Martingale Optimal Transport and Utility Maximization

Royer Guillaume

ECOLE POLYTECHNIQUE

Ecole Doctorale: Mathématiques et Informatique

THESE

pour obtenir le titre de

Docteur de l'Ecole Polytechnique

Spécialité : mathématiques appliquées

Guillaume Royer

Martingale Optimal Transport and Utility Maximization

Transport Optimal Martingale et Problèmes de Maximisation d'Utilité.

Directeur de thèse: Nizar Touzi

préparée au CMAP (Ecole Polytechnique).

Jury :

Directeur :	Nizar Touzi	-	Ecole Polytechnique
Rapporteurs :	Peter Tankov	-	Université Paris VII
	Josef Teichmann	-	ETH Zürich
Examinateurs :	Bruno Bouchard	-	Université Paris Dauphine
	Laurent Denis	-	Université du Maine
	Emmanuel Gobet	-	Ecole Polytechnique
	Pierre Henry-Labordère	-	Société Générale
	Huyen Pham	-	Université Paris VII

Remerciements

Je tiens en premier lieu à remercier mon directeur de thèse Nizar Touzi pour sa disponibilité et les conseils avisés qu'il m'a procuré tout au long de ces années de thèse. Sa rigueur, ses connaissances dans de très nombreux domaines mathématiques et ses idées ont été des facteurs déterminants dans l'achèvement de ma thèse. Son soutien, sa patience et ses encouragements m'ont permis de traverser les moments de doute et d'aller de l'avant. Avoir eu la chance de travailler à ses côtés m'a permis de faire des progrès considérables. Pour tout cela je l'en remercie profondément. Je souhaite aussi remercier chaleureusement Pierre Henry-Labordère pour nos discussions ainsi que pour ses idées très originales m'amenant parfois à m'intéresser à des notions mathématiques très éloignées de mon champ habituel. Merci pour sa disponibilité, et cela malgré son emploi du temps professionnel qui a pu être chargé.

Je voudrais ensuite adresser tous mes remerciements à Peter Tankov et Joseph Teichmann qui ont accepté de rapporter cette thèse. Je leur suis extrêmement reconnaissant du temps qu'ils m'ont consacré ainsi que de leur lecture et leurs remarques approfondies. Je tiens aussi à remercier Bruno Bouchard, Laurent Denis, Emmanuel Gobet et Huyên Pham d'avoir accepté de constituer mon jury de thèse.

Je tiens aussi à remercier très chaleureusement Olivier Guéant pour tous ses conseils depuis maintenant cinq ans, ainsi que pour m'avoir donné la possibilité, avec Nicolas Grandchamps des Raux, d'élargir mon champ de recherche à des thématiques de microstructure des marchés. Je leur en suis très reconnaissant.

Une partie de ce travail de thèse a été écrit en collaboration avec Dylan Possamaï et je le remercie pour toutes nos discussions (animées parfois !) que nous avons pu avoir jusque tard dans la nuit souvent. Elles m'ont permis de faire progresser mes réflexions et nos articles. Je remercie de même Emilie Fabre et Jiang Pu pour leur collaboration dans plusieurs de mes articles produits durant cette thèse.

Le temps passé au Centre de Mathématiques Appliquées à l'école Polytechnique a été un lieu idéal pour le travail et le partage des connaissances. J'ai une pensée particulière pour Alexander Cox, Zhen-Jie Rhen et Xiaolu Tan pour nos conversations mathématiques. Je remercie l'ensemble des membres du laboratoire, en particulier Anis, Chao, Gaoyue, Nabil, Romain, Stefano et Thibaut. L'équipe administrative mérite aussi amplement mes remerciements, en particulier Nassera, Nathalie et Sandra. Je pense aussi à tous ceux que j'ai pu rencontrer pendant ma thèse, notamment lors de conférences lointaines, Anthony, Sébastien, Romuald, Ludovic, Sigrid, Marcel, Roxana, Stephen, Thibaut et Thomas.

Plus personnellement, je tiens à remercier comme il se doit mes camarades de l'ENS pour toutes nos très nombreuses conversations, parfois mathématiques, Julien, Franck, Thibaut et Sylvain, parmi lesquels je serais le premier à lancer ce bal des soutenances! Une petite pensée aussi pour jcc qui le mérite amplement!

Je tiens enfin à remercier mes parents ainsi que mes beaux-parents parents pour leur soutien permanent et leur présence pendant toutes ces annés d'études.

Mes derniers mots sont pour ma compagne Lucienne. Ta présence indispensable, ton soutien sans faille pour tous mes choix, ta gentillesse et ta compréhension dans mes moments de doutes ne seront jamais suffisament remerciés et j'espère pouvoir dans le futur t'apporter à mon tour ces mêmes choses.

Martingale Optimal Transport and Utility Maximization

Abstract: This PhD dissertation presents two independent research topics dealing with contemporary issues from financial mathematics, the second one being composed of two distinct problems.

In the first part we study the question of martingale optimal transport, which comes from the questions of no-arbitrage optimal bounds of liabilities. We first consider the question in discret time of the existence of a martingale law with given marginals. This result was first proved by Strassen [94] and is the starting point of martingale optimal transport. We provide a new proof of this theorem based on utility maximization technics, adapted from a proof of the fundamental theorem of asset pricing by Rogers [83].

We then consider the question of martingale optimal transport in continuous time, introduced in the framework of lookback options by Galichon, Henry-Labordère et Touzi [46]. We first establish a partial duality result concerning the robust superhedging of any contingent claim. For that purpose, we adapt recent technics developed by Neufeld and Nutz [74] in the context of martingale optimal transport. In a second time we study a robust utility maximization of a contingent claim with exponential utility in the context of martingale optimal transport, and we deduce its robust utility indifference price, given that the underlying's dynamic has a constant and well-known sharpe ratio. We prove that this robust utility indifference price is equal to the robust superhedging price.

The second part of this disseration considers first the problem of optimal liquidation of an indivisible asset. We study the advantage that an agent can take from having a dynamic trading strategy in an orthogonal asset. The question of its influence on the optimal liquidation rule is asked. We then provide examples illustrating our results.

The last chapter of this thesis concerns the utility indifference price of a European option in the context of small transaction costs. We use technics developed by Soner and Touzi [90] to obtain an asymptotic expansion of the Merton value functions with and without the option. These expansions are obtained by using homogenization technics. We formally obtain a system of equations verified by the values involved in the expansion and show rigorously that they are solutions. We then deduce an asymptotic expansion of the utility indifference price.

Keywords: Martingale optimal transport, utility maximization, robust superhedging, quasi-sure stochastic analysis, robust utility maximisation, robust indifference pricing, optimal stopping, optimal control, transaction costs, asymptotic expansions, viscosity solutions, homogenization.

Transport Optimal Martingale et Problèmes de Maximisation d'Utilité

Résumé: Cette thèse présente deux principaux sujets de recherche indépendants, le dernier regroupant deux problématiques distinctes.

Dans la première partie nous nous intéressons au problème du transport optimal martingale, dont le but premier est de trouver des bornes de non-arbitrage pour des options quelconques. Nous nous intéressons tout d'abord à la question en temps discret de l'existence d'une loi de probabilité sous laquelle le processus canonique est martingale, ayant deux lois marginales fixées. Ce résultat dû à Strassen [94] est le point de départ pour le problème primal de transport optimal martingale. Nous en donnons une preuve basée sur des techniques financières de maximisation d'utilité, en adaptant une méthode développée par Rogers pour prouver le théorème fondamental d'évaluation d'actif [83]. Ces techniques correspondent à une version en temps discrétisé du transport optimal martingale.

Nous considérons ensuite le problème de transport optimal martingale en temps continu introduit dans le cadre des options lookback par Galichon, Henry-Labordère et Touzi [46]. Nous commencons par établir un résultat de dualité partiel concernant la surcouverture robuste d'une option quelconque. Pour cela nous adaptons au transport optimal martingale des travaux récents de Neufeld et Nutz [74]. Nous étudions ensuite le problème de maximisation d'utilité robuste d'une option quelconque avec fonction d'utilité exponentielle dans le cadre du transport optimal martingale, et en déduisons le prix d'indifférence d'utilité robuste, sous une dynamique où le ratio de sharpe est constant et connu. Nous prouvons en particulier que ce prix d'indifférence d'utilité robuste est égal au prix de surcouverture robuste.

La deuxième partie de cette thèse traite tout d'abord d'un problème de liquidation optimale d'un actif indivisible. Nous étudions la profitabilité de l'ajout d'une stratégie d'achat et de vente d'un actif orthogonal au premier sur la stratégie de liquidation optimale de l'actif indivisible. Nous fournissons ensuite quelques exemples illustratifs.

Le dernier chapitre de cette thèse concerne le problème du prix d'indifférence d'utilité d'une option européenne en présence de petits coûts de transaction. Nous nous inspirons des travaux récents de Soner et Touzi [90] pour obtenir un développement asymptotique des fonctions valeurs des problèmes de Merton avec et sans l'option. Ces développements sont obtenus en utilisant des techniques d'homogénisation. Nous obtenons formellement un système d'équations vérifiées par les composantes du problème et nous vérifions que celles-ci en sont bien solution. Nous en déduisons enfin un développement asymptotique du prix d'indifférence d'utilité souhaité.

Mots-clés: transport optimal martingale, maximisation d'utilité, sur-couverture robuste, analyse stochastique quasi-sûre, maximisation d'utilité robuste, prix d'indifférence d'utilité robuste, arrêt optimal, contrôle optimal, coûts de transaction, développements asymptotiques, solution de viscosité, homogénisation.

Contents

Chapter 1

Introduction

Cette thèse s'articule en deux thématiques indépendantes. Le premier axe traite du transport optimal martingale étudié récemment pour ses nombreuses applications financières, notamment pour obtenir des bornes de non-arbitrage pour le prix des options exotiques. Le deuxième axe regroupe deux chapitres indépendants qui présentent des développements autour du problème de Merton.

La première partie est consacrée à trois contributions au transport optimal martingale. Nous considérons en premier lieu une question de transport optimal martingale en temps discret, à savoir le théoreme de Strassen. Nous en donnons une nouvelle preuve en utilisant des arguments de maximisation d'utilité. Nous nous attardons ensuite sur le transport optimal martingale en temps continu et présentons deux contributions. La première est un théorème de dualité adapté au cadre du transport optimal martingale, résultat technique important pour les applications. La deuxième est l'étude d'un problème de maximisation d'utilité, suivi naturellement de la question du prix d'indifférence d'utilité, d'un portefeuille d'actifs contenant nécessairement une option de type exotique, par des méthodes de type analyse quasi-sûre inspirées du chapitre précédent.

La deuxième partie de cette thèse analyse deux développements autour du problème de gestion de portefeuille de Merton. On s'intéresse tout d'abord à l'intéret que peut avoir un agent souhaitant liquider optimalement un actif indivisible à avoir une stratégie d'achat et de vente d'un actif indépendant du premier, et son influence sur la stratégie de liquidation optimale. Un deuxième sujet concerne la question du prix d'indifférence d'utilité d'une option européenne en présence de petits coûts de transaction. On en détermine un développement asymptotique en utilisant des techniques d'homogénisation.

1.1 Quelques rappels sur le problème de Merton

Le désormais très classique problème de Merton correspond à la formulation première de la question générale de maximisation d'utilité. Un agent a la possibilité d'investir son capital en le répartissant entre plusieurs actifs :un actif sans risque noté S^0 et d actifs risqués $(S^1, ..., S^d)$. Cette stratégie d'investissement π, doublée de sa consommation c, conduit à une valeur de portefeuille (valant X_0 en 0) égale à $X_T^{c,\pi}$ en T.

L'investisseur cherche à résoudre le problème:

$$\sup_{(c,\pi)\in\mathcal{A}} \mathbb{E}\left[\int_0^T e^{-\beta s} U_1(c_s)ds + e^{-\beta T} U_2\left(X_T^{c,\pi}\right)\right].$$

Quelques explications sont nécessaires ici. Tout d'abord U_1 et U_2 sont deux fonctions croissantes et concaves représentant l'utilité correspondante de l'agent relative à sa consommation instantanée en ce qui concerne U_1 et relative à sa richesse finale pour U_2. Les propriétés de ces fonctions d'utilités reflètent pour la croissance le fait que le bonheur augmente avec la consommation, et pour la concavité la propriété de décroissance de l'intéret marginal d'obtenir (ou de consommer) un peu plus d'argent. Le coefficient β représente ici une préférence pour le présent et ne doit pas être confondu avec le taux d'intéret sans risque. Enfin l'ensemble $\mathcal{A}$ est l'ensemble décrivant les possibilités admissible pour le choix des stratégies.

Ce problème très général a été introduit dans le cas $U_1 = 0$, $\beta = 0$, une diffusion de type brownien géométrique pour le prix des actifs risqués, et une utilité finale U_2 de type puissance par Merton dans [71], puis généralisé par Pliska dans [79]. De nombreux développements autour de cette thématique sont apparus alors. L'ajout de coûts de transaction entre les différents actifs, qui nous concerne dans cette thèse, a fait le fruit d'une littérature très importante, débutant par les travaux de Magill et Constantinides [31]. Nous détaillerons ce cas plus tard.

L'utilisation de l'évaluation par indifférence d'utilité s'est généralisée principalement à cause des difficultés rencontrés pour surcouvrir les options les plus simples (Calls) en présence de coûts de transaction. En effet, Davis et Clark [34] ont d'abord conjecturé que la stratégie la moins cher de surréplication du call était le "buy-and-hold" (on achète l'action au départ et on la garde jusqu'a expiration de l'option), conduisant à des coûts de surcouverture prohibitif. En effet il n'y a plus aucun intérêt à payer le prix de l'action, pour être sûr de toucher moins en T... Ce résultat a ensuite été prouvé rigoureusement simultanément par Soner, Shreve et Cvitanic [88] et Levental et Skorohod [68].
Le prix d'indifférence d'utilité correspond à la valeur qu'un agent qui souhaite acheter une option doit payer à un deuxième, pour que ce dernier soit indifférent en terme d'utilité entre avoir l'option assortie de ce montant et ne rien avoir. Cette approche a été initiée par Hodges et Neuberger [58].

1.2 Transport optimal martingale, du discret au continu

1.2.1 Du transport optimal classique au transport optimal martingale

1.2.1.1 le problème de Monge/Monge-Kantorovich

Nous commencons par une brève introduction au transport optimal classique qui a donné lieu à une littérature prolifique. Nous renvoyons le lecteur aux notes de cours d'Ambrosio [4], de Carlier [28], où aux deux livres de Villani [99, 100] pour une présentation complète, ainsi que des applications.

Le transport optimal sous sa forme primitive a été introduit par Gaspard Monge en 1781 sous la forme du problème de "remblais et déblais". On considère un tas de sable et un trou de même volume que le tas de sable. Un ouvrier, qui souhaite rentrer chez lui au plus tôt, cherche le moyen le moins fatiguant pour remplir le trou avec le sable qui se trouve à coté. Modélisons le tas de sable par une région A de l'espace , et B le trou qu'il doit remplir. Transporter du sable du point x au point y lui coûte un effort $c(x, y)$. Il va donc chercher à minimiser la quantité

$$\int_A c(x, T(x))dx,$$

où sa "stratégie" de déplacement T se doit de préserver le volume.

De manière plus rigoureuse à présent, on considère $\mathcal{X}$ et $\mathcal{Y}$ deux espaces polonais munis de leurs tribus boréliennes respectives, notées par la suite $\mathcal{B}(\mathcal{X})$ et $\mathcal{B}(\mathcal{Y})$. Considérons une mesure μ sur $\mathcal{X}$, et une mesure ν sur $\mathcal{Y}$. On suppose ici (et dans tout ce qui suivra) que ces mesures sont des lois de probabilités. Nous considérons pour toute application mesurable $T : \mathcal{X} \to \mathcal{Y}$ la mesure image de μ par T sur $\mathcal{Y}$ notée $T_\# \mu$ (c'est à dire la mesure telle que pour tout borélien B de $\mathcal{Y}$, on a $T_{\#}\mu(B) := \mu\left(T^{-1}(B)\right)$). Une telle application mesurable T est appelée plan de transport. Le problème de Monge consiste à trouver le coût minimal de transport entre μ et ν

$$\mathbf{P}^M := \inf_{T \ s.t. \ T_{\#}\mu=\nu} \int_{\mathcal{X}} c(x, T(y))\mu(dx). \tag{1.2.1}$$

Une relaxation de ce problème consiste en la formulation de Kantorovich. Avec les notations précédentes, on introduit maintenant l'ensemble $\mathcal{P}(\mu,\nu)$ des mesures ayant pour premières et deuxièmes lois marginales respectivement μ et ν. Ainsi on cherche une mesures γ sur $\mathcal{X} \times \mathcal{Y}$ minimisant le problème de coût. On s'intéresse alors au problème:

$$\mathbf{P}^{MK} := \inf_{\gamma \in \mathcal{P}(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\gamma(dx,dy) = \inf_{\gamma\in\mathcal{P}(\mu,\nu)} \mathbb{E}^{\gamma}\left[c(X,Y)\right]. \tag{1.2.2}$$

Comparons maintenant les deux formulations (1.2.1) et (1.2.2). On constate tout d'abord qu'en considérant un plan de transport T, on obtient un élément de $\mathcal{P}(\mu,\nu)$ défini par $\gamma^T(dx,dy) := \mu(dx) \times \delta_{T(x)}(dy)$. Ainsi clairement l'inégalité $\mathbf{P}^{MK} \leq \mathbf{P}^M$ est toujours vraie. L'autre inégalité en revanche fait défaut dans un certain nombre de cas. Considérons le cas extrême suivant: c est une fonction de coût strictement positive (fonction puissance sur $\mathbb{R}_+^*$), μ est la mesure de dirac en 1, et ν est la mesure uniforme sur l'intervale $[0.5, 1.5]$. Alors on voit clairement qu'il n'existe pas de plan de transport conservant la masse entre ces deux mesures. Ainsi $\mathbf{P}^M = \infty$. Or $\mathcal{P}(\mu,\nu)$ est non vide car il contient l'élément $\mu \otimes \nu$, défini par:

$$\forall A \in \mathcal{B}(\mathcal{X}) \text{ et } B \in \mathcal{B}(\mathcal{X}),\ \mu\otimes\nu(A\times B) = \mu(A)\cdot\nu(B).$$

La quantité $\mathbf{P}^{MK}$ est donc nécessairement finie et les deux problèmes (1.2.1) et (1.2.2) ne sont donc pas équivalent.

Dans un registre moins extrême que cet exemple, nous pouvons aussi faire face au problème suivant: on a équivalence entre les deux problèmes, c'est à dire $\mathbf{P}^M = \mathbf{P}^{MK}$, avec existence d'une solution pour $\mathbf{P}^{MK}$, mais pas pour $\mathbf{P}^M$. En effet le problème de Monge-Kantorovich est aussi plus aisé dans sa résolution, dans le sens où l'ensemble $\mathcal{P}(\mu,\nu)$ est faiblement compact. Ainsi dans le cas où la fonction de coût c est continue, la continuité de l'application

$$f:\ \mathcal{P}(\mu,\nu) \to \mathbb{R}:\ \gamma \mapsto \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\gamma(dx,dy),$$

devient aisée (sous des hypothèses d'intégrabilité de c). On obtient alors facilement le résultat suivant:

Theorem 1.2.1. *Supposons que c soit continue, alors il existe des solutions au problème de Monge-Kantorovich $\boldsymbol{P}^{MK}$.*

Nous renvoyons le lecteur au théorème 5.10 du livre de Villani [100] pour une revue des conditions que l'on peut mettre sur c pour obtenir ce résultat.

Le résultat central, qui fera l'objet de la version martingale du transport optimal, est la formulation duale de Kantorovich du problème de Monge/Kantorovich (1.2.2). Nous commencons par une caractérisation simple de l'ensemble $\mathcal{P}(\mu,\nu)$.

Pour deux éléments $\phi \in \mathbb{L}^1(\mu)$ et $\psi \in \mathbb{L}^1(\nu)$, on définit l'élément $\phi \oplus \psi$ par:

$$\forall (x,y) \in \mathcal{X}\times\mathcal{Y},\quad \phi\oplus\psi(x,y) = \phi(x) + \psi(y).$$

Si gamma est une mesure sur $\mathcal{X}\times\mathcal{Y}$, alors on a équivalence entre (i) et (ii) où:

(i) $\gamma \in \mathcal{P}(\mu, \nu)$,

(ii) $\forall \phi \in \mathbb{L}^1(\mu)$ et $\psi \in \mathbb{L}^1(\nu)$, $\int_{\mathcal{X}\times\mathcal{Y}} \phi \oplus \psi(x,y)\gamma(dx,dy) = \int_{\mathcal{X}} \phi(x)\mu(dx) + \int_{\mathcal{Y}} \psi(y)\nu(dy)$.

Nous pouvons par cette constatation introduire la formulation duale de Kantorovich, formulation linéaire contrairement à la formulation primale précédement introduite. Définissons tout d'abord l'ensemble $\mathcal{D}^K$ par:

$$\mathcal{D}^K := \big\{(\phi,\psi) \in \mathbb{L}^1(\mu) \times \mathbb{L}^1(\nu),\ \phi \oplus \psi \leq c\big\}.$$

Le dual de Kantorovich est alors défini par:

$$\mathbf{D}^K := \sup_{(\phi,\psi)\in\mathcal{D}^K} \mu(\phi) + \nu(\psi), \quad \text{où } \mu(\phi) := \int_{\mathcal{X}} \phi(x)\mu(dx) \text{ et } \nu(\psi) := \int_{\mathcal{Y}} \psi(y)\nu(dy). \tag{1.2.3}$$

Le bien-fondé de cette notion vient du souhait d'obtenir le résultat suivant:

Theorem 1.2.2 (Résultat souhaité)**.**

$$\boldsymbol{P}^{MK} = \boldsymbol{D}^K.$$

De plus il existe une solution au problème duale $\boldsymbol{D}^K$.

Ce résultat ne va pas être vrai dans le cas général. Néanmoins nous avons toujours l'inégalité $\mathbf{P}^{MK} \geq \mathbf{D}^K$. En effet, en prenant $\gamma \in \mathcal{P}(\mu,\nu)$ et $(\phi,\psi) \in \mathcal{D}^K$, nous avons aisémment:

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\gamma(dx,dy) \geq \int_{\mathcal{X}\times\mathcal{Y}} \phi \oplus \psi(x,y)\gamma(dx,dy) = \int_{\mathcal{X}} \phi(x)\mu(dx) + \int_{\mathcal{Y}} \psi(y)\nu(dy).$$

L'inégalité inverse sera obtenue sous certaines hypothèses sur la fonction de coût c, en ce qui concerne l'égalité entre problème primal et problème dual, et sur les espaces $\mathcal{X}$ et $\mathcal{Y}$ en ce qui concerne l'existence d'une solution optimale au problème dual. Nous renvoyons le lecteur au livre de Villani [100] pour la preuve d'un tel résultat.

1.2.1.2 Formulation en temps discret

A la problématique de Monge-Kantorovich précédente, nous ajoutons une contrainte supplémentaire. Nous imposons désormais que les probabilités admissibles pour le problème primal soient martingales dans un sens que nous allons expliciter.

On appelle désormais (X,Y) le processus canonique sur l'espace $\mathcal{X}\times\mathcal{Y}$. On dit que $\mathbb{P} \in \mathcal{P}(\mu,\nu)$ est martingale si on a:

$$\mathbb{E}^{\mathbb{P}}[Y|X] = X,\ \mathbb{P}-\text{p.s.}$$

On note alors:

$$\mathcal{M}(\mu,\nu) := \big\{\mathbb{P} \in \mathcal{P}(\mu,\nu),\ \mathbb{E}^{\mathbb{P}}[Y|X] = X,\ \mathbb{P}-\text{p.s.}\big\}.$$

La formulation primale du problème de transport sous contrainte martingale pour une fonction de coût c s'écrit donc naturellement:

$$\mathbf{P}^m := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}[c(X,Y)]. \tag{1.2.4}$$

Notons que nous adoptons maintenant la formulation de maximisation, et non de minimisation comme dans le transport classique. Cette nouvelle formule est plus adaptée aux applications financières que l'on va considérer par la suite.

En ce qui concerne la formulation duale, nous devons introduire un terme supplémentaire, comparé au problème dual (1.2.3), correspondant à la contrainte de martingale. Ainsi pour toute fonction $h : \mathbb{R}^d \to \mathbb{R}^d$, et $(\phi, \psi) \in \mathbb{L}^1(\mu) \times \mathbb{L}^1(\nu)$, nous définissons

$$\left(h^{\otimes} + \phi \oplus \psi\right)(x, y) := h(x)(y - x) + \phi(x + \psi(y).$$

Introduisons ensuite:

$$\mathcal{D}^m := \left\{(h, \phi, \psi) \in \mathcal{A},\ h^{\otimes} + \phi \oplus \psi \geq c,\ \mathcal{M}(\mu, \nu) - q.s.\right\},$$

où l'on dit que pour un ensemble de probabilités $\mathcal{Q}$, une assertion est vraie $\mathcal{Q}$ quasi-sûrement (notée $\mathcal{Q}$-q.s.) si elle est vraie $\mathbb{P}$-presque sûrement pour tout $\mathbb{P}$ dans $\mathcal{Q}$. L'ensemble $\mathcal{A}$ des triplets admissibles est un sous-ensemble de $\mathbb{L}^0(\mathbb{R})^3$ (nous ne détaillons pas ici les conditions que $\mathcal{A}$ doit vérifier, conditions qui pourront varier en fonction des problèmes considérés). Nous observons ici que cette définition dépend fortement de l'espace de lois de probabilités que l'on souhaite considérer. Nous pourrions obtenir aussi la formulation relaxée suivante: pour une famille de lois de probabilités $\mathcal{P}$ sur $\mathcal{X} \times \mathcal{Y}$ contenant l'ensemble $\mathcal{M}(\mu, \nu)$ (où un sous ensemble de $\mathcal{M}(\mu, \nu)$), nous définissons alors:

$$\mathcal{D}^m(\mathcal{P}) := \left\{(h, \phi, \psi) \in \mathcal{A},\ h^{\otimes} + \phi \oplus \psi \geq c,\ \mathcal{P} - q.s.\right\}. \tag{1.2.5}$$

Cette formulation sera celle privilégiée lorsque l'on introduira le transport optimal sous contrainte martingale en temps continu.

Le problème dual s'exprime donc:

$$\mathbf{D}^m := \inf_{(h,\phi,\psi) \in \mathcal{D}^m} \mu(\phi) + \nu(\psi). \tag{1.2.6}$$

Ainsi la question naturelle de l'égalité entre problème primal (1.2.4) et dual (1.2.6) se pose, similairement à l'équivalent de transport classique. L'inégalité triviale dans le cas classique, à savoir $\mathbf{P} \leq \mathbf{D}$, reste vraie ici (elle est inversée ici par rapport à celle présentée en section 1.2.1.1, dûe à notre formulation de maximisation). En effet en utilisant les mêmes arguments, on obtient facilement

$$\mathbf{P}^m \leq \mathbf{D}^m.$$

L'autre inégalité est beaucoup moins évidente. Un certain nombre de travaux récents s'intéressent à cette question. Citons notamment les papiers de Beiglböck, Henry-Labordère et Penkner [9], Beiglböck et Juillet [10] et Henry-Labordère et Touzi [52].

Une question fondamentale, qui n'est pas un problème pour le transport classique, est de savoir si l'ensemble $\mathcal{M}(\mu, \nu)$ est non vide. Remarquons tout d'abord qu'en considérant deux lois de probabilité sur $\mathbb{R}^d$ μ et ν, nous ne pouvons avoir $\mathcal{M}(\mu, \nu)$ et $\mathcal{M}(\nu, \mu)$ non vides simultanément que dans un cas particulier, lorsque $\mu = \nu$. Ce résultat sera une conséquence triviale de ce que l'on nommera par la suite l'ordre convexe pour les mesures (cf partie 1.2.2 et le chapitre 2). De même nous pouvons très bien avoir simultanément ces deux ensembles de lois de probabilités vide. Tout ceci sera détaillé en détail dans le chapitre 2.

Ensuite, toutes les questions naturelles que l'on se pose dans le cadre du transport classique peuvent apparaitre ici:

- Y a-t-il existence d'un maximiseur pour le problème primal, c'est à dire une probabilité $\mathbb{P}^* \in \mathcal{M}(\mu, \nu)$ telle que $\mathbf{P}^m = \mathbb{E}^{\mathbb{P}^*}\left[c(X, Y)\right]$?

- Y a-t-il existence d'un triplet minimiseur pour le problème dual, c'est à dire un triplet $(h^*, \phi^*, \psi^*) \in \mathcal{D}^m$ telle que $\mathbf{D}^m = \int_{\mathcal{X}} \phi(x)\mu(dx) + \int_{\mathcal{Y}} \psi(y)\nu(dy)$?
- Dans quels cas peut-on obtenir des formules explicites pour ces optimiseurs?

Remarquons que nous avons fait le choix ici d'introduire le transport martingale pour uniquement deux lois marginales. Bien sûr nous pourrions considérer le cas n-marginal en considérant le processus canonique $(X_i)_{1\le i\le n}$ et en imposant les contraintes μ_i comme lois marginales aux dates $i = 1, ..., n$.

1.2.1.3 Formulation en temps continu, sur l'espace des fonctions continues.

Le problème de transport sous contrainte martingale précédement introduit admet aussi naturellement une version en temps continu. En réalité il n'y a pas unicité de la formulation continue, nous allons expliquer cela.
Nous choisissons ici de ne considérer que des lois de probabilités sur l'espace des fonctions continues sur $[0, T]$ à valeurs dans $\mathbb{R}^d$ valant 0 en 0, noté Ω. Ainsi toute loi de probabilité considérée induira naturellement des trajectoires continues pour tout ω. Nous munissons Ω de la filtration naturelle $\mathbb{F}$ définie par $\mathcal{F}_t := \sigma\{\omega_s,\ s \le t\}$ pour tout $0 \le t \le T$, et le processus canonique associé $B_t(\omega) := \omega_t$ pour tout $0 \le t \le T$.

On se fixe une loi μ sur $\mathbb{R}^d$. Nous considérons désormais une famille générale de lois sur Ω notée $\mathcal{M}$ sous laquelle nous savons que le processus canonique est martingale pour tout $\mathbb{P}$ dans $\mathcal{M}$. Notons alors $\mathcal{M}(\mu)$ le sous ensemble de $\mathcal{M}$ composé des probabilités $\mathbb{P}$ ayant pour T-marginale μ, c'est à dire telle que:

$$\forall f \in C(\mathbb{R}^d, \mathbb{R}),\ \mathbb{E}^{\mathbb{P}}[f(B_T)] = \int_{\mathbb{R}^d} f(x)\mu(dx).$$

Considérons désormais une fonction mesurable $\xi : \Omega \to \mathbb{R}$. Le problème primal de transport s'écrit donc naturellement:

$$\mathbf{P} := \sup_{\mathbb{P}\in\mathcal{M}(\mu)} \mathbb{E}^{\mathbb{P}}[\xi]. \tag{1.2.7}$$

En ce qui concerne le problème dual, nous nous devons de définir proprement l'équivalent de (1.2.5). Ici nous n'avons considéré qu'une seule loi terminale, sachant qu'a l'instant initial nous savons déjà que la répartition est la fonction de dirac en 0. Tout le problème ici vient de la définition de la composante dynamique. Pour cela nous considérons que la composante dynamique $h \in \mathcal{H}$ est un processus adapté et telle que

$$(h \cdot B)_. := \int_0^{\cdot} h_s dB_s$$

est bien définie et qu'elle admet de plus une version pouvant être définie trajectoire par trajectoire, assurant ainsi que cette quantité nous donnera bien une valeur qui sera valable (et unique!) quelle que soit la probabilité $\mathbb{P} \in \mathcal{M}$. Ainsi d'importantes difficultés dites d'aggrégation apparaissent naturellement ici. Le problème du choix de $\mathcal{M}$ prend tout son sens ici. Le choix de l'espace $\mathcal{H}$ sera donc forcément en relation directe avec l'ensemble $\mathcal{M}$. Ces questions d'aggrégation sont principalement traitées par la définition d'une intégrale stochastique trajectoire par trajectoire. Les principales références sur le sujet sont tout d'abord le calcul d'Itô sans probabilités de Föllmer [45], et suivent les travaux de Karandikar [65], puis Denis et Martini [39], Nutz [76] et enfin l'approche de Dolinsky et Soner [40]. La formulation que nous utiliserons est basée sur l'analyse quasi-sûre venant des équations différentielles stochastiques rétrogrades du second ordre (2EDSRs) de Soner, Touzi et Zhang [92, 93].

Admettons désormais que sous $\mathcal{M}$, nous puissions définir une intégrale stochastique. Alors le problème dual devient donc:

$$\mathbf{D} := \inf\Big\{\int_{\mathbb{R}^d}\lambda(x)\mu(dx),\ \ \exists(h,\lambda)\ \ (h\cdot B)_T + \lambda(B_T) \geq \xi\ \ \mathcal{M}-q.s.\Big\}. \tag{1.2.8}$$

Comme première conséquence de cette formulation, nous insistons sur l'importance de la définition propre de l'espace $\mathcal{H}$ des stratégies h admissibles. Typiquement nous imposerons que l'intégrale stochastique $(h\cdot B)_.$ soit sur-martingale pour toute probabilité $\mathbb{P}\in\mathcal{M}$. Il est en effet illusoire d'espérer obtenir un résultat intéressant si l'on souhaite avoir la propriété de martingale pour tout $\mathbb{P}\in\mathcal{M}$, pour un ensemble $\mathcal{M}$ assez large (qui représente l'incertitude de modèle pour un marché financier). De même autoriser les martingales locales strictes poserait ici problème sachant qu'on peut sur-couvrir n'importe quelle fonction ζ avec une martingale locale, de manière triviale (la valeur $\mathbf{D}$ serait alors trivialement $-\infty$).

Une première question importante, correspondant au résultat de dualité obtenu pour les équations différentielles stochastiques rétrogrades du second ordre développées par Soner, Touzi et Zhang [91, 92, 93], concerne un premier résultat de dualité, simplement partiel dans le cadre du transport optimal. On cherche en effet à savoir si le problème ci-dessus (1.2.8) est égal au problème dual partiel:

$$\mathbf{D}^p := \inf_{\lambda}\sup_{\mathbb{P}\in\mathcal{M}}\mathbb{E}^{\mathbb{P}}\left[\xi - \lambda(B_T) + \mu(\lambda)\right]. \tag{1.2.9}$$

Le théorème de dualité souhaité dans le cadre du transport optimal martingale en temps continu est donc naturellement:

$$\mathbf{D} = \mathbf{D}^p = \mathbf{P}. \tag{1.2.10}$$

Notons que, comme dans le cadre discret, l'inégalité $\mathbf{D}\geq\mathbf{D}^p\geq\mathbf{P}$ sera ici triviale, sous réserve d'avoir proprement défini l'ensemble des stratégies admissibles.

Ce résultat général de dualité est obtenu lorsque ξ est de la forme $g(B_T, M_T)$ où $M_T := \sup_{s\leq T} B_s$ est le maximum de la trajectoire à la date T, avec des hypothèses de croissance sur g par Galichon, Henry-Labordère et Touzi [46], et Henry-Labordère, Obloj, Spoida et Touzi dans le cas multimarginal [51]. Ils utilisent des techniques de contrôle stochastique permettant de retrouver des résultats obtenus dans un cadre d'arrêt optimal, et ayant pour particularité d'avoir existence des optimiseurs, à savoir une probabilité $\mathbb{P}$ optimale pour le problème primal (1.2.7), et une stratégie λ optimale pour le problème dual partiel (1.2.9). L'égalité entre (1.2.8) et (1.2.9) est obtenue par des techniques d'analyse quasi-sûre.
Suivant ce résultat, Dolinsky et Soner [40] ont montré la validité du résultat de dualité (1.2.10), dans un cadre général incluant des fonctions de coûts ξ de la forme $g(B_T, M_T, \int_0^T B_s ds)$, sous des conditions de régularité sur g (lipschitz).

1.2.1.4 Interprétation financière

Cette formulation de transport optimal martingale correspond à la situation financière suivante. On considère un marché financier composé d'un actif risqué S et d'un actif sans risque. L'actif risqué est ici supposé avoir une dynamique martingale, et l'actif sans risque ne verse pas d'intérêts. On suppose de plus qu'à la date initiale $t_0 = 0$, pour toute valeur d'exercice K (ou strike), le prix des calls de maturités respectives t_1 (notée 1) et t_2 (notée 2) et de strike K est connu (où t_1 et t_2 sont deux dates précisés telles que $0 < t_1 < t_2$). Un agent possédant une richesse initiale X_0 est autorisé à prendre une position semi-statique, c'est à dire que sa position à la date t_2 sera composée:

- d'une composante statique, faite d'achats et de ventes de calls de maturités t_1 et t_2, achats qu'il aura effectué à la date 0,
- d'une composante dynamique correspondant au nombre d'actions qu'il souhaite détenir à chaque date.

Sa richesse X_2 à la date terminale sera donc donnée par:

$$X_2 = X_0 + h_0(S_0)(S_1 - S_0) + h_1(S_1)(S_2 - S_1) + \phi(S_1) - \mu(\phi) + \psi(S_2) - \nu(\psi),$$

où à la date initiale, sa position en calls de maturité 1 (resp 2) est notée ϕ (resp ψ) pour un coût $\mu(\phi)$ (resp $\nu(\psi)$). Il a de plus choisi à la date initiale d'acheter h_0 actions, et de réajuster sa position à la date 1 pour en obtenir h_1.

En temps continu introduit dans la partie 1.2.1.3, nous notons le processus canonique S. La composante statique est notée λ et la composante dynamique h est un processus adapté. La richesse X au temps T devient donc:

$$X_T = X_0 + \int_0^T h_s dS_s + \lambda(S_T) - \mu(\lambda).$$

Nous détaillons maintenant la significations des prix $\mu(\phi)$ et $\nu(\psi)$ (on se concentre ici sur le premier, le raisonnement étant le même pour $\nu(\psi)$). Cette réflexion est dûe à Breeden et Litzenberger[27]. On sait que les prix des calls de maturité 1 de tous strikes sont connus. Sous l'hypothèse supplémentaire de linéarité et de continuité de la fonctionnelle de prix (notée F par la suite) des options on peut alors déduire de manière unique le prix à la date initiale 0 de n'importe quelle option européenne à la date 1 (c'est à dire une option où l'on délivre une quantité d'argent dépendant uniquement de la valeur S_1). En effet considérons tout d'abord le cas d'un payoff d'option ϕ deux fois continument dérivable, alors la formule d'intégration par partie classique (connue en mathématiques financières comme étant la formule de Carr-Madan (voir Carr et Chou [29]) nous dit que:

$$\phi(S_1) = \phi(S_0) + (S_1 - S_0)\lambda'(S_0) + \int_{-\infty}^{S_0} (K - S_1)^+ \lambda''(K)dK + \int_{S_0}^{\infty} (S_1 - K)^+ g''(K)dK.$$

La linéarité nous dit donc que le prix de l'option de payoff λ vérifie donc:

$$F(\phi) = \phi(S_0) + \phi'(S_0)F(\cdot - S_0) + \int_{-\infty}^{S_0} p(K)\phi''(K)dK + \int_{S_0}^{\infty} c(K)\phi''(K)dK,$$

où $p(K)$ (resp $c(K)$) est le prix du put (resp du call) de strike K, c'est à dire $p(K) = F((K - \cdot)^+)$ (resp $c(K) = F((\cdot - K)^+)$). Le prix du put de strike K est déterminé de manière unique grâce à la parité call-put vérifiée ici (linéarité et non-arbitrage): $p(K) = c(K) + K - S_0$. Le prix $F(\cdot - S_0)$ est évidemment nul ici par non-arbitrage.

En remarquant ensuite que la fonction c (et donc p) est convexe, elle est alors deux fois dérivable au sens des mesures. On constate aussi que c et p ont même dérivée seconde. On note cette mesure μ et on obtient alors par intégration par partie que:

$$\int_{-\infty}^{S_0} p(K)\phi''(K)dK + \int_{S_0}^{\infty} c(K)\phi''(K)dK = \int_{\mathbb{R}} \phi(x)\mu(dx).$$

Ainsi le prix de l'option délivrant le payoff $g(S_1)$ est donné à la date initiale par la valeur $\mu(\phi)$, où μ est donnée par le prix des calls.

L'extension à un payoff quelconque (irrégulier) est immédiate ici grâce à la continuité de la fonctionnelle des prix F. En effet par densité des fonctions régulières dans l'espace des fonctions mesurables, on a immédiatement pour tout payoff ϕ mesurable (et intégrable) que:

$$F(\phi) = \mu(\phi).$$

Remark 1.2.1. *Cette interprétation n'est plus valable en pratique en dimension supérieure à 1. En effet l'hypothèse de la connaissance de toutes les options européennes sur une action est relativement raisonnable, par la liquidité des calls et des puts et le nombre de strikes traités, mais si on considère un panier de deux actions* (S^1, S^2)*, alors très peu (voire aucune dans la plupart des cas) d'options européennes prenant en compte ces deux valeurs simultanément sont cotées (où simplement traitées), rendant impossible de déterminer une quelconque loi jointe du couple* (S^1, S^2) *induite par le marché.*

1.2.2 Une nouvelle preuve du théorème de Strassen par maximisation d'utilité

1.2.2.1 Littérature existante

Le but de ce chapitre est de prouver le résultat suivant:

Theorem 1.2.3 (Théorème de Strassen)**.** *Soient* μ *et* ν *deux mesures de probabilité sur* $\mathbb{R}^d$ *telles que* $\int |x|\mu(dx) + \int |y|\nu(dy) < \infty$*, alors* $\mathcal{M}(\mu, \nu) \neq \varnothing$ *si et seulement si* $\mu \preceq \nu$ *pour l'ordre convexe, c'est à dire:*

$$\mu(g) \leq \nu(g) \text{ pour toute fonction convexe } g.$$

Ce résultat dû à Strassen fournit une condition nécessaire et suffisante sur les mesures μ et ν pour qu'il existe une probabilité martingale joignant les deux. C'est le point de départ pour toute considération sur le transport optimal martingale.

La preuve de Strassen dans [94] utilise des arguments de topologie générale. Il prouve ce résultat en construisant une topologie particulière sur les mesures sur $\mathbb{R}^d$ et en appliquant le théorème de Hahn-Banach. Notons aussi l'existence de preuves alternatives via la théorie de l'arrêt optimal mais se restreignant à la dimension 1 (voir Hobson & Pedersen [57]).

La version en temps continu de ce théorème (appelé le théorème de Kellerer [67]) a donné lieu à de nouvelles preuves adaptables de facto en temps discret (voir Hirsch et Roynette par exemple [55], ou Hirsch, Profeta, Roynette et Yor [53, 54]). Le théorème de Kellerer[67] soutient que pour qu'il existe un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P})$ tel que pour une famille $(\mu_t)_{0 \leq t \leq 1}$ de lois de probabilités sur $\mathbb{R}^d$, $\mathbb{P}$ ait pour t-marginale la mesure μ_t pour tout $0 \leq t \leq 1$ (c'est à dire $\forall$ $f \in C^0$, $t \in [0,1]$, $\mathbb{E}^{\mathbb{P}}[f(S_t)] = \mu_t(f)$), alors il faut et il suffit que:

(i) pour tout $0 \leq t \leq 1$, $\int |x|\mu_t(dx)$

(ii) pour toute fonction convexe g, l'application $t \mapsto \mu_t(g)$ est croissante sur $[0, 1]$.

Une telle famille (μ_t) est appelée peacock. Citons par exemple le résultat de Hirsch et Roynette [55] qui utilise une construction explicite à partir de la formule de la volatilité implicite de Dupire [41] dans le cas où la famille (μ_t) est continue dans un sens à déterminer, et utilisent un argument de passage à la limite pour retomber sur le cas général.

1.2.2.2 Résultats et contributions

Notre démarche ici s'inspire grandement d'une méthode d'obtention du théorème fondamental de l'évaluation par arbitrage, qui pour un processus adapté $(S_i)_{i\in\{0,1,..,N\}}$ établit le lien entre non-arbitrage et existence d'une probabilité martingale. Un arbitrage est défini comme un processus adapté θ (et admissible en un certain sens technique que l'on ne précise pas ici) tel que:

$$\sum_{i=0}^{N-1}\theta_i\cdot(S_{i+1}-S_i)\geq 0\ \mathbb{P}-p.s.\ \text{ et } \mathbb{P}\left[\sum_{i=0}^{N-1}\theta_i\cdot(S_{i+1}-S_i)>0\right]>0.$$

On note $\theta\cdot S_N:=\sum_{i=0}^{N-1}\theta_i\cdot(S_{i+1}-S_i)$.

Ce résultat a donné lieu à de nombreuses preuves dont notamment celle de Rogers [83] qui va nous intéresser ici. Il considère le problème de maximisation d'utilité sous la probabilité $\mathbb{P}$ suivant:

$$V(\mathbb{P}):=\sup_\theta \mathbb{E}^{\mathbb{P}}\left[-e^{-\theta\cdot S_N}\right].$$

Si $\mathbb{P}$ était une probabilité martingale pour le processus S, alors on aurait par l'inégalité de Jensen pour un sens, et considérant le processus $\theta=0$ pour l'autre, que $V(\mathbb{P})=-1$.

Dans le cas général, on va chercher une stratégie optimale θ^*, maximisant le problème $V(\mathbb{P})$. Si un tel maximiseur existe, alors la probabilité $d\mathbb{Q}:=ce^{-\theta^*\cdot S_N}\cdot d\mathbb{P}$, où $0<c<\infty$ est une constante de renormalisation, va être une probabilité martingale pour $\mathbb{P}$. En effet formellement nous avons pour tout θ admissible et ε réel,

$$J^\theta(\varepsilon):=c^{-1}\mathbb{E}^{\mathbb{Q}}\left[-e^{-\varepsilon\theta\cdot S_N}\right]=\mathbb{E}^{\mathbb{P}}\left[-e^{-(\theta^*+\varepsilon\theta)\cdot S_N}\right]\leq\mathbb{E}^{\mathbb{P}}\left[-e^{-\theta^*\cdot S_N}\right]=c^{-1}=J^\theta(0).$$

La condition du premier ordre sur J^θ nous indique donc que sa dérivée première (formellement toujours) en $\varepsilon=0$ est nulle, c'est à dire:

$$\mathbb{E}^{\mathbb{Q}}\left[\theta\cdot S_N\right]=0.$$

Ceci étant vrai pour tout processus admissible θ, cela nous donne la propriété de martingale du processus S sous $\mathbb{Q}$.

Notre approche est moralement la même en ce sens que l'on va considérer, en reprenant les notations de la section 1.2.1, le problème de maximisation d'utilité:

$$V^{\mu,\nu}:=\sup_{(h,\phi,\psi)}\mathbb{E}^{\mathbb{P}}\left[-e^{-\left(h^\otimes+\phi\oplus\psi\right)(X,Y)-\mu(\phi)-\nu(\psi)}\right].$$

Dans le cadre du théorème de Strassen, nous nous attendons alors à trouver un triplet optimal (h^*,ϕ^*,ψ^*), nous permettant de définir une probabilité

$$d\mathbb{Q}:=c^{-1}e^{-\left(h^{*\otimes}+\phi^*\oplus\psi^*\right)(X,Y)-\mu(\phi^*)-\nu(\psi^*)}\cdot d\mathbb{P}.$$

Les conditions du premier ordre sur respectivement h^*, ϕ^* et ψ^* devraient fournir les conditions respectives pour la probabilité $\mathbb{Q}$ de martingalité du processus canonique (c'est à dire $\mathbb{E}^{\mathbb{Q}}[Y|X]=X$ $\mathbb{Q}$-p.s. , loi marginale μ pour la variable X, et loi marginale ν pour la variable Y.
Ce raisonnement formel ne va pas pouvoir être développé rigoureusement. En effet le choix de la probabilité de départ $\mathbb{P}$ sous lequel faire notre maximisation d'utilité est complètement non trivial ici. Pour obtenir un triplet maximiseur, il faudrait donc qu'une probabilité $\mathbb{Q}$ dans $\mathcal{M}(\mu,\nu)$ ait une densité par rapport à la probabilité de départ $\mathbb{P}$. De plus il va être relativement aisé de

maximiser dans chaque direction, c'est à dire à (ϕ, ψ) donnés, trouver $h^*(\phi, \psi)$ optimal, où bien à (h, ψ) donnés, trouver $\phi^*(h, \psi)$ optimal etc... Mais la maximisation simultanée est bien plus complexe et ne permet pas d'obtenir des formules explicites pour les optimiseurs.
Pour palier à ces difficultés, notre approche de la résolution est complètement différente de celle adoptée par Rogers dans [83]. Nous nous concentrons dans un premier temps sur le cas où μ est à support fini, pour lequel on va construire une probabilité $\mathbb{P}$ sous laquelle établir la maximisation d'utilité. On montre que dans ce cas on va avoir que le supremum est strictement négatif, et en dégager l'existence d'un élément de $\mathcal{M}(\mu, \nu)$. On va ensuite passer au cas général μ quelconque en fabriquant une suite $(\mu_n)_{n\in\mathbb{N}}$ telle que $\mu_n \preceq \mu$ pour tout n et $\mu_n \to \mu$ pour la convergence faible. On déduit ensuite la non trivialité de $\mathcal{M}(\mu_n, \nu)$ pour tout n grâce à la première étape. Enfin on termine en considérant une suite d'éléments $\mathbb{Q}_n$ dans $\mathcal{M}(\mu_n, \nu)$ pour construire un élément $\mathbb{Q}$ dans $\mathcal{M}(\mu, \nu)$.

1.2.3 Surcouverture robuste d'options exotiques dans le cadre du transport optimal martingale

1.2.3.1 Littérature existante

Nous nous intéressons ici à démontrer un théorème de sur-couverture robuste adapté aux exigences du transport optimal martingale. Réutilisant les notations de la section 1.2.1.3, nous souhaitons avoir égalité entre (1.2.8) et (1.2.9) pour une classe aussi large que possible de fonctions ζ. Les problèmes techniques évoqués dans la section 1.2.1.3 prennent ici tout leur sens.
Cette dualité, appelée partielle auparavant, revient à l'étude de la surcouverture robuste avec incertitude de modèle. Elle a été introduite par Avellaneda, Levy et Paras [6] puis Lyons [69]. Les travaux de Denis et Martini [39] ont ensuite permis une formulation quasi-sûre du problème d'incertitude de modèle.

La formuation du problème de surcouverture robuste d'une option ζ, dans une classe de modèles $\mathcal{P}$ revient donc à faire le lien entre le problème primal

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\zeta], \tag{1.2.11}$$

et le problème dual:

$$\inf\left\{X_0 \ s.t. \ \exists H \in \mathcal{H} \ : \ X_0 + \int_0^T H_s dB_s \geq \zeta, \ \mathcal{P} - q.s.\right\}. \tag{1.2.12}$$

Ici contrairement à ce qui a été présenté en section 1.2.1.3, nous ne supposons pas nécessairement que les éléments de $\mathcal{P}$ soient des probabilités martingales.

Cette dernière formulation sous forme de cible stochastique fait apparaitre plus précisément les difficultés techniques auxquelles nous faisons face. Tout d'abord la question centrale d'aggrégation, c'est à dire de définition trajectorielle de l'intégrale stochastique. Cela revient à l'étude de l'ensemble des stratégies admissibles $\mathcal{H}$, mais aussi de l'ensemble $\mathcal{P}$. Plusieurs travaux sur ces questions d'aggrégations méritent d'être mentionnés, notamment Bichteler [15], Follmer [45], Karandikar [65] et Nutz [76].
Les résultats d'aggrégation que l'on utilisera par la suite sont ceux introduit par Soner et al dans le cadre des 2EDSRs. La question du choix de $\mathcal{P}$ est intimement liée à ces questions d'aggrégation et nous choisirons une formulation dite forte qui permettra entre autre d'avoir une propriété de représentation des martingales cruciales dans notre résolution.

Des résultats récents dûs à Nutz et Van Handel [77] et Neufeld et Nutz [74] fournissent une selection d'outils permettant d'avoir le résultat de dualité souhaité dans des conditions de régularités minimes sur ζ, à savoir de la semi-analyticité. Les résultats précédents ne permettaient d'obtenir cette dualité qu'en ayant des hypothèses très fortes, de type lipschitz. Dans ces deux articles, les auteurs définissent un ensemble de trois conditions sur $\mathcal{P}$ qui, réunies, assurent la dualité. La première condition est de type mesurabilité, la deuxième est une propriété d'invariance et la troisième une condition de stabilité assurant la sélection mesurable nécessaire à la programmation dynamique.

Récemment El Karoui et Tan [42, 43] ont obtenu le même type de résultats, pour le même degré de généralité sur la régularité de ζ, en utilisant la théorie des capacités.

1.2.3.2 Motivation et nouveaux Résultats

Dans le cadre du transport optimal et sa formulation en termes d'analyse quasi-sûre introduite par Galichon et al [46], la famille de probabilités considérée vérifie en particulier que le processus canonique doit être martingale sous toute probabilité dans la famille. Le souhait d'obtenir un principe de programmation dynamique pour le problème primal (1.2.11) nous pousse à introduire pour toute probabilité la notion de distribution de probabilité conditionnelle régulière (appelée par son anagramme anglais RCPD par la suite) introduite par Stroock et Varadhan [95]. La propriété de mesurabilité de la version dynamique de (1.2.11) va être ici vérifiée en utilisant les techniques introduites par Nutz et van Handel [77] et Neufeld et Nutz [74].

Malheureusement l'ensemble de lois de probabilié induit par cette RCPD ne va pas vérifier les conditions classiques de sélection mesurable nécessaire ici, nous le prouvons dans ce cas particulier. Cette difficulté technique nous fait perdre la propriété centrale de programmation dynamique telle qu'elle est obtenue dans [77] et [74]. Nous obtenons néanmoins une version plus faible de cette programmation dynamique pour toute probabilité: on obtient donc des propriétés presques sûres, avec des définition qui dépendent des probabilités alors que les résultats de [77] et [74] permettent d'avoir une formulation dépendant uniquement de la trajectoire. Les hypothèses d'intégrabilités de [74] sont ici relaxées pour coller aux besoin du transport optimal: on va requérir uniquement de l'uniforme intégrabilité sur ζ^+.
La programmation dynamique dans cette version probabilité par probabilité va néanmoins être suffisante pour dérouler les outils désormais classiques conduisant au résultat de surréplication. En effet la propriété nécessaire de surmartingale de la version dynamique de (1.2.11) est vérifiée, et nous obtenons par un théorème de décomposition des surmartingales et des techniques d'aggrégation une caractérisation explicite de la stratégie H de surréplication.

Nous déduisons de cette méthode une famille de conditions légèrement modifiée par rapport à celles introduites par Nutz et Van Handel [77] et Neufeld et Nutz [74], permettant une caractérisation de l'espace des lois de probabilités admissibles $\mathcal{P}$ adaptée à des contextes très larges, incluant donc naturellement le transport optimal martingale.

1.2.4 Prix d'indifférence d'utilité robuste d'options exotiques avec critère d'utilité exponentiel

1.2.4.1 Littérature existante

Nous nous intéressons désormais à la version robuste du problème de maximisation d'utilité présenté en section 1.1. Un agent économique cherche à optimiser son choix d'investissement, de sorte qu'il lui soit au plus profitable en moyenne dans le pire des scénarios qu'il pense pouvoir se produire. Formellement son incertitude sur le modèle se matérialise par un ensemble de lois de probabilités $\mathcal{P}$ et il cherche donc à optimiser à horizon fini T, ayant une option $-\xi$ en portefeuille et une stratégie dynamique de trading X^π, la quantité:

$$\sup_{\pi} \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^\pi - \xi\right)\right].$$

Cette question a donné lieu à une littérature très importante durant les 15 dernières années. Citons tout d'abord l'approche première qui consistait à considérer que l'ensemble $\mathcal{P}$ était constitué de lois de probabilité absolument continues par rapport à une probabilité de référence $\mathbb{P}_0$. Les techniques de dualités convexes sont employées pour résoudre le problème, et nous nous inspirerons de ces techniques pour notre résultat. Citons les travaux Gilboa et Schmeidler[48], Biagini et Frittelli [12, 13], Biagini Frittelli et Grasselli [14], Bordigoni Matousi et Schweizer [20], l'article "aux 6 auteurs" [36], Kabanov et Striker [62], Schachermayer[85], etc...

Les premiers à considérer une famille $\mathcal{P}$ contenant des lois mutuellement singulières furent Denis et Kervarec [38]. Ils supposent que $\mathcal{P}$ est une famille relativement compacte et qu'il n'y a pas d'option en portefeuille. Ils déduisent alors un théorème de min/max ainsi que l'existence d'une stratégie, et d'un élément $\mathbb{P}$ de $\mathcal{P}$ optimaux. Leurs résultats sont valables pour un large spectre de fonctions d'utilités. Des travaux de Tevzadze, Toronjadze et Uzunashvili [97] considérant un problème similiaire pour fonctions d'utilités exponentielle et puissance, ainsi que pour un critère de type moyenne-variance.
Plus récemment Matoussi, Possamai et Zhou [70] se sont intéressés au même problème en montrant qu'il revenait à la résolution d'une équation différentielle stochastique rétrograde du second ordre, et obtiennent des résultats pour les utilités puissance, logarithme et exponentielle. Les bornes hautes et basses sur la volatilité jouent encore une fois un rôle central dans leur approche. De plus leur approche impose une régularité assez forte sur les options considérées.
Enfin Bouchard, Moreau et Nutz [24] traitent ce problème en considérant de l'incertitude à la fois sur la dérive et sur la volatilité pour considérer le problème de couverture par quantile. Ils utilisent une approche mixant des techniques de cible stochastique et de solutions de viscosités, induisant naturellement des restrictions de régularité sur le payoff.

1.2.4.2 Motivation et nouveaux Résultats

On suppose que le prix de l'action suit une dynamique de la forme:

$$dS_t = S_t\left(b\sigma_t^2 dt + \sigma_t dW_t\right).$$

Notre étude est motivée par le problème de maximisation d'utilité robuste sans aucune borne sur la volatilité. Ainsi le processus σ aura pour seule contrainte d'être à valeurs dans $\mathbb{R}_+$. Le coefficient b qui apparait dans la dérive est constant et correspond au ratio de Sharpe. Nous supposons que b est observé par l'investisseur. Cela représente la situation financière où le gestionnaire de portefeuille n'a aucune idée sur la valeur de la volatilité, mais croit fermement en son calcul du ratio de Sharpe.
Nous cherchons ici à maximiser de manière robuste l'utilité d'un portefeuille contenant une option ξ fixée (pour laquelle on suppose uniquement de la mesurabilité) ainsi que des stratégies semi-statiques, à optimiser par l'agent. Cet ajout des stratégies statiques correspondent à l'explication financière détaillée dans la section 1.2.1.4 que l'on ne rappelle pas ici. Les fonctions d'utilités considérées sont de type exponentiel. Le problème s'écrit donc:

$$\sup_{(H,\lambda)} \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[U\left(\int_0^T H_t dS_t + \lambda(S_T) - \mu(\lambda) - \xi\right)\right].$$

Nous obtenons pour résultat principal que cette maximisation d'utilité robuste fournit une valeur égale au prix de sur-couverture robuste corrigé de l'entropie de la loi μ donnée par le prix des options. Ceci étant vrai pour tout payoff ξ, nous obtenons alors que le prix d'indifférence d'utilité robuste de n'importe quelle option est exactement égal au prix de surcouverture robuste.
Pour obtenir ce résultat nous nous inspirons des résultats obtenus dans le chapitre 3. Pour une stratégie statique λ fixée, nous utilisons les techniques de dualité convexes classiques pour ce problème pour obtenir une première borne. L'analyse quasi-sûre adaptée au problème de transport optimal martingale, et détaillée dans le chapitre 3, nous permet alors de construire une solution optimale pour la stratégie de trading dynamique et ainsi obtenir l'inégalité inverse. La maximisation sur les stratégies statiques se fait alors aisément.

1.3 Deux situations de gestion de portefeuille

Dans cette partie, nous traitons de deux questions indépendantes de gestion de portefeuille.

1.3.1 Liquidation optimale d'un actif indivisible avec investissement indépendant

1.3.1.1 Problématique et littérature existante

On s'intéresse à présent à un investisseur souhaitant liquider un actif indivisible Y et ayant une richesse initiale x. Celui ci peut représenter une usine, un appartement, ou simplement une obligation avec un nominal élevé. La dynamique supposée markovienne de ce bien est donnée par:

$$dY_t = Y_t\left(\mu(Y_t)dt + \sigma(Y_t)dB_t\right). \tag{1.3.1}$$

On suppose ici que le critère d'utilité de l'investisseur est donné par une fonction croissante concave U et celui s'intéresse donc naturellement à la quantité:

$$\sup_{\tau} \mathbb{E}\left[U(Y_\tau)\right], \tag{1.3.2}$$

où τ est un temps d'arrêt représentant sa stratégie de liquidation, qu'il souhaite choisir de manière optimale.

Mais cet investisseur qui a accès au marché, et plus particulièrement à un actif risqué S, indépendant de Y, va pouvoir dans un même temps investir de manière dynamique son capital initial x_0 en cet actif risqué S, s'offrant ainsi la possibilité d'augmenter son utilité par rapport à (1.3.2). Avec une stratégie de portefeuille dynamique π et une richesse à la date t associée $X_t^\pi := x_0 + \int_0^t \pi_s dS_s$, le nouveau critère est donc naturellement:

$$\sup_{\pi,\tau} \mathbb{E}\left[U(X_\tau^\pi + Y_\tau)\right]. \tag{1.3.3}$$

Ce problème mixte contrôle optimal/arrêt optimal fournit évidemment une fonction valeur supérieure à (1.3.2) (en considérant la stratégie d'investissement nulle $\pi = 0$ dans (1.3.3), on retombe immédiatement sur le problème associé (1.3.2)). La profitabilité de (1.3.3) comparée à (1.3.2) est donc en question ici.

Cette question a été étudiée par Henderson et Hobson [50] dans le cas particulier d'un critère d'utilité de type puissance, et pour une dynamique de type Black-Scholes (c'est à dire avec μ et σ constants) pour l'actif indivisible Y. Ils utilisent des techniques d'arrêt optimal pour résoudre les problèmes (1.3.2) et (1.3.3) en considérant les stratégies d'arrêt définies par des temps d'atteinte de barrières, puis en optimisant sur ces barrières. Ils établissent une liste des résultats obtenus en fonction des valeurs attribuées au paramètre de la fonction puissance. Ils retombent sur le résultat correspondant à l'utilité exponentielle par un raisonnement à la limite lorsque le paramètre de la fonction puissance tend vers 0.

1.3.1.2 Nouveaux Résultats

Notre approche se veut plus générale que celle d'Henderson et Hobson [50]. Nous ne traitons plus uniquement le cas d'une diffusion de type Black-Scholes pour l'actif indivisible, mais une diffusion markovienne générale (1.3.1). Nous nous intéressons aussi au cas d'une fonction d'utilité quelconque. Contrairement à [50], nous utilisons une approche systématique de contrôle stochastique, en considérant la formulation dynamique associée à (1.3.3), corrigée du changement de variable $y := R(z)$ où R est une fonction d'échelle de la diffusion Y, c'est à dire une transformation telle que $R(Y).$ est martingale.

$$\bar{V}(x,z) := \sup_{\tau,\pi} \mathbb{E}\left[U(R(Y_\tau^{R^{-1}(z)}) + X_\tau^{x,\pi})\right]. \tag{1.3.4}$$

Nous en déduisons l'équation aux dérivées partielles associée:

$$\min\left\{-\bar{V}_{zz}, -\bar{V}_{xx}, \bar{V} - U(x+R(z))\right\} = 0.$$

Un candidat à la solution se doit donc d'être partiellement concave par rapport à x et par rapport à z, tout en étant supérieur à $U(x+R(z))$. Ce candidat est donc tout naturellement la plus petite fonction majorant $U(x+R(z))$ et qui soit à la fois partiellement concave par rapport à x et par rapport à z (mais pas nécessairement par rapport au couple). Cette solution est donnée par la limite (possiblement infinie) $\bar{U}^\infty$ de la suite croissante $\left(\bar{U}^n\right)$ définie par:

$$\bar{U}^0(x,z) = U(x+R(z)),\ \bar{U}^{2n+1} = \left(U^{2n}\right)^{\text{conc}_z},\ \bar{U}^{2n+2} = \left(U^{2n+1}\right)^{\text{conc}_x},$$

où conc_x (resp conc_z) désigne l'enveloppe concave par rapport à x (resp z).

Des résultats et techniques classiques de la théorie des solutions de viscosité nous prouvent ensuite que la fonction valeur $\bar{V}$ va naturellement être au dessus de $\bar{U}^\infty$. L'autre inégalité va nous être fournie en utilisant les propriétés de concavité partielle de $\bar{U}^\infty$, nous assurant ainsi que $\bar{U}^\infty$ est bien solution du problème (1.3.4).

Nous exhibons enfin une stratégie ε-optimale $\tau_\varepsilon, \pi_\varepsilon$ du problème (1.3.4), construite à partir de la compréhension de la suite de fonction $\left(\bar{U}^n\right)$ comme suite d'enveloppes concave. Nous obtenons enfin sous une hypothèse technique l'existence d'une stratégie optimale pour le problème (1.3.4).

Ainsi il devient aisé de faire le parallèle entre le problème où l'on s'interdit toute stratégie d'investissement indépendant (1.3.2) et celle où on les autorise (1.3.3). En effet on observe que le problème (1.3.2) correspond exactement à la solution $\bar{U}^1$. L'équivalence entre les deux problèmes est donc réduite à l'égalité $\bar{U}^1 = \bar{U}^\infty$. Les exemples d'Henderson et Hobson [50] mettent en avant des situations pour lesquelles l'ajout de l'investissement indépendant n'apporte rien, et certaines où il va être profitable.

Nos résultats généralisent donc ceux d'Henderson et Hobson [50] et fournissent pour leur cas particulier une solution plus simple à calculer, que nous détaillons. En effet la solution explicite du problème (1.3.4) (ou (1.3.3), les deux sont équivalents) se calcule simplement comme une suite d'enveloppes concave. Notons que dans leurs applications, cette suite d'enveloppe concave devient stationnaire dans le pire des cas pour $n = 2$.

1.3.2 Prix d'indifférence d'utilité d'une option européenne pour petits coûts de transaction

1.3.2.1 Littérature existante

Le premier résultat important de prix d'indifférence d'utilité a été obtenu par Davis, Panas et Zariphopoulou [35], où ils montrent que le problème de pricing d'une option européenne avec coûts de transaction proportionnels revient à résoudre deux problèmes de contrôle stochastique dont les fonctions valeurs sont les uniques solutions de viscosités d'inégalités variationnelles quasi-linéaires. Mentionnons aussi Constantinides et Zariphopoulou [32, 33] qui ont déduit des bornes sur le prix des options. Puis, à partir des travaux de Barles et Soner [8] calculant rigoureusement le comportement limite de la fonction valeur lorsque les coûts de transaction et la tolérance au risque de l'investisseur tendent vers 0, beaucoup d'articles se sont intéressés à ces régimes limite. Les difficultés numériques de résolution du problème de [35] justifient amplement l'utilisation de ces développements asymptotiques. Citons Whalley et Willmott [101] en ce qui concerne un développement asymptotique formel pour petits coûts de transaction, et Bouchard [21] et Bouchard, Kabanov et Touzi [23] pour une preuve rigoureuse du comportement du prix d'indifférence lorsque l'aversion au risque tend vers l'infini. La première preuve rigoureuse du résultat de [101] a été obtenue en dimension 1 par Shreve et Soner [87]. Plusieurs résultats rigoureux [17, 47, 59, 60, 84] et formels [5, 49, 63, 64] s'en sont suivi. Les travaux de Bichuch [16] seront discutés de façon détaillée par la suite. Le problème multidimensionnel, qui présente des difficultés techniques liées aux frontières libre (voir Shreve et Soner [86, 89] et Chen et Dai [30]) est resté hors de portée jusqu'aux travaux de Bichuch et Shreve [18] où ils se sont intéressés à un marché contenant deux actifs risqués suivant des dynamiques browniennes arithmétiques.
Plus récemment le problème général a été résolu par Soner et Touzi [90], en reliant cette question du développement asymptotique pour petits coûts de transaction à la théorie de l'homogénisation. En effet le premier terme du développement se trouve être expliqué comme la valeur propre associée à l'équation de programmation dynamique d'un problème de contrôle stochastique ergodique. Ceci permet une preuve rigoureuse, basée sur des techniques d'homogénisation, même si ce problème s'éloigne quelque peu du problème d'homogénisation classique, notamment à cause d'une variable dite "rapide" qui n'apparait pas dans les équations originelles. Leur approche limitée à la dimension 1, permet néanmoins de traiter une classe importante de dynamiques markoviennes pour l'actif risqué, ainsi que des fonctions d'utilités complètement générales (la littérature se limite très souvent aux fonction d'utilités puissances et à des browniens géométriques pour les diffusions). Ces résultats ont ensuite été étendus au cas multidimensionnel par Possamai, Soner et Touzi [81].
Dès lors, leurs techniques ont été utilisées par Altarovici, Muhle-Karbe et Soner [3] pour traiter le cas à coûts de transaction fixés, par Bouchard Moreau et Soner [25] pour le problème de couverture sous contraintes de pertes et Moreau Muhle-Karbe et Soner [72] pour obtenir un modèle d'impact sur le prix.

