THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Couverture d'options dans un marché avec impact et schémas numériques pour les EDSR basés sur des systèmes de particules

École Doctorale de Dauphine — ED 543

Spécialité **Sciences**

COMPOSITION DU JURY :

M. Bruno BOUCHARD
Université Paris-Dauphine
Directeur de thèse

M. Frédéric ABERGEL
CentraleSupélec
Président du jury

M. Dirk BECHERER
Humboldt-Universität zu Berlin
Rapporteur

M. Huyên PHAM
Université Paris-Diderot
Membre du jury

M. Xiaolu TAN
Université Paris-Dauphine
Membre du jury

M. Nizar TOUZI
Ecole Polytechnique
Rapporteur

Soutenue le 09.10.2017
par Yiyi ZOU

Dirigée par **Bruno BOUCHARD**

DAUPHINE UNIVERSITÉ PARIS | PSL RESEARCH UNIVERSITY PARIS

Acknowledgments

Firstly, I would like to express my profuse and sincere gratitude to my supervisor Prof. Bruno Bouchard, for his continuous guidance during my PhD study and for his infinite patience. The research reported in this dissertation would not have been possible without his support and help. I could not have imagined having better and friendly advisor to complete this three years journey.

I would also like to thank Prof. Nizar Touzi and Prof. Dirk Becherer for reviewing this dissertation. I also extend my thanks to the rest of my PhD committee: Prof. Frédéric Abergel, Prof. Huyên Pham and Dr. Xiaolu Tan, for accepting to be part of the jury. My thanks also to my collaborator, Dr. Grégoire Loeper, for the inspiring research topic on market impact and liquidity cost he has shared with us, as well as the many constructive discussions he has brought.

I am extremely grateful to the École Doctoral de Dauphine and to CEREMADE, for the financial support they have provided all through my PhD study and for their friendliness. I am also grateful to the LFA team in CREST, to Prof. Jean-Michel Zokoian without whom I would not have been able to spend the three years with lovely colleagues and friends.

Last but not the least, I would like to thank my family: my parents and my husband for supporting me spiritually throughout my PhD and my life in general.

Contents

Acknowledgments i

1 Présentation générale **1**
1.1 Modèles avec impact 1
1.1.1 Motivation générale 1
1.1.2 Description des dynamiques 2
1.1.3 Un-covered options 3
1.1.4 Covered options 5
1.2 BDE Branching 7
1.2.1 Motivation et problème lié aux polynômes 7
1.2.2 Algorithme 8
1.2.3 Contribution 8
1.3 Publications 9

2 Introduction **11**
2.1 Impact model 11
2.1.1 General motivation 11
2.1.2 Dynamics description 12
2.1.3 Un-covered options 13
2.1.4 Covered options 15
2.2 BDE Branching 17
2.2.1 Motivation and Problems related to polynomials 17
2.2.2 Algorithm 17
2.2.3 Contribution 18
2.3 Publications 19

3 Almost-sure hedging with permanent price impact **21**
3.1 Portfolio and price dynamics 23
3.1.1 Impact rules 23
3.1.2 Discrete rebalancing from a continuous signal and continuous time trading limit 24
3.1.3 Jumps and large orders splitting 28
3.2 Super-hedging of a European claim 31
3.2.1 Super-hedging price 32
3.2.2 Dynamic programming 33
3.2.3 Pricing equation 34
3.2.4 An example: the fixed impact case 37
3.2.5 Proof of the pde characterization 37
3.3 Appendix 45

4 Hedging of covered options with linear market impact and gamma constraint **47**
4.1 Model and hedging problem 48

4.1.1 Continuous time trading dynamics 48
4.1.2 Hedging equation and gamma constraint 50
4.2 Viscosity solution characterization 52
4.2.1 A sequence of smooth supersolutions 53
4.2.2 Supersolution property for the weak formulation 63
4.2.3 Conclusion of the proof and construction of almost optimal strategies 67
4.3 Adding a resilience effect 68
4.4 Numerical approximation and examples 69
4.4.1 Finite difference scheme 69
4.4.2 Numerical examples: the fixed impact case 70
4.5 Appendix 72

5 Numerical approximation of BSDEs using local polynomial drivers and branching processes 73
5.1 Introduction 73
5.2 Numerical method for a class of BSDE based on branching processes 74
5.2.1 Local polynomial approximation of the generator 75
5.2.2 Picard iteration with doubly reflected BSDEs 76
5.2.3 A branching diffusion representation for $\bar{Y}^m$ 78
5.2.4 The numerical algorithm 79
5.3 Example of application 82
5.4 Appendix 86
5.4.1 Technical lemmas 86
5.4.2 More on the error analysis for the abstract numerical approximation 89

List of Figures 91

Bibliography 91

Chapter 1

Présentation générale

Contents

1.1 Modèles avec impact **1**
1.1.1 Motivation générale 1
1.1.2 Description des dynamiques 2
1.1.3 Un-covered options 3
1.1.4 Covered options 5
1.2 BDE Branching **7**
1.2.1 Motivation et problème lié aux polynômes 7
1.2.2 Algorithme 8
1.2.3 Contribution 8
1.3 Publications **9**

1.1 Modèles avec impact

1.1.1 Motivation générale

La théorie classique de la valorisation des produits dérivés se base sur trois hypothèses principales: marché complet, absence de coût de transaction et liquidité parfaite. À partir de ces hypothèses, la réplication dynamique fournit une façon intuitive de valoriser ces produits dérivés dans un marché sans arbitrage. Empiriquement, ces hypothèses sont rarement vérifiées dans le marché, en particulier pour les transactions dont le notionel est important et pour les actifs peu liquides.

De nombreux travaux sont dédiés au relâchement de ces hypothèses. Certains traitent le coût de liquidité sans prendre en compte l'impact permanent sur le sous-jacent, voir par exemple [17]. Le prix du sous-jacent, fonction du volume de transaction instantané, ne dépend pas de l'historique de la négociation. Comme la transaction peut (à la limite) être exécutée au prix marginal, l'introduction du coût de liquidité n'affecte pas le prix de couverture. Lorsqu'une contrainte de liquidité supplémentaire est imposée, par exemple en imposant une borne sur le "gamma" de la stratégie, l'équation de valorisation présente alors une forme quadratique supplémentaire, voir [18]. D'autres travaux se concentrent sur la forme analytique de la dynamique des prix en établissant l'équilibre de l'offre et de la demande, sans tenir compte du coût de liquidité: par exemple [25], [45] et [46]. Le prix du sous-jacent est ainsi modifié par les activités de différents joueurs dans le marché. Cet impact génère une non-linéarité dans l'équation de valorisation du dérivé, qui provient toutefois d'un terme de volatilité modifié.

Plus récemment, [1] et [37] ont introduit une nouvelle approche dans laquelle la dynamique des prix se compose d'un processus de diffusion classique et d'un terme à l'impact linéaire. Ce modèle intègre à la fois le coût de liquidité et l'impact sur le prix du sous-jacent, et ce de manière très naturel. Lorsque la solution d'une certaine équation non-linéaire est régulière, ils montrent qu'une réplication parfaite est possible. Cependant, leur déduction de l'équation de valorisation ainsi que de la stratégie de réplication s'appuit sur une approche par vérification, qui non seulement impose des conditions de régularité fortes, mais qui en outre ne permet pas de montrer que prix de sur-réplication et prix de réplication coïncident. Ainsi notre problématique est, sans présumer d'aucune condition de régularité sur la fonction de prix du produit dérivé, d'établir l'équation de valorisation en suivant l'approche par cible stochastique classique. Comme nous le verrons, ceci nécessite des modifications importantes des approches par cible stochastique classiques.

1.1.2 Description des dynamiques

Nous modélisons l'impact permanent de la négociation sur le prix du sous-jacent par la fonction d'impact f: la variation du prix du sous-jacent dûe à la négociation de $\delta \in \mathbb{R}$, en nombre d'unités d'actif risqué, est $\delta f(x)$, où x est le prix du sous-jacent avant la transaction. Le coût pour acheter un tel nombre d'actifs est

$$\delta x + \frac{1}{2}\delta^2 f(x) = \delta \int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota,$$

où la quantité

$$\int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota$$

est le coût moyen pour acheter chaque unité d'actif.

La dynamique du sous-jacent entre deux transactions consécutives est décrite par une diffusion

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t.$$

Nous appelons une stratégie de négociation le processus définit comme

$$Y = Y_{0-} + \int_0^\cdot b_s ds + \int_0^\cdot a_s dW_s + \int_0^\cdot \int \delta\nu(d\delta, ds),$$

où a et b sont des processus prévisibles à valeurs dans $\mathbb{R}$ satisfaisant

$$(a, b) \in \mathcal{A} := \cup_k \mathcal{A}_k,$$
$$\mathcal{A}_k := \{(a, b) \text{ prévisible à valeurs dans } \mathbb{R} : |(a, b)| \leq k \; dt \times d\mathbb{P} - \text{a.e.}\},$$

et $\nu \in \mathcal{U} := \cup_k \mathcal{U}_k$, est un contrôle dans l'espace des mesures positives à valeurs dans $\mathbb{N}$. Ici, $\mathcal{U}_k$ est l'ensemble des mesures aléatoires à valeurs dans $\{0, \cdots, k\}$ avec support $[-k, k] \times [0, T]$.

Le processus Y doit être considéré comme un signal de négociation, l'objectif du trader étant d'avoir en permanence Y unités d'actif risqué dans son portefeuille. Le composant ν modélise les sauts du nombre d'actif risqué au moment d'établir un delta initial quand on en a besoin, ou au moment de liquider le portefeuille à la maturité (on verra qu'entre ces deux dates, la stratégie ne saute pas).

Nous déduisons les dynamiques du sous-jacent X et de la richesse totale V (la somme du cash et de la valeur des actions) par la condition d'autofinancement. Ceci est d'abord réalisé en considérant que l'on passe un ordre de taille $Y_{t_{i+1}} - Y_{t_i}$ aux dates t_i, avant de passer à la limite quand la fréquence de balancement tend vers l'infini, i.e. $t_{i+1} - t_i \to 0$. De ce point de vue, notre modèle est issu d'une formulation à partir de stratégies simples.

Proposition 1.1.1. *Soient $(a, b, \nu) \in \mathcal{A} \times \mathcal{U}$, alors les dynamiques de X et de V correspondant à des rebalancements en continu sont*

$$\begin{aligned} X &= X_{0-} + \int_0^{\cdot} \sigma(X_s) dW_s + \int_0^{\cdot} f(X_s) dY_s^c + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s)) ds \\ &\quad + \int_0^{\cdot} \int \Delta \mathrm{x}(X_{s-}, \delta) \nu(d\delta, ds) \\ V &= V_{0-} + \int_0^{\cdot} Y_s dX_s^c + \frac{1}{2} \int_0^{\cdot} a_s^2 f(X_s) ds \\ &\quad + \int_0^{\cdot} \int (Y_{s-} \Delta \mathrm{x}(X_{s-}, \delta) + \mathfrak{I}(X_{s-}, \delta)) \, \nu(d\delta, ds) \end{aligned}$$

où

$$\begin{aligned} \mathrm{x}(x, y) &:= x + \int_0^y f(\mathrm{x}(x, s)) ds, \;\; \Delta \mathrm{x}(x, y) = \mathrm{x}(x, y) - x \\ \mathfrak{I}(x, z) &:= \int_0^z s f(\mathrm{x}(x, s)) ds, \quad \textit{pour } x, y, z \in \mathbb{R}. \end{aligned}$$

Dans le processus X ci-dessus, en plus de sa dynamique endogène, les deux termes $f dY^c$ et $a(\sigma f')$ proviennent de l'impact de la négociation et de la covariation quadratique entre Y et X. La quantité $\Delta \mathrm{x}(x, y)$ représente le saut du prix du sous-jacent dû à l'achat immédiat du nombre y d'actifs. Quant à la richesse V, le terme $\frac{1}{2} a^2 f$ est le résultat net du coût de transaction et de l'appréciation de l'action grâce à l'achat de celle-ci. Un rebalancement immédiat d'un nombre y d'actifs engendre non seulement un saut du type $Y_- \Delta \mathrm{x}$ par la variation du cours des actions détenues avant l'achat, mais aussi un terme $\mathfrak{I}$ dû à l'ajout de y nouvelles actions dans le portefeuille. À partir de ces dynamiques, nous proposons de définir le problème de sur-réplication pour deux types d'options. Selon que l'option est *"covered"* ou pas, les résultats sont différents, les techniques de preuve également.

1.1.3 Un-covered options

1.1.3.1 Définition du problème

Cette section traite du problème de sur-réplication d'options *"un-covered"*. Le *trader* de ces options, partant d'une position nulle sur le sous-jacent, construit son "delta" initial pour commencer sa stratégie. Il délivre un certain montant $g_0(X_T)$ de cash et un nombre $g_1(X_T)$ d'actions à maturité. Notons qu'à la date terminale, cela correspond à un montant de cash égal à $G(X_T)$ où

$$G(x) := \inf\{y \mathrm{x}(x, y) + g_0(\mathrm{x}(x, y)) - \mathfrak{I}(x, y) : \text{ t.q. } y = g_1(\mathrm{x}(x, y))\}, \; x \in \mathbb{R}.$$

Le terme $\mathrm{x}(x, y)$ correspond à la valeur *post-trade* du sous-jacent s'il vaut x avant, et si y est la quantité achetée. $\mathfrak{I}(x, y)$ est l'impact de la négociation sur la valeur du portefeuille, ce qui est le coût de liquidité.

1.1.3.2 Difficulté technique principale

La difficulté principale consiste à établir le principe de la programmation dynamique, qui forme une relation entre la valeur du portefeuille optimal et le prix de sur-réplication à travers les processus de contrôle, notamment X, Y, voir dans [48], [50] ou [41]. Cependant, la formulation usuelle ne peut servir à montrer l'équation de valorisation dans notre modèle d'impact. La raison étant que dans un modèle où X, Y, V ont tous un terme linéaire en

la variable de contrôle b, l'optimization s'effectue sur une droite, un problème singulier! Cette difficulté est surmontée par un changement de variable, ce qui permet d'énoncer le principe de la programmation dynamique en fonction du processus de valeur liquidative, i.e. le prix obtenu après que l'opérateur ait liquidé ses positions immédiatement:

$$\hat{X}^{t,z,\gamma} := \mathrm{x}(X^{t,x,\gamma}, -Y^{t,z,\gamma})$$

où $z = (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$ avec le prix initial du sous-jacent $X_{0-} = x$, la position initiale y et la richesse initiale du portefeuille v. Nous définissons le prix correspondant de sur-réplication - $w(t, \hat{X}_t)$, qui désigne le minimum de la richesse initiale afin de couvrir la créance négociable à la maturité, quand on part d'une position nulle sur le sous-jacent.

Ces changements de variable nous évitent la singularité dans le problème d'optimization, ils permettent aussi d'en déduire l'équation de valorisation d'une façon formelle. Cependant, les sauts introduits par les processus $\hat{X}$ et w exigent un peu plus de délicatesse quant à l'énonce du principe de la programmation dynamique. En fait, dans la démonstration de la proprété de solution de viscosité, nous avons besoins que la variable de contrôle (a, b, ν) soit bornée. Il faut par conséquent construire l'ensemble de variables de contrôle d'une manière prudente telle que cet ensemble soit fermé et permette d'utiliser un argument de sélection mesurable.

Soit Γ l'ensemble des variables de contrôle $\Gamma := \cup_{k\geq 1}\Gamma_k$ où

$$\Gamma_k := \{(a, b, \nu) \in \mathcal{A}_k \times \mathcal{U}_k : |Y| \leq k\}.$$

Une stratégie $\gamma \in \Gamma$ est appelée *admissible* si elle permet aux processus $(X_T^{t,z,\gamma}, Y_T^{t,z,\gamma}, V_T^{t,z,\gamma})$ de sur-répliquer (g_0, g_1). Soit w_k le minimum de la richesse initiale qui assure l'existence des stratégies admissibles dans l'ensemble Γ_k. Remarquons que la valeur à T de la fonction $w := \inf_{k\geq 1} w_k$ est égale au prix de sur-réplication G.

Notre principe de la programmation dynamique se décrit de la manière suivante:

Proposition 1.1.2. *Fixons $(t, x, v) \in [0, T] \times \mathbb{R} \times \mathbb{R}$.*

1. *Si $v > w(t, x)$, alors il existe $\gamma \in \Gamma$ et $y \in \mathbb{R}$ tel que*

$$V_\theta^{t,z,\gamma} \quad \geq \quad w(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma}),$$

   pour tout temps d'arrêt $\theta \geq t$, où $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

2. *Fixons $k \geq 1 Si\ v < w_{2k+2}(t, x)$, alors il n'existe pas $\gamma_k \in \Gamma_k$, $y \in [-k, k]$ et un temps d'arrêt $\theta \geq t$ tel que*

$$V_\theta^{t,z,\gamma} > w_k(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$$

   avec $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

Une fois le principe de la programmation dynamique établi, il reste à suivre les étapes classiques de l'approche de cible stochastique afin de déduire l'équation de valorisation.

1.1.3.3 Contribution

Nous avons démontré que l'équation de valorisation est *quasi-linéaire*.

Soient $\hat{\mu}$ et $\hat{\sigma}$ les processus du dérive et de la volatilité décalés, définis par

$$\hat{\mu}(x', y') := \frac{1}{2}[\partial^2_{xx}\mathrm{x}\sigma^2](\mathrm{x}(x', y'), -y') \text{ and } \hat{\sigma}(x', y') := (\sigma\partial_x\mathrm{x})(\mathrm{x}(x', y'), -y').$$

Définissons également les opérateur qui agissent sur des fonctions régulières

$$\begin{aligned}\hat{F}\varphi &:= -\partial_t\varphi - \hat{\mu}\partial_x[\varphi + \mathfrak{I}] - \frac{1}{2}\hat{\sigma}^2\partial^2_{xx}[\varphi + \mathfrak{I}] \\ \hat{y}[\varphi](t,x) &:= \mathrm{x}^{-1}(x, x + f(x)\partial_x\varphi(t,x)) \ .\end{aligned}$$

Alors le prix du dérivé est une solution de viscosité de

$$F\varphi\mathbf{1}_{[0,T[} + (\varphi - G)\mathbf{1}_{\{T\}} = 0 \ \text{ on } [0,T]\times\mathbb{R} \ ,$$

où $F\varphi(t,x) := \hat{F}\varphi(t,x,\hat{y}[\varphi](t,x))$. Nous avons ensuite établi le principe de comparaison, d'où l'unicité de la solution. Sous des conditions appropriées, l'équation admet une solution dans $C^{1,2}([0,T)\times\mathbb{R})\cap C^0([0,T]\times\mathbb{R})$. Dans ce cas-là, la stratégie de couverture est donnée par

$$Y = \mathrm{x}^{-1}(\hat{X}, \hat{X} + (f\partial_x\varphi)(\cdot,\hat{X})) - Y_{T-}\mathbb{1}_{\{T\}}$$

Cette stratégie, malgré son apparence compliquée, est actuellement une sorte de *delta-réplication* décalée. Dans l'exemple de modèle Bachelier avec impact constant égale à λ,

$$Y = \partial_x\varphi(\cdot, X - \lambda Y),$$

qui est le delta prise en $X - \lambda Y$ (voir Section 3.2.4).

1.1.4 Covered options

1.1.4.1 Définition du problème

Une situation similaire à ce qui a été traité dans la section précédente est la valorisation des options europénnes *"covered"*. Le terme *"covered"* implique que l'acheteur de l'option délivre à la création du portefeuille de couverture le delta initial dont le *trader* a besoin, et accepte un mélange d'actions et de liquide à la maturité (choisi par le trader). L'indifférence de l'acheteur entre les actions et le liquide au début et à la fin de la réplication dynamique élimine le coût engendré par la couverture initiale et finale. Le problème de sur-réplication des options "covered" est formellement défini comme

$$\mathrm{v}(t,x) := \inf\{v = c + yx \ : \ (c,y)\in\mathbb{R}^2 \text{ s.t. } \mathcal{G}(t,x,v,y)\neq\emptyset\},$$

où $\mathcal{G}(t,x,v,y)$ est l'ensemble des variables de contrôle (a,b) tel que $\phi := (y,a,b)$ satisfait

$$V_T^{t,x,v,\phi} \geq g(X_T^{t,x,\phi}).$$

La définition ci-dessus ressemble significativement au problème précédent, à la différence que le terme $\mathfrak{I}$ n'apparait plus: il ne servait qu'à gérer le saut de la richesse lié au premier achat du sous-jacent. Nous avons vu dans la section précédente que les seuls moments où la stratégie de couverture saute sont respectivement le début et la fin de la réplication. La possibilité d'acquérir ou de délivrer les actions aux clients en ces deux moments élimine $\mathfrak{I}$ dans notre problématique.

1.1.4.2 Difficulté technique principale

Malgré sa ressemblance à la première situation, nous ne pouvons toutefois démontrer le principe de la programmation dynamique dans les deux directions en raison de l'intéraction forte entre la stratégie de réplication et le processus du prix de sous-jacent. La première partie du principe de la programmation dynamique peut encore être formulée sous une forme faible comme dans [19, Section 5]. Nous établissons ensuite la propriété de sur-solution en suivant essentiellement la même méthode, voir Section 4.2.2. Le problème

est que, à une date intermédiaire θ, la quantité Y_θ de titres détenue n'a aucune raison de correspondre avec celle $\hat{Y}_\theta$ nécessaire pour pouvoir continuer la stratégie de couverture jusqu'en T, même si la richesse est elle suffisante. Dans [51], les auteurs ont réussi à contourner cette difficulté par un argument très habile de concaténation continue. Cette démarche toutefois ne fonctionne plus dans notre modèle du fait de l'impact. En suivant leur approche, on modifierait la trajectoire de prix et donc la cible $\hat{Y}$. Pour cette raison là, nous utilisons la technique de régularisation dévelopée dans [13] afin de démontrer la propriété de sous-solution, ce qui évite la nécessité d'utiliser le principe de la programmation dynamique. L'idée de cette technique est de construire tout d'abord une suite de sur-solutions régulières qui, par un argument de vérification, fournit une borne supérieure sur le prix de sur-réplication. Puisque cette suite de sur-solutions converge vers la solution de l'équation de valorisation souhaitée, un principe de comparaison implique que leur limite n'est rien d'autre que le prix de sur-réplication (car nous savons déjà que celui-ci est sur-solution).

La deuxième différence du cas de *"un-covered"* options est l'introduction de la contrainte de *gamma* dans ces options *"covered"*. Un calcul formel (voir [37] et [1]) montre que l'équation de valorisation est non-linéaire du type

$$-\partial_t \varphi - \frac{1}{2} \frac{\sigma^2}{1 - f \partial^2_{xx} \varphi} \partial^2_{xx} \varphi = 0.$$

Pour que cette équation ne soit pas mal-posée à cause du terme $1 - f\partial^2_{xx}\varphi$ dans le dénominateur, nous devons imposer que $1 - f\partial^2_{xx}\varphi$ ne change pas de signe, d'où la contrainte de gamma

$$\bar{\gamma} \geq \partial^2_{xx} \varphi$$

dans laquelle $\bar{\gamma} : \mathbb{R} \to \mathbb{R}$ est une application bornée par

$$\iota \leq \bar{\gamma} \leq 1/f - \iota$$

pour $\iota > 0$. Cette contrainte nous permet d'assurer que le dénominateur reste positif

$$1 - f\partial^2_{xx}\varphi > 0.$$

Une interprétation intuitive est détailée dans Section 4.1.2. Remarquons que notre contrainte est légèrement différente de celle traitée dans les travaux de [19], [48] et [50], où leur gamma est borné par des constantes.

1.1.4.3 Contribution

Nous avons démontré que l'équation de valorisation est non-linéaire du type *parabolique*.

Définissons l'opérateur de variation qui comprend la contrainte gamma:

$$F[\varphi] := \min \left\{ -\partial_t \varphi - \frac{1}{2} \frac{\sigma^2}{1 - f \partial^2_{xx} \varphi} \partial^2_{xx} \varphi \ , \ \bar{\gamma} - \partial^2_{xx} \varphi \right\} \tag{1.1.1}$$

Le prix du dérivé est une solution de viscosité de $F[\varphi] = 0$ sur $[0, T) \times \mathbb{R}$. Quant à la condition terminale, la contrainte dans ce domaine $[0, T) \times \mathbb{R}$ se propage tout au long de l'axe temporel jusqu'à la maturité T, ce qui modifie le pay-off g en la fonction $\hat{g}$, définie par:

$$\hat{g} := (g - \bar{\Gamma})^{\text{conc}} + \bar{\Gamma},$$

où $\bar{\Gamma} \in \mathcal{C}^2$ avec $\partial^2_{xx} \bar{\Gamma} = \bar{\gamma}$. L'exposant "conc" désigne l'enveloppe concave de la fonction.

Nous obtenons au finale la description de la valorisation du prix de dérivé sur le domaine entier $[0,T]\times\mathbb{R}$:

$$F[\varphi]\mathbf{1}_{[0,T)}+(\varphi-\hat{g})\mathbf{1}_{\{T\}}=0 \quad \text{on } [0,T]\times\mathbb{R}.$$

Nous avons également obtenu la stratégie de couverture optimale à ϵ près en même temps. En plus de proposer un schéma numérique pour cette équation, nous illustrons la façon dont la fonction d'impact f a changé le prix du dérivé en appliquant notre résultat au modèle Bachelier. À la fin de ce chapitre, nous avons montré qu'ajouter de la résilience ne change pas l'équation de valorisation obtenue.

1.2 BDE Branching

1.2.1 Motivation et problème lié aux polynômes

Le deuxième objectif de cette thèse est d'étudier les schémas numériques pour les équations différentielles stochastiques rétrogrades (EDSR). El Karoui, Peng et Quenez [24] soulignent que l'EDSR est étroitement liée au problème de la valorisation et de la réplication. De nombreuses approches numériques ont été suggérées au sujet de la résolution de ce système d'équations. Les travaux de [15], [53] et [2] par exemple, partis de la technique de discrétisation en temps, reposent sur l'estimation de l'espérance conditionelle, une méthode rétrograde.

Un autre type d'algorithm numérique, proposé et étudié dans [27], [29] et [28], est basé sur un procédé purement "forward", qui permet déviter l'estimation de l'espérance conditionelle. Le facteur clé de cet algorithme "forward" est la représentation probabiliste en termes de "processus de branchement", une généralisation de la représentation de Feynman-Kac. En fait, étant donnée une EDSR avec un générateur polynomial

$$\begin{aligned} X_\cdot &= X_0+\int_0^\cdot \mu(X_s)ds+\int_0^\cdot \sigma(X_s)dW_s \\ Y_\cdot &= g(X_T)+\int_\cdot^T \sum_{k\geq 0} p_k(Y_t)^k dt-\int_\cdot^T Z_t dW_t \end{aligned}$$

où $(p_k)_{k\geq 0}$ dont des poids et W un mouvement brownien, la solution $Y_\cdot$ coincide avec celle de l'équation différentielle partielle de Kolmogorov-Petrovskii-Piskunov (KPP)

$$\begin{aligned} \partial_t u(t,x)+\mathcal{L}u(t,x)+\sum_{k\geq 0} p_k u^k(t,x)=0 \\ u(T,x)=g(x) \end{aligned}$$

où $\mathcal{L}$ est l'opérateur Dynkin associé au processus X. Chercher un algorithm numérique pour $Y_\cdot$ revient à résoudre l'équation de KPP numériquement. Cette dernière pourrait être exprimée de façon "forward" et probabiliste par le processus de branchement, à l'instar de [27], [29] et [28]. Soient $\mathcal{K}_t^n$ l'ensemble des particules de la génération n vivantes à l'instant t, $\bar{\mathcal{K}}_t^n$ les particules de la génération n nées avant t. Ces particules sont indexées par k et notées X^k. Par ailleurs, définissons

$$\mathcal{K}_t:=\cup_{n\geq 1}\mathcal{K}_t^n,\quad \bar{\mathcal{K}}_t:=\cup_{n\geq 1}\bar{\mathcal{K}}_t^n,$$

qui désignent respectivement les particules vivantes à et nées avant t. Soient ρ la densité de la loi de survie de ces particules, $\bar{F}$ est 1 moins la fonction de répartition. Alors la solution $u(.,.)$ est égale à l'espérance de produits

$$u(t,x) = \mathbb{E}\Big(\prod_{k\in\mathcal{K}_T}\frac{g(X_T^k)}{\overline{F}(\delta_k)}\Big)\Big(\prod_{k\in\bar{\mathcal{K}}_T\setminus\mathcal{K}_T}\frac{1}{\rho(\delta_k)}\Big).$$

où δ_k est la longévité de la particule k. Alors que les travaux cités ci-dessus se concentrent sur les générateurs de forme polynômiale, nous nous intéressons à la généralisation de cette démarche au générateur arbitraire Lipschistzien.

