d p_{\zeta_k^\star, j}^u Y_{\tau_k^\star}^j - c_{\zeta_k^\star, u} \right\} \right\}, \quad (6.2.9)$$

recall that $\mathscr{C}$ is ordered.
In the following lemma, we show, since $\mathcal{D}$ has non-empty interior, that the number of switches (hence the cost) required to leave any point on the boundary of $\mathcal{D}$ is square integrable, following the strategy $\phi^\star$. This result will be used to prove that the cost associated to $\phi^\star$ satisfies $\mathbb{E}\left[\left(A_t^{\phi^\star}\right)^2 |\mathcal{F}_t\right]$ is almost surely finite.

Lemma 6.2.1. *Suppose that Assumption 6.2.1 is satisfied. For $y \in \mathcal{D}$, we define*

$$S(y) = \left\{1 \leq i \leq d : y_i = \sup_{u \in \mathscr{C}} \left\{\sum_{j=1}^d p_{i,j}^u y_j - c_{i,u}\right\}\right\}.$$

Let $\mathcal{U} = (u_i^k)_{k \geq 0, i \in S(y)}$ be a family of elements of $\mathscr{C}$ satisfying $y_i = \sum_{j=1}^d p_{i,j}^{u_i^k} y_j - c_{i,u_i^k}$. Consider the time-inhomogeneous Markov Chain X on $S(y) \cup \{0\}$ defined by, for $k \geq 0$ and $i, j \in S(y)^2$,

$$\begin{aligned}
\mathbb{P}(X_{k+1} = j | X_k = i) &= p_{i,j}^{u_i^k}, \\
\mathbb{P}(X_{k+1} = 0 | X_k = i) &= 1 - \sum_{j \in S(y)} p_{i,j}^{u_i^k}, \\
\mathbb{P}(X_{k+1} = 0 | X_k = 0) &= 1, \\
\mathbb{P}(X_{k+1} = i | X_k = 0) &= 0
\end{aligned}$$

Then 0 is accessible from every $i \in S(y)$, meaning that X is an absorbing Markov Chain.
Moreover, let $N(y, \mathcal{U}) = \inf\{n \geq 0 : X_k = 0\}$. Then $N(y, \mathcal{U}) \in L^2(\mathbb{P}^i)$ for all $i \in S(y)$, where $\mathbb{P}^i$ is the probability satisfying $\mathbb{P}^i(X_0 = i) = 1$.

Proof. Assume that there exists $i \in S(y)$ from which 0 is not accessible. Then every communicating class accessible from i is included in $S(y)$. In particular, there exists a recurrent class $S' \subset S(y)$. For all $k \geq 0$ and $i \in S'$, we have $p_{i,j}^{u_i^k} = 0$ if $j \notin S'$ since S' is recurrent. Moreover, since $S' \subset S(y)$, we obtain, for all $k \geq 0$ and $i \in S'$, by definition of $S(y)$,

$$y_i = \sum_{j \in S'} p_{i,j}^{u_i^k} y_j - c_{i,u_i^k}.$$

Since S' is a recurrent class, the matrix $P^k = (p_{i,j}^{u_i^k})_{i,j \in S'}$ is stochastic and irreducible. By definition of $\mathcal{D}$, we have $\mathcal{D} \subset \mathbb{R}^{d-|S'|} \times \left\{y \in \mathbb{R}^{|S'|} : y_i \geq \sum_{j \in S'} p_{i,j}^{u_i^k} y_j - c_{i,u_i^k}, i \in S'\right\} = \mathbb{R}^{d-|S'|} \times \mathcal{D}'$.
Let $i_0 \in S'$. The computations of Proposition 6.3.2 and Theorem 6.3.2 below show that $\mathcal{D}' \cap \{y_{i_0} = 0\}$ is a simplex of $\mathbb{R}^{|S'|-1}$ and its $|S'|$ extreme points are the $x^i, i \in S'$, where x^i is the unique solution to

$$\begin{aligned}
x_l^i &= \sum_{j \in S'} p_{l,j}^{u_l^k} x_j^i - c_{l,u_l^k}, l \neq i, \\
x_{i_0}^i &= 0.
\end{aligned}$$

Since the point $y \in \mathcal{D}$ satisfies $y_i = \sum_{j\in S'} p^{u_i^k}_{i,j} y_j - c_{i,u_i^k}$ for all $i \in S'$, we obtain $x^i = y - y_{i_0}\sum_{j\in S'} e_j$, and $\mathcal{D}' \cap \{y_{i_0} = 0\} = \{y - y_{i_0}\sum_{i=1}^d e_i\}$.
Hence we have proved that $\mathcal{D}'$ is a line in $\mathbb{R}^{|S'|}$, and $|S'| \geq 2$ as $p^u_{i,i} \neq 1$ for all $u \in \mathscr{C}$ and $i \in \{1,\dots,d\}$. Thus $\mathcal{D} \subset \mathbb{R}^{d-|S'|} \times \mathcal{D}'$ gives a contradiction with the fact that $\mathcal{D}$ has non-empty interior.
For $k \geq 0$, let $P^k = (p^{u_i^k}_{i,j})_{i,j\in S(y)}$ and for $n \geq 0$, let $P^{(n)} = P^0 \cdots P^n$. It is known ([KS81], Theorem 3.1.1 and Corollary 3.1.2) that $\sum_{n\geq 0} P^{(n)}$ converges, let $N := \sum_{n\geq 0} P^{(n)}$. Then it is easy to observe that $T_k \in L^2(\mathbb{P}^i)$ for all $i \in S(y)$ ([KS81], Theorem 3.3.5). □

Lemma 6.2.2. *1. For any admissible strategy $\phi \in \mathscr{A}^i_t$, $\mathcal{M}^\phi$ is a square integrable martingale with $\mathcal{M}^\phi_t = 0$. Moreover, $\mathcal{A}^\phi$ is increasing and satisfies $\mathcal{A}^\phi_T \in L^2(\mathcal{F}^\infty_T)$. In addition,*

$$\mathbb{E}\left[\left(\sum_{k\geq 0}\left(Y^{\zeta_{k+1}}_{\tau_{k+1}} - \mathbb{E}\left[Y^{\zeta_{k+1}}_{\tau_{k+1}}|\mathcal{F}^k_{\tau_{k+1}}\right]\right)1_{\{\tau_{k+1}\leq t\}}\right)^2 |\mathcal{F}^0_t\right] < +\infty. \quad (6.2.10)$$

2. The strategy $\phi^\star$ is admissible.

Proof. 1. Let $\phi \in \mathscr{A}^i_t$. Using (6.2.6) and (6.2.7), we have, for all $s \in [t,T]$,

$$\mathcal{A}^\phi_s = \sum_{k\geq 0}\left(Y^{\zeta_k}_{\tau_{k+1}} - \sum_{j=1}^d p^{\alpha_{k+1}}_{\alpha_k,j} Y^j_{\tau_{k+1}} + c_{\zeta_k,\alpha_{k+1}}\right)1_{\{t<\tau_{k+1}\leq s\}},$$

which is increasing since each summand is positive as $Y \in \mathcal{D}$.
We have, for $t \leq s \leq T$,

$$\mathcal{Y}^\phi_s - \mathcal{Y}^\phi_t = \sum_{k\geq 0}\left(Y^{\zeta_k}_{\tau_{k+1}\wedge s} - Y^{\zeta_k}_{\tau_k\wedge s}\right) + \sum_{k\geq 0}\left(Y^{\zeta_{k+1}}_{\tau_{k+1}} - Y^{\zeta_k}_{\tau_{k+1}}\right)1_{\{t<\tau_{k+1}\leq s\}}. \quad (6.2.11)$$

Using (6.2.2), we get, for all $k \geq 0$,

$$\begin{aligned} &Y^{\zeta_k}_{\tau_{k+1}\wedge s} - Y^{\zeta_k}_{\tau_k\wedge s} \\ &= -\int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s} f^{\zeta_k}(u, Y^{\zeta_k}_u, Z^{\zeta_k}_u)\mathrm{d}u + \int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s} Z^{\zeta_k}_u \mathrm{d}W_u - \int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s} \mathrm{d}K^{\zeta_k}_u. \end{aligned}$$

We also have, using (6.2.7), for all $k \geq 0$,

$$\begin{aligned} &Y^{\zeta_{k+1}}_{\tau_{k+1}} - Y^{\zeta_k}_{\tau_{k+1}} \\ &= \left(Y^{\zeta_{k+1}}_{\tau_{k+1}} - \mathbb{E}\left[Y^{\zeta_{k+1}}_{\tau_{k+1}}|\mathcal{F}^k_{\tau_{k+1}}\right]\right) - \left(Y^{\zeta_k}_{\tau_{k+1}} - \sum_{j=1}^d p^{\alpha_{k+1}}_{\zeta_k,j} Y^j_{\tau_{k+1}} + c_{\zeta_k,\alpha_{k+1}}\right) + c_{\zeta_k,\alpha_{k+1}}. \end{aligned}$$

Plugging the two previous equalities into (6.2.11), we get:

$$
\begin{aligned}
\mathcal{Y}_s^\phi - \mathcal{Y}_t^\phi =& \sum_{k\geq 0}\left(-\int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s} f^{\zeta_k}(u, Y_u^{\zeta_k}, Z_u^{\zeta_k})\mathrm{d}u + \int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s} Z_u^{\zeta_k}\mathrm{d}W_u - \int_{\tau_k\wedge s}^{\tau_{k+1}\wedge s}\mathrm{d}K_u^{\zeta_k}\right)\\
&+ \sum_{k\geq 0}\left(Y_{\tau_{k+1}}^{\zeta_{k+1}} - \mathbb{E}\Big[Y_{\tau_{k+1}}^{\zeta_{k+1}}|\mathcal{F}_{\tau_{k+1}}^k\Big]\right)1_{\{t<\tau_{k+1}\leq s\}} + A_s^\phi - A_t^\phi\\
&- \sum_{k\geq 0}\left(Y_{\tau_{k+1}}^{\zeta_k} - \sum_{j=1}^d p_{\zeta_k,j}^{\alpha_{k+1}} Y_{\tau_{k+1}}^j + c_{\zeta_k,\alpha_{k+1}}\right)1_{\{t<\tau_{k+1}\leq s\}}.
\end{aligned}
$$

By definition of $Y^\phi, Z^\phi, K^\phi, \mathcal{M}^\phi, \mathcal{A}^\phi$, we obtain, for all $s\in[t,T]$,

$$
\begin{aligned}
\mathcal{Y}_s^\phi =& \xi^{a_T} + \int_s^T f^{a_u}(u, \mathcal{Y}_u^\phi, \mathcal{Z}_u^\phi)\mathrm{d}u - \int_s^T \mathcal{Z}_u^\phi\mathrm{d}W_u - \int_s^T\mathrm{d}\mathcal{M}_u^\phi\\
&- \int_s^T \mathrm{d}A_u^\phi + \Big[\Big(\mathcal{A}_T^\phi + \mathcal{K}_T^\phi\Big) - \Big(\mathcal{A}_s^\phi + \mathcal{K}_s^\phi\Big)\Big]. \qquad (6.2.12)
\end{aligned}
$$

For $n\geq 1$, we consider the admissible strategy $\phi_n = (\zeta_0, (\tau_k^n, \alpha_k^n)_{k\geq 0})$ defined by $\zeta_0^n = i = \zeta_0, \tau_k^n = \tau_k, \alpha_k^n = \alpha_k$ for $k\leq n$, and $\tau_k^n = T+1$ for all $k>n$. We set $\mathcal{Y}_s^n := \mathcal{Y}_s^{\phi_n}, \mathcal{Z}_s^n := \mathcal{Z}_s^{\phi_n}, \ldots$ for all $s\in[t,T]$.
By (6.2.12) applied to the strategy ϕ^n, we get, recalling that $\mathcal{A}_t^n = 0$,

$$
\begin{aligned}
\mathcal{A}_{\tau_n\wedge T}^n =& \mathcal{Y}_t^n - \mathcal{Y}_{\tau_n\wedge T}^n - \int_t^{\tau_n\wedge T} f^{a_s^n}(s, \mathcal{Y}_s^n, \mathcal{Z}_s^n)\mathrm{d}s + \int_t^{\tau_n\wedge T}\mathcal{Z}_s^n\mathrm{d}W_s\\
&+ \int_t^{\tau_n\wedge T}\mathrm{d}\mathcal{M}_s^n + \int_t^{\tau_n\wedge T}\mathrm{d}A_s^n - \int_t^{\tau_n\wedge T}\mathrm{d}\mathcal{K}_s^n.
\end{aligned}
$$

We obtain, for a constant $\Lambda>0$,

$$
\begin{aligned}
\mathbb{E}\big[|\mathcal{A}_{\tau_n\wedge T}^n|^2\big] \leq& \Lambda\Bigg(\mathbb{E}\Bigg[|\mathcal{Y}_t^n|^2 + |\mathcal{Y}_{\tau_n\wedge T}^n|^2 + \int_t^{\tau_n\wedge T}|f(s, \mathcal{Y}_s^n, \mathcal{Z}_s^n)|^2\mathrm{d}s\\
&+ \int_t^{\tau_n\wedge T}|\mathcal{Z}_s^n|^2\mathrm{d}s + \int_t^{\tau_n\wedge T}\mathrm{d}[\mathcal{M}^n]_s + (A_T^\phi - A_t^\phi)^2 + (\mathcal{K}_T^n)^2\Bigg]\Bigg).
\end{aligned}
$$

We have

$$
\mathbb{E}\big[|\mathcal{Y}_r^n|^2\big] \leq \sum_{j=1}^d \mathbb{E}\big[|Y_r^j|^2\big] = \mathbb{E}\big[|Y_r|^2\big] \leq \mathbb{E}\Bigg[\sup_{t\leq r\leq T}|Y_r|^2\Bigg] = \|Y\|_{\mathbb{S}_d^2(\mathbb{F}^0)}^2,
$$

$$
\mathbb{E}\Bigg[\int_t^{\tau_n\wedge T}|f(s, \mathcal{Y}_s^n, \mathcal{Z}_s^n)|^2\mathrm{d}s\Bigg] \leq 4L^2T\|Y\|_{\mathbb{S}_d^2(\mathbb{F}^0)}^2 + 4L^2\|Z\|_{\mathbb{H}_d^2(\mathbb{F}^0)} + 2\|f(\cdot,0,0)\|_{\mathbb{H}_d^2(\mathbb{F}^0)}^2,
$$

and

$$
\mathbb{E}\big[(\mathcal{K}_T^n)^2\big] \leq \mathbb{E}\big[|K_T|^2\big].
$$

Thus, by these estimates and the fact that $A_T^\phi - A_t^\phi \in L^2(\mathcal{F}_T^\infty)$ as ϕ is admissible, there exists a constant $\Lambda_1>0$ such that

$$
\mathbb{E}\big[|\mathcal{A}_{\tau_n\wedge T}^n|^2\big] \leq \Lambda_1 + \Lambda\mathbb{E}\Bigg[\int_t^{\tau_n\wedge T}\mathrm{d}[\mathcal{M}^n]_s\Bigg]. \qquad (6.2.13)
$$

Using (6.2.12) applied to ϕ^n and Itô's formula between t and $\tau_n \wedge T$, since $\mathcal{M}^n$ is a square integrable martingale orthogonal to W and $A^n, \mathcal{A}^n, \mathcal{K}^n$ are non-decreasing and non-negative, we get

$$\begin{aligned}
&\mathbb{E}\Big[|\mathcal{Y}_t^n|^2 + \int_t^{\tau_n\wedge T} |\mathcal{Z}_s^n|^2 \mathrm{d}s + \int_t^{\tau_n\wedge T} \mathrm{d}[\mathcal{M}^n]_s\Big] \\
&= \mathbb{E}\Big[|\mathcal{Y}_{\tau_n\wedge T}^n|^2 + 2\int_t^{\tau_n\wedge T} \mathcal{Y}_s^n f^{a_s^n}(s,\mathcal{Y}_s^n,\mathcal{Z}_s^n)\mathrm{d} - 2\int_t^{\tau_n\wedge T} \mathcal{Y}_s^n \mathrm{d}A_s^n \\
&\quad +2\int_t^{\tau_n\wedge T} \mathcal{Y}_s^n \mathrm{d}\mathcal{A}_s^n + 2\int_t^{\tau_n\wedge T} \mathcal{Y}_s^n \mathrm{d}\mathcal{K}_s^n\Big] \\
&\le \mathbb{E}\big[|\mathcal{Y}_{\tau_n\wedge T}^n|^2\big] + 2\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n f^{a_s^n}(s,\mathcal{Y}_s^n,\mathcal{Z}_s^n)|\mathrm{d}s\Big] + 2\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}A_s^n\Big] \\
&\quad + 2\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}\mathcal{A}_s^n\Big] + 2\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}\mathcal{K}_s^n\Big]. \qquad (6.2.14)
\end{aligned}$$