1.3.2.2 Motivation et nouveaux Résultats

Ce travail se place dans le contexte général initié dans le cas unidimensionnel par Soner et Touzi [90], puis dans le cas multidimensionnel par Possamai, Soner et Touzi [81]. Notre objectif principal consiste à prouver rigoureusement un développement asymptotique du prix d'indifférence d'utilité d'une option européenne pour des modèles de diffusions markoviens généraux, et pour n'importe quelle fonction d'utilité. Les travaux en ce sens réalisés jusqu'alors sont ceux de Bichuch [16], puis très récemment ceux de Bouchard, Moreau et Soner [25]. Dans [16] l'auteur considère le cas d'une utilité de type exponentielle permettant d'obtenir, grâce aux propriétés d'échelle, simplement et explicitement le prix de la fonction valeur du problème de contrôle. De plus son approche requiert des hypothèses assez fortes notamment sur la dérivabilité de l'option (qui doit être C^4), sur la dimension (qui ne peut excéder 1) et sur la dynamique Black-Scholes du prix de l'actif. Bouchard, Moreau et Soner se restreignent à la dimension 1, même si leurs travaux doivent pouvoir se généraliser aux dimensions supérieures. Leur méthode consiste à considérer directement le développement du prix, alors que la nôtre implique de calculer tout d'abord le développement asymptotique de la fonction valeur, puis d'en déduire le développement asymptotique du prix. De plus leur méthode nécessite, comme pour Bichuch [16], une régularité très forte de la fonction valeur.

Nous nous plaçons pour notre étude dans le cadre multidimensionnel et nous ne supposons que de la continuité pour le payoff de l'option. Notre approche pour obtenir un développement du prix d'indifférence d'utilité consiste à obtenir des développements asymptotiques pour les fonctions valeurs avec et sans l'option considérée, puis à en déduire le prix d'indifférence d'utilité. Pour obtenir les développements asymptotiques des fonctions valeurs, nous suivons la méthode introduite dans [90], puis dans [81] pour le cas multidimensionnel, en utilisant les techniques d'homogénisation. Nous obtenons formellement un système d'équations vérifiées par les composantes du problème et nous vérifions que ceux-ci en sont bien solution. Le prix d'indifférence d'utilité suit alors naturellement.
Nous concluons en montrant comment nos techniques permettent d'obtenir les résults de [16] pour une utilité exponentielle. Nous détaillons aussi le cas des fonctions d'utilité puissance, même si là, notre approche reste formelle.

Part I

Martingale optimal transport: discrete and continuous issues

Chapter 2

A utility maximization proof of Strassen's Theorem

Contents

2.1 Introduction

This chapter is dedicated to a new proof of Strassen's theorem. This theorem concerns the existence of a martingale probability measure with given marginals. This theorem states that for two probability measures μ and ν on $\mathbb{R}^d$, there exists a martingale law with first marginal μ and second marginal ν if and only if μ and ν are in convex order (this will be defined later). Its original proof is due to Strassen in [94]. This is an application of Hahn-Banach Theorem under a suitable topology. This proof fails from having any significant financial meaning, while the result is the starting point of the martingale optimal transport motivated by obtaining no-arbitrage bounds for prices of derivatives.

Our approach consists in using utility maximizations results to obtain the existence of a martingale measure with given marginals. This method is mainly inspired from a proof of the Fundamental Theorem of Asset Pricing (FTAP) by Rogers [83]. This theorem links the no-arbitrage condition and the existence of an equivalent martingale measure. The method of Rogers consists in finding some optimal investment strategy and then deriving from this strategy a martingale law equivalent to the law under which the utility maximization is solved. We adapt this method to the setting of Strassen's Theorem and face some very different problems than the ones of Roger's method to the proof of FTAP.

In Section 2.2 , we describe precisely the problem and build the right setup of the utility maximization involved. In Section 2.3 we derive the result in the case where μ has finite support. We provide examples showing that the utility maximization in the general case is completely non-trivial, in particular in the way to define the probability law under which we maximize our utility. In Section 2.4 we use a convergence argument to obtain the general result of Strassen's theorem. Finally in Section we build an algorithm to obtain an element of $\mathcal{M}(\mu,\nu)$.

2.2 Main result

2.2.1 The probabilistic framework

We consider $\Omega := \mathbb{R}^d \times \mathbb{R}^d$. We denote by (X,Y) the canonical process on $\mathbb{R}^d \times \mathbb{R}^d$ and $\mathcal{P}_{\mathbb{R}^{2d}}$ the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$.
Let μ and ν be two probability measures on $\mathbb{R}^d$. We denote by $\mathcal{P}(\mu,\nu)$ the set:

$$\mathcal{P}(\mu,\nu) := \left\{\mathbb{P} \in \mathcal{P}_{\mathbb{R}^{2d}} :\ X \sim_{\mathbb{P}} \mu \text{ and } Y \sim_{\mathbb{P}} \nu\right\}.$$

We then introduce the subset of $\mathcal{P}(\mu,\nu)$ consisting of martingale probabilities:

$$\mathcal{M}(\mu,\nu) := \left\{\mathbb{P} \in \mathcal{P}(\mu,\nu) :\ \mathbb{E}^{\mathbb{P}}\left[Y|X\right] = X,\ \mu - \text{a.s.}\right\}.$$

We first see easily that $\mathcal{P}(\mu,\nu)$ is non-empty since $\mu \otimes \nu$ is one of its elements. The non-emptyness of $\mathcal{M}(\mu,\nu)$ is much more complicated. Indeed a first condition to obtain a martingale law is the integrability of the pair (X,Y). So a first necessary condition to obtain that $\mathcal{M}(\mu,\nu)$ is non empty is that:

$$\int |x|\mu(dx) + \int |y|\nu(dy) < \infty. \tag{2.2.1}$$

In all the following, for any measure μ on $\mathbb{R}^d$, we will denote by $\mathbb{L}^1(\mu)$ the subset of $\mathbb{L}^0(\mathbb{R}^d,\mathbb{R})$ consisting of all elements f such that $\int |f(x)|\mu(dx) < \infty$, and $\mathbb{L}^0(\mathbb{R}^d) := \mathbb{L}^0(\mathbb{R}^d,\mathbb{R}^d)$.

2.2.2 Strassen's Theorem

We show here the following theorem:

Theorem 2.2.1. *Assume that μ and ν are two probability measures on $\mathbb{R}^d$ with $\int |x|\mu(dx) + \int |y|\nu(dy) < \infty$, then $\mathcal{M}(\mu,\nu) \neq \emptyset$ if and only if $\mu \preceq \nu$ in convex order, i.e.*

$$\mu(g) \leq \nu(g) \textit{ for all convex function } g.$$

If $\mathcal{M}(\mu,\nu) \neq \emptyset$, then we have easily that $\mu(g) \leq \nu(g)$ for all convex function g. Indeed consider $\mathbb{P} \in \mathcal{M}(\mu,\nu)$ and g convex with $\nu(g)$ finite, then by Jensen's inequality for conditionnal expectations, we have:

$$\nu(g) = \mathbb{E}^{\mathbb{P}}\left[g(Y)\right] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[g(Y)|X\right]\right] \geq \mathbb{E}^{\mathbb{P}}\left[g\left(\mathbb{E}^{\mathbb{P}}\left[Y|X\right]\right)\right] = \mathbb{E}^{\mathbb{P}}\left[g(X)\right] = \mu(g).$$

The above inequality holds even if the quantities are note finite. Indeed for g convex and finite, we know that there exists ℓ affine such that $g(x) \geq \ell(x)$ for all x. Then since $\int |x|\nu(dx) < \infty$, we have that $\nu(g) \geq \nu(\ell)$ and $\nu(\ell)$ finite. So $\nu(g)$ (and similarly $\mu(g)$) is finite or $+\infty$ and the previous calculation holds, even if $\nu(g) = \infty$.

We then need to prove the reverse inequality. For that purpose, we first explore the particular case where μ has finite support in section 2.3, and then extend the result to a general measure μ in section 2.4.

2.2.3 The mechanism of utility maximization: heuristic and technical results

The original proof of Theorem 2.2.1 is an application of Hahn-Banach Theorem. Our aim is to establish a new proof by considering a utility maximization framework.

For $(h, \phi, \psi) \in \mathbb{L}^0(\mathbb{R}^d) \times \mathbb{L}^1(\mu) \times \mathbb{L}^1(\nu)$, we define:

$$h^{\otimes}(x, y) := h(x) \cdot (y - x), \quad \phi^{\mu}(x) := \phi(x) - \mu(\phi), \quad \phi^{\nu}(y) := \phi(y) - \nu(\psi),$$

and:

$$\phi \oplus \psi(x, y) := \phi(x) + \psi(y).$$

Assume further that $e^{-(h^{\otimes} + \phi \oplus \psi)} \in \mathbb{L}^1(\mu \otimes \nu)$, then the function

$$\psi_{\mu}^{*}(h, \phi)(y) := \ln \left(\int e^{-\left(h^{\otimes} + \phi^{\mu}\right)(x,y)} \mu(dx) \right)$$

is well defined and we have:

$$\begin{aligned} \int e^{-(h^{\otimes} + \phi \oplus \psi)(x,y)} \mu \otimes \nu(dx, dy) &= \int e^{-\psi^{\nu}(y)} \left(\int e^{-(h^{\otimes} + \phi^{\mu})(x,y)} \mu(dx) \right) \nu(dy) \\ &= \int e^{\psi_{\mu}^{*}(h,\phi)(y) - \psi^{\nu}(y)} \nu(dy). \end{aligned}$$

We consider then the problem:

$$V(\mu, \nu) := \sup_{(h,\phi,\psi) \in \mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h, \phi, \psi), \tag{2.2.2}$$

where

$$J^{\mu,\nu}(h, \phi, \psi) := \mathbb{E}^{\mu \otimes \nu} \left[U((h^{\otimes} + \phi^{\mu} \oplus \psi^{\nu})(X, Y)) \right], \quad U(x) := -e^{-x},$$

$\mathbb{E}^{\mu \otimes \nu}$ is the expectation operator under the product measure $\mu \otimes \nu$, and the set of admissible strategies $\mathbf{D}(\mu, \nu)$ is given by:

$$\begin{aligned} \mathbf{D}(\mu, \nu) := \big\{ &(h, \phi, \psi) \in \mathbb{L}^0(\mathbb{R}^d) \times \mathbb{L}^1(\mu) \times \mathbb{L}^1(\nu), \ h^{\otimes} + \phi \oplus \psi \in \mathbb{L}^1(\mu \otimes \nu), \\ &e^{-(h^{\otimes} + \phi \oplus \psi)} \in \mathbb{L}^1(\mu \otimes \nu), \ \text{and} \ \psi^{*}(h, \phi) \in \mathbb{L}^1(\nu) \big\}. \end{aligned}$$

In all the following, we will abuse notation by denoting that $h \in \mathbf{D}(\mu, \nu)$ (resp $(h, \phi) \in \mathbf{D}(\mu, \nu)$) if there exists (ϕ, ψ) (resp ψ) such that $(h, \phi, \psi) \in \mathbf{D}(\mu, \nu)$ and we will say that h (resp (h, ϕ)) is admissible.

Remark 2.2.1. *Following Rogers [83], we want to find some $\mathbb{P}^* \in \mathcal{M}(\mu, \nu)$ with $\mathbb{P}^*$ absolutely continuous w.r.t. the measure used for the utility maximisation, here $\mu \otimes \nu$. As we will explain later on, this utility maximization makes sens in the particular case where μ has finite support. Then any probability distribution with support on supp$(\mu) \times$ supp(ν) will be absolutely continuous with respect to $\mu \otimes \nu$, which will be the key here to obtain an element of $\mathcal{M}(\mu, \nu)$.*

Our setup corresponds to the following financial situation. The financial market is composed of a family of tradable assets corresponding to a d-dimensional vector S with price at time t_1 is $S_{t_1} = X$ and price at time $t_2 > t_1$ is $S_{t_2} = Y$, plus a non risky asset normalized here to unity. An investor can then take a dynamic position which gives wealth at time t_2:

$$W_{t_2}^{h} := W_0 + h_0(S_0)(X - S_0) + h_1(X)(Y - X),$$

where W_0 is its starting wealth and where h_0 (resp h_1) corresponds to the position in the asset vector S at time $t = 0$ (resp $t = t_1$) he decides to take.
Moreover we assume that calls and puts of all strikes for maturity t_1 and t_2 are tradable so that, under the assumption of a linear pricing funcionnal, we have by [27] that any european position ϕ of maturity t_1 (resp ψ of maturity t_2) has the no arbitrage price $\mu(\phi)$ (resp $\nu(\psi)$, where μ (resp ν) is the law of S at time t_1 (resp t_2) identified from the calls and puts prices. The investor is then allowded to take any european position $\phi \in \mathbb{L}^1(\mu)$, $\psi \in \mathbb{L}^1(\nu)$ and then his wealth at time t_2 is given by:

$$W_{t_2}^{h,\phi,\psi} = W_0 + h_0(S_0)(X - S_0) + h_{t_1}(X)(Y - X) + \phi(X) - \mu(\phi) + \psi(Y) - \nu(\psi).$$

Notice that the component $h_0(S_0)(X - S_0)$ can be added to the position ϕ since we observe S_0 at time 0. However, since S_0 is deterministic, this would be redundant with $\phi(X)$.

We detail here some preliminary results, which explain the mechanism of the maximization. These results holds even if $\mu \preceq \nu$ is not verified and indeed we do not use this property on the following proofs. The first lemmas explain how we can easily improve a given strategy $(h, \phi, \psi) \in \mathbf{D}(\mu, \nu)$ into one direction:

Lemma 2.2.1. *Let μ and ν be two measures on $\mathbb{R}^d$. For every $(h_0, \phi_0) \in \mathbf{D}(\mu, \nu)$ we have:*

$$\sup_{\psi \text{ s.t. } (h_0,\phi_0,\psi)\in\mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h_0, \phi_0, \psi) \;=\; J^{\mu,\nu}(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0)) = -e^{\nu(\psi^*_\mu(h_0,\phi_0))} < 0.$$

Proof. For any $(h_0, \phi_0, \psi) \in \mathbf{D}(\mu, \nu)$ we have:

$$\begin{aligned} J^{\mu,\nu}(h_0, \phi_0, \psi) &= -\int e^{-(h_0^\otimes + \phi_0^\mu + \psi^\nu)(x,y)} \mu(dx)\nu(dy) \\ &= -\int \nu(dy) e^{-\psi^\nu(y)} \left(\int e^{-(h_0^\otimes + \phi_0^\mu)(x,y)} \mu(dx) \right) \end{aligned}$$

Since h_0 and ϕ_0 are admissible, we see that $e^{\psi^*_\mu(h_0,\phi_0)(\cdot) - \psi(\cdot)}$ is well defined and in $\mathbb{L}^1(\nu)$ and:

$$J^{\mu,\nu}(h_0, \phi_0, \psi) = \int -e^{\psi^*_\mu(h_0,\phi_0)(xy - \psi^\nu(y)} \nu(dy).$$

Then by Jensen's Inequality, we have for every ψ such that $(h_0, \phi_0, \psi) \in \mathbf{D}(\mu, \nu)$:

$$J^{\mu,\nu}(h_0, \phi_0, \psi) \le -e^{\nu(\psi^*(h_0,\phi_0)) - \nu(\psi^\nu))} = -e^{\nu(\psi^*)}.$$

Since $(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0))$ is not necessarily in $\mathbf{D}(\mu, \nu)$, we construct a sequence (ψ_n) such that $(h_0, \phi_0, \psi_n) \in \mathbf{D}(\mu, \nu)$ and $J(h_0, \phi_0, \psi_n) \to J(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0))$ as $n \to \infty$.

We consider ψ_0 such that $(h_0, \phi_0, \psi_0) \in \mathbf{D}(\mu, \nu)$. For $n \in \mathbb{N}$, we define $\psi_n := \psi^*_\mu(h_0, \phi_0) \vee (\psi_0 - n)$. Clearely, $\psi_0 - n \le \psi_n \le \psi^*_\mu(h_0, \phi_0)$, so that $\psi_n \in \mathbb{L}^1(\nu)$, and $\psi_n(y) \to \psi^*_\mu(h_0, \phi_0)(y)$ for all y when $n \to \infty$. We also have that:

$$|\psi_n - \psi^*_\mu(h_0, \phi^0)| = |\psi_0 - \psi^*_\mu(h_0, \phi_0) - n|\mathbf{1}_{\psi^*_\mu(h_0,\phi_0) - \psi_0 \le -n} \le |\psi^*_\mu(h_0, \phi_0)| + |\psi_0|.$$

We verify now that $(h_0, \phi_0, \psi_n) \in \mathbf{D}(\mu, \nu)$. Since $\psi^*_\mu(h_0, \phi_0) \in \mathbb{L}^1(\nu)$, and

$$|\psi_0 - \psi^*_\mu(h_0, \phi_0) - n|\mathbf{1}_{\psi^*_\mu(h_0,\phi_0) - \psi_0 \le -n}(y) \to 0$$

when $n \to \infty$ for all y, we have by dominated convergence theorem that $\nu(|\psi_n - \psi^*_\mu(h_0,\phi_0)|) \to 0$ when $n \to \infty$ and in particular

$$\nu(\psi_n) \underset{n\to\infty}{\longrightarrow} \nu(\psi^*_\mu(h_0,\phi_0)).$$

Moreover, since $\psi_n - \psi_0 \in \mathbb{L}^1(\nu)$, we have

$$h_0^\otimes + \phi_0 \oplus \psi_n = h_0^\otimes + \psi_n - \psi_0 + \phi_0 \oplus \psi_0 \in \mathbb{L}^1(\mu\otimes\nu).$$

Then by construction, we also have that

$$0 \le e^{-(h_0^\otimes + \phi_0 \oplus \psi_n)} \; = \; e^{-(h_0^\otimes + \phi_0\oplus\psi_0) + (\psi_0 - \psi_n)} \le e^n e^{-(h_0^\otimes + \phi_0 \oplus \psi_0)} \in \mathbb{L}^1(\mu\otimes\nu).$$

We then have the admissibility of (h_0, ϕ_0, ψ_n). Finally, we obtain that:

$$-\infty < J(h_0,\phi_0,\psi_n) = \int -e^{\psi^*_\mu(h_0,\phi_0)(y) - \psi_n(y) + \nu(\psi_n)}\nu(dx).$$

And since $-\psi^*_\mu(h_0,\phi_0)(y) + \psi_n(y) \searrow 0$ as $n \nearrow \infty$, we have by monotone convergence theorem that

$$\int -e^{\psi^*_\mu(h_0,\phi_0)(y) - \psi_n(y)}\nu(dy) \nearrow \int -e^0 \nu(dy) = -1 \quad \text{as } n \searrow \infty,$$

so that

$$\lim_{n\to\infty} J^{\mu,\nu}(h_0,\phi_0,\psi_n) = J^{\mu,\nu}(h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)).$$

□

We now provide the crucial optimality property of the maximizer ψ^*_μ.

Definition 2.2.1. *Let μ and ν be two probability measures on $\mathbb{R}^d$, then for any $(h,\phi,\psi) \in \mathbf{D}(\mu,\nu)$, we denote by $\mathbb{P}^{h,\phi,\psi}$ the probability measure defined by:*

$$\frac{d\mathbb{P}^{h,\phi,\psi}}{d\mu\otimes\nu} = \frac{e^{-(h^\otimes + \phi^\mu \oplus \psi^\nu)(X,Y)}}{\mathbb{E}^{\mu\otimes\nu}\left[e^{-(h^\otimes+\phi^\mu\oplus\psi^\nu)(X,Y)}\right]}.$$

The following proposition shows that if we consider the improved strategy ψ^*_μ, then the law induced by this strategy has a particular form. This corresponds to the first order condition of optimality detailed in [83].

Proposition 2.2.1. *Let μ and ν be two probability measures on $\mathbb{R}^d$, then for any $(h_0,\phi_0,\psi_0) \in \mathbf{D}(\mu,\nu)$, we have that $\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}$ has for second marginal law ν, i.e.*

$$\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}[Y \in A] = \nu(A), \quad \textit{for all } A \in \mathcal{B}_{\mathbb{R}^d}.$$

Proof. Assume to the contrary that there exists ψ continuous and bounded such that $\mathbb{E}^{\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}}[\psi(Y)] > \nu(\psi)$. Then since ψ is bounded, we have that $y \mapsto \psi^\nu(y)e^{-\varepsilon\psi^\nu(y)}$ is bounded, so that by the dominated convergence theorem :

$$g : \varepsilon \mapsto \mathbb{E}^{\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}}\left[-e^{-\varepsilon\psi^\nu(Y)}\right]$$

is C^∞ , with first derivative

$$g'(\varepsilon) = \mathbb{E}^{\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}}\left[\psi^\nu(Y)e^{-\varepsilon\psi^\nu(Y)}\right].$$

We then have $g(0) = -1$, $g'(0) = \mathbb{E}^{\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}}[\psi^\nu(Y)] > 0$, so that for $\varepsilon > 0$ small enough, we have

$$\mathbb{E}^{\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}}\left[-e^{-\varepsilon\psi^\nu(Y)}\right] > -1.$$

By the definition of $\mathbb{P}^{h_0,\phi_0,\psi^*_\mu(h_0,\phi_0)}$, this is equivalent to:

$$J^{\mu,\nu}(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0) + \varepsilon\psi) > J^{\mu,\nu}(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0)).$$

Since $(h_0, \phi_0, \psi^*_\mu(h_0, \phi_0) \in \mathbf{D}(\mu, \nu)$, this contradicts the optimality of $\psi^*_\mu(h_0, \phi_0)$.

□

These techniques will be used in the next sections to obtain the required property of the probability measures involved. One could also consider some maximization over the parameter ϕ or h. Since they are not involved in the following, we do not detail here how to do it. Notice that the maximization over ϕ is very similar to the maximization over ψ, by considering the analogue $\phi^*_\nu(h, \psi)$ of $\psi^*_\mu(h, \phi)$ defined by:

$$\phi^*_\nu(h, \psi)(x) := \ln\left(\int e^{-(h^\otimes+\psi^\nu)(x,y)}\nu(dy)\right).$$

Under some additionnal assumptions on the admissibility set $\mathbf{D}(\mu, \nu)$, the probability induced by this maximization would have for first marginal law μ, this is the analogue of Proposition 2.2.1. The maximization over the parameter h is slightly different and one should get into the details of [83] to see how to build a maximizer. Imagine that the maximizer h^* exists, then with similar calculation of Proposition 2.2.1, the probability induced by h^* would be a martingale probability, but we would have no idea of the marginal laws of this probability.

2.3 The result for μ with finite support

One major difference between our approach and the one introduced by Rogers in [83] is that it is difficult, and often impossible, to exhibit a nice pricing measures under which we do our utility maximization. In addition, finding an optimal triplet $(h^*, \phi^*, \psi^*_\mu)$ is not as easy as the construction made in the previous section where we considered only one direction maximizations. If we have this optimal admissible triplet (h^*, ϕ^*, ψ^*) under some probability $\mathbb{P}$, then the proability measure $\mathbb{Q}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}} = ce^{-(h^{*,\otimes}+\phi^{*,\mu}\oplus\psi^{*,\nu})(x,y)}$ for some renormalization constant $c > 0$ would be an element of $\mathcal{M}(\mu, \nu)$. This is morally the consequence of the first order optimal conditions obtained in Proposition 2.2.1 for the second marginal law.
One can then see that $\mathbb{Q}$ and $\mathbb{P}$ have to be equivalent. The most natural, and we formulated our problem in that sense in Section 2.2.3, is to consider for $\mathbb{P}$ the natural measure $\mu \otimes \nu$. But if we consider the particular case of $\mathbb{R}^d = \mathbb{R}$, $\nu = \mu = L([0, 1])$ the lebesgue measure on $[0, 1]$, then the only element of $\mathcal{M}(\mu, \mu)$ is clearly $\mathbb{Q} := \mu(dx)\delta_x(dy)$ (the measure giving weight on the diagonal). We observe also that $\mathbb{Q}$ and $\mu \otimes \mu$ are singular, which makes us lose any hope of obtaining a triplet of optimal strategies (h, ϕ, ψ).

In order to avoid these difficulties, we fist focus on the case of μ with finite support, which will be more convenient for calculus:

Theorem 2.3.1. *Assume that μ has finite support, $\int |y|\nu(dy) < \infty$, and $\mu \preceq \nu$, then we have:*

$$\mathcal{M}(\mu, \nu) \neq \emptyset.$$

We first show that, if μ and ν are in convex order, the problem (2.2.2) is non-degenerate in the following sense:

Proposition 2.3.1. *Assume that μ has finite support and that $\mu \preceq \nu$, then:*

$$V(\mu,\nu) := \sup_{(h,\phi,\psi)\in\mathbf{D}(\mu,\nu)} \mathbb{E}^{\mu\otimes\nu}\left[U\left((h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)\right)\right] < 0 \tag{2.3.1}$$

Proof. We separate the proof in two steps.
Step 1: $V(\mu,\nu) \le V(\mu,\mu)$.
Consider a triplet $(h,\phi,\psi) \in \mathbf{D}(\mu,\nu)$, then we have:

$$\begin{aligned}\mathbb{E}^{\mu\otimes\nu}\left[-e^{-(h(X)(Y-X)+\phi^{\mu}(X)+\psi^{\nu}(Y))}\right] &= \mathbb{E}^{\mu\otimes\nu}\left[\mathbb{E}^{\mu\otimes\nu}\left[-e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}|Y\right]\right]\\ &= \mathbb{E}^{\mu\otimes\nu}\left[-e^{-\psi^{\nu}(Y)}\mathbb{E}^{\mu\otimes\nu}\left[e^{-(h(X)(Y-X)+\phi^{\mu}(X))}|Y\right]\right]\end{aligned}$$

We then denote $f(y) := \int e^{-(h(x)(y-x)+\phi^{\mu}(x))}\mu(dx) = e^{\psi^*_{\mu}(h,\phi)(y)} > 0$. We clearely have:

$$\mathbb{E}^{\mu\otimes\nu}\left[-e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\right] = \mathbb{E}^{\nu}\left[-e^{-\psi^{\nu}(Y)}f(Y)\right].$$

Since μ has finite support, h and ϕ are bounded, and we obtain by bounded convergence theorem that $\psi^*_{\mu}(h,\phi)$ is C^{∞} and convex. Indeed define for all $z \in \mathbb{R}^d$ the function $g_z : \mathbb{R} \to \mathbb{R}^+$:

$$g_z \ : \ t \ \mapsto \ln\left(\int e^{-\left(h^{\otimes}+\phi^{\mu}\right)(x,y+tz)}\mu(dx)\right),$$

we have to show that g_z is convex for all $z \in \mathbb{R}^d$. We obtain that g_z is C^{∞} and

$$g_z''(t) = \frac{\int (h(x)\cdot z)^2 e^{-\left(h^{\otimes}+\phi^{\mu}\right)(x,y+tz)}\mu(dx)f(y+tz) - \left(\int (h(x)\cdot z)e^{-\left(h^{\otimes}+\phi^{\mu}\right)(x,y+tz)}\mu(dx)\right)^2}{f^2(y+tz)}.$$

We obtain by Cauchy-Schwartz inequality that $g_z''(t) \ge 0$. Indeed defining the probability measure $\tilde{\mu}(dx) := \frac{e^{-\left(h^{\otimes}+\phi^{\mu}\right)(x,y+tz)}}{\int e^{-(h^{\otimes}+\phi^{\mu})(\tilde{x},y+tz)}\mu(d\tilde{x})}\mu(dx)$, we have:

$$g_z''(t) = \int (h(x)\cdot z)^2\,\tilde{\mu}(dx) - \left(\int (h(x)\cdot z)\,\tilde{\mu}(dx)\right)^2 \ge 0.$$

So g_z is convex, which is the required result.

We now recall from Lemma 2.2.1 that:

$$\sup_{(h,\phi,\psi)\in\mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h,\phi,\psi) = \sup_{(h,\phi)\in\mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h,\phi,\psi^*_{\mu}(h,\phi)) = \sup_{(h,\phi)\in\mathbf{D}(\mu,\nu)} -e^{\nu(\psi^*_{\mu}(h,\phi))}.$$

Then since $\mu \preceq \nu$, and for $(h,\phi) \in \mathbf{D}(\mu,\nu)$ $\psi^*_{\mu}(h,\phi)$ is convex, we have that $\mu(\psi^*_{\mu}(h,\phi)) \le \nu(\psi^*_{\mu}(h,\phi))$, and then:

$$-e^{\nu(\psi^*_{\mu}(h,\phi))} \le -e^{\mu(\psi^*_{\mu}(h,\phi))}.$$

Now since μ has finite support, each pair $(h,\phi) \in \mathbf{D}(\mu,\nu)$ is bounded, so that $(h,\phi,0)$ is bounded and then in $\mathbf{D}(\mu,\mu)$. We then have that:

$$\sup_{(h,\phi)\in\mathbf{D}(\mu,\nu)} -e^{\nu(\psi^*_{\mu}(h,\phi))} \le \sup_{(h,\phi)\in\mathbf{D}(\mu,\nu)} -e^{\mu(\psi^*_{\mu}(h,\phi))} \le \sup_{(h,\phi)\in\mathbf{D}(\mu,\mu)} -e^{\mu(\psi^*_{\mu}(h,\phi))}.$$

We then obtain:

$$V(\mu,\nu) \le V(\mu,\mu). \tag{2.3.2}$$

Step 2: We now show that $V(\mu,\mu) < 0$.
We define $\mathbb{P}_1(dx,dy) := \mu(dx)\delta_{\{x\}}(dy)$. Since μ has finite support, then the probability measures $\mathbb{P} := \mu \otimes \mu$ and $\mathbb{Q} := \frac{1}{2}\mathbb{P}_1 + \frac{1}{2}\mathbb{P}$ are equivalent, with density $0 < \frac{d\mathbb{P}}{d\mathbb{Q}} < \infty$ on the support of $\mathbb{P}$. The density $\frac{d\mathbb{P}}{d\mathbb{Q}}$ will be denoted as a function of (X,Y): $f(X,Y) := \frac{d\mathbb{P}}{d\mathbb{Q}}$. Then we have for $(h,\psi,\psi) \in \mathbf{D}(\mu,\mu)$:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[-e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\right] &= \mathbb{E}^{\mathbb{Q}}\left[-e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\frac{d\mathbb{P}}{d\mathbb{Q}}\right] \\
&\le \frac{1}{2}\mathbb{E}^{\mathbb{P}_1}\left[-e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\frac{d\mathbb{P}}{d\mathbb{Q}}\right] \\
&= \frac{1}{2}\mathbb{E}^{\mathbb{P}_1}\left[-e^{\ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\right].
\end{aligned}$$

Then, by Jensen's inequality, we have using the fact that f takes finite many values, so that $\int |\ln(f(x,x))|\mu(dx) < \infty$:

$$\mathbb{E}^{\mathbb{P}_1}\left[-e^{\ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(X,Y)}\right] \le -e^{\int \ln(f(x,x))\mu(dx)} < 0,$$

which ends the proof.

□

Remark 2.3.1. *The assumption of μ with finite support is crucial here, since we are facing the problem $V(\mu,\mu)$. Indeed we absolutely need the existence of an integrable density measure $\frac{d\mathbb{P}}{d\mathbb{Q}}$ in step 2, which may not exist in the general case. As an example, consider $\mu = L([0,1])$ the lebesgue measure on $[0,1]$. Using the notations introduced in Step 2, we have that $\mathbb{P}[X=Y]=0$ while $\mathbb{Q}[X=Y]=\frac{1}{2}$, and then $f(X,X)=0$. So we have $\int \ln(f(x,x))\mu(dx) = -\infty$ and then $-e^{\int \ln(f(x,x))\mu(dx)} = 0$, and then we cannot conclude in Step 2.*

Proof of Theorem 2.3.1 We consider a maximizing sequence (h_n,ϕ_n,ψ_n) of (2.3.1). We then consider the probability $\mathbb{P}_n := \mathbb{P}^{h_n,\phi_n,\psi^*_\mu(h_n,\phi_n)}$ defined in Section 2.2.3. We recall from Proposition 2.2.1 that the second marginal law of $\mathbb{P}_n$ is ν.

Now since μ has compact support, the sequence $(\mathbb{P}_n)$ is tight. Indeed if we denote by $\tilde{\mu}_n$ the first marginal law of $\mathbb{P}_n$, we know that the support of μ_n is included in the support of μ. Then consider K^ν_ε a compact of $\mathbb{R}^d$ such that $\nu[Y \notin K^\nu_\varepsilon] \le \varepsilon$, and k such that $\mu[X \notin \bar{B}(0,k)] = 0$, we have:

$$\mathbb{P}_n[(X,Y) \notin \bar{B}(0,k) \times K^\nu_\varepsilon] \le \mu_n[X \notin \bar{B}(0,k)] + \nu[Y \notin K^\nu_\varepsilon] \le \nu[Y \notin K^\nu_\varepsilon] \le \varepsilon.$$

We then have up to extraction that $\mathbb{P}_n$ converges weakly to some $\mathbb{P}$ with second marginal law ν. This last property is obvious since for every continuous and bounded ψ, we have for all n, $\mathbb{E}^{\mathbb{P}_n}[\psi(Y)] = \nu(\psi)$, and $\mathbb{E}^{\mathbb{P}}[\psi(Y)] = \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\psi(Y)] = \nu(\psi)$.

We now show that $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. It only remains to show that $\mathbb{P}$ is a martingale measure with first marginal law μ.

Step 1: We first prove that $\mathbb{P}$ is a martingale measure. Indeed assume to the contrary that $\mathbb{P}$ is not a martingale law, then there exists h bounded such that:

$$\mathbb{E}^{\mathbb{P}}\left[h(X)(Y-X)\right] > 0.$$

Since μ has compact support, there exists a constant k such that $h(x)y+k|y|>0$, $\mu(dx)\times\nu(dy)$-a.s. . And by the integrability assumption on ν, we have for $\psi(y):=k|y|$:

$$h^{\otimes}+\psi^{\nu}\in\mathbb{L}^1(\mu\otimes\nu) \text{ and bounded from below.}$$

We then have:

$$\mathbb{E}^{\mathbb{P}}\left[(h^{\otimes}+\psi^{\nu})(X,Y)\right] > 0. \tag{2.3.3}$$

Define now $g:[0,1]\to\mathbb{R}$:

$$\varepsilon\mapsto g(\varepsilon):=\mathbb{E}^{\mathbb{P}}\left[-e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(X,Y)}\right].$$

We have the existence of a constant $C>0$ such that for all $\varepsilon\in[0,1]$,

$$e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(x,y)}\le C,\ \mu(dx)\otimes\nu(dy)\ a.e.$$

and then

$$\left|(h^{\otimes}+\psi^{\nu})(\cdot,\cdot)e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(\cdot,\cdot)}\right|\le C(h^{\otimes}+\psi^{\nu})(\cdot,\cdot)\in\mathbb{L}^1(\mu\otimes\nu).$$

Then by the dominated convergence theorem, we know that g is C^1 on $[0,1]$, and using (2.3.3), we have $g'(0)>0$. Then there exists $\eta>0$ such that for $\varepsilon>0$ small enough, we have:

$$\mathbb{E}^{\mathbb{P}}\left[-e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(X,Y)}\right] > -1+\eta.$$

We then observe that $(x,y)\mapsto -e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(x,y)}$ is continuous by definition of ψ and using that μ has finite support, and bounded from above by 0 and from below since $(h^{\otimes}+\psi^{\nu})$ is bounded from below. Now since $\mathbb{P}_n$ converges weakly to $\mathbb{P}$, we have for n large enough that:

$$\mathbb{E}^{\mathbb{P}_n}\left[-e^{-\varepsilon(h^{\otimes}+\psi^{\nu})(X,Y)}\right] > -1+\frac{\eta}{2}.$$

i.e.

$$\begin{aligned}\mathbb{E}^{\mu\otimes\nu}\left[-e^{-\left((h_n+\varepsilon h)^{\otimes}+\phi_n^{\mu}\oplus(\psi_{\mu}^{*}(h_n,\phi_n)+\varepsilon\psi)^{\nu}\right)(X,Y)}\right] &> (-1+\frac{\eta}{2})\mathbb{E}^{\mu\otimes\nu}\left[e^{-(h_n^{\otimes}+\phi_n^{\mu}\oplus\psi^{*,\nu}(h_n,\phi_n))(X,Y)}\right]\\ &> (-1+\frac{\eta}{2})\mathbb{E}^{\mu\otimes\nu}\left[e^{-(h_n^{\otimes}+\phi_n^{\mu}\oplus\psi_n^{\nu})(X,Y)}\right].\end{aligned}$$

Since (h_n,ϕ_n,ψ_n) is a maximising sequence of the utility maximization problem, we obtain that for n large enough,

$$E^{\mu\otimes\nu}\left[-e^{-\left((h_n+\varepsilon h)^{\otimes}+\phi_n^{\mu}\oplus(\psi_{\mu}^{*}(h_n,\phi_n)+\varepsilon\psi)^{\nu}\right)(X,Y)}\right] > U(\mu,\nu)+\gamma \tag{2.3.4}$$

for some $\gamma>0$.

Consider n and $\gamma>0$ such that (2.3.4) holds. Now as in the proof of Proposition 2.2.1, we introduce $\psi_m^{*,n}:=\psi_{\mu}^{*}(h_n,\psi_n)\vee(\psi_n-m)$. We see that for all $m>0$, $(h_n+\varepsilon h,\phi_n+\varepsilon\phi,\psi_m^{*,n})$ is admissible. Indeed by construction and since μ has finite support and using that $\int|y|\nu(dy)<\infty$, we have $\psi_m^{*,n}\in\mathbb{L}^1(\nu)$, $\phi_n+\varepsilon\phi\in\mathbb{L}^1(\mu)$, and $(h_n+\varepsilon h)^{\otimes}\in\mathbb{L}^1(\mu\otimes\nu)$, so that $(h_n+\varepsilon h)^{\otimes}+(\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n}\in\mathbb{L}^1(\mu\otimes\nu)$. We recall from the construction of the proof of Lemma 2.2.1 that $(h_n,\phi_n,\psi_m^{*,n})$ is admissible so that

$$e^{-(h_n^{\otimes}+\phi_n\oplus\psi_m^{*,n})}\in\mathbb{L}^1(\mu\otimes\nu).$$

Then by construction we have $h^{\otimes} + \psi$ is bounded by below by some constant C, so that:

$$0 < e^{-((h_n+\varepsilon h)^{\otimes}+(\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n})} \le e^{-\varepsilon C} e^{-(h_n^{\otimes}+\phi_n\oplus\psi_m^{*,n})} \in \mathbb{L}^1(\mu\otimes\nu). \tag{2.3.5}$$

Similarly, we have by Jensen's inequality:

$$\int -\left((h_n+\varepsilon h)^{\otimes} + (\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n}\right)(x,\cdot)\mu(dx) \le \ln\left(\int e^{-((h_n+\varepsilon h)^{\otimes}+(\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n})(x,\cdot)}\mu(dx)\right),$$

and using 2.3.5, we have:

$$\ln\left(\int e^{-((h_n+\varepsilon h)^{\otimes}+(\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n})(x,\cdot)}\mu(dx)\right) \le \ln\left(\int e^{-(h_n^{\otimes}+\phi_n\oplus\psi_m^{*,n})(x,\cdot)}\mu(dx)\right) - \varepsilon C.$$

We then have that $\ln\left(\int e^{-((h_n+\varepsilon h)^{\otimes}+(\phi_n+\varepsilon\phi)\oplus\psi_m^{*,n})(x,\cdot)}\mu(dx)\right) \in \mathbb{L}^1(\nu)$ and then the triplet $(h_n + \varepsilon h, \phi_n+\varepsilon\phi, \psi_m^{*,n}) \in \mathbf{D}(\mu,\nu)$.

We then end the proof as in Lemma 2.2.1 by observing that

$$\left((h_n+\varepsilon h)^{\otimes} + (\phi_n+\varepsilon\phi)\oplus\psi_m\right)(x,y) \searrow \left((h_n+\varepsilon h)^{\otimes} + (\phi_n+\varepsilon\phi)\oplus\psi_\mu^*(h_n,\psi_n)\right)(x,y)$$

as $m \nearrow \infty$ for all (x,y), and $\nu(\psi_m) \to \nu(\psi_\mu^*(h_n,\psi_n))$ when $m\to\infty$, so that by dominated convergence theorem, we have:

$$\lim_{m\to\infty} J^{\mu,\nu}(h_n+\varepsilon h, \phi_n+\varepsilon\phi, \psi_m) = J^{\mu,\nu}(h_n+\varepsilon\phi, \phi_n+\varepsilon\phi, \psi_\mu^*(h_n,\psi_n) > V(\mu,\nu),$$

which is the required contradiction.

Step 2 : We now show that the first marginal law of $\mathbb{P}$ is μ. The proof is very similar to the corresponding one in Proposition 2.2.1. Assume to the contrary that we have ϕ continuous such that $\mathbb{E}^{\mathbb{P}}[\phi(X)] > \mu(\phi)$. Since ϕ is bounded on the support of μ, we have, similarly to step 1, by dominated convergence theorem that for $\varepsilon > 0$ small enough:

$$\mathbb{E}^{\mathbb{P}}\left[-e^{-\varepsilon\phi^{\mu}(X)}\right] > -1+\eta.$$

for some $\eta > 0$. Then since $\mathbb{P}_n$ converges weakly towards $\mathbb{P}$, and $x \mapsto -e^{-\varepsilon\phi^{\mu}(x)}$ is continuous and bounded on the support of μ, we have for n large enough that:

$$\mathbb{E}^{\mathbb{P}_n}\left[-e^{-\varepsilon\phi^{\mu}(X)}\right] > -1+\frac{\eta}{2}.$$

From here, the same corresponding construction made in Step 1 leads the same way to a contradiction, i.e. we find an admissible strategy $(\tilde{h},\tilde{\phi},\tilde{\psi})$ such that

$$J^{\mu,\nu}(\tilde{h},\tilde{\phi},\tilde{\psi}) > V(\mu,\nu),$$

which is the required contradiction.

□

2.4 The general result

We now focus on the proof of Theorem 2.2.1.

For a given measure μ on $\mathbb{R}^d$ with $\int |x|\mu(dx) < \infty$, we first construct a sequence $(\mu_n)_n$ with finite support such that $\mu_n \preceq \mu$ and μ_n converges weakly to μ. This result can be found in [10] in the 1-dimension case. We recall it and we provide an alternative proof for the sake of completeness:

Lemma 2.4.1. *Assume that* $\mu \in \mathcal{P}_{\mathbb{R}^d}$ *with* $\int |x|\mu(dx) < \infty$. *Then there exists a sequence* μ_n *with finite support such that* $\mu_n \preceq \mu$ *and* μ_n *converges weakly to* μ.

Proof. For notation simplification, we focus on the 1 dimension case. The general result can be obtained by a similar construction.
step 1: construction.
The construction is recursive. We define $x_1^1 := \int y\mu(dy)$ and:

$$\mu_1 = \delta_{x=x_1^1}.$$

Then, we introduce $x_1^2 = \frac{1}{p_1^2}\int_{-\infty}^{0} y\mu(dy)$, $x_2^2 = \frac{1}{p_2^2}\int_{0+}^{+\infty} y\mu(dy)$, where $p_1^2 = \int_{-\infty}^{0}\mu(dy)$ and $p_2^2 = \int_{0+}^{+\infty}\mu(dy)$, and:

$$\mu_2 = p_1^2\delta_{x=x_1^2} + p_2^2\delta_{x=x_2^2}.$$

Then similarly, we build $x_{-1}^n = \frac{1}{p_{-1}^n}\int_{-\infty}^{-n} y\mu(dy)$, $x_k^n = \frac{1}{p_k^n}\int_{t_k^n+}^{t_{k+1}^n} y\mu(dy)$ and $p_k^n = \int_{t_k^n+}^{t_{k+1}^n}\mu(dy)$ for $k \in \{0...2n^2-1\}$ and $t_k^n = -n + \frac{k}{n}$, and $x_{2n^2}^n = \frac{1}{p_{2n^2}^n}\int_{t_{2n^2}^n+}^{+\infty} y\mu(dy)$:

$$\mu_n = \sum_{k=-1}^{2n^2} p_k^n\delta_{x=x_k^n}.$$

Step 2: $\mu_n \preceq \mu$.
Consider a convex function g, then we have:

$$\mu(g) = \int_{\mathbb{R}} g(x)\mu(dx) = \int_{-\infty}^{-n} g(x)\mu(dx) + \sum_{k=0}^{2n^2-1}\int_{t_k^n+}^{t_{k+1}^n} g(x)\mu(dx) + \int_{n+}^{+\infty} g(x)\mu(dx).$$

For $k \in \{0,...,2n^2-1\}$, If $p_k^n \neq 0$, by Jensen's inequality we have:

$$\begin{aligned}\int_{t_k^n+}^{t_{k+1}^n} g(x)\mu(dx) &= p_k^n\int_{t_k^n+}^{t_{k+1}^n} g(x)\frac{\mu(dx)}{p_k^n} \\ &\geq p_k^n g(\int_{t_k^n+}^{t_{k+1}^n} x\frac{\mu(dx)}{p_k^n}) = p_k^n g(x_k^n).\end{aligned}$$

Notice that this inequality holds also trivially if $p_k^n = 0$. By the same argument we obtain that:

$$\int_{-\infty}^{-n} g(x)\mu(dx) \geq p_{-1}^n g(x_{-1}^n),$$

and:

$$\int_{n+}^{\infty} g(x)\mu(dx) \geq p_{2n^2}^n g(x_{2n^2}^n).$$

Suming up all these inequality, we obtain:

$$\mu(g) \geq \sum_{k=-1}^{2n^2} p_k^n g(x_k^n) = \mu_n(g).$$

Step 3: μ_n *converges weakly to* μ.
Consider a continuous bounded real-valued function f. Denote by B its bound. Let $\varepsilon > 0$, $N \in \mathbb{N}$

such that $\mu((-\infty,-N]\cup(N,\infty))\le\varepsilon$, and δ be the modulus of uniform continuity of $f|_{[-N,N]}$. Then we have for $n\ge N$:

$$|\mu(f)-\mu_n(f)| \le 2B\varepsilon + \sum_{t_k^n\in[-N,N)} |\int_{t_k^n+}^{t_{k+1}^n} f(x)\mu(dx) - p_k^n f(x_k^n)|$$
$$\le 2B\varepsilon + 2\delta(\frac{1}{n}) \longrightarrow_{n\to+\infty} 2B\varepsilon,$$

which shows that $\mu_n(f)\to\mu(f)$ when $n\to\infty$.

□

Remark 2.4.1. *We can observe that the sequence built here is not necessary increasing for the convex order. A slightly different construction would have provided such a sequence.*

Remark 2.4.2. *An alternative proof of this result can be found in Lemma 5.18 p128 of [73] in the 1-dimensional case with the additionnal condition $\int x^2\mu(dx)<\infty$, by using optimal stopping techniques. Still in dimension 1, we also mention the construction in [7] that build finite support measures by preserving convex order.*

Lemma 2.4.2. *Assume that $\mu\in\mathcal{P}_{\mathbb{R}^d}$ with $\int|x|\mu(dx)<\infty$. Let $(\mu_n)_{n\ge0}$ be the sequence constructed in Lemma 2.4.1 and $(\mathbb{P}^n)$ be a sequence in $\mathcal{P}_{\mathbb{R}^d}$ such that $\mathbb{P}^n\in\mathcal{M}(\mu_n,\nu)$ for any $n\in\mathbb{N}$. Then the sequence $(\mathbb{P}^n)$ converges up to extraction to some $\mathbb{P}\in\mathcal{M}(\mu,\nu)$.*

Proof. Given a sequence $(\mathbb{P}^n)_{n\ge0}$ such that for all n, $\mathbb{P}^n\in\mathcal{M}(\mu_n,\nu)$, we show that the sequence is tight. Indeed, If we consider the compact sets $K_m=\bar{B}(0,m)$ the closed ball of radius m centered in 0 of $\mathbb{R}^d$, we have that:

$$\mathbb{P}^n\left[(X,Y)\notin K_m^2\right]\le\mu_n\left[X\notin K_m\right]+\nu\left[Y\notin K_m\right].$$

For m large enough, we have that for all n, $\mu_n\left[X\notin K_m\right]\le\varepsilon$, and $\nu\left[Y\notin K_m\right]\le\varepsilon$, so that the sequence is tight. By Prokhorov's theorem, it converges weakly up to extraction to some $\mathbb{P}$.

Now consider $\phi\in C_b^0(\mathbb{R}^d,\mathbb{R})$, we have that:

$$\mathbb{E}^{\mathbb{P}}\left[\phi(X)\right]=\lim_{n\to+\infty}\mathbb{E}^{\mathbb{P}^n}\left[\phi(X)\right]=\lim_{n\to+\infty}\mathbb{E}^{\mu_n}\left[X\right]=\mathbb{E}^{\mu}\left[\phi(X)\right],$$

where the last inequality follows from the weak convergence of μ_n towards μ, and then we have $\mathbb{E}^{\mathbb{P}}\left[\phi(X)\right]=\mathbb{E}^{\mu}\left[\phi(X)\right]$ for all $\phi\in C_b^0(\mathbb{R}^d,\mathbb{R})$, which ensures that the first marginal law of $\mathbb{P}$ is μ. Similarly, we obtain that the second marginal law of $\mathbb{P}$ is ν, so that $\mathbb{P}\in\mathcal{P}(\mu,\nu)$. It remains to prove the martingale property.
We want to show that for $h\in C_b^0(\mathbb{R}^d,\mathbb{R})$, we have:

$$\lim_{n\to+\infty}\mathbb{E}^{\mathbb{P}^n}\left[h(X)(Y-X)\right]=\mathbb{E}^{\mathbb{P}}\left[h(X)(Y-X)\right].$$

By the Skorohod representation theorem, we have that there exists a probability space Ω, and random variables $(X_n,Y_n)_{n\ge0}$ and (X_∞,Y_∞), such that (X_n,Y_n) has law $\mathbb{P}^n$, (X_∞,Y_∞) has law $\mathbb{P}$, $(X_n,Y_n)\to(X_\infty,Y_\infty)$ a.s.. We now need to show that $\mathbb{E}\left[h(X_n)(Y_n-X_n)\right]\to\mathbb{E}\left[h(X_\infty)(Y_\infty-X_\infty)\right]$ when $n\to\infty$.
Now since $X_n\sim\mu_n$, μ_n converges weakly to μ, and $Y_n\sim\nu$, we have that the sequence $|h(X_n)(Y_n-X_n)|$ is uniformly integrable, so that $h(X_n)(Y_n-X_n)\to h(X_\infty)(Y_\infty-X_\infty)$ in $\mathbb{L}^1$, and so:

$$\mathbb{E}^{\mathbb{P}}\left[h(X)(X-Y)\right]=0.$$

□

We now sum up the proof of Strassen's Theorem:

Proof of Theorem 2.2.1 Consider μ and ν such that $\int |x|\mu(dx) + \int |y|\nu(dy) < \infty$ and such that $\mu \preceq \nu$. Then we consider the sequence $(\mu_n)_{n\geq 0}$ constructed in Lemma 2.4.1. By Theorem 2.3.1 we know that for any $n \in \mathbb{N}$, $\mathcal{M}(\mu_n,\nu) \neq \emptyset$. Finally considering a sequence $(\mathbb{P}_n)$ such that for all n, $\mathbb{P}_n \in \mathcal{M}(\mu_n,\nu)$. Then by Lemma 2.4.2, we have that up to extraction, $\mathbb{P}_n$ converges weakly to some $\mathbb{P} \in \mathcal{M}(\mu_n,\nu)$ and then

$$\mathcal{M}(\mu,\nu) \neq \emptyset.$$

□

2.5 An algorithm to build an element of $\mathcal{M}(\mu,\nu)$

In this section we discuss the above proof in order to set a formal construction to build an element of $\mathcal{M}(\mu,\nu)$. Such explicit constructions exist and we refer for example to the thesis of D. Baker [7] for a construction based on linear programming.

We will stay formal in order to derive the algorithm. In all the following we will consider measures μ and ν on $\mathbb{R}$, such that $\mu \preceq \nu$. In addition to this assumption, we will consider that μ and ν are with finite support. This last assumption is not restrictive since we can use approximations preserving convex order. For example we consider the $\mathcal{U}$-quantization introduced in [7] and we will recall its property in the Appendix.

2.5.1 A complement to Section 2.2.3

From Section 2.2.3 we know that we can consider an optimal investment strategy ψ^* for the utility maximization problem. This particular property can be extended to the two other components of the strategy h and ϕ, under additionnal technical restrictions similar to the one imposed for the admissibility of ψ^*.

Indeed for $(h,\psi) \in \mathbb{L}^0(\mathbb{R}) \times \mathbb{L}^1(\mu)$, we define:

$$\phi^*_\nu(h,\psi)(x) := \ln\left(\int e^{-\left(h^{\otimes}+\psi^{\nu}\right)(x,y)}\nu(dy)\right).$$

Then we have formally the equivalent of Lemma 2.2.1 and Proposition 2.2.1:

Proposition 2.5.1. *Let μ and ν be two measures on $\mathbb{R}$. For every $(h,\psi) \in \mathbf{D}(\mu,\nu)$, we have:*

$$\sup_{\phi \ s.t. \ (h,\phi,\psi)\in\mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h,\phi,\psi) \ = \ J^{\mu,\nu}(h,\phi^*_\nu(h,\psi),\psi) = -e^{\mu(\phi^*_\nu(h,\psi))} < 0.$$

Moreover $\mathbb{P}^{h,\phi^*_\nu(h,\psi),\psi}$ has for first marginal law μ.

In the same way we are able to derive the existence of a maximizing strategy h^*. Indeed, we observe that:

$$\int\int -e^{-(h^{\otimes}+\phi^{\mu}\oplus\psi^{\nu})(x,y)}\mu(dx)\nu(dy) = \int\left(\int -e^{-h(x)(y-x)-\psi^{\nu}(y)}\nu(dy)\right)e^{-\phi^{\mu}(x)}\mu(dx),$$

so that for every $x \in \mathbb{R}$, the strategy $h(x)$ can be improved by taking $h^*(x)$ where $h^*(x)$ is a maximizer of $f^x : \mathbb{R} \to \mathbb{R}$, where:

$$f_x : h \mapsto \int -e^{-h(y-x)-\psi^\nu(y)}\nu(dy).$$

This application is concave. Indeed, we have formally that:

$$f'_x(h) = \int (y-x)e^{-h(x)(y-x)-\psi^\nu(y)}\nu(dy) \tag{2.5.1}$$

$$f''_x(h) = -\int (y-x)^2 e^{-h(x)(y-x)-\psi^\nu(y)}\nu(dy) \tag{2.5.2}$$

Then under suitable assumptions on the supports of μ and ν we have the existence of a finite maximizer h^*. Then formally we obtain:

Proposition 2.5.2. *Let μ and ν be two measures on $\mathbb{R}$. For every $(\phi,\psi) \in \mathbf{D}(\mu,\nu)$, we have the existence of a measurable and μ-almost everywhere finite $h^*_\nu(\psi)$ such that:*

$$\sup_{h \text{ s.t. } (h,\phi,\psi)\in\mathbf{D}(\mu,\nu)} J^{\mu,\nu}(h,\phi,\psi) \;=\; J^{\mu,\nu}(h^*_\nu(\psi),\phi,\psi).$$

Moreover $\mathbb{P}^{h^*_\nu(\psi),\phi,\psi}$ is a martingale law.

Remark 2.5.1. *Observe that by construction the maximizer h^* should only be a function of ν and ψ, and does not depend on ϕ. Hence it can be computed as the zero of the function f'_x.*

2.5.2 Algorithm and numerical results

We now describe the general algorithm. We start from two measures μ and ν with finite support such that $\mu \preceq \nu$.

The algorithm is basically made of an approximation of $V(\mu,\nu)$ computed in n steps, where n is chosen large enough to get convergence. The idea is to maximise recursively the strategies h, ϕ and ψ by replacing them with the values h^*, ϕ^* and ψ^* defined in Sections 2.2.3 and 2.5.1:
We start with the triplet $(h_0,\phi_0,\psi_0) = (0,0,0)$. Then at step i, we compute $\psi_i = \psi^*_\mu(h_{i-1},\phi_{i-1})$, $h_i = h^*_\nu(\psi_i)$ and finally $\phi_i = \phi^*_\nu(h_i,\psi_i)$.
When this is acceptable, we compute the related probability law $\mathbb{P}^{h_n,\phi_n,\psi_n}$.

Algorithm 1 Caculate an approximation of $V(\mu,\nu)$

```
h ← 0
φ ← 0
ψ ← 0
for i = 1 → n do
    ψ ← ψ*_μ(h, φ)
    h ← h*_ν(ψ)
    φ ← φ*_ν(h, ψ)
V^approx ← J^{μ,ν}(h, φ, ψ)
P^approx = P^{h,φ,ψ}
```

This algorithm may deserve some explanations. The first one is that the sequence $(h_k, \phi_k, \phi_k)_{k\geq 0}$ is not necessarily a maximising sequence. This is not a problem here since by definition, we have $(J^{\mu,\nu}(h_k, \phi_k, \psi_k))$ is a non decreasing sequence, bounded by above by $V(\mu,\nu) < 0$. Then by similar arguments of those developed in Section 2.3, since we consider μ and ν with finite support, any converging sub-sequence of $(\mathbb{P}^{h_k,\phi_k,\psi_k})$ converges weakly to an element of $\mathcal{M}(\mu,\nu)$. Hence even if from n to $n+1$, $\mathbb{P}^{h_n,\phi_n,\psi_n}$ may differ from $\mathbb{P}^{h_{n+1},\phi_{n+1},\psi_{n+1}}$, they both should be elements (or close to elements) of $\mathcal{M}(\mu,\nu)$.
One should also notice that $h^*_\nu(\psi)$ is not uniquely defined by Proposition 2.5.2. But as we already pointed in Remark 2.5.1, we can compute an approximated value for h^* by looking for a zero of an increasing function.

Remark 2.5.2. *One should also notice that for any $k \geq 1$, given the order of maximization that we considered at step k, $\mathbb{P}^{h_k,\phi_k,\psi_k}$ is already a martingale law, and the first marginal law is μ, so that for the final test, one should only check if the second marginal law of $\mathbb{P}^{h_n,\phi_n,\psi_n}$ is ν*

We now present numerical results obtained with this algorithm. We considered μ and ν the following measures obtained in [7] by $\mathcal{U}$-quantization of $\mathcal{N}(0,1)$ and $\mathcal{N}(0,2)$ (the normal distributions of mean 0 and variance 1 and 2):

$$\mu := \frac{1}{10}\sum_{i=1}^{10} \delta_{a_i},$$

$$\nu := \frac{1}{10}\sum_{i=1}^{10} \delta_{b_i},$$

where:

$$\begin{pmatrix} a_1\\ a_2\\ a_3\\ a_4\\ a_5\\ a_6\\ a_7\\ a_8\\ a_9\\ a_{10}\end{pmatrix} = \begin{pmatrix} -1.75498\\ -1.04464\\ -0.67731\\ -0.38650\\ -0.12600\\ 0.12600\\ 0.38650\\ 0.67731\\ 1.04464\\ 1.75498\end{pmatrix}, \qquad \begin{pmatrix} b_1\\ b_2\\ b_3\\ b_4\\ b_5\\ b_6\\ b_7\\ b_8\\ b_9\\ b_{10}\end{pmatrix} = \begin{pmatrix} -2.48192\\ -1.47734\\ -0.95786\\ -0.54659\\ -0.17819\\ 0.17819\\ 0.54659\\ 0.95786\\ 1.47734\\ 2.48192\end{pmatrix}$$

The strategy $(h_{30}, \phi_{30}, \psi_{30})$ obtained is:

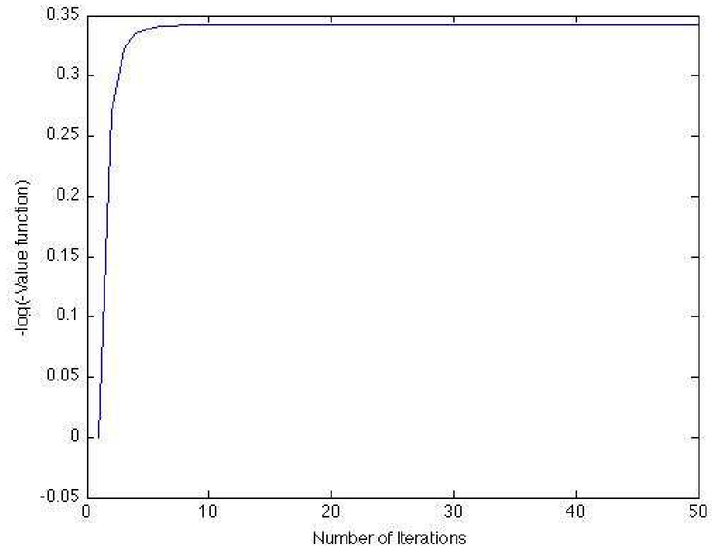

Figure 2.1: Convergence of $J^{\mu,\nu}(h_n, \phi_n, \psi_n)$.

$$\begin{pmatrix} h_{30}(a_1) \\ h_{30}(a_2) \\ h_{30}(a_3) \\ h_{30}(a_4) \\ h_{30}(a_5) \\ h_{30}(a_6) \\ h_{30}(a_7) \\ h_{30}(a_8) \\ h_{30}(a_9) \\ h_{30}(a_{10}) \end{pmatrix} = \begin{pmatrix} 1.8188 \\ 1.0147 \\ 0.6554 \\ 0.3731 \\ 0.1183 \\ -0.1289 \\ -0.3838 \\ -0.6668 \\ -1.0262 \\ -1.8303 \end{pmatrix}, \quad \begin{pmatrix} \phi_{30}(a_1) \\ \phi_{30}(a_2) \\ \phi_{30}(a_3) \\ \phi_{30}(a_4) \\ \phi_{30}(a_5) \\ \phi_{30}(a_6) \\ \phi_{30}(a_7) \\ \phi_{30}(a_8) \\ \phi_{30}(a_9) \\ \phi_{30}(a_{10}) \end{pmatrix} = \begin{pmatrix} -1.3873 \\ -0.3963 \\ -0.0897 \\ 0.0599 \\ 0.1240 \\ 0.1227 \\ 0.0559 \\ -0.0968 \\ -0.4072 \\ -1.4058 \end{pmatrix}, \quad \begin{pmatrix} \psi_{30}(b_1) \\ \psi_{30}(b_2) \\ \psi_{30}(b_3) \\ \psi_{30}(b_4) \\ \psi_{30}(b_5) \\ \psi_{30}(b_6) \\ \psi_{30}(b_7) \\ \psi_{30}(b_8) \\ \psi_{30}(b_9) \\ \psi_{30}(b_{10}) \end{pmatrix} = \begin{pmatrix} 0.7344 \\ -0.2846 \\ -0.6059 \\ -0.7595 \\ -0.8241 \\ -0.8223 \\ -0.7539 \\ -0.5959 \\ -0.2692 \\ 0.7606 \end{pmatrix}$$

The probability law $\mathbb{P}^{h_{30},\phi_{30},\psi_{30}}$ is given by the matrix $(M(ai, b_j))_{1\le i,j\le 10}$ where $M(a_i, b_j) := \mathbb{P}^{h_{30},\phi_{30},\psi_{30}}[(a_i, b_j)]$ where:

$$M = \begin{pmatrix} 0.0512 & 0.0228 & 0.0122 & 0.0067 & 0.0037 & 0.0019 & 0.0009 & 0.0004 & 0.0001 & 0.0000 \\ 0.0218 & 0.0218 & 0.0177 & 0.0136 & 0.0100 & 0.0069 & 0.0045 & 0.0025 & 0.0011 & 0.0001 \\ 0.0122 & 0.0174 & 0.0171 & 0.0152 & 0.0128 & 0.0101 & 0.0074 & 0.0048 & 0.0025 & 0.0005 \\ 0.0070 & 0.0134 & 0.0152 & 0.0152 & 0.0141 & 0.0123 & 0.0100 & 0.0074 & 0.0044 & 0.0011 \\ 0.0040 & 0.0098 & 0.0127 & 0.0141 & 0.0144 & 0.0138 & 0.0123 & 0.0100 & 0.0068 & 0.0022 \\ 0.0022 & 0.0068 & 0.0100 & 0.0123 & 0.0138 & 0.0144 & 0.0141 & 0.0127 & 0.0098 & 0.0040 \\ 0.0011 & 0.0044 & 0.0073 & 0.0100 & 0.0123 & 0.0141 & 0.0152 & 0.0152 & 0.0134 & 0.0070 \\ 0.0005 & 0.0025 & 0.0048 & 0.0074 & 0.0101 & 0.0128 & 0.0152 & 0.0171 & 0.0175 & 0.0122 \\ 0.0001 & 0.0011 & 0.0025 & 0.0045 & 0.0069 & 0.0100 & 0.0136 & 0.0177 & 0.0218 & 0.0218 \\ 0.0000 & 0.0001 & 0.0004 & 0.0009 & 0.0019 & 0.0037 & 0.0067 & 0.0122 & 0.0228 & 0.0512 \end{pmatrix}$$

For the sake of completeness, we also consider a slight modification $\tilde{\mu}$ of the previous measure μ. $\tilde{\mu}$ is characterized as above by the vector:

$$(-1.75498, -1.04464, -0.67731, -0.38650, -0.12600, 0.18600, 0.38650, 0.67731, 1.04464, 1.75498),$$

where we replaced the sixth value of μ by 0.186.

The divergence of the logarithm of the value function, corresponding to the convergence of the value function towards 0, appears to be linear, but quite slow.

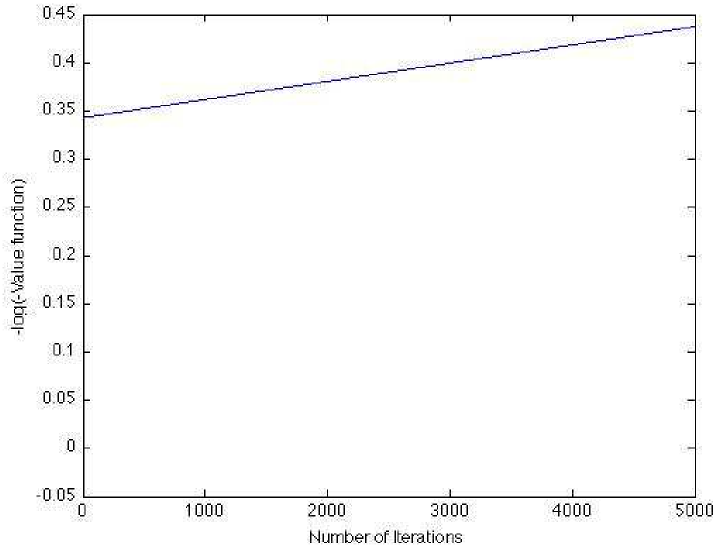

Figure 2.2: Divergence of $J^{\mu,\nu}(h_n,\phi_n,\psi_n)$ for $\tilde{\mu}$ and ν.

Appendix: $\mathcal{U}$-quantization of measures on $\mathbb{R}$

The $\mathcal{U}$-quantization is defined via the quantile function of a measure. For $\mu \in \mathcal{P}(\mathbb{R})$, we denote by F_μ its cumulative distribution function and F_μ^{-1} its generalize inverse, i.e. the function:

$$p \in [0,1] \mapsto F_\mu^{-1}(p) := \inf\big\{x \in \mathbb{R} : p \le F_\mu(x)\big\}.$$

Then we define:

Definition 2.5.1 ($\mathcal{U}$-quantization)**.** *For $n \in \mathbb{N}^*$ and $\mu \in \mathcal{P}(\mathbb{R})$, we define the $\mathcal{U}$-quantization of μ by the measure $U(a_1^\mu, ..., a_n^\mu)$ where:*

$$\begin{cases} U(a_1^\mu, ..., a_n^\mu) = \frac{1}{n}\sum_{i=1}^n \delta_{a_i^\mu}, \\ \quad where \ \ a_i^\mu := n\int_{\frac{i-1}{n}}^{i/n} F^{-1}(u)du. \end{cases}$$

Then we have:

Lemma 2.5.1 (Lemma 2.4.9 in [7])**.** *The $\mathcal{U}$-quantization preserves the convex order, i.e. for μ and ν in $\mathcal{P}(\mathbb{R})$ such that $\mu \preceq \nu$, then we have for all $n \in \mathcal{N}^*$*

$$U(a_1^\mu, ..., a_n^\mu) \preceq U(a_1^\nu, ..., a_n^\nu).$$

Chapter 3

On the robust super-hedging of measurable claims

Contents

3.1 Introduction

An important attention is focused on the problem of robust superhedging in the recent literature. Motivated by the original works of Avellaneda [6] and Lyons [69], the first general formulation of this problem was introduced by Denis and Martini [39] by considering the hedging problem under a non-dominated family of probability measures on the canonical space of continuous trajectories. Since the hedging problem involves stochastic integration, [39] used the capacity theory to develop the corresponding quasi-sure stochastic analysis tools, i.e. stochastic analysis results holding simultaneously under the considered family of non-dominated measures.

The next progress was achieved by Soner, Touzi and Zhang [91] who introduced a restriction of the set of non-dominated measures so as to guarantee that the predictable representation property holds true under each measure. However, [91] placed strong regularity conditions on the random variables of interest in order to guarantee the measurability of the value function of some dynamic version of a stochastic control problem, and to derive the corresponding dynamic programming principle.

By using the notion of measurable analyticity, Nutz and van Handel [77] and Neufeld and Nutz [74] extended the previous results to general measurable claims by introducing some conditions that the non-dominated family of singular measures must satisfy.

The main objective of this chapter is to extend the approach of Neufeld and Nutz [74] so as to introduce some specific additional constraints on the family of probability measures, and to weaken the integrability condition on the random variables of interest. Such an extension is crucially needed in the recent problem of martingale transportation problem [46, 51], where the superhedging problem allows for the static trading of any Vanilla payoff in addition to the dynamic trading of the underlying risky asset. Assuming that the financial market, with this enlarged possibilities of trading, satisfies the no-arbitrage condition leads essentially to the restriction of the family of probability measures to those under which the canonical process is a uniformly integrable martingale. The main problem is that this restriction violates the conditions of [74] on one hand, and that the integrability conditions in [74] are not convenient for the stochastic control approach of [46, 51].