1.2.2 Algorithme

À partir des travaux présentés dans la section précédente, nous aurions souhaité en premier lieu approximer un générateur arbitraire Lipschitzien par une suite de polynômes. Cette approximation est toutefois infaisable en général. La raison est que la solution d'une EDSR avec un générateur polynomial, de degré supérieur ou égal à deux, explosera typiquement en temps fini. Ce fait se complexifie à mesure que le degré du polynôme est plus grand.

Nous proposons ainsi une approximation par une suite de polynômes locaux. Soit $\tilde{f}_{\ell_\circ}(x,y)$ défini comme

$$\tilde{f}_{\ell_\circ} : (x,y) \in \mathbb{R}^d \times \mathbb{R} \mapsto \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y) \ .$$

où les fonctions $(a_{j,l}, \varphi_j)_{\ell \le \ell_\circ, j \le j_\circ}$ sont continues et bornées. Pour une partition de $\mathbb{R}$ donné $(A_i)_i$, les fonctions $(\varphi_j)_{j \le j_\circ}$ peuvent être interprétées comme des noyaux permettant de passer d'un ensemble à l'autre de façon Lipschitzienne. Nous établissons en premier lieu la convergence de la suite de solutions correspondantes à ces générateurs approximés $\tilde{f}$ vers la solution de l'EDSR originaire.

L'argument y apparaît aux deux endroits différents dans la définition du polynôme local $\tilde{f}_{\ell_\circ}$ ci-dessus. Le premier endroit est la partie de monôme, alors que le deuxième est la partie de localization. Afin d'appliquer les résultats en polynôme global, nous allons découpler ces deux rôles de l'argument y. Définissons maintenant $f_{\ell_\circ}(x,y,y')$ par

$$f_{\ell_\circ} : (x,y,y') \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \mapsto \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y') \ .$$

Ce générateur est un polynôme global en y, avec $f_{\ell_\circ}(x,y,y) = \tilde{f}_{\ell_\circ}(x,y)$. Nous utilisons ensuite l'itération de Picard afin d'obtenir la solution d'une telle EDSR. La solution de la m-ème itération est donnée par

$$Y^m_\cdot = g(W_T) + \int_\cdot^T f_{\ell_\circ}(X, Y^m_t, Y^{m-1}_t) dt - \int_\cdot^T Z_t dW_t \ .$$

Pour présenter Y^m de manière probabiliste comme dans la section précédente, nous allons tout d'abord diviser l'intervalle de temps $[0,T]$ en n sous-intervalles délimités par $(t^n_i)_{i \le n}$. La longueur de ces sous-intervalles est choisie de manière à ce que la solution Y^m reste bornée par une constante donnée. Ensuite un argument de récurrence nous permet d'étendre cette réprésentation probabiliste sur l'intervalle entier $[0,T]$.

Contrairement aux schémas classiques de l'itérations de Picard en EDSR, avec par exemple [6], l'estimation précise de l'ensemble des trajectoires de la solution Y à chaque itération n'est plus nécessaire. En effet, si le polynôme local est réparti par une partition $(A_i)_i$ de $\mathbb{R}$, alors il suffit de savoir dans quel ensemble A_i la solution reste aux moments de branchement.

1.2.3 Contribution

Nous définissons récursivement

$$\begin{aligned} v^0 &:= y \quad , \\ v^m(t,x) &:= \mathbb{E}\left[V^m_{t,x}\right], \quad m \ge 1, \end{aligned}$$

où $y : [0,T] \times \mathbb{R}$ est une fonction déterministe bornée, et V^m est définie en fonction de v^{m-1}, v^m par

$$\begin{aligned} V^m_{t,x} &:= \Big(\prod_{k \in \mathcal{K}_{t^n_{i+1}-t}} G^m_{t,x}(k) \Big) \Big(\prod_{k \in \bar{\mathcal{K}}_{t^n_{i+1}-t} \setminus \mathcal{K}_{t^n_{i+1}-t}} A^m_{t,x}(k) \Big), \\ G^m_{t,x}(k) &:= \frac{v^m\big(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}-t}\big)}{\overline{F}(t^n_{i+1} - t - T_{k-})}, \\ A^m_{t,x}(k) &:= \frac{\sum_{j=1}^{j_\circ} a_{j,\xi_k}(X^{x,(k)}_{T_k}) \varphi_j(v^{m-1}(t+T_k, X^{x,(k)}_{T_k}))}{p_{\xi_k}\, \rho(\delta_k)}, \quad \forall (t,x) \in [t^n_i, t^n_{i+1}) \times \mathbf{X}. \end{aligned}$$

dans lesquelles la particule indexée par k est née en T_{k-} et meurt en T_k. Nous avons démontré que la suite v^m converge vers la solution de l'EDSR du générateur f_{ℓ_0}. Nous appliquons à la fin cette méthode à un exemple concret afin d'illustrer la performance de notre algorithme.

1.3 Publications

Tous les travaux dans cette thèse ont été publié ou font l'objet d'une révision dans des journaux académiques.

Bruno Bouchard, Grégoire Loeper, and Yiyi Zou. Almost-sure hedging with permanent price impact. *Finance and Stochastics*, 20(3):741–771, 2016

B Bouchard, G Loeper, and Y Zou. Hedging of covered options with linear market impact and gamma constraint. *arXiv preprint arXiv:1512.07087*, 2015. (en révision mineur pour SIAM SICON)

Bruno Bouchard, Xiaolu Tan, and Yiyi Zou. Numerical approximation of bsdes using local polynomial drivers and branching processes. *arXiv preprint arXiv:1612.06790*, 2016. (soumis)

Chapter 2

Introduction

Contents

2.1 Impact model **11**
2.1.1 General motivation 11
2.1.2 Dynamics description 12
2.1.3 Un-covered options 13
2.1.4 Covered options 15
2.2 BDE Branching **17**
2.2.1 Motivation and Problems related to polynomials 17
2.2.2 Algorithm 17
2.2.3 Contribution 18
2.3 Publications **19**

2.1 Impact model

2.1.1 General motivation

The classical theory of derivatives valuation is based on three principle hypothesis: complete market, absence of transaction cost and perfect liquidity. Based on these hypothesis, the dynamic hedging provides an intuitive way of pricing the derivatives in an arbitrage-free market. Empirically however, these hypothesis are rarely verified, in particular for large transactions on illiquid assets.

Numerous work has been dedicated to the relaxation of these hypothesis. Some considers the liquidity cost without taking into consideration the permanent trading impact on the underlying, see for example [17]. The underlying asset price, a function of the volume of the instantaneous transaction, does not depend on the past trading activity. As the transaction could be executed at the marginal price, the introduction of liquidity cost does not affect the replication price. When a supplementary liquidity constraint is imposed, for example the gamma of the strategy being bounded, the pricing equation presents a supplementary quadratique term, see [18]. Some others are focusing on the analytic form of the price dynamics by establishing the demand-offer equilibrium, without taking the liquidity cost into account, for example [25], [45] et [46]. The underlying price is therefore modified by the trading activity of different players in the market. This impact generates a non-linearity in the pricing equation of the derivatives, which is resulted from a modified volatility term.

More recently, [1] and [37] have conducted a novel approach in which the dynamics of the underlying is composed of a classical diffusion process and linear impact term. This

model integrates at the same time the liquidity cost and the trading impact on the underlying price in a natural manner. When a regular solution of the non-linear equation is available, they have shown that a perfect replication could be then deduced. However, the deduction of such pricing equation as well as of the replication strategy rests on a verification argument. This argument not only relies on some strong regularity conditions, but also could not lead to the equality between the super-replication price and the replication price. Our problematics then becomes, without presuming any regularity condition on the derivatives price function, how to establish the pricing equation by following the classical stochastic target approach. As we will see, this necessitates some important and distinct modifications on the classical stochastic target approach.

2.1.2 Dynamics description

We model the permanent trading impact on the underlying price by the impact function f: the price variation of the underlying due to trading $\delta \in \mathbb{R}$ number of risky asset is $\delta f(x)$, where x is the underlying price before the transaction. The cost of buying such number of asset is

$$\delta x + \frac{1}{2}\delta^2 f(x) = \delta \int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota,$$

where the quantity

$$\int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota$$

is the average cost to purchase one unit of underlying asset.

The dynamics of the underlying between two consecutive transactions is described by a diffusion

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t.$$

We call a trading strategy the process defined as

$$Y = Y_{0-} + \int_0^\cdot b_s ds + \int_0^\cdot a_s dW_s + \int_0^\cdot \int \delta\nu(d\delta, ds),$$

where a et b are predictable processes with values in $\mathbb{R}$ satisfying

$$(a, b) \in \mathcal{A} := \cup_k \mathcal{A}_k,$$
$$\mathcal{A}_k := \{(a, b) \text{ predictable with values in } \mathbb{R} : |(a, b)| \leq k \; dt \times d\mathbb{P} - \text{a.e.}\},$$

and $\nu \in \mathcal{U} := \cup_k \mathcal{U}_k$, is a control variable in the space of positive measures with values in $\mathbb{N}$. Here, $\mathcal{U}_k$ is the set of random measures with values in $\{0, \cdots, k\}$ with support $[-k, k] \times [0, T]$.

The process Y should be regarded as trading signal - the objective of traders is to hold Y units of risky asset in their portfolio. The component ν models the jumps of the number of risky asset at the time of building an initial delta when needed, or at the time of liquidating the portfolio at the maturity (we will see that the strategy does not jump between these two dates).

We deduce the dynamics of the underlying X and of the total wealth V (the sum of the cash and the stock value) via the auto-financing argument. We start by considering passing an order of the size $Y_{t_{i+1}} - Y_{t_i}$ at the discrete time grid t_i, then we obtain the continuous-time dynamics by tending the balancing frequency to infinity, i.e. $t_{i+1} - t_i \to 0$. From this point of view, our model is the result of a formulation from simple strategies.

Proposition 2.1.1. *Let $(a,b,\nu) \in \mathcal{A}\times\mathcal{U}$, then the dynamics of X and of V corresponding to continuous trading strategies are given by*

$$\begin{aligned} X &= X_{0-} + \int_0^{\cdot} \sigma(X_s)dW_s + \int_0^{\cdot} f(X_s)dY_s^c + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s))ds \\ &\quad + \int_0^{\cdot}\int \Delta\mathrm{x}(X_{s-},\delta)\nu(d\delta,ds) \\ V &= V_{0-} + \int_0^{\cdot} Y_s dX_s^c + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X_s)ds \\ &\quad + \int_0^{\cdot}\int (Y_{s-}\Delta\mathrm{x}(X_{s-},\delta) + \mathfrak{I}(X_{s-},\delta))\,\nu(d\delta,ds) \end{aligned}$$

where

$$\begin{aligned} \mathrm{x}(x,y) &:= x + \int_0^y f(\mathrm{x}(x,s))ds,\ \ \Delta\mathrm{x}(x,y) = \mathrm{x}(x,y) - x \\ \mathfrak{I}(x,z) &:= \int_0^z sf(\mathrm{x}(x,s))ds, \quad \textit{pour } x,y,z \in \mathbb{R}. \end{aligned}$$

In the above process of X, in addition to its endogenous dynamics, the two terms fdY^c and $a(\sigma f')$ come from the trading impact and the quadratic covariance between Y et X. The quantity $\Delta\mathrm{x}(x,y)$ represents the jump of the underlying price due to the immediate purchase of y numbers of asset. As for the wealth V, the term $\frac{1}{2}a^2 f$ is the net result of the transaction cost and the underlying appreciation due to the purchase. An immediate trade of y numbers of asset not only generates a jump of the type $Y_-\Delta\mathrm{x}$ by the value variation of the stocks held already before the purchase, but also a term $\mathfrak{I}$ due to the addition of y new stocks in the portfolio. From these dynamics, we propose to define the super-replication problem for two different types of options. According to the fact that the option is "covered" or not, the results are distinguishably different, as well as the technique employed.

2.1.3 Un-covered options

2.1.3.1 Probleme Definition

This section treats the super-replication problem of *"un-covered"* options. The trader of these options, starting from a naught position on the underlying, builds his initial "delta" to launch his strategy. He then delivers certain amount of cash $g_0(X_T)$ and $g_1(X_T)$ numbers of stocks at the maturity. This payoff at the maturity actually corresponds to a total cash amount equal to $G(X_T)$ where

$$G(x) := \inf\{y\mathrm{x}(x,y) + g_0(\mathrm{x}(x,y)) - \mathfrak{I}(x,y) :\ \text{t.q. } y = g_1(\mathrm{x}(x,y))\},\ x \in \mathbb{R}.$$

The term $\mathrm{x}(x,y)$ corresponds to the *post-trade* value of the underlying if its price was x before, and if y is the purchased quantity. $\mathfrak{I}(x,y)$ is the trading impact on the portfolio value, which is nothing but the liquidity cost.

2.1.3.2 Principle technique difficulty

The principle difficulty resides in establishing the Dynamic Programming Principle, which relates the value of the optimal portfolio to the super-replication price, through the control processes, in particular X, Y, see in [48], [50] or [41]. However, the usual formulation could no longer be applied to prove the pricing equation in our model with impact. The reason being that in a model where X, Y, V all have a linear term on the control variable b,

the optimization would be performed on a one-dimension line, a singular optimization problem! This difficulty is overcome by a change of variable, which allows us to announce the Dynamic Programming Principle on function of the process of liquidation value, i.e. the underlying price after the trader has liquidated all his positions immediately:

$$\hat{X}^{t,z,\gamma} := \mathrm{x}(X^{t,x,\gamma}, -Y^{t,z,\gamma})$$

where $z = (\mathrm{x}(x,y), y, v+\mathfrak{I}(x,y))$ with the initial price of the underlying $X_{0-} = x$, the initial position y and the initial portfolio wealth v. We define the corresponding super-replication price- $w(t, \hat{X}_t)$, which designates the minimum of the initial wealth to cover the pay-off at the maturity, when starting from a position naught on the underlying.

This change of variable helps us avoid the singularity in the optimization problem. It also allows us to deduce the pricing equation in a formal way. Meanwhile, the jumps introduced by the processes $\hat{X}$ and w require some more delicacy when announcing the Dynamic Programming Principle. In fact, in the proof of the viscosity solution property, we need that the control variables (a, b, ν) be bounded. One thus has to construct the set of control variables prudently such that this set be closed and thus allow using the argument of measurable selection.

Let Γ be the set of control variables $\Gamma := \cup_{k\geq 1}\Gamma_k$ where

$$\Gamma_k := \{(a,b,\nu) \in \mathcal{A}_k \times \mathcal{U}_k : |Y| \leq k\}.$$

A strategy $\gamma \in \Gamma$ is called *admissible* if it allows the processes $(X_T^{t,z,\gamma}, Y_T^{t,z,\gamma}, V_T^{t,z,\gamma})$ to super-replicate (g_0, g_1). Let w_k be the minimum of the initial wealth that assures the existence of admissible strategies in the set Γ_k. Notice that the $T-$value of the function $w := \inf_{k\geq 1} w_k$ is equal to the super-replication price G.

Our Dynamic Programming Principle is expressed in the way below:

Proposition 2.1.2. *Fix $(t,x,v) \in [0,T] \times \mathbb{R} \times \mathbb{R}$.*

1. *If $v > w(t,x)$, then there exists $\gamma \in \Gamma$ and $y \in \mathbb{R}$ such that*

$$V_\theta^{t,z,\gamma} \quad \geq \quad w(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma}),$$

   for any stopping time $\theta \geq t$, where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

2. *Fix $k \geq 1 Si$ $v < w_{2k+2}(t,x)$, then there does not exist $\gamma_k \in \Gamma_k$, $y \in [-k,k]$ and a stopping time $\theta \geq t$ such that*

$$V_\theta^{t,z,\gamma} > w_k(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$$

   with $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

Once the Dynamic Programming Principle established, it remains only following the classical steps of stochastic target to deduce the pricing equation.

2.1.3.3 Contribution

We have shown the valuation equation is *quasi-linear*.

Let $\hat{\mu}$ et $\hat{\sigma}$ be the processes of the *shifted* drift and volatility, defined as

$$\hat{\mu}(x',y') := \frac{1}{2}[\partial^2_{xx}\mathrm{x}\sigma^2](\mathrm{x}(x',y'), -y') \text{ and } \hat{\sigma}(x',y') := (\sigma\partial_x\mathrm{x})(\mathrm{x}(x',y'), -y').$$

Define the following operator acting on the regular functions

$$\begin{aligned} \hat{F}\varphi &:= -\partial_t\varphi - \hat{\mu}\partial_x[\varphi + \mathfrak{I}] - \frac{1}{2}\hat{\sigma}^2\partial^2_{xx}[\varphi + \mathfrak{I}] \\ \hat{y}[\varphi](t,x) &:= \mathrm{x}^{-1}(x, x + f(x)\partial_x\varphi(t,x)) \ . \end{aligned}$$

Then the price of the derivatives is a viscosity solution of

$$F\varphi \mathbf{1}_{[0,T[} + (\varphi - G)\mathbf{1}_{\{T\}} = 0 \ \text{ on } [0,T] \times \mathbb{R}\,,$$

where $F\varphi(t,x) := \hat{F}\varphi(t,x,\hat{y}[\varphi](t,x))$.

We have then established the comparison principle, based on which is proved the uniqueness of the solution. Under appropriate conditions, the equation admits a solution in $C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$. In this case, the replication strategy is given by

$$Y = \mathrm{x}^{-1}(\hat{X}, \hat{X} + (f\partial_x\varphi)(\cdot, \hat{X})) - Y_{T-}\mathbb{1}_{\{T\}}$$

This strategy, despite its complicated appearance, is actually a kind of *shifted* delta-replication. In the example of Bachelier model with constant trading impact equal to λ,

$$Y = \partial_x\varphi(\cdot, X - \lambda Y),$$

which is the delta taken at $X - \lambda Y$ (see Section 3.2.4).

2.1.4 Covered options

2.1.4.1 Problem definition

A similar situation to what is dealt in the precedent section is the valuation of *"covered"* European options. The term *"covered"* implies that the option buyer delivers the initial delta at the creation of the hedging portfolio needed by the trader, and accepts a mixture of stocks and cash (determined by the trader) at the maturity. The buyer's indifference between the stocks and the cash at the beginning and at the end of the dynamic hedging eliminates the cost generated by the initial and final trading. Thus the problem of super-replication of the "covered" options is formally defined as

$$\mathrm{v}(t,x) := \inf\{v = c + yx \ : \ (c,y) \in \mathbb{R}^2 \text{ s.t. } \mathcal{G}(t,x,v,y) \neq \emptyset\},$$

where $\mathcal{G}(t,x,v,y)$ is the set of control variables (a,b) such that $\phi := (y,a,b)$ satisfies

$$V_T^{t,x,v,\phi} \geq g(X_T^{t,x,\phi}).$$

The above definition resembles significantly to the precedent problem, at the difference that the term $\mathfrak{I}$ does no longer appear: it serves only to manage the jump of the wealth related to the first purchase of the underlying. We have seen in the previous section that the only moments where the hedging strategy jumps are respectively the beginning and the end of the replication. The possibility of acquiring or of delivering the stocks to the clients at these two moments eliminates $\mathfrak{I}$ in our problematics.

2.1.4.2 Principle technique difficulty

Despite its resemblance to the "un-covered" situation, we are no longer able to prove the Dynamic Programming Principle in the two directions, as a result of the strong interaction between the trading strategy and the process of the underlying. The first part of the Principle could still be formulated in the weak form as in [19, Section 5]. We then prove the super-solution property by following essentially the same method, see Section 4.2.2. The problematics for the other direction is that, at any intermediate moment θ, the quantity Y_θ of held asset has no reason to correspond to the $\hat{Y}_\theta$, the one necessary to continue the hedging strategy until time T, even the wealth itself remains sufficient. In [51], the authors succeeded in circumventing this difficulty by an astute argument of continuous concatenation. This method however, no longer works in our model because of

trading impact. By following their approach, one has to modify the path of the underlying price and thus the target Y. For this reason, we are going to use the technique of regularization developed in [13] in order to prove the sub-solution property. The idea of this technique is to construct firstly a series of regular super-solutions which, by verification argument, provides an upper bound on the super-replication price. Now that this series of super-solutions converges towards the solution of the objective valuation equation, the comparison principle implies that their limit is nothing but the super-replication price (because we already know that the price is a super-solution).

The second difference from the case of "un-covered" options is the introduction of the gamma constraint in these "covered" options. A formal calculation shows (see [37] and [1]) that the valuation equation is non-linear, of the type

$$-\partial_t\varphi - \frac{1}{2}\frac{\sigma^2}{1 - f\partial^2_{xx}\varphi}\partial^2_{xx}\varphi = 0.$$

In order that this equation is well-defined due to the term $1 - f\partial^2_{xx}\varphi$ in the denominator, we have to impose that $1 - f\partial^2_{xx}\varphi$ does not change its sign, where comes from the gamma constraint:

$$\bar{\gamma} \geq \partial^2_{xx}\varphi$$

in which $\bar{\gamma} : \mathbb{R} \to \mathbb{R}$ is a function bounded by

$$\iota \leq \bar{\gamma} \leq 1/f - \iota$$

where $\iota > 0$. This constraint assures that the denominator remains positive

$$1 - f\partial^2_{xx}\varphi > 0.$$

An intuitive interpretation is detailed in Section 4.1.2. Notice that our constraint is slightly different from that is dealt in the works of [19], [48] and [50], where their gamma is bounded by some constants.

2.1.4.3 Contribution

We have shown the valuation equation is parabolic non-linear equation.

Define the following variation operator incorporating the gamma constraint:

$$F[\varphi] := \min\left\{-\partial_t\varphi - \frac{1}{2}\frac{\sigma^2}{1 - f\partial^2_{xx}\varphi}\partial^2_{xx}\varphi \ , \ \bar{\gamma} - \partial^2_{xx}\varphi\right\} \tag{2.1.1}$$

The price of the derivatives is a viscosity solution of $F[\varphi] = 0$ on $[0, T) \times \mathbb{R}$. As for the terminal condition, the constraint in the domain of $[0, T) \times \mathbb{R}$ propagates along the temporal axe until the maturity T, which modifies the pay-off g to the the so-call *face-lift* function $\hat{g}$, defined by:

$$\hat{g} := (g - \bar{\Gamma})^{\text{conc}} + \bar{\Gamma},$$

where $\bar{\Gamma} \in \mathcal{C}^2$ with $\partial^2_{xx}\bar{\Gamma} = \bar{\gamma}$. The super-script "conc" designates the concave envelope of the function.

We have proved price function of the derivatives on the entire domain $[0, T] \times \mathbb{R}$ satisfies

$$F[\varphi]\mathbf{1}_{[0,T)} + (\varphi - \hat{g})\mathbf{1}_{\{T\}} = 0 \quad \text{on } [0, T] \times \mathbb{R}.$$

An $\epsilon-$ optimal hedging strategy is obtained at the same time. Not only have we proposed a numerical scheme for this equation, we have also illustrated how the impact function f has changed the price of the derivatives by applying our results to Bachelier model. At the end of this chapter, we have shown that adding resilience to our model does not change the obtained valuation equation.

2.2 BDE Branching

2.2.1 Motivation and Problems related to polynomials

The second objective of this thesis is to study the numerical scheme for the BSDEs. El Karoui, Peng et Quenez [24] has in their work emphasized that the BSDE is closed related to the derivatives valuation and the hedging problems. Numerous numerical methods have been proposed to solve this equation system. The work of [15], [53] and [2] for example, based on the discretization technique, relies on the estimation of the conditional expectation, thus a backward approach.

Another type of numerical algorithm, proposed and studied in [27], [29] and [28], is based on a purely "forward" approach, which avoids the estimation on the conditional expectation. The key factor in this "forward" algorithm is the probabilist representation of the solution in terms of "branching process", a generalization of the Feynman-Kac representation. In fact, given a BSDE with a polynomial driver

$$\begin{aligned} X_\cdot &= X_0 + \int_0^\cdot \mu(X_s)ds + \int_0^\cdot \sigma(X_s)dW_s \\ Y_\cdot &= g(X_T) + \int_\cdot^T \sum_{k\geq 0} p_k (Y_t)^k dt - \int_\cdot^T Z_t dW_t \end{aligned}$$

where $(p_k)_{k\geq 0}$ forms a probability mass and W a Brownian motion, the solution $Y_\cdot$ coincides with that of the PDE of Kolmogorov-Petrovskii-Piskunov (KPP)

$$\begin{aligned} \partial_t u(t,x) + \mathcal{L}u(t,x) + \sum_{k\geq 0} p_k u^k(t,x) = 0 \\ u(T,x) = g(x) \end{aligned}$$

where $\mathcal{L}$ is the Dynkin operator associated to the process X. Proposing a numerical algorithm for $Y_\cdot$ amounts to solving the KPP equation numerically. The latter could be expressed in a "forward" and probabilist way via the branching process, as discussed in [27], [29] and [28]. Let $\mathcal{K}_t^n$ be the set of living particles of $n-$th generation at time t, $\bar{\mathcal{K}}_t^n$ be the set of particles of $n-$th generation that are born before t. These particles are indexed by k and denoted by X^k. Moreover, define

$$\mathcal{K}_t := \cup_{n\geq 1}\mathcal{K}_t^n, \quad \bar{\mathcal{K}}_t := \cup_{n\geq 1}\bar{\mathcal{K}}_t^n,$$

which designates respectively the living particles at and born before t. Let ρ be the survival probability density function of these particles, $\bar{F}$ be 1 minus the cumulative probability function. Then the solution $u(.,.)$ is equal to the expectation of the following product

$$u(t,x) = \mathbb{E}\Big(\prod_{k\in\mathcal{K}_T} \frac{g(X_T^k)}{\overline{F}(\delta_k)}\Big)\Big(\prod_{k\in\bar{\mathcal{K}}_T\setminus\mathcal{K}_T} \frac{1}{\rho(\delta_k)}\Big).$$

where δ_k is the longevity of the particle indexed by k. While the afore-cited work is focused on the driver of the form of polynomial, we are interested in the generalization of this method to an arbitrary Lipschitze driver.

2.2.2 Algorithm

Based on the work presented in the previous section, we look forward to at the first place approximating an arbitrary Lipschitze driver by a series of polynomials. This approximation is however infeasible in general. The reason being that the solution of a BSDE

with polynomial driver, of degree superior or equal to two, would typically explode in a finite time horizon. The situation becomes even more complicated when the degree of the polynomial is higher.

We thus propose an approximation by a series of local polynomials. Let $\tilde{f}_{\ell_\circ}(x,y)$ be defined as

$$\tilde{f}_{\ell_\circ} : (x,y) \in \mathbb{R}^d \times \mathbb{R} \mapsto \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y) \ .$$

where the functions $(a_{j,l}, \varphi_j)_{\ell \leq \ell_\circ, j \leq j_\circ}$ are continuous and bounded. Given a partition $(A_i)_i$ of $\mathbb{R}$, the functions $(\varphi_j)_{j \leq j_\circ}$ could be interpreted as the smooth kernel to move from one subset to another in a Lipschitze way. We establish at the first place the convergence of the series of the solutions corresponding to these approximative drivers $\tilde{f}$ towards the solution of the original BSDE.