We have, using Young's inequality, for some $\epsilon > 0$,

$$\begin{aligned}
\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n f^{a_s^n}(s,\mathcal{Y}_s^n,\mathcal{Z}_s^n)|\mathrm{d}s\Big] &\le \frac{1}{2}\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|^2\mathrm{d}s\Big] + \frac{1}{2}\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |f^{a_s^n}(s,\mathcal{Y}_s^n,\mathcal{Z}_s^n)|^2\mathrm{d}s\Big] \\
&\le T(\frac{1}{2}+2L^2)\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + 2L^2\|Z\|_{\mathbb{H}^2_d(\mathbb{F}^0)} + \|f(\cdot,0,0)\|^2_{\mathbb{H}^2_d(\mathbb{F}^0)}, \\
\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}A_s^n\Big] &\le \frac{1}{2}\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \frac{1}{2}\mathbb{E}\big[(A_T^\phi - A_t^\phi)^2\big], \\
\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}\mathcal{K}_s^n\Big] &\le \frac{1}{2}\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \frac{1}{2}\mathbb{E}\big[|K_T|^2\big],
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\Big[\int_t^{\tau_n\wedge T} |\mathcal{Y}_s^n|\mathrm{d}\mathcal{A}_s^n\Big] &\le \frac{1}{2\epsilon}\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \frac{\epsilon}{2}\mathbb{E}\big[(\mathcal{A}_T^n)^2\big] \\
&\le \frac{1}{2\epsilon}\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \frac{\epsilon}{2}\left(\Lambda_1 + \Lambda\mathbb{E}\Big[\int_t^{\tau_n\wedge T} \mathrm{d}[\mathcal{M}^n]_s\Big]\right).
\end{aligned}$$

Using these estimates together with (6.2.14) gives, for a constant $C_\epsilon > 0$ independent of n,

$$(1-\epsilon\Lambda)\,\mathbb{E}\Big[\int_t^{\tau_n\wedge T} \mathrm{d}[\mathcal{M}^n]_s\Big] \le C_\epsilon\left(\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \|Z\|_{\mathbb{H}^2_d(\mathbb{F}^0)} + \|f(\cdot,0,0)\|^2_{\mathbb{H}^2_d(\mathbb{F}^0)}\right),$$

and choosing $\epsilon = \frac{1}{2\Lambda}$ gives that $\mathbb{E}\Big[\int_t^{\tau_n\wedge T} \mathrm{d}[M^n]_s\Big]$ is upper bounded independently of n. We also get an upper bound independent of n for $\mathbb{E}\big[(\mathcal{A}^n_{\tau_n\wedge T})^2\big]$ by (6.2.13). Since $\int_t^{\tau_n\wedge T} \mathrm{d}[\mathcal{M}^n]_s$ (resp. $|\mathcal{A}^n_{\tau_n\wedge T}|^2$) is non-decreasing to $\int_t^T \mathrm{d}[\mathcal{M}^\phi]_s$ (resp. to $|\mathcal{A}_T^\phi|^2$), we obtain by monotone convergence the result. It is clear that $\mathcal{M}^\phi$ is a martingale satisfying $\mathcal{M}_t^\phi = 0$.

We now prove (6.2.10). Using that $\mathbb{E}\Big[(N_t^\phi)^2|\mathcal{F}_t^0\Big]$ is almost-surely finite as ϕ is admissible, we compute,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{k\geq 0}\left(Y^{\zeta_{k+1}}_{\tau_{k+1}} - \mathbb{E}\Big[Y^{\zeta_{k+1}}_{\tau_{k+1}}|\mathcal{F}^k_{\tau_{k+1}}\Big]\right)1_{\{\tau_{k+1}\leq t\}}\right)^2|\mathcal{F}_t^0\right]\\
&\leq \mathbb{E}\left[\left(\sum_{k\geq 0}\left|Y^{\zeta_{k+1}}_{\tau_{k+1}} - \sum_{j=1}^d p^{\alpha_{k+1}}_{\zeta_k,j}Y^j_t\right|1_{\{\tau_{k+1}\leq t\}}\right)^2|\mathcal{F}_t^0\right]\\
&\leq 4|Y_t|^2\mathbb{E}\Big[(N_t^\phi)^2|\mathcal{F}_t^0\Big] < +\infty.
\end{aligned}$$

2. For $n \geq 1$, we consider the admissible strategy $\phi_n = (\zeta_0, (\tau_k^n, \alpha_k^n)_{k\geq 0})$ defined by $\zeta_0^n = i = \zeta_0, \tau_k^n = \tau_k^\star, \alpha_k^n = \alpha_k^\star$ for $k \leq n$, and $\tau_k^n = T+1$ for all $k > n$. We set $\mathcal{Y}_s^n := \mathcal{Y}_s^{\phi_n}, \mathcal{Z}_s^n := \mathcal{Z}_s^{\phi_n}, \ldots$ for all $s \in [t,T]$.
   By definition of $\tau^\star, \alpha^\star$, recall (6.2.8)-(6.2.9), it is clear that $\mathcal{A}_s^n = 0$ and that $\int_{\tau_k^\star\wedge s}^{\tau_{k+1}^\star\wedge s}\mathrm{d}K_u^{\zeta_k^\star} = 0$ for all $k \leq n$ and $s \in [t,T]$. The identity (6.2.12) for the admissible strategy ϕ^n gives

$$\mathcal{Y}_t^n = \mathcal{Y}^{\tau_n^\star\wedge T} + \int_t^{\tau_n^\star\wedge T} f^{a_s^n}(s,\mathcal{Y}_s^n,\mathcal{Z}_s^n)\mathrm{d}s - \int_t^{\tau_n^\star\wedge T}\mathcal{Z}_s^n\mathrm{d}W_s - \int_t^{\tau_n^\star\wedge T}\mathrm{d}\mathcal{M}_s^n - \int_t^{\tau_n^\star\wedge T}\mathrm{d}A_s^n.$$

   Using similar arguments and estimates as in the precedent proof, we get

$$\mathbb{E}\Big[|A^n_{\tau_n^\star\wedge T} - A_t^n|^2\Big] \leq \Lambda_1 + \Lambda\mathbb{E}\left[\int_t^{\tau_n^\star\wedge T}\mathrm{d}[\mathcal{M}^u]_s\right],$$

   and, for $\epsilon > 0$,

$$(1-\epsilon\Lambda)\,\mathbb{E}\left[\int_t^{\tau_n^\star\wedge T}\mathrm{d}[\mathcal{M}^n]_s\right] \leq C_\epsilon\left(\|Y\|^2_{\mathbb{S}^2_d(\mathbb{F}^0)} + \|Z\|_{\mathbb{H}^2_d(\mathbb{F}^0)} + \|f(\cdot,0,0)\|^2_{\mathbb{H}^2_d(\mathbb{F}^0)}\right).$$

   Choosing $\epsilon = \frac{1}{2\Lambda}$ gives that $\mathbb{E}\Big[\int_t^{\tau_n^\star\wedge T}\mathrm{d}[\mathcal{M}^n]_s\Big]$ and $\mathbb{E}\Big[|A^n_{\tau_n^\star\wedge T} - A_t^n|^2\Big]$ are upper bounded uniformly in n, hence by monotone convergence, we get that $A_T^{\phi^\star} - A_t^{\phi^\star} \in L^2(\mathcal{F}_T^\infty)$.
   It remains to prove that $A_t^{\phi^\star} \in L^2(\mathcal{F}_t^\infty)$. We have $A_t^{\phi^\star} \leq \bar{c}N_t^{\phi^\star}$, and $\mathbb{E}\Big[(N_t^{\phi^\star})^2|\mathcal{F}_t^0\Big] < +\infty$ a.s. is immediate from Lemma 6.2.1, since $\mathbb{E}\Big[(N_t^{\phi^\star})^2|\mathcal{F}_t^0\Big] = \Psi(Y_t)$ with $\Psi(y) = \mathbb{E}^i\left[(N(y,\phi^\star))^2\right], y \in \mathcal{D}$, where $\mathbb{E}^i$ is the expectation under the probability $\mathbb{P}^i$ defined in Lemma 6.2.1.

□

The main theorem is the following:

Theorem 6.2.1. *1. For all $i \in \{1,\ldots,d\}$, $t \in [0,T]$ and $\phi \in \mathscr{A}_t^i$, we have $Y_t^i \geq \mathbb{E}\Big[U_t^\phi - A_t^\phi|\mathcal{F}_t^0\Big]$.*

2. *We have* $Y_t^i = \mathbb{E}\Big[U_t^{\phi^\star} - A_t^{\phi^\star}|\mathcal{F}_t^0\Big]$, *where* $\phi^\star = (i, (\tau_n^\star)_{n\geq 0}, (\alpha_n^\star)_{n\geq 1})$ *is defined in* (6.2.8)-(6.2.9).

Proof. 1. Let $\phi \in \mathscr{A}_t^i$, and consider the identity (6.2.12). Since $\mathcal{M}^\phi$ is a square integrable martingale, orthogonal to W, and since $\mathcal{A}_T^\phi + \mathcal{K}_T^\phi \in L^2(\mathcal{F}_T^\infty)$ and the process $\mathcal{A}^\phi + \mathcal{K}^\phi$ is non-negative and non-decreasing, the comparison Theorem 6.4.6 gives $\mathcal{Y}_t^\phi \geq U_t^\phi$, recall (6.2.1).
Now, we have

$$
\begin{aligned}
\mathcal{Y}_t^\phi &= Y_t^i + \sum_{k\geq 0}\left(Y_t^{\zeta_{k+1}} - Y_t^{\zeta_k}\right) 1_{\{\tau_{k+1}\leq t\}} \\
&= Y_t^i + \sum_{k\geq 0}\left(Y_t^{\zeta_{k+1}} - \mathbb{E}\Big[Y_t^{\zeta_{k+1}}|\mathcal{F}_{\tau_{k+1}}^k\Big]\right) 1_{\{\tau_{k+1}\leq t\}} \\
&\quad - \sum_{k\geq 0}\left(Y_t^{\zeta_k} - \sum_{j=1}^d p_{\zeta_k,j}^{\alpha_{k+1}} Y_{\tau_{k+1}}^j + c_{\zeta_k}^{\alpha_{k+1}}\right) 1_{\{\tau_{k+1}\leq t\}} + A_t^\phi.
\end{aligned}
$$

Since $U_t^\phi \leq \mathcal{Y}_t^\phi$ and $\sum_{k\geq 0}\left(Y_t^{\zeta_k} - \sum_{j=1}^d p_{\zeta_k,j}^{\alpha_{k+1}} Y_{\tau_{k+1}}^j + c_{\zeta_k}^{\alpha_{k+1}}\right) 1_{\{\tau_{k+1}\leq t\}} \geq 0$, we get

$$
\begin{aligned}
U_t^\phi - A_t^\phi &\leq Y_t^i + \sum_{k\geq 0}\left(Y_t^{\zeta_{k+1}} - \mathbb{E}\Big[Y_t^{\zeta_{k+1}}|\mathcal{F}_{\tau_{k+1}}^k\Big]\right) 1_{\{\tau_{k+1}\leq t\}} \\
&\quad - \sum_{k\geq 0}\left(Y_t^{\zeta_k} - \sum_{j=1}^d p_{\zeta_k,j}^{\alpha_{k+1}} Y_{\tau_{k+1}}^j + c_{\zeta_k}^{\alpha_{k+1}}\right) 1_{\{\tau_{k+1}\leq t\}} \\
&\leq Y_t^i + \sum_{k\geq 0}\left(Y_t^{\zeta_{k+1}} - \mathbb{E}\Big[Y_t^{\zeta_{k+1}}|\mathcal{F}_{\tau_{k+1}}^k\Big]\right) 1_{\{\tau_{k+1}\leq t\}}.
\end{aligned}
$$

Using (6.2.10), we can take conditional expectation on both side with respect to $\mathcal{F}_t^0$ to obtain the result.

2. Lemma 6.2.2 shows that the strategy $\phi^\star$ is admissible. Using (6.2.12), since $\mathcal{A}^{\phi^\star} = 0$ and $\int_{\tau_k^\star\wedge s}^{\tau_{k+1}^\star\wedge s} \mathrm{d}K_u^{\zeta_k^\star} = 0$ for all $k \geq 0$, we obtain

$$
\mathcal{Y}_s^{\phi^\star} = \xi^{a_T^\star} + \int_s^T f^{a_u^\star}(u, \mathcal{Y}_u^{\phi^\star}, \mathcal{Z}_u^{\phi^\star})\mathrm{d}u - \int_s^T \mathcal{Z}_u^{\phi^\star}\mathrm{d}W_u - \int_s^T \mathrm{d}\mathcal{M}_u^{\phi^\star} - \int_s^T \mathrm{d}A_u^{\phi^\star}.
$$

By uniqueness Theorem 6.4.4, we get that $\mathcal{Y}_t^{\phi^\star} = U_t^{\phi^\star}$, recall (6.2.1).
We also have

$$
\begin{aligned}
\mathcal{Y}_t^{\phi^\star} &= Y_t^i + \sum_{k\geq 0}\left(Y_t^{\zeta_{k+1}^\star} - Y_t^{\zeta_k^\star}\right) 1_{\{\tau_{k+1}^\star\leq t\}} \\
&= Y_t^i + \mathcal{M}_t^{\phi^\star} + A_t^{\phi^\star},
\end{aligned}
$$

thus $U_t^{\phi^\star} - A_t^{\phi^\star} = \mathcal{Y}_t^{\phi^\star} - A_t^{\phi^\star} = Y_t^i + \mathcal{M}_t^{\phi^\star}$, and taking conditional expectation gives the result.

□

As an immediate consequence, we obtain the uniqueness of the BSDE used to characterise the value function of the control problem.

Corollary 6.2.1. *Under Assumption 6.2.1, there exists a unique solution* $(Y, Z, K) \in \mathbb{S}^2_d(\mathbb{F}^0) \times \mathbb{H}^2_{k\times d}(\mathbb{F}^0) \times \mathbb{A}^2_d(\mathbb{F}^0)$ *to the Obliquely Reflected BSDE* (6.2.2)-(6.2.3)-(6.2.4).

6.3 Existence of Obliquely Reflected BSDEs associated to randomised switching problems

In this section, we address the existence of the Obliquely Reflected BSDE (6.2.2)-(6.2.3)-(6.2.4).

We first give some general properties of the domain $\mathcal{D}$. Then, in an uncontrolled Markovian framework, we give conditions under which there exists a solution.

6.3.1 Properties of the domain of reflection

We first show that the domain $\mathcal{D}$ defined in (6.2.5) is invariant by translation along the vector $(1, \dots, 1)$.

Lemma 6.3.1. *1. For every $x \in \mathcal{D}$, there is a unique decomposition $x = y + z$ where $y \in \mathcal{D} \cap \{y_d = 0\}$ and $z \in \mathbb{R}\left(\sum_{i=1}^d e_i\right)$.*

2. If $x \in \mathcal{D}$, we have $\mathcal{C}(x) \subset \{v \in \mathbb{R}^d : \sum_{i=1}^d v_i = 0\}$.

3. $\mathcal{C}(x) = \mathcal{C}(y)$, where y is from the above decomposition.