The chapter is organized as follows. Section 3.2 introduces the main probabilistic framework. The robust superhedging problem is formulated in Section 3.3, where we also report our main result, together with the comparison to [74]. Section 3.4 contains the proof of the duality result. Finally, some extensions are reported in Section 3.5.

3.2 Preliminaries

3.2.1 Probabilistic framework

Let $\Omega := \{\omega \in C\left([0,T], \mathbb{R}^d\right) : \omega_0 = 0\}$ be the canonical space equipped with the uniform norm $||\omega||_\infty^T := \sup_{0\le t\le T} |\omega_t|$. $\mathcal{F}$ will always be a fixed σ-field on Ω which contains all our filtrations. We then denote B the canonical process, $\mathbb{P}_0$ the Wiener measure, $\mathbb{F} := \{\mathcal{F}_t\}_{0\le t\le T}$ the filtration generated by B and $\mathbb{F}^+ := \{\mathcal{F}_t^+, 0 \le t \le T\}$, the right limit of $\mathbb{F}$ where $\mathcal{F}_t^+ := \cap_{s>t}\mathcal{F}$. We will denote by $\mathbf{M}(\Omega)$ the set of all probability measures on Ω. We also recall the so-called universal filtration $\mathbb{F}^* := \{\mathcal{F}_t^*\}_{0\le t\le T}$ defined as follows

$$\mathcal{F}_t^* := \bigcap_{\mathbb{P}\in\mathbf{M}(\Omega)} \mathcal{F}_t^{\mathbb{P}},$$

where $\mathcal{F}_t^{\mathbb{P}}$ is the usual completion under $\mathbb{P}$.

For any subset E of a finite dimensional space and any filtration $\mathbb{X}$ on $(\Omega, \mathcal{F})$, we denote by $\mathbb{H}^0(E, \mathbb{X})$ the set of all $\mathbb{X}$-progressively measurable processes with values in E. Moreover for all $p > 0$ and for all $\mathbb{P} \in \mathbf{M}(\Omega)$, we denote by $\mathbb{H}^p(\mathbb{P}, E, \mathbb{X})$ the subset of $\mathbb{H}^0(E, \mathbb{X})$ whose elements H satisfy $\mathbb{E}^{\mathbb{P}}\left[\int_0^T |H_t|^p\, dt\right] < +\infty$. The localized versions of these spaces are denoted by $\mathbb{H}^p_{\text{loc}}(\mathbb{P}, E, \mathbb{X})$.

For any subset $\mathcal{P} \subset \mathbf{M}(\Omega)$, a $\mathcal{P}-$polar set is a $\mathbb{P}-$negligible set for all $\mathbb{P} \in \mathcal{P}$, and we say that a property holds $\mathcal{P}-$quasi-surely if it holds outside of a $\mathcal{P}-$polar set. Finally, we introduce the following filtration $\mathbb{G}^{\mathcal{P}} := \{\mathcal{G}_t^{\mathcal{P}}\}_{0\le t\le T}$ which will be useful in the sequel

$$\mathcal{G}_t^{\mathcal{P}} := \mathcal{F}_{t+}^* \vee \mathcal{N}^{\mathcal{P}},\ t < T \text{ and } \mathcal{G}_T^{\mathcal{P}} := \mathcal{F}_T^* \vee \mathcal{N}^{\mathcal{P}},$$

where $\mathcal{N}^{\mathcal{P}}$ is the collection of $\mathcal{P}$-polar sets.

For all $\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F})$, where $\mathbb{S}_d^{>0}$ is the set of positive definite matrices of size $d \times d$, we define the probability measure on $(\Omega, \mathcal{F})$

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

We denote by $\mathcal{M}^S_{\text{loc}}$ the collection of all such probability measures on $(\Omega, \mathcal{F})$. We recall from Karandikar [65] that the quadratic variation process $\langle B\rangle$ is universally defined under any $\mathbb{P} \in$

$\mathcal{M}^S_{\text{loc}}$, and takes values in the set of all nondecreasing continuous functions from $\mathbb{R}_+$ to $\mathbb{S}_d^{>0}$. We will denote its pathwise density with respect to the Lebesgue measure by $\widehat{a}$. Finally we recall from [92] that every $\mathbb{P} \in \mathcal{M}^S_{\text{loc}}$ satisfies the Blumenthal zero-one law and the martingale representation property.

Our focus in this paper will be on the following subset of $\mathcal{M}^S_{\text{loc}}$.

Definition 3.2.1. *$\mathcal{M}^S$ is the sub-class of $\mathcal{M}^S_{loc}$ consisting of all $\mathbb{P} \in \mathcal{M}^S_{loc}$ such that the canonical process B is a $\mathbb{P}-$uniformly integrable martingale.*

3.2.2 Regular conditional probability distributions

In this section, we recall the notion of regular conditional probability distribution (r.c.p.d.), as introduced by Stroock and Varadhan [95]. Let $\mathbb{P} \in \mathbf{M}(\Omega)$ and consider some $\mathbb{F}$-stopping time τ. Then, for every $\omega \in \Omega$, there exists an r.c.p.d. $\mathbb{P}^\omega_\tau$ satisfying:
(i) $\mathbb{P}^\omega_\tau$ is a probability measure on $\mathcal{F}_T$.
(ii) For each $E \in \mathcal{F}_T$, the mapping $\omega \to \mathbb{P}^\omega_\tau(E)$ is $\mathcal{F}_\tau$-measurable.
(iii) $\mathbb{P}^\omega_\tau$ is a version of the conditional probability measure of $\mathbb{P}$ on $\mathcal{F}_\tau$, i.e., for every integrable $\mathcal{F}_T$-measurable r.v. ξ we have $\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_\tau](\omega) = \mathbb{E}^{\mathbb{P}^\omega_\tau}[\xi]$, $\mathbb{P}-$a.s.
(iv) $\mathbb{P}^\omega_\tau(\Omega^\omega_\tau) = 1$, where $\Omega^\omega_\tau := \{\omega' \in \Omega : \ \omega'(s) = \omega(s), \ 0 \le s \le \tau(\omega)\}$.

We next introduce the shifted canonical space and the corresponding notations.

• For $0 \le t \le T$, denote by $\Omega^t := \{\omega \in C([t,T],\mathbb{R}) \ : \ w(t) = 0\}$ the shifted canonical space, B^t the shifted canonical process on Ω^t, $\mathbb{P}^t_0$ the shifted wiener measure, $\mathbb{F}^t$ the shifted filtration generated by B^t.

• For $0 \le s \le t \le T$, $\omega \in \Omega^s$, define the shifted path $\omega^t \in \Omega^t$, $\omega^t_r := \omega_r - \omega_t$ for all $r \in [t,T]$.

• For $0 \le s \le t \le T$, $\omega \in \Omega^s$, define the concatenation path $\omega \otimes_t \widetilde{\omega} \in \Omega^s$ by:

$$(\omega \otimes_t \widetilde{\omega})(r) := \omega_r \mathbf{1}_{[s,t)}(r) + (\omega_t + \widetilde{\omega}_r)\mathbf{1}_{[t,1]}(r) \ \text{ for all } \ r \in [s,T].$$

• For $0 \le s \le t \le T$, for any $\mathcal{F}^s_T$-measurable random variable ξ on Ω^s, and for each $\omega \in \Omega^s$, define the shifted $\mathcal{F}^t_T$-measurable random variable $\xi^{t,\omega}$ on Ω^t by:

$$\xi^{t,\omega}(\widetilde{\omega}) := \xi(\omega \otimes_t \widetilde{\omega}) \ \text{ for all } \ \widetilde{\omega} \in \Omega^t.$$

• The r.c.p.d. $\mathbb{P}^\omega_\tau$ induces naturally a probability measure $\mathbb{P}^{\tau,\omega}$ on $\mathcal{F}^{\tau(\omega)}_T$ such that the $\mathbb{P}^{\tau,\omega}$-distribution of $B^{\tau(\omega)}$ is equal to the $\mathbb{P}^\omega_\tau$-distribution of $\{B_t - B_{\tau(\omega)}, \ t \in [\tau(\omega), T)\}$. It is then clear that for every integrable and $\mathcal{F}_T$-measurable random variable ξ,

$$\mathbb{E}^{\mathbb{P}^\omega_\tau}[\xi] = \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\xi^{\tau,\omega}].$$

For the sake of simplicity, we shall also call $\mathbb{P}^{\tau,\omega}$ the r.c.p.d. of $\mathbb{P}$.

• Finally, we introduce for all $(s,\omega) \in [0,T] \times \Omega$:

$$\mathcal{M}^S_{\text{loc}}(s,\omega) := \left\{ \mathbb{P}^s_0 \circ \Big(\int_s^\cdot \alpha_u^{1/2} dB^s_u \Big)^{-1}, \text{ with } \int_s^T |\alpha_u| du < +\infty, \ \mathbb{P}^s_0 - \text{a.s.} \right\}$$
$$\mathcal{M}^S(s,\omega) := \left\{ \mathbb{P} \in \mathcal{M}^S_{\text{loc}}(s,\omega) \text{ s.t. } B^s \text{ is a uniformly integrable martingale} \right\}.$$

Remark 3.2.1. *We are abusing notations here. In order to suit to the definition of [74] and [77], we should have considered the concatenation of $\mathcal{M}^S_{loc}(s,\omega)$ (resp. $\mathcal{M}^S(s,\omega)$) defined above with the dirac mass on $\omega_{0\le t\le s}$ to ensure that elements of $\mathcal{M}^S_{loc}(s,\omega)$ (resp. $\mathcal{M}^S(s,\omega)$) are probabilities on Ω, and not on Ω^s. The reader should note that the link between these two definition is obvious and we will implicitly identify these families.*

It is clear that the families $(\mathcal{M}^S_{\text{loc}}(s,\omega))_{(s,\omega)\in[0,T]\times\Omega}$ and $(\mathcal{M}^S(s,\omega))_{(s,\omega)\in[0,T]\times\Omega}$ are adapted in the sense that $\mathcal{M}^S_{\text{loc}}(s,\omega) = \mathcal{M}^S_{\text{loc}}(s,\widetilde{\omega})$ and $\mathcal{M}^S(s,\omega) = \mathcal{M}^S(s,\widetilde{\omega})$, whenever $\omega|_{[0,s]} = \widetilde{\omega}|_{[0,s]}$.

3.3 Superreplication and duality

3.3.1 Problem formulation and main results

Throughout this paper, we consider some scalar $\mathcal{G}_T$-measurable random variable ξ. For any $(s,\omega)\in[0,T]\times\Omega$, we naturally restrict the subset $\mathcal{M}^S$ and $\mathcal{M}^S(s,\omega)$ to:

$$\mathcal{M}^S_\xi := \big\{\mathbb{P}\in\mathcal{M}^S:\ \mathbb{E}^{\mathbb{P}}[\xi^-]<+\infty\big\}$$
$$\mathcal{M}^S_\xi(s,\omega) := \big\{\mathbb{P}\in\mathcal{M}^S(s,\omega):\ \mathbb{E}^{\mathbb{P}}[(\xi^{s,\omega})^-]<+\infty\big\}.$$

Notice that such a restriction can be interpreted as suppressing measures which induce arbitrage opportunities in our market.

Our main interest is on the problem of superreplication under model uncertainty and the corresponding dual formulation. Given some initial capital X_0, the wealth process is:

$$X^H_t := X_0 + \int_0^t H_s dB_s,\ t\in[0,T],$$

where $H\in\mathcal{H}^\xi$, the set of admissible trading strategies defined by:

$$\mathcal{H}^\xi := \Big\{H\in\mathbb{H}^0(\mathbb{R}^d,\mathbb{G}^{\mathcal{M}^S})\cap\mathbb{H}^2_{\text{loc}}(\mathbb{P},\mathbb{R}^d,\mathbb{G}^{\mathcal{M}^S}),\ X^H \text{ is a } \mathbb{P}-\text{supermartingale}, \forall\,\mathbb{P}\in\mathcal{P}^\xi_m\Big\}.$$

The main result of this paper is the following.

Theorem 3.3.1. *Let ξ be an upper semi-analytic r.v. with $\sup_{\mathbb{P}\in\mathcal{M}^S}\mathbb{E}^{\mathbb{P}}[\xi^+]<+\infty$. Then*

$$V(\xi) \ :=\ \inf\Big\{X_0:\ X^H_T\ge\xi,\ \ \mathcal{M}^S_\xi-q.s.\ \textit{for some } H\in\mathcal{H}^\xi\Big\}\ =\ \sup_{\mathbb{P}\in\mathcal{M}^S}\mathbb{E}^{\mathbb{P}}[\xi].$$

Moreover, existence holds for the primal problem, i.e. $V(\xi)+\int_0^T H_s dB_s\ge\xi$, $\mathcal{M}^S_\xi-q.s.$ for some $H\in\mathcal{H}^\xi$.

Remark 3.3.1. *Suppose that the random variable ξ^- is $\mathbb{P}-$integrable for all $\mathbb{P}\in\mathcal{M}^S$. Then, $\mathcal{M}^S_\xi=\mathcal{M}^S$, and the corresponding set of admissible strategies $\mathcal{H}^\xi =: \mathcal{H}$ is independent of ξ. Under the condition $\sup_{\mathbb{P}\in\mathcal{M}^S}\mathbb{E}^{\mathbb{P}}[\xi^+]<\infty$, it follows from the previous theorem that:*

$$\inf\big\{X_0:\ X^H_T\ge\xi,\ \ \mathcal{M}^S-q.s.\ \textit{for some } H\in\mathcal{H}\big\}\ =\ \sup_{\mathbb{P}\in\mathcal{M}^S}\mathbb{E}^{\mathbb{P}}[\xi].$$

Remark 3.3.2. *When it comes to which filtration the trading strategies are admissible, we can actually do a little bit better than $\mathbb{G}^{\mathcal{M}^S}$, and consider the universal filtration $\mathbb{F}^*$ completed by the $\mathcal{M}^S$-polar sets, instead of its right limit. Indeed, let X be a process adapted to $\mathbb{G}^{\mathcal{M}^S}$, then following the arguments in Lemma* 2.4 *of [92], we define $\widetilde{X}$ by*

$$\widetilde{X}_t := \limsup_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\int_{t-\varepsilon}^t X_s ds.$$

Then, $\widetilde{X}$ coincides $dt\times\mathbb{P}-$a.e. with X, for any $\mathbb{P}\in\mathcal{M}^S_{loc}$ and is adapted to $\mathbb{F}^*$ completed by the $\mathcal{M}^S$-polar sets. For simplicity, we however refrain from considering this extension.[1]

The problem of superhedging under model uncertainty was first considered by Denis and Martini [37] using the theory of capacities and the quasi-sure analysis. The set of probability measures considered in [37] is larger than $\mathcal{M}^S_{\text{loc}}$, and whether existence of an optimal hedging strategy strategy holds or not in the framework of [37] is still an open problem. Later, Soner, Touzi and Zhang [92] considered the same problem but with a strict subset of $\mathcal{M}^S_{\text{loc}}$ satisfying a separability condition, which allowed them to recover the existence of an optimal strategy. The same approach is adapted in Galichon, Henry-Labordère and Touzi to obtain the duality result of Theorem 3.3.1 for uniformly continuous ξ. Recently, Neufeld and Nutz [74] introduced a new approach which avoids the strong regularity condition on ξ. In the next subsection, we briefly outline their approach and explain why it fails to cover our framework.

3.3.2 The analytic measurability approach

We now introduce the general framework of [77] and [74]. Let $\mathcal{P}$ be a non-empty subset of $\mathcal{M}^S_{\text{loc}}$, with corresponding "shifted" sets $\mathcal{P}(s,\omega)$, satisfying:

Condition 3.3.1. *Let $s\in\mathbb{R}_+$, τ a stopping time such that $\tau\geq s$, $\overline{\omega}\in\Omega$, and $\mathbb{P}\in\mathcal{P}(s,\overline{\omega})$. Set $\theta:=\tau^{s,\overline{\omega}}-s$.*
(i) *The graph $\{(\mathbb{P}',\omega):\omega\in\Omega,\ \mathbb{P}'\in\mathcal{P}(t,\omega)\}\subseteq\mathbf{M}(\Omega)\times\Omega$ is analytic.*
(ii) *We have $\mathbb{P}^{\theta,\omega}\in\mathcal{P}(\tau,\overline{\omega}\otimes_s\omega)$ for $\mathbb{P}$-a.e. $\omega\in\Omega$.*
(iii) *If $\nu:\Omega\to\mathcal{B}(\Omega)$ is an $\mathcal{F}_\theta$-measurable kernel and $\nu(\omega)\in\mathcal{P}(\tau,\overline{\omega}\otimes_s\omega)$ for $\mathbb{P}$-a.e. $\omega\in\Omega$, then the following measure $\overline{\mathbb{P}}\in\mathcal{P}(s,\overline{\omega})$:*

$$\overline{\mathbb{P}}(A) := \iint (\mathbf{1}_A)^{\theta,\omega}(\omega')\nu(d\omega';\omega)\mathbb{P}(dw),\quad A\in\mathcal{F}.$$

Theorem 3.3.2 (Theorem 2.3 in [74])**.** *Suppose $\{\mathcal{P}(s,\omega)\}_{(s,\omega)}$ satisfies Condition 3.3.1. Then, for any upper semi-analytic map ξ with $\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}^{\mathbb{P}}[|\xi|]<+\infty$, we have:*

$$\inf\left\{X_0:\ X_T^H\geq\xi\ \ \mathcal{P}-q.s.\ \text{for some}\ H\in\mathcal{H}\right\} \;=\; \sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}^{\mathbb{P}}[\xi].$$

For the purpose of the application of this result to the problem of martingale optimal transportation, see [46, 51], the last result presents two inconveniences:
- The integrability condition of the previous theorem from [74] turns out to be too strong. The weaker integrability conditions in our Theorem 3.3.1 is crucial for the analysis conducted in [46, 51].
- The set of probability measures of interest is the smaller subset $\mathcal{M}^S$. We shall verify below that

[1]We would like to thank Marcel Nutz for pointing this out.

$\mathcal{M}^S$ satisfies Conditions 3.3.1(i) and (ii), but fails to satisfy (iii). Therefore, we need to extend the results of [74] in order to address the case of $\mathcal{M}^S$.

Example 3.3.1. [$\mathcal{M}^S$ does not satisfy Condition 3.3.1 (iii)] *For simplicity, let $d = 1$. Let $s \in (0,T)$, $t \geq s$, $\overline{\omega} \in \Omega$ and $\mathbb{P} = \mathbb{P}_0^s \in \mathcal{P}_B(s, \overline{\omega})$. Now consider $\omega \in \Omega^s$. The family $(\mathbb{P}_i)_{i\in\mathbb{N}}$ is defined by*

$$\forall i \in \mathbb{N},\ \mathbb{P}_i = \mathbb{P}_0^t \circ \left(\int_0^\cdot \sigma_i^{1/2} dB_u^t\right)^{-1}.$$

where $(\sigma_i)_{i\in\mathbb{N}}$ is a sequence of positive numbers which will be chosen later. We consider the following partition $(E_i)_{i\in\mathbb{N}}$ of $\mathcal{F}_t$

$$\forall i \in \mathbb{N},\ E_i := \{\omega \text{ s.t. } \omega_t \in (-i-1, -i] \cup [i, i+1)\}.$$

We then introduce the $\mathcal{F}_t$-measurable kernel $\nu(\omega)(A) := \sum_{i=0}^{+\infty} \mathbf{1}_{E_i}(\omega)\mathbb{P}_i(A)$, and we define $\overline{\mathbb{P}}$ as in Condition 3.3.1(iii) from $\mathbb{P}$ and ν. We now show that $\mathbb{E}^{\overline{\mathbb{P}}}[|B_T|] = +\infty$ for some convenient choice of the sequence (σ^i). In particular this shows that $\overline{\mathbb{P}} \notin \mathcal{M}^S$.

$$\begin{aligned}
\mathbb{E}^{\overline{\mathbb{P}}}[|B_T|] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\overline{\mathbb{P}}}[|B_T||\mathcal{F}_t]\right] &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\overline{\mathbb{P}}^{t,\overline{\omega}\otimes_s\omega}}[|B_T|]\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\sum_{i=0}^{+\infty} \mathbf{1}_{E_i}(\omega)\mathbb{E}^{\mathbb{P}_i}\left[|B_T^{t,\overline{\omega}\otimes_s\omega}|\right]\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\sum_{i=0}^{+\infty} \mathbf{1}_{E_i}(\omega)\int_{-\infty}^{+\infty} |\sigma_i u + B_t(\omega)| \frac{e^{-u^2/2}}{\sqrt{2\pi}} du\right] = \sum_{i=0}^{+\infty} f_i(\sigma_i),
\end{aligned}$$

where, for all i,

$$f_i(\sigma) := \frac{\sigma}{2\pi}\left(\int_{-i-1}^{-i}\int_{-\infty}^{+\infty}\left|u + \frac{ty}{\sigma}\right| e^{-\frac{u^2+y^2}{2}} dudy + \int_i^{i+1}\int_{-\infty}^{+\infty}\left|u + \frac{ty}{\sigma}\right| e^{-\frac{u^2+y^2}{2}} dudy\right).$$

Notice that $f_i(\sigma) \longrightarrow \infty$, as $\sigma \to \infty$. Then there exists $\sigma_i > 0$ such that $f_i(\sigma_i) \geq 1$. Hence, $\mathbb{E}^{\overline{\mathbb{P}}}[|B_T|] = +\infty$, where $\overline{\mathbb{P}}$ is defined using this family of coefficients.

Proposition 3.3.1. *$\mathcal{M}^S$ and $\mathcal{M}^S_\xi$ verify Condition 3.3.1 (i) and (ii).*

Proof. We only provide the proof for $\mathcal{M}^S$, the result for $\mathcal{M}^S_\xi$ follows by direct adaptation. We first verify Condition 3.3.1 (ii). Let $\mathbb{P} \in \mathcal{M}^S$, and consider an arbitrary $\mathbb{F}$-stopping time τ, and for any $\tau \leq s \leq t$, a $\mathcal{F}_s^\tau$-measurable random variable H. By Lemma A.1 in [66], there exists some $\mathcal{F}_s$-measurable random variable $\tilde{H}$ such that for every ω, $\tilde{H}^{\tau,\omega} = H$. Then, we have for $\mathbb{P}$-a.e. ω

$$\mathbb{E}^{\mathbb{P}^{\tau,\omega}}[|B_t^\tau|] \leq \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[|B_t^{\tau,\omega}|] + |B_\tau(\omega)| = \mathbb{E}_\tau^{\mathbb{P}}[|B_t|](\omega) + |B_\tau|(\omega) < +\infty,$$

where we used the fact that $\mathbb{P} \in \mathcal{M}^S$. Similarly, we have for $\mathbb{P}$-a.e. ω:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^{\tau,\omega}}[HB_t^\tau] = \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\tilde{H}^{\tau,\omega}B_t^{\tau,\omega} - HB_\tau(\omega)] &= \mathbb{E}_\tau^{\mathbb{P}}[\tilde{H}B_t](\omega) - \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[HB_\tau(\omega)] \\
&= \mathbb{E}_\tau^{\mathbb{P}}[\tilde{H}B_s](\omega) - \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[HB_\tau(\omega)] \\
&= \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[HB_s^\tau].
\end{aligned}$$

By the arbitrariness of τ and H, this completes the verification of Condition 3.3.1 (ii).

To verify Condition 3.3.1 (ii), we adapt an argument from [74]. We define the following map

$$\psi : \mathbb{H}^1_{\rm loc}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) \to \mathbf{M}(\Omega), \quad \alpha \mapsto \mathbb{P}^\alpha = \mathbb{P}_0 \circ \left(\int_0^\cdot \alpha_s^{1/2} dB_s \right)^{-1}.$$

From [74] (see Lemmas 3.1 and 3.2), we know that it is sufficient to show that $\mathcal{M}^S \subset \mathbf{M}(\Omega)$ is the image of a Borel space (i.e. a Borel subset of a Polish space) under a Borel map. For that we show that $\mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F})$ is Polish and $\mathbb{H}^1_{\rm loc}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) \subset \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F})$ is Borel. The first part is already given in Lemma 3.1 of [74]. We then need to show that the map $\psi : \mathbb{H}^1_{\rm loc}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) \to \mathbf{M}(\Omega)$ is Borel, which is a direct consequence of Lemma 3.2 in [74].

It then only remains to prove that $\mathbb{H}^1_{\rm m}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) \subset \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F})$ is Borel,where

$$\mathbb{H}^1_{\rm m}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) := \left\{ \alpha \in \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0} \left[|X^\alpha_\tau| \mathbf{1}_{|X^\alpha_\tau| \geq n} \right] \underset{n\to+\infty}{\longrightarrow} 0 \right\}.$$

It is clear that

$$\mathbb{H}^1_{\rm m}(\mathbb{P}_0, \mathbb{S}_d^{>0}, \mathbb{F}) = \bigcap_{p\in\mathbb{N}^*} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{ \alpha \in \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|X^\alpha_\tau| \mathbf{1}_{|X^\alpha_\tau|>n}] \leq \frac{1}{p} \right\}.$$

and

$$\left\{ \alpha \in \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F}) : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|X^\alpha_\tau| \mathbf{1}_{|X^\alpha_\tau|>n}] \leq \frac{1}{p} \right\} = \psi^{-1} \left\{ \mathbb{P} \in \mathcal{M}^S_{\rm loc} : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|B_\tau| \mathbf{1}_{|B_\tau|>n}] \leq \tfrac{1}{p} \right\}.$$

It then suffices to show that for any $n \in \mathbb{N}$, the following function f_n is Borel measurable:

$$f_n \ : \ \mathbb{P} \mapsto \sup_\tau \mathbb{E}^{\mathbb{P}}[|B_\tau| \mathbf{1}_{|B_\tau|>n}].$$

We actually show that this function is lower semi-continuous. For $K, l > 0$, define:

$$\phi(x) = |x| \mathbf{1}_{|x|>n} \quad \text{and} \quad \phi_{K,l}(x) = |x| \wedge K \frac{(|x|-n)^+ - (|x|-n-l)^+}{l}, \quad x \in \mathbb{R}^d.$$

We emphasize that $\phi_{K,l}$ is uniformly continuous and bounded. Let then $\mathbb{P} \in \mathcal{M}^S_{\rm loc}$ and consider some sequence $(\mathbb{P}^i)_{i\geq 0}$ which converges weakly to $\mathbb{P}$. We represent $\mathbb{P}^i$ and $\mathbb{P}$ by α_i and α. Remember that for any $\widetilde{\mathbb{P}} \in \mathcal{M}^S_{\rm loc}$, associated to some $\widetilde{\alpha}$, we have

$$f_n(\widetilde{\mathbb{P}}) = \sup_\tau \mathbb{E}^{\mathbb{P}_0} \left[|X^{\widetilde{\alpha}}_\tau| \mathbf{1}_{|X^{\widetilde{\alpha}}_\tau|>n} \right].$$

The weak convergence of $\mathbb{P}^i$ to $\mathbb{P}$ is equivalent to the convergence in law of X^{α_i} to X^α. Hence, for all τ, we have

$$\lim_{i\to+\infty} \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(X^{\alpha_i}_\tau)] = \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(X^\alpha_\tau)],$$

from which we deduce easily $\liminf_{i\to+\infty} \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(X^{\alpha_i}_\tau)] \geq \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(X^\alpha_\tau)]$. As this is true for all $K > 0$ and $l > 0$, and since the function $\phi_{K,l}$ is non-decreasing in K and non-increasing in l, we also have

$$\liminf_{i\to+\infty} \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi(X^{\alpha_i}_\tau)] \geq \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(X^\alpha_\tau)].$$

Letting K go to $+\infty$ and l to 0 on the right-hand side above, we deduce using monotone convergence and taking supremum in τ

$$\liminf_{i\to+\infty} \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi(X^{\alpha_i}_\tau)] \geq \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi(X^\alpha_\tau)]$$

Then f_n is lower semicontinuous and thus measurable.

□

3.4 The duality result

In this section we show our main result Theorem 3.3.1. We will assume throughout that ξ is upper semi-analytic. For that purpose, we introduce the dynamic version of the dual problem:

$$Y_t(\xi)(\omega) := \sup_{\mathbb{P}\in\mathcal{M}^S(t,\omega)} \mathbb{E}^{\mathbb{P}}[\xi^{t,\omega}], \quad t\in[0,T],\ \omega\in\Omega.$$

We first observe that Y_t is measurable with respect to the universal filtration $\mathcal{F}_t^*$, as a consequence of Step 1 in the proof of Theorem 2.3 in [77], since Condition 3.3.1(i) holds true for $\mathcal{M}^S$.

Lemma 3.4.1. *Let τ be an $\mathbb{F}$-stopping time. Then, for all $\mathbb{P}\in\mathcal{M}^S$:*

$$Y_\tau(\xi) \quad = \quad \operatorname{ess\,sup}^{\mathbb{P}}_{\mathbb{P}'\in\mathcal{M}^S(\tau,\mathbb{P})} \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\tau], \quad \mathbb{P}-a.s.$$

where $\mathcal{M}^S(\tau,\mathbb{P}) = \{\mathbb{P}'\in\mathcal{M}^S \ :\ \mathbb{P}'=\mathbb{P} \text{ on } \mathcal{F}_\tau\}$.

Proof. The inequality $\geq$ is trivial as Y_τ is $\mathcal{F}^*_\tau$-measurable and measures extend uniquely to universal completions, which means that if $\mathbb{P}$ and $\mathbb{P}'$ coincide on $\mathcal{F}_\tau$, they also coincide on $\mathcal{F}^*_\tau$ (see Step 3 of the proof of Theorem 2.3 in [77] for similar arguments). We then focus on $\leq$. Fix some $\mathbb{P}\in\mathcal{M}^S$. We recall that following the same construction as in Step 2 of the proof of Theorem 2.3 in [77], for any $\varepsilon>0$, we can construct a $\mathcal{F}_\tau$-measurable kernel $\nu:\Omega\to\mathbf{M}(\Omega)$ such that:

$$\mathbb{E}^{\nu(\omega)}[\xi^{\tau,\omega}] \geq \big(Y_\tau(\omega)-\varepsilon\big)\mathbf{1}_{\{-\infty<Y_\tau(\omega)<\infty\}} + \varepsilon^{-1}\mathbf{1}_{\{Y_\tau(\omega)=\infty\}} - \infty\mathbf{1}_{\{Y_\tau(\omega)=-\infty\}}, \tag{3.4.1}$$

and such that $\nu(\omega)\in\mathcal{M}^S(\tau,\omega)$, $\mathbb{P}$-a.s.

We then consider the probability $\widetilde{\mathbb{P}}\in\mathcal{M}^S_{\text{loc}}(\tau,\mathbb{P})$ associated to ν through Condition 3.3.1(iii), where this last assertion uses that Condition 3.3.1(iii) is verified for $\mathcal{M}^S_{\text{loc}}$ thanks to Theorem 2.4 in [74]. However there is no guarantee that $\widetilde{\mathbb{P}}$ belongs to $\mathcal{M}^S$, and the rest of this proof overcomes this difficulty by using a suitable approximation.

Step 1: Construction of an approximation ν^n. Define, for all $n\geq 1$,

$$\nu^n(\omega) \quad := \quad \nu(\omega)\mathbf{1}_{\{\mathbb{E}^{\nu(\omega)}[|B^\tau_T|]\leq n\}} + \mathbb{P}^{\tau,\omega}\mathbf{1}_{\{\mathbb{E}^{\nu(\omega)}[|B^\tau_T|]>n\}}.$$

Clearly, ν^n is a measurable kernel, and since $\mathcal{M}^S$ is stable by bifurcation, we also have $\nu^n(\omega)\in\mathcal{M}^S(\tau,\omega)$. Observe that $E_n := \{\omega\in\Omega \ :\ \mathbb{E}^{\nu(\omega)}[|B^\tau_T|]\leq n\}$, $n\geq 1$, is an increasing sequence in $\mathcal{F}_\tau$ with $\mathbb{P}(E_n) \underset{n\to+\infty}{\longrightarrow} 1$, as a consequence of the fact that $\mathbb{E}^{\nu(\omega)}[|B^\tau_T|]<+\infty$, $\mathbb{P}-$a.s. We now define the measure $\overline{\mathbb{P}}^n$ by:

$$\overline{\mathbb{P}}^n(A) := \iint (\mathbf{1}_A)^{\tau,\omega}(\omega')\nu^n(d\omega';\omega)\mathbb{P}(dw), \quad A\in\mathcal{F}.$$

We first show that $\overline{\mathbb{P}}^n\in\mathcal{M}^S(\tau,\mathbb{P})$. The fact that $\overline{\mathbb{P}}^n$ coincides with $\mathbb{P}$ on $\mathcal{F}_\tau$ is clear by construction. Next, we compute that:

$$\begin{aligned}
\mathbb{E}^{\overline{\mathbb{P}}^n}[|B_T|] &= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[|B_T|\mathbf{1}_{E_n} + |B_T|\mathbf{1}_{E_n^c}\right]\\
&= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\nu^n(\omega)}\left[|B_T^{\tau,\omega}|\right]\mathbf{1}_{E_n}\right] + \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\nu^n(\omega)}\left[|B_T^{\tau,\omega}|\right]\mathbf{1}_{E_n^c}\right]\\
&= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\nu(\omega)}\left[|B_T^{\tau,\omega}|\right]\mathbf{1}_{E_n}\right] + \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\mathbb{P}^{\tau,\omega}}\left[|B_T^{\tau,\omega}|\right]\mathbf{1}_{E_n^c}\right]\\
&\leq \mathbb{E}^{\mathbb{P}}\left[\left(\mathbb{E}^{\nu(\omega)}\left[|B_T^\tau|\right]+|B_\tau|\right)\mathbf{1}_{E_n}\right] + \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}^{\tau,\omega}}\left[|B_T^{\tau,\omega}|\right]\mathbf{1}_{E_n^c}\right]\\
&\leq \mathbb{E}^{\mathbb{P}}\left[|B_\tau|\right] + n + \mathbb{E}^{\mathbb{P}}\left[|B_T|\right] < +\infty.
\end{aligned}$$

To prove the martingale property of B under $\overline{\mathbb{P}}^n$, we consider any $0 \leq s \leq t \leq T$ and an arbitrary $\mathcal{F}_s$-measurable random variable H, and we compute that:

$$\begin{aligned}\mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t] &= \mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t\mathbf{1}_{s\leq\tau} + HB_t\mathbf{1}_{s>\tau}]\\ &= \mathbb{E}^{\mathbb{P}}[HB_t\mathbf{1}_{s\leq t\leq\tau}] + \mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t\mathbf{1}_{s\leq\tau}\mathbf{1}_{t>\tau}] + \mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t\mathbf{1}_{s>\tau}],\end{aligned}$$

by the fact that $\overline{\mathbb{P}}^n = \mathbb{P}$ on $\mathcal{F}_\tau$. We continue computing

$$\begin{aligned}\mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t\mathbf{1}_{s>\tau}] &= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{(\overline{\mathbb{P}}^n)^{\tau,\omega}}\left[H^{\tau,\omega}B_t^{\tau,\omega}\right]\mathbf{1}_{s>\tau}\right]\\ &= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\nu^n(\omega)}\left[H^{\tau,\omega}B_t^{\tau} + H^{\tau,\omega}B_\tau(\omega)\right]\mathbf{1}_{s>\tau}\right]\\ &= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[\mathbb{E}^{\nu^n(\omega)}\left[H^{\tau,\omega}B_s^{\tau} + H^{\tau,\omega}B_\tau(\omega)\right]\mathbf{1}_{s>\tau}\right]\\ &= \mathbb{E}^{\overline{\mathbb{P}}^n}\left[HB_s\mathbf{1}_{s>\tau}\right],\end{aligned}$$

where we used the definition of ν^n which ensures that $\nu^n(\omega) \in \mathcal{M}^S(\tau,\omega)$, $\mathbb{P}$-a.s., so that since $H^{\tau,\omega}$ is $\mathcal{F}_s^{\tau(\omega)}$-measurable, we do have $\mathbb{E}^{\nu^n(\omega)}\left[H^{\tau,\omega}B_t^{\tau}\right] = \mathbb{E}^{\nu^n(\omega)}\left[H^{\tau,\omega}B_s^{\tau}\right]$, when $s > \tau(\omega)$. We compute similarly that

$$\mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t\mathbf{1}_{s\leq\tau}\mathbf{1}_{t>\tau}] = \mathbb{E}^{\overline{\mathbb{P}}^n}[HB_\tau\mathbf{1}_{s\leq\tau}\mathbf{1}_{t>\tau}] = \mathbb{E}^{\mathbb{P}}[HB_\tau\mathbf{1}_{s\leq\tau}\mathbf{1}_{t>\tau}].$$

Finally, we have

$$\mathbb{E}^{\mathbb{P}}[HB_t\mathbf{1}_{s\leq t\leq\tau}] = \mathbb{E}^{\mathbb{P}}\left[HB_{t\wedge\tau}\mathbf{1}_{s\leq\tau} - HB_\tau\mathbf{1}_{t>\tau\geq s}\right] = \mathbb{E}^{\mathbb{P}}\left[HB_s\mathbf{1}_{s\leq\tau} - HB_\tau\mathbf{1}_{t>\tau\geq s}\right],$$

since B is a martingale under $\mathbb{P}$. We deduce therefore that,

$$\mathbb{E}^{\overline{\mathbb{P}}^n}[HB_t] = \mathbb{E}^{\overline{\mathbb{P}}^n}[HB_s],$$

which proves by arbitrariness of H that B is a $\overline{\mathbb{P}}^n$-martingale.

Step 2: By (3.4.1), we have for every ω

$$\mathbb{E}^{\nu^n(\omega)}[\xi^{\tau,\omega}] \geq (Y_\tau(\omega) - \varepsilon) \wedge \varepsilon^{-1}\mathbf{1}_{E_n} + \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\xi^{\tau,\omega}]\mathbf{1}_{E_n^c}.$$

Then for any $\omega \in \Omega\backslash\mathcal{N}^{\mathbb{P}}$, for some $\mathbb{P}$-null set $\mathcal{N}^{\mathbb{P}}$

$$\mathbb{E}^{\overline{\mathbb{P}}^n}[\xi|\mathcal{F}_\tau](\omega) \geq (Y_\tau(\omega) - \varepsilon) \wedge \varepsilon^{-1}\mathbf{1}_{E_n}(\omega) + \mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_\tau](\omega)\mathbf{1}_{E_n^c}(\omega).$$

Hence, for any $\omega \in \Omega\backslash\mathcal{N}^{\mathbb{P}}$, for all $n \geq 0$

$$\operatorname*{ess\,sup}^{\mathbb{P}}_{\mathbb{P}'\in\mathcal{M}^S(\tau,\mathbb{P})} \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\tau](\omega) \geq (Y_\tau(\omega) - \varepsilon) \wedge \varepsilon^{-1}\mathbf{1}_{E_n}(\omega) + \mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_\tau](\omega)\mathbf{1}_{E_n^c}(\omega).$$

We emphasize that *a priori*, the right-hand side above is only $\mathcal{F}_\tau^*$-measurable. However, if $\mathbb{P}$ and $\mathbb{P}'$ coincide on $\mathcal{F}_\tau$, they also coincide on $\mathcal{F}_\tau^*$, since measures extend uniquely on universal completions. Therefore the above inequality does indeed hold $\mathbb{P} - a.s.$

Since the sequence E_n increases to Ω (up to some $\mathbb{P}$-null set which we implicitly add to $\mathcal{N}^{\mathbb{P}}$), for any $\omega \in \Omega\backslash\mathcal{N}^{\mathbb{P}}$, there exists $N(\omega) \in \mathbb{N}$ such that if $n \geq N(\omega)$, then $\omega \in E_n$. Therefore, taking n large enough, we have

$$\operatorname*{ess\,sup}^{\mathbb{P}}_{\mathbb{P}'\in\mathcal{M}^S(\tau,\mathbb{P})} \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\tau](\omega) \geq (Y_\tau(\omega) - \varepsilon) \wedge \varepsilon^{-1}. \tag{3.4.2}$$

If $Y_\tau(\omega) = -\infty$, then by the inequality proved at the beginning, the left-hand side above is also equal to $-\infty$. Hence the result in this case. If $Y_\tau(\omega) = +\infty$, then (3.4.2) implies directly that the left-hand side is also $+\infty$ by arbitrariness of $\varepsilon > 0$. Finally, if $Y_\tau(\omega)$ is finite, the desired result follows from (3.4.2) by arbitrariness of ε.

□

We then continue with a version of the tower property in our context.

Proposition 3.4.1. *Let $\mathbb{P}\in\mathcal{M}^S_\xi$, and σ,τ two $\mathbb{F}$-stopping times with $\sigma\leq\tau$. Then, $\mathbb{P}-$a.s.*

$$Y_\sigma(\xi)=\underset{\mathbb{P}'\in\mathcal{P}^\xi_m(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}\left[\underset{\mathbb{P}''\in\mathcal{M}^S_\xi(\tau,\mathbb{P}')}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}''}[\xi|\mathcal{F}_\tau]|\mathcal{F}_\sigma\right]=\underset{\mathbb{P}'\in\mathcal{M}^S(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}\left[\underset{\mathbb{P}''\in\mathcal{M}^S(\tau,\mathbb{P}')}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}''}[\xi|\mathcal{F}_\tau]|\mathcal{F}_\sigma\right],$$

where for any $\mathbb{F}$-stopping time ι and any $\mathbb{P}\in\mathcal{M}^S_\xi$

$$\mathcal{M}^S_\xi(\iota,\mathbb{P}):=\left\{\mathbb{P}'\in\mathcal{M}^S_\xi\ s.t.\ \mathbb{P}'=\mathbb{P}\ on\ \mathcal{F}_\iota\right\}.$$

Proof. We consider $\mathbb{P}\in\mathcal{M}^S_\xi$. Exactly as in the proof of Lemma 3.4.1, we can construct a measurable kernel ν^n from a kernel ν such that:
- ν^n is $\mathcal{F}_\tau$-measurable.
- $\mathbb{P}^n\in\mathcal{M}^S(\tau,\mathbb{P})$ where $\mathbb{P}^n(A)=\iint(\mathbf{1}_A)^{\tau,\omega}(\omega')\nu^n(d\omega';\omega)\mathbb{P}(dw),\quad A\in\mathcal{F}$.
- ν is a $\mathcal{F}_\tau$-measurable kernel such that (3.4.1) holds.
- $E_n=\{\nu^n=\nu\}$ is an increasing sequence such that $\mathbb{P}(E_n)\underset{n\to+\infty}{\longrightarrow}1$.

We then compute for any $\varepsilon>0$

$$\underset{\mathbb{P}'\in\mathcal{M}^S(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\sigma]\ \geq\ \mathbb{E}^{\mathbb{P}^n}[\xi|\mathcal{F}_\tau]\ \geq\ \mathbb{E}^{\mathbb{P}}[(Y_\tau-\varepsilon)\wedge\varepsilon^{-1}\mathbf{1}_{E_n}-\mathbb{E}^{\mathbb{P}}[\xi^-|\mathcal{F}_\tau]\mathbf{1}_{E_n^c}|\mathcal{F}_\sigma],\ \mathbb{P}-a.s.$$

Recall that $\mathbb{E}^{\mathbb{P}}[\xi^-]<\infty$. Then, it follows from the dominated convergence theorem that $\mathbb{E}^{\mathbb{P}}[\xi^-\mathbf{1}_{E_n^c}|\mathcal{F}_\sigma]\longrightarrow0$, as $n\to\infty$, $\mathbb{P}-$a.s. Also, since $Y_\tau\geq-\mathbb{E}^{\mathbb{P}}[\xi^-|\mathcal{F}_\tau]\in\mathbb{L}^1(\mathbb{P})$, it follows from Fatou's lemma that:

$$\mathbb{E}^{\mathbb{P}}[(Y_\tau-\varepsilon)\wedge\varepsilon^{-1}|\mathcal{F}_\sigma]\leq\underset{\mathbb{P}'\in\mathcal{M}^S(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\sigma]\ \mathbb{P}-\text{a.s.}$$

Finally, when $\varepsilon\to0$, we have $\mathbb{P}$-a.s., with the last equality being obvious

$$\mathbb{E}^{\mathbb{P}}[Y_\tau|\mathcal{F}_\sigma]\leq\underset{\mathbb{P}'\in\mathcal{M}^S(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\sigma]=\underset{\mathbb{P}'\in\mathcal{M}^S_\xi(\sigma,\mathbb{P})}{\operatorname{ess\,sup}}\ \mathbb{E}^{\mathbb{P}'}[\xi|\mathcal{F}_\sigma].$$

□

Proposition 3.4.2. *Assume that $\sup_{\mathbb{P}\in\mathcal{M}^S}\mathbb{E}^{\mathbb{P}}[\xi^+]<\infty$. Then, for any $\mathbb{P}\in\mathcal{M}^S_\xi$, the process $\{Y_t(\xi),t\leq T\}$ is a $\mathbb{P}$-supermartingale.*

Proof. In view of Proposition 3.4.1, we already have the tower property. It only remains to show the integrability of $Y_t(\xi)$ for all $t\in[0,T]$. For that we only need to show that $Y_t(\xi^+)$ is integrable. We fix $0\leq t\leq T$ and $\mathbb{P}\in\mathcal{M}^S_\xi$. Then $Y_t(\xi^+)(\omega)\in\mathbb{R}_+\times\{+\infty\}$. Let us then show that the family $\{\mathbb{E}^{\mathbb{P}'}[\xi^+|\mathcal{F}_t],\ \mathbb{P}'\in\mathcal{M}^S(t,\mathbb{P})\}$ is upward directed.

We consider $\mathbb{P}_1$ and $\mathbb{P}_2$ in $\mathcal{M}^S(s,\mathbb{P})$. The set $A:=\{\mathbb{E}^{\mathbb{P}_2}[\xi^+|\mathcal{F}_t]\leq\mathbb{E}^{\mathbb{P}_1}[\xi^+|\mathcal{F}_t]\}$, is $\mathcal{F}_t$-measurable. Then $\overline{\mathbb{P}}:=\mathbb{P}_1\mathbf{1}_A+\mathbb{P}_2\mathbf{1}_{A^c}$ is an element of $\mathcal{M}^S(t,\mathbb{P})$ such that:

$$\mathbb{E}^{\overline{\mathbb{P}}}[\xi^+|\mathcal{F}_t]=\mathbb{E}^{\mathbb{P}_1}[\xi^+|\mathcal{F}_t]\vee\mathbb{E}^{\mathbb{P}_2}[\xi^+|\mathcal{F}_t]$$

We then have an increasing sequence $\mathbb{P}_n$ of $\mathcal{M}^S(t,\mathbb{P})$ such that:

$$\mathbb{E}^{\mathbb{P}_n}[\xi^+|\mathcal{F}_t] \nearrow \operatorname*{ess\,sup}_{\mathbb{P}'\in\mathcal{M}^S(s,\mathbb{P})} \mathbb{E}^{\mathbb{P}'}[\xi^+|\mathcal{F}_t], \ \mathbb{P}-\text{a.s.},$$

and by the monotone convergence theorem $\lim_{n\to+\infty}\mathbb{E}^{\mathbb{P}_n}[\xi^+] = \mathbb{E}^{\mathbb{P}}[Y_t(\xi^+)]$. Hence

$$\mathbb{E}^{\mathbb{P}}[Y_t(\xi^+)] \le \sup_{\mathbb{P}\in\mathcal{M}^S} \mathbb{E}^{\mathbb{P}}[\xi^+] < +\infty, \quad \text{for all } \mathbb{P}\in\mathcal{M}^S.$$

□

We now have all ingredients to follow the classical line of argument for the

Proof of theorem 3.3.1 For the sake of simplicity, the dependence of Y in ξ will be omitted.

(i) We first show that right-limiting process $\overline{Y}_t := Y_{t+}$, $t\le T$, is a $(\mathbb{G}^{\mathcal{M}^S},\mathbb{P})$ supermartingale for all $\mathbb{P}\in\mathcal{M}^S_\xi$. By Proposition 3.4.2 and the fact that for any $\mathbb{P}\in\mathcal{M}^S_{\text{loc}}$, $\overline{\mathbb{F}}^{\mathbb{P}}$ is right-continuous, contains $\mathbb{G}$ and $\mathbb{P}$ has the predictable representation property (see [92]), Y is a $(\mathbb{F}^*,\mathbb{P})$ supermartingale for every $\mathbb{P}\in\mathcal{M}^S_\xi$. Then applying [37] (see Theorem VI.2), we have that $\overline{Y}$ is well defined $\mathbb{P}$-a.s. and $\overline{Y}$ is a right continuous $(\mathbb{G}^{\mathcal{M}^S},\mathbb{P})$ supermartingale for all $\mathbb{P}\in\mathcal{M}^S_\xi$. We also notice the important fact that for all $\mathbb{P}\in\mathcal{M}^S_\xi$ we have $\overline{Y}_t\le Y_t$ $\mathbb{P}$-a.s. In particular, $\overline{Y}_0\le Y_0$, and $\overline{Y}_0$ is constant because $\mathcal{G}_0^{\mathcal{M}^S}$ is trivial.

(ii) We next construct the optimal trading strategy. By the Doob-Meyer decomposition (see Theorem 13 page 115 in [82]), there exists a pair of processes $(H^{\mathbb{P}},K^{\mathbb{P}})$ where $H^{\mathbb{P}}$ belongs to $\mathbb{H}^2_{\text{loc}}(\mathbb{P},\mathbb{R}^d,\mathbb{G}^{\mathcal{M}^S})$ and $K^{\mathbb{P}}$ is $\mathbb{P}-$integrable and non-decreasing, such that:

$$\overline{Y}_t = Y_0 + \int_0^t H^{\mathbb{P}}_s dB_s - K^{\mathbb{P}}_t, \ t\in[0,T], \ \mathbb{P}-\text{a.s.}$$

Since $\overline{Y}$ is right continuous, it follows from Karandikar [65] that the family $H^{\mathbb{P}}$ can be aggregated by some process $\widehat{H}$ in the sense that $\widehat{H} = H^{\mathbb{P}}$, $dt\times d\mathbb{P}$-a.s. for all $\mathbb{P}\in\mathcal{M}^S_\xi$. Thus, for every $\mathbb{P}\in\mathcal{M}^S_\xi$, the local martingale $\int\widehat{H}dB$ is bounded from below by the martingale $\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{G}_t^{\mathcal{M}^S}]$. Hence this is a supermartingale which ensures that $\widehat{H}$ is in $\mathcal{H}^\xi$ and superreplicates the claim ξ $\mathcal{M}^S_\xi$-quasi-surely.

□

3.5 Extensions

3.5.1 The case of $\mathcal{M}^S_b$

In this section, we show that Theorem 3.3.1 together with the previous arguments in its proof, hold for the following example, which is important in the context of second-order BSDEs as introduced in [91]. We recall that $\langle B\rangle$ is well defined pathwise and that its density is denoted by $\widehat{a}$.

Definition 3.5.1. *$\mathcal{M}^S_b$ is the sub-class of $\mathcal{M}^S_{loc}$ consisting of all $\mathbb{P}\in\mathcal{M}^S_{loc}$ such that:*

$$\underline{a}_{\mathbb{P}} \le \widehat{a} \le \overline{a}_{\mathbb{P}}, \quad dt\times d\mathbb{P}-a.s. \text{ for some } \underline{a}_{\mathbb{P}},\overline{a}_{\mathbb{P}}\in\mathbb{S}_d^{>0}.$$

We emphasize that our Example 3.3.1 also shows directly that $\mathcal{P}_b$ does not satisfy Condition 3.3.1(iii). We now prove the three main technical results of this paper in this case.

Proposition 3.5.1. *$\mathcal{M}_b^S$ verifies Condition 3.3.1 (i) and (ii).*

Proof. (ii) has already been proved in [93]. Let us now prove (i). We observe that:

$$\mathcal{M}_b^S = \bigcup_{\underline{a},\bar{a}\in\mathbb{S}_d^{>0}(\mathbb{Q})} \{\mathbb{P}\in\mathcal{M}_{\text{loc}}^S \;:\; \underline{a}\le\hat{a}\le\bar{a}\; dt\times d\mathbb{P}-\text{a.s.}\}.$$

By Proposition 3.1 in [77], we know that all these sets satisfy Condition 3.3.1(i). Since a countable union of analytic set is analytic, we obtain the result.

□

It remains to introduce a suitable sequence of approximations of measurable kernels as in the proof of Lemma 3.4.1 and Proposition 3.4.1.

Proposition 3.5.2. *The results of Lemma 3.4.1, Proposition 3.4.1 and 3.4.2 and Theorem 3.3.1 are valid if we replace $\mathcal{M}^S$ by $\mathcal{M}_b^S$.*

Proof. We only redefine an approximated kernel family adapted to $\mathcal{M}_b^S$, which allows, by the same arguments as in Lemma 3.4.1 and Proposition 3.4.2, to obtain the duality result of Theorem 3.3.1. Let τ be a $\mathbb{F}$-stopping time and ν the $\mathcal{F}_\tau$-measurable kernel obtained by the same construction as in Lemma 3.4.1 and Proposition 3.4.1. For $\mathbb{P}\in\mathcal{M}_b^S$ we are interested in the measure $\overline{\mathbb{P}}$ defined by:

$$\overline{\mathbb{P}}(A) = \iint (\mathbf{1}_A)^{\tau,\omega}(\omega')\,\nu(d\omega';\omega)\,\mathbb{P}(dw).$$

Then $\overline{\mathbb{P}}$ is in $\mathcal{M}_{\text{loc}}^S$ and there is some α s.t. $\overline{\mathbb{P}} = \mathbb{P}_0\circ(X^\alpha_\cdot)^{-1}$. Define $\overline{\mathbb{P}}^n$ by:

$$\overline{\mathbb{P}}^n := \mathbb{P}_0\circ\left(\int_0^\cdot (\alpha_s^{1/2}\mathbf{1}_{s\le\tau} + \pi_n(\alpha_s^{1/2})\mathbf{1}_{s>\tau})dB_s)\right)^{-1},\; n\ge 1,$$

where π_n is the projection on $\overline{\mathcal{B}}_{\mathbb{S}_d^{>0}}(0,n)\backslash\overline{\mathcal{B}}_{\mathbb{S}_d^{>0}}(0,1/n)$, where $\overline{\mathcal{B}}_{\mathbb{S}_d^{>0}}(x,r)$ denotes the closed ball of $\mathbb{S}_d^{>0}$ centered at x with radius r. Then $\overline{\mathbb{P}}^n$ belongs to $\mathcal{P}_b$. Observe also that the sets

$$E_n := \left\{\omega\in\Omega,\; \left(\overline{\mathbb{P}}^n\right)^{\tau,\omega} = \nu(\omega)\right\},$$

are in $\mathcal{F}_\tau$ and define an increasing covering of Ω. We then build the "right" approximation $\widetilde{\mathbb{P}}^n$, ensuring all the convergences in the proofs, by $\widetilde{\mathbb{P}}^n = \overline{\mathbb{P}}^n\mathbf{1}_{E_n} + \mathbb{P}\mathbf{1}_{E_n^c}$. $\widetilde{\mathbb{P}}^n$ is in $\mathcal{M}_b^S$, and associated to the $\mathcal{F}_\tau$-measurable kernel $\tilde{\nu}^n(\omega) := \nu(\omega)\mathbf{1}_{\omega\in E_n} + \mathbb{P}^{\tau,\omega}\mathbf{1}_{\omega\in E_n^c}$. Then we can reproduce exactly the same proofs as in the case of $\mathcal{M}^S$.

□

3.5.2 A general framework

As the reader may have realized, our proofs in the case of $\mathcal{M}^S$ and $\mathcal{P}_b$ are very similar, and essentially rely on the construction of a suitable approximated kernel. In this subsection, we consider a generic subset $\mathcal{P}$ of $\mathcal{M}_{\text{loc}}^S$ (and the corresponding shifted families $\mathcal{P}(s,\omega)$), and we give a general condition, (weaker than Condition 3.3.1) under which our results still hold true. We recall that such a family is said to be stable by bifurcation if for any $\mathbb{F}$-stopping times $0\le\sigma\le\tau$, $\omega\in\Omega$, A $\mathcal{F}_\tau$-measurable, $\mathbb{P}_1$ and $\mathbb{P}_2$ in $\mathcal{P}(\sigma,\omega)$, we have

$$\mathbb{P} = \mathbb{P}_1\mathbf{1}_A + \mathbb{P}_2\mathbf{1}_{A^c}\in\mathcal{P}(\sigma,\omega).$$

Condition 3.5.1. *Let $s \in \mathbb{R}_+$, $\tau \geq s$ a stopping time, $\overline{\omega} \in \Omega$, $\mathbb{P} \in \mathcal{P}(s,\overline{\omega})$ and $\theta := \tau^{s,\overline{\omega}} - s$.*
(i) *The graph $\{(\mathbb{P}',\omega) : \omega \in \Omega,\ \mathbb{P}' \in \mathcal{P}(t,\omega)\} \subseteq \mathbf{M}(\Omega) \times \Omega$ is analytic.*
(ii) *We have $\mathbb{P}^{\theta,\omega} \in \mathcal{P}(\tau, \overline{\omega} \otimes_s \omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*
(iii) *$\mathcal{P}$ is stable by bifurcation.*
(iv) *If $\nu : \Omega \to \mathbf{M}(\Omega)$ is an $\mathcal{F}_\theta$-measurable kernel and $\nu(\omega) \in \mathcal{P}(\tau, \overline{\omega} \otimes_s \omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, then there exists $\nu^n : \Omega \to \mathbf{M}(\Omega)$, which is a $\mathcal{F}_\theta$-measurable kernel such that $\mathbb{P}(\nu^n = \nu) \underset{n\to\infty}{\longrightarrow} 1$ and the following measure $\overline{\mathbb{P}}^n \in \mathcal{P}(s,\overline{\omega})$:*

$$\overline{\mathbb{P}}^n(A) = \iint (\mathbf{1}_A)^{\theta,\omega}(\omega')\nu^n(d\omega';\omega)\mathbb{P}(dw), \quad A \in \mathcal{F}.$$

Remark 3.5.1. *Notice that Condition 3.5.1 is weaker than Condition 3.3.1. Indeed, Condition 3.3.1(iii) implies directly that the set $\mathcal{P}$ is stable by bifurcation. Moreover, considering the constant kernels $\nu^n := \nu$, it also implies Condition 3.5.1(iv). Furthermore, as shown in our previous proofs, the sets $\mathcal{M}^S$ and $\mathcal{P}_b$ satisfy Condition 3.5.1 but not Condition 3.3.1.*

Similarly to our previous notations, we introduce the sets $\mathcal{H}^\xi(\mathcal{P})$ and $\mathcal{P}^\xi$. In this context, we obtain a new version of Theorem 3.3.1:

Theorem 3.5.1. *Let $\mathcal{P}(s,\omega)$ be a family of probability measures satisfying Condition 3.5.1. Let ξ be an upper semi-analytic r.v. with $\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\xi^+] < +\infty$. Then*

$$V(\xi) \ := \ \inf\left\{X_0 :\ X_T^H \geq \xi,\ \ \mathcal{P}^\xi - q.s. \text{ for some } H \in \mathcal{H}^\xi(\mathcal{P})\right\} \ = \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\xi].$$

Moreover, existence holds for the primal problem, i.e. $V(\xi) + \int_0^T H_s dB_s \geq \xi$, $\mathcal{P}^\xi$−q.s. for some $H \in \mathcal{H}^\xi(\mathcal{P})$.

Proof. If we define $\widetilde{E}_n := \left\{\omega \in \Omega \ : \ (\overline{\mathbb{P}}^n)^{\theta,\omega} = \nu(\omega)\right\}$ and then recursively

$$E_0 := \widetilde{E}_0 \text{ and for all } n \geq 1,\ E_n := E_n \bigcup \widetilde{E}_{n-1},$$

then E_n is an increasing sequence such that $\mathbb{P}(E_n) \underset{n\to+\infty}{\longrightarrow} 1$. We can then use the $\mathcal{F}_\tau$-measurable kernel ν^n to define a probability measure $\widetilde{P}^n$ exactly as in the proof of Proposition 3.5.2. We can then use exactly the same arguments as in our previous proofs.

□

Chapter 4

Robust indifference pricing of measurable claims under exponential utility

Contents

4.1 Introduction

In this chapter, we consider the problem of robust indifference pricing of a contingent claim ξ at exercise time T. Namely we consider a financial market consisting of one risky asset S which dynamic is given by:

$$dS_t = S_t \left(b\sigma_t^2 dt + \sigma_t dW_t \right), \tag{4.1.1}$$

where b is a stochastic process, corresponding to the sharpe ratio, and σ is unknown. In addition, we assume that for the maturity T, calls of all strikes are tradable. The agent's purpose, whose utility criterion is of exponential form, is then to maximise the utility of its portfolio containing the liability $-\xi$, and any chosen semi-static trading strategy (π, λ) in the robust setup. Then the agent calculates the corresponding robust utility indifference price of the contingent claim ξ, i.e. denoting $\mathbb{P}^\sigma$ the probability measure such that S follows the dynamic equation (4.1.1), we study:

$$\begin{aligned} \inf \Big\{ p \ : \ & \sup_{\pi,\lambda} \inf_{\sigma} \mathbb{E}^{\mathbb{P}^\sigma} \left[U \left(x + p + \int_0^T \pi_t dS_t + \lambda(S_T) - \mu(\lambda) - \xi \right) \right] \\ & \geq \sup_{\pi,\lambda} \inf_{\sigma} \mathbb{E}^{\mathbb{P}^\sigma} \left[U \left(x + \int_0^T \pi_t dS_t + \lambda(S_T) - \mu(\lambda) \right) \right] \Big\}, \end{aligned}$$

where π is the dynamic component of the strategy, and λ is the static component, which initial cost is given by $\mu(\lambda) := \int \lambda(x)\mu(dx)$, where μ is the measure corresponding to the implied law of S_T given by the call options prices.

Notice that when $b = 0$, our problem leads trivially to the robust super-replication problem, regardless of the nature of the utility function, and the utility indifference price is given by the expectation of the claim. This is an immediate consequence of the Jensen inequality together with the fact that the zero investment strategy in the risky asset is legitimate. Consequently, only the case $b \neq 0$ is of interest.

We next present the first guess solution to this problem, which will turn out to be wrong, in general. Our formulation of model uncertainty implies that, under any fixed admissible measure, the market is complete. By the standard well-known result in financial mathematics, the utility indifference price under this measure coincides with the Black-Schole replication price defined by this risk-neutral expectation of the claim. Then, the first guess solution of our problem is that the robust utility indifference price of the claim should be given by the robust hedging cost. Our main result shows that this intuition is wrong, and that, even in our simple exponential utility framework, the robust utility indifference price turns out to be given by the robust hedging cost of some conveniently modified claim. This is outlined in Section 4.5. In the case of constant Sharpe ratio b, the first guess solution is true due to a very particular simplification induced by the structure of the risk-neutral density. However the corresponding optimal hedging strategy is given by a convenient modification of the robust superhedging strategy.

This chapter is organized as follows. We explore the case of b constant and we show that the robust utility indifference price of a general payoff ξ is given by the corresponding super-replication price. For that purpose, we study the value functions associated to the problem with and without the claim ξ and show that the first value function depends, in the case of exponential utility, of the robust super-replication price of the payoff ξ corrected by the entropy value of the measure μ (we deduce immediatly the value of the second value function by applying the result to $\xi = 0$). Then the robust indifference price follows immediatly. We highlight that this results holds since we can express explicitely the entropy process as a pathwise stochastic integral corrected by a vanilla option (log option).

We then discuss this last result by considering the case of a stochastic observable sharpe ratio process b, which leads to a robust indifference price depending on the super-replication price of the payoff ξ corrected by the entropy, and the super-replication price of entropy. In this case the entropy process fails to have the previous form and then the robust indifference price of ξ is different from its robust super-replication price.

In section 4.2 we present formaly the problem. In section 4.3 we derive a first inequality for the value function by using classical convex duality arguments. In section 4.4, we obtain the second inequality by quasi-sûre analysis and arguments from chapter 3. In section 4.5, we describe the same problem, but with a sharpe ratio depending of the path and explain formaly the result obtained in that case.

4.2 Problem formulation

4.2.1 Probabilistic framework

Let $\Omega := \{\omega \in C([0,T],\mathbb{R}) : \omega_0 = 0\}$ be the canonical space equipped with the uniform norm $||\omega||_\infty^T := \sup_{0\leq t\leq T} |\omega_t|$, B the canonical process, $\mathbb{P}_0$ the Wiener measure, $\mathbb{F} := \{\mathcal{F}_t\}_{0\leq t\leq T}$ the filtration generated by B and $\mathbb{F}^+ := \{\mathcal{F}_t^+, 0 \leq t \leq T\}$, the right limit of $\mathbb{F}$ where $\mathcal{F}_t^+ := \cap_{s>t}\mathcal{F}$. We will denote by $\mathbf{M}(\Omega)$ the set of all probability measures on Ω.

For any subset E of a finite dimensional space and any filtration $\mathbb{X}$ on $(\Omega, \mathcal{F})$, we denote by $\mathbb{H}^0(E, \mathbb{X})$ the set of all $\mathbb{X}$-progressively measurable processes with values in E. Moreover for all $p > 0$ and for all $\mathbb{P} \in \mathbf{M}(\Omega)$, we denote by $\mathbb{H}^p(\mathbb{P}, E, \mathbb{X})$ the subset of $\mathbb{H}^0(E, \mathbb{X})$ whose elements H satisfy $\mathbb{E}^{\mathbb{P}}\left[\int_0^T |H_t|^p\, dt\right] < +\infty$. The localized versions of these spaces are denoted by $\mathbb{H}^p_{\text{loc}}(\mathbb{P}, E, \mathbb{X})$.

For all $\sigma \in \mathbb{H}^2_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$ and $b \in \mathbb{R}$, we define the probability measure on $(\Omega, \mathcal{F})$:

$$\mathbb{P}^\sigma := \mathbb{P}_0 \circ (X^\sigma)^{-1} \text{ where } X_t^\sigma := \int_0^t \sigma_s dB_s + \int_0^t b\sigma_s^2 ds,\ t \in [0,T],\ \mathbb{P}_0 - \text{a.s.}$$

We denote by $\mathcal{P}_S^b$ the collection of all such probability measure on $(\Omega, \mathcal{F})$. We recall from [65] that the quadratic variation process is universally defined under any $\mathbb{P} \in \mathcal{P}_S^b$.
We know by [91] that there exists $\hat{a} > 0$ such that $d\langle B\rangle_t = \hat{a}_t dt\ dt \times \mathbb{P}$-a.s. for all $\mathbb{P} \in \mathcal{P}_S^b$ and it follows from the Levy's characterization of the brownian motion that

$$W^{\mathbb{P}} := \int_0^{\cdot} \hat{a}_t^{-1/2} dB_t - \int_0^{\cdot} b\hat{a}_t^{1/2} dt$$

is a $\mathbb{P}$-brownian motion.

Remark 4.2.1. *Concerning the definition of $\mathcal{P}_S^b$, this probability set is equal to the set $\mathcal{M}_{loc}^S$ when $b = 0$, where $\mathcal{M}_{loc}^S$ was defined in Chapter 3. For $b \neq 0$, we observe easily that an element $\mathbb{P}$ of $\mathcal{P}_S^b$ is not a local martingale measure.*

We next introduce the process:

$$Z_t := e^{-bB_t + \frac{1}{2}b^2\langle B\rangle_t},\ t \in [0,T].$$

Lemma 4.2.1. $\forall \mathbb{P} \in \mathcal{P}_S^b$*, the processes Z, $Z\mathcal{E}(B)$, and ZB are $\mathbb{P}$-local martingales.*

Proof. *(i): Z.* Consider $\mathbb{P} \in \mathcal{P}_S^b$. By Ito's formula, we have under $\mathbb{P}$:

$$\begin{aligned} dZ_t &= -bZ_t dB_t + b^2 Z_t d\langle B\rangle_t \\ &= -b\hat{a}_t^{1/2} Z_t dW_t^{\mathbb{P}} \end{aligned}$$

So Z is a local martingale under $\mathbb{P}$.
(ii): $Z\mathcal{E}(B)$. Using the same notation than in (i), we have by Ito's formula that under $\mathbb{P}$:

$$\begin{aligned} d\left(Z\mathcal{E}(B)\right)_t &= Z_t\mathcal{E}(B)_t dB_t + \mathcal{E}(B) dZ_t + d\langle Z, \mathcal{E}(B)\rangle_t \\ &= Z_t\mathcal{E}(B)_t \hat{a}^{1/2} dW_t^{\mathbb{P}} + Z_t\mathcal{E}(B)_t b\hat{a}_t dt - b\hat{a}_t^{1/2} Z_t\mathcal{E}(B)_t dW_t^{\mathbb{P}} - b\hat{a}_t^{1/2} Z_t\mathcal{E}(B)_t \hat{a}^{1/2} dt \\ &= (1-b) Z_t\mathcal{E}(B)_t \hat{a}^{1/2} dW_t^{\mathbb{P}}. \end{aligned}$$

So that $Z\mathcal{E}(B)$ is a local martingale.
(iii): ZB. Once again, applying Ito's formula, we have:

$$\begin{aligned} d(ZB)_t &= B_t dZ_t + Z_t dB_t + d\langle Z, B\rangle_t \\ &= -b\hat{a}_t^{1/2} B_t Z_t dW_t^{\mathbb{P}} + \hat{a}^{1/2} Z_t dW_t^{\mathbb{P}} + b\hat{a}_t Z_t dt - b\hat{a}_t Z_t dt \\ &= (1 - bB_t)\hat{a}_t^{1/2} Z_t dW_t^{\mathbb{P}}, \end{aligned}$$

which ends the proof.