The argument y appears in two different places in the above definition of the local polynomials $\tilde{f}_{\ell_\circ}$. The first place is the monomial part, while the second is the localization part. In order to apply the results on global polynomial, we are going to decouple these two different roles of the argument y. Define $f_{\ell_\circ}(x,y,y')$ by

$$f_{\ell_\circ} : (x,y,y') \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \mapsto \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y') \ .$$

This driver is a global polynomial on y, with $f_{\ell_\circ}(x,y,y) = \tilde{f}_{\ell_\circ}(x,y)$. We then use the Picard iteration to obtain the solution of such a BSDE. The solution of the m-th iteration is given by

$$Y^m_\cdot = g(W_T) + \int_\cdot^T f_{\ell_\circ}(X, Y^m_t, Y^{m-1}_t) dt - \int_\cdot^T Z_t dW_t \ .$$

To present Y^m in a probabilist way as the in the precedent section, we are going to first of all divide the time interval $[0,T]$ into n sub-intervals delimited by $(t^n_i)_{i \leq n}$. The length of these sub-intervals is chosen in a way such that the solution Y^m remains bounded by a given constant. Then a recurrence argument allows us to expand this probabilist representation over the entire interval $[0,T]$.

On contrast to the classical schemes of Picard iteration applied to BSDEs, for example [6], a precise estimation of the path set of the solution Y on each iteration is no longer required. Actually, if the local polynomial is distributed by a partition $(A_i)_i$ of $\mathbb{R}$, then it suffices to know in which set A_i the solution remains at the branching moments.

2.2.3 Contribution

Define recursively

$$\begin{aligned} v^0 &:= y \quad , \\ v^m(t,x) &:= \mathbb{E}\left[V^m_{t,x}\right], \quad m \geq 1, \end{aligned}$$

where $y : [0,T] \times \mathbb{R}$ is a bounded determinist function, and V^m is defined as function of v^{m-1}, v^m by

$$\begin{aligned} V^m_{t,x} &:= \Big(\prod_{k \in \mathcal{K}_{t^n_{i+1}-t}} G^m_{t,x}(k) \Big) \Big(\prod_{k \in \bar{\mathcal{K}}_{t^n_{i+1}-t} \setminus \mathcal{K}_{t^n_{i+1}-t}} A^m_{t,x}(k) \Big), \\ G^m_{t,x}(k) &:= \frac{v^m\big(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}-t}\big)}{\overline{F}(t^n_{i+1} - t - T_{k-})}, \\ A^m_{t,x}(k) &:= \frac{\sum_{j=1}^{j_\circ} a_{j,\xi_k}(X^{x,(k)}_{T_k}) \varphi_j(v^{m-1}(t+T_k, X^{x,(k)}_{T_k}))}{p_{\xi_k}\, \rho(\delta_k)}, \quad \forall (t,x) \in [t^n_i, t^n_{i+1}) \times \mathbf{X}. \end{aligned}$$

in which the particle indexed by k is born at T_{k-} and dies at T_k. We have shown that the series of v^m converges towards the solution of the BSDE with driver f_{ℓ_0}. We apply at the end this method to a concrete example to illustrate the performance of our algorithm.

2.3 Publications

All work in this thesis has been published or the object of revision in the academic revues:

Bruno Bouchard, Grégoire Loeper, and Yiyi Zou. Almost-sure hedging with permanent price impact. *Finance and Stochastics*, 20(3):741–771, 2016

B Bouchard, G Loeper, and Y Zou. Hedging of covered options with linear market impact and gamma constraint. *arXiv preprint arXiv:1512.07087*, 2015. (en révision mineur pour SIAM SICON)

Bruno Bouchard, Xiaolu Tan, and Yiyi Zou. Numerical approximation of bsdes using local polynomial drivers and branching processes. *arXiv preprint arXiv:1612.06790*, 2016. (soumis)

Chapter 3

Almost-sure hedging with permanent price impact

Contents

3.1	**Portfolio and price dynamics**	**23**
3.1.1	Impact rules	23
3.1.2	Discrete rebalancing from a continuous signal and continuous time trading limit	24
3.1.3	Jumps and large orders splitting	28
3.2	**Super-hedging of a European claim**	**31**
3.2.1	Super-hedging price	32
3.2.2	Dynamic programming	33
3.2.3	Pricing equation	34
3.2.4	An example: the fixed impact case	37
3.2.5	Proof of the pde characterization	37
3.3	**Appendix**	**45**

Introduction

Two fundamental assumptions in the Black and Scholes approach for option hedging are that the price dynamics are unaffected by the hedger's behavior, and that he can trade unrestricted amounts of asset at the instantaneous value of the price process. In other words, it relies on the absence of market impact and of liquidity costs or liquidity constraints. This work addresses the problem of option hedging under a price dynamics model that incorporates directly the hedger's trading activity, and hence that violates those two assumptions.

In the literature, one finds numerous studies related to this topic. Some of them incorporate liquidity costs but no price impact, the price curve is not affected by the trading strategy. In the setting of [17], this does not affect the super-hedging price because trading can essentially be done in a bounded variation manner at the marginal spot price at the origine of the curve. However, if additional restrictions are imposed on admissible strategies, this leads to a modified pricing equation, which exhibits a quadratic term in the second order derivative of the solution, and renders the pricing equation fully non-linear, and even not unconditionally parabolic, see [18] and [51]. Another branch of literature focuses on the derivation of the price dynamics through clearing condition. In the papers [25], [46], [45], the authors work on supply and demand curves that arise

from "reference" and "program" traders (i.e. option hedgers) to establish a modified price dynamics, but do not take into account the liquidity costs, see also [36]. This approach also leads to non-linear pde's, but the non-linearity comes from a modified volatility process rather than from a liquidity cost source term. Finally, the series of papers [48], [50], [41] address the liquidity issue indirectly by imposing bounds on the "gamma" of admissible trading strategies, no liquidity cost or price impact are modeled explicitly.

More recently, [37] and [1] have considered a novel approach in which the price dynamic is driven by the sum of a classical Wiener process and a (locally) linear market impact term. The linear market impact mechanism induces a modified volatility process, as well as a non trivial average execution price. However, the trader starts his hedging with the correct position in stocks and does not have to unwind his final position (this corresponds to "covered" options with delivery). Those combined effects lead to a fully non-linear pde giving the exact replication strategy, which is not always parabolic depending on the ratio between the instantaneous market impact (liquidity costs) and permanent market impact.

In this chapter we build on the same framework as [37], in the case where the instantaneous market impact equals the permanent impact (no relaxation effect), and go one step further by considering the effect of (possibly) unwinding the portfolio at maturity, and of building the initial portfolio. Consequently the spot "jumps" at initial time when building the hedge portfolio, and at maturity when unwinding it (depending on the nature of the payoff - delivery can also be made in stocks). In this framework, we find that the optimal super-replication strategy follows a modified quasi-linear Black and Scholes pde. Although the underlying model is similar to the one proposed by the second author [37], the pricing pde is therefore fundamentally different (quasi-linear vs fully non-linear).

Concerning the mathematical approach, while in [37] the author focused on exhibiting an exact replication strategy by a verification approach, in this work we follow a stochastic target approach and derive the pde from a dynamic programming principle. The difficulty is that, because of the market impact mechanism, the state process must be described by the asset price and the hedger's portfolio (i.e. the amount of risky asset detained by the hedger) and this leads to a highly singular control problem. It is overcome by a suitable change of variable which allows one to reduce to a zero initial position in the risky asset and state a version of the geometric dynamic programming principle in terms of the post-portfolio liquidation asset price process: the price that would be obtained if the trader was liquidating his position immediately.

This chapter is organized as follows. In Section 4.1, we present the impact rule and derive continuous time trading dynamics as limits of discrete time rebalancing policies. The super-hedging problem is set in Section 3.2 as a stochastic target problem. We first prove a suitable version of the geometric dynamic programming and then derive the corresponding pde in the viscosity solution sense. Uniqueness and regularity are established under suitable assumptions. We finally further discuss the case of a constant impact coefficients, to provide a better understanding of the "hedging strategy".

General notations. In the rest of this dissertation, Ω is the canonical space of continuous functions on $\mathbb{R}_+$ starting at 0, $\mathbb{P}$ is the Wiener measure, W is the canonical process, and $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is the augmentation of its raw filtration $\mathbb{F}^\circ = (\mathcal{F}_t^\circ)_{t\geq 0}$. All random variables are defined on $(\Omega, \mathcal{F}_\infty, \mathbb{P})$. $\mathbf{L}_0$ (resp. $\mathbf{L}_2$) denotes the space of (resp. square integrable) $\mathbb{R}^n$-valued random variables, while $\mathbf{L}_0^\lambda$ (resp. $\mathbf{L}_2^\lambda$) stands for the collection of predictable $\mathbb{R}^n$-valued processes ϑ (resp. such that $\|\vartheta\|_{\mathbf{L}_2^\lambda} := \mathbb{E}[\int_0^\infty |\vartheta_s|^2 ds]^{\frac{1}{2}}$). Given a stochastic process ξ, ξ^c refers to its continuous part.

Given a function ϕ, we denote by ϕ' and ϕ'' its first and second order derivatives if they exist. When ϕ depends on several arguments, we use the notations $\partial_x\phi$, $\partial^2_{xx}\phi$ to denote the first and second order partial derivatives with respect to its x-argument, and write

$\partial^2_{xy}\phi$ for the cross second order derivative in its (x, y)-argument. We also denote by $|x|$ the Euclidean norm of $x \in \mathbb{R}^n$, the integer $n \geq 1$ is given by the context. Unless otherwise specified, inequalities involving random variables are taken in the $\mathbb{P}-$a.s. sense. We use the convention $x/0 = \text{sign}(x) \times \infty$ with $\text{sign}(0) = +$.

3.1 Portfolio and price dynamics

This section is devoted to the derivation of our model with continuous time trading. We first consider the situation where a trading signal is given by a continuous Itô process and the position in stock is rebalanced in discrete time. In this case, the dynamics of the stock price and the portfolio wealth are given according to our impact rule. A first continuous time trading dynamic is obtained by letting the time between two consecutive trades vanish. Then we incorporate jumps as the limit of continuous trading on a short time horizon.

We restrict ourselves here to single stock market for simplicity, the extension to a multi-dimensional market is just a matter of notations.

3.1.1 Impact rules

We model the impact of a strategy on the price process through an impact function f: the price variation due to buying a (infinitesimal) number $\delta \in \mathbb{R}$ of shares is $\delta f(x)$, if the price of the asset is x before the trade. The cost of buying this additional δ units is given by

$$\delta x + \frac{1}{2}\delta^2 f(x) = \delta \int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota,$$

in which

$$\int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota$$

should be interpreted as the average cost for each additional unit. Between two times of trading $\tau_1 \leq \tau_2$, the dynamics of the stock is given by the strong solution of the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{3.1.1}$$

All over this paper, we assume that

$$\begin{array}{c} f \in C^2_b \text{ and is (strictly) positive,} \\ (\mu, \sigma, \sigma^{-1}) \text{ is Lipschitz and bounded.} \end{array} \tag{H1}$$

Remark 3.1.1. *a. We restrict here to an impact rule which is linear in the size of the order. However, note that in the following it will only be applied to orders of infinitesimal size (at the limit). One would therefore obtain the same final dynamics* (3.1.23)-(3.1.24) *below by considering a more general impact rule $\delta \mapsto F(x, \delta)$ whenever it satisfies*

$$F(x, 0) = \partial^2_{\delta\delta} F(x, 0) = 0 \quad \text{and} \quad \partial_\delta F(x, 0) = f(x).$$

See Remark 3.1.2 below. Otherwise stated, we only need to consider the value and the slope at $\delta = 0$ of the impact function for our analysis.

b. A typical example of such a function is $F = \Delta\mathrm{x}$ where

$$\Delta\mathrm{x}(x, \delta) := \mathrm{x}(x, \delta) - x\,, \tag{3.1.2}$$

with $\mathrm{x}(x,\cdot)$ *defined as the solution of*

$$\mathrm{x}(x,\cdot) \quad = \quad x + \int_0^{\cdot} f(\mathrm{x}(x,s))ds. \tag{3.1.3}$$

The curve x *has a natural interpretation. For an order of small size* $\Delta\iota$*, the stock price jumps from* x *to* $x + \Delta\iota f(x) \simeq \mathrm{x}(x,\Delta\iota)$*. Passing another order of size* $\Delta\iota$ *moves it again to approximately* $\mathrm{x}(\mathrm{x}(x,\Delta\iota),\Delta\iota) = \mathrm{x}(x,2\Delta\iota)$*. Passing to the limit* $\Delta\iota \to 0$ *while keeping the total trade size* δ *provides asymptotically a price move equal to* $\Delta\mathrm{x}(x,\delta)$*.*

This specific curve will play a central role in our analysis, see Section 3.1.3.

3.1.2 Discrete rebalancing from a continuous signal and continuous time trading limit

We first consider the situation in which the number of shares the trader would like to hold is given by a continuous Itô process Y of the form

$$Y = Y_0 + \int_0^{\cdot} b_s ds + \int_0^{\cdot} a_s dW_s, \tag{3.1.4}$$

where

$$(a,b) \in \mathcal{A} := \cup_k \mathcal{A}_k,$$
$$\mathcal{A}_k := \{(a,b) \in \mathbf{L}_0^{\lambda} : |(a,b)| \leq k \ dt \times d\mathbb{P} - \text{a.e.}\} \text{ for } k > 0.$$

In order to derive our continuous time trading dynamics, we consider the corresponding discrete time rebalancing policy set on a time grid

$$t_i^n := iT/n, \ i = 0, \dots, n, \ n \geq 1,$$

and then pass to the limit $n \to \infty$.

If the trader only changes the composition of his portfolio at the discrete times t_i^n, then he holds $Y_{t_i^n}$ stocks on each time interval $[t_i^n, t_{i+1}^n)$. The number of shares actually held at $t \leq T$ is

$$Y_t^n := \sum_{i=0}^{n-1} Y_{t_i^n} \mathbf{1}_{\{t_i^n \leq t < t_{i+1}^n\}} + Y_T \mathbf{1}_{\{t=T\}} \tag{3.1.5}$$

and the number of purchased shares is

$$\delta_t^n := \sum_{i=1}^{n} \mathbf{1}_{\{t=t_i^n\}} (Y_{t_i^n} - Y_{t_{i-1}^n}).$$

Given our impact rule, the corresponding dynamics for the stock price process is

$$X^n = X_0 + \int_0^{\cdot} \mu(X_s^n)ds + \int_0^{\cdot} \sigma(X_s^n)dW_s + \sum_{i=1}^{n} \mathbf{1}_{[t_i^n,T]} \delta_{t_i^n}^n f(X_{t_i^n -}^n), \tag{3.1.6}$$

in which X_0 is a constant.[1]

To describe the portfolio process, we provide the dynamics of the sum V^n of the amount of cash held and the potential amount $Y^n X^n$ associated to the position in stocks:

$$V^n = \text{cash position } + Y^n X^n. \tag{3.1.7}$$

[1] A resilience effect could be modeled by considering a process X° satisfying (3.1.1) on $[0,T]$ and by replacing $\mu(X)$ in the dynamics of X by a drift of the form $\mu(X) - \psi(X - X^\circ)$ for some bounded function ψ such that $x\psi(x) \geq 0$. We do not consider this possibility here, for sake of simplicity.

Observe that this is not the liquidation value of the portfolio except when $Y^n = 0$, as the liquidation of Y^n stocks will have an impact on the price and does not generate a gain equal to $Y^n X^n$. However, if we keep Y^n in mind, the couple (V^n, Y^n) gives the exact composition in cash and stocks of the portfolio. By a slight abuse of language, we call V^n the portfolio value or wealth process.

Assuming zero risk free rate for ease of notations, the dynamics of V^n is given by

$$V^n = V_0 + \int_0^\cdot Y^n_{s-} dX^n_s + \sum_{i=1}^n \mathbf{1}_{[t^n_i, T]} \frac{1}{2} (\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}), \tag{3.1.8}$$

or equivalently

$$\begin{aligned} V^n = V_0 &+ \sum_{i=1}^n \mathbf{1}_{[t^n_{i-1}, T]} Y_{t^n_{i-1}} (X^n_{\cdot \wedge t^n_i -} - X^n_{t^n_{i-1}}) \\ &+ \sum_{i=1}^n \mathbf{1}_{[t^n_i, T]} \left[\frac{1}{2} (\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}) + Y_{t^n_{i-1}} \delta^n_{t^n_i} f(X^n_{t^n_i -}) \right], \end{aligned} \tag{3.1.9}$$

in which $V_0 \in \mathbb{R}$. The first term on the right-hand side corresponds to the evolution of the portfolio value strictly between two trades: it is given by the number of shares held multiplied by the price increment. The second term comes from transaction cost taking place at the time t^n_i. When a trade of size $\delta^n_{t^n_i}$ occurs at time t^n_i, the cost of buying the stocks is $2^{-1}(\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}) + \delta^n_{t^n_i} X^n_{t^n_i -}$. On the other hand, this trade not only adds $\delta^n_{t^n_i}$ more stocks on top of the existing $Y^n_{t^n_i -} = Y_{t^n_{i-1}}$ units in the portfolio, it also moves the price of the stock to $X^n_{t^n_i}$. The value increment due to the additional position and the price's move is therefore $\delta^n_{t^n_i} X^n_{t^n_i} + Y^n_{t^n_i -}(X^n_{t^n_i} - X^n_{t^n_i -})$. Since $X^n_{t^n_i} - X^n_{t^n_i -} = \delta^n_{t^n_i} f(X^n_{t^n_i -})$, we obtain (3.1.9), a compact version of which is given in (3.1.8).

Our continuous time trading dynamics are obtained by passing to the limit $n \to \infty$, i.e. by considering faster and faster rebalancing strategies.

Proposition 3.1.1. *Let $Z := (X, Y, V)$ where Y is defined as in (3.1.4) for some $(a, b) \in \mathcal{A}$, and (X, V) solves*

$$X = X_0 + \int_0^\cdot \sigma(X_s) dW_s + \int_0^\cdot f(X_s) dY_s + \int_0^\cdot (\mu(X_s) + a_s (\sigma f')(X_s)) ds \tag{3.1.10}$$

and

$$V = V_0 + \int_0^\cdot Y_s dX_s + \frac{1}{2} \int_0^\cdot a_s^2 f(X_s) ds. \tag{3.1.11}$$

Let $Z^n := (X^n, Y^n, V^n)$ be defined as in (3.1.6)-(3.1.5)-(3.1.8). Then, there exists a constant $C > 0$ such that

$$\sup_{[0,T]} \mathbb{E}\left[|Z^n - Z|^2 \right] \le C n^{-1}$$

for all $n \ge 1$.

Proof. This follows standard arguments and we only provide the main ideas. In all proofs of this chapter, we denote by C a generic positive constant which does not depend on n nor $i \le n$, and may change from line to line. We shall use repeatedly (H1) and the fact that a and b are bounded by some constant k, in the $dt \times d\mathbb{P}$-a.e. sense.

a. The convergence of the process Y^n is obvious:

$$\sup_{[0,T]} \mathbb{E}\left[|Y^n - Y|^2 \right] \le C n^{-1}. \tag{3.1.12}$$

For later use, set $\Delta X^n := X - X^n$ and also observe that the estimate

$$\sup_{[t^n_{i-1},t^n_i)} \mathbb{E}\left[|\Delta X^n|^2\right] \le \mathbb{E}\left[|\Delta X^n_{t^n_{i-1}}|^2\right](1+Cn^{-1}) + Cn^{-1} \tag{3.1.13}$$

is standard. We now set

$$\tilde{X}^n_t := X^n_t + A^{i,n}_t + B^{i,n}_t, \;\; t^n_{i-1} \le t < t^n_i,$$

where

$$\begin{aligned} A^{i,n}_t &:= \int_{t^n_{i-1}}^t f(X^n_s)dY_s + \int_{t^n_{i-1}}^t a_s(\sigma f')(X^n_s)ds \\ B^{i,n}_t &:= \int_{t^n_{i-1}}^t (Y_s - Y_{t^n_{i-1}})(\mu f' + \frac{1}{2}\sigma^2 f'')(X^n_s)ds + \int_{t^n_{i-1}}^t (Y_s - Y_{t^n_{i-1}})(\sigma f')(X^n_s)dW_s. \end{aligned}$$

Since $A^{i,n}_{t^n_i} + B^{i,n}_{t^n_i} = \delta^n_{t^n_i} f(X^n_{t^n_i -})$, we have

$$\lim_{t\to t^n_i} \tilde{X}^n_t = X^n_{t^n_i}.$$

Set $\Delta\tilde{X}^n := X - \tilde{X}^n$, $\beta^1 := bf + a\sigma f'$ and $\beta^2 := af$, so that

$$\begin{aligned} d|\Delta\tilde{X}^n_t|^2 &= 2\Delta\tilde{X}^n_t[(\mu+\beta^1_t)(X_t) - (\mu+\beta^1_t)(X^n_t)]dt \\ &\quad + [(\sigma+\beta^2_t)(X_t) - (\sigma+\beta^2_t)(X^n_t) - (Y_t - Y_{t^n_{i-1}})(\sigma f')(X^n_t)]^2 dt \\ &\quad - 2\Delta\tilde{X}^n_t(Y_t - Y_{t^n_{i-1}})(\mu f' + \frac{1}{2}\sigma^2 f'')(X^n_t)dt \\ &\quad + 2\Delta\tilde{X}^n_t[(\sigma+\beta^2_t)(X_t) - (\sigma+\beta^2_t)(X^n_t)]dW_t \\ &\quad - 2\Delta\tilde{X}^n_t(Y_t - Y_{t^n_{i-1}})(\sigma f')(X^n_t)dW_t. \end{aligned}$$

In view of (3.1.12)-(3.1.13), this implies

$$\begin{aligned} \mathbb{E}\left[|\Delta\tilde{X}^n_t|^2\right] &\le \mathbb{E}\left[|\Delta X^n_{t^n_{i-1}}|^2\right] + C\mathbb{E}\left[\int_{t^n_{i-1}}^t (|\Delta\tilde{X}^n_s|^2 + |X_s - X^n_s|^2 + |Y_s - Y_{t^n_{i-1}}|^2)ds\right] \\ &\le \mathbb{E}\left[|\Delta X^n_{t^n_{i-1}}|^2\right](1+Cn^{-1}) + C\mathbb{E}\left[\int_{t^n_{i-1}}^t |\Delta\tilde{X}^n_s|^2 ds + n^{-2}\right], \end{aligned}$$

and therefore

$$\sup_{[t^n_{i-1},t^n_i)} \mathbb{E}\left[|\Delta\tilde{X}^n|^2\right] \;\le\; \mathbb{E}\left[|\Delta X^n_{t^n_{i-1}}|^2\right](1+Cn^{-1}) + Cn^{-2}, \tag{3.1.14}$$

by Gronwall's Lemma. Since $\lim_{t\to t^n_i} \tilde{X}^n_t = X^n_{t^n_i}$, this shows that

$$\mathbb{E}\left[|\Delta X^n_{t^n_i}|^2\right] \le \sup_{[t^n_{i-1},t^n_i)} \mathbb{E}\left[|\Delta\tilde{X}^n|^2\right] \le Cn^{-1} \quad \text{for all } i \le n.$$

Plugging this inequality in (3.1.13), we then deduce

$$\sup_{[t^n_{i-1},t^n_i]} \mathbb{E}\left[|\Delta X^n|^2\right] \;\le\; Cn^{-1} \quad \text{for all } i \le n. \tag{3.1.15}$$

b. We now consider the difference $V - V^n$. It follows from (3.1.9) that

$$\begin{aligned} V^n_{t^n_i} &= V^n_{t^n_{i-1}} + \int_{t^n_{i-1}}^{t^n_i} Y_{t^n_{i-1}} \mu(X^n_s) ds + \int_{t^n_{i-1}}^{t^n_i} Y_{t^n_{i-1}} \sigma(X^n_s) dW_s \\ &\quad + \int_{t^n_{i-1}}^{t^n_i} \left(\frac{1}{2} a^2_s f(X^n_s) + Y_{t^n_{i-1}} a_s (f'\sigma)(X^n_s) \right) ds + \int_{t^n_{i-1}}^{t^n_i} Y_{t^n_{i-1}} f(X^n_s) dY_s \\ &\quad + \int_{t^n_{i-1}}^{t^n_i} \alpha^{1n}_s ds + \int_{t^n_{i-1}}^{t^n_i} \alpha^{2n}_s dW_s \end{aligned}$$

where, by (3.1.12), α^{1n} and α^{2n} are adapted processes satisfying

$$\sup_{[t^n_{i-1}, t^n_i)} \mathbb{E}[|\alpha^{1n}|^2 + |\alpha^{2n}|^2] \le Cn^{-1}.$$

In view of (3.1.12)-(3.1.15), this leads to

$$V^n_{t^n_i} = V^n_{t^n_{i-1}} + V_{t^n_i} - V_{t^n_{i-1}} + \int_{t^n_{i-1}}^{t^n_i} \gamma^{1n}_s ds + \int_{t^n_{i-1}}^{t^n_i} \gamma^{2n}_s dW_s \tag{3.1.16}$$

where γ^{1n} and γ^{2n} are adapted processes satisfying

$$\sup_{[t^n_{i-1}, t^n_i)} \mathbb{E}[|\gamma^{1n}|^2 + |\gamma^{2n}|^2] \le Cn^{-1}. \tag{3.1.17}$$

Set

$$\tilde{V}^n_t := V^n_{t^n_{i-1}} + V_t - V_{t^n_{i-1}} + \int_{t^n_{i-1}}^{t} \gamma^{1n}_s ds + \int_{t^n_{i-1}}^{t} \gamma^{2n}_s dW_s, \quad t^n_{i-1} \le t < t^n_i.$$

Then, by applying Itô's Lemma to $|\tilde{V}^n_t - V_t|^2$, using (3.1.17) and Gronwall's Lemma, we obtain

$$\sup_{[t^n_{i-1}, t^n_i)} \mathbb{E}\left[|\tilde{V}^n - V|^2 \right] \le \mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2 \right] (1 + Cn^{-1}) + Cn^{-2},$$

so that, by the identity $\lim_{t \to t^n_i} \tilde{V}^n_t = V^n_{t^n_i}$, recall (3.1.16), and an induction,

$$\mathbb{E}\left[|V^n_{t^n_i} - V_{t^n_i}|^2 \right] \le Cn^{-1}, \quad i \le n.$$

We conclude by observing that

$$\begin{aligned} \mathbb{E}\left[|V^n_t - V_t|^2 \right] &\le C\mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2 + |V^n_{t^n_{i-1}} - V^n_t|^2 + |V_{t^n_{i-1}} - V_t|^2 \right] \\ &\le C\left(\mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2 \right] + n^{-1} \right), \end{aligned}$$

for $t^n_{i-1} \le t < t^n_i$. □

Remark 3.1.2. *If the impact function $\delta f(x)$ was replaced by a more general C^2_b one of the form $F(x, \delta)$, with $F(x, 0) = \partial^2_{\delta\delta} F(x, 0) = 0$, the computations made in the above proof would only lead to terms of the form $\partial_\delta F(X, 0) dY$ and $a\sigma(X)\partial^2_{x\delta} F(X, 0)$ in place of $f(X)dY$ and $a(\sigma f')(X)$ in the dynamics (4.1.2). Similarly, the term $a^2 f(X)$ would be replaced by $a^2 \partial_\delta F(X, 0)$ in (4.1.3).*

3.1.3 Jumps and large orders splitting

We now explain how to incorporate jumps in our dynamics. Let $\mathcal{U}_k$ denote the set of random $\{0, \cdots, k\}$-valued measures ν supported by $[-k,k]\times[0,T]$ that are adapted in the sense that $t \mapsto \nu(A \times [0,t])$ is adapted for all Borel subset A of $[-k,k]$. We set

$$\mathcal{U} := \cup_{k\geq 0}\, \mathcal{U}_k.$$

Note that an element ν of $\mathcal{U}$ can be written in the form

$$\nu(A,[0,t]) = \sum_{j=1}^{k} \mathbf{1}_{\{(\delta_j,\tau_j)\in A\times[0,t]\}} \tag{3.1.18}$$

in which $0 \leq \tau_1 < \cdots < \tau_k \leq T$ are stopping times and each δ_j is a real-valued $\mathcal{F}_{\tau_j}$-random variable.