Proof. 1. For $x \in \mathcal{D}$, we set $y = x - z$ with $z = x_d \sum_{i=1}^d e_i$. It is clear that $y_d = 0$, let us show that $y \in \mathcal{D}$. If $i \in \{1, \dots, d\}$, we have

$$y_i = x_i - x_d \geq \max_{u\in\mathscr{C}}\left(\sum_{j=1}^d p^u_{i,j} x_j - c_{i,u}\right) - x_d = \max_{u\in\mathscr{C}}\left(\sum_{j=1}^d p^u_{i,j}(x_j - x_d) - c_{i,u}\right)$$
$$= \max_{u\in\mathscr{C}}\left(\sum_{j=1}^d p^u_{i,j} y_j - c_{i,u}\right),$$

thus $y \in \mathcal{D}$. The uniqueness is clear as $z = x_d \sum_{i=1}^d e_i$ necessarily.

2. Let $x \in \mathcal{D}$. The computation above shows that $x \pm \sum_{i=1}^d e_i \in \mathcal{D}$. Let $v \in \mathcal{C}(x)$. We then have, by definition,

$$0 \geq v^\top(x \pm \sum_{i=1}^d e_i - x) = \pm v^\top \sum_{i=1}^d e_i = \pm \sum_{i=1}^d v_i.$$

Thus we get $\sum_{i=1}^d v_i = 0$.

3. Let $x \in \mathcal{D}$. Since $x = y + x_d \sum_{i=1}^d e_i$, it is enough to show that for all $u \in \mathcal{D}$ and all $a \in \mathbb{R}, \mathcal{C}(u) \subset \mathcal{C}(u + a\sum_{i=1}^d e_i)$.
Let $v \in \mathcal{C}(u)$. We have, for all $z \in \mathcal{D}$, since $\sum_{i=1}^d v_i = 0$ and $v^\top(z-u) \le 0$,
$$v^\top(z - (u + a\sum_{i=1}^d e_i)) = v^\top(z-u) - av^\top \sum_{i=1}^d e_i = v^\top(z-u) \le 0.$$
Thus $v \in \mathcal{C}(u + a\sum_{i=1}^d e_i)$.

□

We now assume that there is no control on the switching distribution:

Assumption 6.3.1.
- *$\mathscr{C} = \{0\}$,*
- *The Markov chain with stochastic matrix $P := (p_{i,j})_{1\le i,j\le d} := (p^0_{i,j})_{1\le i,j\le d}$ is irreducible.*

We set $Q = I_d - P$. In this setting, the domain $\mathcal{D}$ is thus defined by:
$$\mathcal{D} = \{x \in \mathbb{R}^d : x_i \ge \sum_{j=1}^d p_{i,j}x_j - c_i\} = \{x \in \mathbb{R}^d : x \ge Px - c\} = \{x \in \mathbb{R}^d : Qx + c \ge 0\}.$$
Since P is irreducible, it is well known (see for example [Bia15], Section 2.5) that for all $1 \le i,j \le d$, the matrix $Q^{(i,j)}$ is invertible, and we have $(-1)^{i+j}\det Q^{(i,j)} = \det Q^{(i,i)} =: \mu_i > 0$. Moreover, $\mu Q = 0$ with $\mu = (\mu_i)_{i=1}^d$, i.e. $\frac{\mu}{\sum_{i=1}^d \mu_i}$ is the invariant probability measure for P. For all $\mathcal{I},\mathcal{J} \subset \{1,\dots,d\}^2$, we set $A^{(\mathcal{I},\mathcal{J})}$ to be the adjunct matrix of $Q^{(\mathcal{I},\mathcal{J})}$. For $1 \le i \le d$, we have in particular $A^{(i,i)}_{j,k} = (-1)^{j+k-1_{\{j>i\}}-1_{\{k>i\}}}\det Q^{(\{i,k\},\{i,j\})}$ for all $(j,k) \in \{1,\dots,d\}\backslash\{i\}$.
For all $1 \le i \ne j \le d$, we define $C^{i,j} := \left((Q^{(j,j)})^{-1}c^{(j)}\right)_i$ and $C^{j,j} = 0$. This can be thought as the mean cost to switch from state i to state j.

We first show that $\mathcal{D} \cap \{y_d = 0\}$ is always closed and bounded in $\{y_d = 0\}$.

Proposition 6.3.1. *For all $x \in \mathcal{D}$ and $1 \le i,j \le d$, we have*
$$-C^{i,j} \le x_i - x_j \le C^{j,i}. \tag{6.3.1}$$
In particular, the set $\mathcal{D} \cap \{y_d = 0\}$ is compact in $\{y \in \mathbb{R}^d : y_d = 0\}$.

Proof. For $j \in \{1,\dots,d\}$ and $x \in \mathbb{R}^d$, we set $\pi_j(x) = (x_i - x_j)_{i\ne j} \in \mathbb{R}^{d-1}$.
Let $x \in \mathcal{D}$ and $j \in \{1,\dots,d\}$. For all $i \in \{1,\dots,d\}, i \ne j$, we have, by definition of $\mathcal{D}$ and since $\sum_{k=1}^d p_{i,k} = 1$,
$$x_i - x_j \ge \sum_{k=1}^d p_{i,k}(x_k - x_j) - c_i.$$
Thus $\pi_j(x)$ satisfies to
$$Q^{(j,j)}\pi_j(x) \ge -c^{(j)}.$$

Since $\left(Q^{(j,j)}\right)^{-1} = \sum_{k\geq 0}\left(P^{(j,j)}\right)^k \geq 0$, we obtain

$$\pi_j(x) \geq -\left(Q^{(j,j)}\right)^{-1} c^{(j)} = -(C^{i,j})_{i\neq j}. \tag{6.3.2}$$

Inequality (6.3.2) gives $x_i - x_j \geq -C^{i,j}$ for all $i \neq j$.
Let $1 \leq i \neq j \leq d$. The precedent reasoning gives $x_i - x_j \geq -C^{i,j}$ and $x_j - x_i \geq -C^{j,i}$, thus (6.3.1) is proved.
Lastly, fix $y \in \mathcal{D} \cap \{y_d = 0\}$. We have by (6.3.1) $-C^{i,d} \leq y_i \leq C^{d,i}$ for all $1 \leq i \leq d-1$, thus $\mathcal{D} \cap \{y_d = 0\}$ is a closed bounded subset of $\{y_d = 0\}$. □

Remark 6.3.1. *The bounds (6.3.1) are optimal: the goal of the next results is to show that, if $\mathcal{D}$ is not empty, each column of the matrix $-C$ is a point of $\mathcal{D}$. In particular, the point $y = (y_1, \dots, y_d) = (-C^{1,d}, \dots, -C^{d-1,d}, 0)^\top$ is is in $\mathcal{D} \cap \{y_d = 0\}$, and it satisfies $-C^{i,d} = y_i - y_d$, i.e. the left inequality in (6.3.1) is saturated for $j = d$.*
More generally, since $\mathcal{D}$ is invariant by translation by $(1, \dots, 1)$, we obtain that the point $y^i = (C^{d,i} - C^{j,i})_{j=1}^d$ is in $\mathcal{D} \cap \{y_d = 0\}$. For each $j \neq i$, we obtain $y^i_j - y^i_i = C^{d,i} - C^{j,i} - C^{d,i} + C^{i,i} = -C^{j,i}$ which saturates the left inequality in (6.3.1). In fact, we prove that these points y^i are the extreme points of $\mathcal{D} \cap \{y_d = 0\}$, which is the simplex generated by these points.

Example 6.3.1. *To illustrate the precedent remark, we give an example in dimension $d = 3$. We consider the following situation*

$$P = \begin{pmatrix} 0 & 0.4 & 0.6 \\ 0.8 & 0 & 0.2 \\ 0.1 & 0.9 & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}.$$

In $\mathbb{R}^2 \simeq \{z = 0\}$, the following graph shows the domain $\mathcal{D} \cap \{z = 0\}$ in red, and the lines $x = -C^{1,3}$, $x = C^{3,1}$, $y = -C^{2,3}$, $y = C^{3,2}$, $y = x + C^{1,2}$ and $y = x - C^{2,1}$ in blue. As insighted in the remark and proved next, the inequalities are optimal and the points $p_1 = (C^{3,1}, C^{3,1} - C^{2,1}, 0)^\top = (\frac{8}{3}, \frac{4}{3}, 0)^\top$, $p_2 = (C^{3,2} - C^{1,2}, C^{3,2}, 0)^\top = (2, 2, 0)^\top$ and $p_3 = (-C^{1,3}, -C^{2,3}, 0)^\top = (\frac{4}{3}, \frac{2}{3}, 0)$ are the extreme points of $\mathcal{D} \cap \{z = 0\}$.

We now give necessary and sufficient conditions for $\mathcal{D}$ to be non-empty (resp. with non-empty interior).

The next technical lemma, whose proof is deferred to the appendix 6.4.2, is used in the sequel to obtain necessary and sufficient conditions for $\mathcal{D}$ to be non-empty (resp. to have non-empty interior).

Lemma 6.3.2. *1. For all $1 \leq i \leq d$, we have*

$$\sum_{k\neq i} p_{i,k} C^{k,i} = \frac{1}{\mu_i} \sum_{k\neq i} \mu_k c_k.$$

2. For all distinct $1 \leq i, j, k \leq d$, we have

$$\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k} = A^{(i,i)}_{j,j} \mu_k = A^{(j,j)}_{i,i} \mu_k.$$

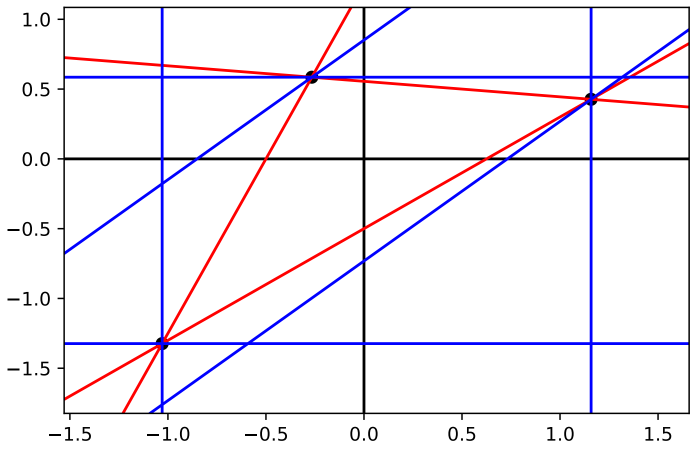

Figure 6.1 – The red domain is the randomised switching domain, the blue lines are given by (6.3.1), the black dots are p_1, p_2, p_3.

3. For all $1 \leq i \neq j \leq d$, *we have*

$$C^{i,j} + C^{j,i} = \frac{A^{(i,i)}_{j,j}}{\mu_i\mu_j}\mu c = \frac{A^{(j,j)}_{i,i}}{\mu_i\mu_j}\mu c$$

Lemma 6.3.3. *The following are equivalent:*

i) There exists $1 \leq i \neq j \leq d$ *such that* $C^{i,j} + C^{j,i} \geq 0$.

ii) We have $\mu c \geq 0$.

iii) We have $C + C^\top \geq 0$.

Similarly:

i) There exists $1 \leq i \neq j \leq d$ *such that* $C^{i,j} + C^{j,i} > 0$.

ii) We have $\mu c > 0$.

iii) We have $C^{i,j} + C^{j,i} > 0$ *for all* $1 \leq i \neq j \leq d$.

Proof. It is an immediate consequence of the previous lemma. □

Proposition 6.3.2. *The domain* $\mathcal{D}$ *is non-empty if and only if the first set of conditions of the previous lemma are satisfied.*
The domain $\mathcal{D}$ *has non-empty interior if and only if the second set of conditions of the previous lemma are satisfied.*

Proof. It is clear from (6.3.1) that $C + C^\top \geq 0$ (resp. $C^{i,j} + C^{j,i} > 0$ for all $i \neq j$) is a necessary condition for having $\mathcal{D}$ non-empty (resp with non-empty interior).
Assuming now that $C + C^\top \geq 0$, we show that each column of $-C$ is in $\mathcal{D}$. To do so,

we prove that $c - QC_j \geq 0$ for all $1 \leq j \leq d$, where C_j is the column j of C. This will prove that $\mathcal{D}$ is non-empty.
Fix $1 \leq j \leq d$. First, we deal with the $d-1$ coordinates $i \neq j$ by showing $c^{(j)} - (QC_j)^{(j)} = 0$.
Since $C^{j,j} = 0$, we have $(QC_j)^{(j)} = Q^{(j,j)} C_j^{(j)}$. Thus,

$$\begin{aligned} c^{(j)} - (QC_j)^{(j)} &= c^{(j)} - Q^{(j,j)} C_j^{(j)} \\ &= c^{(j)} - Q^{(j,j)} \left(Q^{(j,j)}\right)^{-1} c^{(j)} \\ &= 0, \end{aligned} \tag{6.3.3}$$

recall (6.3.2).
We now consider the coordinate j, and we show that $c_j - \sum_{k=1}^d Q_{j,k} C^{k,j} \geq 0$. Since $C^{j,j} = 0$ and $Q_{j,k} = -p_{j,k}$ for $k \neq j$, we have to show $c_j + \sum_{k \neq j} p_{j,k} C^{k,j} \geq 0$. By Lemma (6.3.2), we have $\sum_{k\neq j} p_{j,k} C^{k,j} = \frac{1}{\mu_j} \sum_{k \neq j} \mu_k c_k$. Hence, using that $\mu c \geq 0$ by the previous Lemma, we obtain

$$c_j + \sum_{k \neq j} p_{j,k} C^{k,j} = c_j + \frac{1}{\mu_j} \sum_{k \neq j} \mu_k c_k = \frac{\mu c}{\mu_j} \geq 0. \tag{6.3.4}$$

This proves that if $C + C^\top \geq 0$, then the domain $\mathcal{D}$ is non-empty.

Assume now that $C^{i,j} + C^{j,i} > 0$ for all $1 \leq i \neq j \leq d$, and we prove that $\mathcal{D} \cap \{y_d = 0\}$ has non-empty interior.
Since each column of $-C$ is in $\mathcal{D}$ and $\mathcal{D} = \mathcal{D} + \mathbb{R}(1, \dots, 1)$, we obtain that, for each $1 \leq i \leq d$, the vector $y_i := (C^{d,i} - C^{j,i})_{j=1}^d$ is in $\mathcal{D} \cap \{y_d = 0\}$.
Since $\mathcal{D}$ is convex, the convex hull of the points $y_i, 1 \leq i \leq d$ is included in $\mathcal{D} \cap \{y_d = 0\}$, and it is enough to show that these d points are affinely independent.
We prove that the vectors $\hat{C}_i := (\tilde{C}_i - \tilde{C}_d)^d, 1 \leq i \leq d-1$ are linearly independent in $\mathbb{R}^{d-1} \simeq \mathbb{R}^d \cap \{y_d = 0\}$ by showing that the square matrix M of size $d-1$ whose columns are $\hat{C}_i, 1 \leq i \leq d-1$ is nonsingular. The entries of M are given by $M_{i,j} = C^{i,d} + C^{d,j} - C^{i,j}$ for all $1 \leq i, j \leq d-1$.
In the first part of the proof, we have observed that $-C_i$ satisfies $c^{(i)} - (QC_i)^{(i)} = 0$, for all $1 \leq i \leq d$. Since $Q \sum_{i=1}^d e_i = 0$, this gives $c^{(i)} - Q^{(i,d)} \tilde{C}_i^{(i)} = 0$.
We thus obtain $Q^{(\{i,d\},d)} \hat{C}_i = 0$ for all $1 \leq i \leq d-1$, which can be written as $Q^{(\{i,d\},\{i,d\})} \hat{C}_i^{(i)} = -Q_i^{(i,d)} \hat{C}_{i,i} = -Q_i^{(i,d)} (C^{i,d} + C^{d,i})$, with Q_i being the ith column of Q and $\hat{C}_{i,i}$ the ith coordinate of $\hat{C}_i$.
Finally, for $1 \leq i \neq j \leq d-1$, we easily obtain, by developing $\det Q^{(\{i,d\},\{j,d\})}$ along the ith column,

$$A_{j,i}^{(d,d)} = \sum_{k \neq i,d} A_{j,k}^{(\{i,d\},\{i,d\})} Q_{k,i}.$$

Hence we obtain, for $1 \leq i \neq j \leq d$, denoting $\hat{C}_{i,j}$ the jth coordinate of $\hat{C}_i$,

$$\begin{aligned} \hat{C}_{i,j} &= -\frac{C^{i,d} + C^{d,i}}{\det Q^{(\{i,d\},\{i,d\})}} \sum_{k \neq i,d} A_{j,k}^{(\{i,d\},\{i,d\})} Q_{k,i} \\ &= -\frac{C^{i,d} + C^{d,i}}{\det Q^{(\{i,d\},\{i,d\})}} A_{j,i}^{(d,d)}. \end{aligned}$$

Thus $\det M = \pm \prod_{i=1}^{d-1} \frac{C^{i,d}+C^{d,i}}{\det Q^{(i,d;i,d)}} \det A^{(d,d)} = \pm \prod_{i=1}^{d-1} \frac{C^{i,d}+C^{d,i}}{\det Q^{(\{i,d\},\{i,d\})}} \neq 0.$ □

6.3.2 The uncontrolled Markovian framework

We now introduce a Markovian framework, and prove that a solution to (6.2.2)-(6.2.3)-(6.2.4) exists under Assumption 6.3.1 and a technical copositivity hypothesis, see 6.3.3 below.
To this effect, we rely on the existence theorem obtained in [CR18], which we recall next.