□

Remark 4.2.2. *For this proof we chose to use the weak formulation and the existence of a Brownian motion under each $\mathbb{P}$. One could have also consider that the process $(B - b\langle B\rangle)_.$ is a local martingale under each $\mathbb{P} \in \mathcal{P}_S^b$ and use the Itô's formula to obtain the local martingale property.*

Remark 4.2.3. *We emphasize that Z and $Z\mathcal{E}(B)$ are non-negative local martingales, and so supermartingales. Since we will need the processes to be martingales, one only need to consider elements $\mathbb{P}$ of $\mathcal{P}_S^b$ such that $\mathbb{E}^{\mathbb{P}}[Z_T] = \mathbb{E}^{\mathbb{P}}[Z_T\mathcal{E}(B)_T] = 1$.*

In all the following we will consider for a given $S_0 > 0$ the process :

$$S_t := S_0\mathcal{E}(B)_t. \tag{4.2.1}$$

We shall consider the subset $\mathcal{P}_m^b$ of $\mathcal{P}_S^b$ consisting of all $\mathbb{P} \in \mathcal{P}_S^b$ such that $\mathbb{E}^{\mathbb{P}}[Z_T] = 1$, and such that $M := (Z\mathcal{E}(B))_.$ and $(ZB)_.$ are uniformly integrable $\mathbb{P}$-martingales. Notice that since we are in finite time horizon, and since M is a $\mathbb{P}$-supermartingale for all $\mathbb{P} \in \mathcal{P}_S^b$, the condition that M is a uniformly integrable $\mathbb{P}$-martingale is equivalent to $\mathbb{E}^{\mathbb{P}}[M_T] = 1$.

The process Z corresponds to the density of the risk-neutral probability measures. By our assumption, we assume that the probability $\mathbb{Q} := Z_T \cdot \mathbb{P}$ is well defined. Moreover we impose that under $\mathbb{Q}$ the underlying S is a uniformly integrable martingale. This assumption is in line with the optimal transport setup developed in [46] for the superreplication of lookback options. The last condition (the canonical process is a uniformly integrable martingale) is technical and ensures that the liability $\ln(Z_T)$ is perfectly replicable by a mix of static and dynamic strategies. This is the point of Lemma 4.3.2.

Remark 4.2.4. *One important difference between our setup and the classical optimal transport formulation is that in our context, the canonical process is not the underlying, but is used to define its dynamics.*

We then introduce for any $\mathbb{P} \in \mathcal{P}_m^b$ the measure $\mathbb{Q}$ defined by:

$$\mathbb{Q} := Z_T \cdot \mathbb{P}.$$

And the set of all measures $\mathbb{Q}$ is denoted $\mathcal{M}^{S,b}$. A probability $\mathbb{Q}$ is uniquely defined and under $\mathbb{Q}$ the canonical process follows the law:

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

Finally the set $\mathcal{M}^{S,b}$ is given by:

Proposition 4.2.1. *$\mathcal{M}^{S,b}$ is the set of all $\mathbb{Q} \in \mathcal{M}^S_{loc}$ such that:*
(i) *B is a uniformly integrable $\mathbb{Q}$-martingale,*
(ii) *$\mathcal{E}(B)$ is a uniformly integrable $\mathbb{Q}$-martingale,*
(iii) $\mathbb{E}^{\mathbb{Q}}[Z_T^{-1}] = 1$.

Proof. if $\mathbb{Q} \in \mathcal{M}^{S,b}$, then $\mathbb{Q} = Z_T \cdot \mathbb{P}$ for some $\mathbb{P} \in \mathcal{P}^b_m$. Then we know by Lemma 4.2.1 that $B.$ and $\mathcal{E}(B).$ are uniformly integrable $\mathbb{Q}$-martingale, so that (i) and (ii) are true.
Since $Z_T > 0$ $\mathbb{P}$-a.s., we have that $\mathbb{P} = Z_T^{-1} \cdot \mathbb{Q}$ and $1 = \mathbb{E}^{\mathbb{P}}[1] = \mathbb{E}^{\mathbb{Q}}\left[Z_T^{-1}\right]$, so that (iii) is true.

Now consider $\mathbb{Q} \in \mathcal{M}^S_{\text{loc}}$ such that (i), (ii) and (iii) are true. Then by (iii) the probability measure $\mathbb{P} := Z_T^{-1} \cdot \mathbb{Q}$ is well defined. Under $\mathbb{P}$ the canonical process has the law: $\mathbb{P}_0 \circ (X^\sigma)^{-1}$, where σ is such that $\mathbb{Q} := \mathbb{P}_0 \circ (X^{\sigma,0})^{-1}$. Since $Z_T^{-1} > 0$ $\mathbb{Q}$-a.s., we can consider the process Z_T and we have $\mathbb{P} = Z_T^{-1} \cdot \mathbb{Q}$, so $\mathbb{E}^{\mathbb{P}}[Z_T] = \mathbb{E}^{\mathbb{Q}}[1] = 1$.
It remains to prove that M and (ZB) are uniformly integrable, which is exactly the formulation of (i) an (ii).
We then have that $\mathbb{P} \in \mathcal{P}^b_m$, and so $\mathbb{Q} = Z_T \cdot \mathbb{P}$ is in $\mathcal{M}^{S,b}$.

□

Remark 4.2.5. *It is then easy to see that $\mathcal{M}^{S,b}$ is a subset of $\mathcal{M}^S$ with equality if and only if $b = 0$.*

For any family $\mathcal{P}$ of probability measures, we will say that a property holds $\mathcal{P}$-quasi surely (q.s. for short) if it holds $\mathbb{P}$-almost surely for all $\mathbb{P} \in \mathcal{P}$.

In the following the process $Z_.^{-1}$ will be denoted $\tilde{Z}$.

4.2.2 Financial market description

We consider a financial market consisting of a tradable asset S with dynamic defined by (4.2.1). The law of S under $\mathbb{P}^\sigma$ is the same as the law of S^σ under $\mathbb{P}_0$, where:

$$\frac{dS^\sigma_t}{S^\sigma_t} = \sigma_t dB_t + b\sigma_t^2 dt, \ \mathbb{P}_0 - \text{a.s.}$$

The choice of dynamic corresponds to an asset manager believing that the sharp ratio is constant, but with no idea of the value of the volatility.

In addition to the continuously tradable asset S, we assume that for the maturity T, Calls and Puts of every strikes K are tradable. Under the asssumption of linearity of the pricing functionnal, we know that one can identify the T-marginal of S denoted μ. In our framework μ is a measure on $\mathbb{R}_+$. Then the only no-arbitrage price of a european payoff $\lambda \in \mathbb{L}^1(\mu)$ at time 0 is given by $\mu(\lambda) := \int \lambda(s)\mu(ds)$. In that context, the trader is allowed to take statically any european position.

A strategy for the asset manager is then a pair (H, λ) where λ is the static European position, and H corresponds to the dynamic strategies. The final wealth induced from such a semi-static hedging strategy, starting from X_0, is:

$$X^{H,\lambda}_T := X_0 + \int_0^T H_t dB_t + \lambda(S_T) - \mu(\lambda) = X^H_T + \lambda(S_T) - \mu(\lambda). \tag{4.2.2}$$

In all the following, we will consider a liability ξ that the agent has to deliver at time T, so that its final wealth, considering the choice of a strategy (H, λ) is given by:

$$X_0 + \int_0^T H_t dB_t + \lambda(S_T) - \mu(\lambda) - \xi = X_T^H - \xi^\lambda, \tag{4.2.3}$$

where we denote $\xi^\lambda := \xi - \lambda(S_T) + \mu(\lambda)$ for short.

4.2.3 Static and dynamic strategies

In this section, we set the admissible conditions for the stategies. We consider an upper semi-continuous map ξ and a parameter $\eta > 0$ defined in the next subsection.

For any upper-semianalytic map ζ, we restrict the probability families $\mathcal{P}_m^b$ and $\mathcal{M}^{S,b}$ by:

$$\mathcal{M}_\zeta^{S,b} := \{\mathbb{Q} \in \mathcal{M}^{S,b} : \ \mathbb{E}^{\mathbb{Q}}\left[\zeta^-\right] < +\infty\},$$

$$\mathcal{P}_\zeta^b := \{\mathbb{P} \in \mathcal{P}_m^b : \ \mathbb{Q} = Z_T \cdot \mathbb{P} \in \mathcal{M}_\zeta^{S,b}\}.$$

This restriction is motivated by the exclusion of probabilities which induces arbitrage opportunities.

We now describe the set of static strategies Λ^μ by the subset of $\mathbb{L}^1(\mu)$:

$$\Lambda^\mu := \left\{\lambda \in \mathbb{L}^1(\mu) \ : \ \sup_{\mathbb{Q} \in \mathcal{M}^{S,b}} \mathbb{E}^{\mathbb{Q}}\left[\left(\frac{b^2}{\eta}\ln(S_T) + \xi^\lambda\right)^+\right] < \infty\right\}.$$

It now remains to define the set of admissible dynamic strategies. We define the dynamic version of X_T^H defined in (4.2.2) by:

$$X_\cdot^H := X_0 + \int_0^\cdot H_t dB_t.$$

For all $\lambda \in \Lambda^\mu$, we define the set of admissible dynamic strategies by $\mathcal{H}^\lambda$ where:

$$\mathcal{H}^\lambda := \bigcap_{\mathbb{Q} \in \mathcal{M}_{\xi^\lambda}^{S,b}} \mathcal{H}(\mathbb{Q}), \quad \text{where:}$$

$$\mathcal{H}(\mathbb{Q}) := \{H \in \mathbb{H}^2_{\text{loc}}(\mathbb{Q}, \mathbb{R}, \mathbb{F}_+^*), X^H \text{ is a } \mathbb{Q}-\text{supermartingale}\}.$$

4.2.4 Robust utility maximization and indifference pricing

We recall that the subset of a polish space is said to be analytic if it is the image of a borel subset of another polish space by a borel measurable map. In all the following, we consider an upper-semianalytic map $\xi : \Omega \to \mathbb{R}$, i.e. for all $c \in \mathbb{R}$, $\{\xi > c\}$ is analytic.

We consider the exponential utility function $U := -e^{-\lambda \cdot} : \mathbb{R} \mapsto \mathbb{R}$, non decreasing. Our aim is to study the utility maximisation under volatility uncertainty of a portfolio containing the liability $-\xi$. The formulation of this problem is then:

$$V(\xi, X_0) = \sup_{(H,\lambda) \in \mathcal{H}^\lambda \times \Lambda^\mu} \inf_{\mathbb{P} \in \mathcal{P}_{\xi^\lambda}^b} \mathbb{E}\left[U(X_T^H - \xi^\lambda)\right]. \tag{4.2.4}$$

The problem of robust utility indifferent pricing of the payoff ξ is defined by:

$$\bar{p}(\xi, X_0) := \inf\{p \in \mathbb{R} \ : \ V(\xi, X_0 + p) \geq V(0, X_0)\},$$

where $V(0,\cdot)$ corresponds to the problem (4.2.4) with zero liability.

We clearely have

$$V(\xi, X_0) = \sup_{\lambda\in\Lambda^\mu} V^\lambda(\xi, X_0),$$

where:

$$V^\lambda(\xi, X_0) = \sup_{H\in\mathcal{H}^\lambda} \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^\lambda}} \mathbb{E}\left[U(X_T^H - \xi^\lambda)\right].$$

The whole study will be made on the quantity V^λ and we end all our calculus by taking the supremum over λ in Λ^μ. We introduce

$$\tilde\Lambda^\mu := \left\{\tilde\lambda \in \mathbb{L}^1(\mu) \ : \ \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}} \mathbb{E}^{\mathbb{Q}}\left[\left(\xi^{\tilde\lambda}\right)^+\right] < \infty\right\}.$$

We show here that:

Theorem 4.2.1. *For all $\lambda \in \Lambda^\mu$, $V^\lambda(\xi, X_0) = e^{-\eta X_0} V_0^\lambda(\xi)$ with:*

$$V_0^\lambda(\xi) = -\exp\left(-b^2\ln(S_0) + b^2\mu(\ln) + \eta h^{\tilde\lambda}(\xi)\right),$$

where for any $f \in \tilde\Lambda^\mu$, $h^f(\xi)$ is the robust superhedging price of the payoff ξ^f:

$$h^f(\xi) := \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^f}} \mathbb{E}^{\mathbb{Q}}\left[\xi^f\right], \tag{4.2.5}$$

and $\tilde\lambda$ is given by: $\tilde\lambda(s) := \lambda(s) - \frac{b^2}{\eta}\ln(s)$.

The proof is reported in Sections 4.3 and 4.4.

Introducing $V_0(\xi) := V(0,\xi)$, we clearly have $V(\xi, X_0) = e^{-\eta X_0}V_0(\xi)$. We then deduce from Theorem 4.2.1:

Theorem 4.2.2. *The value function given in (4.2.4) is:*

$$V_0(\xi) = -\exp(-b^2\ln(S_0) + b^2\mu(\ln) + \eta h(\xi)),$$

where h is the robust super-replication price given marginals and is equal to:

$$h(\xi) = \inf_{\lambda\in\tilde\Lambda^\mu} h^\lambda(\xi).$$

Proof. By Theorem 4.2.1, we have:

$$\begin{aligned} V_0(\xi) &= \sup_{\lambda\in\Lambda^\mu} -\exp\left(-b^2\ln(S_0) + \eta \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda + \frac{b^2}{\eta}\ln(S_T)\right]\right) \\ &= -\exp\left(-b^2\ln(S_0) + \eta \inf_{\lambda\in\Lambda^\mu} \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda + \frac{b^2}{\eta}\ln(S_T)\right]\right). \end{aligned}$$

Then denoting $\tilde\lambda(s) := \lambda(s) - \frac{b^2}{\mu}\ln(s)$, we clearly have:

$$\tilde\lambda \in \tilde\Lambda^\mu \Leftrightarrow \lambda \in \Lambda^\mu.$$

And:

$$\mathbb{E}^{\mathbb{Q}}\left[\xi^{\lambda}+\frac{b^2}{\eta}\ln\left(S_T\right)\right]=\mathbb{E}^{\mathbb{Q}}\left[\xi^{\tilde{\lambda}}\right]+\frac{b^2}{\eta}\mu\left(\ln\right).$$

So:

$$V_0(\xi)=-\exp(-b^2\ln(S_0)+b^2\mu(\ln)-\eta h(\xi)),$$

with:

$$h(\xi)=\inf_{\tilde{\lambda}\in\tilde{\Lambda}^{\mu}}\sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^{\tilde{\lambda}}}}\mathbb{E}^{\mathbb{Q}}\left[\xi+\tilde{\lambda}(S_T)-\mu(\tilde{\lambda})\right].$$

□

The robust indifference price is given by:

Corollary 4.2.1. *The robust indifference price of a payoff ξ is independant of X_0 and is equal to the robust super-replication price given marginals:*

$$\bar{p}(\xi)=\inf_{\tilde{\lambda}\in\tilde{\Lambda}^{\mu}}\sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^{\tilde{\lambda}}}}\mathbb{E}^{\mathbb{Q}}\left[\xi^{\tilde{\lambda}}\right].$$

We also emphasize on the link between the utility maximisation problem (4.2.4) and the super-replication problem:

$$\mathcal{U}(\xi):=\inf\left\{X_0\in\mathbb{R},\ s.t.\ \exists(H,\lambda)\in\mathcal{H}^{\lambda}\times\tilde{\Lambda}^{\mu}:\ X_T^{H,\lambda}\geq\xi\ \mathcal{M}^{S,b}_{\xi^{\lambda}}-q.s.\right\}. \tag{4.2.6}$$

As a consequence of Theorem 4.4.1 proved rigorously in section 4.4.2, we have that $\mathcal{U}(\xi)=h(\xi)$ where h is defined in Theorem 4.2.2.
We know that in particular cases, there exists a solution (H,λ) to the super-replication problem (4.2.6), see for example the application to lookback options developed in [46] and [51]. Then in those cases, we can derive the optimal solution to the utility maximisation problem (4.2.4).

Corollary 4.2.2. *Assume that there exists a solution $(H,\lambda)\in\mathcal{H}^{\lambda}\times\tilde{\Lambda}^{\mu}$ of the super-replication problem (4.2.6), then the optimal solution of (4.2.4) exists and is given by:*

$$\begin{cases} H_t^* &= H_t-\frac{b^2-b}{\eta} \\ \lambda^*(s) &= \lambda(s)+\frac{b^2}{\eta}\ln\left(\frac{s}{S_0}\right).\end{cases}$$

4.3 First inequality by duality

In this section, we use the classical approach of utility maximisation by duality, which provides an efficient method to obtain an upper bound for the problem.

In all the following, we fix $\lambda\in\Lambda^{\mu}$. In this section, we want to show that:

$$V_0^{\lambda}(\xi)\leq-\exp(-b^2\ln(S_0)+b^2\mu(\lambda)-\eta h^{\tilde{\lambda}}(\xi)),$$

where we recall that $\tilde{\lambda}(s):=\lambda(s)-\frac{b^2}{\eta}\ln(s)$.

We next introduce $\bar{V}^{\lambda}(\xi,X_0)$ the upper value of the game:

$$\bar{V}^{\lambda}(\xi,X_0):=\inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}}\sup_{H\in\mathcal{H}^{\lambda}(\mathbb{Q})}\mathbb{E}^{\mathbb{P}}\left[U\left(X_T^H-\xi^{\lambda}\right)\right]. \tag{4.3.1}$$

We clearly have that :

$$V^{\lambda}(\xi,X_0)\leq\inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}}\sup_{H\in\mathcal{H}^{\lambda}}\mathbb{E}^{\mathbb{P}}\left[U\left(X_T^H-\xi^{\lambda}\right)\right]\leq\bar{V}^{\lambda}(\xi,X_0). \tag{4.3.2}$$

Remark 4.3.1. *We emphasize here on the different admissible set of continuous strategies in the maximization problems $V_0^\lambda(\xi)$ and $\bar{V}_0^\lambda(\xi)$. Indeed in the robust formulation problem $V_0^\lambda(\xi)$, an admissible strategy H must be defined for any probability $\mathbb{P} \in \mathcal{P}_{\xi^\lambda}$ while the upper value of the game involves a maximization with a portfolio $H^{\mathbb{P}}$ for all $\mathbb{P}$ and we do not require the existence of an aggregator H such that $H = H^{\mathbb{P}}$ $\mathbb{P}$-a.s. for all $\mathbb{P}$.*

We consider the Legendre-Fenchel transform of U. For every positive real number y, we have:

$$\tilde{U}(y) = \sup_{x\in\mathbb{R}} (U(x) - xy) = -\frac{y}{\eta} + \frac{y}{\eta}\ln(\frac{y}{\eta}).$$

The optimal value of x is reached for:

$$x^* = \left(U'\right)^{-1}(y) = -\frac{1}{\eta}\ln(\frac{y}{\eta}).$$

We start by:

Lemma 4.3.1. *we have:*

$$\inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^\lambda}} \inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right) + yZ_T\left(x - \xi^\lambda\right)\right] = \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^\lambda}} \sup_{H\in\mathcal{H}^\lambda(\mathbb{Q})} \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^{x,H} - \xi^\lambda\right)\right].$$

Proof. By definition of $\tilde{U}$ we have for every $(x, y) \in \mathbb{R}\times\mathbb{R}_+$ and $H \in \mathcal{H}^\lambda(\mathbb{Q})$:

$$\tilde{U}\left(yZ_T\right) \geq U\left(X_T^{x,H} - \xi^\lambda\right) - yZ_T\left(X_T^{x,H} - \xi^\lambda\right). \tag{4.3.3}$$

Now taking the expectation, because $X^{x,H}$ is a supermartingale under $\mathbb{Q}$, we have:

$$\mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right)\right] + y\left(x - \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda\right]\right) \geq \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^{x,H} - \xi^\lambda\right)\right].$$

The last inequality is true for all $y \geq 0$ and all $H \in \mathcal{H}^\lambda(\mathbb{Q})$. Then:

$$\inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right)\right] + y\left(x - \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda\right]\right) \geq \sup_{H\in\mathcal{H}^\lambda(\mathbb{Q})} \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^{x,H} - \xi^\lambda\right)\right].$$

When it comes to the equality, we are looking for H and y such that we have equality almost surely in (4.3.3), i.e. we want that:

$$X_T^H - \xi^\lambda = \left(U'\right)^{-1}\left(yZ_T\right),\ \mathbb{P} - \text{a.s.} \tag{4.3.4}$$

We will restrict the search to the H such that X^H is a martingale. A necessary condition to have (4.3.4), is then the equality of expectations:

$$\mathbb{E}^{\mathbb{Q}}\left[\left(U'\right)^{-1}\left(yZ_T\right) + \xi^\lambda\right] = \mathbb{E}^{\mathbb{Q}}\left[X_T^H\right] = x,$$

where $\mathbb{Q} := Z_T \cdot \mathbb{P}$.

Now since $\left(U'\right)^{-1}(z) = -\frac{1}{\eta}\ln\left(\frac{y}{\eta}\right)$, we have:

$$\mathbb{E}^{\mathbb{Q}}\left[\left(U'\right)^{-1}\left(yZ_T\right) + \xi^\lambda\right] = \left(U'\right)^{-1}(y) + \mathbb{E}^{\mathbb{Q}}\left[\left(U'\right)^{-1}\left(Z_T\right) + \xi^\lambda\right].$$

By assumptions on $\mathbb{P}$ and λ, we have that the last quantity is finite and the properties of $\left(U'\right)^{-1}$ ensure the existence and uniqueness of $\hat{y}^{\mathbb{P}}$ such that:

$$\mathbb{E}^{\mathbb{Q}}\left[\left(U'\right)^{-1}\left(\hat{y}^{\mathbb{P}}Z_T\right) + \xi^\lambda\right] = x.$$

Moreover we have by the martingale representation property, the existence of $H^{\mathbb{Q}} \in \mathcal{H}^{\lambda}(\mathbb{Q})$ such that:

$$x + \int_0^T H_t^{\mathbb{Q}} dt = \left(U'\right)^{-1}\left(\hat{y}^{\mathbb{P}} Z_T\right) + \xi^{\lambda}, \quad \mathbb{Q}-\text{a.s.}$$

Since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent martingale measures, the last equality is also true $\mathbb{P}$ almost surely, so that:

$$\inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right)\right] + y\left(x - \mathbb{E}^{\mathbb{Q}}\left[\xi^{\lambda}\right]\right) = \sup_{H\in\mathcal{H}^{\lambda}(\mathbb{Q})} \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^{x,H} - \xi^{\lambda}\right)\right].$$

Finally taking the infimum over all $\mathbb{P} \in \mathcal{P}^b_{\xi^{\lambda}}$, we have:

$$\inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right)\right] + y\left(x - \mathbb{E}^{\mathbb{Q}}\left[\xi^{\lambda}\right]\right) = \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \sup_{H\in\mathcal{H}^{\lambda}(\mathbb{Q})} \mathbb{E}^{\mathbb{P}}\left[U\left(X_T^{x,H} - \xi^{\lambda}\right)\right].$$

□

We now characterize the value of $\bar{V}^{\lambda}$ to obtain an upper bound for the problem:

Proposition 4.3.1. *Assume that $\lambda \in \Lambda^{\mu}$, then:*

$$\bar{V}_0^{\lambda}(\xi) = -\exp\Big(\sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^{\lambda}}} \mathbb{E}^{\mathbb{Q}}\left[\eta\xi^{\lambda} - \ln\left(Z_T\right)\right]\Big).$$

Proof. By Lemma 4.3.1, we have that:

$$\bar{V}^{\lambda}(\xi, X_0) = \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right) + yZ_T\left(X_0 - \xi^{\lambda}\right)\right].$$

The proof allows us to see that for a given $\mathbb{Q} \in \mathcal{M}^{S,b}_{\xi^{\lambda}}$, we have:

$$\inf_{y\geq 0} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(yZ_T\right) + yZ_T\left(X_0 - \xi^{\lambda}\right)\right] = \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(\hat{y}^{\mathbb{P}} Z_T\right) + \hat{y}^{\mathbb{P}} Z_T\left(X_0 - \xi^{\lambda}\right)\right],$$

where $\hat{y}^{\mathbb{P}}$ verifies:

$$\mathbb{E}^{\mathbb{Q}}\left[\left(U'\right)^{-1}\left(\hat{y}^{\mathbb{P}} Z_T\right) + \xi^{\lambda}\right] = X_0,$$

i.e.:

$$-\frac{1}{\eta}\ln\left(\frac{\hat{y}^{\mathbb{P}}}{\eta}\right) = X_0 - \mathbb{E}^{\mathbb{Q}}\left[\xi^{\lambda} - \frac{1}{\eta}\ln\left(Z_T\right)\right]. \tag{4.3.5}$$

Since $\tilde{U}(y) = -\frac{y}{\eta} + \frac{y}{\eta}\ln\left(\frac{y}{\eta}\right)$ and using (4.3.5), we have:

$$\begin{aligned}
\bar{V}^{\lambda}(\xi, X_0) &= \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \mathbb{E}^{\mathbb{P}}\left[\tilde{U}\left(\hat{y}^{\mathbb{P}} Z_T\right) + \hat{y}^{\mathbb{P}} Z_T\left(X_0 - \xi^{\lambda}\right)\right] \\
&= \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \Big\{ -\frac{\hat{y}^{\mathbb{P}}}{\eta} + \hat{y}^{\mathbb{P}}\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{\eta}\ln\left(\hat{y}^{\mathbb{P}} Z_T\right) + X_0 - \xi^{\lambda}\right] \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \Big\{ -\frac{\hat{y}^{\mathbb{P}}}{\eta} \Big\} \\
&= \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^{\lambda}}} \Big\{ -\exp\left(-\eta X_0 + \mathbb{E}^{\mathbb{Q}}\left[\eta\xi^{\lambda} - \ln\left(Z_T\right)\right]\right) \Big\} \\
&= -\exp\Big(-\eta X_0 - \inf_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^{\lambda}}} \mathbb{E}^{\mathbb{Q}}\left[-\eta\xi^{\lambda} + \ln\left(Z_T\right)\right]\Big).
\end{aligned}$$

□

This proposition involves the relative entropy $H(\mathbb{Q}|\mathbb{P})$ of $\mathbb{Q}$ from $\mathbb{P}$ defined by:

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

In the following lemma, we establish the connection between relative entropy and semi-static strategy under the hypothesis that $\mathbb{P}\in\mathcal{P}^b_m$.

Lemma 4.3.2. *For any $\mathbb{P}\in\mathcal{P}^b_m$, we have for $\mathbb{Q}=Z_T\cdot\mathbb{P}$:*

$$\begin{aligned} H(\mathbb{Q}|\mathbb{P}) &:= \mathbb{E}^{\mathbb{P}}\left[Z_T\ln\left(Z_T\right)\right] \\ &= b^2\ln(S_0) - b^2\mathbb{E}^{\mathbb{Q}}\left[\ln\left(S_T\right)\right]. \end{aligned}$$

Moreover we have $\mathcal{P}^b_m$-q.s. that:

$$\ln(Z_T) = b^2\ln(S_0) - b^2\ln(S_T) + (b^2-b)B_T.$$

Proof. We recall from the definitions that $\mathcal{P}^b_m$-q.s.:

$$\ln(Z_T) = -bB_T + \frac{1}{2}b^2\langle B\rangle_T, \text{ and } \ln(S_T) = \ln(S_0) + B_T - \frac{1}{2}\langle B\rangle_T.$$

We then have $\mathcal{P}^b_m$-q.s.

$$\ln(Z_T) = b^2\ln(S_0) - b^2\ln(S_T) + (b^2-b)B_T.$$

We then have for all $\mathbb{P}\in\mathcal{M}^{S,b}$, and $\mathbb{Q}=Z_T\cdot\mathbb{P}$:

$$H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}^{\mathbb{Q}}\left[\ln(Z_T)\right] = b^2\ln(S_0) - b^2\mathbb{E}^{\mathbb{Q}}\left[\ln(S_T)\right].$$

The last equality comes from the fact that the canonical process B is a $\mathbb{Q}$-martingale for all $\mathbb{Q}\in\mathcal{M}^{S,b}$.

□

Remark 4.3.2. *We observe in this proof that the result does not hold if we do not assume that b is constant.*

Remark 4.3.3. *The martingale property of the canonical process allows us to have a perfect replication strategy for the entropy payoff $\ln(Z_.)$ under $\mathbb{Q}$, using static and dynamic strategies. The dynamic strategy is given by $H_t := \frac{b^2-b}{S_t}$ and the position is given by $\lambda(x) := -b^2\ln\left(\frac{x}{S_0}\right)$.*

4.4 Second inequality by quasi-sure analysis

For the second inequality, we use a superreplication argument, by quasi-sure analysis. For that purpose we use the framework introduced in [77] and [74], and then developed in Chapter 3 for the application to the optimal transport setup. This new approach of super-replication results uses the analytic tools and results developed in [11].

4.4.1 Regular Conditional Probability Distribution

In this section, we recall the notion of regular conditional probability distribution (r.c.p.d.), as introduced by Stroock and Varadhan [95]. Let $\mathbb{P} \in \mathbf{M}(\Omega)$ and consider some $\mathbb{F}$-stopping time τ. Then, for every $\omega \in \Omega$, there exists an r.c.p.d. $\mathbb{P}_\tau^\omega$ satisfying:
(i) $\mathbb{P}_\tau^\omega$ is a probability measure on $\mathcal{F}_T$.
(ii) For each $E \in \mathcal{F}_T$, the mapping $\omega \to \mathbb{P}_\tau^\omega(E)$ is $\mathcal{F}_\tau$-measurable.
(iii) $\mathbb{P}_\tau^\omega$ is a version of the conditional probability measure of $\mathbb{P}$ on $\mathcal{F}_\tau$, i.e., for every integrable $\mathcal{F}_T$-measurable r.v. ξ we have $\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_\tau](\omega) = \mathbb{E}^{\mathbb{P}_\tau^\omega}[\xi]$, $\mathbb{P}-$a.s.
(iv) $\mathbb{P}_\tau^\omega(\Omega_\tau^\omega) = 1$, where $\Omega_\tau^\omega := \{\omega' \in \Omega : \ \omega'(s) = \omega(s), \ 0 \leq s \leq \tau(\omega)\}$.

We next introduce the shifted canonical space and the corresponding notations.

- For $0 \leq t \leq T$, denote by $\Omega^t := \{\omega \in C([t,T],\mathbb{R}) \ : \ w(t) = 0\}$ the shifted canonical space, B^t the shifted canonical process on Ω^t, $\mathbb{P}_0^t$ the shifted Wiener measure, $\mathbb{F}^t$ the shifted filtration generated by B^t.
- For $0 \leq s \leq t \leq T$, $\omega \in \Omega^s$, define the shifted path $\omega^t \in \Omega^t$, $\omega_r^t := \omega_r - \omega_t$ for all $r \in [t,T]$.
- For $0 \leq s \leq t \leq T$, $\omega \in \Omega^s$, define the concatenation path $\omega \otimes_t \widetilde{\omega} \in \Omega^s$ by:

$$(\omega \otimes_t \widetilde{\omega})(r) := \omega_r \mathbf{1}_{[s,t)}(r) + (\omega_t + \widetilde{\omega}_r)\mathbf{1}_{[t,1]}(r) \ \text{ for all } \ r \in [s,T].$$

- For $0 \leq s \leq t \leq T$, for any $\mathcal{F}_T^s$-measurable random variable ζ on Ω^s, and for each $\omega \in \Omega^s$, define the shifted $\mathcal{F}_T^t$-measurable random variable $\zeta^{t,\omega}$ on Ω^t by:

$$\zeta^{t,\omega}(\widetilde{\omega}) := \zeta(\omega \otimes_t \widetilde{\omega}) \ \text{ for all } \ \widetilde{\omega} \in \Omega^t.$$

- The r.c.p.d. $\mathbb{P}_\tau^\omega$ induces naturally a probability measure $\mathbb{P}^{\tau,\omega}$ on $\mathcal{F}_T^{\tau(\omega)}$ such that the $\mathbb{P}^{\tau,\omega}$-distribution of $B^{\tau(\omega)}$ is equal to the $\mathbb{P}_\tau^\omega$-distribution of $\{B_t - B_{\tau(\omega)}, \ t \in [\tau(\omega), T)\}$. It is then clear that for every integrable and $\mathcal{F}_T$-measurable random variable ζ,

$$\mathbb{E}^{\mathbb{P}_\tau^\omega}[\zeta] = \mathbb{E}^{\mathbb{P}^{\tau,\omega}}[\zeta^{\tau,\omega}].$$

For the sake of simplicity, we shall also call $\mathbb{P}^{\tau,\omega}$ the r.c.p.d. of $\mathbb{P}$.

- Finally, we introduce for all $(s,\omega) \in [0,T] \times \Omega$:

$$\mathcal{M}_{\text{loc}}^S(s,\omega) := \left\{ \mathbb{P}_0^s \circ \left(\int_s^{\cdot} \alpha_u^{1/2} dB_u^s \right)^{-1}, \text{ with } \int_s^T |\alpha_u| du < +\infty, \ \mathbb{P}_0^s - \text{a.s.} \right\}$$

$$\mathcal{M}^{S,b}(s,\omega) := \Big\{ \mathbb{P} \in \mathcal{P}_S(s,\omega) \text{ s.t. } \mathbb{E}^{\mathbb{P}}[(Z^{s,\omega})^{-1}] = 1, \\ \mathcal{E}(B)^{s,\omega} \text{ and } B^{s\omega} \text{ are u.i. } \mathbb{P} - \text{martingales} \Big\}.$$

Remark 4.4.1. *We are abusing notations here. In order to suit to the definition of [74] and [77], we should have considered the concatenation of $\mathcal{M}_{loc}^S(s,\omega)$ (resp. $\mathcal{M}^{S,b}(s,\omega)$) defined above with the Dirac mass on $\omega_{0 \leq t \leq s}$ to ensure that elements of $\mathcal{M}_{loc}^S(s,\omega)$ (resp. $\mathcal{M}^{S,b}(s,\omega)$) are probabilities on Ω, and not on Ω^s. The reader should note that the link between these two definitions is obvious and we will implicitly identify these families.*

It is clear that the families $(\mathcal{M}_{\text{loc}}^S(s,\omega))_{(s,\omega)\in[0,T]\times\Omega}$ and $(\mathcal{M}^{S,b}(s,\omega))_{(s,\omega)\in[0,T]\times\Omega}$ are adapted in the sense that $\mathcal{M}_{\text{loc}}^S(s,\omega) = \mathcal{M}_{\text{loc}}^S(s,\widetilde{\omega})$ and $\mathcal{M}^{S,b}(s,\omega) = \mathcal{M}^{S,b}(s,\widetilde{\omega})$, whenever $\omega|_{[0,s]} = \widetilde{\omega}|_{[0,s]}$.

4.4.2 The duality result

We consider a generic subset $\mathcal{P}$ of $\mathcal{P}_S$ (and the corresponding shifted families $\mathcal{P}(s,\omega)$). We recall that such a family is said to be stable by bifurcation if for any $\mathbb{F}$-stopping times $0 \le \sigma \le \tau$, $\omega \in \Omega$, A $\mathcal{F}_\tau$-measurable, $\mathbb{P}_1$ and $\mathbb{P}_2$ in $\mathcal{P}(\sigma,\omega)$, we have

$$\mathbb{P} = \mathbb{P}_1 \mathbf{1}_A + \mathbb{P}_2 \mathbf{1}_{A^c} \in \mathcal{P}(\sigma,\omega).$$

We assume moreover that the family $\mathcal{P}(s,\omega)$ satisfies:

Condition 4.4.1. *Let $s \in \mathbb{R}_+$, $\tau \ge s$ a stopping time, $\overline{\omega} \in \Omega$, $\mathbb{P} \in \mathcal{P}(s,\overline{\omega})$ and $\theta := \tau^{s,\overline{\omega}} - s$.*
(i) *The graph $\{(\mathbb{P}',\omega) : \omega \in \Omega,\ \mathbb{P}' \in \mathcal{P}(t,\omega)\} \subseteq \mathbf{M}(\Omega) \times \Omega$ is analytic.*
(ii) *We have $\mathbb{P}^{\theta,\omega} \in \mathcal{P}(\tau, \overline{\omega} \otimes_s \omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*
(iii) *$\mathcal{P}$ is stable by bifurcation.*
(iv) *If $\nu : \Omega \to \mathbf{M}(\Omega)$ is an $\mathcal{F}_\theta$-measurable kernel and $\nu(\omega) \in \mathcal{P}(\tau, \overline{\omega} \otimes_s \omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, then there exists $\nu^n : \Omega \to \mathbf{M}(\Omega)$, which is a $\mathcal{F}_\theta$-measurable kernel such that $\mathbb{P}(\nu^n = \nu) \underset{n\to\infty}{\longrightarrow} 1$ and the following measure $\overline{\mathbb{P}}^n \in \mathcal{P}(s,\overline{\omega})$:*

$$\overline{\mathbb{P}}^n(A) = \iint (\mathbf{1}_A)^{\theta,\omega}(\omega')\nu^n(d\omega';\omega)\mathbb{P}(dw), \quad A \in \mathcal{F}.$$

Similarly to our previous notations, we introduce the sets $\mathcal{H}(\mathcal{P})$ and $\mathcal{P}_\zeta$ for ζ upper semi-analytic. In this context, we obtain the Theorem:

Theorem 4.4.1 (Theorem 5.1 in Chapter 3)**.** *Let $\mathcal{P}(s,\omega)$ be a family of probability measures satisfying Condition 4.4.1. Let ζ be an upper semi-analytic r.v. with $\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\zeta^+] < +\infty$ and $\zeta \in \mathbb{L}^1(\mathbb{P})$ for all $\mathbb{P} \in \mathcal{P}$. Then*

$$V(\zeta) \quad := \quad \inf\Big\{X_0 :\ X_T^H \ge \xi,\ \ \mathcal{P}^\zeta - q.s.\ for\ some\ H \in \mathcal{H}(\mathcal{P})\Big\} \ = \ \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\zeta].$$

Moreover, existence holds for the primal problem, i.e. $V(\zeta) + \int_0^T H_s dB_s \ge \zeta$, $\mathcal{P}^\zeta$−q.s. for some $H \in \mathcal{H}^\zeta(\mathcal{P})$.

In order to apply Theorem 4.4.1, we need to show:

Proposition 4.4.1. *For any $\lambda \in \Lambda^\mu$, the family $\left(\mathcal{M}^{S,b}_{\xi^\lambda}(s,\omega)\right)$ verifies Condition 4.4.1.*

Proof. The proof of 4.4.1 (i) is verified in a separate lemma.
We first verify Condition 4.4.1 (ii). Let $\mathbb{Q} \in \mathcal{M}^{S,b}_{\xi^\lambda}$, and consider an arbitrary $\mathbb{F}$-stopping time τ, and $\mathbb{F}^\tau$-stopping time σ. By Lemma A.1 in [66], there exists some $\mathbb{F}^\tau$-stopping time $\tilde{\sigma}$ such that for every ω, $\tilde{\sigma}^{\tau,\omega} = \sigma$. Then, we have for $\mathbb{Q}$-a.e. ω

$$\mathbb{E}^{\mathbb{Q}^{\tau,\omega}}[|B^\tau_\sigma|] \le \mathbb{E}^{\mathbb{Q}^{\tau,\omega}}[|B^{\tau,\omega}_{\tilde{\sigma}^{\tau,\omega}}|] + |B_\tau(\omega)| = \mathbb{E}^{\mathbb{Q}}_\tau[|B_{\tilde{\sigma}}|](\omega) + |B_\tau|(\omega) < +\infty,$$

where we used the fact that $\mathbb{Q} \in \mathcal{Q}_m B$. Similarly, we have for $\mathbb{Q}$-a.e. ω:

$$\mathbb{E}^{\mathbb{Q}^{\tau,\omega}}[B^\tau_\sigma] = \mathbb{E}^{\mathbb{Q}^{\tau,\omega}}[B^{\tau,\omega}_{\tilde{\sigma}^{\tau,\omega}} - B_\tau(\omega)] = \mathbb{E}^{\mathbb{Q}}_\tau[B_{\tilde{\sigma}}](\omega) - B_\tau(\omega) = 0.$$

By the arbitrariness of σ, this shows that $B^{\tau,\omega}$ is a uniformly integrable martingale. By the exact same argument we have that $e^{B^{\tau,\omega} - \frac{1}{2}\langle B^{\tau,\omega}\rangle}$ is a uniformly integrable martingale.

For the property of $\tilde{Z}^{\tau,\omega}$, since it is a supermartingale, we have $E^{\mathbb{Q}^{\tau,\omega}}[\tilde{Z}^{\tau,\omega}_T] \le 1$. Then we have that for $\mathbb{Q}$-a.e. ω:

$$1 = \mathbb{E}^{\mathbb{Q}}[\tilde{Z}_T] = \mathbb{E}^{\mathbb{Q}}\left[\tilde{Z}_\tau \mathbb{E}^{\mathbb{Q}^{\tau,\omega}}[\tilde{Z}^{\tau,\omega}_T]\right] \le \mathbb{E}^{\mathbb{Q}}[\tilde{Z}_\tau] = 1$$

So we have equality everywhere and $E^{\mathbb{Q}^{\tau,\omega}}[\tilde{Z}_T^{\tau,\omega}] = 1$ for $\mathbb{Q}$-a.e. ω, which ends the proof of Condition 4.4.1 (ii).

The stability by bifurcation is trivial here and corresponds to Condition 4.4.1 (iii). It then allows us to consider for any $\mathcal{F}_\tau$-measurable kernel ν and $\mathbb{Q} \in \mathcal{M}_{\xi^\lambda}^{S,b}$, we define ν^n by:

$$\nu^n(\omega) \quad := \quad \nu(\omega)\mathbf{1}_{\{A_n\}} + \mathbb{Q}^{\tau,\omega}\mathbf{1}_{\{A_n^c\}}.$$

where $A_n := \left\{\mathbb{E}^{\nu(\omega)}[|B_T^\tau|] \leq n\right\}$. This ensures the uniform martingale property of the process B as it was proved in Chapter 3. The martingale property of the process $\tilde{Z}$ and $e^{B_\cdot - \frac{1}{2}\langle B\rangle_\cdot}$ are automatically verified since the only problem involves integrability. Since theses local martingales are uniformly bounded by below by 0, they are super martingales and the integrability condition is automatically verified. The proof of $\mathbb{Q}(A_n) \to 1$ when $n \to \infty$ was also obtained in Chapter 3. This ends the proof of Condition 4.4.1 (iv)

□

Lemma 4.4.1. *For any $\lambda \in \Lambda^\mu$, the family $\left(\mathcal{M}_{\xi^\lambda}^{S,b}(s,\omega)\right)$ satisfies Condition 4.4.1(i).*

Proof. We adapt here arguments from Chapter 3 and [74]. We define the following map

$$\psi : \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \to \mathbf{M}(\Omega), \quad \alpha \mapsto \mathbb{P}^\alpha = \mathbb{P}_0 \circ \left(\int_0^\cdot \alpha_s^{1/2} dB_s\right)^{-1}.$$

From [74] (see Lemmas 3.1 and 3.2), we know that it is sufficient to show that $\mathcal{M}_{\xi^\lambda}^{S,b} \subset \mathbf{M}(\Omega)$ is the image of a Borel space (i.e. a Borel subset of a Polish space) under a Borel map. By Lemma 3.1 of [74], the space $\mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F})$ is Polish and $\mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \subset \mathbb{H}^0(\mathbb{S}_d^{>0}, \mathbb{F})$ is Borel.Also by Lemma 3.2 in [74], the map $\psi : \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \to \mathbf{M}(\Omega)$ is Borel.

It then only remains to prove that $\mathbb{H}^1_{\text{b}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \subset \mathbb{H}^1_{\text{loc}}(\mathbb{R}_+, \mathbb{F})$ is Borel,where

$$\mathbb{H}^1_{\text{b}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) = \mathbb{H}^1_{\tilde{Z}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \cap \mathbb{H}^1_{\text{exp}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \cap \mathbb{H}^1_{\text{m}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}),$$

with:

$$\mathbb{H}^1_{\tilde{Z}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) := \left\{\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{R}_+, \mathbb{F}) \ : \ \mathbb{E}^{\mathbb{P}_0}\left[\exp\left(bX_T^\alpha - \frac{1}{2}b^2\langle X^\alpha\rangle_T\right)\right] = 1\right\},$$

$$\mathbb{H}^1_{\text{exp}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) := \left\{\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{R}_+, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|\mathcal{E}(X^\alpha)_\tau|\mathbf{1}_{|\mathcal{E}(X^\alpha)_\tau|>n}] \underset{n\to+\infty}{\longrightarrow} 0\right\},$$

$$\mathbb{H}^1_{\text{m}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) := \left\{\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{S}_d^{>0}, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}\left[|X_\tau^\alpha|\mathbf{1}_{|X_\tau^\alpha|\geq n}\right] \underset{n\to+\infty}{\longrightarrow} 0\right\}.$$

We already know from Chapter 3 that $\mathbb{H}^1_{\text{m}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$ is Borel.

Step 1: We first prove that $\mathbb{H}^1_{\text{exp}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$ is Borel. Following the proof of Proposition 3.1 in Chapter 3, we have:

$$\mathbb{H}^1_{\text{exp}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) := \bigcap_{p\in\mathbb{N}^*} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{R}_+, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|\mathcal{E}(X^\alpha)_\tau\mathbf{1}_{\mathcal{E}(X^\alpha)_\tau|>n}] \leq \frac{1}{p}\right\}.$$

It then suffies to show that for all $n, p \in \mathbb{N}^*$, the set

$$\left\{\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{R}_+, \mathbb{F}) \ : \ \sup_\tau \mathbb{E}^{\mathbb{P}_0}[|\mathcal{E}(X^\alpha)_\tau\mathbf{1}_{\mathcal{E}(X^\alpha)_\tau|>n}] \leq \frac{1}{p}\right\}$$

is borel. . We then show that the function $f : \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \to \mathbb{R}_+$ is borel, where:

$$f : \alpha \mapsto \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi(\mathcal{E}(X^\alpha)_\tau)],$$

where $\phi(x) = |x|\mathbf{1}_{x>n}$.

For that purpose, we first introduce for any $K, l > 0$ the continuous function

$$\phi_{K,l}(x) := \frac{n+l}{l}(x-n)^+\mathbf{1}_{x<n+l} + x\mathbf{1}_{n+l\leq x<K} + K\mathbf{1}_{K\leq x}.$$

We clearely have $\phi_{K,l}(x) \nearrow \phi(x)$ when $K \nearrow \infty$ and $l \searrow 0$.

We then introduce the function:

$$f_{K,l} : \alpha \mapsto \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(\mathcal{E}(X^\alpha)_\tau)].$$

And finally we denote by π_k the projection on the segment $[0,k]$, and the function:

$$f^k_{K,l} : \alpha \mapsto \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha)})_\tau)].$$

Our aim is to show that $f^k_{K,l}$ is continuous. We recall from Lemma 3.2 in [74] and from [82] Theorem IV.32 p176 that

$$\alpha \mapsto \int_0^\cdot \pi_k(\alpha_s)^{1/2} dB_s$$

is continuous for the topology of uniform convergence on compacts in probability (denoted ucp for short).
By definition of the topology of $\mathbb{H}^0(\mathbb{R}_+, \mathbb{F})$, we also have that

$$\alpha \mapsto \int_0^\cdot \pi_k(\alpha_s) ds$$

is continuous for the topology ucp.
Then we have that

$$\alpha \mapsto \int_0^\cdot \pi_k(\alpha_s)^{1/2} dB_s - \frac{1}{2}\int_0^\cdot \pi_k(\alpha_s) ds$$

is continuous for the topology of ucp.
Now consider a sequence (α_m) and α in $\mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$ such that $\alpha_m \to \alpha$. Then by continuity and boundedness of $\phi_{K,l} \circ \exp$, we have that

$$\phi_{K,l}\left(\mathcal{E}(X^{\pi_k(\alpha_m)})_\cdot\right) \to \phi_{K,l}\left(\mathcal{E}(X^{\pi_k(\alpha)})_\cdot\right)$$

for the topology of ucp.
Hence:

$$\begin{aligned}
&\left|\sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha_m)})_\tau)] - \sup_\tau \mathbb{E}^{\mathbb{P}_0}[\phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha)})_\tau)]\right| \\
&\leq \sup_\tau \mathbb{E}^{\mathbb{P}_0}\left[\left|\phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha_m)})_\tau) - \phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha)})_\tau)\right|\right] \\
&\leq \mathbb{E}^{\mathbb{P}_0}\left[\sup_{0\leq t\leq T}\left|\phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha_m)})_t) - \phi_{K,l}(\mathcal{E}(X^{\pi_k(\alpha)})_t)\right|\right] \underset{m\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Then the application $f^k_{K,l}$ is continuous.

Now recall that by dominated convergence theorem, for $\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$,

$$\int_0^\cdot \pi_k(\alpha_s)^{1/2} dB_s \underset{k\to\infty}{\longrightarrow} \int_0^\cdot \alpha_s^{1/2} dB_s \text{ ucp, and } \int_0^\cdot \pi_k(\alpha_s) ds \underset{k\to\infty}{\longrightarrow} \int_0^\cdot \alpha_s ds \text{ ucp.}$$

By the same argument as above, we obtain that for $\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$:

$$f_{K,l}(\alpha) = \lim_{k\to\infty} f^k_{K,l}.$$

And then $f_{K,l}$ is measurable as the limit of continuous functions.

Finally we obtain easily for $\alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$:

$$f(\alpha) = \lim_{K\nearrow\infty,\ l\searrow 0} \nearrow f_{K,l}(\alpha),$$

and so f is borel, which ends the first part of the proof.

Step 2: We finally prove that $\mathbb{H}^1_{\tilde{Z}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F})$ is Borel. As in step 1, we show that the application $f : \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}) \to \mathbb{R}$ is borel, where

$$f : \alpha \mapsto \mathbb{E}^{\mathbb{P}_0}\left[\exp\left(bX^\alpha_T - \frac{1}{2}b^2\langle X^\alpha\rangle_T\right)\right].$$

Similarly as above, for $K > 0$ and $k \in \mathbb{N}$, we define:

$$f_K : \alpha \mapsto \mathbb{E}^{\mathbb{P}_0}\left[\exp\left(bX^\alpha_T - \frac{1}{2}b^2\langle X^\alpha\rangle_T\right) \wedge K\right],$$

$$f^k_K : \alpha \mapsto \mathbb{E}^{\mathbb{P}_0}\left[\exp\left(bX^{\pi_k(\alpha)}_T - \frac{1}{2}b^2\langle X^{\pi_k(\alpha)}\rangle_T\right) \wedge K\right].$$

Following the same scheme as in step 1, we obtain that f^k_K is continuous, $f_K(\cdot) = \lim_{k\to\infty} f^k_K(\cdot)$ is measurable as the limit of continuous functions, and

$$f(\alpha) = \lim_{K\nearrow\infty} \nearrow f_K(\alpha), \text{ for all } \alpha \in \mathbb{H}^1_{\text{loc}}(\mathbb{P}_0, \mathbb{R}_+, \mathbb{F}).$$

We then have that f is measurable, which ends the proof.

□

4.4.3 The second inequality

We are now able to show the:

Lemma 4.4.2. *Let $\lambda \in \Lambda^\mu$, then:*

$$V^\lambda_0(\xi) \geq -\exp\left(-b^2 \ln(S_0) + \eta \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda + \frac{b^2}{\eta}\ln(S_T)\right]\right).$$

Proof. By Proposition 4.4.1 and Theorem 4.4.1, we know that there exists $H \in \mathcal{H}^\lambda$ such that:

$$\sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda + \frac{b^2}{\eta}\ln(S_T)\right] + \int_0^T H_s dB_s \geq \xi^\lambda + \frac{b^2}{\eta}\ln(S_T), \quad \mathcal{M}^{S,b}_{\xi^\lambda} - \text{q.s.}$$

We also have that:

$$\ln(S_T) - \ln(S_0) = B_T - \frac{1}{2}\langle B\rangle_T,$$

so that, denoting $v^\lambda := \eta X_0 - \eta \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathbb{E}^{\mathbb{Q}}\left[\xi^\lambda + \frac{b^2}{\eta}\ln\left(\frac{S_T}{S_0}\right)\right]$, for any $\mathbb{Q} \in \mathcal{M}^{S,b}_{\xi^\lambda}$, we have:

$$\begin{aligned}
-\exp\left(-v^\lambda\right) &\leq \mathbb{E}^{\mathbb{Q}}\left[-\exp\left(-\eta\left(X^H_T - \xi^\lambda\right) + b^2\ln\left(\frac{S_T}{S_0}\right)\right)\right] \\
&\leq \mathbb{E}^{\mathbb{Q}}\left[-\exp\left(-\eta\left(X^H_T - \xi^\lambda\right) + (b^2-b)B_T\right)\exp\left(bB_T - \frac{b^2}{2}\langle B\rangle_T\right)\right] \\
&\leq \mathbb{E}^{\mathbb{P}}\left[-\exp\left(-\eta\left(X^H_T - \xi^\lambda\right) + (b^2-b)B_T\right)\right].
\end{aligned}$$

We then define $\tilde{H} := H - \frac{b^2-b}{\eta}$. Since B is a $\mathbb{Q}$-martingale and X^H a $\mathbb{Q}$-supermartingale for every $\mathbb{Q} \in \mathcal{M}^{S,b}_{\xi^\lambda}$, we have that $X^{\tilde{H}}$ is a $\mathbb{Q}$-martingale for all $\mathbb{Q} \in \mathcal{M}^{S,b}_{\xi^\lambda}$.
Finally we have:

$$-\exp\left(v^\lambda\right) \leq \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^\lambda}} \mathbb{E}^{\mathbb{P}}\left[-\exp\left(-\eta\left(X^{\tilde{H}}_T - \xi^\lambda\right)\right)\right],$$

which ends the proof.

□

4.5 The stochastic sharpe ratio case

In this section, we extend this utility maximization study to the case of a stochastic sharpe ratio. This includes the particular case of b deterministic. Our aim is to explain how the result of Corollary 4.2.1 fails to be true in a general framework, and discuss how an additionnal term (entropy) appears in the case of exponential utility.

Namely, we consider now that $b(t, \omega_t)$ is a continuous bounded map. In all the following we state formally the main results of the utility maximization value and of the robust utility indifference price, and we are confident that the solving method developped in sections 4.3 and 4.4 is still valid.

4.5.1 Probability framework

For any $\sigma \in \mathbb{H}^2_{\text{loc}}(\mathbb{P}_0, \mathbb{E}_+, \mathbb{F})$, we define the probability measure on $(\Omega, \mathcal{F})$:

$$\mathbb{P}^{\sigma,b} := \mathbb{P}_0 \circ (X^{\sigma,b})^{-1} \text{ where } X^{\sigma,b}_t := \int_0^t \sigma_s dB_s + \int_0^t b_s\sigma^2_s ds,\ t\in[0,T],\ \mathbb{P}_0 - \text{a.s.}$$

We denote by $\mathcal{P}^b_S$ the collection of all such probability measure on $(\Omega, \mathcal{F})$.

Remark 4.5.1. *We may see that we consider here again a strong formulation for the canonical process. One could also use a weak formulation as introduced in [78], which will lead to additionnal difficulties, especially in the existence and uniqueness of those probability laws.*

As in section 4.2.4, we obtain that the processes Z and $Z\mathcal{E}(B)$ are local $\mathbb{P}$-local martingales for all $\mathbb{P} \in \mathcal{P}^b_S$, where $Z = \mathcal{E}(b\cdot B)^{-1}$.
The admissible set of probability $\mathcal{P}^b_m$ consists of all $\mathbb{P} \in \mathcal{P}^b_S$ such that $\mathbb{E}^{\mathbb{P}}[Z_T] = 1$ and $M := Z\mathcal{E}(B)$ is a uniformly integrable $\mathbb{P}$-martingale.

Remark 4.5.2. *One can see that this definition of the admissible set of probability laws differs from the one adopted in section 4.2. This comes from Remark 4.3.2 following Lemma 4.3.2 which stands that we cannot hope to find a semi-static perfect hedge for the entropy function (defined as* $\ln(Z_T)$*) for a general process b. As a consequence of that, we drop the condition of* ZB *uniformly integrable* $\mathbb{P}$*-martingale for all* $\mathbb{P} \in \mathcal{P}_m^b$*.*

The equivalent martingale measure set $\mathcal{M}^{S,b}$ is then similarly consisting of all measures $\mathbb{Q} := Z_T \cdot \mathbb{P}$ for some $\mathbb{P} \in \mathcal{P}_m^b$, and we have:

Proposition-Definition 1. $\mathcal{M}^{S,b}$ *is the set of all* $\mathbb{Q} \in \mathcal{M}_{loc}^S$ *such that:*
(i) $\mathcal{E}(B)$ *is a uniformly integrable* $\mathbb{Q}$*-martingale,*
(ii) $\mathbb{E}^{\mathbb{Q}}[Z_T^{-1}] = 1$.

4.5.2 Main result

Focusing first on the utility maximisation problem, we define the set of admissible strategies for the semi static process $X^{H,\lambda}$.

The set of static strategies Λ^μ is the subset of $\mathbb{L}^1(\mu)$:

$$\Lambda^\mu := \Big\{ \lambda \in \mathbb{L}^1(\mu) \ : \ \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}} \mathbb{E}^{\mathbb{Q}} \left[\left(-\frac{1}{\eta} \ln(Z_T) + \xi^\lambda \right)^+ \right] < \infty \Big\}.$$

For all $\lambda \in \Lambda^\mu$, we define the set of admissible dynamic strategies by $\mathcal{H}^\lambda$ where:

$$\mathcal{H}^\lambda := \bigcap_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\xi^\lambda}} \mathcal{H}^\lambda(\mathbb{Q}), \quad \text{where:}$$

$$\mathcal{H}^\lambda(\mathbb{Q}) := \big\{ H \in \mathbb{H}^0(\mathbb{R}, \mathbb{F}^*_+) \cap \mathbb{H}^2_{\text{loc}}(\mathbb{Q}, \mathbb{R}, \mathbb{F}^*_+), X^H \text{ is a } \mathbb{Q}-\text{supermartingale} \big\}.$$

The robust utility maximisation problem is then given by:

$$V^b(\xi, X_0) = \sup_{(H,\lambda)\in\mathcal{H}^\lambda\times\Lambda^\mu} \inf_{\mathbb{P}\in\mathcal{P}^b_{\xi^\lambda}} \mathbb{E}\left[U(X_T^H - \xi + \lambda(S_T) - \mu(\lambda)) \right], \tag{4.5.1}$$

and we have the:

Theorem 4.5.1. *Let* $U = -e^{-\eta\cdot}$*. Then for all* $\lambda \in \Lambda^\mu$*,* $V^{\lambda,b}(\xi, X_0) = e^{-\eta X_0} V_0^{\lambda,b}(\xi)$ *with:*

$$V_0^{\lambda,b}(\xi) = -\exp\left(\eta h^{\tilde{\lambda}} \left(\xi - \frac{1}{\eta} \ln(Z_T) \right) \right),$$

where for any $f \in \Lambda^\mu$ *and* ζ *upper-semianalytic,* $h^f(\zeta)$ *is the robust super-replication price of the payoff* ζ^f*:*

$$h^f(\zeta) := \sup_{\mathbb{Q}\in\mathcal{M}^{S,b}_{\zeta^f}} \mathbb{E}^{\mathbb{Q}} \left[\zeta^f \right], \tag{4.5.2}$$

Introducing $V_0^b(\xi) := V^b(0, \xi)$, we clearely have $V^b(\xi, X_0) = e^{-\eta X_0} V_0^b(\xi)$. We then deduce immediatly from Theorem 4.5.1:

Theorem 4.5.2. *Let $U = -e^{-\eta \cdot}$. Then the value function V_0^b is:*

$$V_0^b(\xi) = -\exp\left(\eta h\left(\xi - \frac{1}{\eta}\ln(Z_T)\right)\right),$$

where h is the robust super-replication price given marginals and is equal to:

$$h(\xi - \frac{1}{\eta}\ln(Z_T)) = \inf_{\lambda\in\Lambda^\mu} h^\lambda(\xi - \frac{1}{\eta}\ln(Z_T)).$$

The corresponding utility indifference price is then the direct consequence of Theorem 4.5.2:

Corollary 4.5.1. *The robust indifference price $\bar{p}(\xi)$ of a payoff ξ is independant of X_0 and is given by:*

$$\bar{p}(\xi) = h(\xi - \frac{1}{\eta}\ln(Z_T)) - h(-\frac{1}{\eta}\ln(Z_T)).$$

We may insist here that when considering a stochastic sharpe ratio, the result of Theorem 4.2.1 and Corollary 4.2.1 does not hold in this more general context. When b is constant, we saw that the entropy terminal value can be expressed as a pathwise stochastic integral corrected by a vanilla option (log option), leading to simpler formulas than those of Theorem 4.5.2 and Corollary 4.5.1. In this general framework, we fail to obtain simpler formulations, even for the exponential utility maximization. Indeed it is trivial to see that h is sub-linear in the sens that for ξ_1 and ξ_2 upper semi-analytic and admissible, we have:

$$h(\xi_1 + \xi_2) \leq h(\xi_1) + h(\xi_2).$$

The robust indifference price is then express as the difference of the robust super-replication price of ξ corrected by the entropy process Z, and the robust super-replication price of the entropy process Z. This shows in particular that we cannot hope to obtain that the robust utility indifference price is given by the super-replication price, but only one inequality, which is:

$$\bar{p}(\xi) \leq h(\xi).$$

Part II

Some utility maximization problems

Chapter 5

Liquidation of an indivisible asset with independent investment

Contents

5.1 Introduction

This chapter considers a mixed optimal stopping optimal control problem introduced by Hobson and Henderson [50]. The framework of [50] is the following. An investor holds an indivisible asset, with price process defined as a geometric Brownian motion. In addition, a nonrisky asset, normalized to unity, and a financial asset are available for frictionless continuous-time trading. The risky asset price process is a local martingale with zero covariation with the indivisible asset process. The investor's preferences are defined by the expected power utility function. The objective of the risk averse investor is to choose optimally a stopping time for selling the indivisible asset, while optimally continuously trading on the financial market.

In the absence of the indivisible asset, the problem reduces to a pure portfolio investment problem. Since the risky asset price process is a local martingale, it follows from the Jensen inequality that the optimal investment strategy of the risk averse investor consists in not trading the risky asset. One could also consider the martingale assumption as a renormalization to zero of the optimal investment rule. Therefore, the main question raised by [50] is whether this optimal strategy is affected by the optimal liquidation problem of the independent indivisible asset. In the context of the power utility function, [50] shows that the answer to this question depends on the model parameters, and they provide the optimal stopping-investment strategies.

Our objective is to extend the results of [50] in two directions. First, the indivisible asset price process is defined by an arbitrary scalar homogeneous stochastic differential equation. Second, the investor's preferences are characterized by a general expected utility function. In contrast with [50], we use the standard dynamic programming approach to stochastic control and optimal stopping to show that a lower bound is given by the limit of a sequence of functions defined by successive concavifications with respect to each variable. The resulting function is then the smallest majorant of the utility function which is partially concave in each of the variables. This construction of the lower bound induces a maximizing sequence of stopping times and portfolio strategies. This observation allows to prove that this lower bound indeed coincides with the value function. Finally, we prove that this maximizing sequence is weakly compact, and we deduce the existence of an optimal strategy.

The chapter is organized as follows. The problem is formulated in Section 5.2. The main results are stated in Section 5.3. In particular, in Subsection 5.3.2, we specialize the discussion to the original context of [50], and we show that our general results cover their findings. The explicit derivation of the value function is reported in Section 5.4. Finally, Section 5.5 contains the proof of existence of an optimal stopping-investment strategy.

5.2 Problem formulation

Let B be a Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Throughout this paper, we consider an indivisible asset with price process Y^y defined by the stochastic differential equation:

$$dY_t^y = Y_t^y\big[\mu(Y_t^y)dt + \sigma(Y_t^y)dB_t\big], \;\; Y_0^y = y > 0$$

where the coefficients $\mu,\ \sigma\ : \mathbb{R}_*^+ \longrightarrow \mathbb{R}$ are bounded, locally Lipschitz-continuous, and $\sigma > 0$. In particular, this ensures the existence and uniqueness of a strong solution to the previous SDE.

The first objective of the investor is to decide about a optimal stopping time τ for the liquidation of the indivisable asset. We shall denote by $\mathcal{T}$ the collection of all finite $\mathbb{F}-$stopping times.

The financial market also allows for the continuous frictionless trading of a risky security whose price process is a local martingale orthogonal to W. Then assuming a zero interest rate (or, in other words, considering forward prices), the return from a self-financing portfolio strategy is a process X in the set

$$\mathcal{M}^{\perp}(x) := \{X \text{ càdlàg martingale with } X_0 = x, \text{ and } [X, B] = 0\}, \tag{5.2.1}$$

where $[X, B]$ denotes the quadratic covariation process between X and B. In the last admissibility set, the condition $[X, B] = 0$ reflects that the indivisible asset cannot be hedged dynamically by the financial assets, while the martingale condition implies that, in the absence of the indivisible asset, the optimal investment in risky security of a risk-averse agent is zero. Following Hendersen and Hobson [50], our objective is precisely to analyze the impact of the presence of the indivisible asset on this optimal no-trading strategy.

Given a nondecreasing concave function $U : \mathbb{R}^+ \longrightarrow \mathbb{R}\cup\{-\infty\}$ representing the utility function of a risk-averse investor, we consider the problem:

$$V(x, y) \;\; := \sup_{(X,\tau)\in\mathcal{S}(x,y)} \mathbb{E}\big[U(X_\tau + Y_\tau)\big], \;\; (x, y) \in D, \tag{5.2.2}$$

where $D := \{\mathbb{R} \times \mathbb{R}_*^+;\ x + y \geq 0\}$,

$$\mathcal{S}(x, y) := \big\{(X, \tau) \in \mathcal{M}^{\perp}(x) \times \mathcal{T} : (X + Y^y)_{.\wedge\tau} \geq 0 \text{ and } \{U(X_{\tau\wedge n} + Y_{\tau\wedge n})^-\}_{n\geq 0} \text{ is UI}\big\},$$

and UI is an abreviation for uniformly integrable.

We also introduce the corresponding no-trade problem:

$$m(x,y) \quad := \quad \sup_{\tau\in\mathcal{T}(x,y)} \mathbb{E}\big[U(x+Y^y_\tau)\big], \quad (x,y)\in D, \tag{5.2.3}$$

where $\mathcal{T}(x,y) := \{\tau \in \mathcal{T} : (x,\tau) \in \mathcal{S}(x,y)\}$ and we denote by x the process constantly equal to 0.

5.3 Main results

5.3.1 General utility function

We first introduce a suitable change of variable, transforming the process Y^y into a local martingale. This is classically obtained by means of the scale function S of Y^y defined as a solution of:

$$S'(y)y\mu(y) + \frac{1}{2}y^2\sigma^2(y)S''(y) = 0.$$

By additionally requiring that $S'(c) = 1$ and $S(c) = 0$, for some c in the domain of the diffusion Y, this ordinary differential equation induces a uniquely defined continuous one-to-one function $S : (0,\infty) \longrightarrow dom(S) = \big(S(0), S(\infty)\big)$. We denote $R := S^{-1}$ its continuous inverse. Then the process $Z := S(Y)$ is a local martingale satisfying the stochastic differential equation:

$$dZ_t = \tilde{\sigma}(Z_t)dB_t, \quad \text{with} \quad \tilde{\sigma}(z) = R(z)S'(R(z))\sigma(R(z)).$$

From now on, we will work with the process Z instead of Y. We define the corresponding domain

$$\bar{D} \quad := \quad \{(x,z) \in \mathbb{R}\times dom(S) \;:\; x + R(z) \geq 0\},$$

and we introduce the functions:

$$\bar{m}(x,z) := m(x,R(z)), \quad \bar{V}(x,z) := V(x,R(z)) \quad \text{and} \quad \bar{U}(x,z) := U(x+R(z)), \quad (x,z)\in\bar{D}.$$

Notice that $\bar{U}$ is in general not concave w.r.t. z but still concave w.r.t. x. We then introduce

$$\bar{U}_1 := (\bar{U})^{\mathrm{conc}_z},$$

where conc_z denotes the concave envelope w.r.t. z.

Proposition 5.3.1. *Assume that $\bar{U}^1$ is locally bounded, then $m(x,y) = \bar{U}^1(x,S(y))$ for all $(x,y)\in\bar{D}$.*

Proof. We organize the proof in two steps.
Step 1: We first show that $\bar{m} \leq \bar{U}^1$ for any $\delta > 0$. We fix $(x,z)\in\bar{D}$. For $\tau \in \mathcal{T}(x,R(z))$, and θ_n a localising sequence for Z, we define $\tau_n = \tau\wedge\theta_n$. We then have by Jensen's inequality:

$$E\left[\bar{U}(x,Z_{\tau_n})\right] \leq \mathbb{E}\left[\bar{U}^1(x,Z_{\tau_n})\right] \leq \bar{U}^1(x,\mathbb{E}[Z_{\tau_n}]) = \bar{U}^1(x,z).$$

Now we have by Fatou's Lemma that:

$$\liminf_{n\to\infty}\mathbb{E}\left[\bar{U}(x,Z_{\tau_n})^+\right] \geq \mathbb{E}\left[\liminf_{n\to\infty}\bar{U}(x,Z_{\tau_n})^+\right] = \mathbb{E}\left[\bar{U}(x,Z_\tau)^+\right].$$

By the uniform integrability of the family $\{U(x+Y_{\tau\wedge n})^-, n \geq 0\}$, we obtain:

$$\lim_{n\to\infty} \mathbb{E}\left[\bar{U}(x, Z_{\tau_n})^-\right] = \mathbb{E}\left[\bar{U}(x, Z_\tau)^-\right].$$

Then, $\mathbb{E}\left[\bar{U}(x, Z_\tau)\right] \leq \bar{U}^1(x,z)$, and it follows from the arbitrariness of $\tau \in \mathcal{T}(x, R(z))$ that $\bar{m} \leq \bar{U}^1$.