Then, given $(a,b,\nu) \in \mathcal{A}\times\mathcal{U}$, we define the trading signal as

$$Y = Y_{0-} + \int_0^{\cdot} b_s ds + \int_0^{\cdot} a_s dW_s + \int_0^{\cdot}\int \delta\nu(d\delta, ds), \tag{3.1.19}$$

where $Y_{0-} \in \mathbb{R}$.

In view of the previous sections, we assume that the dynamics of the stock price and portfolio value processes are given by (4.1.2)-(4.1.3) when Y has no jump. We incorporate jumps by assuming that the trader splits a large order δ_j into small pieces on a short time interval. This is a natural idea in practice which aims at avoiding generating a too large impact, and thus paying too much liquidity cost. Given the asymptotic already derived in the previous section, we can reduce the trader's behavior to the case where the transaction is done continuously at a constant rate δ_j/ε on $[\tau_j, \tau_j+\varepsilon]$, for some $\varepsilon > 0$. Denote by (X_{0-}, V_{0-}) the initial stock price and portfolio value, then the number of stocks in the portfolio associated to a strategy $(a,b,\nu) \in \mathcal{A}_k \times \mathcal{U}_k$ is given by

$$Y^\varepsilon = Y + \sum_{j=1}^{k} \mathbf{1}_{[\tau_j,T]} \left[-\delta_j + \varepsilon^{-1}\delta_j(\cdot \wedge (\tau_j+\varepsilon) - \tau_j)\right], \tag{3.1.20}$$

and the corresponding stock price and portfolio value dynamics are

$$X^\varepsilon = X_{0-} + \int_0^{\cdot} \sigma(X^\varepsilon_s) dW_s + \int_0^{\cdot} f(X^\varepsilon_s) dY^\varepsilon_s + \int_0^{\cdot} (\mu(X^\varepsilon_s) + a_s(\sigma f')(X^\varepsilon_s))ds \tag{3.1.21}$$

$$V^\varepsilon = V_{0-} + \int_0^{\cdot} Y^\varepsilon_s dX^\varepsilon_s + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X^\varepsilon_s) ds. \tag{3.1.22}$$

When passing to the limit $\varepsilon \to 0$, we obtain the convergence of $Z^\varepsilon := (X^\varepsilon, Y^\varepsilon, V^\varepsilon)$ to $Z = (X,Y,V)$ with (X,V) defined in (3.1.23)-(3.1.24) below. In the following, we only state the convergence of the terminal values, see the proof for a more complete description. It uses the curve x defined in (3.1.3) above, recall also (3.1.2).

Proposition 3.1.2. *Given $(a,b,\nu) \in \mathcal{A}\times\mathcal{U}$, let $Z = (X,Y,V)$ be defined by (3.1.19) and*

$$\begin{aligned} X &= X_{0-} + \int_0^{\cdot} \sigma(X_s) dW_s + \int_0^{\cdot} f(X_s) dY^c_s + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s))ds \\ &\quad + \int_0^{\cdot}\int \Delta\mathrm{x}(X_{s-},\delta)\nu(d\delta, ds) \end{aligned} \tag{3.1.23}$$

$$\begin{aligned} V &= V_{0-} + \int_0^{\cdot} Y_s dX^c_s + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X_s) ds \\ &\quad + \int_0^{\cdot}\int \left(Y_{s-}\Delta\mathrm{x}(X_{s-},\delta) + \mathfrak{I}(X_{s-},\delta)\right)\nu(d\delta, ds) \end{aligned} \tag{3.1.24}$$

where

$$\mathfrak{I}(x,z) := \int_0^z s f(\mathrm{x}(x,s))ds, \quad \textit{for } x,z \in \mathbb{R}. \tag{3.1.25}$$

Set $Z^\varepsilon := (X^\varepsilon, V^\varepsilon, Y^\varepsilon)$. *Then, there exists a constant* $C > 0$ *such that*

$$\mathbb{E}\left[|Z^\varepsilon_{T+\varepsilon} - Z_T|^2\right] \le C(\varepsilon + \mathbb{P}[\sup_{t\le T} \nu(\mathbb{R}, [t, t+\varepsilon]) \ge 2]^{\frac{1}{2}}),$$

for all $\varepsilon \in (0,1)$. *Moreover,*

$$\lim_{\varepsilon\to 0} \mathbb{P}[\sup_{t\le T} \nu(\mathbb{R}, [t, t+\varepsilon]) \ge 2] = 0.$$

Proof. Let ν be of the form (3.1.18) for some $k \ge 0$ and note that the last claim simply follows from the fact that $\{\tau_{j+1} - \tau_j \ge \varepsilon\} \uparrow \Omega$ up to a $\mathbb{P}$-null set for all $j \le k$.
Step 1. We first consider the case where $\tau_{j+1} \ge \tau_j + \varepsilon$ for all $j \ge 1$. Again, the estimate on $|Z^\varepsilon_{T+\varepsilon} - Z_T|$ follows from simple observations and standard estimates, and we only highlight the main ideas. We will indeed prove that for $1 \le j \le k+1$

$$\mathbb{E}\left[\sup_{[\tau_{j-1}+\varepsilon, \tau_j)} |Z - Z^\varepsilon|^2 + \sup_{0\le s\le\varepsilon} \mathbb{E}[|Z_{\tau_j+s} - Z^\varepsilon_{\tau_j+\varepsilon}|^2\right] \le C\varepsilon, \tag{3.1.26}$$

where we use the convention $\tau_0 = 0$ and $\tau_{k+1} = T$. The result is trivial for (Y^ε, Y) since they are equal on each interval $[\tau_{j-1}+\varepsilon, \tau_j)$ and (a,b) is bounded.

a. We first prove a stronger result for (X^ε, X). Fix $p \in \{2,4\}$. Let x^ε be the solution of the ordinary differential equation

$$\mathrm{x}^\varepsilon_t = X_{\tau_j-} + \int_0^t \frac{\delta_j}{\varepsilon} f(\mathrm{x}^\varepsilon_s)ds.$$

Set $\Delta X^\varepsilon := X^\varepsilon - \mathrm{x}^\varepsilon_{\cdot-\tau_j}$. Itô's Lemma leads to

$$\begin{aligned} d(\Delta X^\varepsilon_t)^p = {} & p(\Delta X^\varepsilon_t)^{p-1}\alpha^{1,\varepsilon}_t dt + \frac{p(p-1)}{2}(\Delta X^\varepsilon_t)^{p-2}(\alpha^{2,\varepsilon}_t)^2 dt \\ & + p(\Delta X^\varepsilon_t)^{p-1}\alpha^{2,\varepsilon}_t dW_t \\ & + p\frac{\delta_j}{\varepsilon}(\Delta X^\varepsilon_t)^{p-1}(f(X^\varepsilon_t) - f(\mathrm{x}^\varepsilon_{t-\tau_j}))dt \end{aligned}$$

on $[\tau_j, \tau_j+\varepsilon]$, in which $\alpha^{1,\varepsilon}$ and $\alpha^{2,\varepsilon}$ are bounded processes. The inequality $x^{p-1} \le x^{p-2} + x^p$, the Lipschitz continuity of f and Gronwall's Lemma then imply

$$\sup_{0\le t\le\varepsilon} \mathbb{E}\left[|X^\varepsilon_{\tau_j+t} - \mathrm{x}^\varepsilon_t|^p\right] \le C\mathbb{E}\left[|X^\varepsilon_{\tau_j-} - X_{\tau_j-}|^p + \int_0^\varepsilon |X^\varepsilon_{\tau_j+s} - \mathrm{x}^\varepsilon_s|^{p-2} ds\right].$$

We now use a simple change of variables to obtain

$$\mathrm{x}^\varepsilon_\varepsilon = \mathrm{x}(X_{\tau_j-}, \delta_j) = X_{\tau_j},$$

in which x is defined in (3.1.3), while

$$\sup_{0\le t\le\varepsilon} \mathbb{E}\left[|X_{\tau_j+t} - X_{\tau_j}|^p\right] \le C\varepsilon^{\frac{p}{2}}.$$

Since X and X^ε have the same dynamics on $[\tau_j+\varepsilon, \tau_{j+1})$, this shows that

$$\begin{aligned}\mathbb{E}\left[\sup_{[\tau_j+\varepsilon,\tau_{j+1})}|X_t - X_t^\varepsilon|^p\right] &\le C\mathbb{E}\left[|X_{\tau_j+\varepsilon} - X^\varepsilon_{\tau_j+\varepsilon}|^p\right] \\ &\le C\mathbb{E}\left[|\mathrm{x}^\varepsilon_\varepsilon - X^\varepsilon_{\tau_j+\varepsilon}|^p + |X_{\tau_j+\varepsilon} - X_{\tau_j}|^p\right] \\ &\le C\mathbb{E}\left[|X^\varepsilon_{\tau_j-} - X_{\tau_j-}|^p + \int_0^\varepsilon |X^\varepsilon_{\tau_j+s} - \mathrm{x}^\varepsilon_s|^{p-2}ds + \varepsilon^{\frac{p}{2}}\right].\end{aligned}$$

For $p = 2$, this provides

$$\mathbb{E}\left[\sup_{[\tau_{j-1}+\varepsilon,\tau_j)}|X - X^\varepsilon|^p + \sup_{0\le s\le\varepsilon}\mathbb{E}[|X_{\tau_j+s} - X^\varepsilon_{\tau_j+\varepsilon}|^p\right] \le C\varepsilon^{\frac{p}{2}},$$

by induction over j, and the case $p = 4$ then follows from the above. For later use, note that the estimate

$$\sup_{0\le t\le\varepsilon}\mathbb{E}\left[|X^\varepsilon_{\tau_j+t} - \mathrm{x}^\varepsilon_t|^4\right] \le C\varepsilon^2 \tag{3.1.27}$$

is a by-product of our analysis.

b. The estimate on $V - V^\varepsilon$ is proved similarly. We introduce

$$\mathrm{v}^\varepsilon_t := V_{\tau_j-} + \int_0^t \frac{\delta_j^2}{\varepsilon^2}sf(\mathrm{x}^\varepsilon_s)ds + Y_{\tau_j-}\int_0^t\frac{\delta_j}{\varepsilon}f(\mathrm{x}^\varepsilon_s)ds = V_{\tau_j-} + \int_0^t Y^\varepsilon_s\frac{\delta_j}{\varepsilon}f(\mathrm{x}^\varepsilon_s)ds,$$

and obtain a first estimate by using (3.1.27):

$$\begin{aligned}\mathbb{E}\left[|V^\varepsilon_{\tau_j+t} - \mathrm{v}^\varepsilon_t|^2\right] &\le C\mathbb{E}\left[|V^\varepsilon_{\tau_j-} - V_{\tau_j-}|^2 + \varepsilon + \left(\int_0^\varepsilon \varepsilon^{-1}Y^\varepsilon_{\tau_j+s}\delta_j|X^\varepsilon_{\tau_j+s} - \mathrm{x}^\varepsilon_s|ds\right)^2\right] \\ &\le C\mathbb{E}\left[|V^\varepsilon_{\tau_j-} - V_{\tau_j-}|^2 + \varepsilon\right],\end{aligned}$$

for $0 \le t \le \varepsilon$. Then, we observe that

$$\mathrm{v}^\varepsilon_\varepsilon = V_{\tau_j-} + \mathfrak{I}(X_{\tau_j-}, \delta_j) + Y_{\tau_j-}\Delta\mathrm{x}(X_{\tau_j-}, \delta_j) = V_{\tau_j},$$

while

$$\sup_{0\le t\le\varepsilon}\mathbb{E}\left[|V_{\tau_j+t} - V_{\tau_j}|^2\right] \le C\varepsilon.$$

By using the estimate on $X - X^\varepsilon$ obtained in a., we then show that

$$\mathbb{E}\left[\sup_{[\tau_j+\varepsilon,\tau_{j+1})}|V_t - V^\varepsilon_t|^2\right] \;\le\; C\mathbb{E}\left[|V_{\tau_j+\varepsilon} - V^\varepsilon_{\tau_j+\varepsilon}|^2 + \varepsilon\right],$$

and conclude by using an induction over j.

Step 2. We now consider the general case. Define

$$\tau^\varepsilon_{j+1} = (\varepsilon + \tau^\varepsilon_j)\vee\tau_{j+1}\ ,\ \delta^\varepsilon_{j+1} = \int_{(\tau^\varepsilon_j,\tau^\varepsilon_{j+1}]}\delta\nu(d\delta, dt)\ ,\ j \ge 1,$$

where $(\tau^\varepsilon_1, \delta^\varepsilon_1) = (\tau_1, \delta_1)$. On $E_\varepsilon := \{\min_{j\le k-1}(\tau_{j+1} - \tau_j) \ge \varepsilon\}$, $(\tau^\varepsilon_j, \delta^\varepsilon_j)_{j\ge1} = (\tau_j, \delta_j)_{j\ge1}$. Hence, it follows from Step 1. that

$$\mathbb{E}\left[|Z^\varepsilon_{T+\varepsilon} - Z_T|^2\right] \le C\varepsilon + C\mathbb{E}\left[|\tilde{Z}^\varepsilon_{T+\varepsilon}|^4 + |Z_T|^4\right]^{\frac12}\mathbb{P}[E^c_\varepsilon]^{\frac12},$$

in which $\tilde{Z}^\varepsilon$ stands for the dynamics associated to $(\tau_j^\varepsilon, \delta_j^\varepsilon)_{j\geq 1}$. It now follows from standard estimates that $(\tilde{Z}^\varepsilon_{T+\varepsilon})_{0<\varepsilon\leq 1}$ and Z_T are bounded in $\mathbf{L}^4$. □

We conclude this section with a proposition of some important properties of the functions x and $\mathfrak{I}$ appeared in Proposition 3.1.1. They will be used in the subsequent section.

Proposition 3.1.3. *For all $x, y, \iota \in \mathbb{R}$,*

1. $\mathrm{x}(\mathrm{x}(x,\iota), -y-\iota) = \mathrm{x}(x,-y)$,
2. $f(x)\partial_x \mathrm{x}(x,y) = \partial_y \mathrm{x}(x,y) = f(\mathrm{x}(x,y))$,
3. $\mathfrak{I}(\mathrm{x}(\mathrm{x}(x,\iota),-y-\iota), y+\iota) - \mathfrak{I}(\mathrm{x}(x,-y),y) = y\Delta\mathrm{x}(x,\iota) + \mathfrak{I}(x,\iota)$,
4. $f(x)\partial_x\mathfrak{I}(x,y) + \Delta\mathrm{x}(x,y) = \partial_y\mathfrak{I}(x,y) = yf(\mathrm{x}(x,y))$.

Proof. (i) is an immediate consequence of the Lipschitz continuity of the function f, which ensures uniqueness of the ODE defining x in (3.1.3). More generally, it has the flow property, which we shall use in the following arguments. The assertion (ii) is an immediate consequence of the definition of x: $\mathrm{x}(\mathrm{x}(x,\iota), y-\iota) = \mathrm{x}(x,y)$ for $\iota > 0$ and $\partial_y \mathrm{x}(x,0) = f(x)$, so that differentiating at $\iota = 0$ provides (ii). The identity in (iii) follows from direct computations. As for (iv), it suffices to write that $\mathfrak{I}(\mathrm{x}(x,\iota), y-\iota) = \int_\iota^y (t-\iota) f(\mathrm{x}(x,t))dt$ for $\iota > 0$, and again to differentiate at $\iota = 0$. □

Remark 3.1.3. *It follows from Proposition 3.1.3 that our model allows round trips at (exactly) zero cost. If x is the current stock price, v the wealth, and y the number of shares in the portfolio, then performing an immediate jump of size δ makes (x,y,v) jump to $(\mathrm{x}(x,\delta),\ y+\delta,\ v+y\Delta\mathrm{x}(x,\delta)+\mathfrak{I}(x,\delta))$. Passing immediately the opposite order, we come back to the position $(\mathrm{x}(\mathrm{x}(x,\delta),-\delta),\ y+\delta-\delta,\ v+y\Delta\mathrm{x}(x,\delta)+\mathfrak{I}(x,\delta)+(y+\delta)\Delta\mathrm{x}(\mathrm{x}(x,\delta),-\delta)+\mathfrak{I}(\mathrm{x}(x,\delta),-\delta)) = (x,y,v)$, by Proposition 3.1.3(i)-(iii). This is a desirable property if one wants to have a chance to hedge options perfectly, or more generally to obtain a non-degenerated super-hedging price.*

3.2 Super-hedging of a European claim

We now turn to the super-hedging problem. From now on, we define the admissible strategies as the Itô processes of the form

$$Y = y + \int_0^\cdot b_s ds + \int_0^\cdot a_s dW_s + \int_0^\cdot \int \delta\nu(d\delta, ds) \tag{3.2.1}$$

in which $y \in \mathbb{R}$, $(a,b,\nu) \in \mathcal{A}\times\mathcal{U}$ and Y is essentially bounded. If $|Y| \leq k$ and $(a,b,\nu) \in \mathcal{A}_k \times \mathcal{U}_k$, then we say that $(a,b,\nu) \in \Gamma_k$, $k \geq 1$, and we let

$$\Gamma := \cup_{k\geq 1}\Gamma_k.$$

We will comment in Remark 3.2.1 below the reason why we are restricted to bounded controls.

Given $(t,z) \in \mathrm{D} := [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$, we define

$$Z^{t,z,\gamma} := (X^{t,z,\gamma}, Y^{t,z,\gamma}, V^{t,z,\gamma})$$

as the solution of (3.1.23)-(3.2.1)-(3.1.24) on $[t,T]$ associated to $\gamma \in \Gamma$ and with initial condition $Z^{t,z,\gamma}_{t-} = z$.

3.2.1 Super-hedging price

A European contingent claim is defined by its payoff function, a measurable map $g = (g_0, g_1) : x \in \mathbb{R} \mapsto \mathbb{R}^2$. The first component is the cash-settlement part, i.e. the amount of cash paid at maturity, while g_1 is the delivery part, i.e. the number of stocks to be delivered.

An admissible strategy $\gamma \in \Gamma$ allows to super-hedge the claim associated to the payoff g, starting from the initial condition z at time t if

$$Z_T^{t,z,\gamma} \in \mathrm{G}$$

where

$$\mathrm{G} := \{(x, y, v) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ : \ v - yx \geq g_0(x) \text{ and } y = g_1(x)\}. \tag{3.2.2}$$

Recall that V stands for the frictionless liquidation value of the portfolio, it is the sum of the cash component and the value YX of the stocks held without taking the liquidation impact into account.

Set

$$\mathcal{G}_k(t, z) := \{\gamma \in \Gamma_k \ : \ Z_T^{t,z,\gamma} \in \mathrm{G}\} \ , \ \mathcal{G}(t, z) := \cup_{k\geq 1} \mathcal{G}_k(t, z),$$

and define the super-hedging price as

$$w(t, x) := \inf_{k\geq 1} w_k(t, x) \ \text{ where } \ w_k(t, x) := \inf\{v : \mathcal{G}_k(t, x, 0, v) \neq \emptyset\}.$$

For later use, we precise what are the T-values of these functions.

Proposition 3.2.1. *Define*

$$G_k(x) := \inf\{y\mathrm{x}(x, y) + g_0(\mathrm{x}(x, y)) - \mathfrak{I}(x, y) : |y| \leq k \ s.t. \ y = g_1(\mathrm{x}(x, y))\}, \ x \in \mathbb{R},$$

and $G := \inf_{k\geq 1} G_k$. *Then,*

$$w_k(T, \cdot) = G_k \quad \textit{and} \quad w(T, \cdot) = G. \tag{3.2.3}$$

Proof. Set $z = (x, 0, v)$ and fix $\gamma = (a, b, \nu) \in \Gamma$. By (3.1.23)-(3.1.24), we have

$$Z_T^{T,z,\gamma} = (\mathrm{x}(x, y), y, v + \mathfrak{I}(x, y)) \ \text{ with } y := \int \delta\nu(d\delta, \{T\}).$$

In view of (3.2.2), $Z_T^{T,z,\gamma} \in \mathrm{G}$ is then equivalent to

$$v + \mathfrak{I}(x, y) - y\mathrm{x}(x, y) \geq g_0(\mathrm{x}(x, y)) \ \text{ and } \ y = g_1(\mathrm{x}(x, y)).$$

By definition of w (resp. w_k), we have to compute the minimal v for which this holds for some $y \in \mathbb{R}$ (resp. $|y| \leq k$). □

Remark 3.2.1. *Let us conclude this section with a comment on our choice of the set of bounded controls* Γ.

a. First, this ensures that the dynamics of X, Y *and* V *are well-defined. This could obviously be relaxed by imposing* $\mathbf{L}^2_\lambda$ *bounds. However, note that the bound should anyway be uniform. This is crucial to ensure that the dynamic programming principle stated in Section 3.2.2 is valid, as it uses measurable selection arguments:* $\omega \mapsto \vartheta[\omega] \in \mathbf{L}^\lambda_2$ *does not imply* $\mathbb{E}\left[\|\vartheta[\cdot]\|_{\mathbf{L}^\lambda_2}\right] < \infty$. *See Remark 3.2.2 below for a related discussion.*

b. In the proof of Theorem 3.2.1, we will need to perform a change of measure associated to a martingale of the form $dM = -M\chi^a dW$ *in which* χ^a *may explode at a speed* a^2 *if* a *is not bounded. See Step 1. of the proof of Theorem 3.2.1. In order to ensure that*

this local martingale is well-defined, and is actually a martingale, one should impose very strong integrability conditions on a.

In order to simplify the presentation, we therefore stick to bounded controls. Many other choices are possible. Note however that, in the case $f \equiv 0$, a large class of options leads to hedging strategies in our set Γ, up to a slight payoff smoothing to avoid the explosion of the delta or the gamma at maturity. This implies that, although the perfect hedging strategy may not belong to Γ, at least it is a limit of elements of Γ and the super-hedging prices coincide.

3.2.2 Dynamic programming

Our control problem is a stochastic target problem as that studied in [49]. The aim of this section is to show that it satisfies a version of their geometric dynamic programming principle.

It is important to notice that the value function w is not amenable to dynamic programming per se. The reason being that it assumes a zero initial stock holding at time t, while the position Y_θ will in general not be zero at a later time θ. It is therefore *a priori* not possible to compare the later wealth process V_θ with the corresponding super-hedging price $w(\theta, X_\theta)$. However, a version of the geometric dynamic programming principle can still be obtained if we introduce the process

$$\hat{X}^{t,z,\gamma} := \mathrm{x}(X^{t,z,\gamma}, -Y^{t,z,\gamma}) \tag{3.2.4}$$

which represents the value of the stock immediately after liquidating the stock position.

We refer to Remark 3.2.2 below for the reason why part (ii) of the following dynamic programming principle is stated in terms of $(w_k)_{k\geq 1}$ instead of w.

Proposition 3.2.2 (GDP)**.** *Fix $(t,x,v) \in [0,T] \times \mathbb{R} \times \mathbb{R}$.*

1. *If $v > w(t,x)$ then there exists $\gamma \in \Gamma$ and $y \in \mathbb{R}$ such that*

$$V_\theta^{t,z,\gamma} \quad \geq \quad w(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma}),$$

   for all stopping time $\theta \geq t$, where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

2. *Fix $k \geq 1$. If $v < w_{2k+2}(t,x)$ then we can not find $\gamma \in \Gamma_k$, $y \in [-k,k]$ and a stopping time $\theta \geq t$ such that*

$$V_\theta^{t,z,\gamma} > w_k(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$$

   with $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

Proof. Step 1. In order to transform our stochastic target problem into a time consistent one, we introduce the auxiliary value function corresponding to an initial holding y in stocks:

$$\hat{w}(t,x,y) := \inf_{k\geq 1} \hat{w}_k(t,x,y) \quad \text{where} \quad \hat{w}_k(t,x,y) := \inf\{v : \mathcal{G}_k(t,x,y,v) \neq \emptyset\}.$$

Note that $w_{k+1}(t,x) \leq \inf\{v : \exists\, y \in [-k,k] \text{ s.t. } \mathcal{G}_k(t, \mathrm{x}(x,y), y, v + \mathfrak{I}(x,y)) \neq \emptyset\}$, which follows from (3.1.23)-(3.1.24). Since $\mathrm{x}(\mathrm{x}(x,-y),y) = x$, see Proposition 3.1.3, this implies

$$\hat{w}_k(t,x,y) \geq w_{k+1}(t, \mathrm{x}(x,-y)) + \mathfrak{I}(\mathrm{x}(x,-y), y), \tag{3.2.5}$$

for $|y| \leq k$. Similarly, since $\mathfrak{I}(x,-y) + y\Delta\mathrm{x}(x,-y) = -\mathfrak{I}(\mathrm{x}(x,-y),y)$ by Proposition 3.1.3, we have

$$\hat{w}_{k+1}(t,x,y) \leq w_k(t, \mathrm{x}(x,-y)) + \mathfrak{I}(\mathrm{x}(x,-y), y). \tag{3.2.6}$$

Step 2. a. Assume that $v > w(t,x)$. The definition of w implies that we can find $y \in \mathbb{R}$ and $\gamma \in \mathcal{G}(t,z)$ where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$. By the arguments of [49, Step 1 proof of Theorem 3.1], $V_\theta^{t,z,\gamma} \geq \hat{w}(\theta, X_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, for all stopping times $\theta \geq t$. Then, (3.2.5) applied for $k \to \infty$ provides (i).

b. Assume now that we can find $\gamma \in \Gamma_k$, $y \in [-k,k]$ and a stopping time $\theta \geq t$ such that $V_\theta^{t,z,\gamma} > (w_k + \mathfrak{I})(\theta, \hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$. By (3.2.4)-(3.2.6), $V_\theta^{t,z,\gamma} > \hat{w}_{k+1}(\theta, X_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, and it follows from [49, Step 2 proof of Theorem 3.1] and Corollary 3.3.1 that $v + \mathfrak{I}(x,y) \geq \hat{w}_{2k+1}(t, \mathrm{x}(x,y), y)$. We conclude that (ii) holds by appealing to (3.2.5) and the identities $\mathrm{x}(\mathrm{x}(x,y), -y) = x$ and $\mathfrak{I}(\mathrm{x}(\mathrm{x}(x,y), -y), y) = \mathfrak{I}(x,y)$, see Proposition 3.1.3. □

We conclude this section with purely technical considerations that justify the form of the above dynamic programming principle. They are of no use for the later developments but may help to clarify our approach.