For all $(t,x) \in [0,T] \times \mathbb{R}^q$, let $X^{t,x}$ be the solution to the following SDE:

$$\begin{aligned} \mathrm{d}X_s &= b(s,X_s)\mathrm{d}s + \sigma(s,X_s)\mathrm{d}W_s, s \in [t,T], \\ X_t &= x. \end{aligned}$$

We are interested in the solutions $(Y^{t,x}, Z^{t,x}, K^{t,x}) \in \mathbb{S}^2_d(\mathbb{F}^0) \times \mathbb{H}^2_{d\times k}(\mathbb{F}^0) \times \mathbb{A}^2_d(\mathbb{F}^0)$ of (6.2.2)-(6.2.3)-(6.2.4), where the terminal condition satisfies $\xi = g(X_T^{t,x})$, and the driver satisfies $f(\omega,s,y,z) = \psi(s, X_s^{t,x}(\omega), y, z)$ for some deterministic measurable functions g, ψ. We next give the precise set of Assumptions we need to obtain our results.

6.3.2.1 A general existence result

We recall here the existence result proved in [CR18], see also [DAFH17].

Assumption 6.3.2. *i)* $\psi : [0,T] \times \mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$ *is a measurable function satisfying: there exists* $p \geq 0$ *and* $L \geq 0$ *such that, for all* $(t,x,y,z) \in [0,T] \times \mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^{d\times k}$*, we have*

$$|\psi(t,x,y,z)| \leq L(1 + |x|^p + |y| + |z|).$$

Moreover, $\psi(t,x,\cdot,\cdot)$ *is continuous on* $\mathbb{R}^d \times \mathbb{R}^{d\times k}$ *for all* $(t,x) \in [0,T] \times \mathbb{R}^q$.

ii) $(b,\sigma) : [0,T] \times \mathbb{R}^q \to \mathbb{R}^q \times \mathbb{R}^{q\times k}$ *is a measurable function satisfying, for all* $(t,x,y) \in [0,T] \times \mathbb{R}^q \times \mathbb{R}^q$,

$$\begin{aligned} |b(t,x)| + |\sigma(t,x)| &\leq L(1 + |x|), \\ |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| &\leq L|x-y|. \end{aligned}$$

iii) $g : \mathbb{R}^q \to \mathbb{R}^d$ *is measurable and for all* $(t,x) \in [0,T] \times \mathbb{R}^q$*, we have*

$$|g(t,x)| \leq L(1 + |x|^p).$$

iv) Let $\mathcal{X} = \{\mu(t,x;s,dy), x \in \mathbb{R}^q$ *and* $0 \leq t \leq s \leq T\}$ *be the family of laws of* $X^{t,x}$ *on* $\mathbb{R}^q$*, i.e., the measures such that* $\forall A \in \mathcal{B}(\mathbb{R}^q)$*,* $\mu(t,x;s,A) = \mathbb{P}(X_s^{t,x} \in A)$.
For any $t \in [0,T)$*, for any* $\mu(0,a;t,dy)$*-almost every* $x \in \mathbb{R}^q$*, and any* $\delta \in]0, T-t]$*, there exists an application* $\phi_{t,x} : [t,T] \times \mathbb{R}^d \to \mathbb{R}_+$ *such that:*

(a) $\forall k \geq 1$*,* $\phi_{t,x} \in L^2([t+\delta,T] \times [-k,k]^q; \mu(0,a;s,dy)ds)$*,*

(b) $\mu(t,x;s,dy)ds = \phi_{t,x}(s,y)\mu(0,a;s,dy)ds$ *on* $[t+\delta,T] \times \mathbb{R}^q$.

v) $H : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is a measurable function, and there exists $\eta > 0$ such that, for all $(y, y') \in \mathcal{D} \times \mathbb{R}^d$ and $v \in \mathfrak{n}(y)$, we have

$$v^\top H(t,x,y,z)v \geq \eta,$$
$$|H(t,x,y',z)| \leq L.$$

Moreover, H is continuous on $\mathcal{D}$.

Remark 6.3.2. *Assumption iv) is true as soon as σ is uniformly elliptic, see [HLP97].*

We thus have the following existence result

Theorem 6.3.1 ([CR18], Theorem 4.1)**.** *Under Assumption 6.3.2, there exists a solution $(Y^{t,x}, Z^{t,x}, \Psi^{t,x}) \in \mathbb{S}^2_d(\mathbb{F}^0) \times \mathbb{H}^2_{d\times k}(\mathbb{F}^0) \times \mathbb{H}^2_d(\mathbb{F}^0)$ of the following system*

$$Y_s = g(X_T^{t,x}) + \int_s^T \psi(u, X_u^{t,x}, Y_u, Z_u)\mathrm{d}u - \int_s^T Z_u \mathrm{d}W_u - \int_s^T H(Y_u)\Psi_u \mathrm{d}u, \, s \in [t,T], \tag{6.3.5}$$

$$Y_s \in \mathcal{D}, \;\; \Psi_s \in \mathcal{C}(Y_s), \;\; t \leq s \leq T, \tag{6.3.6}$$

$$\int_t^T 1_{\{Y_s \notin \partial\mathcal{D}\}} |\Psi_s| \mathrm{d}s = 0. \tag{6.3.7}$$

6.3.2.2 The existence theorem

We construct a function $H : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ which satisfies to Assumption 6.3.2 v) and such that

$$H(y)v \in \mathcal{C}_o(y), \tag{6.3.8}$$

for all $y \in \mathcal{D}$ and $v \in \mathcal{C}(y)$, where $\mathcal{C}_o(y)$ is the cone of directions of reflection, given here by

$$\mathcal{C}_o(y) := -\sum_{i=1}^d \mathbb{R}_+ e_i 1_{\{y_i = \max_{u\in\mathscr{C}}\{\sum_{j=1}^d p^u_{ij} y_j - c_{i,u}\}\}}.$$

If Assumption 6.3.2 i), ii), iii), iv) is also satisfied, we obtain the existence of a solution to (6.3.5)-(6.3.6)-(6.3.7). Setting $K^{t,x}_s := -\int_t^s H(Y^{t,x}_u)\Psi^{t,x}_u \mathrm{d}u$ for $t \leq s \leq T$ shows that $(Y^{t,x}, Z^{t,x}, K^{t,x})$ is a solution to (6.2.2)-(6.2.3)-(6.2.4).

In addition to Assumption 6.3.1, we need the following technical assumption to construct H satisfying Assumption 6.3.2 v) and (6.3.8).

Assumption 6.3.3. *For all $1 \leq i,j \leq d$, the matrix $Q^{(j,i)}\left(Q^{(i,i)}\right)^{-1}\left(Q^{(j,i)}\right)^\top$ is strictly copositive, meaning that for all $0 \leq x \in \mathbb{R}^{d-1}, x \neq 0$, we have*

$$x^\top Q^{(j,i)}\left(Q^{(i,i)}\right)^{-1}\left(Q^{(j,i)}\right)^\top x > 0.$$

The main theorem is the following.

Theorem 6.3.2. *Suppose that Assumption 6.3.2 i), ii), iii), iv), Assumption 6.3.1 and Assumption 6.3.3 are satisfied and that $\mathcal{D}$ has non-empty interior.*

Then there exists $H : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ satisfying to 6.3.2 v), and Assumption 6.2.1 is satisfied with $\xi = g(X_T)$ and $f(\omega, s, y, z) = \psi(s, X^{t,x}_s(\omega), y, z)$.

Proof. For each $1 \le i \le d$, we define the following column vectors:

$$y^i := (C^{d,i} - C^{j,i})_{j=1}^d,$$
$$n_i := (-Q_{i,j})_{j=1}^d.$$

The proof is divided into several steps.

Step 1. Since $\mathcal{D}$ has non-empty interior, the previous proposition shows that the points $y^i, 1 \le i \le d$, are affinely independent and that the simplex generated by these points is included in $\mathcal{D} \cap \{y_d = 0\}$. We show that the converse inclusion is also true. For $y \in \mathbb{R}^{d-1} \cap \{y_d = 0\}$, there exists a unique $(\lambda_1, \dots, \lambda_{d-1}) \in \mathbb{R}^{d-1}$ such that $y = \sum_{i=1}^d \lambda_i y^i$, where $\lambda_d = 1 - \sum_{i=1}^{d-1} \lambda_i$. If $y \in \mathcal{D}$, we have, with $\lambda = (\lambda_1, \dots, \lambda_d)^\top$,

$$V = Q \begin{pmatrix} y^1 & \cdots & y^d \end{pmatrix} \lambda + c \ge 0.$$

The computations of the previous proof, see (6.3.3)-(6.3.4), give $Qy^i = \left(\frac{\mu c}{\mu_i} - c_i\right) e_i - \sum_{j \ne i} c_j e_j$, hence:

$$\begin{aligned} V &= \sum_{i=1}^d \left[\left(\frac{\mu c}{\mu_i} - c_i \right) \lambda_i - \sum_{j \ne i} c_j \lambda_j + c_i \right] e_i \\ &= \sum_{i=1}^d \frac{\mu c}{\mu_i} \lambda_i e_i. \end{aligned}$$

Since $V \ge 0$ and $\frac{\mu c}{\mu_i} > 0$, we obtain $\lambda_i \ge 0$ for all $1 \le i \le d$. We have thus proved that y is in the convex hull of the $y^i, 1 \le i \le d$.

Step 2. We compute the outward normal cone $\mathcal{C}(y)$ for each $y = y^i, 1 \le i \le d$:

$$\mathcal{C}(y^i) = \sum_{j \ne i} \mathbb{R}_+ n_j.$$

We fix $1 \le i \le d$.

Step 2.a. First, let us show that $(n_j)_{j \ne i}$ is a basis for $\{y \in \mathbb{R}^d : \sum_{i=1}^d v_i = 0\}$. Let $1 \le i \ne j \le d$. It is clear that $n_j \in \{v \in \mathbb{R}^d : \sum_{k=1}^d v_k = 0\}$. Since it is a hyperplane of $\mathbb{R}^d$ and that the family $(n_j)_{j \ne i}$ has $d-1$ elements, it is enough to show that the vectors are linearly independent. We observe that the matrix whose lines are the $n_j^{(i)}, j \ne i$, is $-Q^{(i)}$. Since P is irreducible, $Q^{(i)}$ is invertible. The vectors $n_j^{(i)}, j \ne i$ form a basis of $\mathbb{R}^{d-1}$, hence the vectors $n_j, j \ne i$ form a basis of $\{v \in \mathbb{R}^d : \sum_{k=1}^d v_k = 0\}$.

Step 2.b. We show that $n_j \in \mathcal{C}(y^i)$ for all $j \ne i$. For any $z \in \mathcal{D}$, by definition of y^i, we have

$$c_j \ge \sum_{k=1}^d p_{jk} z_k - z_j = n_j^\top z,$$
$$c_j = \sum_{k=1}^d p_{jk} y_k^i - y_j^i = n_j^\top y^i.$$

This gives $n_j^\top(z-y^i) = n_j^\top z - n_j^\top y^i \leq 0$, hence $n_j \in \mathcal{C}(y^i)$.
Step 2.c. Conversely, since $(n_j)_{j\neq i}$ is a basis for $\{v \in \mathbb{R}^d : \sum_{i=1}^d v_i = 0\} \ni \mathcal{C}(y^i)$, for $v \in \mathcal{C}(y^i)$ there exists a unique $\lambda = (\lambda_j)_{j\neq i} \in \mathbb{R}^{d-1}$ such that $v = \sum_{j\neq i} \lambda_j n_j = (n_j)_{j\neq i}\lambda$. For all $z \in \mathcal{D}$, we then have:

$$\begin{aligned} 0 &\geq \lambda^\top \left[(n_j)_{j\neq i}\right]^\top (z-y^i) \\ &= -\lambda^\top Q^{(i,\cdot)}(z-y^i) \\ &= -\lambda^\top \left[Q^{(i,\cdot)}z + c^{(i)}\right] \end{aligned}$$

Hence $\lambda^\top \left[Q^{(i,\cdot)}z + c^{(i)}\right] \geq 0$ for all $z \in \mathcal{D}$.
For any $j \neq i$, if one takes $z = y^j \in \mathcal{D}$, by definition of y^j, one gets $Q^{(i,\cdot)}y^j + c^{(i)} = \frac{\mu c}{\mu_j} e_j$, with $\frac{\mu c}{\mu_j} > 0$. Hence

$$0 \leq \lambda^\top \left[Q^{(i,\cdot)}z + c^{(i)}\right] = \lambda_j \frac{\mu c}{\mu_j},$$

which gives $\lambda_j \geq 0$. Hence $v \in \sum_{j\neq i} \mathbb{R}_+ n_j$.

Step 3. We construct $H(y^i)$ satisfying $v^\top H(y^i)v > 0$ for each $1 \leq i \leq d$ and for each $v \in \bigcup_{j=1}^d \mathcal{C}(y^j)$. Fix $1 \leq i \leq d$, and let $B^i \in \mathbb{R}^{(d-1)\times(d-1)}$ be the base change matrix from $(-n_j^{(i)})_{j\neq i}$ to the canonical basis of $\mathbb{R}^{d-1}$. We set $H(y^i) := I^i B^i P^i$, with $I^i : \mathbb{R}^{d-1} \to \mathbb{R}^d$ and $P^i : \mathbb{R}^d \to \mathbb{R}^{d-1}$ the linear maps defined by

$$\begin{aligned} I^i(x_1,\dots,x_{d-1}) &= (x_1,\dots,x_{i-1},0,x_i,\dots,x_{d-1}), \\ P^i(x_1,\dots,x_d) &= (x_1,\dots,x_{i-1},x_{i+1},\dots,x_d). \end{aligned}$$

Let's remark that $(B^i)^\top = (Q^{(i)})^{-1}$, and by construction of $H^i(y^i)$, for all $x \in \mathbb{R}^d$, we have $x^\top H^i(y^i)x = (x^{(i)})^\top B^i x^{(i)} = (x^{(i)})^\top \left(Q^{(i)}\right)^{-1} x^{(i)}$.
Let $0 \neq v \in \mathcal{C}(y^j) = \sum_{k\neq j} \mathbb{R}_+ n_k$ for some $1 \leq j \leq d$. Since $n_k = -Q_k^\top$, there exists a vector $0 \neq \lambda = (\lambda_k)_{k\neq j} \geq 0$ such that $v = -(Q^{(j,\cdot)})^\top \lambda$, where $Q^{(j,\cdot)} \in \mathbb{R}^{d\times(d-1)}$ is obtained by removing the jth line of Q.
Then $v^{(i)} = -\left(Q^{(j,i)}\right)^\top \lambda$, and we seek to prove that

$$0 < (v^{(i)})^\top \left(Q^{(i,i)}\right)^{-1} v^{(i)} = \lambda^\top Q^{(j,i)} \left(Q^{(i,i)}\right)^{-1} \left(Q^{(j,i)}\right)^\top \lambda,$$

which is true by Assumption 6.3.3, as $0 \neq \lambda \geq 0$.