Step 2: For the second inequality we use the PDE characterization of the problem. Let $\bar{m}_*(x,z) := \liminf_{z'\to z,\ (x,z')\in\bar{D}} \bar{m}(x,z')$ be the lower semicontinuous envelop of the function $x \longmapsto m(x,z)$. From Step 1, we have $\bar{U} \leq \bar{m} \leq \bar{U}^1$. Then, by the assumption that $\bar{U}^1$ is locally bounded, it follows that $\bar{m}_*$ is finite. By classical tools of stochastic control, we have that $\bar{m}_*(x,\cdot)$ is a viscosity super-solution of:

$$\min\{u - \bar{U}(x,\cdot), -u_{zz}\} \geq 0,$$

Then $\bar{m}_*(x,z) \geq \bar{U}^1(x,z)$ for all $(x,z) \in \bar{D}$. Combining with Step 1, we have thus proved that $\bar{m} \leq \bar{U}^1 \leq \bar{m}_* \leq \bar{m}$.

□

We next return to our problem of interest V. Notice that $\bar{U}^1$ is in general not concave in x, see the power utility example in Subsection 5.3.2. We remark also that the calculations performed in this context show that $\bar{U}^n$ is not even continuous, in general, as illustrated by the case $1 < \gamma \leq p$ of Proposition 5.3.2 in which we have $\bar{U}^1$ locally bounded but discontinuous in the x variable (discontinuity at $x = 0$).

Since the risky asset price process is a local martingale, the value function is expected to be concave in x, because of the maximization over the trading strategies in the risky asset. We are then naturally lead to introduce a function $\bar{U}^2 := \left(\bar{U}^1\right)^{\mathrm{conc}_x}$ as a further concavification of $\bar{U}^1$ with respect to the $x-$variable, which may again loose the concavity with respect to the $z-$variable. This leads naturally to the following sequence $\left(\bar{U}^n\right)_n$:

$$\bar{U}^0 = \bar{U}, \quad \bar{U}^{2n+1} = \left(\bar{U}^{2n}\right)^{\mathrm{conc}_z}, \quad \bar{U}^{2n+2} = \left(\bar{U}^{2n+1}\right)^{\mathrm{conc}_x}, \quad n \geq 0.$$

The sequence $\left(\bar{U}^n\right)_n$ is clearly non decreasing, and then converges pointwise to a limit $\bar{U}^\infty$ taking values in $\mathbb{R} \cup \{+\infty\}$. It is then easy to check that $\bar{U}^\infty$ is the smallest dominant of $\bar{U}$ which is partially concave in x, and partially concave in z.

The first main result of the paper is the following:

Theorem 5.3.1. *Assume that the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is sufficiently rich in the following sence:*

(H1) Either, there is a Brownian motion W independent of B,

(H2) Or, there is a sequence $(\xi_n)_{n\geq 0}$ of independent uniformly distributed random variables which may be added to enrich the initial filtration.

Then, $V(x,y) = \bar{U}^\infty(x, S(y))$ for all $(x,y) \in D$. In particular, $V = m$ iff $\bar{U}^\infty = \bar{U}^1$. Moreover if $\bar{U}^\infty$ is locally bounded, then it is continuous. If $\bar{U}^\infty$ is not locally bounded, then $\bar{U}^\infty = +\infty$ on the domain.

We next focus on the existence and the characterization of a solution to the problem V. We need to introduce the following assumption:

Assumption 5.3.1. *For all $(x,z) \in int(\bar{D})$, there exists an open subset O of $\bar{D}$ such that $(x,z) \in O$, $\bar{U} = \bar{U}^\infty$ on ∂O and $\bar{O}$ is a compact subset of $int(\bar{D})$.*

Since $\bar{U} \le \bar{U}^n \le \bar{U}^\infty$ for all $n \ge 0$, this assumption implies that:

$$\bar{U}^n = \bar{U} \text{ on } \partial O \quad \text{for all } n \ge 0.$$

Remark 5.3.1. *Assumption 5.3.1 implies that $\bar{U}^\infty$ is locally bounded. To see this, we first observe that $\bar{U}^\infty$ is nondecreasing in x. Indeed, for any $k \ge 0$, assume that $\bar{U}^{2k}$ is nondecreasing in x, then for any $h \ge 0$, we have on $(S(-x), S(+\infty)))$, $\bar{U}^{2k}(x,\cdot) \le \bar{U}^{2k}(x+h,\cdot)$, and therefore $\left(\bar{U}^{2k}(x,\cdot)\right)^{conc_z} \le \left(\bar{U}^{2k}(x+h,\cdot)\right)^{conc_z}$. Since S is nondecreasing, we obtain that the concave envelope of $\bar{U}^{2k}(x+h,\cdot)$ restricted to the domain $(S(-x), S(+\infty)))$ is smaller than the concave envelope of $\bar{U}^{2k}(x+h,\cdot)$ on $(S(-x-h), S(+\infty)))$. So we have that $\bar{U}^{2k+1}$ is nondecreasing w.r.t. x. Then $\left(\bar{U}^{2k+1}\right)^{conc_x}$ is non decreasing w.r.t. x. This monotonicity property is then inherited by the limit $\bar{U}^\infty$. By the same argument, we see that $\bar{U}^\infty$ is nondecreasing in the z variable.*

We now show that $\bar{U}^\infty$ is locally bounded. For any $(x,z) \in int(\bar{D})$, there exists $r > 0$ such that the square of side r centered in (x,z) (denoted $C((x,z),r)$) is in $int(\bar{D})$. By Assumption 5.3.1, there exists $z^* \ge z + r/2$ with $(x+r/2, z^*) \in int(\bar{D})$ such that $\bar{U}(x+r/2,z^*) = \bar{U}^\infty(x+r/2,z^*)$. Then for any $(\tilde{x},\tilde{z}) \in C((x,z),r)$, we have $\bar{U}^\infty(\tilde{x},\tilde{z}) \le \bar{U}^\infty(\tilde{x}, z+r/2) \le \bar{U}^\infty(x+r/2, z+r/2) \le \bar{U}^\infty(x+r/2,z^*) < \infty$.
Similarly, we also have: $\bar{U}^\infty(\tilde{x},\tilde{z}) \ge \bar{U}^\infty(x-r/2, z-r/2) \le \bar{U}(x-r/2, z-r/2) > -\infty$, and then the result.

Theorem 5.3.2. *Let Assumption 5.3.1 hold true, and assume that the filtered probability space satisfies Condition (H2) of Theorem 5.3.1. Then for all $(x,y) \in D$:*

$$V(x,y) = \mathbb{E}[U(X^*_{\tau^*} + Y_{\tau^*})] \quad \textit{for some} \quad (X^*,\tau^*) \in \mathcal{S}(x,y).$$

The optimal strategy (X^*,τ^*) will be characterized as the limit of an explicit sequence. Moreover if $\bar{U}^\infty = \bar{U}^n$ for some n, then (X^*,τ^*) is derived explicitely in Section 5.5.

5.3.2 The power utility case

In [50], the indivisible asset Y^y is defined as a geometric Brownian motion:

$$dY^y_t = Y^y_t(\mu dt + \sigma dB_t), \quad Y^y_0 = y > 0$$

and the agent preferences are characterized by a power utility function with parameter $p \in (0,\infty)$:

$$U_p(x) = \frac{x^{1-p}-1}{1-p}, \quad p \neq 1, \quad \text{and} \quad U_1(x) = \ln(x).$$

Following [50], we introduce the constants γ and $\hat{\gamma}_p$ defined by:

$$\gamma = \frac{2\mu}{\sigma^2} \quad \text{and} \quad \hat{\gamma}_p \in (0, p\wedge 1),\ (p-\hat{\gamma}_p)^p(p+1-\hat{\gamma}_p) - (2p-\hat{\gamma}_p)^p(1-\hat{\gamma}_p) = 0,$$

where the existence and uniqueness of $\hat{\gamma}_p$ follows from direct calculation.

Proposition 5.3.2. *Let* $U = U_p$ *as defined in* (5.4.1). *Then:*
(i) *for* $\gamma \leq 0$*, we have* $\bar{U}^\infty = \bar{U}^0 < \infty$*,*
(ii) *for* $0 < \gamma \leq \hat{\gamma}_p$*, we have* $\bar{U}^\infty = \bar{U}^1 < \infty$*,*
(iii) *for* $\hat{\gamma}_p < \gamma < 1 \wedge p$*, we have* $\bar{U}^\infty = \bar{U}^2 < \infty$ *and* $\bar{U}^1 \neq \bar{U}^2$*,*
(iv) *for* $\gamma \geq p \wedge 1$*,*
(iv-a) $p \leq 1$*, we have* $\bar{U}^\infty = \bar{U}^2 = +\infty$*,*
(iv-b) $p > 1$*, and* $\gamma \leq p$*, we have* $\bar{U}^\infty = \bar{U}^2 < +\infty$*,*
(iv-c) $p > 1$*, and* $\gamma > p$*, we have* $\bar{U}^\infty = \bar{U}^1 < +\infty$*.*

Corollary 5.3.1. *Let* $U = U_p$ *as defined in* (5.4.1). *Then*
(i) $V = m$ *if and only if* $\gamma \leq \hat{\gamma}_p$ *or* $\gamma > p > 1$*,*
(ii) *for* $\gamma < p \wedge 1$*, Assumption 5.3.1 holds true, so that an optimal hedging-stopping strategy exists.*

Remark 5.3.2. *In the present power utility example, Proposition 5.3.2 states in particular that* $\bar{U}^\infty$ *equals either* U^0, U^1*, or* U^2*, whenever* $\bar{U}^\infty < \infty$*. Then, the optimal strategy is directly obtained from Lemma 5.5.2, and there is no need to the limiting argument of Section 5.5.2.*

Remark 5.3.3. *From our explicit calculations, we observe that Assumption 5.3.1 fails in cases (iv-b) and (iv-c) of Proposition 5.3.2. Our explicit calculations in these cases show that* $\bar{U}^\infty$ *is asymptotic to* $\bar{U}$ *near infinity. For this reason, the existence of an optimal strategy is lost.*

The result of Corollary 5.3.1 is in line with the findings of [50], and in fact complements with some missing cases in [50]. Loosely speaking, Corollary 5.3.1 states that when $\gamma \leq \hat{\gamma}_p$ or when $\gamma > p > 1$, the agent is indifferent to do fair investments on the market; the optimal strategy consists in keeping a constant wealth and solving an optimal stopping time problem, i.e. m. Instead, when $\hat{\gamma}_p < \gamma \leq p$, the agent can take advantage of a dynamic management strategy of its portfolio.

Remark 5.3.4. *The methodology used in [50] is the following.*

- They construct a parametric family of stopping rules and admissible martingales by first fixing the portfolio value and waiting until the indivisible asset reaches a certain level, and then fixing the time and optimizing the jump of the portfolio value process.

- For each element of this family, they evaluate the corresponding performance, and optimize over the parameter values.

The rigorous proof follows from a verification argument. Our methodology relies on the standard stochastic control approach which, via a dynamic programming equation, provides a better understanding of V *and justifies the above construction of optimal strategies.*

5.4 Characterizing the value function

In this section, we first prove that $\bar{V} \leq \bar{U}^\infty$. In Subsection 5.4.2, we prove the reverse inequality under Condition (H1) on the probability space. The corresponding result under Condition (H2) will be proved at the end of Subsection 5.5.1.

5.4.1 Upper bound

Lemma 5.4.1. $\bar{U}^\infty$ *is continuous iff it is locally bounded. If* $\bar{U}^\infty$ *is not locally bounded, then* $\bar{U}^\infty = +\infty$ *on the domain.*

Proof. We first study the case of $\bar{U}^\infty$ is locally bounded. Since $\bar{U}^\infty$ is locally bounded, concave w.r.t. x and concave w.r.t. z, we have that $\bar{U}^\infty(x,\cdot)$ and $\bar{U}^\infty(\cdot,z)$ are continuous on their domain, for all x and z.

Now assume on the contrary that there exists $\varepsilon > 0$, $(x,z) \in int(\bar{D})$ and a sequence $(x_n, z_n) \in int(\bar{D})$, $(x_n,z_n) \underset{n\to+\infty}{\longrightarrow} (x,z)$ such that:

$$\forall n \geq 0, \quad |\bar{U}^\infty(x_n,z_n) - \bar{U}^\infty(x,z)| > \varepsilon.$$

Without loss of generality, we assume that:

$$\bar{U}^\infty(x_n,z_n) > \bar{U}^\infty(x,z) + \varepsilon.$$

By continuity of $\bar{U}^\infty(\cdot,z)$, we have for n large enough:

$$\bar{U}^\infty(x_n,z_n) - \bar{U}^\infty(x_n,z) > \frac{\varepsilon}{2}.$$

Without loss of generality, we assume that $\forall n \geq 0$, $z_n \geq z$. We then define $\tilde{z}^n = z - \sqrt{z_n - z}$. Then by convexity of $\bar{U}^\infty(\cdot,z)$, we have:

$$\frac{\bar{U}^\infty(x_n,z) - \bar{U}^\infty(x_n,\tilde{z}_n)}{z - \tilde{z}_n} \geq \frac{\bar{U}^\infty(x_n,z_n) - \bar{U}^\infty(x_n,z)}{z_n - z} > \frac{\varepsilon}{2}\frac{1}{z_n - z}.$$

Then:

$$\bar{U}^\infty(x_n,z) - \bar{U}^\infty(x_n,\tilde{z}_n) > \frac{\varepsilon}{2}\frac{1}{\sqrt{z_n - z}}.$$

Since $(x_n,\tilde{z}_n) \underset{n\to+\infty}{\longrightarrow} (x,z)$, this is a contradiction with the local boundedness of $\bar{U}^\infty$.

Now for the case $\bar{U}^\infty$ not locally bounded, then we have $(x,z) \in int(\bar{D})$ and $(x_n,z_n) \to (x,z)$ such that $\bar{U}(x_n,z_n) \to +\infty$. We then have $c > 0$ such that $(x+c, z+c) \in int(\bar{D})$. Then $\bar{U}^\infty(x+c,z+c) = +\infty$. Indeed, since for every $\tilde{x}$ and $\tilde{z}$, $\bar{U}^\infty(\tilde{x},\cdot)$ and $\bar{U}^\infty(\cdot,\tilde{z})$ are non decreasing on their domain, for n large enough, we have:

$$\bar{U}^\infty(x_n,z_n) \leq \bar{U}^\infty(x_n, z+c) \leq \bar{U}^\infty(x+c, z+c).$$

And then taking the limit, we have $\bar{U}^\infty(x+c,z+c) = +\infty$. Now since $\bar{U}^\infty$ is partially concave w.r.t. x and w.r.t. z, we clearly have $\bar{U}^\infty = +\infty$ on the domain.

□

We now focus on the first inequality in Theorem 5.3.1.

Lemma 5.4.2. *$\bar{V} \leq \bar{U}^\infty$ on $\bar{D}$.*

In order to prove Lemma 5.4.2, we use a regularization argument in the case $\bar{U}^\infty$ locally bounded. By Lemma 5.4.1, $\bar{U}^\infty$ is continuous on the interior of $\bar{D}$. But in general, it is not twice differentiable in each variable. Therefore, we introduce for any $\varepsilon \in (0,1]$:

$$\bar{U}^n_\varepsilon(x,z) = \int_{\bar{D}} \bar{U}^n(\xi,\zeta)\rho_\varepsilon(x-\xi, z-\zeta)d\xi d\zeta, \ (x,z) \in \bar{D}, \text{ for all } n \in [0,\infty], \tag{5.4.1}$$

where for all u in $\mathbb{R}^2$:

$$\rho_\varepsilon(u) = \varepsilon^{-2}\rho(u/\varepsilon) \quad \text{with} \quad \rho(u) = Ce^{-1/(1-|u|^2)}\mathbf{1}_{|u|<1},$$

and C is chosen such that $\int_{\mathbb{R}^2}\rho(u)du = \int_{B(0,1)}\rho(u)du = 1$. Clearly, ρ_ε is C^∞, compactly supported, and ρ_ε converges pointwise to the Dirac mass at zero. We also intoduce for any $\delta > 0$:

$$\bar{U}^n_{\varepsilon,\delta}(x,z) := \bar{U}^n_\varepsilon(x+2\delta, z), \ (x+2\delta, z) \in \bar{D}, \text{ for all } n \in [0,\infty].$$

Lemma 5.4.3. *$\bar{U}_\varepsilon^\infty \underset{\varepsilon\to 0}{\longrightarrow} \bar{U}^\infty$ pointwise on $\bar{D}$, $\bar{U}_\varepsilon^\infty \in C^\infty(\bar{D})$, $\bar{U}_\varepsilon^\infty \geq \bar{U}_\varepsilon$ on $\bar{D}$, and for ε small enough, $\bar{U}_{\varepsilon,\delta}^\infty$ is concave in each variable.*

Proof. The first three claims follow from classical properties of the convolution product together with the non-negativity of ρ_ε and the construction of $\bar{U}^\infty$.
Let us prove the concavity of $\bar{U}_{\varepsilon,\delta}^\infty$ w.r.t. x. The same proof holds for z. For any $\varepsilon < \delta$, we fix x, x' and z such that $(x,z) \in \bar{D}$ and $(x',z) \in \bar{D}$. For $\lambda \in [0,1]$, denote $\hat{x} := \lambda x + (1-\lambda)x'$. Then using the concavity of $\bar{U}^\infty$ in x:

$$\begin{aligned}
\bar{U}_{\varepsilon,\delta}^\infty(\hat{x},z) &= \int_{\mathbb{R}^2} \bar{U}^\infty(\lambda(x+2\delta+\xi)+(1-\lambda)(x'+2\delta+\xi), z+\zeta)\rho_\varepsilon(\xi,\zeta)d\xi d\zeta \\
&\geq \int_{\mathbb{R}^2} \left(\lambda\bar{U}^\infty(x+2\delta+\xi, z+\zeta)+(1-\lambda)\bar{U}^\infty(x'+2\delta+\xi, z+\zeta)\right)\rho_\varepsilon(\xi,\zeta)d\xi d\zeta \\
&= \lambda\bar{U}_{\varepsilon,\delta}^\infty(x,z)+(1-\lambda)\bar{U}_{\varepsilon,\delta}^\infty(x',z).
\end{aligned}$$

□

Proof of Lemma 5.4.2 In the case $\bar{U}^\infty$ not locally bounded, then by Lemma 5.4.1, we have $\bar{U}^\infty = +\infty$ and the result is obvious.

Now assume that $\bar{U}^\infty$ is locally bounded. We proceed in two steps.

Step 1. Let $(\theta_n)_n$ be a localizing sequence for the local martingale Z. We fix $\delta > 0$ and we consider $\varepsilon < \delta$. Let $(X,\tau) \in \mathcal{S}(x,R(z))$ and $\tau_n = \tau \wedge \theta_n$. Clearely we have that (X,τ_n) is in $\mathcal{S}(x,R(z))$. Then by Itô's formula for jump processes:

$$\begin{aligned}
&\bar{U}_{\varepsilon,\delta}^\infty(X_{t\wedge\tau}, Z_{t\wedge\tau_n}) - \bar{U}_{\varepsilon,\delta}^\infty(x,z) = \\
&\int_0^{t\wedge\tau_n} \frac{1}{2}\partial_{xx}\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)d[X,X]_u^c + \int_0^{t\wedge\tau_n} \frac{1}{2}\partial_{yy}\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)\tilde{\sigma}^2(Z_u)du \\
&+ \int_0^{t\wedge\tau_n} \partial_z\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)\tilde{\sigma}(Z_u)dB_u + \int_0^{t\wedge\tau_n} \partial_x\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)dX_u \\
&+ \sum_{0<u\leq t\wedge\tau_n} \left(\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u) - \bar{U}_{\varepsilon,\delta}^\infty(X_{u-},Z_u) - \partial_x\bar{U}_{\varepsilon,\delta}^\infty(X_{u-},Z_u)\Delta X_u\right).
\end{aligned} \tag{5.4.2}$$

Since $\bar{U}_{\varepsilon,\delta}^\infty$ is concave in x and in z, then:

$$\bar{U}_{\varepsilon,\delta}^\infty(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n}) - \bar{U}_{\varepsilon,\delta}^\infty(x,z) \leq \int_0^{t\wedge\tau_n} \partial_z\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)\tilde{\sigma}(Z_u)dB_u + \int_0^{t\wedge\tau_n} \partial_x\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)dX_u.$$

We have for all $(\tilde{x},\tilde{z})$:

$$\begin{aligned}
\bar{U}_{\varepsilon,\delta}(\tilde{x},\tilde{z}) &= \int_{\bar{B}((\tilde{x},\tilde{z}),\varepsilon)} \bar{U}(\tilde{x}+2\delta-u, \tilde{z}-v)\rho_\varepsilon(u,v)dudv \\
&\geq \int_{\bar{B}((\tilde{x},\tilde{z}),\varepsilon)} U(\delta)\,\rho_\varepsilon(u,v)dudv = U(\delta),
\end{aligned}$$

where the last inequality follows from the fact that U is non decreasing and $\tilde{x}+2\delta-u+R(\tilde{z}-v) \geq 0$ on $\bar{B}((\tilde{x},z),\varepsilon)$. By Lemma 5.4.3, this implies:

$$\bar{U}_{\varepsilon,\delta}^\infty(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n}) \geq U(\delta).$$

Since , $|U(\delta)| < \infty$, the local martingale:

$$\int_0^{t\wedge\tau_n} \partial_z\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)\tilde{\sigma}(Z_u)dB_u + \int_0^{t\wedge\tau_n} \partial_x\bar{U}_{\varepsilon,\delta}^\infty(X_u,Z_u)dX_u^c, \quad t \geq 0,$$

is bounded from below so it is a supermartingale. Then it follows from (5.4.2) that:

$$\mathbb{E}[\bar{U}^{\infty}_{\varepsilon,\delta}(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n}] \leq \bar{U}^{\infty}_{\varepsilon,\delta}(x,z).$$

Step 2 Since $\bar{U}^{\infty}_{\varepsilon,\delta}(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n})$ is bounded from below by $U(\frac{\delta}{2})$ and $\bar{U}^{\infty}_{\varepsilon,\delta} \underset{\varepsilon\to 0}{\longrightarrow} \bar{U}^{\infty}_{\delta}$ pointwise, we obtain by Fatou's Lemma that:

$$\mathbb{E}[\bar{U}^{\infty}_{\delta}(X_\tau, Z_\tau)] = \mathbb{E}[\lim_{\substack{t,n\to\infty \\ \varepsilon\to 0}} \bar{U}^{\infty}_{\varepsilon,\delta}(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n})] \leq \liminf_{\substack{t,n\to\infty \\ \varepsilon\to 0}} \mathbb{E}[\bar{U}^{\infty}_{\varepsilon,\delta}(X_{t\wedge\tau_n}, Z_{t\wedge\tau_n})] \leq \bar{U}^{\infty}_{\delta}(x,z),$$

and therefore:

$$\bar{V}(x,z) \leq \bar{U}^{\infty}_{\delta}(x,z) \leq \bar{U}^{\infty}(x+2\delta, z).$$

We finally send $\delta \to 0$ and obtain by continuity of $\bar{U}^{\infty}$ in the x-variable:

$$\bar{V}(x,z) \leq \bar{U}^{\infty}(x,z).$$

□

5.4.2 Lower bound for the value function under (H1)

Under Assumption (H1) on the filtration, it follows that $\mathcal{M}^{\perp}$ is non-trivial, and contains the set:

$$\mathcal{M}^W(x,y) := \{X \ C^0\text{-mart}: \ X_t = X_0 + \int_0^t \phi_s dW_s \text{ for some } \phi \in \mathbb{H}^2_{\text{loc}} \text{ and } X + Y \geq 0 \ a.s.\}.$$

In this subsection, we use the PDE characterization of the problem to obtain the lower bound for the value function. In order to use the classical tools of stochastic control and viscosity solutions we introduce the following simplified problem V^0:

$$V^0(x,y) := \sup_{(X,\tau)\in\mathcal{S}^W(x,y)} \mathbb{E}[U(X^{\alpha,t,x}_\tau + Y^{t,y}_\tau)],$$

where $\mathcal{S}^W(x,y) := \{(X,\tau) \in \mathcal{S}(x,y) : X \in \mathcal{M}^W(x,y)\}$.

Since $\mathcal{M}^W(x,y) \subset \mathcal{M}^{\perp}(x,y)$, we have

$$V^0(x,y) \leq V(x,y).$$

The aim of introducing $\mathcal{A}$ is to use the weak dynamic programming principle introduced in [26]. We recall the definition of the lower semi-continuous envelope:

$$V^0_*(x,y) := \liminf_{\substack{y'\to y \\ x'\to x}} V^0(x,y), \quad (x,y) \in D.$$

By Lemma 5.4.2, we have $U(x+y) \leq V(x,y) \leq \bar{U}^{\infty}(x, R(y))$. Since $\bar{U}^{\infty}$ is locally bounded, so is V. Therefore V^0_* is finite.

We now derive the dynamic programming equation, which will provide us with the lower bound:

Proposition 5.4.1. *Assume that $\bar{U}^{\infty}$ is locally bounded, then $\bar{V}^0_*$ is a viscosity supersolution of:*

$$\min\{-v_{zz}, -v_{xx}, v - \bar{U}\} = 0 \quad on \ \bar{D}.$$

In particular $\bar{V}^0_*$ is partially concave w.r.t x and z.

Proof. We first show that V_*^0 is a viscosity supersolution of:

$$\min\{-\frac{1}{2}y^2\sigma(y)^2 v_{yy}(x,y) - y\mu(y)v_y(x,y); -v_{xx}(x,y); v - U(x+y)\} = 0 \tag{5.4.3}$$

on D. Indeed, it is easy to check that the assumptions of Theorem 4.1 in [26] are verified, so that the following weak dynamic programming principle holds:

$$V^0(x,y) \geq \sup_{(X,\tau)\in\mathcal{S}^W(x,y)} \mathbb{E}\left[V_*^0(X_\theta, Y_\theta^y)\mathbf{1}_{\theta\leq\tau} + U(X_\theta + Y_\theta^y)\mathbf{1}_{\theta>\tau}\right] \text{ for all } \theta \text{ stopping time.}$$

Now take $\phi \in \mathcal{C}^{2,2}(\mathbb{R})$ such that $\min(V_*^0 - \phi) = (V_*^0 - \phi)(x_0, y_0)$. After possibly adding a constant to ϕ, we can assume without loss of generality that:

$$\min(V_*^0 - \phi) = (V_*^0 - \phi)(x_0, y_0) = 0.$$

Let $(x_n, y_n)_{n\geq 0}$ be a sequence such that $(x_n, y_n, V^0(x_n, y_n)) \to (x_0, y_0, V_*^0(x_0, y_0))$ as $n \to \infty$. We can see that selling immediately leads to $V_*^0(x,y) \geq U(x+y)$. Indeed by the continuity of U,

$$V_*^0(x,y) = \liminf_{(x',y')\to(x,y)} V^0(x',y') \geq \liminf_{(x',y')\to(x,y)} U(x'+y') = U(x+y)$$

Let us define $\beta_n := V^0(x_n,y_n) - \phi(x_n,y_n)$ and $(X^n, Y^n) = (x_n + \alpha W, Y^{y_n})$, where α is such that $X^n + Y^n \geq 0$, $\mathbb{P}$-a.s. We consider the following stopping time

$$\theta_n := \inf\{t \geq 0 : \ (t, X_t^n - x_n, Y_t^n - y_n) \notin [0, h_n) \times \alpha B$$

where α is a positive given constant, B is the unit ball of $\mathbb{R}^2$ and

$$h_n := \sqrt{|\beta_n|}\mathbf{1}_{\beta_n\neq 0} + \frac{1}{n}\mathbf{1}_{\beta_n=0},$$

where we recall that $\beta_n \longrightarrow 00$ as $n \to \infty$. By the dynamic programming principle together with Itô's formula, it follows that:

$$\begin{aligned} V^0(x_n,y_n) = \beta_n + \phi(x_n,y_n) &\geq \mathbb{E}[\phi(X_{\theta_n}^n, Y_{\theta_n}^n)] \\ &= \phi(x_n,y_n) + \mathbb{E}[\int_0^{\theta_n}\left(y\mu\phi_y + \frac{1}{2}y^2\sigma^2\phi_{yy} + \frac{1}{2}\alpha\phi_{xx}\right)(X_u^n, Y_u^n)\,du]. \end{aligned}$$

This leads to:

$$\beta_n \ \geq \ \mathbb{E}\Big[\int_0^{\theta_n}\left(y\mu\phi_y + \frac{1}{2}y^2\sigma^2\phi_{yy} + \frac{1}{2}\alpha\phi_{xx}\right)(X_u^n, Y_u^n)\,du\Big]$$

Since μ and σ are locally Lipschitz continuous and have linear growth, one can show the following standard estimate for all $h > 0$:

$$\mathbb{E}\left[\sup_{t\leq s\leq t+h}|Y_s^{y_n} - y_n|^2\right] \leq Ch^2(1+|y_n|^2).$$

This leads to $(X^n, Y^n) \underset{n\to\infty}{\longrightarrow} (x_0 + \alpha W, Y^{y_0})$ $\mathbb{P}$-a.s. For n suffsiciently large and all $\omega \in \Omega$, $\theta(\omega) = h_n$. Moreover by definition of θ_n, the following quantity

$$\frac{1}{h_n}\int_0^{\theta_n}\left(y\mu\phi_y + \frac{1}{2}y^2\sigma^2\phi_{yy} + \frac{1}{2}\alpha\phi_{xx}\right)(X_u^n, Y_u^n)\,du$$

is bounded, uniformly in n. Therefore, by the mean value and the dominated convergence theorem,

$$0 \geq \frac{1}{2}y_0^2\sigma^2(y_0)\phi_{yy}(x_0,y_0) + y^0\mu(y_0)\phi_y(x_0,y_0) + \frac{1}{2}\alpha^2\phi_{xx}(x_0,y_0).$$

By the arbitrariness of $\alpha \in \mathbb{R}$, this implies that $-\phi_{xx}(x_0,y_0) \leq 0$. Hence, V_*^0 is a viscosity supersolution on D of:

$$\min\{-\frac{1}{2}y^2\sigma^2(y)v_{yy} - y\mu(y)v_y;\ -v_{xx};\ v(x,y) - U(x+y)\} = 0.$$

Finally, the supersolution stated in the proposition is a direct consequence of the first step and the change of variable in the theory of viscosity solutions, see e.g. [44]. The partial concavity property follows from Lemmas 6.9 and 6.23 in [98].

□

Corollary 5.4.1. *Assume $\bar{U}^\infty$ is locally bounded. Then for all $(x,y) \in D$, we have:*

$$V(x,y) \geq \bar{U}^\infty(x, S(y)).$$

Proof. We already know that $V(x,y) \geq V^0(x,y) \geq \bar{V}_*^0(x,S(y))$. On the other hand, since $\bar{V}_*^0$ is partially concave w.r.t. x and w.r.t. z, and is a majorant of $\bar{U}$, it follows that $\bar{V}_*^0$ is a majorant of $\bar{U}^\infty$. This completes the proof.

□

5.5 Optimal strategy

We now derive an optimal strategy under Assumption 5.3.1 together with Condition (H2) of Theorem 5.3.1. This will allow also to recover the case $\bar{U}^\infty = +\infty$ since the construction is robust, whenever the concave envelopes are not finite.

5.5.1 Construction of a maximizing sequence under (H2)

We fix $(x,z) \in int(\bar{D})$ and we consider O the open set defined in Assumption 5.3.1. We define the following sequence of stopping times $(\tau^n)_{n\geq 0}$:

Since $\bar{U}^1$ is the concavification of $\bar{U}$ with respect to the z-variable, we introduce the stopping time with frozen x-variable:

$$\tau_1^0 = \inf\{t \geq 0 :\ \bar{U}^1(X_0, Z_t) = \bar{U}^0(X_0, Z_t)\},$$

At time τ_1^0, $Z_{\tau_1^0}$ takes values in $\{z_1, z_2\}$ where $z_1 = \sup\{z \leq Z_0 :\ \bar{U}^1(X_0,z) = \bar{U}(X_0,z)\}$ and $z_2 = \inf\{z \geq Z_0 :\ \bar{U}^1(X_0,z) = \bar{U}(X_0,z)\}$. Notice that z_1 and z_2 are finite, taking values in $\bar{O}$. We then define $X_t = X_0$ for $t < \tau_1^0$ and for $t \geq \tau_1^0$:

$$X_t = \eta(X_0, Z_{\tau_1^0}),$$

where $\mathbb{E}\left[\eta(X_0, Z_{\tau_1^0})|\mathcal{F}_{\tau_1^0-}\right] = X_0$ and:

$$\mathbb{P}\{[\eta(X_0, Z_{\tau_1^0}) = a(X_0, Z_{\tau_1^0})|(X_0, Z_{\tau_1^0})\} = p(X_0, Z_{\tau_1^0})$$

$$\mathbb{P}\{[\eta(X_0, Z_{\tau_1^0}) = b(X_0, Z_{\tau_1^0})|(X_0, Z_{\tau_1^0})\} = 1 - p(X_0, Z_{\tau_1^0})$$

with:

$$
\begin{aligned}
&d(v) := \{x \in \mathbb{R} : (x, v) \in \bar{D}\},\\
&a(u, v) := \inf\{\alpha \in d(v),\ \alpha \geq u : \bar{U}^2(\alpha, v) = \bar{U}^1(\alpha, v)\},\\
&b(u, v) := \sup\{\alpha \in d(v),\ \alpha \leq u : \bar{U}^2(\alpha, v) = \bar{U}^1(\alpha, v)\},
\end{aligned}
$$

and $p(u, v)$ such that :

$$u = p(u, v)a(u, v) + (1 - p(u, v))b(u, v).$$

Similarly, we define a sequence of stopping times $(\tau_i^n)_{0\leq i\leq n+1}$ by $\tau_0^n = 0$, and for $i \in \{1, ..., n+1\}$:

$$\tau_i^n = \inf\{t \geq \tau_{i-1}^n :\ \bar{U}^{2(n-i+1)+1}(X^n_{\tau^n_{i-1}}, Z_t) = \bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, Z_t)\},$$

where the martingale X^n is constructed as follows. Let:

$$
\begin{aligned}
&a_i^n(u, v) := \inf\{\alpha \in d(v),\ \alpha \geq u : \bar{U}^{2(n-i+1)}(\alpha, v) = \bar{U}^{2(n-i+1)-1}(\alpha, v)\},\\
&b_i^n(u, v) := \sup\{\alpha \in d(v),\ \alpha \leq u : \bar{U}^{2(n-i+1)}(\alpha, v) = \bar{U}^{2(n-i+1)-1}(\alpha, v)\}.
\end{aligned}
$$

By Assumption 5.3.1, $(a^n(u, v), v)$ and $(b^n(u, v), v)$ are in $\bar{O}$ and $\bar{U}^{2n-i+1}(\cdot, v)$ is linear on $[a_i^n(u, v), b_i^n(u, v)]$. We then define $p_i^n(u, v) \in [0, 1]$ by:

$$u = p_i^n(u, v)a_i^n(u, v) + (1 - p_i^n(u, v))b_i^n(u, v),$$

so that:

$$\bar{U}^{2(n-i+1)}(u, v) = p_i^n(u, v)\bar{U}^{2(n-i+1)-1}(a_i^n(u, v), v) + (1 - p_i^n(u, v))\bar{U}^{2(n-i+1)-1}(b_i^n(u, v), v).$$

With these notations, we define the process X^n:

$$X_t^n = X_0^n \mathbf{1}_{[0,\tau_1^n)}(t) + \sum_{i=1}^{n-1} \eta_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\mathbf{1}_{[\tau_i^n,\tau_{i+1}^n)}(t) + \eta_n^n(X^n_{\tau^n_{n-1}}, Z_{\tau_n^n})\mathbf{1}_{[\tau_n^n,\infty)}(t),$$

where each r.v. $\eta_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})$ is independant of $\mathcal{F}_{\tau_i^n}$ and has distribution:

$$
\begin{aligned}
&\mathbb{P}\left[\eta_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}) = a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})|\mathcal{F}_{\tau_i^n-}\right] = p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}),\\
&\mathbb{P}\left[\eta_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}) = b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})|\mathcal{F}_{\tau_i^n-}\right] = 1 - p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}).
\end{aligned}
$$

The existence of such r.v. $\{\eta_i^n, i \leq n\}_n$ is guaranteed by Assumption (H2).

Remark 5.5.1. *The measurability of p_i^n, a_i^n and b_i^n is not necessary because it is only involved in a finite number of values at each step.*

Lemma 5.5.1. *Under assumption 5.3.1, $(X^n, \tau^n_{n+1}) \in \mathcal{S}(x, y)$ for all $n \geq 1$.*

Proof. That $[X^n, Z] = 0$ follows from the fact that X is a pure jump process and Z is continuous. We also see that (X^n, Z) takes it values only in a compact K given by assumption 5.3.1, so $\tau^n_{n+1} \in \mathcal{T}$ and the process is non negative. We now prove the martingale property. For all $i \in \{1, ..., n\}$:

- $t \in (\tau_i^n, \tau_{i+1}^n) \Rightarrow \mathbb{E}[X_t^n|\mathcal{F}_{t-}] = X_{t-}^n$

- If $t = \tau_i^n$, then:

$$\begin{aligned}\mathbb{E}[X_t^n|\mathcal{F}_{t-}] &= \mathbb{E}[\eta_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})|\mathcal{F}_{t-}] \\ &= a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\mathbb{E}[\mathbf{1}_{\eta_i^n=a_i^n}|\mathcal{F}_{t-}] + b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\mathbb{E}[1-\mathbf{1}_{\eta_i^n=a_i^n}|\mathcal{F}_{t-}] \\ &= X^n_{\tau^n_{i-1}} = X^n_{t-}\end{aligned}$$

□

The crucial property of the sequence $(X^n, \tau^n_{n+1})_n$ is the following.

Lemma 5.5.2. *For all $n \geq 0$, we have:*

$$\mathbb{E}[\bar{U}(X^n_{\tau^n_{n+1}}, Z_{\tau^n_{n+1}})] = \bar{U}^{2n+1}(x,z). \tag{5.5.1}$$

Proof. We organize the proof in three steps.
Step1: We first show that for all $i \in \{1, ..., n+1\}$, we have:

$$\mathbb{E}\left[\bar{U}^{2(n-i+1)-1}\left(X^n_{\tau_i^n}, Z_{\tau_i^n}\right)\right] = \mathbb{E}\left[\bar{U}^{2(n-i+1)}\left(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}\right)\right]. \tag{5.5.2}$$

Indeed:

$$\begin{aligned}\mathbb{E}\left[\bar{U}^{2(n-i+1)-1}\left(X^n_{\tau_i^n}, Z_{\tau_i^n}\right)\right] &= \mathbb{E}[\bar{U}^{2(n-i+1)-1}\left(a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)\mathbb{E}\left[\mathbf{1}_{\eta_i^n=a_i^n}|X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}\right] \\ &+ \bar{U}^{2(n-i+1)-1}\left(b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)\mathbb{E}\left[\mathbf{1}_{\eta_i^n=b_i^n}|X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}\right]] \\ &= \mathbb{E}[\bar{U}^{2(n-i+1)-1}\left(a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}) \\ &+ \bar{U}^{2(n-i+1)-1}\left(b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)(1-p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})).\end{aligned}$$

Then by definition of the random variables $a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})$ and $b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})$, and the linearity of $\bar{U}^{2(n-i+1)}(\cdot, Z_{\tau_i^n})$ on $\left[b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\right]$, we have:

$$\begin{aligned}\mathbb{E}\left[\bar{U}^{2(n-i+1)-1}(X^n_{\tau_i^n}, Z_{\tau_i^n})\right] &= \mathbb{E}\Big[\bar{U}^{2(n-i+1)}\left(a_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}) \\ &+\bar{U}^{2(n-i+1)}\left(b_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}), Z_{\tau_i^n}\right)(1-p_i^n(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n}))\Big] \\ &= \mathbb{E}\left[\bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\right].\end{aligned}$$

Step 2: We next show that:

$$\mathbb{E}\left[\bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\right] = \mathbb{E}\left[\bar{U}^{2(n-i+1)+1}(X^n_{\tau^n_{i-1}}, Z_{\tau^n_{i-1}})\right]. \tag{5.5.3}$$

We emphasize here that the process X^n takes its values in a finite set. Then the fact that $\sigma > 0$ and continuous ensures that $|\tilde{\sigma}| > c > 0$ on $\text{proj}_z(\bar{O})$ and then if follows that for all i, $\tau_i^n < \infty$ and that $\mathbb{E}[X^n_{\tau_i^n}|X^n_{\tau^n_{i-1}}] = X_{\tau^n_{i-1}}$.
Then we know that $\bar{U}^{2(n-i+1)+1}\left(X^n_{\tau^n_{i-1}}, z\right)$ is linear on H_i^n where:

$$H_i^n := \left\{z > 0 \ : \ \bar{U}^{2(n-i+1)+1}(X^n_{\tau^n_{i-1}}, z) > \bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, z)\right\}.$$

We can now conclude, by definition of τ_i^n that:

$$\begin{aligned}\mathbb{E}\left[\bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\right] &= \mathbb{E}\left[\bar{U}^{2(n-i+1)+1}(X^n_{\tau^n_{i-1}}, Z_{\tau_i^n})\right] \\ &= \mathbb{E}\left[\bar{U}^{2(n-i+1)+1}(X^n_{\tau^n_{i-1}}, Z_{\tau^n_{i-1}})\right].\end{aligned}$$

Step 3: we now prove (5.5.1): Using (5.5.2) and (5.5.3) we have:

$$\begin{aligned}\bar{U}^{2n+1}(x,z) =& \sum_{i=1}^{n} \mathbb{E}\left[\bar{U}^{2(n-i+1)}(X^n_{\tau^n_{i-1}}, Z_{\tau^n_i}) - \bar{U}^{2(n-i+1)-1}(X^n_{\tau^n_i}, Z_{\tau^n_i})\right] \\ &+ \sum_{i=0}^{n} \mathbb{E}\left[\bar{U}^{2(n-i+1)-1}(X^n_{\tau^n_{i-1}}, Z_{\tau^n_i}) - \bar{U}^{2(n-i+1)-2}(X^n_{\tau^n_i}, Z_{\tau^n_{i+1}})\right] \\ &+ \mathbb{E}\left[\bar{U}^0(X^n_{\tau^n_n}, Z_{\tau^n_{n+1}})\right] \\ =& \mathbb{E}\left[\bar{U}^0(X^n_{\tau^n_n}, Z_{\tau^n_{n+1}})\right].\end{aligned}$$

By construction, we have $\tau^n_{n+1} \geq \tau^n_n$ so we have $X^n_{\tau^n_{n+1}} = X^n_{\tau^n_n}$ and then:

$$\bar{U}^{2n+1}(x,z) = \mathbb{E}\left[\bar{U}^0(X^n_{\tau^n_{n+1}}, Z_{\tau^n_{n+1}})\right].$$

□

Proof of Theorem 5.3.1 under (H2) By Lemma 5.4.2, $\bar{V} \leq \bar{U}^\infty$. Then, since the sequence $\left(\bar{U}^n\right)_n$ converges towards $\bar{U}^\infty$, it follows immediately from Lemma 5.5.2 that $(X^n, \tau^n_{n+1})_n$ is a maximizing sequence of strategies.

□

Remark 5.5.2. *Notice that Assumption 5.3.1 and the local boundedness condition of $\bar{U}^\infty$ are not necessary to obtain a maximizing sequence. Indeed we have that the concave envelope f^{conc} of a function f defined on an interval $I \subset \mathbb{R}$ is given by: $\sup_{\substack{y_1 \leq y \leq y_2 \\ y_1, y_2 \in I}} (\lambda(y_1,y_2)f(y_1) + (1-\lambda(y_1,y_2))f(y_2))$, with $\lambda(y_1,y_2) = \frac{y_2 - y}{y_2 - y_1}$, with convention $\lambda(y,\cdot) = 1$ and $\lambda(\cdot,y) = 0$. So we could have considered ε-optimal sequences of coefficients a^n_i and b^n_i rather than optimal ones, which may not exist in the general case, and the proof holds. However the present construction is crucial for the result of the subsequent section.*

5.5.2 Existence of an optimal strategy

Proof of Theorem 5.3.2 Let $(X^n_{\tau^n_{n+1}}, Z_{\tau^n_{n+1}})_{n\geq 0}$ be the sequence defined in Lemma 5.5.2. These pairs of random variables take values in the compact subset $\bar{O}$. We then define μ_n the law of $(X^n_{\tau^n_{n+1}}, Z_{\tau^n_{n+1}})$. This is a sequence of probability distributions with support in the compact subset $\bar{O}$. Then (μ_n) is tight, and by the Prokhorov theorem we may find a subsequence, still renamed (μ_n), which converges to some probability distribution μ with support in $\bar{O}$.
Step 1: We first prove that $\int_{\bar{O}} \bar{U}(\xi,\zeta)d\mu(\xi,\zeta) = \bar{U}^\infty(x,z)$.
Indeed, we have that $\bar{U}$ is continuous on $\bar{D}$ and $\bar{O}$ is a compact of $\bar{D}$, So by Lemma 5.5.2 together with the weak convergence property, we obtain:

$$\bar{U}^\infty(x,z) = \lim_{n\to\infty} \bar{U}^n(x,z) = \lim_{n\to\infty} \int_{\bar{O}} \bar{U}(\xi,\zeta)d\mu^n(\xi,\zeta) = \int_{\bar{O}} \bar{U}(\xi,\zeta)d\mu(\xi,\zeta).$$

Step 2: We next introduce a pair (X^*, τ^*) such that $(X^*_{\tau^*}, Z_{\tau^*}) \sim \mu$.

First, we consider τ^* a $(\sigma(B_{0\leq s\leq t}))_{t\geq 0}$-stopping time such that $Z_{\tau^*} \sim \mu_z$, where $\mu_z(A) := \int_{\mathbb{R}\times A} \mu(dx,dz)$ is the z-marginal law of μ. Such a stopping time exists because μ_z is compactly supported and $\tilde{\sigma} \geq c > 0$ on $\bar{O}$ for some $c > 0$, thanks to the assumption that $\sigma > 0$. This result is proved in [56], section 4.3.

We now consider $f : [0,1]^2 \to K$ a Borel function such that the pushforward measure of the lesbegue measure on $[0,1]^2$ by f is μ and $f(x,y) = (f_1(x,y), f_2(y))$. The existence of this function corresponds to the existence of the conditional probability distribution.
We denote F_{μ_z} the cumulative distribution function of μ_z. ζ denotes a uniform random variable independent of B and we implicitly assume that the filtration $\mathbb{F}$ is rich enough to support that ζ is $\mathcal{F}_{\tau^*}$-measurable and independant of $\mathcal{F}_{\tau^*-}$. In particular, ζ is independent of $\sigma(B_{0\leq s\leq \tau^*})$.
The candidate process X^* is then:

$$\forall t \geq 0, \quad X_t^* := f_1(\zeta, F_{\mu_z}(Z_\tau))\mathbf{1}_{t\geq\tau}.$$

Then we clearely have that $(X^*_{\tau^*}, Z_{\tau^*}) \sim \mu$.

Step 3: It remains to prove that X^* is a martingale in $\mathcal{M}^\perp$.
We easily have that $\mathbb{E}[X^*_{\tau^*}] = X_0$. Indeed, as $X^*_{\tau^*}$ takes values in a compact subset, the weak convergence implies that:

$$\mathbb{E}[X^*_{\tau^*}] = \int x\mu(dx,dz) = \lim_{n\to\infty} \int x\mu^n(dx,dz) = X_0$$

It remains to prove that X^* is independent of $\sigma(B_{0\leq s\leq\tau^*})$. By construction of X^*, we see that:

$$\mathbb{E}[X^*_{\tau^*}|\sigma(B_{0\leq s\leq\tau^*})] = \mathbb{E}[X^*_{\tau^*}|Z_{\tau^*}].$$

Then we have to prove that:

$$\mathbb{E}[X^*_{\tau^*}|Z_{\tau^*}] = X_0,$$

i.e. that for all ϕ bounded continuous function, we have:

$$\mathbb{E}[(X^*_{\tau^*} - X_0)\phi(Z_{\tau^*})] = \int_{\bar{O}} (x - X_0)\phi(z)\mu(dx,dz) = 0.$$

By continuity of ϕ, and the fact that μ is compactly supported, we have that:

$$\begin{aligned}\int_{\bar{O}} (x - X_0)\phi(z)\mu(dx,dz) &= \lim_{n\to+\infty} \int_{\bar{O}} (x - X_0)\phi(z)\mu^n(dx,dz) \\ &= \lim_{n\to+\infty} \mathbb{E}[(X^n_{\tau^n_{n+1}} - X_0)\phi(Z_{\tau^n_{n+1}})].\end{aligned}$$

Then:

$$\begin{aligned}\mathbb{E}[(X^n_{\tau^n_{n+1}} - X_0)\phi(Z_{\tau^n_{n+1}})] &= \mathbb{E}\left[\left(\sum_{i=1}^{n+1} X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right)\phi(Z_{\tau^n_{n+1}})\right] \\ &= \sum_{i=1}^{n+1} \mathbb{E}\left[\mathbb{E}_{\tau^n_i}\left[\left(X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right)\phi(Z_{\tau^n_{n+1}})\right]\right] \\ &= \sum_{i=1}^{n+1} \mathbb{E}\left[\left(X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right)\mathbb{E}_{\tau^n_i}\left[\phi(Z_{\tau^n_{n+1}})\right]\right].\end{aligned}$$

By continuity of Z, we have that $\mathbb{E}_{\tau^n_i}\left[\phi(Z_{\tau^n_{n+1}})\right] = \mathbb{E}_{\tau^n_i-}\left[\phi(Z_{\tau^n_{n+1}})\right]$. And then:

$$\begin{aligned}\mathbb{E}\left[\left(X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right)\mathbb{E}_{\tau^n_i}\left[\phi(Z_{\tau^n_{n+1}})\right]\right] &= \mathbb{E}\left[\left(X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right)\mathbb{E}_{\tau^n_i-}\left[\phi(Z_{\tau^n_{n+1}})\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}_{\tau^n_i-}\left[\phi(Z_{\tau^n_{n+1}})\right]\mathbb{E}_{\tau^n_i-}\left[X^n_{\tau^n_i} - X^n_{\tau^n_{i-1}}\right]\right] \\ &= 0,\end{aligned}$$

where we used the fact that $\mathbb{E}_{\tau^n_i-}\left[X^n_{\tau^n_i}\right] = X^n_{\tau^n_{i-1}}$. This concludes the proof

□

5.6 Appendix: power utility function

Our goal is to compute explicitly the function $\bar{U}^\infty$ in the context of the power utility function of Section 5.3.2.Proposition 5.3.2 then follows immediatly from our explicit calculations.

The scale function S_γ of Y is given up to an affine transformation by

$$S_\gamma(y) = \mathrm{sgn}(1-\gamma)y^{1-\gamma} \ \text{ if } \gamma \neq 1 \text{ and } S_1(y) = \ln(y).$$

Then:

$$R_\gamma(z) := (\mathrm{sgn}(1-\gamma)z)^{\frac{1}{1-\gamma}} \ \text{ if } \gamma \neq 1 \text{ for all } \mathrm{sgn}(1-\gamma)z \in \mathbb{R}_+, \text{ and } R_1(z) = e^z \text{ for all } z \in \mathbb{R}$$

and the process Z is a martingale defined by:

$$Z_t = Z_0 e^{|1-\gamma|\sigma B_t - \frac{1}{2}(1-\gamma)^2\sigma^2 t}, \ Z_0 = \mathrm{sgn}(1-\gamma)Y_0^{1-\gamma}, \ \text{if } \gamma \neq 1.$$

$$Z_t = Z_0 + \sigma B_t, \ Z_0 = \ln(Z_0), \ \text{if } \gamma = 1.$$

For notational convenience, we will stop the dependance of R on γ.

Proof of Proposition 5.3.2 We consider separately several cases.
(i) $\gamma < 1$: Then, the admissible domain of R is $(0, +\infty)$.
(i-1) $p \neq 1$: We first recall the value of the derivatives with respect to z:

$$\partial_z \bar{U}(x,z) = \frac{1}{1-\gamma} z^{\frac{\gamma}{1-\gamma}} \left(x + z^{\frac{1}{1-\gamma}}\right)^{-p}$$

$$\partial_{zz}\bar{U}(x,z) = \frac{1}{(1-\gamma)^2} z^{\frac{2\gamma-1}{1-\gamma}} \left(x + z^{\frac{1}{1-\gamma}}\right)^{-p-1} \left[\gamma\left(x + z^{\frac{1}{1-\gamma}}\right) - pz^{\frac{1}{1-\gamma}}\right]$$

(i-1a)$\gamma > p$: For any x, $\partial_{zz}\bar{U}(x,z) > 0$ for z large enough. Since the domain of this partial function is $(0,\infty)$, and $\bar{U}(x,z) \to +\infty$ when $z \to +\infty$, we have $\bar{U}^1(x,\cdot) = +\infty$. So $\bar{U}^\infty = \bar{U}^1 = +\infty$.
(i-1b) $\gamma = p$: For $x > 0$, $\partial_{zz}\bar{U}(x,z) > 0$ and the same scheme as above leads to $\bar{U}^1(x,z) = +\infty$. For $x \leq 0$, $\partial_{zz}\bar{U}(x,z) \leq 0$ and then $\bar{U}^1(x,z) = \bar{U}(x,z)$.
We then have $\bar{U}^1(x,z) = \bar{U}(x,z)\mathbf{1}_{x\leq 0} + \infty\mathbf{1}_{x>0}$. for $z \in (0,\infty)$, we then study $\bar{U}^1(\cdot,z)$ on $(-z^{\frac{1}{1-\gamma}}, \infty)$. Since $\bar{U}^1 = +\infty$ for x large enough, we have $\bar{U}^2(x,z) = +\infty$ for every (x,z) in the domain. So $\bar{U}^\infty = \bar{U}^2 = +\infty$
(i-1c) $\gamma < p$:
• $\gamma \leq 0$ leads to $\partial_{zz}\bar{U}(x,z) \leq 0$ so that $\bar{U}$ is concave w.r.t. x and z and then $\bar{U}^\infty = \bar{U}$.
• $\gamma > 0$. For $x \leq 0$, we have $\partial_{zz}\bar{U}(x,z) \leq 0$ so that $\bar{U}^1(x,\cdot) = \bar{U}(x,\cdot)$. For $x > 0$, there exists $z(x)$ such that $\partial_{zz}\bar{U}(x,z) > 0$ for $z < z(x)$ and $\partial_{zz}\bar{U}(x,z) \leq 0$ for $z \geq z(x)$. Since $\partial_z\bar{U}(x,z) \to 0$ when $z \to +\infty$, there exists $\tilde{z}(x)$ such that $\bar{U}^1(x,z) = U(x) + z\partial_z\bar{U}(x,\tilde{z}(x))$ for $z \leq \tilde{z}(x)$ and $\bar{U}^1(x,z) = \bar{U}(x,z)$ for $z > \tilde{z}(x)$. We see that $z(x)$ is the unique solution of:

$$\bar{U}(x,z(x)) - U(x) = z(x)\partial_z\bar{U}(x,z(x)).$$

i.e. if we denote $\xi(x) := x^{-1}z(x)^{\frac{1}{1-\gamma}}$, then $\xi(x)$ is the unique solution of:

$$\frac{(1+\xi)^{1-p} - 1}{1-p} = \frac{\xi}{1-\gamma}(1+\xi)^{-p}.$$

We easily observe that $\xi_0 := \xi(x)$ is independant of x and then:

$$\bar{U}^1(x,z) = \bar{U}(x,z)\mathbf{1}_{x\xi_0 \le z^{\frac{1}{1-\gamma}}} + \left(\frac{x^{1-p}-1}{1-p} + zx^{\gamma-p}\frac{\xi_0^\gamma}{1-\gamma}(1+\xi_0)^{-p}\right)\mathbf{1}_{x\xi_0 > z^{\frac{1}{1-\gamma}}}.$$

We focus on the derivation w.r.t. x on the interval $(\frac{z^{\frac{1}{1-\gamma}}}{\xi_0}, +\infty)$. Indeed, on $(-z^{\frac{1}{1-\gamma}}, \frac{z^{\frac{1}{1-\gamma}}}{\xi_0})$ we clearely have $\partial_{xx}\bar{U}^1(x,z) \le 0$.

On this domain, we have:

$$\partial_x\bar{U}^1(x,z) = x^{-p} + \frac{\gamma-p}{1-\gamma}x^{\gamma-p-1}z\xi_0^\gamma(1+\xi_0)^{-p}$$

$$\partial_{xx}\bar{U}^1(x,z) = -px^{-p-1}\left[1 - \frac{(\gamma-p)(\gamma-p-1)}{p(1-\gamma)}zx^{\gamma-1}\xi_0^\gamma(1+\xi_0)^{-p}\right].$$

We now discuss the possible signs of $\partial_{xx}\bar{U}^1$.

We denote for $\xi \in [0,\xi_0]$, the function $\Delta(\xi) := 1 - \frac{(p+1-\gamma)(p-\gamma)}{p(1-\gamma)}\xi_0^\gamma\xi^{1-\gamma}(1+\xi_0)^{-p}$. We are seeking a solution ξ_1 to the equation:

$$\Delta(\xi) = 0.$$

The function Δ is non-increasing with $\Delta(0) = 1$. So we have to discuss whether $\Delta(\xi_0)$ is positive or not. To achieve it, let us introduce the function $\tilde{\Delta}$ defined by:

$$\begin{aligned}\tilde{\Delta} \;:\; \mathbb{R}_*^+ &\longrightarrow \mathbb{R}_*^+ \\ x &\longmapsto 1 - \frac{(p+1-\gamma)(p-\gamma)}{p(1-\gamma)}x(1+x)^{-p}\end{aligned}$$

This is clearely a non-increasing continuous and one-to-one function on $\mathbb{R}_*^+$. And we can see that seeking the sign of $\Delta(\xi_0)$ remains to check the sign of $\tilde{\Delta}(x)$ under the condition $\Theta(x) = 0$. So let us consider now the following non-linear system of equations:

$$\tilde{\Delta}(x) = 0 \quad \text{and} \quad \Theta(x) = 0 \tag{5.6.1}$$

This is equivalent to:

$$\begin{aligned}(1+\xi_0)^{-p} &= \frac{1-\gamma}{1+p-\gamma} \\ 1 + \frac{(1+\xi_0)^{-p}}{1-\gamma}&\left[(\gamma-p)\,\xi_0 - (1-\gamma)\right] = 0\end{aligned}$$

We can see after calculus that the solution of (5.6.1) is $x = \frac{p}{p-\gamma}$. Moreover, for a fixed p, we have:

$$G(\gamma) = 0 \Leftrightarrow \text{ there is a unique solution to (5.6.1).}$$

Since G is a non-decreasing continuous and one-to-one function, it admits a unique solution $\hat{\gamma}_p$. Moreover, we have that G is negative on $\gamma \le \hat{\gamma}_p$ and positive on $\gamma > \hat{\gamma}_p$. This result gives us that:

$\star$ For $\gamma > \hat{\gamma}_p$, G positive implies $\tilde{\Delta}(x)$ negative. It means that $\Delta(\xi_0)$ is negative, so $\bar{U}^1$ is not concave in its first variable and admits an inflexion point to be determined.

$\star$ For $\gamma \le \hat{\gamma}_p$, G negative implies $\tilde{\Delta}(x)$ positive. This means that $\Delta(\xi_0)$ is positive, so $\bar{U}^1$ is concave in its first variable.

We now focus on the case $\gamma > \hat{\gamma}_p$. We are looking for a pair (x_1, x_2) such that $x_1 \le \frac{z^{\frac{1}{1-\gamma}}}{\xi_0} < x_2$ and x_1 maximal such that:

$$\frac{\bar{U}^1(x_2,z) - \bar{U}^1(x_1,z)}{x_2 - x_1} = \partial_x\bar{U}^1(x_2,z) \le \partial_x\bar{U}^1(x_1,z). \tag{5.6.2}$$

This is the characterization of the concave envelope of $\bar{U}^1$ w.r.t. x. We observe that this pair exists since $\partial_x \bar{U}^1(x,z) \to 0$ when $x \to +\infty$ and $\partial_x \bar{U}^1(x,z) \to +\infty$ when $x \to -z^{\frac{1}{1-\gamma}}$.
An other remark is that for any $\lambda > 0$, we have $\frac{\bar{U}^1(\lambda x_2, \lambda^{1-\gamma} z) - \bar{U}^1(\lambda x_1, \lambda^{1-\gamma} z)}{\lambda x_2 - \lambda x_1} = \lambda^{-p} \frac{\bar{U}^1(x_2,z) - \bar{U}^1(x_1,z)}{x_2 - x_1}$ and $\partial_x \bar{U}^1(\lambda x_i, \lambda^{1-\gamma} z) = \lambda^{-p} \partial_x \bar{U}^1(x_i, z)$ for $i \in \{1,2\}$. We then see that there exists ξ_1 and ξ_2 such that for any $(x,z) \in int(\bar{D})$, we have $(x_1, x_2) = (\frac{z^{\frac{1}{1-\gamma}}}{\xi_1}, \frac{z^{\frac{1}{1-\gamma}}}{\xi_2})$.

Finally we can compute the value of $\bar{U}^2$:

$$\begin{aligned}\bar{U}^2(x,z) =& \bar{U}(x,z) \mathbf{1}_{x\xi_1 \le z^{\frac{1}{1-\gamma}}} + \bar{U}^1(x,z) \mathbf{1}_{x\xi_2 \ge z^{\frac{1}{1-\gamma}}} + \left(\bar{U}^1 \left(\frac{z^{\frac{1}{1-\gamma}}}{\xi_2}, z \right) \right. \\ & \left. + \left(x - \frac{z^{\frac{1}{1-\gamma}}}{\xi_2} \right) \partial_x \bar{U}^1 \left(\frac{z^{\frac{1}{1-\gamma}}}{\xi_2}, z \right) \right) \mathbf{1}_{\frac{z^{\frac{1}{1-\gamma}}}{\xi_1} < x \le \frac{z^{\frac{1}{1-\gamma}}}{\xi_2}}.\end{aligned}$$

By construction, $\bar{U}^2$ is concave w.r.t. x. For the concavity w.r.t. z, we already know that $\partial_{zz} \bar{U}^2 \le 0$ out of $\left[(x\xi_2)^{1-\gamma}, (x\xi_2)^{1-\gamma} \right]$. We also obtain by tedious calculations that $\partial_{zz} \bar{U}^2 \le 0$ on $\left((x\xi_2)^{1-\gamma}, (x\xi_2)^{1-\gamma} \right)$, and that $\partial_{z-} \bar{U}^2 \left(x, (x\xi_2)^{1-\gamma} \right) \ge \partial_{x+} \bar{U}^2 \left(x, (x\xi_2)^{1-\gamma} \right)$, and $\partial_{z-} \bar{U}^2 \left(x, (x\xi_1)^{1-\gamma} \right) \ge \partial_{z+} \bar{U}^2 \left(x, (x\xi_1)^{1-\gamma} \right)$, where ∂_{z-} (resp ∂_{z+}) corresponds to the left derivative (resp the right derivative) with respect to z.

(i-2) $p = 1$: The derivatives w.r.t. z are:

$$\partial_z \bar{U}(x,z) = \frac{1}{1-\gamma} z^{\frac{\gamma}{1-\gamma}} \left(x + z^{\frac{1}{1-\gamma}} \right)^{-1},$$

$$\partial_{zz} \bar{U}(x,z) = \frac{1}{(1-\gamma)^2} z^{\frac{2\gamma-1}{1-\gamma}} \left(x + z^{\frac{1}{1-\gamma}} \right)^{-2} \left[\gamma \left(x + z^{\frac{1}{1-\gamma}} \right) - z^{\frac{1}{1-\gamma}} \right].$$

(i-2a) $\gamma \le 0$: In that situation $\partial_{zz} \bar{U} \le 0$ and then $\bar{U}^\infty = \bar{U}$.
(i-2b) $\gamma > 0$: If $x \le 0$, then $\partial_{zz} \bar{U}(x,z) \le 0$ and $\bar{U}^1(x,z) = \bar{U}(x,z)$.

If $x > 0$, there is an inflection point, similarly to the case $\gamma < p$, $p \neq 1$. We find $z(x)$ such that $\partial_{zz} \bar{U}(x,z) > 0$ for $z < z(x)$ and $\partial_{zz} \bar{U}(x,z) \le 0$ for $z \ge z(x)$. Since $\partial_z \bar{U}(x,z) \to 0$ when $z \to +\infty$, there exists $\tilde{z}(x)$ such that $\bar{U}^1(x,z) = U(x) + z \partial_z \bar{U}(x, \tilde{z}(x))$ for $z \le \tilde{z}(x)$ and $\bar{U}^1(x,z) = \bar{U}(x,z)$ for $z > \tilde{z}(x)$. We see that $z(x)$ is the unique solution of:

$$\bar{U}(x, z(x)) - U(x) = z(x) \partial_z \bar{U}(x, z(x)).$$

i.e. if we denote $\xi(x) := x^{-1} z(x)^{\frac{1}{1-\gamma}}$, then $\xi(x)$ is the unique solution of:

$$\ln(1+\xi) = \frac{\xi}{1-\gamma} (1+\xi)^{-1}.$$

We easily observe that $\xi_0 := \xi(x)$ is independant of x and then:

$$\bar{U}^1(x,z) = \bar{U}(x,z) \mathbf{1}_{x\xi_0 \le z^{\frac{1}{1-\gamma}}} + \left(\ln(x) + z x^{\gamma-1} \frac{\xi_0^\gamma}{1-\gamma} (1+\xi_0)^{-1} \right) \mathbf{1}_{x\xi_0 > z^{\frac{1}{1-\gamma}}}.$$

The derivation of $\bar{U}^2$ is similar to the previous case. Indeed, for $x \le \frac{z^{\frac{1}{1-\gamma}}}{\xi_0}$, $\partial_{xx} \bar{U}^1(x,z) \le 0$ by definition of U.
For $x \ge \frac{z^{\frac{1}{1-\gamma}}}{\xi_0}$, we have:

$$\partial_x \bar{U}^1(x,z) = \left[x^{-1} - z x^{\gamma-2} \xi_0^\gamma (1+\xi_0)^{-1} \right],$$

$$\partial_{xx} \bar{U}^1(x,z) = -x^{-2} \left[1 + (2-\gamma) z x^{\gamma-1} \xi_0^\gamma (1+\xi_0)^{-1} \right].$$

The exact same scheme as the one leading to the system of equations (5.6.1) leads to the existence of $\hat{\gamma}_1 \in (0,1)$ such that for $\gamma \leq \hat{\gamma}_1$, we have $\partial_{xx}\bar{U}^1 \leq 0$, and for $\gamma > \hat{\gamma}_1$, there exists an inflexion point.

It remains to solve the case $\gamma > \hat{\gamma}_1$. We are seeking for a pair (x_1, x_2) such that $x_1 \leq \frac{z^{\frac{1}{1-\gamma}}}{\xi_0} < x_2$ with x_1 maximal such that (5.6.2) is true. By the same arguments, there exists ξ_1 and ξ_2 such that for any $z > 0$, we have $(x_1, x_2) = \left(\frac{z^{\frac{1}{1-\gamma}}}{\xi_1}, \frac{z^{\frac{1}{1-\gamma}}}{\xi_2}\right)$ and:

$$\begin{aligned}\bar{U}^2(x,z) =& \bar{U}(x,z)\mathbf{1}_{x\xi_1 \leq z^{\frac{1}{1-\gamma}}} + \bar{U}^1(x,z)\mathbf{1}_{x\xi_2 \geq z^{\frac{1}{1-\gamma}}} \\ &+ \left(\bar{U}^1\left(\frac{z^{\frac{1}{1-\gamma}}}{\xi_2}, z\right) + \left(x - \frac{z^{\frac{1}{1-\gamma}}}{\xi_2}\right)\partial_x \bar{U}^1\left(\frac{z^{\frac{1}{1-\gamma}}}{\xi_2}, z\right)\right)\mathbf{1}_{\frac{z^{\frac{1}{1-\gamma}}}{\xi_1} < x \leq \frac{z^{\frac{1}{1-\gamma}}}{\xi_2}}.\end{aligned}$$

The concavity in z is easily obtained by direct calculations.
(ii) $\gamma = 1$: The admissible domain of R is $(-\infty, \infty)$.
(ii-1) $p \neq 1$: We have:

$$\partial_x \bar{U}(x,z) = e^z\left(x + e^z\right)^{-p},$$

$$\partial_{xx}\bar{U}(x,z) = e^z\left(x + e^z\right)^{-p-1}\left[(x + e^z) - pe^z\right].$$

(ii-1a) $p < 1$: If $x \geq 0$, then $\partial_{zz}\bar{U}(x,z) > 0$ and then since z is unbounded ($\forall z \in \mathbb{R}$, $x + e^z > 0$ if $x \geq 0$), and $\bar{U}(x,\cdot)$ is strictly convex and $\bar{U}(x,z) \to +\infty$ when $z \to +\infty$, we have $\bar{U}^1(x,z) = +\infty$.

For $x < 0$, we have $\partial_{zz}\bar{U}(x,z) \leq 0$ for $z \leq \ln\left(\frac{1-p}{x}\right)$ and $\partial_{zz}\bar{U}(x,z) > 0$ for $z > \left(\frac{1-p}{x}\right)$, and the same argument leads to $\bar{U}^1(x,z) = +\infty$. (ii-1b) $p > 1$: If $x \leq 0$, then $\partial_{zz}\bar{U}(x,z) \leq 0$ and $\bar{U}^1(x,z) = \bar{U}(x,z)$. For $x > 0$, we have $\partial_{zz}\bar{U}(x,z) > 0$ for $z < \ln\left(\frac{x}{p-1}\right)$ and $\partial_{zz}\bar{U}(x,z) \leq 0$ for $x \geq \ln\left(\frac{x}{p-1}\right)$. Since $\bar{U}(x,z) \to U(x) > -\infty$ when $z \to -\infty$, and $\bar{U}(x,z) \to -\frac{1}{1-p}$ when $z \to +\infty$, we have that the concave envelope is always equal to the limit when $z \to +\infty$, i.e. $\bar{U}^1(x,z) = \frac{1}{p-1}$. So:

$$\bar{U}^1(x,z) = \bar{U}(x,z)\mathbf{1}_{x\leq 0} + \frac{1}{p-1}\mathbf{1}_{x>0}.$$

In particular we see that $\bar{U}^1$ is not continuous.