Remark 3.2.2. *Part* (ii) *of Proposition 3.2.2 can not be stated in terms of w. The reason is that the measurable selection techniques can not be used with the set Γ. Indeed, if $\omega \mapsto \gamma[\omega] \in \Gamma$, then the corresponding bounds depend on ω and are not uniform: a measurable family of controls $\{\gamma[\omega], \omega \in \Omega\}$ does not permit to construct an element in Γ. Part* (i) *of Proposition 3.2.2 only relies on a conditioning argument, which can be done within Γ.*

Remark 3.2.3. *A version of the geometric dynamic programming principle also holds for $(\hat{w}_k)_{k\geq 1}$, this is a by-product of the above proof. It is therefore tempting to try to derive a pde for the function $\hat{w}$. However, the fact that the control b appears linearly in the dynamics of (X,Y,V) makes this problem highly singular, and "standard approaches" do not seem to work. We shall see in Lemma 3.2.1 that this singularity disappears in the parameterization $\mathrm{x}(X,-Y)$ used in Proposition 3.2.2. Moreover, hedging implies a control on the diffusion part of the dynamics which is translated into a strong relation between Y and the space gradient $D\hat{w}(\cdot,X,Y)$. This would lead to a pde set on a curve on the coordinates (t,x,y) depending on $D\hat{w}$ (the solution of the pde).*

3.2.3 Pricing equation

In order to understand what would be the partial differential equation w should solve, we state the following key lemma. Although the control b appears linearly in the dynamics of (X,Y,V), this lemma shows that the singularity that may be resulted in does indeed not appear when applying Itô's Lemma to $V - (\varphi + \mathfrak{I})(\cdot, \hat{X}, Y)$, recall (3.2.4). The potential singularity is absorbed by the functions x and $\mathfrak{I}$ (compare with Remark 3.2.3). The proof of this Lemma is postponed to Section 3.2.5.

Lemma 3.2.1. *Fix $(t,x,y,v) \in \mathrm{D}$, $z := (x,y,v)$, $\gamma = (a,b,\nu) \in \Gamma$. Then,*

$$\begin{aligned}\hat{X}^{t,z,\gamma} &= \mathrm{x}(x,-y) \\ &\quad + \int_t^\cdot [\hat{\mu}(\hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma}) + (\partial_x \mathrm{x}\mu - \frac{1}{2}\partial_x \mathrm{x} a_s^2 f f')(X_s^{t,z,\gamma}, -Y_s^{t,z,\gamma})]ds \\ &\quad + \int_t^\cdot \hat{\sigma}(\hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma}) dW_s.\end{aligned}$$

Given $\varphi \in C_b^\infty$, *set* $\mathcal{E}^{t,z,\gamma} := V^{t,z,\gamma} - (\varphi + \mathfrak{I})(\cdot, \hat{X}^{t,z,\gamma}, Y^{t,z,\gamma})$. *Then,*

$$\begin{aligned}\mathcal{E}^{t,z,\gamma} - \mathcal{E}_t^{t,z,\gamma} &= \int_t^\cdot [Y_s^{t,z,\gamma} - \check{Y}_s^{t,z,\gamma}](\mu - f'fa_s^2/2)(X_s^{t,z,\gamma})ds \\ &+ \int_t^\cdot [Y_s^{t,z,\gamma} - \check{Y}_s^{t,z,\gamma}]\sigma(X_s^{t,z,\gamma})dW_s \\ &+ \int_t^\cdot \hat{F}\varphi(s, \hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma})ds\end{aligned}$$

in which

$$\begin{aligned}\check{Y}^{t,z,\gamma} &:= Y^{t,z,\gamma} + \frac{\hat{X}^{t,z,\gamma} - X^{t,z,\gamma}}{f(X^{t,z,\gamma})} + \partial_x\varphi(\cdot, \hat{X}^{t,z,\gamma})\frac{f(\hat{X}^{t,z,\gamma})}{f(X^{t,z,\gamma})} \\ \hat{F}\varphi &:= -\partial_t\varphi - \hat{\mu}\partial_x[\varphi + \mathfrak{I}] - \frac{1}{2}\hat{\sigma}^2\partial_{xx}^2[\varphi + \mathfrak{I}]\end{aligned}$$

and $\hat{\mu}(x', y') := \frac{1}{2}[\partial_{xx}^2 \mathrm{x}\sigma^2](\mathrm{x}(x', y'), -y')$, $\hat{\sigma}(x', y') := (\sigma\partial_x \mathrm{x})(\mathrm{x}(x', y'), -y')$ *for* $(x', y') \in \mathbb{R} \times \mathbb{R}$.

Let us now appeal to Proposition 3.2.2 and apply Lemma 3.2.1 to $\varphi = w$, assuming that w is smooth and that Proposition 3.2.2(i) is valid even if we start from $v = w(t, x)$, i.e. assuming that the inf in the definition of w is a min. With the notations of the above lemma, applying Proposition 3.2.2(i) formally for $\theta = t+$ leads to

$$\begin{aligned}0 &\leq d\mathcal{E}_t^{t,z,\gamma} \\ &= (y - \hat{y})\left\{[\mu - ff'a_t^2/2)(\mathrm{x}(x, y))]dt + \sigma(\mathrm{x}(x, y))dW_t\right\} \\ &\quad + \hat{F}w(t, \hat{x}, y)dt\end{aligned}$$

in which

$$\hat{y} = y + \frac{\hat{x} - \mathrm{x}(x, y)}{f(\mathrm{x}(x, y))} + \partial_x w(t, \hat{x})\frac{f(\hat{x})}{f(\mathrm{x}(x, y))} \quad \text{and} \quad \hat{x} = \mathrm{x}(\mathrm{x}(x, y), -y) = x.$$

Remaining at a formal level, this inequality cannot hold unless $y = \hat{y}$, because $\sigma \neq 0$ and $\hat{F}w(t, x, \hat{y}) = \hat{F}w(t, \hat{x}, y) \geq 0$. This means that w should be a super-solution of

$$F\varphi(t, x) := \hat{F}\varphi(t, x, \hat{y}[\varphi](t, x)) = 0 \tag{3.2.7}$$

where, for a smooth function φ and x^{-1} being the inverse of $\mathrm{x}(x, \cdot)$

$$\hat{y}[\varphi](t, x) := \mathrm{x}^{-1}(x, x + f(x)\partial_x\varphi(t, x)).$$

From (ii) of Proposition 3.2.2, we can deduce formally that the above inequality should actually be an equality, and therefore that w should solve (3.2.7).

In order to give a sense to the above, we assume that

$$\left\{\begin{array}{l} \mathrm{x}(x, \cdot) \text{ is invertible for all } x \in \mathbb{R} \\ (x, z) \in \mathbb{R} \times \mathbb{R} \mapsto \mathrm{x}^{-1}(x, z) \text{ is } C^2. \end{array}\right. \tag{H2}$$

In view of (3.2.3), we therefore expect w to be a solution of

$$F\varphi\mathbf{1}_{[0,T[} + (\varphi - G)\mathbf{1}_{\{T\}} = 0 \text{ on } [0, T] \times \mathbb{R}. \tag{3.2.8}$$

Since w may not be smooth and (ii) of Proposition 3.2.2 is stated in terms of w_k instead of w, we need to refer to the notion of viscosity solutions and the relaxed semi-limits of $(w_k)_{k\geq 1}$. Define

$$w_*(t, x) := \liminf_{(t', x', k) \to (t, x, \infty)} w_k(t', x') \text{ and } w^*(t, x) := \limsup_{(t', x', k) \to (t, x, \infty)} w_k(t', x'),$$

where the limits are taken over $t' < T$ as usual. Note that w_* coincides with the lower-semicontinuous enveloppe of w, as a result of $w = \inf_{k\geq 1} w_k = \lim_{k\to\infty} \downarrow w_k$ by definition.

We are now in position to state the main result of this section. Assume additionally

$$\left\{ \begin{array}{c} G \text{ is continuous and } G_k \downarrow G \text{ uniformly on compact sets.} \\ w_* \text{ and } w^* \text{ are finite on } [0,T]\times\mathbb{R}. \end{array} \right. \tag{H3}$$

The first part of (H3) will be used to obtain the boundary condition. The second part is natural to ensure our problem is not ill-posed.

Theorem 3.2.1 (Pricing equation)**.** *The functions w_* and w^* are respectively a viscosity super- and a subsolution of (3.2.8). If they are bounded and $\inf f > 0$, then $w = w_* = w^*$ and w is the unique bounded viscosity solution of (3.2.8). If in addition G is bounded and C^2 with G, G', G'' Hölder continuous, then $w \in C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$.*

The proof is reported in Section 3.2.5. Let us now discuss the verification counterpart.

Remark 3.2.4 (Verification)**.** *Assume that φ is a smooth solution of (3.2.8) and that we can find $(a,b) \in \mathcal{A}$ such that the following system holds on $[t,T)$:*

$$\begin{array}{rcl} X &=& x + \Delta\mathrm{x}(x, \hat{y}[\varphi](t,x)) + \displaystyle\int_t^\cdot \sigma(X_s)dW_s + \int_0^\cdot f(X_s)dY_s^c \\ && + \displaystyle\int_0^\cdot (\mu(X_s) + a_s(\sigma f')(X_s))ds + \Delta\mathrm{x}(X_{T-}, -Y_{T-})\mathbf{1}_{\{T\}} \\ Y &=& \hat{y}[\varphi](t,x) + \displaystyle\int_t^\cdot b_s ds + \int_t^\cdot a_s dW_s - Y_{T-}\mathbf{1}_{\{T\}} \\ &=& \mathrm{x}^{-1}(\hat{X}, \hat{X} + (f\partial_x\varphi)(\cdot,\hat{X})) - Y_{T-}\mathbf{1}_{\{T\}} \\ \hat{X} &:=& \mathrm{x}(X, -Y) \\ V &=& \varphi(t,x) + \mathfrak{I}(x, \hat{y}[\varphi](t,x)) + \displaystyle\int_t^\cdot Y_s dX_s^c + \frac{1}{2}\int_0^\cdot a_s^2 f(X_s)ds \\ && + (Y_{T-}\Delta\mathrm{x}(X_{T-}, -Y_{T-}) + \mathfrak{I}(X_{T-}, -Y_{T-}))\mathbf{1}_{\{T\}}. \end{array}$$

a. Note that $\hat{X}_t = \mathrm{x}(X_t, -Y_t) = \mathrm{x}(\mathrm{x}(x, \hat{y}[\varphi](t,x)), -\hat{y}[\varphi](t,x)) = x$, recall Proposition 3.1.3(i), so that $Y_t = \hat{y}[\varphi](t,x) = \mathrm{x}^{-1}(\hat{X}_t, \hat{X}_t + (f\partial_x\varphi)(t,\hat{X}_t))$. We therefore need to find (a,b) such that $X = \mathrm{x}(\hat{X}, Y) = \hat{X} + (f\partial_x\varphi)(\cdot,\hat{X})$. This amounts to solving:

$$\begin{array}{rcl} \sigma(X) + f(X)a &=& \hat{\sigma}(\hat{X}, Y)\partial_x\psi(\cdot,\hat{X}) \\ f(X)b + (\mu + a\sigma f')(X) &=& (\hat{\mu}(\hat{X}, Y) + (\partial_x\mathrm{x}\mu - \frac{1}{2}\partial_x\mathrm{x}a_s^2 ff')(X, -Y))\partial_x\psi(\cdot, X) \\ && + \frac{1}{2}\hat{\sigma}^2(\hat{X}, Y)\partial_{xx}^2\psi(\cdot,\hat{X}) \end{array}$$

where $\psi(t,x) := x + (f\partial_x\varphi)(t,x)$. Since $f > 0$, this system has a solution. Moreover, $(a,b) \in \mathcal{A}$ under additional smoothness and boundedness assumption.
b. Let $\check{Y}$ be as in Lemma 3.2.1 for the above dynamics. Since $X = \mathrm{x}(\hat{X}, Y) = \hat{X} + (f\partial_x\varphi)(\cdot,\hat{X})$ on $[t,T)$ by construction, we have $\check{Y} = Y$ on $[t,T)$. Then it follows from Lemma 3.2.1 and (3.2.7)-(3.2.8) that

$$V_{T-} = \varphi(T, \hat{X}_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}) = G(\hat{X}_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}).$$

Since $X_T = \hat{X}_{T-}$ and $Y_{T-}\Delta\mathrm{x}(X_{T-}, -Y_{T-}) + \mathfrak{I}(X_{T-}, -Y_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}) = 0$, see Proposition 3.1.3, this implies $V_T = G(X_T)$. The hedging strategy thence consists in taking an initial position $Y_t = \hat{y}[\varphi](t,x)$ and then applying the control (a,b) up to T. A final trade is performed at T. In particular, the number of stocks Y is continuous on (t,T).

3.2.4 An example: the fixed impact case

In this section, we consider the simple case of a constant impact function f: $f(x) = \lambda > 0$ for all $x \in \mathbb{R}$. This is certainly a too simple model, but this allows us to highlight the structure of our result as the pde simplifies in this case. Indeed, for

$$\mathrm{x}(x,y) = x + y\lambda \quad \text{and} \quad \mathfrak{I}(x,y) = \frac{1}{2}y^2\lambda,$$

we have

$$\hat{\mu}(x,y) = 0 \ , \ \hat{\sigma}(x,y) := \sigma(x + y\lambda) \ , \ \hat{y}[\varphi] := \partial_x\varphi.$$

The pricing equation is given by a local volatility model in which the volatility depends on the hedging price itself, and therefore on the claim (g_0, g_1) to be hedged:

$$0 = -\partial_t\varphi(t,x) - \frac{1}{2}\sigma^2(x + \partial_x\varphi\lambda)\partial^2_{xx}\varphi(t,x).$$

As for the process Y in the verification argument of Remark 3.2.4, it is given by

$$Y = \partial_x\varphi(\cdot, \hat{X}) = \partial_x\varphi(\cdot, X - \lambda Y).$$

This shows that the hedging strategy (if well-defined) consists in following the usual Δ-hedging strategy but for a $\Delta = \partial_x\varphi$ computed at the liquidation price of the stock $\hat{X}$ defined in (3.2.4).

Note that the pricing equation is reduced to the usual heat equation when σ is constant. This fact, showing the limitation of the fixed impact model, is expected. To explain this degeneration, let us consider the simpler case $g_1 = 0, \mu = 0$ and use the notations of Remark 3.2.4. Since σ is constant, the strategy Y does not affect the coefficients in the dynamics of X, it just adds an extra shift λdY each time we buy or sell. The imposition of $Y_T = 0$ (after the final jump) and $Y_{t-} = 0$ leads to a null total impact: $X_T = X_{t-} + \sigma(W_T - W_t)$. As for the wealth process, we have

$$\begin{aligned} V_T &= \varphi(t,x) + \frac{1}{2}Y_t^2\lambda + \int_t^T Y_s dX_s^c + \frac{1}{2}\int_t^T a_s^2\lambda ds - Y_{T-}^2\lambda + \frac{1}{2}Y_{T-}^2\lambda \\ &= \varphi(t,x) + \int_t^T Y_s\sigma dW_s + \frac{1}{2}\lambda(Y_t^2 - Y_{T-}^2) + \int_t^T \lambda Y_s dY_s^c + \frac{1}{2}\int_t^T a_s^2\lambda ds \\ &= \varphi(t,x) + \int_t^T Y_s\sigma dW_s. \end{aligned}$$

In other words, the liquidation costs are cancelled: when buying, the trader pays a cost and increases the price; when selling back, he pays a cost again but sells at a higher price. If there is no impact on the underlying dynamics X and f is constant, these two directions perfectly offset. On contrast, the hedging strategy is affected: $Y = \partial_x\varphi(\cdot, X - \lambda Y)$ on $[0,T)$.

3.2.5 Proof of the pde characterization

3.2.5.1 The key lemma

We first provide the proof of our key result.

Proof of Lemma 3.2.1. To alleviate the notations, we omit the super-scripts.

a. We first observe from Proposition 3.1.3(i) that $\mathrm{x}(X, -Y)$ has continuous paths, while Proposition 3.1.3(ii) implies that $f\partial_x\mathrm{x} - \partial_y\mathrm{x} = 0$ (and therefore $f'\partial_x\mathrm{x} + f\partial^2_{xx}\mathrm{x} - \partial^2_{xy}\mathrm{x} = 0$). Using Itô's Lemma, this leads to

$$\begin{aligned} d\mathrm{x}(X_s, -Y_s) &= (\mu - \frac{1}{2}a_s^2 ff')(X_s)\partial_x\mathrm{x}(X_s, -Y_s)ds + \sigma(X_s)\partial_x\mathrm{x}(X_s, -Y_s)dW_s \\ &\quad + \frac{1}{2}\left[\sigma^2\partial^2_{xx}\mathrm{x} - a_s^2 f\partial^2_{xy}\mathrm{x} + a_s^2\partial^2_{yy}\mathrm{x}\right](X_s, -Y_s)ds. \end{aligned}$$

We now use the identity $f\partial^2_{xy}\mathrm{x} - \partial^2_{yy}\mathrm{x} = 0$, which also follows from Proposition 3.1.3(ii), to simplify the above expression into

$$\begin{aligned} d\mathrm{x}(X_s, -Y_s) &= [\partial_x \mathrm{x}(\mu - \frac{1}{2}a_s^2 ff') + \frac{1}{2}\partial^2_{xx}\mathrm{x}\sigma^2](X_s, -Y_s)ds \\ &\quad + (\sigma\partial_x \mathrm{x})(X_s, -Y_s)dW_s. \end{aligned}$$

b. Similarly, it follows from Proposition 3.1.3(iii) that $V - \mathfrak{I}(\hat{X}, Y)$ has continuous paths, and so does $\mathcal{E}$ by a. Before applying Itô's lemma to derive the dynamics of $\mathcal{E}$, observe that $\partial_y\mathfrak{I}(\mathrm{x}(x,-y),y) = yf(\mathrm{x}(\mathrm{x}(x,-y),y)) = yf(x)$ and that $\partial^2_{yy}\mathfrak{I}(\mathrm{x}(x,-y),y) = y(ff')(x) + f(x)$. Also note that $\hat{\sigma}(\mathrm{x}(x,-y),y) = \sigma(x)\partial_x\mathrm{x}(x,-y)$. Then, using the dynamics of $\hat{X}$ derived above, we obtain

$$\begin{aligned} d\mathcal{E}_s =& (Y_s - \check{Y}_s)\sigma(X_s)dW_s + (Y_s - \check{Y}_s)[\mu - \frac{1}{2}a_s^2(ff')](X_s)ds + \hat{F}\varphi(s, \hat{X}_s, Y_s)ds \\ &+ a_s\sigma(X_s)[Y_s f'(X_s) - \partial_x\mathrm{x}(X_s, -Y_s)\partial^2_{xy}\mathfrak{I}(\hat{X}_s, Y_s)]ds, \end{aligned}$$

where

$$\check{Y} := \partial_x(\varphi + \mathfrak{I})(\cdot, \hat{X}, Y)\partial_x\mathrm{x}(X, -Y).$$

By Proposition 3.1.3(ii)(iv), $f(x)\partial^2_{xy}\mathfrak{I}(x,y) = \partial_y[yf(\mathrm{x}(x,y)) - \Delta\mathrm{x}(x,y)] = y(f'f)(\mathrm{x}(x,y))$. Since $\partial_x\mathrm{x}(x,-y) = f(\mathrm{x}(x,-y))/f(x)$, see Proposition 3.1.3(ii), it follows that

$$\partial_x\mathrm{x}(X,-Y)\partial^2_{xy}\mathfrak{I}(\mathrm{x}(X,-Y),Y) = Yf'(X),$$

which implies

$$d\mathcal{E}_s = (Y_s - \check{Y}_s)\sigma(X_s)dW_s + (Y_s - \check{Y}_s)[\mu - \frac{1}{2}a_s^2(ff')](X_s)ds + \hat{F}\varphi(s, \hat{X}_s, Y_s)ds.$$

We now deduce from Proposition 3.1.3 that

$$\begin{aligned} \partial_x\mathfrak{I}(\hat{X}, Y) &= \frac{-\Delta\mathrm{x}(\hat{X}, Y) + Yf(\mathrm{x}(\hat{X}, Y))}{f(\hat{X})} = \frac{\hat{X} - X + Yf(X)}{f(\hat{X})} \\ \partial_x\mathrm{x}(X, -Y) &= f(\hat{X})/f(X), \end{aligned}$$

so that

$$\check{Y} = \partial_x\varphi(\cdot, \hat{X})\frac{f(\hat{X})}{f(X)} + \frac{\hat{X} - X}{f(X)} + Y.$$

□

3.2.5.2 Super- and subsolution properties

We now prove the super- and subsolution properties of Theorem 3.2.1.

Supersolution property. We first prove the supersolution property. It follows from similar arguments as in [10]. Let φ be a C^∞_b function, and $(t_o, x_o) \in [0,T] \times \mathbb{R}$ be a strict (local) minimum point of $w_* - \varphi$ such that $(w_* - \varphi)(t_o, x_o) = 0$.

a. We first assume that $t_o < T$ and $F\varphi(t_o, x_o) < 0$, and work towards a contradiction. In view of (3.2.7),

$$\hat{F}\varphi(t,x,y) < 0 \text{ if } (t,x) \in B \text{ and } |y - \hat{y}[\varphi](t,x)| \le \varepsilon,$$

for some open ball $B \subset [0,T[\times\mathbb{R}$ which contains (t_o, x_o), and some $\varepsilon > 0$. Since x^{-1} is continuous, this implies that

$$\hat{F}\varphi(t,x,y) < 0 \text{ if } (t,x) \in B \text{ and } |x + \partial_x\varphi(t,x)f(x) - \mathrm{x}(x,y)| \le \varepsilon f(\mathrm{x}(x,y)), \tag{3.2.9}$$

after possibly changing B and ε. Let $(t_n, x_n)_n$ be a sequence in B that converges to (t_o, x_o) and such that $w(t_n, x_n) \to w_*(t_o, x_o)$ (recall that w_* coïncides with the lower-semicontinuous envelope of w). Set $v_n := w(t_n, x_n) + n^{-1}$. It follows from Proposition 3.2.2(i) that we can find $(a^n, b^n, \nu^n) = \gamma_n \in \Gamma$ and $y_n \in \mathbb{R}$ such that

$$V_{\theta_n}^{t_n,z_n,\gamma_n} \geq w(\theta_n, \hat{X}_{\theta_n}^{t_n,z_n,\gamma_n}) + \mathfrak{I}(\hat{X}_\theta^{t_n,z_n,\gamma_n}, Y_{\theta_n}^{t_n,z_n,\gamma_n}), \tag{3.2.10}$$

where $z_n := (\mathrm{x}(x_n, y_n), y_n, v_n + \mathfrak{I}(x_n, y_n))$ and θ_n is the first exit time after t_n of $(\cdot, \hat{X}^{t_n,z_n,\gamma_n})$ from B (note that $\hat{X}_{t_n}^{t_n,z_n,\gamma_n} = \mathrm{x}(\mathrm{x}(x_n, y_n), -y_n) = x_n$). In the following, we use the simplified notations X^n, $\hat{X}^n$, V^n and Y^n for the corresponding quantities indexed by (t_n, z_n, γ_n). Since (t_o, x_o) reaches a strict minimum $w_* - \varphi$, this implies

$$V_{\theta_n}^n \geq \varphi(\theta_n, \hat{X}_{\theta_n}^n) + \mathfrak{I}(\hat{X}_\theta^n, Y_{\theta_n}^n) + \iota, \tag{3.2.11}$$

for some $\iota > 0$. Let $\check{Y}^n$ be as in Lemma 3.2.1 and observe that

$$\check{Y}^n - Y^n = \frac{\hat{X}^n + \partial_x\varphi(\cdot, \hat{X}^n) f(\hat{X}^n) - \mathrm{x}(\hat{X}^n, Y^n)}{f(\mathrm{x}(\hat{X}^n, Y^n))}. \tag{3.2.12}$$

Set

$$\chi^n := \frac{(\mu - f'f(a_s^n)^2/2)(X^n)}{\sigma(X^n)} + \frac{\hat{F}\varphi(\cdot, \hat{X}^n, Y^n)}{(Y^n - \check{Y}^n)\sigma(X^n)} \mathbf{1}_{|Y^n - \check{Y}^n| \geq \varepsilon}$$

and consider the measure $\mathbb{P}^n$ defined by

$$\frac{d\mathbb{P}_n}{d\mathbb{P}} = M_{\theta_n}^n \text{ where } M^n = 1 - \int_{t_n}^{\cdot \wedge \theta_n} M_s^n \chi_s^n dW_s.$$

Then, it follows from (3.2.11), Lemma 3.2.1, (3.2.9) and (3.2.12) that

$$\begin{aligned} \iota &\leq \mathbb{E}^{\mathbb{P}_n}[V_{\theta_n}^n - (\varphi + \mathfrak{I})(\theta_n, \hat{X}_{\theta_n}^n, Y_{\theta_n}^n)] \\ &\leq v_n + \mathfrak{I}(x_n, y_n) - (\varphi + \mathfrak{I})(t_n, \mathrm{x}(\mathrm{x}(x_n, y_n), -y_n), y_n) \\ &= v_n - \varphi(t_n, x_n). \end{aligned}$$

The right-hand side goes to 0, which is the required contradiction.

b. We now explain how to modify the above proof for the case $t_o = T$. After possibly replacing $(t, x) \mapsto \varphi(t, x)$ by $(t, x) \mapsto \varphi(t, x) - \sqrt{T - t}$, we can assume that $\partial_t\varphi(t, x) \to \infty$ as $t \to T$, uniformly in x on each compact set. Then (3.2.9) still holds for B of the form $[T - \eta, T) \times B(x_o)$ in which $B(x_o)$ is an open ball around x_o and $\eta > 0$ small. Assume that $\varphi(T, x_o) < G(x_o)$. Then, after possibly changing $B(x_o)$, we have $\varphi(T, \cdot) \leq G - \iota_1$ on $B(x_o)$, for some $\iota_1 > 0$. Then, with the notations of a., we deduce from (3.2.3)-(3.2.10) that

$$V_{\theta_n}^n \geq \varphi(\theta_n, \hat{X}_{\theta_n}^n) + \mathfrak{I}(\hat{X}_\theta^n, Y_{\theta_n}^n) + \iota_1 \wedge \iota_2,$$

in which $\iota_2 := \min\{(w_* - \varphi)(t, x) : (t, x) \in [t_o - \eta, T) \times \partial B(x_o)\} > 0$ and θ_n is now the minimum between T and the first time after t_n at which $\hat{X}^n$ exists $B(x_o)$. The contradiction is then deduced from the same arguments as above. □

Subsolution property. We now turn to the subsolution property. Again the proof is close to [10], except that we have to account for the specific form of the dynamic programming principle stated in Proposition 3.2.2(ii). Let φ be a C_b^∞ function, and $(t_o, x_o) \in [0, T] \times \mathbb{R}$ be a strict (local) maximum point of $w^* - \varphi$ such that $(w^* - \varphi)(t_o, x_o) = 0$. By [4, Lemma 4.2], we can find a sequence $(k_n, t_n, x_n)_{n \geq 1}$ such that $k_n \to \infty$, (t_n, x_n) is a local maximum point of $w_{k_n}^* - \varphi$ and $(t_n, x_n, w_{k_n}(t_n, x_n)) \to (t_o, x_o, w^*(t_o, x_o))$.

a. As above, we first assume that $t_o < T$. Set $\varphi_n(t,x) := \varphi(t,x) + |t-t_n|^2 + |x-x_n|^4$ and assume that $F\varphi(t_o,x_o) > 0$. Then, $F\varphi_n > 0$ on a open neighborhood B of (t_o,x_o) which contains (t_n,x_n), for all n large enough. Since we are going to localize the dynamics, we can modify φ_n, σ, μ and f in such a way that they are identically equal to 0 outside a compact $A \supset B$. It then follows from Remark 3.2.4 a. that, after possibly changing $n \geq 1$, we can find $(b^n, a^n) \in \mathcal{A}_{k_n}$ such that the following admits a strong solution:

$$\begin{aligned}
X^n &= x_n + \Delta \mathrm{x}(x_n, \hat{y}[\varphi_n](t_n,x_n)) + \int_{t_n}^{\cdot} \sigma(X_s^n) dW_s + \int_{t_n}^{\cdot} f(X_s^n) dY_s^{n,c} \\
&\quad + \int_{t_n}^{\cdot} (\mu(X_s) + a_s^n (\sigma f')(X_s^n)) ds \\
Y^n &= \hat{y}[\varphi_n](t_n,x_n) + \int_{t_n}^{\cdot} b_s^n ds + \int_{t_n}^{\cdot} a_s^n dW_s \\
&= \mathrm{x}^{-1}(\hat{X}^n, \hat{X}^n + (f\partial_x \varphi_n)(\cdot, \hat{X}^n)) \\
\hat{X}^n &:= \mathrm{x}(X^n, -Y^n) \\
V^n &= v_n + \mathfrak{I}(x_n, \hat{y}[\varphi_n](t_n,x_n)) + \int_{t_n}^{\cdot} Y_s^n dX_s^{n,c} + \frac{1}{2} \int_{t_n}^{\cdot} (a_s^n)^2 f(X_s^n) ds.
\end{aligned}$$