Step 4. We now $H : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that Assumption 6.3.2 v) and (6.3.8) are satisfied. Since $\mathcal{D} \cap \{y_d = 0\}$ is the simplex generated by the d points $y_i, 1 \leq i \leq d$, for each point $y \in \mathcal{D} \cap \{y_d = 0\}$, there exists a unique vector $\lambda \in \mathbb{R}^d, \sum_{i=1}^d \lambda_i = 1, \lambda_i \geq 0$ for all i, satisfying to $y = \sum_{i=1}^d \lambda_i y_i$. We then set $H(y) = \sum_{i=1}^d \lambda_i H(y_i)$. Since $\mathcal{D} \cap \{y_d = 0\}$ is a simplex, it is clear that $\mathcal{C}(y) \subset \bigcup_{j=1}^d \mathcal{C}(y^j)$ for each boundary point y, and hence from Step 3. we obtain that

$$v^\top H(y)v \geq = \sum_{i=1}^d \lambda_i v^\top H(y^i)v \geq \sum_{i=1}^d \lambda_i \eta = \eta.$$

Moreover, it is easy to check that (6.3.8) is satisfied on the boundary points of $\mathcal{D}$ satisfying $y_d = 0$.
Finally, we set $H(x) = H(x - x_d \sum_{i=1}^d e_i)$ for $x \in \mathcal{D}$ and $H(x) = H(\pi(x))$ for $x \in \mathbb{R}^d$, where $\pi : \mathbb{R}^d \to \mathcal{D}$ is the orthogonal projection onto the convex set $\mathcal{D}$. Using Lemma (6.3.1), it is clear that Assumption 6.3.2 v) and (6.3.8) are satisfied on $\mathcal{D}$. □

6.4 Appendix

6.4.1 Study of the filtration $\mathbb{F}^\infty$

We fix a strategy $\phi \in \Phi$ and we study the filtrations $\mathbb{F}^i, i \geq 0$ and $\mathbb{F}^\infty$ which are constructed in subsection (6.2.1).
For each $n \geq 0$, we define a new filtration $\mathbb{G}^n = (\mathcal{G}^n_t)_{t\geq 0}$ by the relations $\mathcal{G}^0_t = \mathcal{F}^0_t$ and for $n \geq 1$, $\mathcal{G}^n_t = \mathcal{F}^0_t \vee \sigma(X_i, i \leq n) = \mathcal{G}^{n-1}_t \vee \sigma(X_n)$.

6.4.1.1 Representation Theorems

The goal of this section is to derive Integral Representation Theorems for the filtrations $\mathbb{F}^i, i \geq 0$ and $\mathbb{F}$.
We first recall, see [AJ17]:

Theorem 6.4.1 (Lévy). *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a filtered probability space with $\mathbb{F}$ non necessarily right-continuous. Let $\xi \in \mathcal{F}$ and X a $\mathbb{F}-$supermartingale.*

1. *We have $\mathbb{E}[\xi|\mathcal{F}_t] \to \mathbb{E}[\xi|\mathcal{F}_\infty]$ a.s. and in L^1, as $t \to \infty$.*
2. *If t_n decreases to t, we have $X_{t_n} \to X_{t^+}$ a.s. and in L^1 as $n \to \infty$.*

In particular, if $X_t = \mathbb{E}[\xi|\mathcal{F}_t]$, we get that $\mathbb{E}[\xi|\mathcal{F}_{t_n}] \to \mathbb{E}[\xi|\mathcal{F}_{t^+}]$ a.s. and in L^1 as $n \to \infty$, for t_n decreasing to t.

We now recall an important notion of coincidence of filtrations between two stopping times, introduced in [AJ17]. This will be useful for our purpose in the sequel.
Let S, T two random times, which are stopping times for two filtrations $\mathbb{H}^1 = (\mathcal{H}^1_t)_{t\geq 0}$ and $\mathbb{H}^2 = (\mathcal{H}^2_t)_{t\geq 0}$. We set

$$[\![S, T[\![:= \{(\omega, s) \in \Omega \times \mathbb{R}_+ : S(\omega) \leq s < T(\omega)\}.$$

We say that $\mathbb{H}^1$ and $\mathbb{H}^2$ *coincide* on $[\![S, T[\![$ if

1. for each $t \geq 0$ and each $\mathcal{H}^1_t$-measurable variable ξ, there exists a $\mathcal{H}^2_t$-measurable χ variable such that $\xi 1_{S\leq t<T} = \chi 1_{S\leq t<T}$,
2. for each $t \geq 0$ and each $\mathcal{H}^2_t$-measurable variable χ, there exists a $\mathcal{H}^1_t$-measurable variable ξ such that $\chi 1_{S\leq t<T} = \xi 1_{S\leq t<T}$.

We now study the right-continuity of the filtration $\mathbb{G}^n$ for some $n \geq 0$. Using its specific structure, it is easy to compute conditional expectations. Lévy's theorem then allows to obtain the right-continuity.

Lemma 6.4.1. *Let $n \geq 0$.*

1. *If $\xi \in L^1(\mathcal{F}^0_\infty)$ and $\xi' \in L^1(\sigma(X_i, 1 \leq i \leq n))$, then for $t \geq 0$, we have $\mathbb{E}[\xi\xi'|\mathcal{G}^n_t] = \mathbb{E}\big[\xi|\mathcal{F}^0_t\big]\,\xi'$.*

2. *$\mathbb{G}^n$ is right-continuous.*

Proof. 1. If $F \in \mathcal{F}^0_t$ and $F' \in \sigma(X_i, 1 \leq i \leq n)$, we have, by independence,

$$\begin{aligned}\mathbb{E}\big[\xi\xi' 1_{F\cap F'}\big] &= \mathbb{E}[\xi 1_F]\,\mathbb{E}\big[\xi' 1_{F'}\big]\\ &= \mathbb{E}\big[\mathbb{E}\big[\xi|\mathcal{F}^0_t\big]\,1_F\big]\,\mathbb{E}\big[\xi' 1_{F'}\big]\\ &= \mathbb{E}\big[\xi'\mathbb{E}\big[\xi|\mathcal{F}^0_t\big]\,1_{F\cap F'}\big]\,.\end{aligned}$$

Since $\{F \cap F', F \in \mathcal{F}^0_t, F' \in \sigma(X_i, 1 \leq i \leq n)\}$ is a π-system generating $\mathcal{G}^n_t$, the result follows by a monotone class argument.

2. Let $t \geq 0$ and t_m decreasing to t. We have, using Lévy's Theorem, the previous point and the right-continuity of $\mathbb{F}^0$,

$$\begin{aligned}\mathbb{E}\big[\xi\xi'|\mathcal{G}^n_{t+}\big] &= \lim_m \mathbb{E}\big[\xi\xi'|\mathcal{G}^n_{t_m}\big]\\ &= \lim_m \xi'\mathbb{E}\big[\xi|\mathcal{F}^0_{t_m}\big]\\ &= \xi'\mathbb{E}\big[\xi|\mathcal{F}^0_t\big]\\ &= \mathbb{E}\big[\xi\xi'|\mathcal{G}^n_t\big]\,.\end{aligned}$$

By a monotone class argument, we have $\mathbb{E}\big[\xi|\mathcal{G}^n_{t+}\big] = \mathbb{E}[\xi|\mathcal{G}^n_t]$ for all bounded $\mathcal{G}^n_\infty$-measurable ξ, hence the right-continuity of $\mathbb{G}^n$.

□

Using the previous Lemma, we show how to compute conditional expectations in $\mathbb{F}$ and $\mathbb{F}^n$ for all $n \geq 0$, and show that these filtrations are right-continuous.

Proposition 6.4.1. 1. *For all $m \geq n \geq 0$, $\mathbb{F}^n$, $\mathbb{F}^m$ and $\mathbb{F}^\infty$ coincide on $[\![0, \tau_{n+1}[\![$. For all $n \geq 0$, $\mathbb{F}^n$ and $\mathbb{G}^n$ coincide on $[\![\tau_n, +\infty[\![$.*

2. *For all $n \geq 0$ and $t \geq 0$, we have, for $\xi \in L^1(\mathcal{F}^{n+1}_\infty)$:*

$$\mathbb{E}\big[\xi|\mathcal{F}^{n+1}_t\big] = \mathbb{E}[\xi|\mathcal{F}^n_t]\,1_{t<\tau_{n+1}} + \mathbb{E}\big[\xi|\mathcal{G}^{n+1}_t\big]\,1_{\tau_{n+1}\leq t}. \tag{6.4.1}$$

Let $t \geq 0$ such that $\sum_{n=0}^{+\infty} \mathbb{P}(\tau_n \leq t < \tau_{n+1}) = 1$. Then, for $\xi \in L^1(\mathcal{F}^\infty_\infty)$,

$$\mathbb{E}[\xi|\mathcal{F}^\infty_t] = \sum_{n=0}^{+\infty} \mathbb{E}[\xi|\mathcal{F}^n_t]\,1_{\tau_n\leq t<\tau_{n+1}}.$$

3. *For all $n \geq 0$, $\mathbb{G}^n$ is right-continuous.*

4. *The filtration $\mathbb{G}$ is right-continuous on $[0, T]$.*

Proof. 1. Let $t \geq 0$ be fixed.
If ξ is $\mathcal{F}^n_t$-measurable, since $\mathcal{F}^n_t \subset \mathcal{F}^m_t \subset \mathcal{F}^\infty_t$ for $m \geq n$, taking $\chi = \xi$ gives a $\mathcal{F}^m_t$-measurable (resp. $\mathcal{F}^\infty_t$-measurable) random variable such that $\xi 1_{t<\tau_{n+1}} =$

$\chi 1_{t<\tau_{n+1}}$.
Conversely, if χ is a $\mathcal{F}_t^m$-measurable random variable, then

$$\chi = f(\tilde{\chi}, X_1 1_{\tau_1 \le t}, \dots, X_m 1_{\tau_m \le t}),$$

for a measurable f and a $\mathcal{F}_t^0$-measurable variable $\tilde{\chi}$. Since $X_m 1_{\tau_m \le t} = 0$ on $\{t < \tau_{n+1}\}$ when $m \ge n$, one gets:

$$\begin{aligned}\chi 1_{t<\tau_{n+1}} &= f(\tilde{\chi}, X_1 1_{\tau_1 \le t}, \dots, X_n 1_{\tau_n \le t}, 0, \dots, 0) 1_{t<\tau_{n+1}} \\ &=: \xi 1_{t<\tau_{n+1}},\end{aligned}$$

where ξ is $\mathcal{F}_t^n$-measurable.
Last, let χ be a $\mathcal{F}_t^\infty$-measurable variable. Then $\chi = f(\tilde{\chi}, X_{i_1} 1_{\tau_{i_1} \le t}, \dots, X_{i_N} 1_{\tau_{i_N} \le t})$ for some $N \ge 0$ and $1 \le i_1 \le \dots \le i_N$, and the same arguments applies.

The proof of the second claim is straightforward as one remarks that for $t \ge 0$ and $n \ge 1$, the equality $f(\xi, X_1, \dots, X_n) 1_{\tau_n \le t} = f(\xi, X_1 1_{\tau_1 \le t}, \dots, X_n 1_{\tau_n \le t}) 1_{\tau_n \le t}$ holds, since the random times $\tau_i, i \ge 0$ are non-decreasing.

2. Let $n \ge 0$ and $\xi \in L^1(\mathcal{F}_\infty^{n+1})$.
Since $\mathbb{F}^n$ and $\mathbb{F}^{n+1}$ coincide on $[\![0, \tau_{n+1}[\![$, we have $\mathbb{E}\big[\xi|\mathcal{F}_t^{n+1}\big] 1_{t<\tau_{n+1}} = \tilde{\xi} 1_{t<\tau_{n+1}}$ for a $\mathcal{F}_t^n$-measurable variable $\tilde{\xi}$. In particular, the left hand side is also $\mathcal{F}_t^n$-measurable. Hence $\mathbb{E}\big[\xi|\mathcal{F}_t^{n+1}\big] 1_{t<\tau_{n+1}} = \mathbb{E}\big[\mathbb{E}\big[\xi|\mathcal{F}_t^{n+1}\big] 1_{t<\tau_{n+1}}|\mathcal{F}_t^n\big] = \mathbb{E}[\xi|\mathcal{F}_t^n] 1_{t<\tau_{n+1}}$.
Similarly, since $\mathbb{F}^{n+1}$ and $\mathbb{G}^{n+1}$ coincide on $[\![\tau_{n+1}, +\infty[\![$, we have $\mathbb{E}\big[\xi|\mathcal{G}_t^{n+1}\big] 1_{\tau_{n+1} \le t} = \hat{\xi} 1_{\tau_{n+1} \le t}$ for a $\mathcal{F}_t^{n+1}$-measurable variable $\hat{\xi}$. In particular, the left hand side is $\mathcal{F}_t^{n+1}$-measurable. Hence $\mathbb{E}\big[\xi|\mathcal{G}_t^{n+1}\big] 1_{\tau_{n+1} \le t} = \mathbb{E}\big[\mathbb{E}\big[\xi|\mathcal{G}_t^{n+1}\big] 1_{\tau_{n+1} \le t}|\mathcal{F}_t^{n+1}\big] = \mathbb{E}\big[\xi|\mathcal{F}_t^{n+1}\big] 1_{\tau_{n+1} \le t}$.
Let $t \ge 0$ such that $\sum_n \mathbb{P}(\tau_n \le t < \tau_{n+1}) = 1$. We have, since $\mathbb{G}$ and $\mathbb{G}^n$ coincide on $[\![0, \tau_{n+1}[\![$, using the same arguments as before,

$$\begin{aligned}\mathbb{E}[\xi|\mathcal{G}_t] &= \sum_n \mathbb{E}[\xi|\mathcal{G}_t] 1_{\tau_n \le t < \tau_{n+1}} \\ &= \sum_n \mathbb{E}[\xi|\mathcal{G}_t^n] 1_{\tau_n \le t < \tau_{n+1}}.\end{aligned}$$

3. We prove by induction that $\mathbb{F}^n$ is right-continuous. Since $\mathbb{F}^0$ is the augmented Brownian filtration, the result is true for $n = 0$.
Assume now that $\mathbb{F}^{n-1}, n \ge 1$, is right-continuous. Let $t \ge 0, \xi \in L^1(\mathcal{F}_\infty^n)$ and $(t_m)_{m \ge 0}$ such that $t_m \ge t_{m+1}$ and $\lim_m t_m = t$. We have, using the previous point and the right-continuity of $\mathbb{F}^{n-1}$ and $\mathbb{G}^n$:

$$\begin{aligned}\mathbb{E}[\xi|\mathcal{F}_{t+}^n] &= \lim_m \mathbb{E}\big[\xi|\mathcal{F}_{t_m}^n\big] \\ &= \lim_m \mathbb{E}\big[\xi|\mathcal{F}_{t_m}^n\big] 1_{t_m<\tau_n} + \mathbb{E}\big[\xi|\mathcal{F}_{t_m}^n\big] 1_{\tau_n \le t_m} \\ &= \lim_m \mathbb{E}\big[\xi|\mathcal{F}_{t_m}^{n-1}\big] 1_{t_m<\tau_n} + \mathbb{E}\big[\xi|\mathcal{G}_{t_m}^n\big] 1_{\tau_n \le t_m} \\ &= \mathbb{E}\big[\xi|\mathcal{F}_t^{n-1}\big] 1_{t<\tau_n} + \mathbb{E}[\xi|\mathcal{G}_{t_m}] 1_{\tau_n \le t} \\ &= \mathbb{E}[\xi|\mathcal{F}_t^n].\end{aligned}$$

4. Let $t < T, \xi \in L^1(\mathcal{F}^\infty_\infty), (t_m)_{m\geq 0}$ such that $T > t_m > t_{m+1}$ and $\lim_m t_m = t$. We have, by Lévy's Theorem and the first point,

$$\begin{aligned}
\mathbb{E}[\xi|\mathcal{F}^\infty_{t^+}] &= \lim_m \mathbb{E}\big[\xi|\mathcal{F}^\infty_{t_m}\big] \\
&= \lim_m \sum_{n=0}^{+\infty} \mathbb{E}\big[\xi|\mathcal{F}^\infty_{t_m}\big]\, 1_{\tau_n \leq t_m < \tau_{n+1}} \\
&= \lim_m \sum_{n=0}^{+\infty} \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big]\, 1_{\tau_n \leq t_m < \tau_{n+1}}.
\end{aligned}$$

Fix $\omega \in \Omega$. We have that $t_m < \hat{T} < \tau_{N+1}(\omega)$, hence

$$\begin{aligned}
\mathbb{E}[\xi|\mathcal{F}^\infty_{t^+}](\omega) &= \lim_m \sum_{n=0}^{+\infty} \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big](\omega) 1_{\tau_n(\omega) \leq t_m < \tau_{n+1}(\omega)} \\
&= \lim_m \sum_{n=0}^{N(\omega)+1} \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big](\omega) 1_{\tau_n(\omega) \leq t_m < \tau_{n+1}(\omega)} \\
&= \sum_{n=0}^{N(\omega)+1} \lim_m \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big](\omega) 1_{\tau_n(\omega) \leq t_m < \tau_{n+1}(\omega)} \\
&= \sum_{n=0}^{+\infty} \lim_m \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big](\omega) 1_{\tau_n(\omega) \leq t_m < \tau_{n+1}(\omega)}.
\end{aligned}$$

Finally using the right-continuity of each $\mathbb{F}^n$, we get

$$\begin{aligned}
\mathbb{E}[\xi|\mathcal{F}^\infty_{t^+}] &= \sum_{n=0}^{+\infty} \lim_m \mathbb{E}\big[\xi|\mathcal{F}^n_{t_m}\big]\, 1_{\tau_n \leq t_m < \tau_{n+1}} \\
&= \sum_{n=0}^{+\infty} \mathbb{E}[\xi|\mathcal{F}^n_t]\, 1_{\tau_n \leq t < \tau_{n+1}} \\
&= \mathbb{E}[\xi|\mathcal{F}^\infty_t]\,,
\end{aligned}$$

which proves that $\mathbb{F}^\infty$ is right-continuous on $[0, T]$.