The calculation of $\bar{U}^2$ is easier than in the previous cases. For a fixed $z \in \mathbb{R}$. We study $\bar{U}^1(\cdot,z)$ on $(-e^z, \infty)$. $\bar{U}^1(\cdot,z)$ is non decreasing, constant on $[0,\infty)$ and concave on $(-e^z, 0)$, with $\bar{U}^1(-e^z, z) = -\infty$. So there exists $x_0 \in (-e^z, 0)$ such that $\partial_x\bar{U}^1(x_0,z) = \frac{\bar{U}^1(0,z) - \bar{U}^1(x_0,z)}{-x_0}$, and $\bar{U}^2(\cdot,z)$ is linear on $(-x_0, 0)$ and $\bar{U}^2(x,z) = \bar{U}^1(x,z)$ elsewhere. x_0 is easily given by $x_0 = -\frac{e^z}{p}$ and then:

$$\begin{aligned}\bar{U}^2(x,z) =& \bar{U}(x,z)\mathbf{1}_{x\leq -\frac{e^z}{p}} - \frac{1}{1-p}\mathbf{1}_{x\geq 0} \\ &+ \left(\bar{U}\left(-\frac{e^z}{p}, z\right) + \left(x + \frac{e^z}{p}\right)e^{-pz}\left(1 - \frac{1}{p}\right)^{-p}\right)\mathbf{1}_{x\in\left(-\frac{e^z}{p}, 0\right)}.\end{aligned}$$

The partial concavity w.r.t. z is then trivial and we have $\bar{U}^\infty = \bar{U}^2$.
(ii-2) $p = 1$: we have:

$$\partial_z\bar{U}(x,z) = \left(1 + xe^{-z}\right)^{-1},$$

$$\partial_{zz}\bar{U}(x,z) = xe^{-z}\left(1 + xe^{-z}\right)^{-2}.$$

For $x > 0$, we have $\partial_{zz}\bar{U}(x,z) > 0$ and then as above, since $\bar{U}(x,z) \to \infty$ when $z \to \infty$, we have $\bar{U}^1(x,z) = \infty$.
For $x \geq 0$, we have $\partial_{zz}\bar{U}(x,z) \leq 0$ and then $\bar{U}^1(x,z) = \bar{u}(x,z)$. Summing up:

$$\bar{U}^1(x,z) = \bar{U}(x,z)\mathbf{1}_{x\leq 0} + \infty\mathbf{1}_{x>0}.$$

As a consequence, we see that:

$$\bar{U}^2 = +\infty.$$

(iii) $\gamma > 1$: The admissible domain of R is $(-\infty, 0)$. For any p, the partial derivatives w.r.t. z are given by:

$$\partial_z\bar{U}(x,z) = \frac{1}{\gamma-1}(-z)^{\frac{\gamma}{1-\gamma}}\left(x + (-z)^{\frac{1}{1-\gamma}}\right)^{-p},$$

$$\partial_{zz}\bar{U}(x,z) = \frac{1}{(\gamma-1)^2}(-z)^{\frac{2\gamma-1}{1-\gamma}}\left(x + (-z)^{\frac{1}{1-\gamma}}\right)^{-p-1}\left[\gamma\left(x + (-z)^{\frac{1}{1-\gamma}}\right) - p(-z)^{\frac{1}{1-\gamma}}\right].$$

(iii-1) $p \leq 1$: For any x, $\partial_{zz}\bar{U}(x,z) > 0$ for z large enough and $\bar{U}(x,z) \to +\infty$ when $z \to 0$ so that $\bar{U}^1(x,z) = +\infty$.
(iii-2) $1 < p < \gamma$: For $x \geq 0$, we have $\partial_z\bar{U}(x,z) \to 0$ when $z \to -\infty$ and $\bar{U}(x,z) \to \frac{1}{p-1}$ when $z \to 0$, so $\bar{U}^1(x,z) = \frac{1}{p-1}$.
For $x < 0$, for $z \leq -\left(\frac{\gamma}{p-\gamma}x\right)^{1-\gamma}$, $\partial_{zz}\bar{U}(x,z) \leq 0$ and for $z > -\left(\frac{\gamma}{p-\gamma}x\right)^{1-\gamma}$, $\partial_{zz}\bar{U}(x,z) > 0$. Since $\bar{U}(x,z) \to \frac{1}{p-1}$ when $z \to 0$, there exists z_0 such that $-z_0\partial_z\bar{U}(x,z_0) = \frac{1}{p-1} - \bar{U}(x,z_0)$. Similarly to the case $\gamma < 1$, z_0 verifies $(-z_0)^{\frac{1}{1-\gamma}} = -x\xi_0$ with $\xi_0 = \frac{\gamma-1}{\gamma-p}$.
We then have:

$$\begin{aligned}\bar{U}^1(x,z) =& \bar{U}(x,z)\mathbf{1}_{\{-x\xi_0 > (-z)^{\frac{1}{1-\gamma}}\}} + \frac{1}{p-1}\mathbf{1}_{\{x\geq 0\}} \\ &+ z(-x)^{\gamma-p}\frac{(p-1)^{-p}}{(\gamma-p)^{\gamma-p}}(\gamma-1)^{\gamma-1}\mathbf{1}_{\{0<-x\xi_0\leq(-z)^{\frac{1}{1-\gamma}}\}}\end{aligned}$$

The concavity of $\bar{U}^1$ w.r.t. x is then straightforward.
(iii-3) $p \geq \gamma$: For $x \leq 0$, $\partial_{zz}\bar{U}(x,z) \leq 0$ and $\bar{U}^1(x,z) = \bar{U}(x,z)$.
For $x > 0$, there is an inflexion point. Now since $\partial_z\bar{U}(x,z) \to 0$ when $z \to -\infty$, we have $\bar{U}^1(x,z) = \frac{1}{p-1}$. So:

$$\bar{U}^1(x,z) = \bar{U}(x,z)\mathbf{1}_{x\leq 0} + \frac{1}{p-1}\mathbf{1}_{x>0}.$$

We now search $\bar{U}^2$. For any $z \in (-\infty, 0)$, $\bar{U}^1(\cdot, z)$ is concave on $(-(-z)^{\frac{1}{1-\gamma}}, 0)$ and constant on $[0,\infty)$, and discontinuous at $x = 0$. We are looking for $x_0 \in (-(-z)^{\frac{1}{1-\gamma}}, 0)$ such that:

$$\bar{U}^1(0,z) - \bar{U}(x_0,z) = -x_0\partial_x\bar{U}(x_0,z).$$

The solution is given by $x_0 = \frac{1-p}{p}(-z)^{\frac{1}{1-\gamma}}$ and we have:

$$\begin{aligned}\bar{U}^2(x,z) =& \bar{U}(x,z)\mathbf{1}_{x<\frac{1-p}{p}(-z)^{\frac{1}{1-\gamma}}} + \frac{1}{p}\mathbf{1}_{x>0} \\ &+ \left((-z)^{\frac{1-p}{1-\gamma}} + \left(x + \frac{p-1}{p}(-z)^{\frac{1}{1-\gamma}}\right)p^p(-z)^{\frac{-p}{1-\gamma}}\right)\mathbf{1}_{\frac{1-p}{p}(-z)^{\frac{1}{1-\gamma}}\leq x<0}.\end{aligned}$$

The concavity of $\bar{U}^2$ w.r.t. z is easily verified.

□

Chapter 6

General indifference pricing of European options with small transaction costs

Contents

6.1 Introduction

This chapter remains in the context of the general approach initiated by [90], and our main goal is to provide rigorous asymptotic expansions of the utility indifference price of European contingent claims in general Markovian, multidimensional models and with general utility functions. To the best of our knowledge, the only related papers in the literature are [16] and the very recent manuscript [25]. However, the level of generality we consider seems to be new, in particular since both these works are restricted to the one dimension case. Furthermore, [16] is restricted to exponential utilities, because their scaling properties allow to deduce directly and completely explicitly the price from the value function of the control problem. Hence, it suffices to obtain the expansion for the value function to obtain the expansion for the price, whereas in our case, even though we follow the same approach, the expansion for the price cannot be deduced so easily. Moreover, our method of proof allows to weaken strongly the assumptions made in [16], since, for instance, we only need to assume continuity of the option payoffs we consider, while [16] needed C^4 regularity. When compared with [25], even though we think that their approach could reasonably be extended to the same multi-dimensional setting as ours, the methods with which they approach the problem is different from ours, since they attack directly the expansion for the price, while we first start with an expansion for the value function. Besides, the set of assumptions under which their result for the utility indifference price holds true also implies strong regularity for the payoff functions, which makes our result more general.

The rest of the chapter is organized as follows. In Section 6.2, we present succinctly the markets we consider, with and without frictions, and we follow the general approach of [90] to give formal asymptotics for both the value function and the utility indifference price. Section 6.3 is then devoted to the main results of the chapter, as well as the general assumptions under which we will be working and the proof of the expansion for the price. Then, in Section 6.4, we discuss some particular examples and compare our result with the existing literature. Finally, Section 6.5 provides the proof of all the technical results of the chapter.

6.2 General setting

In this section we describe the problem and recall the way to obtain formal asymptotics.

6.2.1 Financial market with frictions

We work on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a d-dimensional Brownian motion W. For a fixed time horizon $T > 0$, and for any $t \in [0, T]$, we define the filtration $\mathbb{F}^t := (\mathcal{F}^t_s)_{t \le s \le T}$ to be the completed natural filtration of the process W^t, defined by

$$W^t_s := W_s - W_t,\ \ s \in [t, T].$$

For notational simplicity, we let $\mathbb{F} := \mathbb{F}^0$. The financial market consists of a non-risky asset S^0 and d risky assets with price process $\{S_t = (S^1_t, \dots, S^d_t), t \in [0, T]\}$ given by the stochastic differential equations (SDEs),

$$\frac{dS^0_t}{S^0_t} = r(t, S_t)dt,\quad \frac{dS^i_t}{S^i_t} = \mu^i(t, S_t)dt + \sum_{j=1}^{d} \sigma^{i,j}(t, S_t)dW^j_t,\quad 1 \le i \le d,$$

where $r : [0, T] \times \mathbb{R}^d \to \mathbb{R}_+$ is the instantaneous interest rate and $\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathcal{M}_d(\mathbb{R})$ are the coefficients of instantaneous mean return and volatility, satisfying the standing assumptions:

$$r, \mu, \sigma \text{ are bounded and Lipschitz, and } (\sigma\sigma^T)^{-1} \text{ is bounded.}$$

In particular, this guarantees the existence and the uniqueness of a strong solution to the above stochastic differential equations (SDEs).

The portfolio of an investor is represented by the dollar value X invested in the non-risky asset, the vector process $Y = (Y^1, \dots, Y^d)$ of the value of the positions in each risky asset, and a short position in a European option represented by some payoff function $g : \mathbb{R}^d \longrightarrow \mathbb{R}_+$, that he has to hold until the final time T. Starting from any time $t \in [0,T]$, these state variables are controlled by the choices of the total amount of transfers $L_s^{i,j}$, $0 \leq i,j \leq d$, from the i-th to the j-th asset cumulated between time t and s. Naturally, the control processes $\{L_s^{i,j}, s \geq t\}$ are defined as RCLL, nondecreasing, $\mathbb{F}^t$-progressively measurable processes with $L_{t^-}^{i,j} = 0$ and $L^{i,i} \equiv 0$.

In addition to the trading activity, the investor consumes between time t and T at a rate determined by a nonnegative $\mathbb{F}^t$-progressively measurable process $\{c_s, t \leq s \leq T\}$. Here c_s represents the rate of consumption in terms of the non-risky asset S^0, which means that the investor can only consume from the bank account. Such a pair $\nu := (c, L)$ is called a *consumption-investment strategy*. For any $t \in [0,T]$ and any initial position $(X_{t^-}, Y_{t^-}) = (x,y) \in \mathbb{R} \times \mathbb{R}^d$, the portfolio positions of the investor are given by the following state equation

$$dX_u = \big(r(u,S_u)X_u - c_u\big)du + \mathbf{R}^0(dL_u), \quad \text{and} \quad dY_u^i = Y_u^i \frac{dS_u^i}{S_u^i} + \mathbf{R}^i(dL_u), \quad i = 1,\dots,d,$$

where

$$\mathbf{R}^i(\ell) := \sum_{j=0}^{d} \big(\ell^{j,i} - (1+\varepsilon^3\lambda^{i,j})\ell^{i,j}\big), \quad i = 0,\dots,d, \text{ for all } \ell \in \mathcal{M}_{d+1}(\mathbb{R}_+),$$

is the change of the investor's position in the i-th asset induced by a transfer policy ℓ, given a structure of proportional transaction costs $\varepsilon^3\lambda^{i,j}$ for any transfer from asset i to asset j. Here, $\varepsilon > 0$ is a small parameter, $\lambda^{i,j} \geq 0$, $\lambda^{i,i} = 0$, for all $i,j = 0,\dots,d$, and the scaling ε^3 is chosen to state the expansion results simpler. In some instances, we may forbid transactions between certain assets by setting the corresponding transaction costs to $+\infty$, however we will always allow transactions from and to the bank account, that is to say that we always assume

$$\lambda^{i,0} + \lambda^{0,i} < +\infty, \; i = 1,\dots,d.$$

For simplicity, we will also denote

$$\mathcal{I} := \Big\{(i,j) \in \{0,\dots,d\}^2, \lambda^{i,j} < +\infty\Big\}.$$

Let $(X,Y)^{\nu,t,s,x,y}$ denote the controlled state process. A consumption-investment strategy ν is said to be *admissible* for the initial position (t,s,x,y), if the induced state process is well defined and satisfies the solvency condition $(X,Y)_s^{\nu,t,s,x,y} \in K_\varepsilon$, for all $s \in [t,T]$, $\mathbb{P}-$a.s., where the solvency region is defined by:

$$K_\varepsilon := \Big\{(x,y) \in \mathbb{R} \times \mathbb{R}^d : \; (x,y) + \mathbf{R}(\ell) \in \mathbb{R}_+^{1+d} \text{ for some } \ell \in \mathcal{M}_{d+1}(\mathbb{R}_+)\Big\}.$$

The corresponding set of admissible strategies is denoted by $\Theta^\varepsilon(t,s,x,y)$. For given initial positions $S_t = s \in \mathbb{R}_+^d$, $X_{t^-} = x \in \mathbb{R}$, $Y_{t^-} = y \in \mathbb{R}^d$ and given $\nu \in \Theta^\varepsilon(t,s,x,y)$, we denote by $S^{t,s}$, $X^{t,s,x,y,\nu}$ and $Y^{t,s,x,y,\nu}$ the corresponding prices and state processes. The consumption-investment problem is then the following maximization problem,

$$v^{\varepsilon,g}(t,s,x,y) := \sup_{\nu \in \Theta^\varepsilon(t,s,x,y)} \mathbb{E}_t\left[\int_t^T \kappa e^{-\int_t^\xi k(\iota,S_\iota)d\iota} U_1(c_\xi)d\xi + e^{-\int_t^T k(\xi,S_\xi)d\xi}\mathcal{U}_2^{\varepsilon,g}\right], \tag{6.2.1}$$

where $\kappa \in \{0,1\}$ is here so that we can consider simultaneously the problems with or without consumption and where $k : [0,T] \times \mathbb{R}^d$, $U_1 : (0,\infty) \mapsto \mathbb{R}$ and $\mathcal{U}_2$ is defined by

$$\mathcal{U}_2^{\varepsilon,g} := U_2\left(\ell^\varepsilon\left(X_T^{t,s,x,y,\nu}, Y_T^{t,s,x,y,\nu}\right) - g(S_T^{t,s})\right),$$

for some function $U_2 : \mathbb{R} \mapsto \mathbb{R}$ and a liquidation function $\ell^\varepsilon : \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ defined by

$$\ell^\varepsilon(x,y) := x + \sum_{j=1}^{d} \left(\frac{y_j}{1+\varepsilon^3 \lambda^{j,0}} \mathbf{1}_{y_j \geq 0} + (1+\varepsilon^3 \lambda^{0,j}) y_j \mathbf{1}_{y_j<0} \right).$$

We emphasize here that the choice of such a liquidation function implies that at time T, the investor will liquidate all his positions on the risky assets to only have cash. Moreover we assume that U^1 and U^2 are utility functions which are C^2, increasing and strictly concave. We also denote the convex conjugate of U_1 by,

$$\widetilde{U}_1(\tilde{c}) \quad := \quad \sup_{c>0} \left\{ U^1(c) - c\tilde{c} \right\}, \qquad \tilde{c} \in \mathbb{R},$$

and by $\mathrm{Supp}(\tilde{U}^1)$ its support, that is to say the points $\tilde{c} \in \mathbb{R}$ such that $\tilde{U}^1(\tilde{c}) < +\infty$.

6.2.2 The Merton problem without frictions

The Merton value function $v^g := v^{0,g}$ corresponds to the limiting case $\varepsilon = 0$, where there is no transaction costs. In this case, there is no longer any need to keep track of the transfers between the different assets, and we can take as a state variable the total wealth obtained by aggregating the positions on all the assets. We therefore define $Z := X + Y \cdot \mathbf{1}_d$, where $\mathbf{1_d} \in \mathbb{R}^{\mathbf{d}}$ is a vector of ones. The dynamics of Z are then given by

$$dZ_t = (r(t,S_t)Z_t - c_t)\,dt + \sum_{i=1}^{d} Y_t^i \left(\frac{dS_t^i}{S_t^i} - r(t,S_t)dt \right). \tag{6.2.2}$$

In this context, for any $t \in [0,T]$, the set of admissible investment-consumption strategies starting from time t corresponds to the $\mathfrak{v} := (c,Y) \in \mathbb{R}_+ \times \mathbb{R}^d$ such that the corresponding wealth process Z is well-defined and remains $\mathbb{P} - a.s.$ non-negative between t and T. For initial conditions $S_t = s$ and $Z_t = z$, we denote this set by $\Theta^0(t,s,z)$. Moreover, for any $\mathfrak{v} \in \Theta^0(t,s,z)$, we denote by $S^{t,s}$ and $Z^{t,s,z,\mathfrak{v}}$ the corresponding stock prices and wealth process. The value function of the Merton problem is then

$$v^g(t,s,z) := \sup_{\mathfrak{v} \in \Theta^0(t,s,z)} \mathbb{E}_t \left[\int_t^T \kappa e^{-\int_t^\xi k(\iota,S_\iota)d\iota} U_1(c_\xi) d\xi + e^{-\int_t^T k(\xi,S_\xi)d\xi} \, \mathcal{U}_2^{0,g} \right], \tag{6.2.3}$$

where we have defined

$$\mathcal{U}_2^{0,g} := U_2\left(Z_T^{t,s,z,\mathfrak{v}} - g(S_T^{t,s}) \right).$$

We will assume in the following that v^g is smooth, so that it is the unique classical solution of the HJB equation

$$kv^g - rzv_z^g - \mathcal{L}^0 v^g - \kappa \widetilde{U}_1(v_z^g) - \sup_{y \in \mathbb{R}^d} \left\{ y \cdot \left((\mu - r\mathbf{1}_d) v_z^g + \sigma\sigma^{\mathrm{T}} \mathbf{D}_{sz} v^g \right) + \frac{1}{2} |\sigma^{\mathrm{T}} y|^2 v_{zz}^g \right\} = 0$$
$$v^g(T,s,z) = U_2(z - g(s)), \tag{6.2.4}$$

where $\mathbf{D}_{sz} := \frac{\partial}{\partial z} \mathbf{D}_s$, and

$$\mathcal{L}^0 := \frac{\partial}{\partial t} + \mu \cdot \mathbf{D}_s + \frac{1}{2} \mathrm{Tr}\left[\sigma\sigma^{\mathrm{T}} \mathbf{D}_{ss} \right], \tag{6.2.5}$$

with for $i,j = 1,\dots,d$,

$$\mathbf{D}_s^i := s^i \frac{\partial}{\partial s^i}, \; \mathbf{D}_{ss}^{i,j} := s^i s^j \frac{\partial^2}{\partial s^i \partial s^j}, \; \mathbf{D}_s = (\mathbf{D}_s^i)_{1 \le i \le d}, \text{ and } \mathbf{D}_{ss} := (\mathbf{D}_{ss}^{i,j})_{1 \le i,j \le d}.$$

Moreover, for a smooth scalar functions $(t,s,x,y) \in [0,T] \times \mathbb{R}_+^d \times \mathbb{R} \times \mathbb{R}^d \longmapsto \psi(t,s,x,y)$ and $(t,s,z) \in [0,T] \times \mathbb{R}_+^d \times \mathbb{R}_+ \longmapsto \varphi(t,s,z)$ we set

$$\psi_x := \frac{\partial \psi}{\partial x} \in \mathbb{R}, \ \psi_y := \frac{\partial \psi}{\partial y} \in \mathbb{R}^d, \ \varphi_z := \frac{\partial \varphi}{\partial z} \in \mathbb{R}.$$

The optimal consumption and positioning in the various assets are defined by the functions $\mathbf{c}^g(t,s,z)$ and $\mathbf{y}^g(t,s,z)$ obtained as the maximizers of the Hamiltonian in the HJB equation. For any $s \in \mathbb{R}_+^d$ and any $z \in \mathbb{R}_+$

$$\mathbf{c}^g(t,s,z) := -\kappa \widetilde{U}_1'\left(v_z^g(t,s,z)\right) = \kappa \left(U_1'\right)^{-1}\left(v_z^g(t,s,z)\right) \tag{6.2.6}$$

$$-v_{zz}^g(t,s,z)\sigma\sigma^{\mathrm{T}}(t,s)\mathbf{y}^g(t,s,z) = (\mu - r\mathbf{1}_d)(t,s)v_z^g(t,s,z) + \sigma\sigma^{\mathrm{T}}(t,s)\mathbf{D}_{sz}v^g(t,s,z). \tag{6.2.7}$$

6.2.3 The utility indifference price

We are interested in the so-called utility indifference price of the European option g, in both models with or without frictions. They are defined respectively by:

$$p^{\varepsilon,g}(t,s,x) := \inf\left\{p \in \mathbb{R} : v^{\varepsilon,g}(t,s,x+p,0) \geq v^{\varepsilon,0}(t,s,x,0)\right\} \tag{6.2.8}$$

$$p^g(t,s,x) := \inf\left\{p \in \mathbb{R} : v^g(t,s,x+p) \geq v^0(t,s,x)\right\}, \tag{6.2.9}$$

where $v^{\varepsilon,0}$ and v^0 correspond respectively to the value functions of the problems (6.2.1) and (6.2.4) without the option, that is to say when $g = 0$. Notice also that we consider here that the initial endowments of the investor are in cash only. This is purely for simplicity and all our results could be easily generalized if we allow the investor to have a non-zero position on the risky assets for the problem with frictions.

6.2.4 Dynamic programming

The dynamic programming equation corresponding to the singular stochastic control problem $v^{\varepsilon,g}$ involves the following differential operators. Let:

$$\mathcal{L} \ := \ \frac{\partial}{\partial t} + \mu \cdot (\mathbf{D}_s + \mathbf{D}_y) + r\mathbf{D}_x + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^{\mathrm{T}}\left(\mathbf{D}_{yy} + \mathbf{D}_{ss} + 2\mathbf{D}_{sy}\right)\right], \tag{6.2.10}$$

and for $i,j = 1,\dots,d$,

$$\mathbf{D}_x := x\frac{\partial}{\partial x}, \ \mathbf{D}_y^i := y^i\frac{\partial}{\partial y^i}, \ \mathbf{D}_{yy}^{i,j} := y^iy^j\frac{\partial^2}{\partial y^i \partial y^j}, \ \mathbf{D}_{sy}^{i,j} := s^iy^j\frac{\partial^2}{\partial s^i \partial y^j}$$
$$\mathbf{D}_y = (\mathbf{D}_y^i)_{1\leq i\leq d}, \ \mathbf{D}_{yy} := (\mathbf{D}_{yy}^{i,j})_{1\leq i,j\leq d}, \ \mathbf{D}_{sy} := (\mathbf{D}_{sy}^{i,j})_{1\leq i,j\leq d}.$$

Theorem 6.2.1. *Assume that $v^{\varepsilon,g}$ is locally bounded, then it is a viscosity solution of*

$$\begin{cases} \min\limits_{(i,j)\in\mathcal{I}}\left\{kv^{\varepsilon,g} - \mathcal{L}v^{\varepsilon,g} - \kappa\widetilde{U}_1(v_x^{\varepsilon,g}), \ \Lambda_{i,j}^{\varepsilon}\cdot(v_x^{\varepsilon,g}, v_y^{\varepsilon,g})\right\} = 0, \ (t,s,x,y) \in [0,T) \times \mathbb{R}_+^d \times K_\varepsilon \\ v^{\varepsilon,g}(T,s,x,y) = U_2\left(\ell^\varepsilon(x,y) - g(s)\right), \ (s,x,y) \in \mathbb{R}_+^d \times K_\varepsilon, \end{cases} \tag{6.2.11}$$

where $\Lambda_{i,j}^{\varepsilon} := e_i - e_j + \varepsilon^3\lambda^{i,j}\ e_i, \ 0 \leq i,j \leq d$. Moreover $v^{\varepsilon,g}$ converges to the Merton value function v^g, as ε tends to zero.

Let us point out that the result as stated above does not seem to be present in the literature (at least as far as we know) on the subject. Several related results, can be found however, for instance with infinite time-horizon and without consumption (see Kabanov and Safarian [61]), or when consumption and transfers between the risky assets are not allowed (see Akian, Menaldi and Sulem [1] or Akian, Séquier and Sulem [2]). Nonetheless, this is a classical result and does not lie at the heart of our analysis. We will therefore refrain from writing its proof.

6.2.5 Formal Asymptotics for the value function

Based on [90] and [81], we postulate the following expansion for $(t, s, x, y) \in [0, T) \times \mathbb{R}^d_+ \times K_\varepsilon$

$$v^{\varepsilon,g}(t, s, x, y) = v^g(t, s, z) - \varepsilon^2 u^g(t, s, z) - \varepsilon^4 w^g(t, s, z, \xi) + \circ(\varepsilon^2), \tag{6.2.12}$$

where we recall that $z = x + y \cdot \mathbf{1}_d$ and we define the "fast" variable $\xi \in \mathbb{R}^d$ by

$$\xi^i := \xi^i_\varepsilon(t, s, x, y) = \frac{y^i - \mathbf{y}^{g,i}(t, s, z)}{\varepsilon},\ 1 \le i \le d,$$

with the additional useful convention $\xi^0 = 0$. We now derive the key equations verified by u^g and w^g, from the dynamic programming equation (6.2.11). The easiest part corresponds to the gradient constraint in (6.2.11). By straightforward formal calculations, we have for $(i, j) \in \mathcal{I}$

$$\Lambda^\varepsilon_{i,j} \cdot (v^{\varepsilon,g}_x, v^{\varepsilon,g}_y) = \varepsilon^3 \left(\lambda^{i,j} v^g_z + (e_i - e_j) \cdot D_\xi w^g\right) + \circ(\varepsilon^3) = \varepsilon^3 \left(\lambda^{i,j} v^g_z + w^g_{\xi_i} - w^g_{\xi_j}\right) + \circ(\varepsilon^3).$$

We now explore the drift condition in (6.2.11). Thank to the linearity of $\mathcal{L}$, we decompose the calculation in several parts. First of all we have using (6.2.3) and (6.2.7)

$$\begin{aligned} kv^g - \mathcal{L}v^g - \kappa\tilde{U}_1(v^g_x) &= kv^g - \partial_t v^g - \mu \cdot \mathbf{D}_s v^g - rzv^g_z + y \cdot (r\mathbf{1}_d - \mu)\, v^g_z \\ &\quad - \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^T\left(\mathbf{D}_{yy}v^g + \mathbf{D}_{ss}v^g + 2\mathbf{D}_{sy}v^g\right)\right] - \kappa\tilde{U}_1(v^g_x) \\ &= (\mathbf{y}^g - y) \cdot \left((\mu - r\mathbf{1}_d)\, v^g_z + \sigma\sigma^T\mathbf{D}_{sz}v^g\right) + \frac{1}{2}|\sigma^T\mathbf{y}^g|^2 v^g_{zz} - \frac{1}{2}|\sigma^T y|^2 v^g_{zz} \\ &= -\, v^g_{zz}(\mathbf{y}^g - y) \cdot \sigma\sigma^T\mathbf{y}^g + \frac{1}{2}\left(|\sigma^T\mathbf{y}^g|^2 v^g_{zz} - |\sigma^T y|^2 v^g_{zz}\right) \\ &= -\, \frac{1}{2}\left|\sigma^T(\mathbf{y}^g - y)\right|^2 v^g_{zz} = -\frac{\varepsilon^2}{2}\left|\sigma^T\xi\right|^2 v^g_{zz}. \end{aligned} \tag{6.2.13}$$

Similarly, we obtain by straightforward but tedious calculations that

$$\varepsilon^4 (kw^g - \mathcal{L}w^g) = \frac{\varepsilon^4}{2}\mathrm{Tr}\left[\mathbf{D}_{yy}w^g + \mathbf{D}_{ss}w^g + 2\mathbf{D}_{sy}w^g\right] + \circ(\varepsilon^2) = \frac{\varepsilon^2}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T w^g_{\xi\xi}\right] + \circ(\varepsilon^2),$$

where the diffusion coefficient is given by

$$\alpha^g(t, s, z) := \left[\left(I_d - \mathbf{y}^g_z(t, s, z)\mathbf{1}^T_d\right)\mathrm{diag}[\mathbf{y}^g(t, s, z)] - (\mathbf{y}^g_s)^T(t, s, z)\mathrm{diag}[s]\right]\sigma(t, s). \tag{6.2.14}$$

This calculation highlights the role played by the so-called fast variable ξ. Indeed any of the second order derivatives of w^g with respect to s or y, corresponds to a second-order derivative of $\hat{w}^g$ scaled by $1/\varepsilon^2$. These terms are then exactly of the same order as the one obtained above

Finally it is obvious that, using the definition of $\mathbf{c}^g$ in (6.2.6):

$$\kappa\tilde{U}_1(v^{\varepsilon,g}_x) - \kappa\tilde{U}_1(v^g_x) + \kappa\mathbf{c}^g(v^{\varepsilon,g}_x - v^g_x) = \kappa\tilde{U}_1(v^{\varepsilon,g}_x) - \kappa\tilde{U}_1(v^g_x) - \varepsilon^2\kappa\mathbf{c}^g u^g_z + \circ(\varepsilon^2) = \circ(\varepsilon^2).$$

Combining these approximations and putting them into the drift condition of (6.2.11), we obtain that u^g must be solution of the second corrector equation:

$$\begin{cases} \mathcal{A}^g u^g = a^g(t, s, z),\ (t, s, z) \in [0, T) \times (0, +\infty)^{d+1} \\ u^g(T, s, z) = 0,\ (s, z) \in (0, +\infty)^{d+1}, \end{cases} \tag{6.2.15}$$

where the differential operator $\mathcal{A}^g$ is defined by

$$\mathcal{A}^g u := ku - \mathcal{L}^0 u - \left(rz + \mathbf{y}^g \cdot (\mu - r\mathbf{1}_d) - \kappa\mathbf{c}^g\right)u_z - \frac{1}{2}|\sigma^{\mathrm{T}}\mathbf{y}^g|^2\, u_{zz} - \sigma\sigma^{\mathrm{T}}\mathbf{y}^g \cdot \mathbf{D}_{sz}u,$$

and the function a^g is the second component of the solution (w^g, a^g) of the first corrector equation:

$$\max_{(i,j)\in\mathcal{I}} \max\left\{\frac{|\sigma(t,s)\xi|^2}{2} v^g_{zz}(t,s,z) - \frac{1}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T(t,s,z) w^g_{\xi\xi}(t,s,z,\xi)\right] + a^g(t,s,z)\ ,\right.$$
$$\left. -\lambda^{i,j} v^g_z(t,s,z) + \frac{\partial w^g}{\partial \xi^i}(t,s,z,\xi) - \frac{\partial w^g}{\partial \xi^j}(t,s,z,\xi)\right\} = 0,\ \xi\in\mathbb{R}^d. \qquad (6.2.16)$$

Remark 6.2.1. *Notice that naturally we consider (6.2.16) only on $[0,T)\times\mathbb{R}_+\times\mathbb{R}_+$, because since the value function is known at time T, its expansion is trivial and takes the form*

$$\begin{aligned} v^{g,\varepsilon}(T,s,x,y) &= U_2\left(\ell^\varepsilon(x,y) - g(s)\right)\\ &= U_2\Big(z - g(s) - \varepsilon^3\sum_{j=1}^d \Big(\lambda^{j,0}y_j^+ + \lambda^{0,j}y_j^-\Big)\Big)\\ &= v^g(T,s,z) - \varepsilon^3\sum_{j=1}^d \Big(\lambda^{j,0}y_j^+ + \lambda^{0,j}y_j^-\Big) U_2'\left(z-g(s)\right) + \circ(\varepsilon^3), \end{aligned}$$

where we used the fact that U^2 is C^2 and where the expansion is locally uniform in (s,x,y) since all the functions appearing are continuous.

Since we enforce that the function u^g solution of the second corrector equation (6.2.15) is null at time T, it would seem reasonable to think that the expansion (6.2.12) also holds at time T. However, as we will see in our proofs, this will usually only be true if the Merton value function and the corresponding optimal strategy are smooth enough at time T. If explosions are allowed at time T (which, as pointed out in Section 6.4, can happen for the derivatives of $\mathbf{y}^g$ if g is a Call option), then the remainder in the expansion (6.2.12) can become unbounded near T. In the previous works by Bichuch [16] and Bouchard, Moreau and Soner [25], strong regularity on v^g up to time T was assumed (which implies then that the payoff g has to be regular), in order to prevent $\mathbf{y}^g$ and several of its derivatives from exploding at T. With our method however, this is no longer needed. We refer the reader to Section 6.3.1 for more details on our assumptions.

Finally, we recall from [90] and [81] the following normalization. Set

$$\eta^g(t,s,z) := -\frac{v^g_z(t,s,z)}{v^g_{zz}(t,s,z)},\ \rho := \frac{\xi}{\eta^g(t,s,z)},\ \overline{w}^g(t,s,z,\rho) := \frac{w^g(t,s,z,\eta^g(t,s,z)\rho)}{\eta^g(t,s,z)v^g_z(t,s,z)},$$
$$\overline{a}^g(t,s,z) := \frac{a^g(t,s,z)}{\eta^g(t,s,z)v^g_z(t,s,z)},\ \bar{\alpha}^g(t,s,z) := \frac{\alpha^g(t,s,z)}{\eta^g(t,s,z)},$$

so that the corrector equations with variable $\rho\in\mathbb{R}^d$ have the form,

$$\begin{cases} \displaystyle\max_{(i,j)\in\mathcal{I}}\max\Big\{-\frac{\left|\sigma^T(t,s)\rho\right|^2}{2} - \frac{1}{2}\mathrm{Tr}\left[\bar{\alpha}^g(\bar{\alpha}^g)^T(t,s,z)\overline{w}^g_{\rho\rho}(t,s,z,\rho)\right] + \overline{a}^g(t,s,z)\ ;\\ \displaystyle\qquad -\lambda^{i,j} + \frac{\partial\overline{w}^g}{\partial\rho_i}(t,s,z,\rho) - \frac{\partial\overline{w}^g}{\partial\rho_j}(t,s,z,\rho)\Big\} = 0 \\ \mathcal{A}^g u^g(t,s,z) = v^g_z(t,s,z)\eta^g(t,s,z)\overline{a}^g(t,s,z). \end{cases} \qquad (6.2.17)$$

We emphasize that the first corrector equation (6.2.17) is an equation for the variable ξ, (t,s,z) are only parameters. Moreover, the wellposedness of this equation has been obtained in [81]. We recall below the properties of $\overline{w}^g$ that we will use. Before stating the result, let us define the following closed convex subset of $\mathbb{R}^d$, and the corresponding support function

$$C := \left\{\rho\in\mathbb{R}^d : -\lambda^{j,i}\le\rho_i-\rho_j\le\lambda^{i,j},\ (i,j)\in\mathcal{I}\right\},\quad \delta_C(\rho) := \sup_{u\in C} u\cdot\rho,\quad \rho\in\mathbb{R}^d,$$

with the convention that $\rho_0 = 0$.

Proposition 6.2.1. *Assume that $\bar{\alpha}^g(t,s,z)$ is non-degenerate for any $(t,s,z) \in [0,T) \times \mathbb{R}_+^d \times \mathbb{R}_+$. Then there exists a unique solution $(\overline{w}^g, \overline{a}^g)$ of the equation (6.2.17), such that $\overline{a}^g \in \mathbb{R}_+$, $\rho \longmapsto \overline{w}^g(t,s,z,\rho)$ is C^1 in $\mathbb{R}^d$ with a Lipschitz gradient, such that the following growth condition is satisfied*

$$\lim_{|\rho| \to +\infty} \frac{\overline{w}^g(t,s,z,\rho)}{\delta_C(\rho)} = 1,$$

and such that $\overline{w}^g(\cdot, 0) = 0$. Moreover, for any $(t,s,z) \in [0,T) \times \mathbb{R}_+^d \times \mathbb{R}_+$

- *$\overline{w}^g(t,s,z,\cdot)$ is convex.*

- *The set $\mathcal{O}_0^g(t,s,z) := \left\{\rho \in \mathbb{R}^d,\ \overline{w}_\rho^g(t,s,z,\rho) \in \mathrm{int}(C)\right\}$ is open and bounded, the map $\rho \longmapsto \overline{w}^g(t,s,z,\rho)$ is C^∞ on $\mathcal{O}_0^g(t,s,z)$ and $\overline{w}^g(t,s,z,\rho)$ attains its minimum in ρ at some point $\rho^*(t,s,z)$ in $\mathcal{O}_0^g(t,s,z)$.*

- *There is a constant $M > 0$ such that $0 \leq \overline{w}_{\rho\rho}^g(t,s,z,\rho) \leq M \mathbf{1}_{\overline{\mathcal{O}}_0^g(t,s,z)}(\rho)$ for a.e. $\rho \in \mathbb{R}^d$.*

Of course, under suitable regularity assumptions on η^g and v_z^g, the function w^g satisfies similar properties.

6.2.6 Formal asymptotics for the utility indifference price

We now develop an expansion for $p^{\varepsilon,g}$, using the expansion of $v^{\varepsilon,g}$ defined in (6.2.12). We first recall that, at least formally

$$v^{\varepsilon,g}(t,s,x,y) = v^g(t,s,x,y) - \varepsilon^2 u^g(t,s,z) + \circ(\varepsilon^2)$$
$$v^{\varepsilon,0}(t,s,x,y) = v^0(t,s,x,y) - \varepsilon^2 u^0(t,s,z) + \circ(\varepsilon^2).$$

Then, at least if $v^{\varepsilon,g}$ is increasing with respect to x, $p^{\varepsilon,g}(t,s,x)$ should be such that:

$$v^{\varepsilon,g}(t,s,x+p^{\varepsilon,g}(t,s,x),0) = v^{\varepsilon,0}(t,s,x,0).$$

We conjecture (and we will prove under natural assumptions) that $p^{\varepsilon,g}$ should satisfy the following expansion

$$p^{\varepsilon,g}(t,s,x) = p^g(t,s,x) + \varepsilon^2 h^g(t,s,x) + \circ(\varepsilon^2), \tag{6.2.18}$$

for some function h^g to be determined. Using (6.2.12), we obtain formally

$$\begin{aligned} v^0(t,s,x) - \varepsilon^2 u^0(t,s,x) + \circ(\varepsilon^2) &= v^g(t,s,x+p^{\varepsilon,g}(t,s,x),0) - \varepsilon^2 u^g(t,s,x+p^{\varepsilon,g}(t,s,x)) + \circ(\varepsilon^2) \\ &= v^g(t,s,x+p^g(t,s,x)) + \varepsilon^2 v_x^g(t,s,x+p^g(t,s,x)) h^g(t,s,x) \\ &\quad - \varepsilon^2 u^g(t,s,x+p^g(t,s,x)) + \circ(\varepsilon^2). \end{aligned}$$

Since by definition we have $v^0(t,s,x,0) = v^g(t,s,x+p^g(t,s,x))$, we deduce:

$$h^g(t,s,x) = \frac{u^g(t,s,x+p^g(t,s,x)) - u^0(t,s,x)}{v_x^g(t,s,x+p^g(t,s,x))}.$$

6.3 Main results

We recall from [90] the following notations. For any function $f(s,x,y)$, we define the change of variable:

$$\hat{f}(t,s,z,\xi) := f\big(t,s,z - \varepsilon\xi \cdot \mathbf{1}_d - \mathbf{y}^g(t,s,z) \cdot \mathbf{1}_d, \varepsilon\xi + \mathbf{y}^g(t,s,z)\big).$$

We then define

$$\bar{u}^{\varepsilon,g}(t,s,x,y) \;:=\; \frac{v^g(t,s,z) - v^{\varepsilon,g}(t,s,x,y)}{\varepsilon^2},\; s \in \mathbb{R}^d_+,\;\; (x,y) \in K_\varepsilon, \tag{6.3.1}$$

and its relaxed semi-limits:

$$u^{g,*}(t,s,x,y) := \overline{\lim_{(\varepsilon,t',s',x',y')\longrightarrow(0,t,s,x,y)}} \bar{u}^{\varepsilon,g}(t',s',x',y'),$$

$$u^g_*(t,s,x,y) := \underline{\lim_{(\varepsilon,t',s',x',y')\longrightarrow(0,t,s,x,y)}} \bar{u}^{\varepsilon,g}(t',s',x',y').$$

Finally, we introduce:

$$u^{\varepsilon,g}(t,s,x,y) := \bar{u}^{\varepsilon,g}(t,s,x,y) - \varepsilon^2 w^g(t,s,z,\xi),\; s \in \mathbb{R}^d_+,\; (x,y) \in K_\varepsilon.$$

6.3.1 Assumptions

In all the following, we consider payoff functions g and functions r, μ and σ such that the following four assumptions hold.

Assumption 6.3.1 (Smoothness of $\mathbf{y}^g$, v^g, $\mathbf{y}^0$ and v^0)**.** *For $\vartheta = 0$ or g, we have*

(i) *The map $v^\vartheta(t,s,z)$ is $C^{1,2,2}$ in $[0,T)\times(0,+\infty)^{d+1}$ and $C^{0,0,0}$ in $[0,T]\times(0,+\infty)^{d+1}$. Moreover, for any $(t,s)\in[0,T]\times(0,+\infty)^d$, the map $z\longmapsto v^\vartheta(t,s,z)$ is C^1 in $(0,+\infty)$ and we have*

$$v^\vartheta_z(t,s,z) > 0,\;\; (t,s,z) \in [0,T]\times(0,+\infty)^{d+1},$$

and

$$\left|v^\vartheta_{zz}\right|(t,s,z) \le \frac{C(s,z)}{(T-t)^{1-\mu}},\;\; (t,s,z) \in [0,T)\times(0,+\infty)^{d+1},$$

for some continuous function C and some $\mu \in (0,1]$.

(ii) *The map $\mathbf{y}^\vartheta(t,s,z)$ is $C^{1,2,2}$ in $[0,T)\times(0,+\infty)^{d+1}$ and $C^{0,0,0}$ in $[0,T]\times(0,+\infty)^{d+1}$. Moreover, for any $(t,s)\in[0,T]\times(0,+\infty)^d$, the map $z\longmapsto \mathbf{y}^\vartheta(t,s,z)$ is C^1 in $(0,+\infty)$ and there exist some constants $(c_0,c_1,\eta)\in(0,+\infty)\times(0,+\infty)\times(0,1]$ such that for any $(t,s,z)\in[0,T]\times(0,+\infty)^{d+1}$*

$$\mathbf{y}^{\vartheta,i}_z(t,s,z) > 0,\;\; c_0 \le \mathbf{y}^\vartheta_z(t,s,z)\cdot\mathbf{1}_d \text{ and } \left[\alpha^\vartheta(\alpha^\vartheta)^{\mathrm{T}}\right](t,s,z) \ge c_1 I_d,\;\; 1\le i\le d,$$

and for any $(t,s,z)\in[0,T)\times(0,+\infty)^{d+1}$

$$\left[\left|\mathbf{y}^\vartheta_t\right| + \left|\mathbf{y}^\vartheta_s\right| + \left|\mathbf{y}^\vartheta_{zz}\right| + \left|\mathbf{y}^\vartheta_{sz}\right| + \left|\mathbf{y}^\vartheta_{ss}\right|\right](t,s,z) \le \frac{C(s,z)}{(T-t)^{1-\eta}},$$

for some continuous function C.

Remark 6.3.1. *It can be readily checked that if it happens that $\mathbf{y}^\vartheta$ does not depend on z, then even though Assumption 6.3.1(ii) does not hold (since $\mathbf{y}^\vartheta_z = 0$), all our subsequent proofs still go through. It will be important for us later on when we treat the case of exponential utility in Section 6.4.1.*

Remark 6.3.2. *We assumed here that the first-order derivatives of v^ϑ and $\mathbf{y}^\vartheta$ with respect to z are well defined at T, unlike the other derivatives which may not exist at T. This basically due to the so-called remainder estimate that we obtain in Lemma 6.5.4, since these terms are the only ones which appear in conjunction with $\tilde{U}^1$ and its derivatives. We may have let them explode at* *time T with a certain speed, but we would then have needed to control the growth at infinity of $\tilde{U}^1$ and its derivatives. The above assumptions being already technical, we refrained from doing so, but we insist on the fact that in particular examples, our general conditions may be readily improved simply by looking at the remainder estimate obtained and using it in the proof of the viscosity subsolution property at the boundary in Section 6.5.4.*

We now state an assumption on the regularity of the solution of the first corrector equation with respect to the parameters (t, s, z).

Assumption 6.3.2 (First corrector equation: regularity on the parameters)**.** *For $\vartheta = 0$ or g, the set $\mathcal{O}_0^\vartheta(t, s, z)$ (see Proposition 6.2.1) as well as $a^\vartheta(t, s, z)$ and $\rho^*(t, s, z)$ are continuous in $(t, s, z) \in [0, T) \times (0, +\infty)^{d+1}$. Moreover, both w^ϑ and $\tilde{w}^\vartheta(\cdot, \xi) := w^\vartheta(\cdot, \xi) - w^\vartheta(\cdot, \eta^\vartheta(\cdot)\rho^*(\cdot))$ are $C^{1,2,2}$ in $[0, T) \times (0, +\infty)^{d+1}$ and satisfy for any $(t, s, z, \xi) \in [0, T) \times (0, +\infty)^{d+1} \times \mathbb{R}^d$*

$$(|\varpi_t| + |\varpi_s| + |\varpi_{ss}| + |\varpi_z| + |\varpi_{sz}| + |\varpi_{zz}|)\,(t, s, z, \xi) \le C(t, s, z)\,(1 + |\xi|) \tag{6.3.2}$$

$$(|\varpi_\xi| + |\varpi_{s\xi}| + |\varpi_{z\xi}|)\,(t, s, z, \xi) \le C(t, s, z), \tag{6.3.3}$$

for $\varpi = w^\vartheta$ or $\tilde{w}^\vartheta$ and for some continuous function $C(t, s, z)$.

The above assumption can be readily verified in dimension $d = 1$ for which the functions w^g and a^g are given explicitly in terms of the Merton value function and its derivatives. However, it would be a very difficult task to verify it in the general framework considered here. Our intention is simply to state directly what are the kind of regularity we must assume to recover the expansions, and then these can be checked on particular examples. For further reference, we also insist on the fact that by definition, the function $\tilde{w}^g$ is non-negative.

A fundamental step in any homogenization proof is to show that the correctors are uniformly locally bounded. In our context, this means that we need to show that $\bar{u}^{\varepsilon,g}$ is locally uniformly bounded. Since by definition it is a positive quantity, we only need an upper bound. We put this as an assumption.

Assumption 6.3.3 (Local bound of $\bar{u}^g$)**.** *The family of functions $\bar{u}^{\varepsilon,g}$ is locally uniformly bounded from above.*

Of course, one could argue that we are avoiding a major problem here. However, exactly as for the previous assumption, given the level of generality we are working with, verifying that it holds for generic models goes beyond the scope of this paper. Let us instead explain how one can expect to recover it on particular examples, by sketching the general approach. First of all, notice that we can without loss of generality only consider the case where all the $\lambda^{i,j} = +\infty$ for $1 \le i, j \le d$ and where $\kappa = 0$ (*i.e.* no consumption allowed). Indeed, the corresponding value function is clearly smaller than $v^{\varepsilon,g}$, and thus the corresponding $\bar{u}^{\varepsilon,g}$ is greater than the one for which we want to find an upper bound. Hence, it suffices to consider this case.

The first step is then to construct a regular viscosity sub-solution to the dynamic programming equation (6.2.11) which has the form

$$V^{\varepsilon,K}(t, s, z, \xi) := v^g(t, s, z) - \varepsilon^2 K u^g(t, s, z) - \varepsilon^4 w^g(t, s, z, \xi),$$

where u^g and w^g are the solutions to the corrector equations and where K is a large constant.

Indeed, using comparison for (6.2.11), this would then imply that $V^{\varepsilon,K} \le v^{\varepsilon,g}$, from which we can immediately deduce the required upper bound for $\bar{u}^{\varepsilon,g}$.

Of course, the first problem would then be that we are not sure that u^g and w^g are smooth. For u^g, since it is the solution to a linear PDE, this could be readily checked as soon as we have enough regularity on a^g. However, for w^g as soon as $d > 1$, we cannot reasonably expect it to be more than C^1 in ξ, since variational inequalities with gradient constraints in dimension greater than 1 are generally not C^2. Nonetheless, this issue can easily be solved by replacing w^g by a function $W^g(t,s,z,\xi) = \widehat{w}^g(\eta^g(t,s,z)\rho)/(\eta^g(t,s,z)v_z^g(t,s,z))$, where $\widehat{w}^g$ is the first component of the solution $(\widehat{w}^g, \widehat{a}^g)$ to the following equation,

$$\max_{0\leq i\leq d} \max \left\{ -\frac{c_1^*|\rho|^2}{2} - \frac{c_2^*}{2}\Delta\widehat{w}(\rho) + \widehat{a} \ , \ -\widehat{\lambda}^i + \frac{\partial \widehat{w}}{\partial \rho_i}(\rho) \ , \ -\tilde{\lambda}^i - \frac{\partial \widehat{w}}{\partial \rho_i}(\rho) \right\} = 0, \tag{6.3.4}$$

with the normalization $\widehat{w}(0) = 0$ and given positive constants $c_1^*, c_2^*, \widehat{\lambda}^i, \tilde{\lambda}^i$ such that for some constant $M > 0$

$$c_1^* I_d \geq \sigma\sigma^T, \qquad c_2^* I_d \geq \bar{\alpha}^g(\bar{\alpha}^T)^g, \qquad \widehat{\lambda}^i = \tilde{\lambda}^i = M\overline{\lambda} := M \max_{(i,j)\in\mathcal{I}} \lambda^{i,j}.$$

Then, as shown in [81], the unique solution $\widehat{w}^g$ is given as, $\widehat{w}^g(\rho) = \sum_{i=1}^d \ \widetilde{w}_i^g(\rho_i)$, where $\widetilde{w}_i^g$ is the explicit solution of the corresponding one-dimensional problem, which is known explicitly and is C^2.

To prove the viscosity sub solution property, one can then argue exactly as in the proof of Lemma 3.1 in [81]. This proof is made under assumptions ensuring homotheticity in z of the functions appearing, but the general approach will be valid in other cases as well, albeit with more complicated computations. For instance, in the case where U_2 is an exponential utility, and the frictionless market is the Black-Scholes model, the dependence in (t,s,z) of all quantities involved is known explicitly (see the formulas in Section 6.4 for details). Basically, one has to check that for M large enough the gradient constraints

$$\Lambda_{i,0}^\varepsilon \cdot (V_x^{\varepsilon,K}, V_{y_i}^{\varepsilon,K}) \leq 0, \quad \text{holds whenever} \quad M\overline{\lambda} + \frac{\partial \widehat{w}}{\partial \rho_i}(\rho) \leq 0.$$

Similarly,

$$\Lambda_{0,i}^\varepsilon \cdot (V_x^{\varepsilon,K}, V_{y_i}^{\varepsilon,K}) \leq 0, \quad \text{holds whenever} \quad -M\overline{\lambda} + \frac{\partial \widehat{w}}{\partial \rho_i}(\rho) \leq 0.$$

Then, by choosing K large enough, one has to show that the diffusion operator in (6.2.11) applied to $V^{\varepsilon,K}$ is a non-positive quantity, which would then give the desired result.

Since we assumed that $\bar{u}^{\varepsilon,g}$ is uniformly locally bounded, we can define for $(t_0,s_0,x_0,y_0) \in [0,T]\times(0,\infty)^d\times\mathbb{R}\times\mathbb{R}^d$ with $x_0 + y_0\cdot\mathbf{1}_d > 0$

$$b(t_0,s_0,x_0,y_0) := \sup\left\{u^{\varepsilon,g}(t,s,x,y) \ : (t,s,x,y) \in B_{r_0}(t_0,s_0,x_0,y_0), \ \varepsilon \in (0,\varepsilon_0]\right\}. \tag{6.3.5}$$

Then using the continuity of w^g, there exists $r_0(t_0,s_0,x_0,y_0) > 0$ and $\varepsilon_0(t_0,s_0,x_0,y_0) > 0$ such that $b(t_0,s_0,x_0,y_0) < \infty$.

Our final assumption ensures that we have a comparison theorem for the second corrector equation.

Assumption 6.3.4 (Second corrector equation: comparison)**.** *For $\vartheta = g$ or 0, there exists a set of functions $\mathcal{C}$ which contains $u^{*,\vartheta}$ and u_*^ϑ and such that $u_1 \geq u_2$ on $[0,T]\times(0,+\infty)^{d+1}$, whenever u_1 (resp. u_2) is a l.s.c. (resp. u.s.c.) viscosity super-solution (resp. sub-solution) of (6.2.15) in $\mathcal{C}$.*

Once again, we will not attempt to verify this assumption. Nonetheless, we insist on the fact that the PDE (6.2.15) is linear, so that we can reasonably expect that a comparison theorem on the class of functions with polynomial growth will hold as soon as a^g itself has polynomial growth.

6.3.2 The results

Theorem 6.3.1 (Convergence of $u^{\varepsilon,g}$). *Under assumptions 6.3.1, 6.3.2, 6.3.3 and 6.3.4, the sequence $\bar{u}^{\varepsilon,g}$ converges locally uniformly to a function u^g depending only on (t,s,z) and which is the unique viscosity solution of* (6.2.15).

The proof is relegated to Section 6.5.

Theorem 6.3.2 (Expansion of the utility indifference price). *Under assumptions 6.3.1, 6.3.2, 6.3.3 and 6.3.4, we have for all (t,s,x):*

$$\frac{p^{\varepsilon,g}(t,s,x)-p^g(t,s,x)-\varepsilon^2 h^g(t,s,x)}{\varepsilon^2}\longrightarrow 1, \text{ locally uniformly as } \varepsilon\longrightarrow 0,$$

where

$$h^g(t,s,x):=\frac{u^g(t,s,x+p^g(t,s,x))-u^0(t,s,x)}{v_z^g(t,s,x+p^g(t,s,x))}.$$

Proof.

Step 1: We first show that p^g is continuous in (t,s,z). Indeed since v^g and v^0 are C^2 and v^g is partially strictly concave w.r.t. z, we have that $v_z^g>0$. The continuity of p^g follows easily. Indeed assume on the contrary that there exists (t_0,s_0,x_0), $\varepsilon>0$ and a sequence $(t_n,s_n,x_n)\longrightarrow(t_0,s_0,x_0)$ such that

$$|p^g(t_n,s_n,x_n)-p^g(t_0,s_0,x_0)|>\varepsilon.$$

W.l.o.g., we can assume that $p^g(t_n,s_n,x_n)>p^g(t_0,s_0,x_0)+\varepsilon$. Then we have by definition of p^g and the fact that v^g is increasing w.r.t. the x-variable that for all $n\geq 0$

$$v^g(t_n,s_n,x_n+p^g(t_0,s_0,x_0)+\varepsilon)<v^0(t_n,s_n,x_n), \tag{6.3.6}$$

and

$$v^g(t_0,s_0,x_0+p^g(t_0,s_0,x_0)+\varepsilon)>v^0(t_0,s_0,x_0). \tag{6.3.7}$$

Then by continuity of v^g and v^0 we obtain from (6.3.6) that

$$v^g(t_0,s_0,x_0+p^g(t_0,s_0,x_0)+\varepsilon)\geq v^0(t_0,s_0,x_0),$$

which contradicts (6.3.7).

Step 2: Let $(t_0,s_0,x_0)\in[0,T]\times\mathbb{R}_+^d\times\mathbb{R}_+$. We consider $r>0$ such that on $\bar{B}_r(t_0,s_0,x_0)$, the quantity $u^{\varepsilon,g}(t,s,x+p^g(t,s,x))$ and $u^{\varepsilon,0}(t,s,x)$ converge uniformly to respectively $u^g(t,s,x+p^g(t,s,x))$ and $u^0(t,s,x)$. Notice that the existence of r is guaranteed by the result of Theorem 6.3.1, together with the fact that p^g is continuous. We use the notations p^g (resp. h^g) for $p^g(t,s,x)$

(resp. $h^g(t,s,x)$) for simplicity. For any $\delta \in (-1,1)$, we have uniformly on $\bar{B}_r(t_0,s_0,x_0)$:

$$\begin{aligned} v^{\varepsilon,g}(t,s,x+p^g+\varepsilon^2 h^g+\varepsilon^2\delta,0) &= v^g(t,s,x+p^g)+\varepsilon^2\left(h^g+\delta\right)v_x^g(t,s,x+p^g) \\ &\quad -\varepsilon^2 u(t,s,x+p^g)+\circ(\varepsilon^2) \\ &= v^0(t,s,x)+\varepsilon^2\left(h^g+\delta\right)v_z^g(t,s,x+p^g) \\ &\quad -\varepsilon^2 u(t,s,x+p^g)+\circ(\varepsilon^2) \\ &= v^0(t,s,x)-\varepsilon^2 u(t,s,x)+\varepsilon^2\delta v_z^g(t,s,x+p^g)+\circ(\varepsilon^2) \\ &= v^{\varepsilon,0}(t,s,x)+\varepsilon^2\delta v_z^g(t,s,x+p^g)+\circ(\varepsilon^2). \end{aligned}$$

Hence, the following holds uniformly on $\bar{B}_r(t_0,s_0,x_0)$

$$\frac{v^{\varepsilon,g}(t,s,x+p^g+\varepsilon^2 h^g+\varepsilon^2\delta)-v^{\varepsilon,0}(t,s,x)}{\varepsilon^2}=\delta v_z^g(t,s,x+p^g)+\circ(1). \tag{6.3.8}$$

We now claim that for any $\delta>0$, there is some $\varepsilon^*(\delta)$ such that we have for $\varepsilon\le\varepsilon^*(\delta)$ that on $\bar{B}_r(t_0,s_0,x_0)$:

$$p^g(t,s,x)+\varepsilon^2 h^g(t,s,x)-\varepsilon^2\delta\le p^{\varepsilon,g}(t,s,x)\le p^g(t,s,x)+\varepsilon^2 h^g(t,s,x)+\varepsilon^2\delta,$$

which implies directly the required result.

Remains to prove the claim. Assume on the contrary that we have $\delta>0$ and $(\varepsilon_n,t_n,s_n,x_n)$, where for all $n\ge 0$, $(t_n,s_n,x_n)\in\bar{B}_r(t_0,s_0,x_0)$ and $\varepsilon_n\longrightarrow 0$, such that, for example, for all $n\ge 0$

$$p^{\varepsilon_n,g}(t_n,s_n,x_n)>p^g(t_n,s_n,x_n)+\varepsilon_n^2 h^g(t_n,s_n,x_n)+\varepsilon_n^2\delta.$$

Then by definition of $p^{\varepsilon,g}$, we have that

$$v^{\varepsilon_n,g}(t_n,s_n,x_n+p^g+\varepsilon_n^2 h^g+\varepsilon_n^2\delta)-v^{\varepsilon_n,0}(t_n,s_n,x_n)\le 0,$$

which contradicts (6.3.8) for n large enough, i.e. ε_n small enough. The other inequality can be shown similarly.

□

6.4 Examples and applications

In this Section we will specialize our discussion to a simpler case, in order to highlight how our method allows not only to recover existing results but to go beyond them. Throughout the Section, we place ourselves in dimension $d=1$, and assume a Black-Scholes dynamic for the risky asset

$$\frac{dS_t}{S_t}=\mu dt+\sigma dW_t,\ \mathbb{P}-a.s.$$

Moreover, the interest rate r is assumed to be constant and the investor aims at solving the following versions of the stochastic control problems (6.2.1) and (6.2.3)

$$v^{\varepsilon,g}(t,s,x,y):=\sup_{(c,L)\in\Theta^\varepsilon(t,s,x,y)}\mathbb{E}_t\left[\int_t^T\kappa U_1(c_\xi)d\xi+U_2\left(\ell^\varepsilon\left(X_T^{t,s,x,y},Y_T^{t,s,x,y}\right)-g(S_T)\right)\right], \tag{6.4.1}$$

and

$$v^g(t,s,z):=\sup_{(c,\theta)\in\Theta^0(t,s,z)}\mathbb{E}_t\left[\int_t^T\kappa U_1(c_\xi)d\xi+U_2\left(Z_T^{\theta,t,s,z}-g(S_T)\right)\right], \tag{6.4.2}$$

corresponding to the case $k = 0$ in (6.2.1).

We will now show, for particular choices of utility functions, that we can calculate almost explicitly all the quantities involved in the asymptotic expansion (6.2.18), as well as check that all our assumptions hold under certain explicit conditions. For further reference and use, we recall that in this setting, the so-called Black-Scholes price of the claim g, denoted by V^g is given by

$$V^g(t,s) := \mathbb{E}_t^{\mathbb{Q}}\left[e^{-r(T-t)}g(S_T^{t,s})\right],$$

where $\mathbb{Q}$ is the so-called risk neutral pricing measure defined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(-\frac{\mu - r}{\sigma}W_T\right).$$

Moreover, it is a well known result that as soon as g has polynomial growth at infinity, V^g is also the unique (classical) solution to the following PDE

$$-V_t^g - rsV_s^g - \frac{1}{2}\sigma^2 s^2 V_{ss}^g + rV^g = 0,\ (t,s) \in [0,T) \times \mathbb{R}_+, \qquad V^g(T,\cdot) = g(\cdot). \tag{6.4.3}$$

Finally, we recall that in dimension 1, [90] gave an explicit solution to the second corrector equation (see their Section 4.1),

$$w^g(t,s,z,\xi) = \begin{cases} \left(v_z^g\left[-\frac{\sigma^2}{12(\eta^g)^2(\alpha^g)^2}\xi^4 + \frac{\sigma^2}{2(\eta^g)^2(\alpha^g)^2}\xi_0^2\xi^2 + \frac{\lambda^{1,0}-\lambda^{0,1}}{2}\xi\right]\right)(t,s,z), \text{ if } |\xi| \le \xi_0 \\ v_z^g(t,s,z)\left[-\frac{3}{16}\left(\lambda^{1,0}+\lambda^{0,1}\right)\xi_0 - \lambda^{0,1}\xi\right], \text{ if } \xi \le -\xi_0 \\ v_z^g(t,s,z)\left[-\frac{3}{16}\left(\lambda^{1,0}+\lambda^{0,1}\right)\xi_0 + \lambda^{1,0}\xi\right], \text{ if } \xi \ge \xi_0, \end{cases} \tag{6.4.4}$$

where

$$\xi_0 := \xi_0(t,s,z) := \eta^g(t,s,z)\left(\frac{3}{4}\frac{(\alpha^g)^2(t,s,z)}{(\eta^g)^2(t,s,z)\sigma^2}\left(\lambda^{1,0}+\lambda^{0,1}\right)\right)^{1/3},$$

which in turn allows us to have an explicit form for the function $a^g(t,s,z)$ in terms of the Merton value function

$$a^g(t,s,z) = \frac{\sigma^2 v_z^g(t,s,z)}{2\eta^g(t,s,z)}\xi_0^2(t,s,z), \tag{6.4.5}$$

where we remind the reader that the diffusion coefficient α is given by

$$\alpha^g(t,s,z) = \sigma\big((1-\mathbf{y}_z^g(t,s,z))\mathbf{y}^g(t,s,z) - s\mathbf{y}_s^g(t,s,z)\big).$$

6.4.1 Exponential Utility

Let us assume in this subsection that

$$U_1(x) = U_2(x) = -e^{-\gamma x},\ \text{for some } \gamma > 0.$$

6.4.2 Derivation of the expansion

We start by giving the solution to the Merton problem corresponding to $\varepsilon = 0$. In the case $\kappa = 0$ (no consumption), the solution can be found for instance in [16] (see also the references therein). The generalization to the consumption case is an easy (but lengthy) exercise, so that we omit its proof.

Proposition 6.4.1. *The value function for the stochastic control problem (6.4.2) is given for any* $(t,s,z)\in[0,T]\times\mathbb{R}_+\times\mathbb{R}_+$ *by*

$$v^g(t,s,z)=-\exp\big(-\gamma v_1(t)\,(z-V^g(t,s))+v_2(t)\big),$$

where

$$\begin{aligned}
v_1(t) &:= \frac{r}{\kappa+e^{-r(T-t)}(r-\kappa)}\\
v_2(t) &:= \frac{1}{\kappa e^{r(T-t)}+r-\kappa}\left[\frac{\mu-r}{2\sigma^2}(\kappa-r)(T-t)+\kappa\frac{\log^2\left(\kappa e^{r(T-t)}+r-\kappa\right)-\log^2(r)}{2}\right.\\
&\qquad\left.-\kappa\left(e^{r(T-t)}-1\right)\left(\frac{\mu-r}{2\sigma^2}+\log(r)-2\right)-r(T-t)e^{r(T-t)}\right)\Bigg].
\end{aligned}$$

Moreover, the optimal trading strategy and consumption are given by

$$\begin{aligned}
\mathbf{y}^g(t,s) &:= sV^g_s(t,s)+\frac{\mu-r}{\gamma\sigma^2}v_1^{-1}(t)\\
\mathbf{c}^g(t,s,z) &:= \kappa\left(-\frac{1}{\gamma}\log(v_1(t))+v_1(t)(z-V^g(t,s))+v_2(t)\right).
\end{aligned}$$

Remark 6.4.1. *It can be checked directly that when* $\kappa=0$*, the above reduces to the formula given in Remark 3.4 of [16].*

Using the above Proposition, we recover the expected result that the utility indifference price $p^g(t,s,z)$ does not depend on z in this case, and is simply given by the Black-Scholes price $V^g(t,s)$ of the contingent claim g. We refer the reader to Theorem 1 and Section 3 in [35] for further details on this general result.

Next, using (6.4.5), we deduce immediately that the function a^g is given in this case by

$$a^g(t,s,z)=-\gamma^2v_1^2(t)v^g(t,s,z)\frac{\sigma^2}{2}\left(\frac{3}{4}\left(\lambda^{0,1}+\lambda^{1,0}\right)\left(\frac{\mu-r}{\sigma^2}-\gamma v_1(t)s^2V^g_{ss}(t,s)\right)^2\right)^{\frac{2}{3}}.$$

Since the dependence of a^g in z only comes from v^g, it is natural to expect that the solution u^g to the first corrector equation in this case admits the factorized form

$$u^g(t,s,z)=:-v^g(t,s,z)\tilde{u}^g(t,s).$$

Direct calculations using the PDE (6.2.16) show that the function $\tilde{u}^g$ should then satisfy

$$\begin{cases}-\tilde{u}^g_t-rs\tilde{u}^g_s-\dfrac{1}{2}\sigma^2\tilde{u}^g_{ss}+\kappa v_1(t)\tilde{u}^g=\gamma^2v_1^2\dfrac{\sigma^2}{2}\left(\dfrac{3}{4}\left(\lambda^{0,1}+\lambda^{1,0}\right)\left(\dfrac{\mu-r}{\sigma^2}-\gamma v_1s^2V^g_{ss}\right)^2\right)^{\frac{2}{3}}\\ \tilde{u}^g(T,\cdot)=0.\end{cases}\tag{6.4.6}$$

Finally, the expansion (6.2.18) takes the form

$$p^{\varepsilon,g}(t,s)=V^g(t,s)+\frac{\varepsilon^2}{\gamma v_1(t)}\left(\tilde{u}^g(t,s)-\tilde{u}^0(t,s)\right)+\circ(\varepsilon^2),$$

which is exactly the same as the one given in Corollary 3.8 of [16] in the case $\kappa=0$.

Let us now give sufficient conditions under which all the above calculations are rigorous and under which Assumptions 6.3.1, 6.3.2 and 6.3.4 are satisfied. Concerning Assumption 6.3.3, given the length of the paper, we do not verify it here, but we would like to mention that in addition to the general sketch of proof given in Section 6.3.1, the fact that we are in dimension $d = 1$ opens up another way to prove the result by constructing an almost optimal strategy for the problem with friction. Indeed, in this case, this can be achieved by using a solution to the Skorokhod problem with reflection on the boundary $\mathcal{O}^g(t,s,z)$ which, as we will see below is actually regular (unlike when $d \geq 2$ where we know nothing about its regularity in general). This approach was used by Bouchard, Moreau and Soner in [25].