In the above, we have set $v_n := w_{k_n}(t_n,x_n) - n^{-1}$. Observe that the construction of Y^n ensures that it coincides with the corresponding process $\check{Y}^n$ of Lemma 3.2.1. Also note that $\hat{X}^n_{t_n} = \mathrm{x}(\mathrm{x}(x_n,y_n), -y_n) = x_n$, and let θ_n be the first time after t_n at which $(\cdot, \hat{X}^n)$ exists B. By applying Itô's Lemma, using Lemma 3.2.1 and the fact that $F\varphi_n \geq 0$ on B, we obtain

$$V^n_{\theta_n} \geq (\varphi_n + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + v_n - \varphi_n(t_n,x_n).$$

Let $2\varepsilon := \min\{|t-t_o|^2 + |x-x_o|^4,\ (t,x) \in \partial B\}$. For n large enough, the above implies

$$V^n_{\theta_n} \geq (w_{k_{n-1}} + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + \varepsilon + \iota_n,$$

where $\iota_n := (\varphi_n - w_{k_{n-1}})(t_{n-1}, x_{n-1}) + v_n - \varphi_n(t_n,x_n)$ converges to 0. Hence, we can find n such that

$$V^n_{\theta_n} > (w_{k_{n-1}} + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}).$$

Now observe that we can change the subsequence $(k_n)_{n\geq 1}$ in such a way that $k_n \geq 2k_{n-1} + 2$. Then, $v_n = w_{k_n}(t_n,x_n) - n^{-1} < w_{2k_{n-1}+2}(t_n,x_n)$, which leads to a contradiction to Proposition 3.2.2(ii).

b. It remains to consider the case $t_o = T$. As in Step 1., we only explain how to modify the argument used above. Let (v_n, k_n, t_n, x_n) be as in a. We now set $\varphi_n(t,x) := \varphi(t,x) + \sqrt{T-t} + |x-x_n|^4$. Since $\partial_t \varphi_n(t,x) \to -\infty$ as $t \to T$, we can find n large enough so that $F\varphi_n \geq 0$ on $[t_n, T) \times B(x_o)$ in which $B(x_o)$ is an open ball around x_o. Assume that $\varphi(T,x_o) > G(x_o) + \eta$ for some $\eta > 0$. Then, after possibly changing $B(x_o)$, we can assume that $\varphi_n(T,\cdot) \geq G + \eta$ on $B(x_o)$. We now use the same construction as in a. but with θ_n defined as the minimum between T and the first time where $\hat{X}^n$ exists $B(x_o)$. We obtain

$$V^n_{\theta_n} \geq (\varphi_n + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + v_n - \varphi_n(t_n,x_n).$$

Let $2\varepsilon := \min\{|x-x_o|^4,\ x \in \partial B(x_o)\}$. For n large enough, the above implies

$$V^n_{\theta_n} \geq w_{k_{n-1}}(\theta_n, \hat{X}^n_{\theta_n}) \mathbf{1}_{\theta_n < T} + G(\hat{X}^n_{\theta_n}) \mathbf{1}_{\theta_n = T} + \mathfrak{I}(\hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + \varepsilon \wedge \eta + \iota_n,$$

where ι_n converges to 0. By (3.2.3) and (H3),

$$V^n_{\theta_n} > w_{k_{n-1}}(\theta_n, \hat{X}^n_{\theta_n}) + \mathfrak{I}(\hat{X}^n_{\theta_n}, Y^n_{\theta_n}),$$

for n large enough. We conclude as in a. □

3.2.5.3 Comparison

In all this section, we work under the additional condition

$$\inf f > 0. \tag{3.2.13}$$

Direct computations (use (3.2.7) and Proposition 3.1.3) show that $\hat{F}\varphi$ is of the form

$$\hat{F}\varphi = -\partial_t\varphi - B(\cdot, f\partial_x\varphi)\partial_x\varphi - \frac{1}{2}A^2(\cdot, f\partial_x\varphi)\partial_{xx}\varphi - L(\cdot, f\partial_x\varphi) \tag{3.2.14}$$

where A, B and $L : (t, x, p) \in [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous functions.

Let Φ be a solution of the ordinary differential equation

$$\Phi'(t) = f(\Phi(t)),\ t \in \mathbb{R}. \tag{3.2.15}$$

Then, Φ is a bijection on $\mathbb{R}$ (as f is Lipschitz and $1/f$ is bounded) and the following is an immediate consequence of the definition of viscosity solutions.

Lemma 3.2.2. *Let v be a supersolution (resp. subsolution) of (3.2.8). Fix $\rho > 0$. Then,*

$$\tilde{v}(t, x) = e^{\rho t} v(t, \Phi(x))$$

is a supersolution (resp. subsolution) of

$$\begin{aligned} 0 \;=\; & \rho\varphi - \partial_t\varphi - \left[B(\Phi, e^{-\rho t}\partial_x\varphi)/f(\Phi) - \frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x\varphi)f'(\Phi)/f(\Phi)^2\right]\partial_x\varphi \\ & -\frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x\varphi)\partial_{xx}\varphi/f(\Phi)^2 - e^{\rho t}L(\Phi, e^{-\rho t}\partial_x\varphi) \end{aligned} \tag{3.2.16}$$

with the terminal condition

$$\varphi(T, \cdot) = e^{\rho T}G(\Phi). \tag{3.2.17}$$

To prove that comparison holds for (3.2.8), it suffices to prove that it holds for (3.2.16)-(3.2.17). The latter is a direct consequence of the following standard result. We provide here the complete proof without further precise reference.

Theorem 3.2.2. *Let $\mathcal{O}$ be an open subset of $\mathbb{R}$, u (resp. v) be a upper-semicontinuous subsolution (resp. lower-semicontinuous supersolution) on $[0, T) \times \mathcal{O}$ of:*

$$\rho\varphi - \partial_t\varphi - \bar{B}(\cdot, e^{-\rho t}\partial_x\varphi)\partial_x\varphi - \frac{1}{2}\bar{A}^2(\cdot, e^{-\rho t}\partial_x\varphi)\partial_{xx}\varphi - e^{\rho t}\bar{L}(\cdot, e^{-\rho t}\partial_x\varphi) = 0 \tag{3.2.18}$$

where $\rho > 0$ is constant, $\bar{A}, \bar{B}$ and $\bar{L} : (t, x, p) \in [0, T] \times \mathcal{O} \times \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous functions. Suppose that u and v are bounded and satisfy $u \le v$ on the parabolic boundary of $[0, T) \times \mathcal{O}$, then $u \le v$ on the closure of $[0, T] \times \mathcal{O}$.

Proof. Suppose to the contrary that

$$\sup_{[0,T]\times\mathcal{O}} (u - v) > 0.$$

Define, for $n > 0$,

$$\Theta_n := \sup_{(t,x,y)\in[0,T)\times\mathcal{O}^2} \left(u(t, x) - v(t, y) - \frac{n}{2}|x - y|^2 - \frac{1}{2n}|x|^2\right).$$

Then there exists $\iota > 0$, such that $\Theta_n \geq \iota$ for n large enough. Since u and v are bounded and $u \leq v$ on the parabolic boundary of the domain, there exists $(t_n, x_n, y_n) \in [0,T) \times \mathcal{O}^2$ which achieves the above supremum.

As usual, apply Ishii's Lemma combined with the sub- and super-solution properties of u and v, and the Lipschitz continuity of $\bar{A}, \bar{B}$ and $\bar{L}$ leads to, for some constant C which does not depend on n and $p_n := n(x_n - y_n)$,

$$\begin{aligned}
\rho(u(t_n,x_n) - v(t_n,y_n)) &\leq [\bar{B}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{B}(y_n, e^{-\rho t_n}p_n)]p_n \\
&\quad + \frac{1}{n}x_n \bar{B}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) \\
&\quad + \frac{3n}{2}[\bar{A}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{A}(y_n, e^{-\rho t_n}p_n)]^2 \\
&\quad + \frac{1}{2n}\bar{A}^2(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) \\
&\quad + e^{\rho t_n}\left(\bar{L}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{L}(y_n, e^{-\rho t_n}p_n)\right) \\
&\leq C\left(n(x_n - y_n)^2 + |x_n - y_n| + \frac{1}{n}x_n^2 + \frac{1}{n}\right)
\end{aligned}$$

In view of Lemma 3.2.3 below, and $\rho > 0, u(t_n,x_n) - v(t_n,y_n) \geq \Theta_n \geq \iota$, the above leads to a contradiction for n large enough. □

We conclude with the proof of the technical lemma used in our arguments above.

Lemma 3.2.3. *Let Ψ be a bounded upper-semicontinuous function on $[0,T] \times \mathbb{R}^2$, and Ψ_i, $i = 1, 2$, be two non-negative lower-semicontinuous functions on $\mathbb{R}$ such that $\{\Psi_1 = 0\} = \{0\}$. For $n > 0$, set*

$$\Theta_n := \sup_{(t,x,y) \in [0,T] \times \mathbb{R}^2} \left(\Psi(t,x,y) - n\Psi_1(x-y) - \frac{1}{n}\Psi_2(x)\right)$$

and assume that there exists $(\hat{t}_n, \hat{x}_n, \hat{y}_n) \in [0,T] \times \mathbb{R}^2$ such that:

$$\Theta_n = \Psi(\hat{t}_n, \hat{x}_n, \hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n).$$

Then, after possibly passing to a subsequence,

1. $\lim_{n\to\infty} n\Psi_1(\hat{x}_n - \hat{y}_n) = 0$ *and* $\lim_{n\to\infty} \frac{1}{n}\Psi_2(\hat{x}_n) = 0$.
2. $\lim_{n\to\infty} \Theta_n = \sup_{(t,x) \in [0,T] \times \mathcal{O}} \Psi(t,x,x)$.

Proof. For later use, set $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ and note that we can extend Ψ as a bounded upper-semicontinuous function on $[0,T] \times \bar{\mathbb{R}}^2$. Set $M := \sup_{(t,x) \in [0,T] \times \mathbb{R}} \Psi(t,x,x)$, and select a sequence $(t_n, x_n)_{n \geq 1}$ such that

$$\lim_{n\to\infty} \Psi(t_n, x_n, x_n) = M \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n}\Psi_2(x_n) = 0.$$

Let C be a upper-bound for Ψ. Then,

$$\begin{aligned}
C - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) &\geq \Psi(\hat{t}_n, \hat{x}_n, \hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \\
&\geq \Psi(t_n, x_n, x_n) - \frac{1}{n}\Psi_2(x_n) \\
&\geq M - \varepsilon_n
\end{aligned}$$

where $\epsilon_n \to 0$. Since Ψ_1 and Ψ_2 are non-negative, letting $n \to \infty$ in the above inequality leads to

$$\lim_{n\to\infty} \Psi_1(\hat{x}_n - \hat{y}_n) = 0$$

which implies $\lim_{n\to\infty}(\hat{x}_n - \hat{y}_n) = 0$ by the assumption $\{\Psi_1 = 0\} = \{0\}$.

After possibly passing to a subsequence, we can then assume that $\lim_{n\to\infty} \hat{x}_n = \lim_{n\to\infty} \hat{y}_n = \hat{x} \in \bar{\mathbb{R}}$ and that $\lim_{n\to\infty} \hat{t}_n = \hat{t} \in [0,T]$. The upper semi-continuity ofΨ along with the above leads to

$$\begin{aligned}
& M - \liminf_{n\to\infty} \left(n\Psi_1(\hat{x}_n - \hat{y}_n) + \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\
\geq\ & \Psi(\hat{t},\hat{x},\hat{x}) - \liminf_{n\to\infty} \left(n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\
\geq\ & \limsup_{n\to\infty} \left(\Psi(\hat{t}_n,\hat{x}_n,\hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\
\geq\ & M,
\end{aligned}$$

and our claim follows. □

Remark 3.2.5. *It follows from the above that, whenever they are bounded, e.g. if G is bounded, then $w_* \geq w^*$. Since by construction $w_* \leq w \leq w^*$, the three functions are equal to the unique bounded viscosity solution of* (3.2.8).

3.2.5.4 Smoothness

We conclude here the proof of Theorem 3.2.1 by showing that existence of a smooth solution holds when

$$\inf f > 0,\ G \text{ is bounded and } C^2 \text{ with } G, G', G'' \text{ Hölder continuous.} \tag{3.2.19}$$

Note that the assumptions $\inf f > 0$ and (H1) imply that Φ^{-1} is C^2, recall (3.2.15). Hence, by the same arguments as in Section 3.2.5.3, existence of a $C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$ solution to (3.2.16)-(3.2.17) implies the existence of a $C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$ solution to (3.2.8). As for (3.2.16)-(3.2.17), this is a consequence of [35, Thm 14.24], under (H1) and (3.2.19).

It remains to show that the solution can be taken bounded, then the comparison result of Section 3.2.5.3 will imply that w is this solution. Again, it suffices to work with (3.2.16)-(3.2.17). Let φ be a $C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$ solution of (3.2.16)-(3.2.17). Let $S^{t,x}$ be defined by

$$S_s^{t,x} = x + \int_t^s \mu_S(s, S_s^{t,x})ds + \int_t^s \sigma_S(s, S_s^{t,x})dW_s,\ s \geq t,$$

where

$$\begin{aligned}
\mu_S &:= B(\Phi, e^{-\rho t}\partial_x\varphi)/f(\Phi) - \frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x\varphi)f'(\Phi)/f(\Phi)^2 \\
\sigma_S &:= A(\Phi, e^{-\rho t}\partial_x\varphi)/f(\Phi).
\end{aligned}$$

Note that although the coefficients of the stochastic differential equation may only be locally Lipschitz, they are bounded (recall (H1) and (3.2.19)), which suffices to define a solution by the standard localization procedure. Since σ_S is bounded, Itô's Lemma implies that

$$\varphi(t,x)e^{-\rho t} = \mathbb{E}\left[G(\Phi(S_T^{t,x})) + \int_t^T L(\Phi(X_s^{t,x}), e^{-\rho s}\partial_x\varphi(s, X_s^{t,x}))ds \right].$$

Now G and L are bounded, by (H1) and (3.2.19), φ is bounded as well. □

Remark 3.2.6. *We refer to [34] for conditions under which additional smoothness of the solution can be proven.*

3.3 Appendix

We report here the measurability property used in the course of Proposition 3.2.2.

In the following, $\mathcal{A}_k$ is viewed as a closed subset of the Polish space $\mathbf{L}_2^\lambda$ endowed with the usual strong norm topology $\|\cdot\|_{\mathbf{L}_2^\lambda}$.

We regard an element $\nu \in \mathcal{U}_k$ as a measurable map $\omega \in \Omega \mapsto \nu(\omega) \in \mathcal{M}_k$ where $\mathcal{M}_k$ denotes the set of non-negative Borel measures on $\mathbb{R} \times [0,T]$ with total mass less than k, endowed with the topology of weak convergence. This topology is generated by the norm

$$\|m\|_{\mathcal{M}} := \sup\{\int_{\mathbb{R}\times[0,T]} \ell(\delta, s) m(d\delta, ds) : \ell \in \mathrm{Lip}_1\},$$

in which Lip_1 denotes the class of 1-Lipschitz continuous functions bounded by 1, see e.g. [8, Proposition 7.2.2 and Theorem 8.3.2]. Then, $\mathcal{U}_k$ is a closed subset of the space $\mathbf{M}_{k,2}$ of $\mathcal{M}_k$-valued random variables. $\mathbf{M}_{k,2}$ is made complete and separable by the norm

$$\|\nu\|_{\mathbf{M}_2} := \mathbb{E}\left[\|\nu\|_{\mathcal{M}}^2\right]^{\frac{1}{2}}.$$

See e.g. [22, Chap. 5]. We endow the set of controls Γ_k with the natural product topology

$$\|\gamma\|_{\mathbf{L}_2^\lambda \times \mathbf{M}_2} := \|\vartheta\|_{\mathbf{L}_2^\lambda} + \|\nu\|_{\mathbf{M}_2}, \quad \text{for } \gamma = (\vartheta, \nu).$$

As a closed subset of the Polish space $\mathbf{L}_2^\lambda \times \mathbf{M}_{k,2}$, Γ_k is a Borel space, for each $k \geq 1$. See e.g. [7, Proposition 7.12].

The stability result stated below is proved by using standard estimates.

Proposition 3.3.1. *For each $k \geq 1$, there exists a real constant $c_k > 0$ such that*

$$\|Z_T^{t_1,z_1,\gamma_1} - Z_T^{t_2,z_2,\gamma_2}\|_{\mathbf{L}_2} \leq c_k \left(|t_1 - t_2|^{\frac{1}{2}} + |z_1 - z_2| + \|\gamma_1 - \gamma_2\|_{\mathbf{L}_2^\lambda \times \mathbf{M}_2}\right),$$

for all $(t_i, z_i, \gamma_i) \in \mathrm{D} \times \Gamma_k$, $i = 1, 2$.

A direct consequence is the continuity of $(t, z, \gamma) \in \mathrm{D} \times \Gamma_k \mapsto Z_T^{t,z,\gamma}$, therefore it is measurable.

Corollary 3.3.1. *For each $k \geq 1$, the map $(t, z, \gamma) \in \mathrm{D} \times \Gamma_k \mapsto Z_T^{t,z,\gamma} \in \mathbf{L}_2$ is Borel-measurable.*

Chapter 4

Hedging of covered options with linear market impact and gamma constraint

Contents

4.1 Model and hedging problem **48**
4.1.1 Continuous time trading dynamics 48
4.1.2 Hedging equation and gamma constraint 50
4.2 Viscosity solution characterization **52**
4.2.1 A sequence of smooth supersolutions 53
4.2.2 Supersolution property for the weak formulation 63
4.2.3 Conclusion of the proof and construction of almost optimal strategies 67
4.3 Adding a resilience effect **68**
4.4 Numerical approximation and examples **69**
4.4.1 Finite difference scheme 69
4.4.2 Numerical examples: the fixed impact case 70
4.5 Appendix **72**

Introduction

Inspired by [1, 39], we considered a financial market with permanent price impact in the previous chapter, in which the impact function behaves as a linear function around the origin in the number of traded stocks. This class of models is dedicated to the pricing and hedging of derivatives under situations of non-negligible delta-hedging. In fact, the number of stocks required for hedging purpose becomes comparable to the average daily volume traded on the underlying asset. As a consequence, the delta-hedging strategy has an impact on the price dynamics, and also incurs liquidity costs. These models, incorporating both effects while maintaining the completeness of the market, lead to exact replication strategies. As in perfect market models, this approach provides an approximation of the real market conditions and hence can be used by practitioners to design a suitable hedge in a systematic way. Thus eliminating the need to rely on any ad hoc risk criterion.

It is shown in Chapter 3 that the price function of the optimal super-replicating strategy no longer solves a linear parabolic equation, as in the classical case, rather a quasi-linear one. The hedging strategy in this case, essentially follows a modified delta-hedging

rule where the delta is computed at the "unperturbed" value of the underlying, i.e., the one the underlying would have been if the trader's position were liquidated immediately. The approach employed and the results obtained differ substantially from [1,39], in which the control problem is applied to the hedging of *covered options*. The hedging of covered options refers to situations where the buyer of the option delivers at inception the required initial delta position, and accepts a mix of stocks (at their current market price) and cash as payment of the final claim. The buyer's indifference between stock and cash eliminates the cost incurred by the initial and final hedge. Quite surprisingly, this is not a genuine approximation of the problem studied in the previous chapter. The question of the initial and final hedge is fundamental, to the point that the structure of the pricing question is completely different: in Chapter 3 the equation is quasi-linear, whereas it is fully non-linear in [1,39]. In addition, as opposed to the previous chapter, authors in [1,39] use a verification argument to build an exact replication strategy. Due to the special form of the non-linearity, the equation is ill-posed when the solution does not satisfy a gamma-type constraint.

The aim of this chapter is to provide a direct characterization via stochastic target techniques, and to incorporate right from the beginning a gamma constraint on the hedging strategy. The super-solution property can be proved by essentially following the arguments of [19]. The sub-solution characterization is much more difficult to obtain. This is a second main difference from Chapter 3, in which classical geometric dynamic programming and viscosity solutions techniques could be used, once an appropriate change of variable was performed. In the current setting, however unlike in [19], we could not prove the required geometric dynamic programming principle. The underlying reason being the strong interaction between the hedging strategy and the underlying price process due to the market impact. Instead, we use the smoothing technique developed in [13]. We construct a sequence of smooth super-solutions which, by a verification argument, provide upper-bounds on the super-hedging price. As they converge to a solution of the targeted pricing equation, a comparison principle argument implies that their limit is the super-hedging price. A by-product of this construction is the explicit ε-optimal hedging strategies. We also provide the comparison principle and a numerical resolution scheme. To begin with, our analysis takes a simplified approach by restricting the models to only have permanent price impact. Later in Section 4.3, we show why adding a resilience effect does not affect our analysis. Note that this is because the resilience effect considered here has no quadratic variation. This is in contrast to [1], in which the resilience can break the parabolicity of the equation, and renders the exact replication non optimal.

4.1 Model and hedging problem

This section is dedicated to the description of the gamma constraint. We also explain in detail how the pricing equation can be obtained and state our main result.

4.1.1 Continuous time trading dynamics

We adopt the same model setting as in Chapter 3, however with slight modification on the conditions satisfied by the coefficient σ and on the definition of admissible strategy.

Instead of basing our following work on the hypothesis (H1) set previously, we assume throughout this chapter

$$\begin{array}{c} f \in C_b^2 \text{ and } \inf f > 0, \\ (\mu, \sigma) \text{ is Lipschitz and bounded, } \inf \sigma > 0. \end{array} \tag{4.1.1}$$

The above regularity assumptions are used in Chapter 3 to derive the continuous time trading dynamics. The lower bound on σ is used later on, in particular to express the

hedging policy in terms of a gamma, which is crucial for our analysis, see (4.1.4) and the equation just before. Relaxing these assumptions in the form of local conditions or by only assuming that f is C^1 with Lipschitz derivative should be feasible. This however would significantly increase the complexity of our proofs and we leave this to future researches.

Recall the number of shares the trader would like to hold is described by a continuous Itô process Y of the form

$$Y = Y_0 + \int_0^{\cdot} b_s ds + \int_0^{\cdot} a_s dW_s.$$

In addition to the conditions imposed on (a, b) in Chapter 3, that is (a, b) being progressively measurable and essentially bounded, we require here (a, b) to be an element of $\mathcal{A}_k^\circ$ consisting of continuous, $\mathbb{F}$-adapted processes which satisfy

$$a = a_0 + \int_0^{\cdot} \beta_s ds + \int_0^{\cdot} \alpha_s dW_s$$

where (α, β) is continuous, $\mathbb{F}$-adapted, and $\zeta := (a, b, \alpha, \beta)$ is essentially bounded by k and such that

$$\mathbb{E}\left[\sup\left\{|\zeta_{s'} - \zeta_s|,\ t \le s \le s' \le s + \delta \le T\right\} |\mathcal{F}_t^\circ\right] \le k\delta$$

for all $0 \le \delta \le 1$ and $t \in [0, T - \delta]$.

We then define

$$\mathcal{A}^\circ := \cup_k \mathcal{A}_k^\circ.$$

The above additional restriction on (a, b) will be necessary for our results in Section 4.2.2.

For readers' convenience, we recall here the continuous time trading dynamics derived in the previous chapter, see Proposition 3.1.1.

Proposition 1. *Let $Z := (X, Y, V)$ where Y is defined as in (3.1.4) for some $(a, b) \in \mathcal{A}^\circ$, and (X, V) solves*

$$\begin{aligned} X &= X_0 + \int_0^{\cdot} \sigma(X_s) dW_s + \int_0^{\cdot} f(X_s) dY_s + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s)) ds \\ &= X_0 + \int_0^{\cdot} \sigma_X^{a_s}(X_s) dW_s + \int_0^{\cdot} \mu_X^{a_s, b_s}(X_s) ds \end{aligned} \tag{4.1.2}$$

with

$$\sigma_X^{a_s} := (\sigma + a_s f)\ ,\ \ \mu_X^{a_s, b_s} := (\mu + b_s f + a_s \sigma f'),$$

and

$$V = V_0 + \int_0^{\cdot} Y_s dX_s + \frac{1}{2} \int_0^{\cdot} a_s^2 f(X_s) ds. \tag{4.1.3}$$

Let $Z^n := (X^n, Y^n, V^n)$ be defined as in (3.1.6)-(3.1.5)-(3.1.8). Then, there exists a constant $C > 0$ such that

$$\sup_{[0,T]} \mathbb{E}\left[|Z^n - Z|^2\right] \le C n^{-1}$$

for all $n \ge 1$.

Remark 1. *Note that in this work we restrict ourselves to a permanent price impact, no resilience effect is modeled. We shall explain in Section 4.3 below why taking resilience into account does not affect our analysis. See in particular Proposition 10.*

4.1.2 Hedging equation and gamma constraint

Given $\phi = (y, a, b) \in \mathbb{R} \times \mathcal{A}^\circ$ and $(t, x, v) \in [0, T] \times \mathbb{R} \times \mathbb{R}$, we write $(X^{t,x,\phi}, Y^{t,\phi}, V^{t,x,v,\phi})$ for the solution of (4.1.2)-(3.1.4)-(4.1.3) associated to the control (a, b) with time-t initial condition (x, y, v).

In this chapter, we consider covered options, in the sense that the trader is given at the initial time t the number of shares $Y_t = y$ required to launch his hedging strategy and can pay the option's payoff at T in cash and stocks (evaluated at their time-T value). Therefore, he does not exert any immediate impact at time t nor T due to the initial building or final liquidation of his position in stocks. Recalling that V stands for the sum of the position in cash and the number of held shares multiplied by their price, the super-hedging price at time t of the option with payoff $g(X_T^{t,x,\phi})$ is defined as

$$\mathrm{v}(t, x) := \inf\{v = c + yx \ : \ (c, y) \in \mathbb{R}^2 \text{ s.t. } \mathcal{G}(t, x, v, y) \neq \emptyset\},$$

in which $\mathcal{G}(t, x, v, y)$ is the set of elements $(a, b) \in \mathcal{A}^\circ$ such that $\phi := (y, a, b)$ satisfies

$$V_T^{t,x,v,\phi} \geq g(X_T^{t,x,\phi}).$$

In order to understand what the associated partial differential equation is, let us first rewrite the dynamics of Y in terms of X:

$$dY_t^{t,\phi} = \gamma_Y^{a_t}(X_t^{t,x,\phi})dX_t^{t,x,\phi} + \mu_Y^{a_t,b_t}(X_t^{t,x,\phi})dt$$

with

$$\gamma_Y^a := \frac{a}{\sigma + fa} \quad \text{and} \quad \mu_Y^{a,b} := b - \gamma_Y^a \mu_X^{a,b}. \tag{4.1.4}$$

Assuming that the hedging strategy is to track the super-hedging price, as in classical complete market models, then one should have $V^{t,x,v,\phi} = \mathrm{v}(\cdot, X^{t,x,\phi})$. If v is smooth, recalling (4.1.2)-(4.1.3) and applying Itô's lemma twice implies

$$Y^{t,\phi} = \partial_x \mathrm{v}(\cdot, X^{t,x,\phi}) \ , \ \gamma_Y^a(X^{t,x,\phi}) = \partial_{xx}^2 \mathrm{v}(\cdot, X^{t,x,\phi}), \tag{4.1.5}$$

and

$$\frac{1}{2}a^2 f(X^{t,x,\phi}) = \partial_t \mathrm{v}(\cdot, X^{t,x,\phi}) + \frac{1}{2}(\sigma_X^a)^2(X^{t,x,\phi})\partial_{xx}^2 \mathrm{v}(\cdot, X^{t,x,\phi}). \tag{4.1.6}$$

Then, the right-hand side of (4.1.5) combined with the definition of γ_Y^a leads to

$$a = \frac{\sigma \partial_{xx}^2 \mathrm{v}(\cdot, X^{t,x,\phi})}{1 - f\partial_{xx}^2 \mathrm{v}(\cdot, X^{t,x,\phi})} \ , \ \sigma_X^a = \frac{\sigma}{1 - f\partial_{xx}^2 \mathrm{v}(\cdot, X^{t,x,\phi})},$$

and (4.1.6) simplifies to

$$\left[-\partial_t \mathrm{v} - \frac{1}{2}\frac{\sigma^2}{(1 - f\partial_{xx}^2 \mathrm{v})}\partial_{xx}^2 \mathrm{v}\right](\cdot, X^{t,x,\phi}) = 0 \quad \text{on } [t, T). \tag{4.1.7}$$

This is precisely the pricing equation obtained in [1, 39].