□

Lemma 6.4.2. *Let $n \geq 0$ and $\xi \in L^1(\mathcal{F}^n_\infty)$. Let σ be a $\mathbb{F}^n-$stopping time. We have:*

$$\mathbb{E}\big[\xi|\mathcal{F}^{n+1}_\sigma\big] = \mathbb{E}[\xi|\mathcal{F}^n_\sigma]\,.$$

Proof. Assume first that $\sigma = s$ is deterministic.
Let $\tilde{\xi} = \psi(\chi, X_{n+1} 1_{\tau_{n+1} \leq s})$ be a $\mathcal{F}^{n+1}_s$-measurable bounded variable, where χ is $\mathcal{F}^n_s$-measurable. We need to show

$$\mathbb{E}\Big[\xi\tilde{\xi}\Big] = \mathbb{E}\Big[\mathbb{E}[\xi|\mathcal{F}^n_s]\,\tilde{\xi}\Big]\,.$$

We have, with $\hat{\psi}(y) := \int_x \psi(y, x)\mathbb{P}_{X_{n+1}}(\mathrm{d}x)$:

$$\begin{aligned}
\mathbb{E}\big[\mathbb{E}[\xi|\mathcal{G}^n_s]\,\psi(\chi, X_{n+1}1_{\tau_{n+1}\leq s})\big] &= \mathbb{E}\big[\mathbb{E}[\xi|\mathcal{G}^n_s]\,\psi(\chi, 0)1_{s<\tau_{n+1}}\big] + \mathbb{E}\big[\mathbb{E}[\xi|\mathcal{G}^n_s]\,\psi(\chi, X_{n+1})1_{\tau_{n+1}\leq s}\big] \\
&= \mathbb{E}\big[\xi\psi(\chi, 0)1_{s<\tau_{n+1}}\big] + \mathbb{E}\Big[\xi\hat{\psi}(\chi)1_{\tau_{n+1}\leq s}\Big]\,,
\end{aligned}$$

and the same computation with ξ instead of $\mathbb{E}[\xi|\mathcal{G}^n_s]$ gives the same result.
Let σ be a $\mathbb{F}^n$-stopping time, and let $\xi_s = \mathbb{E}[\xi|\mathcal{F}^n_s] = \mathbb{E}\big[\xi|\mathcal{F}^{n+1}_s\big]$. Since $\mathbb{F}^n$ (or $\mathbb{F}^{n+1}$) is right-continuous, there exists a right-continuous modification of $(\xi_s)_{s\geq 0}$. Applying Doob's Theorem twice gives $\xi_\sigma = \mathbb{E}[\xi|\mathcal{F}^n_\sigma]$ and $\xi_\sigma = \mathbb{E}\big[\xi|\mathcal{F}^{n+1}_\sigma\big]$, hence the result. □

We are now in position to prove a Integral Representation Theorem in the filtrations $\mathbb{F}^n$, for all $n \geq 0$.

Proposition 6.4.2. *Let $n \geq 0$ and $\xi \in L^2(\mathcal{F}^n_T)$. Then there exists a $\mathbb{G}^n-$predictable process ϕ such that*

$$\xi = \mathbb{E}\big[\xi|\mathcal{F}^n_{T\wedge\tau_n}\big] + \int_{T\wedge\tau_n}^T \psi_s \mathrm{d}W_s.$$

Proof. We prove the Theorem by induction on $n \geq 0$, following ideas from [Ame00]. The case $n = 0$ is the usual Martingale Representation Theorem in the augmented Brownian filtration $\mathbb{F}^0$.
Assume now that the statement is true for all $\xi \in L^2(\mathcal{F}^{n-1}_T)$ $(n \geq 1)$. Let $\xi \in L^2(\mathcal{F}^n_T)$. Since $\mathcal{F}^n_T = \mathcal{F}^{n-1}_T \vee \sigma(X_n 1_{\tau_n \leq T})$, we get that $\xi = \lim_{m\to\infty} \xi_m$ in $L^2(\mathcal{F}^n_T)$, with $\xi_m = \sum_{i=1}^{l_m} \chi^i_m \zeta^i_m$ and $(\chi^i_m, \zeta^i_m) \in L^\infty(\mathcal{F}^{n-1}_T)\times L^\infty(\sigma(X_n 1_{\tau_n\leq T}))$ for all $m \geq 0$ and $1 \leq i \leq l_m$. By induction, there exists $\mathbb{F}^{n-1}$-predictable processes $\psi^{i,m}$ such that $\chi^i_m = \mathbb{E}\Big[\chi^i_m|\mathcal{F}^{n-1}_{T\wedge\tau_{n-1}}\Big] + \int_{T\wedge\tau_{n-1}}^T \psi^{i,m}_s dW_s$. Since τ_n is a $\mathbb{F}^{n-1}$-stopping time with $\tau_n \geq \tau_{n-1}$, we get:

$$\chi^i_m = \mathbb{E}\big[\chi^i_m|\mathcal{F}^{n-1}_{T\wedge\tau_n}\big] + \int_{T\wedge\tau_n}^T \psi^{i,m}_s \mathrm{d}W_s.$$

Since $\xi^i_m \in L^\infty(\sigma(X_n 1_{\tau_n\leq T})) \subset L^2(\mathcal{F}^n_{T\wedge\tau_n})$, we get

$$\zeta^i_m \int_{T\wedge\tau_n}^T \psi^{i,m}_s \mathrm{d}W_s = \int_{T\wedge\tau_n}^T \zeta^i_m \psi^{i,m}_s \mathrm{d}W_s.$$

In addition, since χ^i_m is $\mathcal{F}^{n-1}_T$-measurable and $\zeta^i_m \in L^2(\mathcal{F}^n_{T\wedge\tau_n})$, we get, by the previous Lemma,

$$\zeta^i_m \mathbb{E}\big[\chi^i_m|\mathcal{F}^{n-1}_{T\wedge\tau_n}\big] = \zeta^i_m \mathbb{E}\big[\chi^i_m|\mathcal{F}^n_{T\wedge\tau_n}\big] = \mathbb{E}\big[\chi^i_m\zeta^i_m|\mathcal{F}^n_{T\wedge\tau_n}\big].$$

Summing over $1 \leq i \leq l_m$ gives:

$$\begin{aligned}
\xi_m &= \sum_{i=1}^{l_m} \chi^i_m \zeta^i_m \\
&= \sum_{i=1}^{l_m} \mathbb{E}\big[\chi^i_m\zeta^i_m|\mathcal{F}^n_{T\wedge\tau_n}\big] + \sum_{i=1}^{l_m} \int_{T\wedge\tau_n}^T \zeta^i_m \psi^{i,m}_s \mathrm{d}W_s \\
&= \mathbb{E}\big[\xi_m|\mathcal{F}^n_{T\wedge\tau_n}\big] + \int_{T\wedge\tau_n}^T \psi^m_s \mathrm{d}W_s,
\end{aligned}$$

where $\psi^m := \sum_{i=1}^{l_m} \psi^{i,m}_s \zeta^i_m$.
Finally, since $\xi_m \to \xi$ in $L^2(\mathcal{F}^n_T)$, we get that $\mathbb{E}\big[\xi_m|\mathcal{F}^n_{T\wedge\tau_n}\big] \to \mathbb{E}\big[\xi|\mathcal{F}^n_{T\wedge\tau_n}\big]$ in $L^2(\mathcal{F}^n_T)$, hence $\int_{T\wedge\tau_n}^T \psi^m_s dW_s$ converges to a limit $\int_{T\wedge\tau_n}^T \psi_s dW_s$ for a $\mathbb{F}^n$-predictable process ψ. □

Theorem 6.4.2. *Let $0 \leq T \leq +\infty$ and $\xi \in L^2(\mathcal{G}^n_T)$. For all $0 \leq k \leq n$, there exists $\mathbb{F}^k$-predictable processes ψ^k such that:*

$$\begin{aligned}\xi &= \mathbb{E}[\xi] + \sum_{k=0}^{n-1} \int_{T\wedge\tau_k}^{T\wedge\tau_{k+1}} \psi^k_s \mathrm{d}W_s + \int_{T\wedge\tau_n}^{T} \psi^n_s \mathrm{d}W_s \\ &\quad + \sum_{k=0}^{n-1} \left(\mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{T\wedge\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^{k}_{T\wedge\tau_{k+1}}\Big] \right) \\ &= \mathbb{E}[\xi] + \int_0^T \Psi^n_s \mathrm{d}W_s + \sum_{k=0}^{n-1} \left(\mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{T\wedge\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^{k}_{T\wedge\tau_{k+1}}\Big] \right),\end{aligned}$$

with $\Psi^n_t := \sum_{k=0}^{n-1} \psi^k_t 1_{T\wedge\tau_k \leq t \leq T\wedge\tau_{k+1}} + \psi^n_t 1_{T\wedge\tau_n \leq t \leq T}$.

Proof. This is an immediate consequence of the previous Theorem. □

Last, we extend this Theorem to obtain an Integral Representation Theorem in $\mathbb{F}^\infty$. We now fix $\xi \in L^2(\mathcal{F}^\infty_T)$ and consider the filtration $\mathbb{A} = (\mathcal{A}_n)_{n\in\mathbb{N}}$ defined by $\mathcal{A}_n := \mathcal{F}^n_T$. We have $\mathcal{A}_\infty = \bigvee_n \mathcal{A}_n = \mathcal{F}^\infty_T$. By Lévy's Theorem, we get

$$\mathbb{E}[\xi|\mathcal{F}^n_T] = \mathbb{E}[\xi|\mathcal{A}_n] \to \mathbb{E}[\xi|\mathcal{A}_\infty] = \xi, \text{ a.s.}. \tag{6.4.2}$$

For all $n \geq 0$, since $\mathcal{F}^n_T \subset \mathcal{F}_T$, we can write:

$$\begin{aligned}\mathbb{E}[\xi|\mathcal{F}^n_T] &= \mathbb{E}[\xi] + \sum_{k=0}^{n-1} \int_{T\wedge\tau_k}^{T\wedge\tau_{k+1}} \psi^{n,k}_s \mathrm{d}W_s + \int_{T\wedge\tau_n}^{T} \psi^{n,n}_s \mathrm{d}W_s \\ &\quad + \sum_{k=0}^{n-1} \left(\mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{T\wedge\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^{k}_{T\wedge\tau_{k+1}}\Big] \right)\end{aligned}$$

Lemma 6.4.3. *We have $\psi^{n,k} = \psi^{k,k}$ on $[T\wedge\tau_k, T\wedge\tau_{k+1})$, for all $n \geq k$.*

Proof. It follows easily by induction, comparing $\mathbb{E}\big[\xi|\mathcal{F}^k_T\big]$ and $\mathbb{E}\big[\mathbb{E}[\xi|\mathcal{F}^n_T]\,|\mathcal{F}^k_T\big]$ and using Itô's isometry. □

For all $n \geq 0$, we define $\psi^n := \psi^{n,n}$. Thus we have, for all $n \geq 0$,

$$\begin{aligned}\mathbb{E}[\xi|\mathcal{G}^n_T] &= \mathbb{E}[\xi] + \sum_{k=0}^{n-1} \int_{T\wedge\tau_k}^{T\wedge\tau_{k+1}} \psi^k_s \mathrm{d}W_s + \int_{T\wedge\tau_n}^{T} \psi^n_s \mathrm{d}W_s \\ &\quad + \sum_{k=0}^{n-1} \left(\mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{T\wedge\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^{k}_{T\wedge\tau_{k+1}}\Big] \right)\end{aligned}$$

We set, for $0 \leq s \leq T$,

$$\begin{aligned}\Psi_s &= \sum_{k=0}^{+\infty} \psi^k_s 1_{T\wedge\tau_k \leq s < T\wedge\tau_{k+1}}, \\ \Psi^n_s &= \Psi_s 1_{s \leq T\wedge\tau_{n+1}} + \psi^n_s 1_{T\wedge\tau_{n+1} < s}, \text{ and} \\ \Delta^k_s &:= \mathbb{E}\Big[\xi|\mathcal{F}^{k+1}_{s\wedge\tau_{k+1}}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}^{k}_{s\wedge\tau_{k+1}}\Big],\end{aligned}$$

so that

$$\mathbb{E}[\xi|\mathcal{F}_T^n] = \mathbb{E}[\xi] + \int_0^T \Psi_s^n \mathrm{d}W_s + \sum_{k=0}^{n-1} \Delta_T^k.$$

Theorem 6.4.3 (Integral Representation Theorem for $\mathbb{F}^\infty$)**.** *For $\xi \in L^2(\mathcal{F}_T^\infty)$, we have*

$$\xi = \mathbb{E}[\xi] + \int_0^T \Psi_s \mathrm{d}W_s + \sum_{k=0}^{+\infty} \Delta_T^k.$$

Proof. By definition of $N = N_T^\phi$, we have $T < \tau_{n+1}$ on $\{n \geq N\}$, see Section 6.2. Thus,

$$\begin{aligned} 1_{N\leq n} \int_0^T \Psi_s^n \mathrm{d}W_s &= \left(\int_0^{T\wedge\tau_{n+1}} \Psi_s \mathrm{d}W_s + \int_{T\wedge\tau_{n+1}}^T \psi_s^n \mathrm{d}W_s \right) 1_{N\leq n} \\ &= 1_{N\leq n} \int_0^T \Psi_s \mathrm{d}W_s. \end{aligned}$$

Moreover, if $k \geq n$, we have, since $T \wedge \tau_{k+1} = T$,

$$\begin{aligned} \Delta_T^k 1_{N\leq n} &= \left(\mathbb{E}\Big[\xi|\mathcal{F}_{T\wedge T_{k+1}}^{k+1}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}_{T\wedge_{k+1}}^{k}\Big] \right) 1_{N\leq n} \\ &= \left(\mathbb{E}\Big[\xi|\mathcal{F}_T^{k+1}\Big] - \mathbb{E}\Big[\xi|\mathcal{F}_T^{k}\Big] \right) 1_{N\leq n}. \end{aligned}$$