Proposition 6.4.2. *In the framework of this section, if we assume that for $\vartheta = g$ or 0* (i) *There exists a constant $c_0 > 0$ such that*

$$\left| \frac{\mu - r}{\gamma \sigma^2 v_1(t)} - s^2 V^\vartheta_{ss}(t,s) \right| \geq c_0, \ (t,s) \in [0,T) \times (0,+\infty).$$

(ii) *V^ϑ is $C^{1,4}$ in $[0,T) \times (0,+\infty)$ and continuous on $[0,T] \times (0,+\infty)$ and there exists $\eta \in (0,1]$ such that*

$$\left[\left| V^\vartheta_s \right| + \left| V^\vartheta_{ts} \right| + \left| V^\vartheta_{sss} \right| \right] (t,s) \leq \frac{C(s)}{(T-t)^{1-\eta}}, \ (t,s) \in [0,T) \times (0,+\infty),$$

and there exists $\nu \in (1/4, 1]$ such that

$$\left| V^\vartheta_{ss} \right| (t,s) \leq \frac{C(s)}{(T-t)^{1-\nu}}, \ (t,s) \in [0,T) \times (0,+\infty),$$

for some continuous function C. Then Assumptions 6.3.1, 6.3.2 and 6.3.4 are satisfied

Proof. We start with Assumption 6.3.1. First of all, it is clear that v^ϑ is $C^{1,2,2}$ in $[0,T) \times (0,+\infty)^2$ and continuous in $[0,T] \times (0,+\infty)^2$, since V^ϑ is and v^1 and v^2 are C^∞ on $[0,T]$. Moreover, we have that v^ϑ is actually C^∞ in $z \in (0,+\infty)$ for every $(t,s) \in [0,T] \times (0,+\infty)$. In particular, v^ϑ_{zz} is bounded on $(t,s,z) \in [0,T] \times (0,+\infty)^2$ and

$$v^\vartheta_z(t,s,z) = \gamma v^1(t) \exp\left(-\gamma v_1(t) \left(z - V^\vartheta(t,s) \right) + v_2(t) \right) > 0, \ (t,s,z) \in [0,T] \times (0,+\infty)^2.$$

Next, notice that $\mathbf{y}^\vartheta$ does not depend on z (see Remark 6.3.1) and that

$$\alpha^\vartheta(t,s) = \sigma \left(\frac{\mu - r}{\gamma \sigma^2 v_1(t)} - s^2 V^\vartheta_{ss}(t,s) \right),$$

so that we clearly have $(\alpha^\vartheta)^2(t,s) \geq c_1$ for some $c_1 > 0$. Then, the estimates on the derivatives of $\mathbf{y}^\vartheta$ are immediate consequences of the assumed estimates on the derivatives of V^ϑ. Hence Assumption 6.3.1 is satisfied.

Let us now look at Assumption 6.3.2. We have in this framework

$$a^\vartheta(t,s,z) = \frac{\sigma^2 v^\vartheta_z(t,s,z)}{2} \gamma v^1(t) \left(\frac{3\left(\lambda^{1,0} + \lambda^{0,1}\right)}{4\gamma^2 (v^1)^2(t)} \left(\frac{\mu - r}{\gamma \sigma^2 v_1(t)} - s^2 V^\vartheta_{ss}(t,s) \right)^2 \right)^{2/3},$$

which implies that a^ϑ is continuous in $[0,T) \times (0,+\infty)^2$. Then, using (6.4.4), the required estimates and regularity in (t,s,z) for w^ϑ and $\mathcal{O}^\vartheta$ are direct consequences of the fact that V^ϑ is $C^{1,4}$ in $[0,T) \times (0,+\infty)^2$ and that v^ϑ is C^∞ in $z \in (0,+\infty)$ for every $(t,s) \in [0,T] \times (0,+\infty)$. Next, $\rho^*(t,s,z)$ is a solution to a cubic equation so that it has the same regularity as its coefficients, which then implies that $\tilde{w}^\vartheta$ also satisfies the required regularity and estimates. Hence Assumption 6.3.2 is satisfied.

Finally, concerning Assumption 6.3.4, as mentioned before, obtaining a comparison theorem for viscosity solutions with polynomial growth is a classical result. Moreover, in this particular case, it is easy to check using Feynman-Kac formula that the PDE (6.4.6) has a unique smooth solution which admits the following probabilistic representation

$$u^g(t,s,z) = -Bv^g(t,s,z)\mathbb{E}^{\mathbb{Q}}\left[\int_t^T v_1^2(u)e^{-\kappa\int_t^u v_1(\xi)d\xi}\left(\frac{\mu-r}{\sigma^2} - \gamma v_1(u)(S_u^{t,s})^2 V_{ss}^g(u,S_u^{t,s})\right)^{\frac{4}{3}} du\right],$$

where

$$B := \frac{\gamma^2\sigma^2}{2}\left(\frac{3}{4}(\lambda^{0,1}+\lambda^{1,0})\right)^{\frac{2}{3}}.$$

Notice that when $g=0$, this can actually be further simplified to obtain

$$u^0(t,s,z) = -Bv^0(t,s,z)\left(\frac{\mu-r}{\sigma^2}\right)^{\frac{4}{3}}\int_t^T v_1^2(u)e^{-\kappa\int_t^u v_1(\xi)d\xi}du.$$

Of course, for all this to be meaningful, the above expectation should be finite, which is once again an implicit assumption on the payoff g. It is easy to show that a sufficient condition for this to be true is that there exist some $\beta \in (0,3/4)$ and $k \geq 0$ such that

$$\left|V_{ss}^g\right|(t,s) \leq \frac{C(s)}{(T-t)^\beta}.$$

□

6.4.2.1 Discussion on the Assumptions in this setting

As we have seen above, the fact that the diffusion coefficient α^g should not be equal to 0 translates directly in our setting into

$$\frac{\mu-r}{\gamma\sigma^2}v_1^{-1}(t) - s^2V_{ss}^g(t,s) \neq 0.$$

This is an implicit assumption on the payoff g, which may not be satisfied if $s^2V_{ss}^g$ can become arbitrarily big, which would be the case for a Call option for instance, since this quantity explodes to $+\infty$ as t goes to T, when we are at the money forward (*i.e.* $s = Ke^{-r(T-t)}$). This condition also naturally appears in the recent work of Bouchard, Moreau and Soner [25], and under a stronger form in [16] (see Assumption 3.2). However, we would like to insist on the fact that in our approach, we do not need to assume regularity on the payoff g directly (except continuity) but on its Black-Scholes price which is much more regular in general. Hence, our assumptions are less restrictive than the ones in [25] and [16]. We would also like to point out that the quantity of interest here is then $s^2V_{ss}^g(t,s)$, which is the so-called *activity rate* of *portfolio Gamma* which plays a central role in the formal asymptotics obtained by Kallsen and Muhle-Karke in [63, 64].

Notice also that a Call option does not satisfy the assumption that the third order derivative of its Black-Scholes price does not explode at time T at a speed strictly less than $(T-t)^{-1}$, however, we believe that this condition can be improved by maybe using other test functions in our proof of the sub solution property at the boundary in Section 6.5.4. This, as pointed out in [16], leads to conjecture that the expansion should also hold in the case of Call options. We leave this problem for future research. However, if one considers a Digital option $g(s) = \mathbf{1}_{s\geq K}$, then one can readily check that the function $\tilde{u}^g$ becomes infinite, which shows that the expansion cannot hold in this case, and that the corresponding first order term, (if it exists) goes to 0 more slowly than ε^2. We emphasize that the exact same phenomenon was already highlighted by Possamaï, Soner and Touzi [80] in a market where the frictions came from the absence of infinite liquidity. Moreover, the techniques of proof used in this paper to show the expansion for Call option can certainly be adapted in our setting.

6.4.3 Power utility

Let us assume in this subsection that

$$U_1(x) = U_2(x) = \frac{x^\gamma}{\gamma}, \text{ for some } \gamma \in (0,1).$$

We would like to point out immediately that such a utility function is not covered by the results we prove in this paper. Indeed, we assumed that U_2 had to be defined on the whole real line. This assumption is a very important one, and without it, the expansion result can actually be completely wrong. Indeed, let us assume, in the framework considered in this section, that we want to compute $p^{\varepsilon,g}(0,s,0,0)$ for a Call option when $\kappa = 0$. Then, since any negative final wealth for the investor leads to a utility equal to $-\infty$, the investor has to use a trading strategy which guarantees him a final wealth which is greater, $\mathbb{P} - a.s.$, than $g(S_T^{0,s})$. In other words, the investor has to at least super-replicate the Call option. However, as recalled in the introduction, it is a well known result that the only super-replicating strategy in a market with proportional transaction costs is the trivial buy-and-hold strategy, whose cost is $(1+\varepsilon^3\lambda^{0,1})s$. However, in the frictionless market, since it is complete, the corresponding super-replication price is actually the Black-Scholes price $V^g(0,s)$. Hence, the utility indifference price $p^{\varepsilon,g}$ does not converge to p^g as ε goes to 0, and our general result is therefore false in this case.

However, it could be possible that, denoting by $V^{\varepsilon,g,SR}(t,s)$ the super-replication price at time t of the claim g when $S_t = s$, our results remain valid when we let the investor start with an initial wealth x which is "sufficiently" above $V^{\varepsilon,g,SR}$ [1]. Of course, this would then mean that the corresponding HJB equation for $v^{\varepsilon,g}$ has to be complemented with a new boundary condition, which is *a priori*, far from trivial to deduce, since it is roughly linked to the problem of utility maximization under stochastic target constraints, which was considered by Bouchard, Elie and Imbert in [22]. We acknowledge that this remains a conjecture, but since the calculations for the expansions are easy enough, we give them anyway. We will not try to verify all our assumptions in this setting, since this can basically be done by assuming sufficient regularity on V^g and V^0, exactly as in the exponential utility case.

We start by giving the solution to the Merton problem corresponding to $\varepsilon = 0$. In the case $\kappa = 0$ and $g = 0$ (no consumption and no claim), the problem was already solved in Merton's seminal paper [71]. The generalization to the consumption case and claim case is an easy (but lengthy) exercise, so that we omit its proof.

Proposition 6.4.3. *The value function for the stochastic control problem* (6.4.2) *is given by for any* $(t,s,z) \in [0,T] \times \mathbb{R}_+ \times [V^g(t,s), +\infty)$

$$v^g(t,s,z) = \left(e^{\frac{\gamma}{1-\gamma}A(T-t)} \left(1 + \kappa\frac{1-\gamma}{A\gamma}\right) - \kappa\frac{1-\gamma}{A\gamma}\right)^{1-\gamma} \frac{(z - V^g(t,s))^\gamma}{\gamma},$$

where

$$A := r + \frac{(\mu - r)^2}{2\sigma^2(1-\gamma)}.$$

[1] We point out that in the recent paper by Bouchard, Moreau and Soner [25], this problem was already pointed out, and they therefore only considered the null payoff when dealing with power utilities. But their framework did not allow for a utility indifference price depending on the initial wealth of the investor, unlike in our setting, so much so that their result do not necessarily imply that our conjecture is false.

Moreover, the optimal trading strategy and consumption are given by

$$\mathbf{y}^g(t,s,z) := sV^g_s(t,s) + \frac{\mu - r}{(1-\gamma)\sigma^2}(z - V^g(t,s))$$

$$\mathbf{c}^g(t,s,z) := \kappa \frac{z - V^g(t,s)}{e^{\frac{\gamma}{1-\gamma}A(T-t)}\left(1 + \kappa\frac{1-\gamma}{A\gamma}\right) - \kappa\frac{1-\gamma}{A\gamma}}.$$

The above Proposition shows immediately that the utility indifference price $p^g(t,s,z)$ is again equal to the Black-Scholes price $V^g(t,s)$ of the contingent claim g.

Next, using (6.4.5), we deduce immediately that the function a^g is given in this case by

$$a^g(t,s,z) = \tilde{A}(t)(z - V^g(t,s))^{\gamma - \frac{10}{3}}\left(\tilde{B}(z - V^g(t,s)) - s^2 V^g_{ss}(t,s)\right)^{\frac{4}{3}},$$

where

$$\tilde{A} := (1-\gamma)^{\frac{7}{3}}\frac{\sigma^2}{2}\left(\frac{3}{4}(\lambda^{0,1} + \lambda^{1,0})\right)^{\frac{2}{3}}\left(e^{\frac{\gamma}{1-\gamma}A(T-t)}\left(1 + \kappa\frac{1-\gamma}{A\gamma}\right) - \kappa\frac{1-\gamma}{A\gamma}\right)^{1-\gamma}$$

$$\tilde{B} := \frac{\mu - r}{\sigma^2(1-\gamma)}\left(1 - \frac{\mu - r}{\sigma^2(1-\gamma)}\right).$$

In this case, there does not seem to be any clear factorization for the function u^g when $g \neq 0$. However, exactly as in the exponential utility case, u^g admits the following probabilistic representation (provided that all the quantities are well-defined, which is once again an implicit assumption on the payoff g)

$$u^g(t,s,z) = \mathbb{E}\left[\int_t^T a^g\left(u, S^{t,s}_u, Z^{t,s,z,\mathbf{y}^g,\mathbf{c}^g}_u\right)du\right].$$

In the case where $g = 0$, it can be verified directly that $u^0(t,s,z) = \tilde{u}^0(t)z^{\gamma-2}$, where

$$\tilde{u}^0(t) := \tilde{B}^{\frac{4}{3}}\int_t^T e^{\int_t^u \zeta(r)dr}\tilde{A}(u)du,$$

where

$$\zeta(t) := r(\gamma - 2) + \frac{(\mu - r)^2(2-\gamma)(1+\gamma)}{2\sigma^2(1-\gamma)^2} + \kappa\frac{2-\gamma}{e^{\frac{\gamma}{1-\gamma}A(T-t)}\left(1 + \kappa\frac{1-\gamma}{A\gamma}\right) - \kappa\frac{1-\gamma}{A\gamma}}.$$

Finally, the asymptotic expansion (6.2.18) should then take the following form

$$p^\varepsilon(t,s,z) = V^g(t,s) + \frac{\varepsilon^2 z^{\gamma-1}(u^g(t,s,z + V^g(t,s)) - u^0(t,s,z))}{\left(e^{\frac{\gamma}{1-\gamma}A(T-t)}\left(1 + \kappa\frac{1-\gamma}{A\gamma}\right) - \kappa\frac{1-\gamma}{A\gamma}\right)^{1-\gamma}} + \circ(\varepsilon^2).$$

6.5 Proof of Theorem 6.3.1

We would like to point out immediately to the reader that several of the proofs below (especially the proofs of the viscosity sub and super-solution properties inside the domain) are very close to the ones given in [81]. Nonetheless, they also provide some corrections to small gaps that we identified in [81], and are made under assumptions which are a little bit more general (in particular, we no longer require the upper bound for $\mathbf{y}^g$ in their Assumption 3.1) and we therefore think that they can be of interest. However, the proof of the viscosity sub-solution property at the boundary is new, and the derivation of the remainder estimate has to be done with a lot more precision than in their case, because of the possible explosions at the boundary.

6.5.1 First properties and derivatives estimates

Denote by L the upper bound of the set C, we define:

$$\bar{\lambda} := \max_{(i,j)\in\mathcal{I}} \lambda^{i,j}, \quad \underline{\lambda} := \min_{(i,j)\in\mathcal{I}} \lambda^{i,j}.$$

We would like to mention that for notational simplicity, we state all the results of this section for $u^{\varepsilon,g}$ and $v^{\varepsilon,g}$, but they of course still hold true for $u^{\varepsilon,0}$ and $v^{\varepsilon,0}$. That being said, we have first the following easy result, whose proof can be found in [81] for instance

Lemma 6.5.1. *Let* $(t,s,x,y) \in [0,T) \times (0,\infty)^d \times K_\varepsilon$. *Then*

$$u^{\varepsilon,g}(t,s,x,y) \geq -\varepsilon L v_z^g(t,s,z) \left| y - \mathbf{y}^g(t,s,z) \right|,$$

so that under Assumption 6.3.3 we obtain that:

$$0 \leq u_*^g(t,s,x,y) \leq u^{g*}(t,s,x,y) < \infty.$$

We start by a technical Lemma, which will be used in the proof of Lemma 6.5.3. The proof follows exactly the same arguments as the ones given in [81], with some modifications due to the fact that, unlike in [81], we do not assume any upper bound for $\mathbf{y}_z^g$. We therefore provide them for the sake of completeness.

Lemma 6.5.2. *Under assumption 6.3.1, 6.3.2 and 6.3.3, the gradient of* $\hat{v}^{\varepsilon,g}$ *exists almost everywhere and there exists a universal constant* A *such that for all* $(t,s,z,\xi) \in [0,T)\times(0,+\infty)^{d+1}\times\mathbb{R}^d$, *we can find some* $\varepsilon^* := \varepsilon^*(t,s,z) > 0$ *such that*

$$|\hat{v}_\xi^{\varepsilon,g}|(t,s,z,\xi) \leq A\varepsilon^4 |\hat{v}^{\varepsilon,g}|(t,s,z,\xi), \text{ for } \varepsilon \leq \varepsilon^*, \text{ and } \hat{v}_z^{\varepsilon,g}(t,s,z,\xi) \leq \gamma^\varepsilon(t,s,z,\xi), \ \forall \varepsilon > 0,$$

where

$$\begin{aligned}
\gamma^\varepsilon(t,s,z,\xi) := {} & D(t,s,z) v_z^g(t,s,z-\varepsilon) + \varepsilon \left| 1 - \mathbf{y}_z^g(t,s,z) \cdot \mathbf{1}_d \right| u^{\varepsilon,g}(t,s,x-\varepsilon,y) \\
& + \varepsilon \sum_{i=1}^d \mathbf{y}_z^{g,i}(t,s,z) u^{\varepsilon,g}(t,s,x,y-\varepsilon e_i) + \varepsilon^3 C(t,s,z-\varepsilon)\left(1+|\xi|\right) D(t,s,z) \\
& + \varepsilon^3 C(t,s,z-\varepsilon) \left| 1 - \mathbf{y}_z^g(t,s,z)\cdot \mathbf{1}_d \right| \frac{\left| \mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon) \right|}{\varepsilon} \\
& + \varepsilon^3 C(t,s,z-\varepsilon) \sum_{i=1}^d \mathbf{y}_z^{g,i}(t,s,z) \frac{\left| \mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon) - \varepsilon e_i \right|}{\varepsilon},
\end{aligned}$$

where C *is the function appearing in Assumption 6.3.2 and where*

$$D(t,s,z) := \left| 1 - \mathbf{y}_z^g(t,s,z)\cdot\mathbf{1}_d \right| + \mathbf{y}_z^g(t,s,z)\cdot\mathbf{1}_d.$$

Proof. *Step 1: first estimate.* By Theorem 6.2.1, we have for all $1 \leq i \leq d$ in the viscosity sense that

$$\Lambda_{i,0}^\varepsilon \cdot (v_x^{\varepsilon,g}, v_y^{\varepsilon,g}) \geq 0 \text{ and } \Lambda_{0,i}^\varepsilon \cdot (v_x^{\varepsilon,g}, v_y^{\varepsilon,g}) \geq 0.$$

We deduce immediately from the definition of $\hat{v}^{\varepsilon,g}$ that for all $1 \leq i \leq d$

$$\frac{\varepsilon^4 \lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}} \hat{v}_z^{\varepsilon,g}(t,s,z,\xi) - \frac{\varepsilon^3\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}} \mathbf{y}_z^g(t,s,z)\cdot\hat{v}_\xi^{\varepsilon,g}(t,s,z,\xi) + \hat{v}_{\xi^i}^{\varepsilon,g}(t,s,z,\xi) \geq 0, \tag{6.5.1}$$

and

$$\varepsilon^4\lambda^{0,i}\hat{v}_z^{\varepsilon,g}(t,s,z,\xi)-\varepsilon^3\lambda^{0,i}\mathbf{y}_z^g(t,s,z)\cdot\hat{v}_\xi^{\varepsilon,g}(t,s,z,\xi)-\hat{v}_{\xi^i}^{\varepsilon,g}(t,s,z,\xi)\geq 0. \tag{6.5.2}$$

Now since we have by Assumption 6.3.1 that for all $1\leq i\leq d$, $\mathbf{y}_z^{g,i}(t,s,z)>0$, we have, by multiplying (6.5.1) by $\mathbf{y}_z^{g,i}$and summing for all $1\leq i\leq d$ that in the viscosity sense

$$\left(1-\varepsilon^3\sum_{i=1}^d\frac{\mathbf{y}_z^{g,i}(t,s,z)\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}\right)\mathbf{y}_z^g(t,s,z).\hat{v}_\xi^{\varepsilon,g}(t,s,z,\xi)\geq-\varepsilon^4\sum_{i=1}^d\frac{\lambda^{i,0}\mathbf{y}_z^{g,i}(t,s,z)}{1+\varepsilon^3\lambda^{i,0}}\hat{v}_z^{\varepsilon,g}(t,s,z,\xi). \tag{6.5.3}$$

Now, we know that there exists a $\varepsilon^*(t,s,z)$ such that

$$1-\varepsilon^3\sum_{i=1}^d\frac{\mathbf{y}_z^{g,i}(t,s,z)\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}\geq 0,\ \text{for } \varepsilon\leq\varepsilon^*(t,s,z), \tag{6.5.4}$$

so that in the viscosity sense, we have for $\varepsilon\leq\varepsilon^*(t,s,z)$

$$\mathbf{y}_z^g(t,s,z).\hat{v}_\xi^{\varepsilon,g}(t,s,z,\xi)\geq-\frac{\sum_{i=1}^d\frac{\lambda^{i,0}\mathbf{y}_z^{g,i}(t,s,z)}{1+\varepsilon^3\lambda^{i,0}}}{1-\varepsilon^3\sum_{i=1}^d\frac{\mathbf{y}_z^{g,i}(t,s,z)\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}}\varepsilon^4\hat{v}_z^{\varepsilon,g}(t,s,z,\xi). \tag{6.5.5}$$

Using this estimate in (6.5.2), we deduce

$$\begin{aligned}\hat{v}_{\xi^i}^{\varepsilon,g}(t,s,z,\xi)&\leq\lambda^{0,i}\varepsilon^4\left(1+\frac{\varepsilon^3\sum_{i=1}^d\frac{\lambda^{i,0}\mathbf{y}_z^{g,i}(t,s,z)}{1+\varepsilon^3\lambda^{i,0}}}{1-\varepsilon^3\sum_{i=1}^d\frac{\mathbf{y}_z^{g,i}(t,s,z)\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}}\right)\hat{v}_z^{\varepsilon,g}(t,s,z,\xi)\\&\leq\varepsilon^4\overline{\lambda}\left(1+\frac{1-c_0}{c_0}\right)\hat{v}_z^{\varepsilon,g}(t,s,z,\xi)\\&=\varepsilon^4\frac{\overline{\lambda}}{c_0}\hat{v}_z^{\varepsilon,g}(t,s,z,\xi),\ \text{for } \varepsilon\leq\varepsilon^*(t,s,z),\end{aligned}$$

where we used Assumption 6.3.1 and the fact that the map $x\longmapsto x/(1-x)$ is non-decreasing. Similarly, using (6.5.5) in (6.5.1) leads to

$$\begin{aligned}\hat{v}_{\xi^i}^{\varepsilon,g}(t,s,z,\xi)&\geq-\frac{\varepsilon^4\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}\left(1+\frac{\varepsilon^3\sum_{i=1}^d\frac{\lambda^{i,0}\mathbf{y}_z^{g,i}(t,s,z)}{1+\varepsilon^3\lambda^{i,0}}}{1-\varepsilon^3\sum_{i=1}^d\frac{\mathbf{y}_z^{g,i}(t,s,z)\lambda^{i,0}}{1+\varepsilon^3\lambda^{i,0}}}\right)\hat{v}_z^{\varepsilon,g}(t,s,z,\xi)\\&\geq-\varepsilon^4\frac{\overline{\lambda}}{c_0}\hat{v}_z^{\varepsilon,g}(t,s,z,\xi),\ \text{for } \varepsilon\leq\varepsilon^*(t,s,z).\end{aligned}$$

Now since by the concavity of $v^{\varepsilon,g}$ in (x,y), we know that its gradient exists almost everywhere and since by Assumption 6.3.1, $\mathbf{y}^g$ is twice continuously differentiable, we have that $\hat{v}_z^{\varepsilon,g}$ exists almost everywhere and we have for $\varepsilon\leq\varepsilon^*(t,s,z)$

$$\left|\hat{v}_\xi^{\varepsilon,g}\right|\leq A\varepsilon^4\hat{v}_z^{\varepsilon,g},\ \text{where } A:=\frac{\overline{\lambda}}{c_0}. \tag{6.5.6}$$

Step 2: second estimate. We now estimate $\hat{v}_z^{\varepsilon,g}$. We first notice that, remembering that $v^{\varepsilon,g}$ is clearly non-decreasing with respect to x and to y^i for $i=1,\dots,d$

$$\begin{aligned}\hat{v}_z^{\varepsilon,g}(t,s,z,\xi)&=(1-\mathbf{y}_z^g(t,s,z)\cdot\mathbf{1}_d)v_x^{\varepsilon,g}(t,s,x,y)+\mathbf{y}_z^g(t,s,z)\cdot v_y^{\varepsilon,g}(t,s,x,y)\\&\leq\left|1-\mathbf{y}_z^g(t,s,z)\cdot\mathbf{1}_d\right|v_x^{\varepsilon,g}(t,s,x,y)+\sum_{i=1}^d\mathbf{y}_z^{g,i}(t,s,z)v_{y^i}^{\varepsilon,g}(t,s,x,y).\end{aligned} \tag{6.5.7}$$

Then by concavity of $v^{\varepsilon,g}$ in x and of v^g in z and since $v^{\varepsilon,g} \le v^g$, we have:

$$\begin{aligned} v_x^{\varepsilon,g}(t,s,x,y) &\le \frac{v^g(t,s,x,y) - v^{\varepsilon,g}(t,s,x-\varepsilon,y)}{\varepsilon} \\ &\le v_z^g(t,s,z-\varepsilon) + \frac{v^g(t,s,z-\varepsilon) - v^{\varepsilon,g}(t,s,x-\varepsilon,y)}{\varepsilon}. \end{aligned}$$

Then by definition of $u^{\varepsilon,g}$, we have:

$$v_x^{\varepsilon,g}(t,s,x,y) \le v_z^g(t,s,z-\varepsilon) + \varepsilon\left(u^{\varepsilon,g}(t,s,x-\varepsilon,y) + \varepsilon^2 w^g(t,s,z-\varepsilon,\xi_\varepsilon)\right),$$

where

$$\xi_\varepsilon := \frac{y - \mathbf{y}^g(t,s,z-\varepsilon}{\varepsilon} = \xi + \frac{\mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon)}{\varepsilon}.$$

Then we recall from the estimate of w^g given by Assumption 6.3.2 that:

$$\begin{aligned} |w^g(t,s,z-\varepsilon,\xi_\varepsilon)| &\le C(t,s,z-\varepsilon)(1+|\xi_\varepsilon|) \\ &\le C(t,s,z)\left(1+|\xi| + \varepsilon^{-1}|\mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon)|\right), \end{aligned}$$

for some continuous positive function C. Hence, we deduce

$$\begin{aligned} v_x^{\varepsilon,g}(t,s,x,y) &\le v_z^g(t,s,z-\varepsilon) + \varepsilon u^{\varepsilon,g}(t,s,x-\varepsilon,y) \\ &\quad + \varepsilon^3 C(t,s,z-\varepsilon)\left(1+|\xi| + \frac{|\mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon)|}{\varepsilon}\right). \end{aligned}$$

Now following the same arguments, we also have for all $1 \le i \le d$:

$$v_{y^i}^{\varepsilon,g}(t,s,x,y) \le v_z^g(t,s,z-\varepsilon) + \frac{v^g(t,s,z-\varepsilon) - v^{\varepsilon,g}(t,s,x,y-\varepsilon e_i)}{\varepsilon},$$

and

$$\begin{aligned} v_{y^i}^{\varepsilon,g}(t,s,x,y) &\le v_z^g(t,s,z-\varepsilon) + \varepsilon u^{\varepsilon,g}(t,s,x,y-\varepsilon e_i) \\ &\quad + \varepsilon^3 C(t,s,z-\varepsilon)\left(1+|\xi| + \frac{|\mathbf{y}^g(t,s,z) - \mathbf{y}^g(t,s,z-\varepsilon) - \varepsilon e_i|}{\varepsilon}\right). \end{aligned}$$

Plugging the estimates for $v_x^{\varepsilon,g}$ and $v_{y^i}^{\varepsilon,g}$ in (6.5.7), we obtain immediately

$$\hat{v}_z^{\varepsilon,g}(t,s,z,\xi) \le \gamma^\varepsilon(t,s,z,\xi).$$

□

Lemma 6.5.3. *Under assumption 6.3.1, 6.3.2 and 6.3.3, $u^{g,*}$ and u_*^g are only functions of (t,s,z). Furthermore, we have:*

$$\begin{aligned} u_*^g(t,s,z) &= \underline{\lim}_{(\varepsilon,t',s',z')\longrightarrow(0,t,s,z)} \bar{u}^{\varepsilon,g}\left(t',s',z' - \mathbf{y}^g(t',s',z')\cdot\mathbf{1}_d, \mathbf{y}^g(t',s',z')\right) \\ u^{g,*}(t,s,z) &= \overline{\lim}_{(\varepsilon,t',s',z')\longrightarrow(0,t,s,z)} \bar{u}^{\varepsilon,g}\left(t',s',z' - \mathbf{y}^g(t',s',z')\cdot\mathbf{1}_d, \mathbf{y}^g(t',s',z')\right). \end{aligned}$$

Proof. We split the proof in two parts:

Step 1. We first show the lemma for $t \in [0,T)$. The result is a consequence of the gradient constraints in (6.2.3) thanks to which we obtained the estimates of Lemma 6.5.2. By definition of $\hat{u}^{\varepsilon,g}$, we have that for every (t,s,z,ξ), there exists $\varepsilon^*(t,s,z)$ such that for any $\varepsilon \le \varepsilon^*(t,s,z)$

$$|\hat{u}_\xi^{\varepsilon,g}|(t,s,z,\xi) \le \varepsilon^{-2}|\hat{v}_\xi^{\varepsilon,g}(t,s,z,\xi)| + \varepsilon^2|w_\xi^g(t,s,z,\xi)| \le \varepsilon^2\left(A\gamma^\varepsilon(t,s,z,\xi) + C(t,s,z)\right),$$

where the second inequality (and the constant A) comes from Lemma 6.5.2. Then for any $\xi_0 \in \mathbb{R}^d$ such that $1 - \sum_{i=1}^d \xi_0^i = 0$, we have for $\varepsilon \leq \varepsilon^*(t,s,z)$

$$\left| \left(\sum_{i=1}^d \xi_0^i e_i - e_0 \right) \cdot (u_x^{\varepsilon,g}, u_y^{\varepsilon,g}) \right| = \frac{1}{\varepsilon} \left| \xi_0 \cdot \hat{u}_\xi^{\varepsilon,g} \right| \leq \varepsilon |\xi_0| \left(A\gamma^\varepsilon(t,s,z,\xi) + C(t,s,z) \right).$$

Next, we remind the reader that $u^{\varepsilon,g}$ is locally bounded. Fix therefore some (t_0, s_0, x_0, y_0), a $r_0 > 0$ small such that $u^{\varepsilon,g}$, and the continuous functions γ^ε and C are bounded uniformly on $B_{r_0}(t_0, s_0, x_0, y_0)$.

Now recall also that $\varepsilon^*(t_0, s_0, z_0)$ is defined (see (6.5.4)) such that

$$1 - \varepsilon^3 \sum_{i=1}^d \frac{\mathbf{y}_z^{g,i}(t_0, s_0, z_0)\lambda^{i,0}}{1 + \varepsilon^3 \lambda^{i,0}} \geq 0, \text{ for } \varepsilon \leq \varepsilon^*(t_0, s_0, z_0).$$

However, since the left-hand side above goes to 1 as ε goes to 0 and since it is continuous in (t,s,z), then reducing ε if necessary, this inequality will also hold for any $(t,s,x,y) \in B_{r_0}(t_0, s_0, x_0, y_0)$

Therefore, we can find a constant K independent of ε and large enough such that for all $\xi_0 \in \mathbb{R}^d$ such that $1 - \sum_{i=1}^d \xi_0^i = 0$, the maps

$$t \longmapsto u^{\varepsilon,g}(t, s, x - t, y + t\xi^0) + \varepsilon K t \text{ and } t \longmapsto -u^{\varepsilon,g}(t, s, x - t, y + t\xi^0) + \varepsilon K t,$$

are non-decreasing. Then by definition, we obtain that $u^{g,*}$ and u_*^g are independent of the ξ-variable.

Step 2. The previous proof does not hold at $t = T$ since the gradient constraints verified by w^g may not hold at T, since w^g may not be defined there. By definition of the relaxed semi limit, we have for (T, s_0, x_0, y_0)

$$u_*^g(T, s_0, x_0, y_0) = l_1(s_0, x_0, y_0) \wedge l_2(s_0, x_0, y_0)$$

where

$$\begin{aligned} l_1(s_0, x_0, y_0) &:= \liminf_{(\varepsilon,s,x,y) \longrightarrow (0,s_0,x_0,y_0)} \bar{u}^{\varepsilon,g}(T, s, x, y) \\ l_2(s_0, x_0, y_0) &:= \liminf_{(\varepsilon,t,s,x,y) \longrightarrow (0,T,s_0,x_0,y_0), t \neq T} \bar{u}^{\varepsilon,g}(t, s, x, y). \end{aligned}$$

We consider separately these two terms. Freezing the variable $t = T$, we obtain that

$$\begin{aligned} l_1(s_0, x_0, y_0) &= \liminf_{(\varepsilon,s,x,y) \longrightarrow (0,s_0,x_0,y_0)} \bar{u}^{\varepsilon,g}(t, s, x, y) \\ &= \lim_{(\varepsilon,s,x,y) \longrightarrow (0,s_0,x_0,y_0)} \frac{U_2(z - g(s)) - U_2\left(\ell^\varepsilon(x,y) - g(s)\right)}{\varepsilon^2} = 0. \end{aligned}$$

Then by Step 1, we know that

$$\begin{aligned} l_2(s_0, x_0, y_0) &= \liminf_{(\varepsilon,t,s,x,y) \longrightarrow (0,T,s_0,x_0,y_0), t \neq T} \bar{u}^{\varepsilon,g}(t, s, x, y) \\ &= \liminf_{(t,s,x,y) \longrightarrow (T,s_0,x_0,y_0), t \neq T} u_*^g(t, s, x, y) \\ &= \liminf_{(t,s,z) \longrightarrow (T,s_0,z_0), t \neq T} u_*^g(t, s, z), \end{aligned}$$

so that we obtain the required result for u_*^g. The same arguments leads to the result for $u^*(T, s, x, y)$, so we omit them.

□

6.5.2 The remainder estimate

We now isolate an important estimate introduced in [90] and [81], which will be of crucial importance in the proofs of sub and super-solutions properties below. Following the seminal work of Evans [?] on the perturbed test function technique, it will be convenient for us to consider, for a test function ϕ of the second corrector equation (6.2.15), potential test functions ψ for (6.2.11) of the form

$$v^g(t,s,z) - \varepsilon^2 \tilde{\phi}^\varepsilon(t,s,z) - \varepsilon^4 \varpi(t,s,z,\xi),$$

where $\tilde{\phi}^\varepsilon$ will be a perturbation of ϕ, and ϖ a smooth function close to w^g. The aim of the following Lemma is to provide a detailed estimate of the remainder terms in the expansion of the parabolic part of (6.2.11) when applied to such a function, which was formally obtained in Section 6.2.5. We emphasize here that unlike in [81], we want to have a very precise estimate, in particular when it comes to the derivatives of $\mathbf{y}^g$ which appear. Indeed, as mentioned in Remark 6.2.1, these derivatives may explode at time T, which will cause some difficulties when proving viscosity solution properties at the terminal time in the subsequent sections. Such a problem was not present in [81] which considered only the infinite horizon case.

Lemma 6.5.4. *Let $\Psi^\varepsilon(t,s,x,y) := v^g(t,s,z) - \varepsilon^2\phi(t,s,z) - \varepsilon^4\hat{\varpi}(t,s,z,\xi)$, with smooth ϕ and such that ϖ satisfies the same estimates as w^g in Assumption 6.3.2. We then have*

$$\begin{aligned}\mathcal{I}(\Psi^\varepsilon)(t,s,x,y) &:= \left(k(t,s)\Psi^\varepsilon - \mathcal{L}\Psi^\varepsilon - \widetilde{U}^1(\Psi^\varepsilon_x)\right)(t,s,x,y)\\ &= \varepsilon^2\Big[-\frac{1}{2}\left|\sigma(t,s)\xi\right|^2 v^g_{zz}(t,s,z) + \frac{1}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T(t,s,z)\hat{\varpi}_{\xi\xi}(t,s,z,\xi)\right]\\ &\qquad - \mathcal{A}^g\phi(t,s,z) + \mathcal{R}^\varepsilon(\phi,\hat{\varpi})(t,s,z,\xi)\Big],\end{aligned}$$

where $\mathcal{R}^\varepsilon(\phi,\hat{\varpi}) := \mathcal{R}^\varepsilon(\phi,\hat{\varpi})(t,s,z,\xi)$ verifies

$$\begin{aligned}|\mathcal{R}^\varepsilon(\phi,\hat{\varpi})| \leq &\Big[K\left(\varepsilon|\xi| + \varepsilon^2|\xi|^2\right)\left(1+|\mathbf{y}^g|+|\mathbf{y}^g|^2\right)\left(1+|\mathbf{y}^g_t|+|\mathbf{y}^g_s|+|\mathbf{y}^g_z|+|\mathbf{y}^g_{zz}|+|\mathbf{y}^g_{sz}|+|\mathbf{y}^g_{ss}|\right)\\ &\times\left(1+|\phi_z|+|\phi_{zz}|+|\phi_{sz}|+\varepsilon^4\mathfrak{R}^\varepsilon(\varpi)\right) + \varepsilon^4K\left(1+\zeta^\varepsilon(t,s,z,\xi)\right)\left(1+|\mathbf{y}^g_z|+|\mathbf{y}^g_z|^2\right)\\ &\times\left(1+|\hat{\varpi}_z|+|\phi_z|^2+\varepsilon^4|\hat{\varpi}_z|^2+\varepsilon^2|\hat{\varpi}_\xi|^2+\varepsilon^{-1}|\hat{\varpi}_\xi|\right)\Big](t,s,z),\end{aligned}$$

where K is a positive continuous function which depends only on r, μ, σ, $\tilde{U}_1$ and v^g and where the quantities $\mathfrak{R}^\varepsilon(\varpi)$ and ζ^ε are defined in the proof.

Proof. For notational simplicity, we will omit the dependence of the coefficients in the parameters. We have:

$$\begin{aligned}\mathcal{I}(\Psi^\varepsilon)(t,s,x,y) &= k\Psi^\varepsilon - \mathcal{L}\Psi^\varepsilon - \kappa\tilde{U}_1\left(\Psi^\varepsilon_x\right)\\ &= \left[kv^g - \mathcal{L}v^g - \kappa\tilde{U}_1\left(v^g_x\right)\right] - \varepsilon^2\left(k\phi - \mathcal{L}\phi\right) - \varepsilon^4\left(k\hat{\varpi} - \mathcal{L}\hat{\varpi}\right)\\ &\quad + \kappa\left[\tilde{U}_1\left(v^g_x\right) - \tilde{U}_1\left(\Psi^\varepsilon_x\right)\right].\end{aligned}$$

We now consider separately every term. We first recall from Section 6.2.5 that:

$$kv^g - \mathcal{L}v^g - \kappa\tilde{U}_1\left(v^g_x\right) = -\frac{1}{2}\left|\sigma^T(\mathbf{y}^g - y)\right|^2 v^g_{zz}$$

Similarly to the previous calculations, we have

$$\begin{aligned}k\phi - \mathcal{L}\phi &= k\phi - \mathcal{L}^0\phi - rz\phi_z - y\cdot\left[\left(\mu - r\mathbf{1}_d\right)\phi_z + \sigma\sigma^T\mathbf{D}_{sz}\phi\right] - \frac{1}{2}|\sigma^T y|^2\phi_{zz}\\ &= \mathcal{A}^g\phi - \kappa\mathbf{c}^g\phi_z + (\mathbf{y}^g - y)\cdot\left(\mu - r\mathbf{1}_d\right)\phi_z - \frac{1}{2}\phi_{zz}\left(|\sigma^T y|^2 - |\sigma^T\mathbf{y}^g|^2\right)\\ &\quad - (y - \mathbf{y}^g)\cdot\sigma\sigma^T\mathbf{D}_{sz}\phi.\end{aligned}$$

Define then

$$\mathcal{R}_\phi := (\mathbf{y}^g - y) \cdot (\mu - r\mathbf{1}_d)\, \phi_z - \frac{1}{2}\phi_{zz}\left(|\sigma^T y|^2 - |\sigma^T \mathbf{y}^g|^2\right) - (y - \mathbf{y}^g)\cdot \sigma\sigma^T \mathbf{D}_{sz}\phi.$$

We clearly have that

$$\begin{aligned}|\mathcal{R}_\phi(t,s,z,\xi)| &\le \varepsilon\left(|\xi|\,|\mu - r\mathbf{1}_d|\,|\phi_z| + \frac{|\sigma^2|}{2}\left(2|\mathbf{y}^g||\xi| + \varepsilon|\xi|^2\right)|\phi_{zz}| + |\sigma|^2|\xi||\mathbf{D}_{sz}\phi|\right)\\ &\le K_1(t,s,z)\left(\varepsilon|\xi| + \varepsilon^2|\xi|^2\right)(1+|\mathbf{y}^g|)\left(1 + |\phi_z| + |\phi_{zz}| + |\phi_{sz}|\right),\end{aligned}$$

where K_1 is a positive continuous function which depends only on σ, r and μ. The third term is more tedious. We sum up the calculations here

$$|\varpi_y| \le |\hat{\varpi}_z| + \frac{1}{\varepsilon}\left(1 + |\mathbf{y}^g_z|\right)|\hat{\varpi}_\xi|, \quad |\varpi_x| \le |\hat{\varpi}_z| + \frac{1}{\varepsilon}|\mathbf{y}^g_z||\hat{\varpi}_\xi|,$$

$$|\varpi_s| \le |\hat{\varpi}_s| + \frac{1}{\varepsilon}|\mathbf{y}^g_s||\hat{\varpi}_\xi|, \quad |\varpi_t| \le |\hat{\varpi}_t| + \frac{1}{\varepsilon}|\mathbf{y}^g_t||\hat{\varpi}_\xi|,$$

$$|\varpi_{yy}| \le \mathrm{Const}\left(|\hat{\varpi}_{zz}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_z|)\,|\hat{\varpi}_{z\xi}| + |\mathbf{y}^g_{zz}||\hat{\varpi}_\xi|\right)\right),$$

$$|\varpi_{ys}| \le \mathrm{Const}\left(|\hat{\varpi}_{sz}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_z|)\,|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{sz}||\hat{\varpi}_\xi| + |\mathbf{y}^g_s||\hat{\varpi}_{z\xi}|\right)\right),$$

$$|\varpi_{ss}| \le \mathrm{Const}\left(|\hat{\varpi}_{ss}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_s|)\,|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{ss}||\hat{\varpi}_\xi|\right)\right).$$

We deduce that

$$\mathrm{Tr}\left[\sigma\sigma^T\left(\mathbf{D}_{yy} + \mathbf{D}_{ss} + 2\mathbf{D}_{sy}\right)\varpi\right] = \frac{1}{\varepsilon^2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\hat{\varpi}_{\xi\xi}\right] + \widetilde{\mathcal{R}}_2(\varpi),$$

where

$$\begin{aligned}\widetilde{\mathcal{R}}_2(\varpi) \le\ & \mathrm{Const}\left(|\mathbf{y}^g|^2 + \varepsilon^2|\xi^2|\right)\left(|\hat{\varpi}_{zz}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_z|)\,|\hat{\varpi}_{z\xi}| + |\mathbf{y}^g_{zz}||\hat{\varpi}_\xi|\right)\right)\\ &+ \mathrm{Const}|s|\left(|\mathbf{y}^g| + \varepsilon|\xi|\right)\left(|\hat{\varpi}_{sz}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_z|)\,|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{sz}||\hat{\varpi}_\xi| + |\mathbf{y}^g_s||\hat{\varpi}_{z\xi}|\right)\right)\\ &+ \mathrm{Const}|s|^2\left(|\hat{\varpi}_{ss}| + \frac{1}{\varepsilon}\left((1+|\mathbf{y}^g_s|)\,|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{ss}||\hat{\varpi}_\xi|\right)\right).\end{aligned}$$

We therefore deduce

$$-\varepsilon^4\left(k\varpi - \mathcal{L}\varpi\right)(t,s,x,y) = \frac{\varepsilon^2}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\hat{\varpi}_{\xi\xi}\right] + \varepsilon^4\mathcal{R}_2(\varpi),$$

where

$$\begin{aligned}
\mathcal{R}_2(\varpi) \le\ & |k||\hat{\varpi}| + |\hat{\varpi}_t| + \frac{1}{\varepsilon}|\mathbf{y}^g_t||\hat{\varpi}_\xi| + |\mu||s|\left(|\hat{\varpi}_s| + \frac{1}{\varepsilon}|\mathbf{y}^g_s||\hat{\varpi}_\xi|\right) \\
& + |r|\left(|z| + |\mathbf{y}^g| + \varepsilon|\xi|\right)\left(|\hat{\varpi}_z| + \frac{1}{\varepsilon}|\mathbf{y}^g_z||\hat{\varpi}_\xi|\right) \\
& + |\mu|\left(|\mathbf{y}^g| + \varepsilon|\xi|\right)\left(|\hat{\varpi}_z| + \frac{1}{\varepsilon}\left(1 + |\mathbf{y}^g_z|\right)|\hat{\varpi}_\xi|\right) \\
& + \text{Const}\left(|\mathbf{y}^g|^2 + \varepsilon^2|\xi|^2\right)\left(|\hat{\varpi}_{zz}| + \frac{1}{\varepsilon}\left(\left(1 + |\mathbf{y}^g_z|\right)|\hat{\varpi}_{z\xi}| + |\mathbf{y}^g_{zz}||\hat{\varpi}_\xi|\right)\right) \\
& + \text{Const}|s|\left(|\mathbf{y}^g| + \varepsilon|\xi|\right)\left(|\hat{\varpi}_{sz}| + \frac{1}{\varepsilon}\left(\left(1 + |\mathbf{y}^g_z|\right)|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{sz}||\hat{\varpi}_\xi| + |\mathbf{y}^g_s||\hat{\varpi}_{z\xi}|\right)\right) \\
& + \text{Const}|s|^2\left(|\hat{\varpi}_{ss}| + \frac{1}{\varepsilon}\left(\left(1 + |\mathbf{y}^g_s|\right)|\hat{\varpi}_{s\xi}| + |\mathbf{y}^g_{ss}||\hat{\varpi}_\xi|\right)\right) \\
\le\ & K_2(t,s,z)\left(1 + \varepsilon|\xi| + \varepsilon^2|\xi|^2\right)\left(1 + |\mathbf{y}^g| + |\mathbf{y}^g|^2\right) \\
& \times \left(1 + |\mathbf{y}^g_t| + |\mathbf{y}^g_s| + |\mathbf{y}^g_z| + |\mathbf{y}^g_{zz}| + |\mathbf{y}^g_{sz}| + |\mathbf{y}^g_{ss}|\right)\mathfrak{R}^\varepsilon(\varpi),
\end{aligned}$$

where $K_2(t,s,z)$ is a positive continuous function which depends only on r and μ and where

$$\mathfrak{R}^\varepsilon(\varpi) := |\hat{\varpi}| + |\hat{\varpi}_t| + |\hat{\varpi}_s| + |\hat{\varpi}_z| + |\hat{\varpi}_{zz}| + |\hat{\varpi}_{sz}| + |\hat{\varpi}_{ss}| + \varepsilon^{-1}\left(|\hat{\varpi}_\xi| + |\hat{\varpi}_{z\xi}| + |\hat{\varpi}_{s\xi}|\right).$$

Summarizing up, we have that the remainder estimate $\mathcal{R}^\varepsilon(\phi,\hat{\varpi})$ denoted $\mathcal{R}$ for short here verifies:

$$|\mathcal{R}^\varepsilon(\phi,\hat{\varpi})|(t,s,z,\xi) \le \left[|\mathcal{R}_\phi| + |\mathcal{R}_2(\varpi)| + |\tilde{U}_1(v^g_x) - \tilde{U}_1(\psi^\varepsilon_x) + \varepsilon^2\mathbf{c}^g\phi_z|\right](t,s,z,\xi).$$

We now estimate the last term involved. Recall that $\mathbf{c}^g = -\tilde{U}_1'(v^g_z(t,s,z))$, we have, omitting the dependence in (t,s,z,ξ)

$$\mathcal{R}_{\tilde{U}_1} := \tilde{U}_1(\Psi^\varepsilon_x) - \tilde{U}_1(v^g_x) - \varepsilon^2\mathbf{c}^g\phi_z = \tilde{U}_1(\Psi^\varepsilon_x) - \tilde{U}_1(v^g_x) + (\Psi^\varepsilon_x - v^g_z)\tilde{U}_1'(v^g_z) + r_1,$$

where $r_1 := \varepsilon^4\varpi_x\tilde{U}_1'(v^g_z)$ verifies

$$|r_1| \le \varepsilon^4\left|\tilde{U}_1'(v^g_z)\right|\left(|\hat{\varpi}_z| + \frac{1}{\varepsilon}|\mathbf{y}^g_z||\hat{\varpi}_\xi|\right).$$

Then we have, using that $\tilde{U}$ is concave

$$\left|\tilde{U}_1(\Psi^\varepsilon_x) - \tilde{U}_1(v^g_x) + (\Psi^\varepsilon_x - v^g_z)\tilde{U}_1'(v^g_z)\right| \le |\Psi^\varepsilon_x - v^g_x|\left|\tilde{U}_1'(\Psi^\varepsilon_x) - \tilde{U}_1'(v^g_x)\right|.$$

Then since $\tilde{U}_1$ is C^2 we have that

$$\left|\tilde{U}_1'(\Psi^\varepsilon_x) - \tilde{U}_1'(v^g_x)\right| \le |\Psi^\varepsilon_x - v^g_x|\,\zeta^\varepsilon(t,s,z,\xi),\ \text{where } \zeta^\varepsilon(t,s,z,\xi) := \sup_{m\in K^\varepsilon(t,s,z,\xi)}\left|\tilde{U}_1''(m)\right|,$$

where $K^\varepsilon(t,s,z,\xi) := \text{Supp}(\tilde{U}_1) \cap \{m \in \mathbb{R},\ |m| \le \mathfrak{H}^\varepsilon(t,s,z,\xi)\}$, with

$$\mathfrak{H}^\varepsilon(t,s,z,\xi) := |v^g_z| + \varepsilon^2\left(|\phi_z| + \varepsilon^2\left(|\hat{\varpi}_z| + \frac{1}{\varepsilon}|\mathbf{y}^g_z||\hat{\varpi}_\xi|\right)\right).$$

Then we obtain

$$\begin{aligned}
\left|\mathcal{R}_{\tilde{U}_1}\right| \le\ & \varepsilon^4\left|\tilde{U}_1'(v^g_z)\right|\left(|\hat{\varpi}_z| + \frac{1}{\varepsilon}|\mathbf{y}^g_z||\hat{\varpi}_\xi|\right) \\
& + \varepsilon^4\text{Const}\left(|\phi_z|^2 + \varepsilon^4\left(|\hat{\varpi}_z|^2 + \frac{1}{\varepsilon^2}|\mathbf{y}^g_z|^2|\hat{\varpi}_\xi|^2\right)\right)\zeta^\varepsilon(t,s,z,\xi) \\
\le\ & \varepsilon^4 K_3(t,s,z)\left(1 + \zeta^\varepsilon(t,s,z,\xi)\right)\left(1 + |\mathbf{y}^g_z| + |\mathbf{y}^g_z|^2\right) \\
& \times\left(1 + |\hat{\varpi}_z| + |\phi_z|^2 + \varepsilon^4|\hat{\varpi}_z|^2 + \varepsilon^2|\hat{\varpi}_\xi|^2 + \frac{1}{\varepsilon}|\hat{\varpi}_\xi|\right),
\end{aligned}$$

where K_3 is a positive continuous function which depends only on $\tilde{U}_1$ and v^g. Finally, we can conclude that

$$\begin{aligned}|\mathcal{R}| \leq\ & K(t,s,z)\left(\varepsilon|\xi| + \varepsilon^2|\xi|^2\right)\left(1 + |\mathbf{y}^g| + |\mathbf{y}^g|^2\right)\\ &\times \left(1 + |\mathbf{y}^g_t| + |\mathbf{y}^g_s| + |\mathbf{y}^g_z| + |\mathbf{y}^g_{zz}| + |\mathbf{y}^g_{sz}| + |\mathbf{y}^g_{ss}|\right)\left(1 + |\phi_z| + |\phi_{zz}| + |\phi_{sz}| + \varepsilon^4\mathfrak{R}^\varepsilon(\varpi)\right)\\ &+ \varepsilon^4 K(t,s,z)\left(1 + \zeta^\varepsilon(t,s,z,\xi)\right)\left(1 + |\mathbf{y}^g_z| + |\mathbf{y}^g_z|^2\right)\\ &\times \left(1 + |\hat{\varpi}_z| + |\phi_z|^2 + \varepsilon^4|\hat{\varpi}_z|^2 + \varepsilon^2|\hat{\varpi}_\xi|^2 + \varepsilon^{-1}|\hat{\varpi}_\xi|\right),\end{aligned}$$

where K is a positive continuous function which depends only on r, μ, σ, $\tilde{U}_1$ and v^g.

□

6.5.3 Viscosity subsolution on $[0,T) \times \mathbb{R}^d \times \mathbb{R}_+$

We focus here on the interior of the domain. Consider $(t_0, s_0, z_0) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}_+$ and $\phi \in C^2([0,T) \times \mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ such that for all $(t,s,z) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}_+ \setminus \{(t_0,s_0,z_0)\}$:

$$0 = (u^{g,*} - \phi)(t_0,s_0,z_0) > (u^{g,*} - \phi)(t,s,z).$$

We want to show that $\mathcal{A}^g\phi(t_0,s_0,z_0) - a^g(t_0,s_0,z_0) \leq 0$. We separate the proof in 4 steps.

Step 1: By Lemma 6.5.3, there exists a sequence $(t^\varepsilon, s^\varepsilon, z^\varepsilon) \longrightarrow (t_0,s_0,z_0)$ when $\varepsilon \longrightarrow 0$ such that

$$\hat{u}^{\varepsilon,g}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) \underset{\varepsilon\longrightarrow 0}{\longrightarrow} u^{g,*}(t_0,s_0,z_0).$$

Then we have that $l^\varepsilon_* := \hat{u}^{\varepsilon,g}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) - \phi(t^\varepsilon, s^\varepsilon, z^\varepsilon) \longrightarrow 0$ and $(x^\varepsilon, y^\varepsilon) \longrightarrow (x_0, y_0)$ where

$$(x^\varepsilon, y^\varepsilon) := \left(z^\varepsilon - \mathbf{y}^g(t^\varepsilon, s^\varepsilon, z^\varepsilon)\cdot \mathbf{1}_d, \mathbf{y}^g(t^\varepsilon, s^\varepsilon, z^\varepsilon)\right),$$

and

$$(x_0, y_0) := \left(z_0 - \mathbf{y}^g(t_0,s_0,z_0)\cdot \mathbf{1}_d, \mathbf{y}^g(t_0,s_0,z_0)\right).$$

Now since u^ε is locally bounded from above, there exists $r_0 > 0$ and $\varepsilon_0 > 0$ depending on (t_0, s_0, x_0, y_0) such that:

$$b_* := \sup\left\{u^{\varepsilon,g}(t,s,x,y),\ (t,s,x,y) \in B_{r_0}(t_0,s_0,x_0,y_0),\ \varepsilon \in (0,\varepsilon_0]\right\} < +\infty,$$

where, reducing r_0 if necessary, the ball is strictly included in the interior of the domain, and where, reducing ε_0 if necessary, we can assume w.l.o.g. that $(t^\varepsilon, s^\varepsilon, x^\varepsilon, y^\varepsilon) \in B_{r_0}(t_0,s_0,x_0,y_0)$ for $\varepsilon \leq \varepsilon_0$.

We now build a test function from ϕ for $v^{\varepsilon,g}$ in order to apply the dynamic programming equation associated to $v^{\varepsilon,g}$. We define for every $(\varepsilon,\delta) \in (0,1]^2$ the function $\hat{\psi}^{\varepsilon,\delta}$ and the corresponding $\psi^{\varepsilon,\delta}$ by:

$$\hat{\psi}^{\varepsilon,\delta}(t,s,z,\xi) := v^g(t,s,z) - \varepsilon^2\left(l^\varepsilon_* + \phi(t,s,z) + \hat{\Phi}^\varepsilon(t,s,z,\xi)\right) - \varepsilon^4(1+\delta)\tilde{w}^g(t,s,z,\xi),$$

where $\hat{\Phi}^{\varepsilon,\delta}$ is defined by

$$\hat{\Phi}^\varepsilon(t,s,z,\xi) := c\left((t-t^\varepsilon)^4 + (s-s^\varepsilon)^4 + (z-z^\varepsilon)^4 + \varepsilon^4(\tilde{w}^g)^4(t,s,z,\xi)\right),$$

and c is a constant chosen large enough so that for $\varepsilon \leq \varepsilon_0$

$$\Phi^\varepsilon \geq 1 + b_* - \phi, \text{ on } B_{r_0}(t_0,s_0,x_0,y_0)\backslash B_{r_0/2}(t_0,s_0,x_0,y_0). \tag{6.5.8}$$

Notice that c_0 is independent of ε. The constant δ will be fixed later. We also emphasize that by assumption, w^g and $\tilde{w}^g$ are only C^1 in ξ on the whole domain.

Step 2: We now show that for ε and δ small enough, the difference $(v^{\varepsilon,g} - \psi^{\varepsilon,\delta})$ has a local minimizer in $B_0 := B_{r_0}(t_0, s_0, x_0, y_0)$. Indeed it is sufficient to show that $I^{\varepsilon,\delta}$ has a local minimizer where:

$$\begin{aligned}
I^{\varepsilon,\delta}(t,s,x,y) &:= \frac{v^{\varepsilon,g}(t,s,x,y) - \psi^{\varepsilon,\delta}(t,s,x,y)}{\varepsilon^2} \\
&= -u^{\varepsilon,g}(t,s,x,y) + l_*^\varepsilon + \phi(t,s,z) + \Phi^\varepsilon(t,s,x,y) + \varepsilon^2\delta\tilde{w}^g(t,s,z,\xi) \\
&\quad - \varepsilon^2 w^g(t,s,z,\eta^g(t,s,z)\rho^*(t,s,z)).
\end{aligned}$$

Now since w^g and $\rho^*(t,s,z)$ are continuous, $\tilde{w}^g$ is non-negative and using (6.5.8), for $\delta > 0$ small enough and $\varepsilon \le \varepsilon_0$, we have for any $(t,s,x,y) \in \partial B_0$:

$$I^{\varepsilon,\delta}(t,s,x,y) \ge -u^{\varepsilon,g}(t,s,x,y) + l_*^\varepsilon + 1 + b_* - \varepsilon^2 w^g(t,s,z,\eta^g(t,s,z)\rho^*(t,s,z)) \ge \frac{1}{2} + l_*^\varepsilon > 0,$$

for ε small enough. Now since $I^{\varepsilon,\delta}(t^\varepsilon, s^\varepsilon, x^\varepsilon, y^\varepsilon) \longrightarrow 0$ when $\varepsilon \longrightarrow 0$, this implies that for ε small enough, $I^{\varepsilon,\delta}$ has a local minimizer $(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{x}^\varepsilon, \tilde{y}^\varepsilon)$ in B_0 and we introduce:

$$\tilde{z}^\varepsilon := \tilde{x}^\varepsilon + \tilde{y}^\varepsilon \cdot \mathbf{1}_d, \text{ and } \tilde{\xi}^\varepsilon := \frac{\tilde{y}^\varepsilon - \mathbf{y}^g(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{z}^\varepsilon)}{\varepsilon}.$$

To summarize, we have:

$$\min_{B_0}(\hat{v}^{\varepsilon,\delta} - \hat{\psi}^{\varepsilon,\delta}) = (\hat{v}^{\varepsilon,\delta} - \hat{\psi}^{\varepsilon,\delta})(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{z}^\varepsilon, \tilde{\xi}^\varepsilon), \text{ with } |\tilde{t}^\varepsilon - t_0| + |\tilde{s}^\varepsilon - s_0| + |\tilde{z}^\varepsilon - z_0| \le r_0, \quad \left|\tilde{\xi}^\varepsilon\right| \le \frac{r_1}{\varepsilon},$$

for some constant r_1. Now since $\psi^{\varepsilon,\delta}$ is at least C^1, we have that by the dynamic programming equation verified by $v^{\varepsilon,g}$ that:

$$\Lambda_{i,j}^\varepsilon \cdot \left(\psi_x^{\varepsilon,\delta}, \psi_y^{\varepsilon,\delta}\right)(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{x}^\varepsilon, \tilde{y}^\varepsilon) \ge 0, \text{ for } (i,j) \in \mathcal{I}. \tag{6.5.9}$$

Step 3: Our aim in this section is to show that for ε small enough, $\psi^{\varepsilon,\delta}$ is actually C^2 in ξ. Thank to Proposition 6.2.1, it is enough to show that for ε small enough we have:

$$\tilde{\rho}^\varepsilon := \frac{\tilde{\xi}^\varepsilon}{\eta^g(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{z}^\varepsilon)} \in \mathcal{O}_0^g(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{z}^\varepsilon),$$

where $\mathcal{O}_0^g(t,s,z)$ is the open set introduced in Proposition 6.2.1. Assume on the contrary that there exists $\varepsilon_n \longrightarrow 0$ such that for n large enough $\tilde{\rho}^{\varepsilon_n} \notin \mathcal{O}_0^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n})$. Then since $\bar{w}^g$ is C^1 and thanks to (6.2.17), we have:

$$-\lambda^{i_0^n,j_0^n} v_z^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}) + \left(\tilde{w}^g_{\xi^{i_0^n}} - \tilde{w}^g_{\xi^{j_0^n}}\right)(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\xi}^{\varepsilon_n}) = 0 \text{ for some } (i_0^n, j_0^n) \in \mathcal{I}.$$

We obtain then by boundedness of $\left(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \varepsilon_n\tilde{\xi}^{\varepsilon_n}\right)_n$, (6.5.9) and using Assumption 6.3.1 (and in particular the constant c_0 introduced there)

$$\begin{aligned}
&- 4c_0\varepsilon_n^2(\varepsilon_n\tilde{w}^g)^3(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\xi}^{\varepsilon_n})\left(\tilde{w}^g_{\xi^{i_0^n}} - \tilde{w}^g_{\xi^{j_0^n}}\right)(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\xi}^{\varepsilon_n}) \\
&+ \varepsilon_n^3 v_z^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n})\left[\lambda^{i_0^n,j_0^n} - (1+\delta)(\overline{w}^g_{\xi^{i_0^n}} - \overline{w}^g_{\xi^{j_0^n}})(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\rho}^{\varepsilon_n})\right] + \circ(\varepsilon_n^3) \ge 0.
\end{aligned}$$

And by positivity of $\tilde{w}^g$, we have:

$$\begin{aligned}
0 &\le -4c_0\lambda^{i_0^n,j_0^n} v_z^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n})\varepsilon_n^2(\varepsilon_n\tilde{w}^g)^3(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\xi}^{\varepsilon_n}) - \delta\lambda^{i_0^n,j_0^n}\varepsilon_n^3 v_z^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}) + \circ(\varepsilon_n^3) \\
&\le -\delta\lambda^{i_0^n,j_0^n}\varepsilon_n^3 v_z^g(\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}) + \circ(\varepsilon_n^3),
\end{aligned}$$

which leads to a contradiction when n goes to $+\infty$.

Step 4: Since $\psi^{\varepsilon,\delta}$ is smooth enough, we are now able to use it as a test function for the parabolic operator in (6.2.11). By the supersolution property of $v^{\varepsilon,g}$, we have:

$$kv^{\varepsilon,g} - \mathcal{L}\psi^{\varepsilon,\delta} - \widetilde{U}_1(\psi_x^{\varepsilon,\delta})(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{x}^\varepsilon, \tilde{y}^\varepsilon) \geq 0.$$

Since $(t,s,z) \mapsto \mathcal{O}_0^g(t,s,z)$ is continuous by Assumption 6.3.2 and since $(\tilde{t}^\varepsilon, \tilde{s}^\varepsilon, \tilde{z}^\varepsilon)$ is bounded, we know that $(\tilde{\xi}^\varepsilon)_\varepsilon$ is bounded. By standard results of the theory of viscosity solutions, we then have a sequence $(\varepsilon_n)_n$ such that $\varepsilon_n \longrightarrow 0$ and such that

$$(t_n, s_n, z_n, \xi_n) := (\tilde{t}^{\varepsilon_n}, \tilde{s}^{\varepsilon_n}, \tilde{z}^{\varepsilon_n}, \tilde{\xi}^{\varepsilon_n}) \longrightarrow (t_0, s_0, z_0, \tilde{\xi}),$$

for some $\tilde{\xi} \in \mathbb{R}^d$. We then have

$$\begin{aligned}
&-\frac{1}{2}v_{zz}^g(t_n,s_n,z_n)\left|\sigma^T(t_n,s_n)\xi_n\right|^2 + \frac{1}{2}(1+\delta)\mathrm{Tr}\left[\alpha^g(\alpha^g)^T(t_n,s_n,z_n)w_{\xi\xi}^g(s_n,z_n,\xi_n)\right]\\
&-\mathcal{A}^g\phi(t_n,s_n,z_n) - \mathcal{A}^g\Phi^{\varepsilon_n}(t_n,s_n,x_n,y_n) + \mathcal{R}^{\varepsilon_n}(\phi+\Phi^{\varepsilon_n},(1+\delta)\tilde{w}^g)(t_n,s_n,z_n,\xi_n) \geq 0,
\end{aligned}$$

where the remainder term $\mathcal{R}^{\varepsilon_n}(\phi+\Phi^{\varepsilon_n},(1+\delta)\tilde{w}^g)(t_n,s_n,z_n,\xi_n)$ is controlled using the result of Lemma 6.5.4.

We know that w^g is C^2 at the points (t_n,s_n,z_n,ξ_n) but not necessarily at $(t_0,s_0,z_0,\tilde{\xi})$, which might be so that $\tilde{\rho} := \tilde{\xi}/(\eta^g(t_0,s_0,z_0)) \in \partial\mathcal{O}_0^g(t,s,z)$. Now we remind the reader that by definition of w^g and since $\rho_n \in \mathcal{O}^g(t_n,s_n,z_n,\xi_n)$

$$-\frac{1}{2}v_{zz}^g(t_n,s_n,z_n)\left|\sigma^T(t_n,s_n)\xi_n\right|^2 + \frac{1}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T w_{\xi\xi}^g\right](t_n,s_n,z_n,\xi_n) = -a^g(t_n,s_n,z_n),$$

so that:

$$\begin{aligned}
&a^g(t_,s_n,z_n) - \mathcal{A}^g\phi(t_n,s_n,z_n) - \mathcal{A}^g\Phi^{\varepsilon_n}(t_n,s_n,x_n,y_n) + \delta\Big(a^g(t_n,s_n,z_n)\\
&+\frac{1}{2}v_{zz}^g(t_n,s_n,z_n)\left|\sigma^T(t_n,s_n)\xi_n\right|^2\Big) + \mathcal{R}^{\varepsilon_n}(\phi+\Phi_n^\varepsilon,(1+\delta)\tilde{w}^g)(t_n,s_n,z_n,\xi_n) \geq 0.
\end{aligned}$$

Therefore, $w_{\xi\xi}^g$ no longer appears directly in the above equation, except in the remainder $\mathcal{R}^{\varepsilon_n}(\phi+\Phi_n^\varepsilon,(1+\delta)\tilde{w}^g)$ for which it is implicitly understood that we do the same transformation. Now by continuity of the map $(t,s,z) \longmapsto a^g(t,s,z)$ stated in Assumption 6.3.2, and since we clearly have that $\mathcal{R}^{\varepsilon_n}(\phi+\phi^{\varepsilon_n},(1+\delta)\tilde{w}^g)(t_n,s_n,z_n,\xi_n) \longrightarrow 0$ (recall that we are away from T here, so that none of the quantities in the upper bound given in Lemma 6.5.4 can explode) and $\mathcal{A}^g\Phi^{\varepsilon_n}(t_n,s_n,x_n,y_n) \longrightarrow 0$ when $n \longrightarrow \infty$, Φ^ε and all its derivatives go to 0. Finally, we obtain

$$a^g(t_0,s_0,z_0) - \mathcal{A}^g\phi(t_0,s_0,z_0) + \delta\left(a^g(t_0,s_0,z_0) - \frac{1}{2}v_{zz}^g(t_0,s_0,z_0)|\sigma^T(t_0,s_0)\tilde{\xi}|^2\right) \geq 0.$$

Recall that $\tilde{\xi}$ may depend on δ but is uniformly bounded. Then we can send δ to 0 to obtain the required result. □

6.5.4 Viscosity subsolution on $\{T\} \times \mathbb{R}^d \times \mathbb{R}_+$

In contrast with the previous section, the use of $u^{g,\varepsilon}$ is not necessary here, and we will therefore concentrate only on $\bar{u}^{\varepsilon,g}$.