Equation (4.1.7) needs to be considered with some precautions due to the singularity at $f\partial_{xx}^2 \mathrm{v} = 1$. One needs to enforce that $1 - f\partial_{xx}^2 \mathrm{v}$ does not change sign. We choose to restrict the solutions to satisfy $1 - f\partial_{xx}^2 \mathrm{v} > 0$. Having the opposite inequality would imply that a does not have the same sign as $\partial_{xx}^2 \mathrm{v}$, so that having sold a convex payoff, one would sell when the stock goes up and buy when it goes down, a very counter-intuitive fact.

In the following, we impose that the constraint

$$-k \leq \gamma_Y^a(X^{t,x,\phi}) \leq \bar{\gamma}(X^{t,x,\phi}) \ , \quad \text{on } [t, T] \ \mathbb{P} - \text{a.e.}, \tag{4.1.8}$$

should hold for some $k \geq 0$, in which $\bar{\gamma}$ is a bounded continuous map satisfying

$$\iota \leq \bar{\gamma} \leq 1/f - \iota, \quad \text{for some } \iota > 0. \tag{4.1.9}$$

We now denote by $\mathcal{A}_{k,\bar{\gamma}}(t,x)$ the collection of elements $(a,b) \in \mathcal{A}_k^\circ$ such that (4.1.8) holds. Define

$$\mathcal{A}_{\bar{\gamma}}(t,x) := \cup_{k\geq 0}\mathcal{A}_{k,\bar{\gamma}}(t,x),$$

and let $\mathrm{v}_{\bar{\gamma}}$ be defined as v but with

$$\mathcal{G}_{\bar{\gamma}}(t,x,v,y) := \mathcal{G}(t,x,v,y) \cap \mathcal{A}_{\bar{\gamma}}(t,x)$$

in place of $\mathcal{G}(t,x,v,y)$. More precisely,

$$\mathrm{v}_{\bar{\gamma}}(t,x) := \inf\{v = c + yx \ : \ (c,y) \in \mathbb{R}^2 \text{ s.t. } \mathcal{G}_{\bar{\gamma}}(t,x,v,y) \neq \emptyset\}. \tag{4.1.10}$$

Then, the equation (4.1.7) has to be modified to take the gamma constraint into account. This equation needs to impose that the second derivative is lower that the bound $\bar{\gamma}$. On the other hand, the above informal analysis shows that the pricing function $\mathrm{v}_{\bar{\gamma}}$ needs at least to be a super-solution of (4.1.7) to guarantee that a hedging strategy can be found. Then, the equation associated to the gamma constraint should read

$$F[\mathrm{v}_{\bar{\gamma}}] := \min\left\{-\partial_t \mathrm{v}_{\bar{\gamma}} - \frac{1}{2}\frac{\sigma^2}{1 - f\partial^2_{xx}\mathrm{v}_{\bar{\gamma}}}\partial^2_{xx}\mathrm{v}_{\bar{\gamma}} \ , \ \bar{\gamma} - \partial^2_{xx}\mathrm{v}_{\bar{\gamma}}\right\} = 0 \ \text{ on } [0,T)\times\mathbb{R}. \tag{4.1.11}$$

As for the T-boundary condition, we know that $\mathrm{v}_{\bar{\gamma}}(T,\cdot) = g$ by definition. However, as usual, the constraint on the gamma in (4.1.11) should propagate up to the boundary and g has to be replaced by its face-lifted version $\hat{g}$, defined as the smallest function above g that is a viscosity super-solution of the equation $\bar{\gamma} - \partial^2_{xx}\varphi \geq 0$. It is obtained by considering any twice continuously differentiable function $\bar{\Gamma}$ such that $\partial^2_{xx}\bar{\Gamma} = \bar{\gamma}$, and then setting

$$\hat{g} := (g - \bar{\Gamma})^{\mathrm{conc}} + \bar{\Gamma},$$

in which the superscript conc stands for concave envelope, cf. [48, Lemma 3.1].[1] We expect

$$\mathrm{v}_{\bar{\gamma}}(T-,\cdot) = \hat{g} \ \text{ on } \mathbb{R}.$$

From now on, assume that

$$\begin{array}{c} \hat{g} \text{ is uniformly continuous,} \\ g \text{ is lower-semicontinuous, } g^- \text{ is bounded and } g^+ \text{ has linear growth.} \end{array} \tag{4.1.12}$$

We are now in a position to state our main result. In the sequel,

$$\mathrm{v}_{\bar{\gamma}}(T,x) \ \text{ stands for } \lim_{\substack{(t',x') \to (T,x) \\ t'<T}} \mathrm{v}_{\bar{\gamma}}(t',x')$$

whenever it is well defined.

Theorem 1. *The value function $\mathrm{v}_{\bar{\gamma}}$ is continuous with linear growth. Moreover, $\mathrm{v}_{\bar{\gamma}}$ is the unique viscosity solution with linear growth of*

$$F[\varphi]\mathbf{1}_{[0,T)} + (\varphi - \hat{g})\mathbf{1}_{\{T\}} = 0 \quad \textit{on } [0,T]\times\mathbb{R}. \tag{4.1.13}$$

We conclude this section with additional remarks.

[1] Obviously, adding an affine map to $\bar{\Gamma}$ does not change the definition of $\hat{g}$.

Remark 2. *Note that $\hat{g}$ can be uniformly continuous without g being continuous. Take for instance $g(x) = \mathbf{1}_{\{x \geq K\}}$ with $K \in \mathbb{R}$, and consider the case where $\bar{\gamma} > 0$ is a constant. Then, $\hat{g}(x) = [\mathbf{1}_{\{x \geq x_o\}} \frac{\bar{\gamma}}{2}(x - x_o)^2] \wedge 1$ with $x_o := K - (2/\bar{\gamma})^{\frac{1}{2}}$.*

Remark 3. *The map $\hat{g}$ inherits the linear growth of g. Indeed, let $c_0, c_1 \geq 0$ be constants such that $|g(x)| \leq w(x) := c_0 + c_1|x|$. Since $\hat{g} \geq g$ by construction, we have $\hat{g}^- \leq w$. On the other hand, since $\bar{\gamma} \geq \iota > 0$ by (4.1.9), it follows from the arguments in [48, Lemma 3.1] that $\hat{g} \leq (w - \tilde{\Gamma})^{\text{conc}} + \tilde{\Gamma}$, in which $\tilde{\Gamma}(x) = \iota x^2/2$. Now, one can easily check by direct computations that*

$$(w - \tilde{\Gamma})^{\text{conc}} = (w - \tilde{\Gamma})(x_o)\mathbf{1}_{[-x_o, x_o]} + (w - \tilde{\Gamma})\mathbf{1}_{[-x_o, x_o]^c}$$

with $x_o := c_1/\iota$. Hence, $(w - \tilde{\Gamma})^{\text{conc}} + \tilde{\Gamma}$ has the same linear growth as w.

Remark 4. *As will appear in the rest of our analysis, one could very well introduce a time dependence in the impact function f and in $\bar{\gamma}$. Another interesting question studied in [39] and [38] concerns the smoothness of the solution and how the constraint on $\partial^2_{xx}\mathrm{v}$ gets naturally enforced by the fast diffusion arising when $1 - f\partial^2_{xx}\mathrm{v}$ is close to 0.*

Remark 5 (Existence of a smooth solution to the original partial differential equation)**.** *When the pricing equation (4.1.13) admits smooth solutions (cf. [39] and [38]) that allow to use the verification theorem, then one can construct exact replication strategies from the classical solution. By the comparison principle of Theorem 3 below, this shows that the value function is the classical solution of the pricing equation, and that the optimal strategy exists and is an exact replication strategy of the option with payoff function $\hat{g}$. We will explain in Remark 12 below how almost optimal super-hedging strategies can be constructed explicitly even when no smooth solution exists.*

Remark 6 (Monotonicity in the impact function)**.** *Note that the map $\lambda \in \mathbb{R} \mapsto \frac{\sigma^2(x)M}{1-\lambda M}$ is non-decreasing on $\{\lambda : \lambda M < 1\}$, for all $(t, x, M) \in [0, T] \times \mathbb{R} \times \mathbb{R}$. Let us now write $\mathrm{v}_{\bar{\gamma}}$ as $\mathrm{v}^f_{\bar{\gamma}}$ to emphasize its dependence on f, and consider another impact function $\tilde{f}$ satisfying the same requirements as f. We denote by $\mathrm{v}^{\tilde{f}}_{\bar{\gamma}}$ the corresponding super-hedging price. Then, the above considerations combined with Theorem 6 and the comparison principle of Theorem 3 below imply that $\mathrm{v}^{\tilde{f}}_{\bar{\gamma}} \geq \mathrm{v}^f_{\bar{\gamma}}$ whenever $\tilde{f} \geq f$ on $\mathbb{R}$. The same implies that $\mathrm{v}^f_{\bar{\gamma}} \geq \mathrm{v}$ in which v solves the heat-type equation*

$$-\partial_t \varphi - \frac{1}{2}\sigma^2 \partial^2_{xx}\varphi = 0 \quad \text{on } [0, T) \times \mathbb{R},$$

with terminal condition $\varphi(T, \cdot) = g$ (recall that $\hat{g} \geq g$). See Section 4.4.2 for a numerical illustration of this fact.

4.2 Viscosity solution characterization

In this section, we provide the proof of Theorem 6. Our strategy is the following.

1. First, we adapt the partial differential equation smoothing technique used in [13] to provide a smooth supersolutions $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}$ of (4.1.13) on $[\delta, T] \times \mathbb{R}$, with $\epsilon > 0$, from which super-hedging strategies can be constructed by a standard verification argument. In particular, $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}} \geq \mathrm{v}_{\bar{\gamma}}$ on $[\delta, T] \times \mathbb{R}$. Moreover, this sequence has a uniform linear growth and converges to a viscosity solution $\bar{\mathrm{v}}_{\bar{\gamma}}$ of (4.1.13) as $\delta, \epsilon \to 0$ and $K \to \infty$. See Section 4.2.1.

2. Second, we construct a lower bound $\underline{v}_{\bar\gamma}$ for $v_{\bar\gamma}$ that is a supersolution of (4.1.13). It is obtained by considering a weak formulation of the super-hedging problem and following the arguments of [19, Section 5] based on one side of the geometric dynamic programming principle, see Section 4.2.2. It is shown that this function has linear growth as well.

3. We can then conclude by using the above results and the comparison principle for (4.1.13) of Theorem 3 below: $\underline{v}_{\bar\gamma} \geq \bar{v}_{\bar\gamma}$ however $\underline{v}_{\bar\gamma} \leq v_{\bar\gamma} \leq \bar{v}_{\bar\gamma}$. As a result, $v_{\bar\gamma} = \bar{v}_{\bar\gamma} = \underline{v}_{\bar\gamma}$ and $v_{\bar\gamma}$ is a viscosity solution of (4.1.13), and has linear growth.

4. Our comparison principle, Theorem 3 below, allows us to conclude that $v_{\bar\gamma}$ is the unique solution of (4.1.13) with linear growth.

As already mentioned in the introduction, unlike [19], we could not prove the required geometric dynamic programming principle that could directly lead to a subsolution property (thus avoiding to use the smoothing technique mentioned in 1. above). This is due to the strong interaction between the hedging strategy and the underlying price process through the market impact. Such a feedback effect is not present in [19].

4.2.1 A sequence of smooth supersolutions

We first construct a sequence of smooth supersolutions $\bar{v}^{\epsilon,K,\delta}_{\bar\gamma}$ of (4.1.13) which appears to be an upper bound on the super-hedging price v_γ, by a simple verification argument. For this, we adapt the methodology introduced in [13]: we first construct a viscosity solution of a version of (4.1.13) with shaken coefficients (in the terminology of [32]) and then smooth it out with a kernel. The main difficulty here is that our terminal condition $\hat{g}$ is unbounded, unlike [13]. This requires additional non trivial technical developments.

4.2.1.1 Construction of a solution for the operator with shaken coefficients

We start with the construction of the operator with shaken coefficients. Given $\epsilon > 0$ and a (uniformly) strictly positive continuous map κ with linear growth defined later on, let us introduce a family of perturbations of the operator appearing in (4.1.13):

$$F^\epsilon_\kappa(t,x,q,M) := \min_{x' \in D^\epsilon_\kappa(x)} \min\left\{ -q - \frac{\sigma^2(x')M}{2(1-f(x')M)}, \bar\gamma(x') - M \right\},$$

where

$$D^\epsilon_\kappa(x) := \{x' \in \mathbb{R} : (x-x')/\kappa(x') \in [-\epsilon,\epsilon]\}. \tag{4.2.1}$$

For ease of notation, we set

$$F^\epsilon_\kappa[\varphi](t,x) := F^\epsilon_\kappa(t,x,\partial_t\varphi(t,x),\partial^2_{xx}\varphi(t,x)),$$

whenever φ is smooth.

Remark 7. *For later use, note that the map $M \in (-\infty,\bar\gamma(x)] \mapsto \frac{\sigma^2(x)M}{2(1-f(x)M)}$ is non-decreasing and convex for each $x \in \mathbb{R}$, recall (4.1.9). Hence, $(q,M) \in \mathbb{R}\times(-\infty,\bar\gamma(x)] \mapsto F^\epsilon_\kappa(\cdot,q,M)$ is concave and non-increasing in M, for all $\epsilon \geq 0$. This is fundamental for our smoothing approach to work.*

We also modify the original terminal condition $\hat{g}$ by using an approximating sequence whose elements are affine for large values of $|x|$.

Lemma 1. *For all $K > 0$ there exists a uniformly continuous map $\hat{g}_K$ and $x_K \geq K$ such that*

- *$\hat{g}_K$ is affine on $[x_K, \infty)$ and on $(-\infty, -x_K]$*
- *$\hat{g}_K = \hat{g}$ on $[-K, K]$*
- *$\hat{g}_K \geq \hat{g}$*
- *$\hat{g}_K - \bar{\Gamma}$ is concave for any C^2 function $\bar{\Gamma}$ satisfying $\partial^2_{xx}\bar{\Gamma} = \bar{\gamma}$.*

Moreover, $(\hat{g}_K)_{K>0}$ is uniformly bounded by a map with linear growth and converges to $\hat{g}$ uniformly on compact sets.

Proof. Fix a C^2 function $\bar{\Gamma}^\circ$ satisfying $\partial^2_{xx}\bar{\Gamma}^\circ = \bar{\gamma}$. By definition, $\hat{g} - \bar{\Gamma}^\circ$ is concave. Let us consider an element Δ^+ (resp. Δ^-) of its super-differential at K (resp. $-K$). Set

$$\begin{aligned}\hat{g}^\circ_K(x) :=& \hat{g}(x)\mathbf{1}_{[-K,K]}(x)\\ &+ \left[\hat{g}(K) + (\Delta^+ + \partial_x\bar{\Gamma}^\circ(K))(x-K)\right]\mathbf{1}_{(K,\infty)}(x)\\ &+ \left[\hat{g}(-K) + (\Delta^- + \partial_x\bar{\Gamma}^\circ(-K))(x+K)\right]\mathbf{1}_{(-\infty,-K)}(x).\end{aligned}$$

Consider now another C^2 function $\bar{\Gamma}$ satisfying $\partial^2_{xx}\bar{\Gamma} = \bar{\gamma}$. Since $\bar{\Gamma}^\circ$ and $\bar{\Gamma}$ differ only by an affine map, the concavity of $\hat{g}^\circ_K - \bar{\Gamma}$ is equivalent to that of $\hat{g}^\circ_K - \bar{\Gamma}^\circ$. The concavity of the latter follows from the definition of $\hat{g}^\circ_K$, as the superdiffential of $\hat{g}^\circ_K - \bar{\Gamma}^\circ$ is non-increasing by construction. In particular, $\hat{g}^\circ_K - \bar{\Gamma}^\circ \geq \hat{g} - \bar{\Gamma}^\circ$ and therefore $\hat{g}^\circ_K \geq \hat{g}$.

We finally define $\hat{g}_K$ by

$$\hat{g}_K = \min\{\hat{g}^\circ_K, (2c_0 + c_1|\cdot| - \bar{\Gamma}^\circ)^{\mathrm{conc}} + \bar{\Gamma}^\circ\}, \tag{4.2.2}$$

with $c_0 > 0$ and $c_1 \geq 0$ such that

$$-c_0 \leq \hat{g}(x) \leq c_0 + c_1|x|,\ x \in \mathbb{R},$$

recall Remark 3. The function $\hat{g}_K$ has the same linear growth as $2c_0 + c_1|\cdot|$, by the same reasoning as in Remark 3. Since $2c_0 > c_0$, $\hat{g}_K = \hat{g}^\circ_K = \hat{g}$ on $[-K, K]$. Furthermore, as the minimum of two concave functions is concave, so is $\hat{g}_K - \bar{\Gamma}$ for any C^2 function $\bar{\Gamma}$ satisfying $\partial^2_{xx}\bar{\Gamma} = \bar{\gamma}$. The other assertions are immediate. □

We now set

$$\hat{g}^\epsilon_K := \hat{g}_K + \epsilon \tag{4.2.3}$$

and consider the equation

$$F^\epsilon_\kappa[\varphi]\mathbf{1}_{[0,T)} + (\varphi - \hat{g}^\epsilon_K)\mathbf{1}_{\{T\}} = 0. \tag{4.2.4}$$

We then choose κ and $\epsilon_\circ \in (0, 1)$ such that

$$\begin{gathered}\kappa \in C^\infty \text{ with bounded derivatives of all orders,}\\ \inf\kappa > 0 \text{ and } \kappa = |\hat{g}_K| + 1 \text{ on } (-\infty, -x_K] \cup [x_K, \infty),\\ -1/\epsilon_\circ < \partial_x\kappa < 1/\epsilon_\circ,\end{gathered} \tag{4.2.5}$$

in which $x_K \geq K$ is defined in Lemma 1. We omit the dependence of κ on K for ease of notations.

Remark 8. *For later use, note that the condition $|\partial_x\kappa| < 1/\epsilon_\circ$ ensures that the map $x \mapsto x + \epsilon\kappa(x)$ and $x \mapsto x - \epsilon\kappa(x)$ are uniformly strictly increasing for all $0 \leq \epsilon \leq \epsilon_\circ$. Also observe that $x_n \to x$ and $x'_n \in D^\epsilon_\kappa(x_n)$, for all n, imply that x'_n converges to an element $x' \in D^\epsilon_\kappa(x)$, after possibly passing to a subsequence. In particular, F^ϵ_κ is continuous.*

When $\kappa \equiv 1$ and $\hat{g}_K^\epsilon \equiv \hat{g}+\epsilon$, (4.2.4) corresponds to the operator in (4.1.13) with shaken coefficients, in the traditional terminology of [32]. The function κ will be used below to handle the potential linear growth at infinity of $\hat{g}$. The introduction of the additional approximation $\hat{g}_K^\epsilon$ is motivated by the fact that the proof of Proposition 4 below requires an affine behavior at infinity. As already mentioned, these additional complications do not appear in [13] because their terminal condition is bounded.

We now prove that (4.2.4) admits a viscosity solution that remains above the terminal condition $\hat{g}$ on a small time interval $[T-c_\epsilon^K, T]$. As already mentioned, we will later smooth this solution out with a regular kernel to provide a smooth supersolution of (4.1.13).

Proposition 2. *For all $\epsilon \in [0, \epsilon_\circ]$ and $K > 0$, there exists a unique continuous viscosity solution $\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}$ of (4.2.4) that has linear growth. It satisfies*

$$\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K} \geq \hat{g}_K + \epsilon/2, \qquad \text{on } [T - c_\epsilon^K, T] \times \mathbb{R}, \tag{4.2.6}$$

for some $c_\epsilon^K \in (0, T)$.
Moreover, $\{[\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}]^+, \epsilon \in [0, \epsilon_\circ], K > 0\}$ is bounded by a map with linear growth, and $\{[\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}]^-, \epsilon \in [0, \epsilon_\circ], K > 0\}$ is bounded by $\sup g^-$.

Proof. The proof is mainly a modification of the usual Perron's method, see [21, Section 4].
a. We first prove that there exists two continuous functions $\bar{w}$ and $\underline{w}$ with linear growth that are respectively super- and subsolution of (4.2.4) for any $\epsilon \in [0, \epsilon_\circ]$.

Since $\hat{g}_K^\epsilon = \hat{g}_K + \epsilon \geq g$ by Lemma 1, it suffices to set

$$\underline{w} := \inf g > -\infty,$$

see (4.1.12). To construct a supersolution $\bar{w}$, let us fix $\eta \in (0, \iota \wedge \inf f^{-1})$ with ι as in (4.1.9), set $\tilde{\Gamma}(x) = \eta x^2/2$ and define $\tilde{g} = (\hat{g}_K^{\epsilon_\circ} - \tilde{\Gamma})^{\mathrm{conc}} + \tilde{\Gamma}$. Then, $\tilde{g} \geq \hat{g}_K^{\epsilon_\circ}$, while the same reasoning as in Remark 3 implies that $\tilde{g}$ shares the same linear growth as $\hat{g}_K^{\epsilon_\circ}$, see (4.2.3) and Lemma 1. We then define $\bar{w}$ by

$$\bar{w}(t,x) = \tilde{g}(x) + 1 + (T-t)A$$

in which

$$A := \sup \frac{\sigma^2 \bar{\gamma}}{2(1 - f\bar{\gamma})}.$$

The constant A is finite, and $\bar{w}$ has the same linear growth as $\tilde{g}$, see (4.1.1)-(4.1.9). Since a concave function is a viscosity supersolution of $-\partial^2_{xx}\varphi \geq 0$, we deduce that $\tilde{g}$ is a viscosity supersolution of $\eta - \partial^2_{xx}\varphi \geq 0$. Then, $\bar{w}$ is a viscosity supersolution of

$$\min\left\{-\partial_t\varphi - A\ ,\ \eta - \partial^2_{xx}\varphi\right\} \geq 0.$$

Since $\bar{\gamma} \geq \iota \geq \eta$, it remains to use Remark 7 to conclude that $\bar{w}$ is a supersolution of (4.2.4).
b. We now express (4.2.4) as a single equation over the whole domain $[0,T] \times \mathbb{R}$ using the following definitions

$$\begin{aligned}
F_{\kappa,+}^{\epsilon,K}(t,x,r,q,M) &:= F_\kappa^\epsilon(t,x,q,M)\mathbf{1}_{[0,T)} + \max\left\{F_\kappa^\epsilon(t,x,q,M), r - \hat{g}_K^\epsilon(x)\right\}\mathbf{1}_{\{T\}} \\
F_{\kappa,-}^{\epsilon,K}(t,x,r,q,M) &:= F_\kappa^\epsilon(t,x,q,M)\mathbf{1}_{[0,T)} + \min\left\{F_\kappa^\epsilon(t,x,q,M), r - \hat{g}_K^\epsilon(x)\right\}\mathbf{1}_{\{T\}}.
\end{aligned}$$

As usual $F_{\kappa,\pm}^{\epsilon,K}[\varphi](t,x) := F_{\kappa,\pm}^{\epsilon,K}(t,x,\varphi(t,x),\partial_t\varphi(t,x),\partial^2_{xx}\varphi(t,x))$. Recall that the formulations in terms of $F_{\kappa,\pm}^{\epsilon,K}$ lead to the same viscosity solutions as (4.2.4) (see Lemma 2 in the

Appendix). This is the formulation to which we apply Perron's method. In view of a., the functions $\underline{w}$ and $\bar{w}$ are sub- and supersolution of $F^{\epsilon,K}_{\kappa,-}=0$ and $F^{\epsilon,K}_{\kappa,+}=0$. Define:

$$\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}} := \sup\{v \in \mathrm{USC} : \underline{w} \le v \le \bar{w} \text{ and } v \text{ is a subsolution of } F^{\epsilon,K}_{\kappa,-}=0\},$$

in which USC denotes the class of upper-semicontinuous maps. Then, the upper- (resp. lower-) semicontinuous envelope $(\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}})^*$ (resp. $(\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}})_*$) of $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ is a viscosity subsolution of $F^{\epsilon,K}_{\kappa,-}[\varphi]=0$ (resp. supersolution of $F^{\epsilon,K}_{\kappa,+}[\varphi]=0$) with linear growth, recall the continuity property of Remark 8 and see e.g. [21, Section 4]. The comparison result of Theorem 3 stated below implies that

$$(\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}})^* = (\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}})_*, \quad \text{on } [0,T]\times\mathbb{R}.$$

Hence, $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ is a continuous viscosity solution of (4.2.4), recall Lemma 2. By construction, it has linear growth. Uniqueness in this class follows from Theorem 3 again.

c. It remains to prove (4.2.6). For this, we need a control on the behavior of $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ as $t \to T$. It is enough to obtain it for a lower bound $v_{\epsilon,K}$ that we first construct. Let φ be a test function such that

$$(\text{strict}) \min_{[0,T)\times\mathbb{R}} (\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}} - \varphi) = (\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}} - \varphi)(t_0,x_0)$$

for some $(t_0,x_0)\in[0,T)\times\mathbb{R}$. By the supersolution property,

$$\min_{x'\in D^\epsilon_\kappa(x_0)} \{\bar{\gamma}(x') - \partial^2_{xx}\varphi(t_0,x_0)\} \ge 0.$$

Recalling (4.1.1) and (4.1.9), this implies that, for $x'\in D^\epsilon_\kappa(x_0)$,

$$1 - f(x')\partial^2_{xx}\varphi(t_0,x_0) \ge \iota f(x') \ge \iota \inf f =: \tilde{\iota} > 0.$$

Using the supersolution property and the above inequalities yields

$$\begin{aligned}
0 &\le \min_{x'\in D^\epsilon_\kappa(x_0)} \left\{ -\partial_t\varphi(t_0,x_0) - \frac{\sigma^2(x')\partial^2_{xx}\varphi(t_0,x_0)}{2(1-f(x')\partial^2_{xx}\varphi(t_0,x_0))} \right\} \\
&\le \min_{x'\in D^\epsilon_\kappa(x_0)} \left\{ -\partial_t\varphi(t_0,x_0) - \frac{\sigma^2(x')\left[\partial^2_{xx}\varphi(t_0,x_0) - \bar{\gamma}(x_0)\right]}{2(1-f(x')\partial^2_{xx}\varphi(t_0,x_0))} \right\} \\
&\le -\partial_t\varphi(t_0,x_0) - \frac{\tilde{\sigma}^2\partial^2_{xx}\varphi(t_0,x_0)}{2\tilde{\iota}} + \frac{\tilde{\sigma}^2\bar{\gamma}(x_0)}{2\tilde{\iota}}
\end{aligned}$$

where $\tilde{\sigma} := \sup\sigma$.