Applying (6.4.1) to $\chi = \mathbb{E}\Big[\xi|\mathcal{F}_T^{k+1}\Big]$, we get

$$\chi = \mathbb{E}\Big[\chi|\mathcal{F}_T^{k+1}\Big] = \mathbb{E}\Big[\chi|\mathcal{F}_T^{k}\Big] 1_{T<\tau_{k+1}} + \mathbb{E}\Big[\chi|\mathcal{G}_T^{k+1}\Big] 1_{\tau_{k+1}\leq T}.$$

Since $T < \tau_{n+1} \leq \tau_{k+1}$ on $\{N \leq n\}$, we finally obtain

$$\mathbb{E}\Big[\xi|\mathcal{F}_T^{k+1}\Big] 1_{N\leq n} = \chi 1_{N\leq n} = \mathbb{E}\Big[\chi|\mathcal{F}_T^{k}\Big] 1_{N\leq n} = \mathbb{E}\Big[\xi|\mathcal{F}_T^{k}\Big] 1_{N\leq n},$$

which gives $\Delta_T^k 1_{N\leq n} = 0$.
Thus:

$$\begin{aligned} \mathbb{E}[\xi|\mathcal{F}_T^n] 1_{N\leq n} &= \left(\mathbb{E}[\xi] + \int_0^T \Psi_s^n \mathrm{d}W_s + \sum_{k=0}^{n-1} \Delta_T^k \right) 1_{N\leq n} \\ &= \left(\mathbb{E}[\xi] + \int_0^T \Psi_s \mathrm{d}W_s + \sum_{k=0}^{+\infty} \Delta_T^k \right) 1_{N\leq n}. \end{aligned}$$

Since $1_{N\leq n} \to 1$ a.s. as $n \to \infty$ as $N = N_T^\phi$ and ϕ is an admissible strategy, see Section 6.2, we get, sending n to $+\infty$, recall (6.4.2),

$$\xi = \mathbb{E}[\xi] + \int_0^T \Psi_s \mathrm{d}W_s + \sum_{k=0}^{+\infty} \Delta_T^k.$$

□

Remark 6.4.1. *We have:*

$$\left[\int_0^\cdot \Psi_s dW_s, \sum_{k=0}^{+\infty} \Delta^k\right]_t = 0,$$
$$\left[\int_0^\cdot \Psi_s dW_s\right]_t = \int_0^t \Psi_s^2 ds,$$
$$\left[\sum_{k=0}^{+\infty} \Delta^k\right]_t = \sum_{\tau_{k+1}\le t} |\Delta_k|^2.$$

Thus the martingales $\int_0^\cdot \Psi_s dW_s$ and $\sum_k \Delta^k$ are orthogonal.

6.4.1.2 Backward Stochastic Differential Equations

We now consider Backward Stochastic Differential Equations. Let $\mathbb{F}$ be one of the filtrations $\mathbb{F}^i, i \geq 0$ or $\mathbb{F}^\infty$. Let ξ be a $\mathcal{F}_T$-measurable variable and $f : \Omega \times \mathbb{R}^d \times \mathbb{R}^{k\times d} \times \mathbb{R}^d \to \mathbb{R}^d$. We assume here that ξ and f are standard parameters [EKPQ97]:

- $\xi \in L^2(\mathcal{F}_T)$,
- $f(\cdot, 0, 0) \in \mathbb{H}^2(\mathbb{R}^d)$,
- There exists $C > 0$ such that

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq C\left(|y_1 - y_2| + |z_1 + z_2|\right).$$

Under these hypothesis, since $\mathbb{F}$ is right-continuous, one can prove ([EKPQ97], Theorem 5.1):

Theorem 6.4.4. *There exists a unique solution $(Y, Z, M) \in \mathcal{S}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathbb{R}^{k\times d}) \times \mathbb{H}^2(\mathbb{R}^d)$ such that M is a martingale with $M_0 = 0$, orthogonal to the Brownian motion, and satisfying*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s - \int_t^T \mathrm{d}M_s.$$

One can easily deal with linear BSDEs in $\mathbb{F}$, and the specific form of its solutions allows to prove a Comparison Theorem. The proofs follow closely [EKPQ97], Theorem 2.2.

Theorem 6.4.5. *Let (b, c) be a bounded $(\mathbb{R} \times \mathbb{R}^n)$-valued predictable process and let $a \in \mathbb{H}^2(\mathbb{R})$. Let $\xi \in L^2(\mathcal{F}_T)$ and let $(Y, Z, M) \in \mathcal{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}^k) \times \mathbb{H}^2(\mathbb{R})$ be the unique solution to*

$$Y_t = \xi + \int_t^T \left(a_s Y_s + b_s Z_s + c_s\right)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s - \int_t^T \mathrm{d}M_s.$$

Let $\Gamma \in \mathbb{H}^2(\mathbb{R})$ the solution to

$$\Gamma_t = 1 + \int_0^t \Gamma_s a_s \mathrm{d}s + \int_0^t \Gamma_s b_s \mathrm{d}W_s.$$

Then, for all $t \in [0,T]$*, one has almost surely,*

$$Y_t = \Gamma_t^{-1}\mathbb{E}\Big[\Gamma_T\xi + \int_t^T \Gamma_s c_s \mathrm{d}s|\mathcal{F}_t\Big].$$

Proof. We fix $t \in [0,T]$ and we apply Itô's formula to the process $Y_t\Gamma_t$:

$$\mathrm{d}\left(Y_t\Gamma_t\right) = Y_{t^-}\mathrm{d}\Gamma_t + \Gamma_{t^-}\mathrm{d}Y_t + \mathrm{d}\left[Y,\Gamma\right]_t.$$

Since Γ is continuous, we get $[Y,\Gamma]_t = \langle Y^c, \Gamma^c\rangle_t + \sum_{s\le t}\left(\Delta Y_s\right)\left(\Delta\Gamma_s\right) = \langle Y^c,\Gamma\rangle_t$, thus,

$$\mathrm{d}\left(Y_t\Gamma_t\right) = \Gamma_t\left(b_tY_t + Z_t\right)\mathrm{d}W_t + \Gamma_t\mathrm{d}M_t - \Gamma_t c_t \mathrm{d}t.$$

We define a martingale by $N_t = \int_0^t \Gamma_s(b_sY_s + Z_s)\mathrm{d}W_s + \int_0^t \Gamma_s\mathrm{d}M_s$, and the previous equality gives

$$Y_T\Gamma_T = Y_t\Gamma_t - \int_t^T \Gamma_s c_s\mathrm{d}s + N_T - N_t.$$

Taking conditional expectation with respect to $\mathcal{F}_t$ on both sides gives the result. □

Theorem 6.4.6. *Let* (ξ, f) *and* (ξ', f') *two standard parameters. Let* $(Y,Z,M) \in \mathcal{S}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathbb{R}^{k\times d}) \times \mathbb{H}^2(\mathbb{R}^d)$ *(resp.* (Y',Z',M')*) the solution associated with* (ξ, f) *(resp.* (ξ', f')*). Assume that*

- $\xi \ge \xi'$,
- $f(Y',Z',M') \ge f'(Y',Z',M')$.

Then $Y_t \ge Y'_t$ *almost surely for all* $t \in [0,T]$.

Proof. Since f is Lipschitz, we consider the bounded processes a, b and c defined by:

$$\begin{aligned}
a_t &= \frac{f(t,Y_t,Z_t) - f(t,Y'_t,Z_t)}{(Y_t - Y'_t)}1_{Y_t \neq Y'_t},\\
b^i_t &= \frac{f(t,Y'_t,\tilde{Z}^{i-1}_t) - f(t,Y'_t,\tilde{Z}^i_t)}{Z^i_t - (Z')^i_t}1_{Z^i_t\neq (Z')^i_t} \quad (0 \le i \le k),\\
c_t &= f(Y',Z',M') - f'(Y',Z',M'),
\end{aligned}$$

where $\tilde{Z}^i_t = ((Z')^1_t, \dots, (Z')^i_t, Z^{i+1}_t, Z^k_t), 0 \le i \le k$. Setting $\delta Y_t = Y_t - Y'_t, \delta Z_t = Z_t - Z'_t$ and $\delta M_t = M_t - M'_t$, we observe that $(\delta Y, \delta Z, \delta M)$ is the solution to the following linear BSDE:

$$\delta Y_t = \delta Y_T + \int_t^T \left(a_s\delta Y_s + b_s\delta Z_s + c_s\right)\mathrm{d}s - \int_t^T \delta Z_s\mathrm{d}W_s - \int_t^T \delta M_s.$$

Using the previous Theorem, we get $Y_t = \Gamma_t^{-1}\mathbb{E}\Big[\delta Y_T\Gamma_T + \int_t^T \Gamma_s c_s\mathrm{d}s|\mathcal{F}_t\Big]$. By definition, Γ is a strictly positive process, and δY_t and c are positive by hypothesis, hence $Y_t \ge 0$. □

6.4.2 Proof of Lemma 6.3.2

Proof. 1. We have, with L_i being the ith line of Q,

$$\mu_i \sum_{k\neq i} p_{i,k} C^{k,i} = -L_i^{(i)} A^{(i,i)} c^{(i)}$$
$$= -\sum_{l\neq i} Q_{i,l} \sum_{k\neq i} A_{l,k}^{(i,i)} c_k$$
$$= -\sum_{k\neq i} c_k \sum_{l\neq i} Q_{i,l} A_{l,k}^{(i,i)},$$

so it is enough to show that $\mu_k = -\sum_{l\neq i} Q_{i,l} A_{l,k}^{(i,i)}$ for all $k \neq i$.
We have:

$$-\sum_{l\neq i} Q_{i,l} A_{l,k}^{(i,i)} = -Q_{i,k} A_{k,k}^{(i,i)} - \sum_{l\neq i,k} Q_{i,l} A_{l,k}^{(i,i)}$$
$$= -Q_{i,k} A_{i,i}^{(k,k)} - \sum_{l\neq i,k} Q_{i,l} (-1)^{l+1-1_{\{l>i\}}-1_{\{k>i\}}} \det Q^{(\{i,k\},\{i,l\})}.$$

Fix $k \neq i$ and $l \neq k, i$. We compute $\det Q^{(\{i,k\},\{i,l\})}$. Adding to column k every other column of $Q^{(\{i,k\},\{i,l\})}$, column k becomes $-X_i^{(i,k)} - X_l^{(i,k)}$, where X_l is the column l of Q, recall that $\sum_{i=1}^d X_l = 0$. Using the linearity of the determinant with respect to column k and moving column k to column i (resp. l), we get, after working out the sign and using that $Q_{i,i} = -\sum_{l\neq i} Q_{i,l}$,

$$-\sum_{l\neq i} Q_{i,l} A_{l,k}^{(i,i)} = -Q_{i,k} A_{i,i}^{(k,k)} - \sum_{l\neq i,k} Q_{i,l} A_{i,i}^{(k,k)} + \sum_{l\neq i,k} Q_{i,l} A_{l,i}^{(k,k)}$$
$$= -\sum_{l\neq i} Q_{i,l} A_{i,i}^{(k,k)} + \sum_{l\neq i,k} Q_{i,l} A_{l,i}^{(k,k)}$$
$$= Q_{i,i} A_{i,i}^{(k,k)} + \sum_{l\neq i,k} Q_{i,l} A_{l,i}^{(k,k)}$$
$$= \sum_{l\neq k} Q_{i,l} A_{l,i}^{(k,k)} = \mu_k.$$

2. By definition, we have

$$A_{i,k}^{(j,j)} = (-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}} \det Q^{(\{j,k\},\{i,j\})},$$
$$A_{j,k}^{(i,i)} = (-1)^{j+k-1_{\{j>i\}}-1_{\{k>i\}}} \det Q^{(\{i,k\},\{i,j\})}.$$

Let $L_l, 1 \leq l \leq d$ be the lines of the matrix Q with columns i, j deleted. We remark that moving the jth line of the matrix $Q^{(\{i,k\},\{i,j\})}$ to the ith row gives a matrix $\tilde{Q}^{(\{i,k\},\{i,j\})}$ such that

$$\det \tilde{Q}^{(\{i,k\},\{i,j\})} = (-1)^{i+j-1-1_{\{i<k<j\}}-1_{\{j<k<i\}}} \det Q^{(i,k;i,j)}$$

Moreover, each row of $\tilde{Q}^{(i,k;i,j)}$ is equal to the corresponding row of $Q^{(j,k;i,j)}$, except for row i: the ith line of $\tilde{Q}^{(i,k;i,j)}$ is L_j, whereas the ith line of $Q^{(j,k;i,j)}$ is L_i. We thus have:

$$\begin{aligned}\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k} &= (-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}}\mu_i \det Q^{(j,k;i,j)} \\ &\quad + (-1)^{j+k-1_{\{j>i\}}-1_{\{k>i\}}}(-1)^{i+j-1-1_{\{i<k<j\}}-1_{\{j<k<i\}}}\mu_j \det \tilde{Q}^{(i,k;i,j)}.\end{aligned}$$

It is easy to check that $(-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}} = (-1)^{j+k-1_{\{j>i\}}-1_{\{k>i\}}}(-1)^{i+j-1-1_{\{i<k<j\}}-1_{\{j<k<i\}}}$, and by multilinearity of the determinant, we obtain:

$$\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k} = (-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}} \det R^{(i,j,k)},$$

where row i of $R^{(i,j,k)}$ is $\mu_i L_i + \mu_j L_j$ and row l of $R^{(i,j,k)}$ is L_l for each $l \neq i,j,k$. Since $\mu Q = 0$, we obtain $\mu_i L_i + \mu_j L_j = -\sum_{l\neq i,j}\mu_l L_l$. Using the multilinearity of the determinant again:

$$\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k} = -(-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}} \sum_{l\neq i,j} \mu_l R^{(i,j,k,l)},$$

where row i of $R^{i,j,k,l}$ is L_l, and row m of $R^{i,j,k,l}$ is L_m for all $m \neq i,j,k$. For all $l \neq i,j,k$, we have $\det R^{(i,j,k,l)} = 0$ as its lth row is also L_l. Thus

$$\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k} = -(-1)^{i+k-1_{\{i>j\}}-1_{\{k>j\}}} \mu_k R^{(i,j,k,k)}.$$

Moving line i of $R^{(i,j,k,k)}$ (containing L_k) to row k gives the matrix $Q^{(i,j;i,j)}$, and an easy sign analysis gives the result.

3. Let $1 \leq i \neq j \leq d$. We have, by definition,

$$\begin{aligned}C^{i,j} + C^{j,i} &= \left(\left(Q^{(j)}\right)^{-1} c^{(j)}\right)_i + \left(\left(Q^{(i)}\right)^{-1} c^{(i)}\right)_j \\ &= \frac{1}{\mu_j}\left(A^{(j,j)}c^{(j)}\right)_i + \frac{1}{\mu_i}\left(A^{(i,i)}c^{(i)}\right)_j \\ &= \frac{1}{\mu_j}\sum_{k\neq j} A^{(j,j)}_{i,k} c_k + \frac{1}{\mu_i}\sum_{k\neq i} A^{(i,i)}_{j,k} c_k \\ &= \frac{1}{\mu_j} A^{(j,j)}_{i,i} c_i + \frac{1}{\mu_i} A^{(i,i)}_{j,j} c_j + \sum_{k\neq i,j} \frac{\mu_i A^{(j,j)}_{i,k} + \mu_j A^{(i,i)}_{j,k}}{\mu_i\mu_j} c_k.\end{aligned}$$

Using the previous point and the fact that $A^{(j,j)}_{i,i} = A^{(i,i)}_{j,j}$, we get

$$\begin{aligned}C^{i,j} + C^{j,i} &= A^{(j,j)}_{i,i}\left(\frac{\mu_i c_i + \mu_j c_j}{\mu_i\mu_j} + \sum_{k\neq i,j}\frac{\mu_k c_k}{\mu_i\mu_j}\right) \\ &= \frac{A^{(j,j)}_{i,i}}{\mu_i\mu_j}\mu c.\end{aligned}$$

□

Bibliography

[Ach+13] Yves Achdou, Guy Barles, Hitoshi Ishii, and Grigorii Lazarevich Litvinov. *Hamilton-Jacobi equations: approximations, numerical analysis and applications*. Vol. 10. Springer, 2013.