Let $(s_0,z_0,\phi) \in (0,+\infty)^{d+1} \times C^2\left((0,+\infty)^{d+1}\right)$ be such that

$$0 = (u^{g,*}-\phi)(T,s_0,z_0) > (u^{g,*}-\phi)(t,s,z),\ \forall (t,s,z) \in [0,T] \times (0,+\infty)^{d+1} \setminus \{(T,s_0,z_0)\}.$$

By definition of viscosity solutions, we want to deduce that $\phi(T, s_0, z_0) \leq 0$.

Assume on the contrary that $\phi(T, s_0, z_0) > 2\delta$ for some $\delta > 0$. Then we have for $r_0 > 0$ small enough,

$$\phi(t, s, z) > \delta, \ \forall (t, s, z) \in [T - r_0, T] \times B_{r_0}(s_0, z_0). \tag{6.5.10}$$

Let us then consider a sequence $(t_\varepsilon, s_\varepsilon, z_\varepsilon)$ converging to (T, s_0, z_0) such that $\hat{\bar{u}}^{\varepsilon,g}(t_\varepsilon, s_\varepsilon, z_\varepsilon, 0) \longrightarrow u^{g,*}(T, s_0, z_0)$. We introduce:

$$l_*^\varepsilon := \hat{\bar{u}}^{\varepsilon,g}(t_\varepsilon, s_\varepsilon, z_\varepsilon, 0) - \phi(t_\varepsilon, s_\varepsilon, z_\varepsilon) \underset{\varepsilon \longrightarrow 0}{\longrightarrow} 0. \tag{6.5.11}$$

By assumption 6.3.3, there exists $0 < r_1 < r_0$ such that:

$$b^* := \sup \left\{ \hat{\bar{u}}^{g,\varepsilon}(t, s, z, 0), \ (t, s, z) \in [T - r_1, T] \times B_{r_1}(s_0, z_0) \right\} < +\infty.$$

We will denote for simplicity $B_1 := [T - r_1, T] \times B_{r_1}(s_0, z_0)$. We split the proof in two parts.

Step 1: We first show that there is some ε_0 such that $t_\varepsilon < T$ for any $\varepsilon \leq \varepsilon_0$. Assume on the contrary that we have a sequence $\varepsilon_n \longrightarrow 0$ such that $\hat{\bar{u}}^{\varepsilon,g}(t_{\varepsilon_n}, s_{\varepsilon_n}, z_{\varepsilon_n}) \longrightarrow u^{g,*}(T, s_0, z_0)$ and such that $t_{\varepsilon_n} = T$ for countably many n. Extracting a further subsequence if necessary, we can assume without loss of generality that the sequence $(t_{\varepsilon_n})_n$ is actually stationary at T. We then have

$$\begin{aligned} \bar{u}^{g,\varepsilon}(T, s_{\varepsilon_n}, x_{\varepsilon_n}, y_{\varepsilon_n}) &= \frac{v^g(T, s_{\varepsilon_n}, z_{\varepsilon_n}) - v^{\varepsilon,g}(T, s_{\varepsilon_n}, x_{\varepsilon_n}, y_{\varepsilon_n})}{\varepsilon^2} \\ &= \frac{U_2(z_{\varepsilon_n} - g(s_{\varepsilon_n})) - U_2(\ell^{\varepsilon_n}(x_{\varepsilon_n}, y_{\varepsilon_n}) - g(s_{\varepsilon_n}))}{\varepsilon^2}, \end{aligned}$$

where $(x_{\varepsilon_n}, y_{\varepsilon_n}) := (z_{\varepsilon_n} - \mathbf{y}^g(T, s_{\varepsilon_n}, z_{\varepsilon_n}), \mathbf{y}^g(T, s_{\varepsilon_n}, z_{\varepsilon_n}))$.

Since $\mathbf{y}^g(T, \cdot, \cdot)$ is continuous by Assumption 6.3.1, we have by definition of ℓ^ε:

$$\left| \frac{\ell^{\varepsilon_n}(x_{\varepsilon_n}, y_{\varepsilon_n}) - z_{\varepsilon_n}}{\varepsilon_n^3} \right| = \left| \frac{1}{\varepsilon_n^3} \sum_{j=1}^d \left(y_j \left(\frac{1}{1 + \varepsilon_n^3 \lambda^{j,0}} - 1 \right) \mathbf{1}_{y_j \geq 0} + \varepsilon_n^3 \lambda^{0,j} y_j \mathbf{1}_{y_j < 0} \right) \right| \leq C,$$

for some constant C independent of n and ε.

Since U^2 is C^1, we deduce that

$$\frac{U_2(z_{\varepsilon_n} - g(s_{\varepsilon_n})) - U_2(\ell^{\varepsilon_n}(x_{\varepsilon_n}, y_{\varepsilon_n}) - g(s_{\varepsilon_n}))}{\varepsilon_n^2} \longrightarrow 0,$$

as $n \longrightarrow +\infty$, wich contradicts (6.5.10) and (6.5.11).

Step 2: Similarly as in Section 6.5.3 and in [90] and [81], we build a test function ψ^ε for $v^{g,\varepsilon}$. Let $\overline{p} \in (0, 1)$ be a constant which will be fixed later. We define ψ^ε by

$$\hat{\psi}^\varepsilon(t, s, z, \xi) := v^g(t, s, z) - \varepsilon^2(l_*^\varepsilon + \phi(t, s, z) + \Phi^\varepsilon(t, s, z)) - \varepsilon^4 \hat{\varpi}(\xi),$$

with

$$\Phi^\varepsilon(t, s, z) := l_0(T - t)^{\overline{p}} + l_1 \left((s - s_\varepsilon)^2 + (z - z_\varepsilon)^2 \right), \ \hat{\varpi}(\xi) := |\xi|^2,$$

for some constants l_1 and l_0. By definition, we have $\hat{\bar{u}}^{\varepsilon,g}(t, s, z, 0) \leq b^*$ for all $(t, s, z) \in B_1$. We now choose l_1 large enough and l_0 so that on $B_1 \backslash B_2$ where $B_2 := [T - \frac{r_1}{2}, T] \times B_{\frac{r_1}{2}}(s_0, z_0)$, we have

$$\phi(t, s, z) + \Phi^\varepsilon(t, s, z) + \varepsilon^2 \xi^2 \geq 2 + b^*.$$

We then have that $v^\varepsilon - \psi^\varepsilon$ has a local minimizer in B_1. Indeed on ∂B_1, for ε small enough, since $l_*^\varepsilon \longrightarrow 0$, we have:

$$\begin{aligned}\frac{v^{\varepsilon,g}(t,s,x,y)-\psi^\varepsilon(t,s,x,y)}{\varepsilon^2} &= -u^{\varepsilon,g}(t,s,x,y)+l_*^\varepsilon+\phi(t,s,z)+\Phi^\varepsilon(t,s,z)+(y-\mathbf{y}^g(t,s,z))^2\\ &\geq -b^*+l_*^\varepsilon+2+b^*>0.\end{aligned}$$

Since $v^{\varepsilon,g}(t_\varepsilon,s_\varepsilon,x_\varepsilon,y_\varepsilon)-\psi^\varepsilon(t_\varepsilon,s_\varepsilon,x_\varepsilon,y_\varepsilon)=0$, we then have the existence of a local minimizer $(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde x_\varepsilon,\tilde y_\varepsilon)\in B_1$. We denote by $(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon,\tilde \xi_\varepsilon)$ the corresponding minimizer after the usual change of variable. We also recall that by classical results on viscosity solutions, we have $(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon)\longrightarrow(T,s_0,z_0)$ as ε goes to 0.

Now by the viscosity supersolution property of v^ε at $(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde x_\varepsilon,\tilde y_\varepsilon)$, we have, since we recall that we do have $\tilde t_\varepsilon<T$

$$\min_{(i,j)\in\mathcal{I}}\left\{k\psi^\varepsilon-\mathcal{L}\psi^\varepsilon-\kappa\tilde U_1(\psi_x^\varepsilon),\ \Lambda_{i,j}^\varepsilon\cdot\left(\psi_x^\varepsilon,\psi_y^\varepsilon\right)\right\}(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde x_\varepsilon,\tilde y_\varepsilon)\geq 0. \tag{6.5.12}$$

Step 3: We now show that there exists $\hat\varepsilon$ such that for $\varepsilon\leq\hat\varepsilon$ the sequence $(\tilde\xi_\varepsilon)_{0<\varepsilon\leq\hat\varepsilon}$ is bounded. Since the sequence $(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon,\varepsilon\tilde\xi_\varepsilon)$ is bounded, we indeed easily compute that the gradient constraints in (6.5.12) implies for $(i,j)\in\mathcal{I}$

$$\Lambda_{i,j}^\varepsilon\cdot\left(\psi_x^\varepsilon,\psi_y^\varepsilon\right)\left(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon,\tilde\xi_\varepsilon\right)=\varepsilon^3\left(\lambda^{i,j}v_z^g(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon)-2(e_i-e_j)\cdot\tilde\xi_\varepsilon\right)+\circ(\varepsilon^3)\geq 0.$$

Then for $i=0$ and $j\geq 1$, we obtain, since $\lambda^{0,j}\in\mathcal{I}$ for any $j\geq1$, that for ε small enough

$$\tilde\xi_\varepsilon^j\geq-\lambda^{0,j}v_z^g(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon)>-\mathrm{Const},$$

where $\mathrm{Const}>0$ is uniform in ε. Then for $i\geq1$ and $j=0$, we obtain that for ε small enough

$$\tilde\xi_\varepsilon^i\leq\lambda^{i,0}v_z^g(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon)<\mathrm{Const}.$$

Hence $(\tilde\xi^\varepsilon)$ is bounded for ε small enough.

Step 4: We now deduce from (6.5.12) and Lemma 6.5.4 that at point $(\tilde t^\varepsilon,\tilde s^\varepsilon,\tilde x^\varepsilon,\tilde y^\varepsilon)$:

$$\varepsilon^2\left(-\frac{v_{zz}^g}{2}\left|\sigma^T\tilde\xi_\varepsilon\right|^2+\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\right]-\mathcal{A}^g(l_*^\varepsilon+\phi+\Phi^\varepsilon)+\mathcal{R}^\varepsilon(l_*^\varepsilon+\phi+\Phi^\varepsilon,\hat\varpi)\right)\geq 0. \tag{6.5.13}$$

Since for ε small enough, $(\tilde\xi_\varepsilon)$ is bounded, and since $\hat\varpi$ only depends on ξ and since Φ^ε and all its derivatives with respect to s and z are bounded, we obtain by Lemma 6.5.4 and using Assumption 6.3.1 that for some $\mathrm{Const}>0$

$$\left|\mathcal{R}^\varepsilon(l_*^\varepsilon+\phi+\Phi^\varepsilon,\hat\varpi)\right|(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon,\tilde\xi_\varepsilon)\leq\varepsilon\frac{\mathrm{Const}}{(T-\tilde t_\varepsilon)^{1-\eta}}.$$

Now by definition of $\mathcal{A}^g$ and Φ^ε, we observe easily that

$$\mathcal{A}^g(l_*^\varepsilon+\phi+\Phi^\varepsilon)(\tilde t_\varepsilon,\tilde s_\varepsilon,\tilde z_\varepsilon)=\frac{\overline{p}l_0}{(T-\tilde t_\varepsilon)^{1-\overline{p}}}+r^\varepsilon,$$

where r^ε is bounded near 0, so that by (6.5.13) and Assumption 6.3.1(i), we obtain

$$-\frac{\overline{p}l_0}{(T-\tilde t_\varepsilon)^{1-\overline{p}}}+\frac{\mathrm{Const}}{(T-\tilde t_\varepsilon)^{1-\mu}}+\varepsilon\frac{\mathrm{Const}}{(T-\tilde t_\varepsilon)^{1-\eta}}+\tilde r^\varepsilon\geq 0,$$

where

$$\tilde r^\varepsilon:=-r^\varepsilon+\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\right],$$

is bounded near 0. Choosing $\overline{p}=(\eta\wedge\mu)/2$, this leads to a contradiction for $\varepsilon>0$ small enough, since $\tilde t_\varepsilon$ goes to T. □

6.5.5 Viscosity supersolution

We are interested in this section in the supersolution part. We first note that since $\bar{u}^{\varepsilon,g} \geq 0$, the supersolution property on $\{T\} \times \mathbb{R}^d \times \mathbb{R}_+$ is indeed trivial. We then only focus on the interior of the domain. Our aim is then to show:

Proposition 6.5.1. *Under Assumptions 6.3.1, 6.3.2 and 6.3.3, u^g_* is a viscosity supersolution of the second corrector equation (6.2.15) on $[0,T) \times (0,+\infty)^{d+1}$.*

We first recall some crucial properties proved in [81], that we shall use in the proof of Proposition 6.5.1. The first one concerns a regular approximation of $\tilde{w}^g$ by convolution. Consider $\upsilon : \mathbb{R}^d \longrightarrow \mathbb{R}$ a positive, even, C^∞ kernel with support in $B_1(0)$. We then define for $m > 0$:

$$\tilde{w}^{g,m}(\cdot,\xi) := \int_{\mathbb{R}^d} \upsilon^m(\zeta)\, \tilde{w}^g(\cdot,\xi-\zeta)d\zeta,$$

where $\upsilon^m(x) := m^{-d}\upsilon(x/m)$. The proof of the following lemma can be found in [81]:

Lemma 6.5.5. *Under Assumtion 6.3.2, we have for any $m > 0$ that:*
(i) *$\tilde{w}^{g,m}$ is C^2, convex in ξ and for any $(t,s,z,\xi) \in [0,T) \times (0,+\infty)^{d+1} \times \mathbb{R}^d$,*

$$0 \leq \tilde{w}^{g,m}(t,s,z,\xi) \leq Lv^g_z(t,s,z)(1+m)(1+|\xi|).$$

(ii) *$\tilde{w}^{g,m}$ is smooth in $(t,s,z) \in [0,T) \times (0,+\infty)^{d+1}$, and satisfies the following estimates, uniformly in m,*

$$\begin{aligned}
&\left(|\tilde{w}^{g,m}| + |\tilde{w}^{g,m}_t| + |\tilde{w}^{g,m}_s| + |\tilde{w}^{g,m}_{ss}| + |\tilde{w}^{g,m}_z| + |\tilde{w}^{g,m}_{sz}| + |\tilde{w}^{g,m}_{zz}|\right)(\cdot,\xi) \leq C(\cdot)(1+m)\,(1+|\xi|)\\
&\left(\left|\tilde{w}^{g,m}_\xi\right| + \left|\tilde{w}^{g,m}_{s\xi}\right| + \left|\tilde{w}^{g,m}_{z\xi}\right|\right)(\cdot,\xi) \leq C(\cdot)\\
&\left|\tilde{w}^{g,m}_{\xi\xi}\right|(\cdot,\xi) \leq C(\cdot)\mathbf{1}_{\xi\in\overline{\mathcal{B}}^g(\cdot)},
\end{aligned} \tag{6.5.14}$$

where $C(t,s,z)$ is a continuous function depending on the Merton value function and its derivatives, and $\mathcal{B}^g(t,s,z)$ is some ball with a continuous radius, centered at 0.

(iii) *For every $(i,j) \in \mathcal{I}$ and every $(t,s,z,\xi) \in [0,T) \times (0,+\infty)^{d+1} \times \mathbb{R}^d$*

$$-\lambda^{i,j}v^g_z(t,s,z) + \tilde{w}^{g,m}_{\xi_i}(t,s,z,\xi) - \tilde{w}^{g,m}_{\xi_j}(t,s,z,\xi) \leq 0.$$

(iv) *For every $(t,s,z,\xi) \in [0,T) \times (0,+\infty)^{d+1} \times \mathbb{R}^d$, we have*

$$\frac{1}{2}v^g_{zz}(t,s,z)\int_{\mathbb{R}^d} \upsilon^m(\zeta)\left|\sigma(t,s)^T(\xi-\zeta)\right|^2 d\zeta - \frac{1}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\tilde{w}^{g,m}_{\xi\xi}\right](t,s,z,\xi) + a^g(t,s,z) \leq 0.$$

To build a test function in the proof of Proposition 6.5.1 we will also use the following result.

Lemma 6.5.6. *For any $\delta \in (0,1)$ and any $\nu > 0$, there exists $a^\delta := a^{\delta,\nu} > 1$ and a function $h^{\delta,\nu} : \mathbb{R}^d \longrightarrow [0,1]$ such that $h^{\delta,\nu}$ is C^∞, $h^{\delta,\nu} = 1$ on $B_1(0)$ and $h^{\delta,\nu} = 0$ on $B_{a^\delta}(0)^c$. Moreover, for any $1 \leq i,j \leq d$ and for any $\xi \in \mathbb{R}^d$*

$$\left|h^{\delta,\nu}_{\xi_i}(\xi)\right| \leq \frac{\nu\delta}{3},\ |\xi|\,|h^{\delta,\nu}_{\xi_i}| \leq \nu\delta,\ \text{and}\ |\xi|\,|h^{\delta,\nu}_{\xi\xi}| + |h^{\delta,\nu}_{\xi\xi}| \leq C^*,$$

for some constant C^* independent of δ.

This Lemma and its proof can be found in [81].

We conclude these preliminary results with the following useful lemma.

Lemma 6.5.7. *For any $\delta \in (0,1)$, $\nu > 0$ and $m > 0$, the map $\Upsilon := \tilde{w}^{g,m} h^{\delta,\nu}$ is smooth and satisfies the following estimates*

$$\left(|\Upsilon| + |\Upsilon_t| + |\Upsilon_s| + |\Upsilon_{ss}| + |\Upsilon_z| + |\Upsilon_{sz}| + |\Upsilon_{zz}|\right)(t,s,z,\xi) \le C(t,s,z)(1+m)\left(1+|\xi|\right) \mathbf{1}_{|\xi| \le a^\delta}$$
$$\left(|\Upsilon_\xi| + |\Upsilon_{s\xi}| + |\Upsilon_{z\xi}|\right)(t,s,z,\xi) \le 4C(t,s,z)\left(1 + (1+m)(1+|\xi|)\frac{\nu\delta\sqrt{d}}{3}\right) \mathbf{1}_{|\xi| \le a^\delta}$$
$$|\Upsilon_{\xi\xi}(t,s,z,\xi)| \le C(t,s,z)\left(1 + 2\frac{\nu\delta\sqrt{d}}{3} + C^*(1+m)(1+|\xi|)\right) \mathbf{1}_{|\xi| \le a^\delta}, \tag{6.5.15}$$

where $C(t,s,z)$ and C^* were introduced in Lemmas 6.5.5 and 6.5.6.

The proof is easy by direct calculations, using the results of Lemmas 6.5.5 and 6.5.6.

Proof of Proposition 6.5.1.

Consider $(t_0, s_0, z_0) \in [0,T) \times (0,+\infty)^{d+1}$ and ϕ, C^2 such that for all $(t,s,z) \in [0,T) \times (0,+\infty)^{d+1} \backslash \{(t_0,s_0,z_0)\}$:

$$0 = (u^g_* - \phi)(t_0,s_0,z_0) < (u^g_* - \phi)(t,s,z).$$

We want to show that $\mathcal{A}^g\phi(t_0,s_0,z_0) - a^g(t_0,s_0,z_0) \ge 0$. Assume on the contrary that:

$$\mathcal{A}^g\phi(t_0,s_0,z_0) - a^g(t_0,s_0,z_0) < 0, \tag{6.5.16}$$

Then there exists $r_0 > 0$ such that $\mathcal{A}^g\phi(t,s,z) - a^g(t,s,z) \le 0$ on $B_{r_0}(t_0,s_0,z_0)$.

We proceed in 5 steps. The first two steps consist in defining a test function for the dynamic programming equation (6.2.11). The third one is devoted to prove that the gradient constraint for this test function is not binding, so that the parabolic part is. The last two steps lead to the required contradiction of (6.5.16).

Step 1: By Lemma 6.5.3, there exists a sequence $(t^\varepsilon, s^\varepsilon, z^\varepsilon) \longrightarrow (t_0,s_0,z_0)$ when $\varepsilon \longrightarrow 0$ such that

$$\hat{u}^{\varepsilon,g}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) \underset{\varepsilon \longrightarrow 0}{\longrightarrow} u^g_*(t_0,s_0,z_0).$$

Then we have that $l^\varepsilon_* := \hat{u}^{\varepsilon,g}(t^\varepsilon,s^\varepsilon,z^\varepsilon,0) - \phi(t^\varepsilon,s^\varepsilon,z^\varepsilon) \longrightarrow 0$ and $(x^\varepsilon, y^\varepsilon) \longrightarrow (x_0,y_0)$, as ε goes to 0, where

$$(x^\varepsilon, y^\varepsilon) := \left(z^\varepsilon - \mathbf{y}^g(t^\varepsilon,s^\varepsilon,z^\varepsilon) \cdot \mathbf{1}_d, \mathbf{y}^g(t^\varepsilon,s^\varepsilon,z^\varepsilon)\right),$$

and

$$(x_0,y_0) := \left(z_0 - \mathbf{y}^g(t_0,s_0,z_0) \cdot \mathbf{1}_d, \mathbf{y}^g(t_0,s_0,z_0)\right).$$

We then consider $\varepsilon_0 > 0$ such that for all $\varepsilon \le \varepsilon_0$

$$|t^\varepsilon - t_0| + |s^\varepsilon - s_0| + |z^\varepsilon - z_0| \le \frac{r_0}{4}, \text{ and } |l^\varepsilon_*| \le 1.$$

Consider next a constant $q_0 > 0$ such that:

$$\sup_{(t,s,z) \in B_{r_0/2}(t_0,s_0,z_0)} \left\{\phi(t,s,z) + C(t,s,z)\right\} + 3 \le q_0 \left(\frac{r_0}{12}\right)^4,$$

where $C(t,s,z)$ is the continuous function appearing in (6.5.14). We then introduce:

$$\phi^\varepsilon(t,s,z) := \phi(t,s,z) - q_0\left(|t-t^\varepsilon|^4 + |z-z^\varepsilon|^4 + |s-s^\varepsilon|^4\right).$$

We then have for $\varepsilon \le \varepsilon_0$ and $(t,s,z) \in \partial B_{r_0/2}(t_0,s_0,z_0)$ that

$$|t-t^\varepsilon| + |z-z^\varepsilon| + |s-s^\varepsilon| \ge \frac{r_0}{4},$$

and thus that

$$|t-t^\varepsilon|^4 + |z-z^\varepsilon|^4 + |s-s^\varepsilon|^4 \ge \frac{1}{81}\left(|t-t^\varepsilon| + |z-z^\varepsilon| + |s-s^\varepsilon|\right)^4 \ge \left(\frac{r_0}{12}\right)^4.$$

Then on $\partial B_{r_0/2}(t_0,s_0,z_0)$, we have:

$$\begin{aligned}\phi^\varepsilon(t,s,z) + l^*_\varepsilon + C(t,s,z) &= \phi(t,s,z) + C(t,s,z) - q_0\left(|t-t^\varepsilon|^4 + |z-z^\varepsilon|^4 + |s-s^\varepsilon|^4\right) + l^*_\varepsilon \\ &\le q_0\left(\frac{r_0}{12}\right)^4 - 3 - q_0\left(|t-t^\varepsilon|^4 + |z-z^\varepsilon|^4 + |s-s^\varepsilon|^4\right) + 1 \\ &\le q_0\left(\left(\frac{r_0}{12}\right)^4 - |t-t^\varepsilon|^4 - |z-z^\varepsilon|^4 - |s-s^\varepsilon|^4\right) - 2 \le -2. \quad (6.5.17)\end{aligned}$$

Consider next the function $\Phi^\varepsilon := \phi^\varepsilon - \phi + l^*_\varepsilon$. By linearity of the operator $\mathcal{A}^g$, and Assumption 6.3.1, we have that there exists $\varepsilon^0 > 0$ such that $b < \infty$, where

$$b := \sup\left\{\,|\mathcal{A}^g\Phi^\varepsilon|\,(t,s,z),\ \varepsilon \le \varepsilon^0,\ (t,s,z) \in \bar{B}_{r_0/2}(t_0,s_0,z_0)\right\}. \quad (6.5.18)$$

Throughout the rest of the proof, we let $m \in (0,1]$. Now for any $\delta \in (0,1)$ and $\nu > 0$, let $h^{\delta,\nu}$ be the function defined by Lemma 6.5.6, and introduce a parameter $\xi^* := 1 \vee \tilde{\xi}_0 \vee \tilde{\xi}^*$, where $\tilde{\xi}_0 > 0$ is greater than $\eta^g(t_0,s_0,z_0)$ times the diameter of $\mathcal{O}^g(t_0,s_0,z_0)$ and large enough so that for every $|\xi| \ge \tilde{\xi}_0$, $\tilde{w}^{g,m}_{\xi\xi}(t,s,z,\xi) = 0$, for every $(t,s,z) \in B_{r_0/2}(t_0,s_0,z_0)$, and $\tilde{\xi}^*$ is such that for any $\xi \in B_{\tilde{\xi}^*}(0)^c$ and $(t,s,z) \in \bar{B}_{r_0/2}(t_0,s_0,z_0)$, we have

$$-\frac{1}{2}v^g_{zz}\left|\sigma^T\xi\right|^2 - \mathcal{A}^g(\phi + \Phi^\varepsilon) > \frac{1}{2}\mathrm{Tr}\left[\alpha^g(\alpha^g)^T\right]C(t,s,z)\left(C^* + \frac{\sqrt{d}\delta\lambda}{4L}\right) + 1, \quad (6.5.19)$$

where C^* is the constant introduced in Lemma 6.5.6 and $C(t,s,z)$ is the function introduced in Lemma 6.5.5 (we remind the reader that they are both uniform in m).

Define then $H(\xi) := h^{\delta,\nu}\left(\frac{\xi}{\xi^*}\right)$ and the test function

$$\hat{\psi}^{\varepsilon,\delta,m}(t,s,z,\xi) := v^g(t,s,z) - \varepsilon^2(\phi^\varepsilon(t,s,z) + l^*_\varepsilon) - \varepsilon^4(1-\delta)\tilde{w}^{g,m}(t,s,z,\xi)H(\xi).$$

Step 2: In this part, we introduce a second modification of the test function. Introduce

$$I^{\varepsilon,\delta,m}(t,s,z,\xi) := \varepsilon^{-2}\left(\hat{v}^{\varepsilon,g} - \hat{\psi}^{\varepsilon,\delta,m}\right)(t,s,z,\xi),$$

we want to show that $I^{\varepsilon,\delta,m}$ has a local maximizer on the interior of the domain. By definition,

$$I^{\varepsilon,\delta,m}(t,s,z,\xi) = \phi^\varepsilon(t,s,z) - \hat{u}^{\varepsilon,g}(t,s,z,\xi) + l^\varepsilon_* - \varepsilon^2 w^g(t,s,z,\xi) + \varepsilon^2(1-\delta)H(\xi)\tilde{w}^{g,m}(t,s,z,\xi).$$

Recall that for $\xi = 0$, $w^g(\cdot,\cdot,\cdot,0) = 0$, so that by definition of l^ε_*, we have

$$I^{\delta,\varepsilon,m}(t^\varepsilon,s^\varepsilon,z^\varepsilon,0) = \varepsilon^2(1-\delta)\tilde{w}^{g,m}(t^\varepsilon,s^\varepsilon,z^\varepsilon,0),$$

which goes to 0 as ε goes to 0, uniformly in $m \in (0,1]$, because of the uniform bounds given by Lemma 6.5.5. Hence, there exists ε_1 such that for any $\varepsilon \leq \varepsilon_0 \wedge \varepsilon^0 \wedge \varepsilon_1$,

$$I^{\delta,\varepsilon,m}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) \geq -1. \tag{6.5.20}$$

Using successively that $v^{\varepsilon,g} \leq v^g$, $0 \leq \tilde{w}^{g,m}(t,s,z,\xi) \leq 2C(t,s,z)(1+|\xi|)$ (with $C(t,s,z)$ still being the continuous function appearing (6.5.14) and where we used the fact that $m \in (0,1]$) and $0 \leq H(\xi) \leq \mathbf{1}_{|\xi| \leq a^\delta \xi^*}$, we have

$$\begin{aligned} I^{\varepsilon,\delta,m}(t,s,z,\xi) &\leq \phi^\varepsilon(t,s,z) + l_*^\varepsilon + \varepsilon^2(1-\delta)H(\xi)\tilde{w}^{g,m}(t,s,z,\xi) \\ &\leq \phi^\varepsilon(t,s,z) + l_*^\varepsilon + 2\varepsilon^2(1-\delta)C(t,s,z)H(\xi)(1+|\xi|) \\ &\leq \phi^\varepsilon(t,s,z) + l_*^\varepsilon + 2\varepsilon^2 C(t,s,z)(1+a^\delta\xi^*), \end{aligned}$$

where a^δ is the constant introduced in Lemma 6.5.6. Then for any $\varepsilon \leq \varepsilon^\delta := \sqrt{2(1+a^\delta\xi^*)}$, we have

$$I^{\varepsilon,\delta,m}(t,s,z,\xi) \leq \phi^\varepsilon(t,s,z) + l_*^\varepsilon + C(t,s,z).$$

Introduce then $Q_{(t_0,s_0,z_0)} := \left\{(t,s,z,\xi),\ (t,s,z) \in \bar{B}_{r_0/2}(t_0,s_0,z_0)\right\}$. The above implies in particular that for $\varepsilon \leq \varepsilon_0 \wedge \varepsilon^0 \wedge \varepsilon_1 \wedge \varepsilon^\delta$

$$\mathfrak{I}(\varepsilon,\delta,m) := \sup_{(t,s,z,\xi) \in Q_{(t_0,s_0,z_0)}} I^{\varepsilon,\delta,m}(t,s,z,\xi) < \infty.$$

Moreover, using (6.5.17), we deduce that for $(t,s,z,\xi) \in \partial Q_{(t_0,s_0,z_0)}$ and for any $\varepsilon \leq \varepsilon_0 \wedge \varepsilon^0 \wedge \varepsilon_1 \wedge \varepsilon^\delta$

$$I^{\varepsilon,\delta,m}(t,s,z,\xi) \leq -2. \tag{6.5.21}$$

We can now consider a sequence $(\hat{t}_n, \hat{s}_n, \hat{z}_n, \hat{\xi}_n)$ in $\operatorname{int}\left(Q_{(t_0,s_0,z_0)}\right)$ such that

$$I^{\varepsilon,\delta,m}(\hat{t}_n, \hat{s}_n, \hat{z}_n, \hat{\xi}_n) \geq \mathfrak{I}(\varepsilon,\delta,m) - \frac{1}{2n}.$$

It is now time to penalize the test function to obtain the existence of an interior maximiser, which is not obvious with our previous construction. We consider $f : \mathbb{R} \longrightarrow [0,1]$, smooth such that $f(0) = 1$ and $f(x) = 0$ if $x \geq 1$. Define

$$\hat{\psi}^{\varepsilon,\delta,m,n}(t,s,z,\xi) := \hat{\psi}^{\varepsilon,\delta,m}(t,s,z,\xi) - \frac{\varepsilon^2}{n} f\left(|\xi - \hat{\xi}_n|\right).$$

Consider then

$$I^{\varepsilon,\delta,m,n}(t,s,z,\xi) := \varepsilon^{-2}\left(\hat{v}^{\varepsilon,g} - \hat{\psi}^{\varepsilon,\delta,m,n}\right)(t,s,z,\xi) = I^{\varepsilon,\delta,m}(t,s,z,\xi) + \frac{1}{n} f\left(|\xi - \hat{\xi}_n|\right).$$

By definition of $(\hat{t}_n, \hat{s}_n, \hat{z}_n, \hat{\xi}_n)$, we have for any $(t,s,z,\xi) \in Q_{(t_0,s_0,z_0)}$

$$I^{\varepsilon,\delta,m,n}(\hat{t}_n, \hat{s}_n, \hat{z}_n, \hat{\xi}_n) = I^{\varepsilon,\delta,m}(\hat{t}_n, \hat{s}_n, \hat{z}_n, \hat{\xi}_n) + \frac{1}{n} \geq I^{\varepsilon,\delta,m}(t,s,z,\xi) + \frac{1}{2n}.$$

Notice that for $|\xi - \hat{\xi}_n| \geq 1$, we have $I^{\varepsilon,\delta,m,n}(t,s,z,\xi) = I^{\varepsilon,\delta,m}(t,s,z,\xi)$. For n large enough, we then have that

$$\sup_{(t,s,z,\xi) \in Q_{(t_0,s_0,z_0)}} I^{\varepsilon,\delta,m,n}(t,s,z,\xi) = \sup_{(t,s,z,\xi) \in Q^n_{(t_0,s_0,z_0)}} I^{\varepsilon,\delta,m,n}(t,s,z,\xi),$$

where $Q^n_{(t_0,s_0,z_0)} := \left\{(t,s,z,\xi), |\xi - \hat{\xi}_n| \leq 1,\ (t,s,z) \in Q_{(t_0,s_0,z_0)}\right\}$ is compact. Then since $I^{\varepsilon,\delta,m,n}$ is continuous, this implies the existence of $(t_n, s_n, z_n, \xi_n) \in Q^n_{(t_0,s_0,z_0)}$ which maximises

$I^{\varepsilon,\delta,m,n}$. We also observe that $(t_n, s_n, z_n, \xi_n) \in \mathrm{int}\left(Q_{(t_0,s_0,z_0)}\right)$. Indeed, it is clear that we have for $\varepsilon \le \varepsilon_0 \wedge \varepsilon^0 \wedge \varepsilon_1 \wedge \varepsilon^\delta$

$$I^{\varepsilon,\delta,m,n}(t_n, s_n, z_n, \xi_n) \ge I^{\varepsilon,\delta,m,n}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) \ge I^{\varepsilon,\delta,m}(t^\varepsilon, s^\varepsilon, z^\varepsilon, 0) \ge -1,$$

and for $(t, s, z, \xi) \in \partial Q_{(t_0,s_0,z_0)}$, we have using (6.5.21)

$$I^{\varepsilon,\delta,m,n}(t, s, z, \xi) \le I^{\varepsilon,\delta,m}(t, s, z, \xi) + \frac{1}{n} \le -2 + \frac{1}{n} < -1, \text{ for } n > 1.$$

We then have for $n > 1$ and $\varepsilon \le \varepsilon_0 \wedge \varepsilon^0 \wedge \varepsilon_1 \wedge \varepsilon^\delta$, that by the viscosity subsolution property of $v^{\varepsilon,g}$ at the point (t_n, s_n, z_n, ξ_n) (with corresponding (t_n, s_n, x_n, y_n))

$$\min_{(i,j)\in\mathcal{I}} \left\{ k\psi^{\varepsilon,\delta,m,n} - \mathcal{L}\psi^{\varepsilon,\delta,m,n} - \widetilde{U}_1(\psi_x^{\varepsilon,\delta,m,n}),\ \Lambda_{i,j}^\varepsilon \cdot (\psi_x^{\varepsilon,\delta,m,n}, \psi_y^{\varepsilon,\delta,m,n}) \right\} \le 0. \tag{6.5.22}$$

Step 3. We now show that for ε small enough, and n large enough,

$$D^{i,j} := \Lambda_{i,j}^\varepsilon \cdot (\psi_x^{\varepsilon,\delta,m,n}, \psi_y^{\varepsilon,\delta,m,n})(t_n, s_n, x_n, y_n) > 0 \text{ for all } (i,j) \in \mathcal{I}.$$

It is easy to compute that for $(i,j) \in \mathcal{I}$, we have

$$D^{i,j} = \varepsilon^3 G^\varepsilon - E^\varepsilon - F^{\varepsilon,n},$$

with

$$G^\varepsilon := \left[\lambda^{i,j} v_z^g(t_n, s_n, z_n) - (1-\delta)(\tilde{w}^{g,m} H)_\xi(t_n, s_n, z_n, \xi_n).(e_i - e_j)\right],$$

$$\begin{aligned} E^\varepsilon :=& \lambda^{i,j}\varepsilon^7(1-\delta)(\tilde{w}_z^{g,m} H)(t_n, s_n, z_n, \xi_n) + \lambda^{i,j}\left[\varepsilon^5(\phi_z(t_n, s_n, z_n) - 4q_0|z_n - z^\varepsilon|^3)\right. \\ &\left. + \varepsilon^6(1-\delta)(\tilde{w}^{g,m} H)_\xi(t_n, s_n, z_n, \xi_n).(e_i - \mathbf{y}_z^g(t_n, s_n, z_n))\right], \end{aligned}$$

and

$$F^{\varepsilon,n} := \frac{\varepsilon}{n} \frac{f'(|\xi_n - \hat{\xi}_n|)}{|\xi_n - \hat{\xi}_n|} \left(\xi_n - \hat{\xi}_n\right) \cdot \left(e_i - e_j + \lambda^{i,j}\varepsilon^3(e_i - \mathbf{y}_z^g(t_n, s_n, z_n))\right).$$

Then by the properties of $h^{\delta,\nu}$ obtained in Lemma 6.5.6 and the estimates of Lemmas 6.5.5 and 6.5.7, we have:

$$\begin{aligned} |E^\varepsilon| \le& \lambda^{i,j}\varepsilon^5\Big[|\phi_z|(t_n, s_n, z_n) + 4q_0\,|z_n - z^\varepsilon|^3 + 2\varepsilon^2 C(t_n, s_n, z_n)(1 + |\xi_n|)\mathbf{1}_{|\xi_n| < a^\delta\xi^*} \\ &+ \varepsilon C_1(t_0, s_0, z_0) C(t_n, s_s, z_n)\left(1 + 2(1 + |\xi_n|)\frac{\nu\delta\sqrt{d}}{3\xi^*}\right)\mathbf{1}_{|\xi_n| \le a^\delta\xi^*}\Big] \\ \le& C_2(t_0, s_0, z_0)\varepsilon^5\left[1 + \varepsilon\left(1 + \nu\delta a^\delta\right) + \varepsilon^2 a^\delta\xi^*\right] \end{aligned}$$

for some functions $C_1(t_0, s_0, z_0)$, $C_2(t_0, s_0, z_0)$ which depend on $\mathbf{y}_z^g$, ϕ_z and the function C.

Similarly, recalling that $|\xi_n - \hat{\xi}_n| \le 1$, we obtain easily for some constant $C_3(t_0, s_0, z_0)$, which depends on $\mathbf{y}_z^g$

$$|F^{\varepsilon,n}| \le C_3(t_0, s_0, z_0)\frac{\varepsilon}{n}.$$

We then study G^ε. By Lemma 6.5.5(i) and (iii), we have

$$\begin{aligned} G^\varepsilon &= \lambda^{i,j} v_z^g - (1-\delta)(\tilde{w}_{\xi_i}^{g,m} - \tilde{w}_{\xi_j}^{g,m})H - (1-\delta)w^m(H_{\xi_i} - H_{\xi_j}) \\ &\ge \lambda^{i,j} v_z^g - \lambda^{i,j}(1-\delta)v_z^g - (1-\delta)Lv_z^g(1 + |\xi_n|)\left(|H_{\xi_i}| + |H_{\xi_j}|\right) \\ &\ge \lambda^{i,j} v_z^g\left(\delta - \frac{L(1 + |\xi_n|)}{\underline{\lambda}}(1-\delta)\left(|H_{\xi_i}| + |H_{\xi_j}|\right)\right). \end{aligned}$$

We then fix the value of ν of $h^{\nu,\delta}$ by $\nu := \frac{3\lambda}{8L}$, so that by Lemma 6.5.6, we have for all $0 \le i \le d$, $(1+|\xi|)H_{\xi_i}(\xi) \le \frac{\lambda\delta}{2L}$, and

$$G^\varepsilon \ge \lambda^{i,j}\delta^2 v_z^g.$$

Notice that the choice of ν only depends on $\underline{\lambda}$ and L, so that the previous constants are not affected by our choice of ν. Since the sequence (t_n, s_n, z_n) lives in a compact set and $v_z^g > 0$ and continuous, we obtain for some constant $C_4(t_0, s_0, z_0) > 0$ that

$$D^{i,j} \ge \varepsilon^3\lambda^{i,j}\delta^2 C_4(t_0,z_0,z_0) - C_2(t_0,s_0,z_0)\varepsilon^5\left[1+\varepsilon\left(1+\nu\delta a^\delta\right)+\varepsilon^2 a^\delta\xi^*\right] - C_3(t_0,s_0,z_0)\frac{\varepsilon}{n}.$$

Then for some constant $\tilde{C}$ and some $\tilde{\varepsilon}^\delta$, we have for all $\varepsilon \le \tilde{\varepsilon}^\delta$ and all $n \ge \tilde{C}\varepsilon^{-5}$ that

$$\varepsilon^3\lambda^{i,j}\delta^2 C_4(t_0,z_0,z_0) - C_2(t_0,s_0,z_0)\varepsilon^5\left[1+\varepsilon\left(1+\nu\delta a^\delta\xi^*\right)+\varepsilon^2 a^\delta\xi^*\right] - C_3(t_0,s_0,z_0)\frac{\varepsilon}{n} > 0,$$

so that $D^{i,j} > 0$.

By the arbitrariness of the pair $(i,j) \in \mathcal{I}$, we then obtain that for ε small enough and n large enough

$$\varepsilon^2 J^{\varepsilon,\delta,m,n} := \left(k\psi^{\varepsilon,\delta,m,n} - \mathcal{L}\psi^{\varepsilon,\delta,m,n} - \widetilde{U}_1(\psi_x^{\varepsilon,\delta,m,n})\right)(t_n,s_n,x_n,y_n) \le 0. \tag{6.5.23}$$

Step 4: We now estimate the remainder associated to (6.5.23). Following the calculations of Lemma 6.5.4, we have

$$\begin{aligned} J^{\varepsilon,\delta,m,n} = &-\frac{1}{2}v_{zz}^g(t_n,s_n,z_n)\left|\sigma^T(t_n,s_n)\xi_n\right|^2 - \mathcal{A}^g(\phi+\Phi^\varepsilon)(t_n,s_n,z_n) + \mathcal{R}^{\varepsilon,\delta,m,n} \\ &+\frac{1}{2}(1-\delta)\mathrm{Tr}\left[\alpha^g(\alpha^g)^T(t_n,s_n,z_n)\left(\tilde{w}^{g,m}H\right)_{\xi\xi}(s_n,z_n,\xi_n)\right], \end{aligned} \quad ,$$

where for all $0 < \varepsilon \le 1$,

$$|\mathcal{R}^{\varepsilon,\delta,m,n}| \le C^\delta\left(\varepsilon(|\xi_n| + \varepsilon|\xi_n|^2 + \varepsilon^4 + \varepsilon^3 + \frac{\varepsilon^2}{n}\right), \tag{6.5.24}$$

where C^δ is a uniform constant depending only of δ. Indeed, we know by Lemma 6.5.7 that $\tilde{w}^{g,m}H$ has the required estimates for the evaluation of the remainder estimate in Lemma 6.5.4 (it is easy to do the correspondance between $\tilde{w}^{g,m}h^{\delta,\nu}$ and $\tilde{w}^{g,m}H$). Then by Lemma 6.5.4 and using the fact that the quantity $\varepsilon\xi = y - \mathbf{y}^g$ is bounded on the ball $B_{r_0/2}(t_0,s_0,z_0)$, we see that, uniformly in m and δ, we have

$$\mathfrak{H}^\varepsilon(t,s,z,\xi) \le |v_z^g| + \varepsilon^2\left(|\phi_z| + |\Phi_z^\varepsilon| + \varepsilon^2\left(|\tilde{w}_z^{g,m}| + \frac{1}{\varepsilon}|\mathbf{y}_z^g|\left|\left(\tilde{w}^{g,m}H\right)_\xi\right| + \frac{1}{n\varepsilon^3}|\mathbf{y}_z^g|K\right)\right),$$

where $\mathfrak{H}^\varepsilon$ was introduced in the proof of Lemma 6.5.4, and K_1 is a uniform constant. Then the quantity $\zeta^\varepsilon(t,s,z,\xi)$ is uniformly bounded on $B_{r_0/2}(t_0,s_0,z_0)$, uniformly in m, but not in δ. Similarly, we obtain easily that the quantity $\varepsilon^4\mathfrak{R}^\varepsilon\left((\tilde{w}^{g,m}H) + \frac{1}{\varepsilon^2 n}f\right)$ is bounded by some constant depending only on (t_0,s_0,z_0) and δ. Hence we have for some function $\tilde{K}^\delta$ depending on $\mathbf{y}^g$ (and its derivatives), the constant $C(t,s,z)$ introduced in Lemma 6.5.5, K_1 and δ that, uniformly in m:

$$\left|\mathcal{R}^{\varepsilon,\delta,m,n}\right| \le \tilde{K}^\delta(t_0,s_0,z_0)(\varepsilon|\xi_n| + \varepsilon^2|\xi_n|^2) + \varepsilon^4\tilde{K}^\delta(t_0,s_0,z_0)(1 + \frac{1}{\varepsilon} + \frac{1}{n\varepsilon^2}).$$

We then need to show that (ξ_n) remains bounded as ε goes to 0. Indeed, we have from Lemma 6.5.6 and Lemma 6.5.5:

$$\left|(\tilde{w}^{g,m}H)_{\xi\xi}\right|(t,s,z,\xi) \leq C(t,s,z)\left(1+2\frac{\nu\delta\sqrt{d}}{3\xi^*}+\frac{2C^*}{(\xi^*)^2}(1+|\xi|)\right)\mathbf{1}_{|\xi|\leq a^\delta\xi^*} \leq \check{C}(t,s,z),$$

where $\check{C}$ is a continuous functions depending on the function C appearing in Lemma 6.5.5 and ξ^*. Now since we only consider elements $(t,s,z)\in Q_{(t_0,s_0,z_0)}$ compact, we have by (6.5.23):

$$\begin{aligned}(-v^g_{zz}(t_n,s_n,z_n))\frac{\left|\sigma^T\xi_n\right|^2}{2}+\mathcal{R}^{\varepsilon,\delta,m,n} &\leq \left(\mathcal{A}^g(\phi+\Phi^\varepsilon)+\left|\alpha^g(\alpha^g)^T\right|\check{C}\right)(t_n,s_n,z_n)\\ &\leq \sup_{(t,s,z)\in Q_{(t_0,s_0,z_0)}}\left(\mathcal{A}^g(\phi+\Phi^\varepsilon)+\left|\alpha^g(\alpha^g)^T\right|\check{C}\right)\\ &\leq \text{Const}.\end{aligned}$$

Then with (6.5.24), we obtain for

$$\check{C}_1 := \sup_{(t,s,z)\in Q_{(t_0,s_0,z_0)}}\check{C}(t,s,z) \text{ and } C_0 := \inf_{(t,s,z)\in Q_{(t_0,s_0,z_0)}}(-v^g_{zz})(t,s,z) > 0,$$

since v^g is strictly concave in z and smooth, that

$$C_0|\xi_n|^2 - C^\delta\left(\varepsilon(|\xi_n|+\varepsilon|\xi_n|^2+\varepsilon^4+\varepsilon^3+\frac{\varepsilon^2}{n}\right)\leq \check{C}_1.$$

Assume next that $|\xi_n|$ goes to ∞ (up to a subsequence) as ε goes to 0. Then, the left-hand side above would go to ∞ which contradict the fact that it is bounded. Then the sequence (ξ_n) is bounded by some $\hat{\xi}^\delta$, depending only on δ, and not on m since we then know that $\mathcal{R}^{\varepsilon,\delta,m,n}\longrightarrow 0$ when $\varepsilon\longrightarrow 0$ and $n\longrightarrow\infty$, uniformly in $m\in(0,1]$.

Step 5: Following [81], we show that $|\xi_n|\leq\xi^*$ for n large enough (uniform in m, but not in δ), where ξ^* was introduced in Step 1 of the proof. We consider now n large enough and ε small enough so that for all $m\in(0,1]$, $|\mathcal{R}^{\varepsilon,\delta,m,n}|\leq 1$. Assume on the contrary that $|\xi_n|>\xi^*$. By our choice of ξ^*, we know that $\tilde{w}^{g,m}_{\xi\xi}(\cdot,\xi_n)=0$ on $Q_{(t_0,s_0,z_0)}$. In the following we omit the dependance in the parameters which will be at any step (t_n,s_n,z_n,ξ_n). From (6.5.23), we have, using in particular the fact that $\xi^*\geq 1$

$$\begin{aligned}-\frac{1}{2}v^g_{zz}\left|\sigma^T\xi_n\right|^2-\mathcal{A}^g(\phi+\Phi^\varepsilon) &\leq -\frac{1-\delta}{2}\text{Tr}\left[\alpha^g(\alpha^g)^T(\tilde{w}^{g,m}H)_{\xi\xi}\right]-\mathcal{R}^{\varepsilon,\delta,m,n}\\ &\leq \text{Tr}\left[\alpha^g(\alpha^g)^T\right]C(t_n,s_n,z_n)\left(C^*+\frac{\sqrt{d}\delta\underline{\lambda}}{4L}\right)+1,\end{aligned}$$

where we used Lemmas 6.5.5 and 6.5.6. Furthermore, we remind the reader that the function C is the one introduced in Lemma 6.5.5(ii), and the constant C^* is the one introduced in Lemma 6.5.6.

This last inequality is in contradiction with (6.5.19), so that we actually have $|\xi_n|\leq\xi^*$. Now since (ξ_n) is a bounded sequence, we have by classical results of the theory of viscosity solution that, up to extraction, there exists $\bar{\xi}^{\delta,m}$, with $\left|\bar{\xi}^{\delta,m}\right|\leq 1$, such that $(t_n,s_n,z_n,\xi_n)\to(t_0,s_0,z_0,\bar{\xi}^{\delta,m})$ when $n\to\infty$. Recalling that $H=1$ on $\bar{B}_{\xi^*}(0)$, and H is C^2, we obtain by (6.5.23), (6.5.24) and Lemma 6.5.5(iv) that at the point $(t_0,s_0,z_0,\bar{\xi}^{\delta,m})$:

$$\begin{aligned}0 &\geq -\frac{1}{2}v^g_{zz}\left|\sigma^T\bar{\xi}^{\delta,m}\right|^2+\frac{1-\delta}{2}\text{Tr}\left[\alpha^g(\alpha^g)^T\tilde{w}^{g,m}_{\xi\xi}\right]-\mathcal{A}^g\phi\\ &\geq -\mathcal{A}^g\phi+(1-\delta)a^g-\delta v^g_{zz}\frac{|\sigma^T\bar{\xi}^{\delta,m}|^2}{2}+\frac{1-\delta}{2}\int_{\mathbb{R}^d}v^m(\zeta)\left|\sigma\zeta\right|^2d\zeta.\end{aligned}$$

Now since $|\bar{\xi}^{\delta,m}|$ is bounded by ξ^* independent of δ and m, we obtain sending δ and m to 0 that $\mathcal{A}^g\phi(t_0, s_0, z_0) - a^g(t_0, s_0, z_0) \geq 0$, which contradicts (6.5.16). □

Bibliography

[1] Akian, M., Menaldi, J.L., Sulem, A. (1996). On an investment-consumption model with transaction costs, *SIAM Journal on Control and Optimization*, 34(1):329–364. (Cited on page 99.)

[2] Akian, M., Sequier, P., Sulem, A. (1995). A finite horizon multidimensional portfolio selection problem with singular transactions, *Proceedings of the 34th IEEE Conference on Decision and Control*, 3:2193–2198. (Cited on page 99.)

[3] Altarovici, A., Muhle-Karbe, J., Soner, H.M. (2013). Asymptotics for fixed transaction costs, preprint, *arXiv:1306.2802*. (Cited on page 16.)

[4] Ambrosio, L. (2000)*Lecture Notes on Optimal Transport Problems*, Lecture Notes in Mathematics (Cited on page 2.)

[5] Atkinson, C., Mokkhavesa, S. (2004). Multi-asset portfolio optimization with transaction cost, *Applied Mathematical Finance*, 11:95–123. (Cited on page 16.)

[6] Avellaneda, M., Lévy, A., Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities, *Applied Mathematical Finance*, 2:73–88. (Cited on pages 11 and 39.)

[7] Baker, D. (2012). Martingales with specified marginals, PhD Thesis. (Cited on pages 32, 33, 35 and 37.)

[8] Barles, G., Soner, H.M. (1998). Option pricing with transaction costs and a nonlinear Black- Scholes equation, *Finance and Stochastics*, 2:369–397. (Cited on page 16.)

[9] Beiglböck, M., Henry-Labordère, P., Penkner, F. (2013). Model-independent Bounds for Option Prices: A Mass-Transport Approach, preprint arXiv:1106.5929, Finance and Stochastics, to appear. (Cited on page 5.)

[10] Beiglböck, M., Juillet, N. (2012). On a problem of optimal transport under marginal martin- gale constraints, preprint, arXiv:1208.1509. (Cited on pages 5 and 30.)

[11] Bertsekas, D. P., Shreve, S. E. (1978). Stochastic Optimal Control. The Discrete- Time Case. Academic Press, New York. (Cited on page 63.)

[12] Biagini, S. and Frittelli M. (2005). Utility Maximisation in Incomplete Markets for Unbounded Processes, Finance and Stochastics, Vol 9-4, pp. 493-517. (Cited on page 13.)

[13] Biagini, S. and Frittelli M. (2007). The supermartingale property of the optimal wealth process for general semimartingales, Finance and Stochastics, Vol 11-2, pp. 253-266. (Cited on page 13.)

[14] Biagini, S., Frittelli M. and Grasselli, M. (2011). Indifference Price with General Semimartingales, Mathematical Finance, Vol. 21, No. 3, pp. 423-446. (Cited on page 13.)

[15] Bichteler, K. (1981). Stochastic integration and L^p-theory of semimartingales. *The Annal of Probability*, 9(1): 19-89. (Cited on page 11.)

[16] Bichuch, M. (2011). Pricing a contingent claim liability with transaction costs using asymptotic analysis for optimal investment, *Finance and Stochastics*, to appear, *arXiv:1112.3012*. (Cited on pages 16, 17, 95, 96, 101, 108, 109 and 111.)

[17] Bichuch, M. (2012). Asymptotic analysis for optimal investment in finite time with transaction costs, *SIAM Journal on Financial Mathematics*, 3:433–458. (Cited on page 16.)

[18] Bichuch, M., Shreve, S.E. (2013). Utility maximization trading two futures with transaction costs, *SIAM Journal on Financial Mathematics*, 4:26–85. (Cited on page 16.)

[19] Billingsley, P. (1999). *Convergence of Probability Measures*, Wiley Series in Probability and Statistics. (Not cited.)

[20] Bordigoni, G., Matoussi, A., and Schweizer, M. (2007). A stochastic control approach to a robust utility maximization problem, "Stochastic analysis and applications", Abel Symp, 2:125-151. Springer, Berlin. (Cited on page 13.)

[21] Bouchard, B. (2000). Option pricing via utility maximization in the presence of transaction costs: an asymptotic analysis, *Cahiers de Mathématiques du CEREMADE*, 2000-0006, *http://basepub.dauphine.fr/handle/123456789/6707*. (Cited on page 16.)

[22] Bouchard, B., Elie, R., Imbert, C. (2009). Optimal control under stochastic target constraints, *SIAM Journal on Control and Optimization*, 48(5):3501–3531. (Cited on page 112.)

[23] Bouchard, B., Kabanov, Yu.M., Touzi, N. (2001). Option pricing by large risk aversion utility under transaction costs, *Decisions in Economics and Finance*, 24:127–136. (Cited on page 16.)

[24] Bouchard, B., Moreau, L., Nutz, M. (2012) Stochastic Target Games with Controlled Loss, Annals of Applied Probability, to appear. (Cited on page 13.)

[25] Bouchard, B., Moreau, L., Soner, H.M. (2013). Hedging under an expected loss constraint with small transaction costs, preprint, *arXiv:1309.4916*. (Cited on pages 16, 95, 96, 101, 110, 111 and 112.)

[26] Bouchard, B., Touzi, N. (2011). Weak Dynamic Programming Principle for Viscosity Solutions. *SIAM Journal on Control and Optimization*, 49, 3, 948-962. (Cited on pages 83 and 84.)

[27] Breeden, D.T., Litzenberger, R.H. (1978). Prices of state-contingent claims implicit in options prices, J. Business, 51, 621651. (Cited on pages 8 and 24.)

[28] Carlier, G. (2012) *Optimal Transportation and Economic Applications*. Lecture Notes. (Cited on page 2.)

[29] Carr, P. and Chou, A. (1997). Breaking barriers, Risk 10,9, 139-145. (Cited on page 8.)

[30] Chen, X., Dai, M. (2013). Characterization of optimal strategy for multi-asset investment and consumption with transaction costs, *SIAM Journal on Financial Mathematics*, to appear, $http://papers.ssrn.com/sol3/papers.cfm?abstract\_id = 2175413$. (Cited on page 16.)

[31] Constantinides, G.M., Magill, M.J.P. (1976). Portfolio selection with transaction costs, J. Econ. Theory, 49:33-83. (Cited on page 2.)

[32] Constantinides, G.M., Zariphopoulou, T. (1999). Bounds on prices of contingent claims in an inter-temporal economy with proportional transaction costs and general preferences, *Finance and Stochastics*, 3:345–369. (Cited on page 16.)

[33] Constantinides, G.M., Zariphopoulou, T. (2001). Bounds on derivative prices in an inter-temporal setting with proportional transaction costs and multiple securities, *Mathematical Finance*, 11:331–346. (Cited on page 16.)

[34] Davis, M.H.A., Clark, J.M.C. (1994). A note on super-replicating strategies, *Philosophical Transactions: Physical Sciences and Engineering*, 347(1684):485–494. (Cited on page 2.)

[35] Davis, M.H.A., Panas, V.G., Zariphopoulou, T. (1993). European option pricing with transaction costs, *SIAM Journal on Control and Optimization*, 31(2):470–493. (Cited on pages 16 and 109.)

[36] Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M. and Stricker, C. (2001) Exponential Hedging and Entropic Penalties, Mathematical Finance 12, pp. 99-123. (Cited on page 13.)

[37] Dellacherie, C., Meyer, P.A. (1982). Probabilities and potential B, *North Holland, Amsterdam*. (Cited on pages 43 and 49.)

[38] Denis, L., and Kervarec, M. (2007). Utility functions and optimal investment in non-dominated models, preprint. (Cited on page 13.)

[39] Denis, L., Martini, C. (2006) *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, *Annals of Applied Probability*, 16(2):827–852. (Cited on pages 6, 11 and 39.)

[40] Dolinsky, Y. and Soner, H.M. (2012). Martingale optimal transport and robust hedging in continuous time, Probability Theory and Related Fields, to appear. (Cited on pages 6 and 7.)

[41] Dupire, B. Pricing with a smile, Risk, 7 (1994), pp. 18-20. (Cited on page 9.)

[42] El Karoui, N. and Tan, X. (2013) Capacities, measurable selection and dynamic programming Part I: abstract framework, preprint. (Cited on page 12.)

[43] El Karoui, N. and Tan, X. (2013) Capacities, measurable selection and dynamic programming Part II: application in stochastic control problems, preprint. (Cited on page 12.)

[44] Fleming, W.H., Soner, H.M. (1986). *Controlled Markov Processes and Viscosity Solutions.* Applications of Mathematics 25. Springer-Verlag, New York. (Cited on page 85.)

[45] Föllmer, H. (1981). Calcul d'Itô sans probabilités. *Séminaire de Probabilités*, XV: 09,143-150, Lecture Notes in Mathematics 850. (Cited on pages 6 and 11.)

[46] Galichon, A., Henry-Labordère, P., Touzi, N. (2013). A stochastic control approach to no-arbitrage bounds given marginals, with an application to Lookback options, *Annals of Applied Probability*, to appear. (Cited on pages iii, v, 7, 12, 40, 43, 56 and 60.)

[47] Gerhold, S., Muhle-Karbe, J., Schachermayer, W. (2012). Asymptotics and duality for the Davis and Norman problem, *Stochastics*, 84:625–641. (Cited on page 16.)

[48] Gilboa, I. and Schmeidler, D. (1989). Maximin expected utility with a non-unique prior. Journal of Mathematical Economics, 18: 141-153. (Cited on page 13.)

[49] Goodman, J., Ostrov, D.N. (2010). Balancing small transaction costs with loss of optimal allocation in dynamic stock trading strategies, *SIAM Journal on Applied Mathematics*, 70:1977–1998. (Cited on page 16.)

[50] Henderson, V., Hobson, D. (2008). An explicit solution for an optimal stopping/optimal control problem which models an asset sale. *Annals of Applied Probability* 18(5), 1681-1705. (Cited on pages 14, 15, 75, 76, 79 and 80.)

[51] Henry-Labordère, P., Oblój, J., Spoida, P., Touzi, N. (2012). Maximum maximum of martingales given finitely-many intermediate marginals, preprint. (Cited on pages 7, 40, 43 and 60.)

[52] Henry-Labordère, P. and Touzi, N. (2013). An explicit Version of the Brenier Theorem in Martingale Optimal Transportation. Preprint. (Cited on page 5.)

[53] Hirsch, F., Profeta, C., Roynette, B. and Yor, M. (2011) Constructing self-similar martingales via two Skorokhod embeddings, *Séminaire de Probabilités* XLIII, Lecture Notes in Mathematics 2011, pp 451-503. (Cited on page 9.)

[54] Hirsch, F., Profeta, C., Roynette, B. and Yor, M. (2011) *Peacocks and Associated Martingales, with Explicit Constructions*, Springer. (Cited on page 9.)

[55] Hirsch, F. and Roynette, B. (2012) A new proof of Kellerer's theorem. ESAIM: PS 16, p48-60. (Cited on page 9.)

[56] D. Hobson, The skorokhod embedding problem and model-independent bounds for option prices, Paris-Princeton Lectures on Mathematical Finance, Lecture Notes in Mathematics, Springer. (Cited on page 89.)

[57] Hobson, D., Pedersen, J. (2002). The minimum maximum of a continuous martingale with given initial and terminal laws, *Annals of probability, pp. 978-999. (Cited on page 9.)*

[58] Hodges, S., Neuberger, A. (1989). Option replication of contingent claims under transaction costs, Review of Futures Markets, 8:222–239. (Cited on page 2.)

[59] Janeček, K., Shreve, S.E. (2004). Asymptotic analysis for optimal investment and consumption with transaction costs, Finance and Stochastics, 8:181–206. (Cited on page 16.)

[60] Janeček, K., Shreve, S.E. (2010). Futures trading with transaction costs, Ilinois Journal of Mathematics, 54(4):1239–1284. (Cited on page 16.)

[61] Kabanov, Yu.M., Safarian, M. (2009). Markets with transaction costs, Springer-Verlag. (Cited on page 99.)

[62] Kabanov, Y., Stricker, C. (2002). On the optimal portfolio for the exponential utility maximization: remarks to the six-author paper. Mathematical Finance, 12, 2,125-134 (Cited on page 13.)

[63] Kallsen, J., Muhle-Karbe, J. (2012). Option pricing and hedging with small transaction costs, Mathematical Finance, to appear, arXiv:1209.2555. (Cited on pages 16 and 111.)

[64] Kallsen, J., Muhle-Karbe, J. (2013). The general structure of optimal investment and consumption with small transaction costs, preprint, arXiv:1303.3148. (Cited on pages 16 and 111.)

[65] Karandikar, R. (1995). On pathwise stochastic integration, Stochastic Processes and Their Applications, 57:11–18. (Cited on pages 6, 11, 40, 49 and 55.)

[66] Kazi-Tani, N., Possamaï, D., Zhou, C. (2012). Second order BSDEs with jumps, part II: existence and applications, preprint, . (Cited on pages 44 and 65.)

[67] Kellerer, H.G. (1972). Markov-Komposition und eine Anwendung auf Martingale. Math. Ann., p99-122 (Cited on page 9.)

[68] Levental, S., Skorohod, A.V. (1997). On the possibility of hedging options in the presence of transaction costs, Annals of Applied Probability, 7:410–443. (Cited on page 2.)

[69] Lyons, T.J. (1995). Uncertain volatility and the risk-free synthesis of derivatives, Applied Mathematical Finance 2(2):117–133. (Cited on pages 11 and 39.)

[70] Matoussi, A., Possamaï, D. and Zhou, C. (2013). Robust utility maximization in non-dominated models with 2BSDEs, Mathematical Finance, to appear. (Cited on page 13.)

[71] Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time case, Journal of Economic Theory, 3:373-413. (Cited on pages 2 and 112.)

[72] Moreau, L., Muhle-Karbe, J., Soner, H.M. (2013). Portfolio choice with small market impact, preprint. (Cited on page 16.)

[73] Morters, P., Peres, Y. :Brownian motion, Cambridge Series in Statistical and Probabilistic Mathematics, Volume 30. (Cited on page 32.)

[74] Neufeld, A., Nutz, M. (2013). Superreplication under volatility uncertainty for measurable claims, Electronic Journal of Probability, Vol. 18, No. 48, pp. 1-14. (Cited on pages iii, v, 12, 39, 40, 42, 43, 44, 45, 46, 63, 64, 66 and 67.)

[75] Neveu, J. (1975). Discrete-Parameter Martingales, North-Holland, Amsterdam. (Not cited.)

[76] Nutz, M. (2012). Pathwise Construction of Stochastic Integrals, Electronic Communications in Probability, Vol. 17, No. 24, pp. 1-7. (Cited on pages 6 and 11.)

[77] Nutz, M., van Handel, R. (2012). Constructing sublinear expectations on path space, preprint, Stochastic Processes and their Applications, Vol. 123, No. 8, pp. 3100–3121. (Cited on pages 12, 39, 42, 43, 46, 50, 63 and 64.)

[78] Pham, T. and Zhang, J. (2012) Two Person Zero-sum Game in Weak Formulation and Path Dependent Bellman-Isaacs Equation, preprint, arXiv:1209.6605. (Cited on page 69.)

[79] Pliska, S.R. (1986). A stochastic calculus model of continuous trading: optimal portfolios, Math. Operations Research, 11:371-382. (Cited on page 2.)

[80] *Possamaï, D., Soner, H.M., Touzi, N. (2012). Large liquidity expansions of superhedging costs, Asymptotic Analysis, 79(1-2):45–64. (Cited on page 111.)*

[81] *Possamaï, D., Soner, H.M., Touzi, N. (2013). Homogenization and asymptotics for small transaction costs: the multidimensional case, preprint, arXiv:1212.6275. (Cited on pages 16, 17, 100, 101, 105, 113, 114, 118, 124, 126, 127 and 132.)*

[82] *Protter, P. (2004). Stochastic integration and differential equations, Springer-Verlag. (Cited on pages 49 and 67.)*

[83] *Rogers, L.C.G. (1994). Equivalent martingale measures and no-arbitrage. Stochastics and Stochastics Reports 51, 41-49. (Cited on pages iii, v, 10, 11, 21, 23, 25 and 26.)*

[84] *Rogers, L.C.G., (2004). Why is the effect of proportional transaction costs $O(\delta^{2/3})$?, in Mathematics of Finance, Contemp. Math. 351, G. Yin and Q. Zhang, eds., AMS, Providence, RI, 303–308. (Cited on page 16.)*

[85] *Schachermayer, W. (2004). Utility Maximisation in Incomplete Markets, Springer Lecture Notes in Mathematics, Vol. 1856, pp. 225-288. (Cited on page 13.)*

[86] *Shreve, S.E., Soner, H.M. (1991). A free boundary problem related to singular stochastic control: parabolic case, Communications in Partial Differential Equations, 16:373–424. (Cited on page 16.)*

[87] *Shreve, S.E., Soner, H.M. (1994). Optimal investment and consumption with transaction costs, Annals of Applied Probability, 4:609–692. (Cited on page 16.)*

[88] *Shreve, S.E., Soner, H.M., Cvitanić, J. (1995). There is no nontrivial hedging portfolio for option pricing with transaction costs, Annals of Applied Probability, 5(2):327–355. (Cited on page 2.)*

[89] *Soner, H.M., Shreve, S.E. (1989). Regularity of the value function of a two-dimensional singular stochastic control problem, SIAM Journal on Control and Optimization, 27:876–907. (Cited on page 16.)*

[90] *Soner, H.M., Touzi, N. (2013). Homogenization and asymptotics for small transaction costs, SIAM Journal on Control and Optimization, 51(4):2893–2921. (Cited on pages iii, v, 16, 17, 95, 96, 100, 101, 102, 108, 118 and 124.)*

[91] *Soner, H.M., Touzi, N., Zhang, J. (2011). Wellposedness of second-order BSDEs, Probability Theory and Related Fields, 153(1-2):149–190. (Cited on pages 7, 39, 49 and 55.)*

[92] *Soner, H.M., Touzi, N., Zhang, J. (2011). Quasi-sure stochastic analysis through aggregation, Electronic Journal of Probability, 16(67):1844–1879. (Cited on pages 7, 41, 43 and 49.)*

[93] *Soner, H.M., Touzi, N., Zhang, J. (2013). Dual formulation of second order target problems, The Annals of Applied Probability 23, Number 1, 308-347. (Cited on pages 7 and 50.)*

[94] *Strassen, V. (1965). The existence of probability measures with given marginals. Ann. Math. Statist., 36:423439. (Cited on pages iii, v, 9 and 21.)*

[95] Stroock, D.W., Varadhan, S.R.S. (1979). Multidimensional diffusion processes, Springer. (Cited on pages 12, 41 and 64.)

[96] Tan, X. and Touzi, N. (2013). Optimal transportation under controlled stochastic dynamics, Annals of Probability 2013, Vol. 41, No 5,3201-3240. (Not cited.)

[97] Tevzadze, R., Toronjadze, T., Uzunashvili, T. (2012). Robust utility maximization for diffusion market model with misspecified coefficients, Finance and Stochastics, to appear. (Cited on page 13.)

[98] Touzi, N. (2012). Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. Fields Institute Monographs, Vol. 29. (Cited on page 85.)

[99] Villani, C. (2003). Topics in Optimal Transportation, Graduate Studies in Mathematics series, Vol. 58, American Mathematical Society. (Cited on page 2.)

[100] Villani, C. (2009). Optimal Transport. Old and New, Volume 338 of Grundlehren der Mathematischen Wissenschaften, Springer. (Cited on pages 2, 3 and 4.)

[101] Whalley, A.E., Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing under transaction costs, Mathematical Finance 7:307–324. (Cited on page 16.)