[AJ17] Anna Aksamit and Monique Jeanblanc. *Enlargement of filtration with finance in view*. Springer, 2017.

[Ame00] Jürgen Amendinger. "Martingale representation theorems for initially enlarged filtrations". In: *Stochastic Processes and their Applications* 89.1 (2000), pp. 101–116.

[BBC16] Bruno Bouchard, Géraldine Bouveret, and Jean-François Chassagneux. "A backward dual representation for the quantile hedging of Bermudan options". In: *SIAM Journal on Financial Mathematics* 7.1 (2016), pp. 215–235.

[BC17] Géraldine Bouveret and Jean-François Chassagneux. "A comparison principle for PDEs arising in approximate hedging problems: application to Bermudan options". In: *Applied Mathematics & Optimization* (2017), pp. 1–23.

[BCR19] Cyril Bénézet, Jean-François Chassagneux, and Christoph Reisinger. "A numerical scheme for the quantile hedging problem". In: *arXiv preprint arXiv:1902.11228* (2019).

[BDR95] Guy Barles, Ch Daher, and Marc Romano. "Convergence of numerical schemes for parabolic equations arising in finance theory". In: *Mathematical Models and Methods in Applied Sciences* 5.1 (1995), pp. 125–143.

[Bén+19] Cyril Bénézet, Jérémie Bonnefoy, Jean-François Chassagneux, Shuoqing Deng, Camilo Garcia Trillos, and Lionel Lenôtre. "A sparse grid approach to balance sheet risk measurement". In: *ESAIM: Proceedings and Surveys* 65 (2019), pp. 236–265.

[BER15] Bruno Bouchard, Romuald Elie, and Antony Réveillac. "BSDEs with weak terminal condition". In: *The Annals of Probability* 43.2 (2015), pp. 572–604.

[BET09] Bruno Bouchard, Romuald Elie, and Nizar Touzi. "Stochastic target problems with controlled loss". In: *SIAM Journal on Control and Optimization* 48.5 (2009), pp. 3123–3150.

[BG04] Hans-Joachim Bungartz and Michael Griebel. “Sparse Grids”. In: *Acta Numerica* 13.May (2004), p. 147. ISSN: 0962-4929. DOI: 10.1017/S0962492904000182.

[Bia15] Philippe Biane. “Polynomials associated with finite Markov chains”. In: *In Memoriam Marc Yor-Séminaire de Probabilités XLVII*. Springer, 2015, pp. 249–262.

[BJ07] Guy Barles and Espen Jakobsen. “Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations”. In: *Mathematics of Computation* 76.260 (2007), pp. 1861–1893.

[BM07] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

[BNV12] Bruno Bouchard and Thanh Nam Vu. “A stochastic target approach for P&L matching problems”. In: *Mathematics of Operations Research* 37.3 (2012), pp. 526–558.

[Bok+09] Olivier Bokanowski, Benjamin Bruder, Stefania Maroso, and Hasnaa Zidani. “Numerical approximation for a superreplication problem under gamma constraints”. In: *SIAM Journal on Numerical Analysis* 47.3 (2009), pp. 2289–2320.

[BS73] Fischer Black and Myron Scholes. “The pricing of options and corporate liabilities”. In: *Journal of political economy* 81.3 (1973), pp. 637–654.

[BS91] Guy Barles and Panagiotis E Souganidis. “Convergence of approximation schemes for fully nonlinear second order equations”. In: *Asymptotic Analysis* 4.3 (1991), pp. 271–283.

[Car08] René Carmona. *Indifference pricing: theory and applications*. Princeton University Press, 2008.

[Car+15] Bob Carpenter, Matthew D Hoffman, Marcus Brubaker, Daniel Lee, Peter Li, and Michael Betancourt. “The stan math library: Reverse-mode automatic differentiation in C++”. In: *arXiv preprint arXiv:1509.07164* (2015).

[CEK12] Jean-François Chassagneux, Romuald Elie, and Idris Kharroubi. “Discrete-time approximation of multidimensional BSDEs with oblique reflections”. In: *Ann. Appl. Probab.* 22.3 (2012), pp. 971–1007. ISSN: 1050-5164. DOI: 10.1214/11-AAP771. URL: http://dx.doi.org/10.1214/11-AAP771.

[CIL92] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. “User’s guide to viscosity solutions of second order partial differential equations”. In: *Bulletin of the American Mathematical Society* 27.1 (1992), pp. 1–67.

[CK93] Jaksa Cvitanic and Ioannis Karatzas. “Hedging Contingent Claims with Constrained Portfolios”. In: *Ann. Appl. Probab.* 3.3 (Aug. 1993), pp. 652–681. DOI: 10.1214/aoap/1177005357. URL: https://doi.org/10.1214/aoap/1177005357.

[CK96] Jakša Cvitanic and Ioannis Karatzas. “Backward stochastic differential equations with reflection and Dynkin games”. In: *The Annals of Probability* 24.4 (1996), pp. 2024–2056.

[CL84] Michael G. Crandall and Pierre-Louis Lions. “Two approximations of solutions of Hamilton-Jacobi equations”. In: *Mathematics of Computation* 43 (1984), p. 1.

[CR18] Jean-François Chassagneux and Adrien Richou. “Obliquely Reflected Backward Stochastic Differential Equations”. In: (2018). <hal-01761991>.

[DAFH17] Tiziano De Angelis, Giorgio Ferrari, and Saïd Hamadène. “A note on a new existence result for reflected BSDEs with interconnected obstacles”. In: *arXiv preprint arXiv:1710.02389* (2017).

[DHP09] Boualem Djehiche, Saïd Hamadène, and Alexandre Popier. “A finite horizon optimal multiple switching problem”. In: *SIAM Journal on Control and Optimization* 48.4 (2009), pp. 2751–2770.

[DP97] Richard W.R. Darling and Etienne Pardoux. “Backwards SDE with random terminal time and applications to semilinear elliptic PDE”. In: *The Annals of Probability* 25.3 (1997), pp. 1135–1159.

[DRZ18] Roxana Dumitrescu, Christoph Reisinger, and Yufei Zhang. “Approximation schemes for mixed optimal stopping and control problems with nonlinear expectations and jumps”. In: *arXiv preprint arXiv:1803.03794* (2018).

[DS94] Freddy Delbaen and Walter Schachermayer. “A general version of the fundamental theorem of asset pricing”. In: *Mathematische annalen* 300.1 (1994), pp. 463–520.

[Duf88] Darrell Duffie. “Security markets: Stochastic models”. In: (1988).

[Dum16] Roxana Dumitrescu. “BSDEs with nonlinear weak terminal condition”. In: *arXiv preprint arXiv:1602.00321* (2016).

[Dum+17] Roxana Dumitrescu, Romuald Elie, Wissal Sabbagh, and Chao Zhou. “BSDEs with weak reflections and partial hedging of American options”. In: *arXiv preprint arXiv:1708.05957* (2017).

[EK14] Romuald Elie and Idris Kharroubi. “BSDE representations for optimal switching problems with controlled volatility”. In: *Stochastics and Dynamics* 14.03 (2014), p. 1450003.

[EK+97] Nicole El Karoui, Christophe Kapoudjian, Etienne Pardoux, Shige Peng, and Marie-Claire Quenez. “Reflected solutions of backward SDEs, and related obstacle problems for PDEs”. In: *the Annals of Probability* (1997), pp. 702–737.

[EKPQ97] Nicole El Karoui, Shige Peng, and Marie-Claire Quenez. “Backward stochastic differential equations in finance”. In: *Mathematical finance* 7.1 (1997), pp. 1–71.

[EKQ95] Nicole El Karoui and Marie-Claire Quenez. “Dynamic programming and pricing of contingent claims in an incomplete market”. In: *SIAM journal on Control and Optimization* 33.1 (1995), pp. 29–66.

[FG15] Nicolas Fournier and Arnaud Guillin. “On the rate of convergence in Wasserstein distance of the empirical measure”. In: *Probability Theory and Related Fields* 162.3-4 (2015), pp. 707–738.

[FK97] Hans Föllmer and Dmitry Kramkov. "Optional decompositions under constraints". In: *Probability Theory and Related Fields* 109.1 (1997), pp. 1–25.

[FL99] Hans Föllmer and Peter Leukert. "Quantile hedging". In: *Finance and Stochastics* 3.3 (1999), pp. 251–273.

[Gev+18] Hugo Gevret, Nicolas Langrené, Jerome Lelong, Xavier Warin, and Aditya Maheshwari. *STochastic OPTimization library in C++*. Research Report. EDF Lab, May 2018. URL: https://hal.archives-ouvertes.fr/hal-01361291.

[GJ10] Michael B Gordy and Sandeep Juneja. "Nested simulation in portfolio risk measurement". In: *Management Science* 56.10 (2010), pp. 1833–1848.

[GP15] Emmanuel Gobet and Stefano Pagliarani. "Analytical approximations of BSDEs with nonsmooth driver". In: *SIAM Journal on Financial Mathematics* 6.1 (2015), pp. 919–958.

[GPP96] Anne Gegout Petit and Etienne Pardoux. "Equations différentielles stochastiques rétrogrades réfléchies dans un convexe". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 57.1-2 (1996), pp. 111–128.

[GRR15] Anouar M Gassous, Aurel Răşcanu, and Eduard Rotenstein. "Multivalued backward stochastic differential equations with oblique subgradients". In: *Stochastic Processes and their Applications* 125.8 (2015), pp. 3170–3195.

[HJ07] Saïd Hamadène and Monique Jeanblanc. "On the starting and stopping problem: application in reversible investments". In: *Mathematics of Operations Research* 32.1 (2007), pp. 182–192.

[HK79] J Michael Harrison and David M Kreps. "Martingales and arbitrage in multiperiod securities markets". In: *Journal of Economic theory* 20.3 (1979), pp. 381–408.

[HLP97] Saïd Hamadène, Jean-Pierre Lepeltier, and Shige Peng. "BSDEs with continuous coefficients and stochastic differential games". In: *Pitman Research Notes in Mathematics Series* (1997), pp. 115–128.

[HP81] J Michael Harrison and Stanley R Pliska. "Martingales and stochastic integrals in the theory of continuous trading". In: *Stochastic processes and their applications* 11.3 (1981), pp. 215–260.

[HT10] Ying Hu and Shanjian Tang. "Multi-dimensional BSDE with oblique reflection and optimal switching". In: *Probab. Theory Related Fields* 147.1-2 (2010), pp. 89–121. ISSN: 0178-8051. DOI: 10.1007/s00440-009-0202-1. URL: http://dx.doi.org/10.1007/s00440-009-0202-1.

[HZ10] Saïd Hamadène and Jianfeng Zhang. "Switching problem and related system of reflected backward SDEs". In: *Stochastic Processes and their applications* 120.4 (2010), pp. 403–426.

[JK05] Espen R. Jakobsen and Kenneth H. Karlsen. "Continuous dependence estimates for viscosity solutions of integro-PDEs". In: *Journal of Differential Equations* 212.2 (2005), pp. 278–318.

[JPR19] Espen R. Jakobsen, Athena Picarelli, and Christoph Reisinger. "Improved order 1/4 convergence for piecewise constant policy approximation of stochastic control problems". In: *arXiv preprint arXiv:1901.01193* (2019).

[Kar89] Ioannis Karatzas. "Optimization problems in the theory of continuous trading". In: *SIAM Journal on Control and Optimization* 27.6 (1989), pp. 1221–1259.

[Kry00] Nikolaj V. Krylov. "On the rate of convergence of finite-difference approximations for Bellmans equations with variable coefficients". In: *Probability Theory and Related Fields* 117.1 (2000), pp. 1–16.

[Kry99] Nikolaj V. Krylov. "Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies". In: *Electronic Journal of Probability* 4.2 (1999), pp. 1–19.

[KS81] John G Kemeny and James Laurie Snell. *Finite Markov Chains: With a New Appendix" Generalization of a Fundamental Matrix"*. Springer, 1981.

[Mer73] Robert C. Merton. "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1 (1973), pp. 141–183. ISSN: 00058556. URL: http://www.jstor.org/stable/3003143.

[Mor11] Ludovic Moreau. "Stochastic target problems with controlled loss in jump diffusion models". In: *SIAM Journal on Control and Optimization* 49.6 (2011), pp. 2577–2607.

[MS90] Robert C Merton and Paul Anthony Samuelson. "Continuous-time finance". In: (1990).

[Par97] E Pardoux. "Generalized discontinuous backward stochastic differential equations". In: *Pitman Research Notes in Mathematics Series* (1997), pp. 207–219.

[Pic13] Alois Pichler. "Evaluations of risk measures for different probability measures". In: *SIAM Journal on Optimization* 23.1 (2013), pp. 530–551.

[Pro56] Yu V Prokhorov. "Convergence of random processes and limit theorems in probability theory". In: *Theory of Probability & Its Applications* 1.2 (1956), pp. 157–214.

[Ram02] Sundareswaran Ramasubramanian. "Reflected backward stochastic differential equations in an orthant". In: *Proceedings Mathematical Sciences* 112.2 (2002), pp. 347–360.

[Rei18] Christoph Reisinger. "The non-locality of Markov chain approximations to two-dimensional diffusions". In: *Mathematics and Computers in Simulation* 143 (2018), pp. 176–185.

[RF16] Christoph Reisinger and Peter A. Forsyth. "Piecewise constant policy approximations to Hamilton-Jacobi-Bellman equations". In: *Applied Numerical Mathematics* 103 (2016), pp. 27–47.

[ST02a] H Mete Soner and Nizar Touzi. "Dynamic programming for stochastic target problems and geometric flows". In: *Journal of the European Mathematical Society* 4.3 (2002), pp. 201–236.

[ST02b] H Mete Soner and Nizar Touzi. "Stochastic target problems, dynamic programming, and viscosity solutions". In: *SIAM Journal on Control and Optimization* 41.2 (2002), pp. 404–424.

[TL94] Shanjian Tang and Xunjing Li. "Necessary conditions for optimal control of stochastic systems with random jumps". In: *SIAM Journal on Control and Optimization* 32.5 (1994), pp. 1447–1475.

[War16] Xavier Warin. "Some non-monotone schemes for time dependent Hamilton–Jacobi–Bellman equations in stochastic control". In: *Journal of Scientific Computing* 66.3 (2016), pp. 1122–1147.

Titre : Étude de méthodes numériques pour la couverture partielle et problèmes de switching avec incertitude sur les coûts

Résumé : Nous apportons dans cette thèse quelques contributions à l'étude théorique et numérique de certains problèmes de contrôle stochastique, ainsi que leurs applications aux mathématiques financières et à la gestion des risques financiers. Ces applications portent sur des problématiques de valorisation et de couverture faibles de produits financiers, ainsi que sur des problématiques réglementaires. Nous proposons des méthodes numériques afin de calculer efficacement ces quantités pour lesquelles il n'existe pas de formule explicite. Enfin, nous étudions les équations différentielles stochastiques rétrogrades liées à de nouveaux problèmes de switching, avec incertitude sur les coûts.

Mots clefs : Couverture partielle, EDSRs, EDPs non-linéaires, Solutions de viscosité, Schémas de différences finies monotones, Grilles Sparses, Mesures de risque, Contrôle optimal stochastique, Switching optimal, Grossissement de Filtration, EDSRs Obliquement Réfléchies

Title : Study of numerical methods for partial hedging and switching problems with costs uncertainty

Abstract : In this thesis, we give some contributions to the theoretical and numerical study to some stochastic optimal control problems, and their applications to financial mathematics and risk management. These applications are related to weak pricing and hedging of financial products and to regulation issues. We develop numerical methods in order to compute efficiently these quantities, when no closed formulae are available. We also study backward stochastic differential equations linked to some new switching problems, with costs uncertainty.

Keywords : Quantile Hedging, BSDEs, Non-Linear PDEs, Viscosity solutions, Monotone Finite Difference schemes, Sparse Grids, Risk Measures, Stochastic Optimal Control, Optimal Switching, Enlargement of Filtration, Obliquely Reflected BSDEs