Denote by $v_{\epsilon,K}$ the unique viscosity solution of

$$\left\{ -\partial_t\varphi - \frac{\tilde{\sigma}^2\partial^2_{xx}\varphi}{2\tilde{\iota}} + \frac{\tilde{\sigma}^2\bar{\gamma}}{2\tilde{\iota}} \right\} \mathbf{1}_{[0,T)} + (\varphi - \hat{g}^\epsilon_K)\mathbf{1}_{\{T\}} = 0. \tag{4.2.7}$$

The comparison principle for (4.2.7) and the Feynman-Kac formula imply that

$$\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}(t,x) \ge v_{\epsilon,K}(t,x) = \mathbb{E}\left[-\int_0^{T-t} \frac{\tilde{\sigma}^2\bar{\gamma}(S^x_r)}{2\tilde{\iota}}\, dr + \hat{g}^\epsilon_K(S^x_{T-t}) \right]$$

where

$$S^x = x + \frac{\tilde{\sigma}}{\sqrt{\tilde{\iota}}} W.$$

It remains to show that (4.2.6) holds for $v_{\epsilon,K}$ in place of $\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K}$. The argument is standard. Since $\hat g_K$ is uniformly continuous, see Lemma 1, we can find $B_\varepsilon^K > 0$ such that

$$\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right| \mathbf{1}_{\{|S^x_{T-t}-x|\le B_\varepsilon^K\}} \le \varepsilon$$

for all $\varepsilon > 0$. We now consider the case $|S^x_{T-t} - x| > B_\varepsilon^K$. Let $C > 0$ denote a generic constant that does not depend on (t, x) but can change from line to line. Because $\hat g_K$ is affine on $[x_K, \infty)$ and on $(-\infty, -x_K]$, see Lemma 1,

$$\mathbb{E}\left[\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right| \mathbf{1}_{\{S^x_{T-t}\ge x_K\}}\right] \le C(T-t)^{\frac12} \text{ if } x \ge x_K,$$

and

$$\mathbb{E}\left[\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right| \mathbf{1}_{\{S^x_{T-t}\le -x_K\}}\right] \le C(T-t)^{\frac12} \text{ if } x \le -x_K.$$

On the other hand, if $x < x_K$, then by linear growth of $\hat g_K^\epsilon$

$$\begin{aligned}
&\mathbb{E}\left[\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right| \mathbf{1}_{\{S^x_{T-t}\ge x_K\}}\mathbf{1}_{\{|S^x_{T-t}-x|\ge B_\varepsilon^K\}}\right]\\
&\le \mathbb{E}\left[\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right|^2\right]^{\frac12} \mathbb{P}\left[|S^x_{T-t} - x| \ge |x_K - x| \vee B_\varepsilon^K\right]^{\frac12}\\
&\le C\frac{(1+|x|)(T-t)^{\frac12}}{|x_K - x| \vee B_\varepsilon^K} \le \frac{C}{B_\varepsilon^K}(T-t)^{\frac12}.
\end{aligned}$$

The (four) remaining cases are treated similarly, and we obtain

$$\mathbb{E}\left[\left|\hat g_K^\epsilon(S^x_{T-t}) - \hat g_K^\epsilon(x)\right|\right] \le \frac{C}{B_\varepsilon^K}(T-t)^{\frac12} + \varepsilon.$$

The fact $\bar\gamma$ is bounded shows that for $t \in [T-1, T]$

$$|v_{\epsilon,K}(t,x) - \hat g_K^\epsilon(x)| \le \frac{C}{B_\varepsilon^K}(T-t)^{\frac12} + \varepsilon.$$

Hence the required result for $v_{\epsilon,K}$. Since $\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K} \ge v_{\epsilon,K}$, this concludes the proof of (4.2.6). □

For later use, note that, by stability, $\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K}$ converges to a solution of (4.1.13) when $\epsilon \to 0$ and $K \to \infty$.

Proposition 3. *As $\epsilon \to 0$ and $K \to \infty$, $\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K}$ converges to a function $\bar{\mathrm{v}}_{\bar\gamma}$ that is the unique viscosity solution of (4.1.13) with linear growth.*

Proof. The family of functions $\{\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K}, \epsilon \in (0, \epsilon_\circ], K > 0\}$ is uniformly bounded by a map with linear growth, see Proposition 2. In view of the comparison result of Theorem 3 below, it suffices to apply [3, Theorem 4.1]. □

Remark 9. *The bounds on $\bar{\mathrm{v}}_{\bar\gamma}$ can be made explicit, which can be useful to design a numerical scheme, see Section 4.4.1 below. First, as a by-product of the proof of Proposition 2, $\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K} \ge \inf g$. Passing to the limit as $\epsilon \to 0$ and $K \to \infty$ leads to*

$$\bar{\mathrm{v}}_{\bar\gamma} \ge \inf g =: \underline{w}.$$

We have also obtained that

$$\bar{\mathrm{v}}_{\bar\gamma}^{\epsilon,K} \le (\hat g_K^{\epsilon_\circ} - \tilde\Gamma)^{\mathrm{conc}} + \tilde\Gamma + 1 + A$$

in which $x \mapsto \tilde{\Gamma}(x) = \eta x^2/2$ *for some* $\eta \in (0, \iota \wedge \inf f^{-1})$ *with* ι *as in* (4.1.9)*, and* $A := T\sup(\sigma^2\bar{\gamma}/[2(1-f\bar{\gamma})])$*. On the other hand,* (4.2.2) *implies*

$$\hat{g}_K^{\epsilon_\circ} \leq 1 + (2c_0 + c_1|\cdot| - \bar{\Gamma}^\circ)^{\text{conc}} + \bar{\Gamma}^\circ$$

for $\bar{\Gamma}^\circ$ *such that* $\partial^2_{xx}\bar{\Gamma}^\circ = \bar{\gamma}$*. Then,*

$$\begin{aligned}
\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K} &\leq \left(1 + (2c_0 + c_1|\cdot| - \bar{\Gamma}^\circ)^{\text{conc}} + \bar{\Gamma}^\circ - \tilde{\Gamma}\right)^{\text{conc}} + \tilde{\Gamma} + 1 + A \\
&\leq \left(1 + (2c_0 + c_1|\cdot| - \tilde{\Gamma})^{\text{conc}} + \tilde{\Gamma} - \tilde{\Gamma}\right)^{\text{conc}} + \tilde{\Gamma} + 1 + A \\
&= \left(1 + 2c_0 + c_1|\cdot| - \tilde{\Gamma}\right)^{\text{conc}} + \tilde{\Gamma} + 1 + A =: \bar{w}
\end{aligned}$$

and

$$\bar{\mathrm{v}}_{\bar{\gamma}} \leq \bar{w}.$$

The function $\bar{w}$ *defined above can be computed explicitly by arguing as in Remark 3.*

Also note that (4.2.2) *and the arguments of Remark 3 imply that there exists a constant* $C > 0$ *such that*

$$\limsup_{|x|\to\infty} |\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}(x)|/(1 + |\hat{g}_K(x)|) \leq C, \text{ for all } \epsilon \in [0, \epsilon_\circ] \text{ and } K > 0. \tag{4.2.8}$$

4.2.1.2 Regularization and verification

Prior to applying our verification argument, it remains to smooth out the function $\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}$. The smoothing technique is similar to that in [13, Section 3], but here again the fact that $\hat{g}$ may not be bounded incurs additional difficulties. We need in this circumstance to use a kernel with a space dependent window.

We first fix a smooth kernel

$$\psi_\delta := \delta^{-2}\psi(\cdot/\delta)$$

in which $\delta > 0$ and $\psi \in C_b^\infty$ is a non-negative function with the closure of its support $[-1, 0] \times [-1, 1]$ that integrates to 1, and such that

$$\int y\psi(\cdot, y)dy = 0. \tag{4.2.9}$$

Let us set

$$\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K,\delta}(t,x) := \int_{\mathbb{R}\times\mathbb{R}} \bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}([t']^+, x')\frac{1}{\kappa(x)}\psi_\delta\left(t' - t, \frac{x' - x}{\kappa(x)}\right)dt'dx'. \tag{4.2.10}$$

We recall that κ enters into the definition of F_κ^ϵ and satisfies (4.2.5).

The following shows that $\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K,\delta}$ is a smooth supersolution of (4.1.13) with a space gradient admitting bounded derivatives. This is due to the space dependent rescaling of the window by κ and will be crucial for our verification arguments.

Proposition 4. *For all* $0 < \epsilon < \epsilon_\circ$ *and* $K > 0$ *large enough, there exists* $\delta_\circ > 0$ *such that* $\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K,\delta}$ *is a* C^∞ *supersolution of* (4.1.13) *for all* $0 < \delta < \delta_\circ$*. It has linear growth and* $\partial_x\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K,\delta}$ *has bounded derivatives of any order.*

Proof. a. It follows from (4.2.5) and (4.2.8) that

$$\limsup_{|x|\to\infty} |\bar{\mathrm{v}}_{\bar{\gamma}}^{\epsilon,K}(x)|/(1 + |\kappa(x)|) < \infty.$$

Direct computations and (4.2.5) then show that $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}$ has linear growth and that all derivatives of $\partial_x \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}$ are uniformly bounded.

b. We now prove the supersolution property inside the parabolic domain. Since the proof is very close to that of [13, Theorem 3.3], we only provide the arguments that require to be adapted, and refer to their proof for other elementary details. Fix $\ell > 0$ and set

$$v_\ell(t,x) := \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(t, (-\ell) \vee x \wedge \ell).$$

We omit the superscripts that are superfluous in this proof. Given $k \geq 1$, set

$$v_{\ell,k}(z) := \inf_{z' \in [0,T]\times\mathbb{R}} \left(v_\ell(z') + k|z - z'|^2 \right).$$

Since v_ℓ is bounded and continuous, the infimum in the above is achieved by a point $\hat z_{\ell,k}(z) = (\hat t_{\ell,k}(z), \hat x_{\ell,k}(z))$, and $v_{\ell,k}$ is bounded, uniformly in $k \geq 1$. This implies that we can find $C_\ell > 0$, independent of k, such that

$$|z - \hat z_{\ell,k}(z)|^2 \leq C_\ell/k =: (\rho_{\ell,k})^2. \tag{4.2.11}$$

Moreover, a simple change of variables argument shows that, if φ is a smooth function such that $v_{\ell,k} - \varphi$ achieves a minimum at $z \in [0,T) \times (-\ell, \ell)$, then

$$(\partial_t\varphi, \partial_x\varphi, \partial^2_{xx}\varphi)(z) \in \bar{\mathcal{P}}^- v_\ell(\hat z_{\ell,k}(z)),$$

where $\bar{\mathcal{P}}^- v_\ell(\hat z_{\ell,k}(z))$ denotes the *closed* parabolic subjet of v_ℓ at $\hat z_{\ell,k}(z)$, see e.g. [21] for the definition. Then, Proposition 2 implies that $v_{\ell,k}$ is a supersolution of

$$\min_{x' \in D^\epsilon_\kappa(\hat x_{\ell,k}(z))} \min \left\{ -\partial_t\varphi(z) - \frac{\sigma^2(x')\partial^2_{xx}\varphi(z)}{2(1 - f(x')\partial^2_{xx}\varphi(z))}, \bar\gamma(x') - \partial^2_{xx}\varphi(z) \right\} \geq 0,$$

$z \in [\rho_{\ell,k}, T - \rho_{\ell,k}) \times (-\ell + \rho_{\ell,k}, \ell - \rho_{\ell,k})$. We next deduce from (4.2.11) that $x' \in D^{\epsilon/2}_\kappa(x)$ implies

$$-\frac{\epsilon}{2}\kappa(x') - C_\ell/k^{\frac12} \leq \hat x_{\ell,k}(t,x) - x' \leq \frac{\epsilon}{2}\kappa(x') + C_\ell/k^{\frac12}.$$

Recall $\inf\kappa > 0$, the above inequality shows that $x' \in D^\epsilon_\kappa(\hat x_{\ell,k}(t,x))$ for k large enough with respect to ℓ. Hence, $v_{\ell,k}$ is a supersolution of

$$\min_{x' \in D^{\epsilon/2}_\kappa} \min \left\{ -\partial_t\varphi - \frac{\sigma^2(x')\partial^2_{xx}\varphi}{2(1 - f(x')\partial^2_{xx}\varphi)}, \bar\gamma(x') - \partial^2_{xx}\varphi \right\} \geq 0$$

on $[\rho_{\ell,k}, T - \rho_{\ell,k}) \times (-\ell + \rho_{\ell,k}, \ell - \rho_{\ell,k})$.

We now argue as in [30]. Since $v_{\ell,k}$ is semi-concave, there exist $\partial^{2,abs}_{xx} v_{\ell,k} \in L^1$ and a Lebesgue-singular negative Radon measure $\partial^{2,sing}_{xx} v_{\ell,k}$ such that

$$\partial^2_{xx} v_{\ell,k}(dz) = \partial^{2,abs}_{xx} v_{\ell,k}(z)dz + \partial^{2,sing}_{xx} v_{\ell,k}(dz) \text{ in the distribution sense}$$

and

$$(\partial_t v_{\ell,k}, \partial_x v_{\ell,k}, \partial^{2,abs}_{xx} v_{\ell,k}) \in \bar{\mathcal{P}}^- v_{\ell,k} \text{ a.e. on } [\rho_k, T - \rho_k] \times (-\ell + \rho_{\ell,k}, \ell - \rho_{\ell,k}),$$

see [31, Section 3]. The above implies that

$$\min_{x' \in D^{\epsilon/2}_\kappa} \min \left\{ -\partial_t v_{\ell,k} - \frac{\sigma^2(x')\partial^{2,abs}_{xx} v_{\ell,k}}{2(1 - f(x')\partial^{2,abs}_{xx} v_{\ell,k})}, \bar\gamma(x') - \partial^{2,abs}_{xx} v_{\ell,k} \right\} \geq 0$$

a.e. on $[\rho_{\ell,k}, T-\rho_{\ell,k}) \times (-\ell+\rho_{\ell,k}, \ell-\rho_{\ell,k})$, or equivalently, by (4.2.1),

$$\min\left\{-\partial_t v_{\ell,k} - \frac{\sigma^2(x)\partial^{2,abs}_{xx} v_{\ell,k}}{2(1-f(x)\partial^{2,abs}_{xx} v_{\ell,k})}, \bar\gamma(x) - \partial^{2,abs}_{xx} v_{\ell,k}\right\}(t',x') \geq 0$$

for all x and for a.e. $(t',x') \in [\rho_{\ell,k}, T-\rho_{\ell,k}) \times (-\ell+\rho_{\ell,k}, \ell-\rho_{\ell,k})$ such that $2|x'-x| \leq \epsilon\kappa(x)$. Take $0 < \delta < \varepsilon/2$. Integrating the previous inequality with respect to (t',x') with the kernel function $\psi_\delta(\cdot,\cdot/\kappa)/\kappa$, using the concavity and monotonicity property of Remark 7 and the fact that $\partial^{2,sing}_{xx} v_{\ell,k}$ is non-positive, we obtain

$$\min\left\{-\partial_t v^\delta_{\ell,k} - \frac{\sigma^2\partial^2_{xx} v^\delta_{\ell,k}}{2(1-f\partial^2_{xx} v^\delta_{\ell,k})}, \bar\gamma - \partial^2_{xx} v^\delta_{\ell,k}\right\} \geq 0 \tag{4.2.12}$$

on $[\rho_{\ell,k}+\delta, T-\rho_{\ell,k}) \times (-x^-_{\ell,k}, x^+_{\ell,k})$, in which

$$v^\delta_{\ell,k}(t,x) := \int_{\mathbb{R}\times\mathbb{R}} v_{\ell,k}([t']^+, x')\frac{1}{\kappa(x)}\psi_\delta\left(t'-\cdot, \frac{x'-\cdot}{\kappa(x)}\right)dt'dx'$$

and

$$x^+_{\ell,k} + \frac{\delta}{2}\kappa(x^+_{\ell,k}) = \ell - \rho_{\ell,k} \text{ and } -x^-_{\ell,k} - \frac{\delta}{2}\kappa(-x^-_{\ell,k}) = -\ell + \rho_{\ell,k}.$$

The above are well defined, see Remark 8. By Remark 8 and (4.2.11), $\pm x^\pm_{\ell,k} \to \pm\infty$ and $\rho_{\ell,k} \to 0$ as $k \to \infty$ and then $\ell \to \infty$. Moreover, $v^\delta_{\ell,k} \to \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}$ as $k \to \infty$ and then $\ell \to \infty$, and the derivatives also converge. Hence, (4.2.12) implies that $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}$ is a supersolution of (4.1.13) on $[\delta, T) \times \mathbb{R}$.

c. We conclude by discussing the boundary condition at T. From Proposition 2,

$$\bar{\mathrm{v}}^{\epsilon,K}_{\bar\gamma} \geq \hat g_K + \epsilon/2, \qquad \text{on } [T-c^K_\epsilon, T] \times \mathbb{R}.$$

The uniform continuity of $\hat g$ implies that of $\hat g_K$, see (4.1.12), therefore $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(T,\cdot) \geq \hat g_K$ on the compact set $[-2x_K, 2x_K]$ for $\delta > 0$ small enough with respect to ϵ. Now observe that $x \geq 2x_K$ and $|x'-x| \leq \delta\kappa(x)$ imply that $x' \geq 2x_K(1-\delta c^K_1) - \delta c^K_0$ in which c^K_1 and c^K_0 are constants. This actually follows from the affine behavior of κ on $[x_K, \infty)$, see (4.2.5) and Lemma 1. For δ small enough, we then obtain $x' \geq x_K$. Since $\hat g_K$ is affine on $[x_K, \infty)$, and ψ is symmetric in its second argument, see (4.2.9), it follows that

$$\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(T,x) \geq \int_{\mathbb{R}\times\mathbb{R}} \hat g_K(x')\frac{1}{\kappa(x)}\psi_\delta\left(t'-T, \frac{x'-x}{\kappa(x)}\right)dt'dx' = \hat g_K(x)$$

for all $x \geq 2x_K$. This also holds for $x \leq -2x_K$ by the same arguments. □

We can now use a verification argument and provide the main result of this section.

Theorem 2. *Let $\bar{\mathrm{v}}_{\bar\gamma}$ be defined as in Proposition 3. It has linear growth. Moreover, $\bar{\mathrm{v}}_{\bar\gamma} \geq \mathrm{v}_{\bar\gamma}$ on $[0,T] \times \mathbb{R}$.*

Proof. The linear growth property has already been stated in Proposition 3. We now show that $\bar{\mathrm{v}}_{\bar\gamma} \geq \mathrm{v}_{\bar\gamma}$ by applying a verification argument to $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}$. From now on $0 < \epsilon \leq \epsilon_\circ$ in which $\epsilon_\circ$ is as in (4.2.5). The parameters $K, \delta > 0$ are chosen as in Proposition 4.

Fix $(t,x) \in (0,T)\times\mathbb{R}$ and $\delta \in (0, t\wedge\epsilon)$. Let (X,Y,V) be defined as in (4.1.2)-(3.1.4)-(4.1.3) with $(x, \partial_x\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(t,x), \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(t,x) - \partial_x\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(t,x)x)$ as initial condition at t, and for the Markovian controls

$$\hat a = \left(\frac{\sigma\partial^2_{xx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}}{1-f\partial^2_{xx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}}\right)(\cdot, X)$$

$$\hat b = \left(\frac{\partial^2_{tx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma} + \partial^2_{xx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(\mu+\hat a\sigma f') + \frac{1}{2}\partial^3_{xxx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}(\sigma+\hat a f)^2}{1-f\partial^2_{xx}\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar\gamma}}\right)(\cdot, X).$$

By definition of F, (4.1.9) and (4.1.1), the above is well-defined as the denominators are always bigger than $\inf f\iota > 0$. All the involved functions being bounded and Lipschitz, see Proposition 4, it is easy to check that a solution to the corresponding stochastic differential equation exists, and that $(\hat{a}, \hat{b}) \in \mathcal{A}^\circ$. Direct computations then show that $Y = \partial_x \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(\cdot, X)$. Moreover, the fact that $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}$ is a supersolution of $F[\varphi] = 0$ on $[t,T] \times \mathbb{R}$ ensures that the gamma constraint (4.1.8) holds, for some $k \geq 1$, and that

$$-\partial_t \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(\cdot, X) - \frac{1}{2}\sigma(X)\hat{a} \geq 0 \ \text{ on } [t,T).$$

The last inequality combined with the definition of $\hat{a}$ implies

$$\begin{aligned}\frac{1}{2}f(X)\hat{a}^2 &\geq \partial_t \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(\cdot, X) + \frac{1}{2}(\sigma(X) + f(X)\hat{a})\hat{a} \\ &= \partial_t \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(\cdot, X) + \frac{1}{2}(\sigma^{\hat{a}}_X(X))^2 \partial^2_{xx} \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(\cdot, X) \ \text{ on } [t,T).\end{aligned}$$

Based on all the elements above

$$\begin{aligned} V_T &= \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(t,x) + \frac{1}{2}\int_t^T f(X_u)\hat{a}_u^2\, du + \int_t^T \partial_x \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(u, X_u)\, dX_u \\ &\geq \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(t,x) + \int_t^T d\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(u, X_u) \\ &= \bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(T, X_T) \geq g(X_T), \end{aligned}$$

in which the last inequality follows from Proposition 4 again.

It remains to pass to the limit $\delta, \epsilon \to 0$. By Proposition 2, $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ is continuous, so that $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}$ converges pointwise to $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ as $\delta \to 0$. By Proposition 3, $\bar{\mathrm{v}}^{\epsilon,K}_{\bar{\gamma}}$ converges pointwise to $\bar{\mathrm{v}}_{\bar{\gamma}}$ as $\epsilon \to 0$ and $K \to \infty$. In view of the above this implies the required result: $\bar{\mathrm{v}}_{\bar{\gamma}} \geq \mathrm{v}_{\bar{\gamma}}$. □

Remark 10. *Note that, in the above proof, we have constructed a super-hedging strategy in $\mathcal{A}_{k,\bar{\gamma}}(t,x)$ and starting with $|Y_t| \leq k$, for some $k \geq 1$ which can be chosen in a uniform way with respect to (t,x), while $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}$ has linear growth.*

4.2.1.3 Comparison principle

We provide here the comparison principle used several times above. Before stating it, we make the following observation, based on direct computations. Recall (4.1.1) and (4.1.9).

Proposition 5. *Fix $\rho > 0$. Consider the map*

$$(t,x,M) \in [0,T] \times \mathbb{R} \times \mathbb{R} \mapsto \Psi(t,x,M) = \frac{\sigma^2(x)M}{2(e^{\rho t} - f(x)M)}.$$

Then, $M \mapsto \Psi(t,x,M)$ is continuous, uniformly in (t,x), on

$$O := \{(t,x,M) \in [0,T] \times \mathbb{R} \times \mathbb{R} : M \leq e^{\rho t}\bar{\gamma}(x)\}.$$

Moreover, there exists $L > 0$ such that $x \mapsto \Psi(t,x,M)$ is L-Lipschitz on O.

Theorem 3. *Fix $\epsilon \in [0, \epsilon_\circ]$. Let U (resp. V) be a upper semicontinuous viscosity subsolution (resp. lower semicontinuous supersolution) of $F^\epsilon_\kappa = 0$ on $[0,T) \times \mathbb{R}$. Assume that U and V have linear growth and that $U \leq V$ on $\{T\} \times \mathbb{R}$, then $U \leq V$ on $[0,T] \times \mathbb{R}$.*

Proof. Set $\hat{U}(t,x) := e^{\rho t}U(t,x)$, $\hat{V}(t,x) := e^{\rho t}V(t,x)$. Then, $\hat{U}$ and $\hat{V}$ are respectively sub- and supersolution of

$$\min_{x' \in D_\kappa^\epsilon} \min \left\{ \rho\varphi - \partial_t\varphi - \frac{\sigma^2(x')\partial_{xx}\varphi}{2(e^{\rho t} - f(x')\partial_{xx}\varphi)}, e^{\rho t}\bar{\gamma}(x') - \partial_{xx}\varphi \right\} = 0 \tag{4.2.13}$$

on $[0,T) \times \mathbb{R}$. For later use, note that the infimum over D_κ^ϵ is achieved in the above, by the continuity of the involved functions.

If $\sup_{[0,T]\times\mathbb{R}}(\hat{U} - \hat{V}) > 0$, then one can find $\lambda \in (0,1)$ such that $\sup_{[0,T]\times\mathbb{R}}(\hat{U} - \hat{V}_\lambda) > 0$ with $\hat{V}_\lambda := \lambda\hat{V} + (1-\lambda)w$, in which

$$w(t,x) := (T-t)A + (c_0^U + c_1^U|\cdot| - \frac{\iota}{4}|\cdot|^2)^{\mathrm{conc}}(x) + \frac{\iota}{4}|x|^2$$

with c_0^U, c_1^U two constants such that $e^{\rho T}|U| \le c_0^U + c_1^U|\cdot|$ and

$$A := \frac{1}{2}\sup \frac{\sigma^2}{1 - \frac{\iota}{2}f}\frac{\iota}{2},$$

where $\iota > 0$ is as in (4.1.9). Note that

$$\hat{V}_\lambda(T,\cdot) \ge \hat{U}(T,\cdot), \tag{4.2.14}$$

and that

$$\begin{array}{c} w \text{ is a viscosity supersolution of (4.2.13)} \\ \hat{V}_\lambda \text{ is a viscosity supersolution of } \lambda\bar{\gamma} + (1-\lambda)\frac{\iota}{2} - \partial^2_{xx}\varphi \ge 0. \end{array} \tag{4.2.15}$$

Moreover, by Remark 7, $\hat{V}_\lambda$ is a supersolution of (4.2.13). Define for $\varepsilon > 0$ and $n \ge 1$

$$\Theta_n^\varepsilon := \sup_{(t,x,y)\in[0,T]\times\mathbb{R}^2} \left[\hat{U}(t,x) - \hat{V}_\lambda(t,y) - \left(\frac{\varepsilon}{2}|x|^2 + \frac{n}{2}|x-y|^2 \right) \right] =: \eta > 0, \tag{4.2.16}$$

in which the last inequality holds for $n > 0$ large enough and $\varepsilon > 0$ small enough. Denote by $(t_n^\varepsilon, x_n^\varepsilon, y_n^\varepsilon)$ the point at which this supremum is achieved. From (4.2.14), we see that $t_n^\varepsilon < T$, and, by standard arguments, see e.g., [21, Proposition 3.7],

$$\lim_{n\to\infty} n|x_n^\varepsilon - y_n^\varepsilon|^2 = 0. \tag{4.2.17}$$

On the other hand, Ishii's lemma implies the existence of $(a_n^\varepsilon, M_n^\varepsilon, N_n^\varepsilon) \in \mathbb{R}^3$ such that

$$\begin{array}{c} (a_n^\varepsilon, \varepsilon x^\varepsilon + n(x_n^\varepsilon - y_n^\varepsilon), M_n^\varepsilon) \in \bar{\mathcal{P}}^{2,+}\hat{U}(t_n^\varepsilon, x_n^\varepsilon) \\ (a_n^\varepsilon, -n(x_n^\varepsilon - y_n^\varepsilon), N_n^\varepsilon) \in \bar{\mathcal{P}}^{2,-}\hat{V}_\lambda(t_n^\varepsilon, y_n^\varepsilon), \end{array}$$

where $\bar{\mathcal{P}}^{2,+}$ and $\bar{\mathcal{P}}^{2,-}$ denote as usual the *closed* parabolic super- and subjets, see [21], and

$$\begin{pmatrix} M_n^\varepsilon & 0 \\ 0 & -N_n^\varepsilon \end{pmatrix} \le R_n^\varepsilon + \frac{1}{n}(R_n^\varepsilon)^2 = 3n \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 3\varepsilon + \frac{\varepsilon^2}{n} & -\varepsilon \\ -\varepsilon & 0 \end{pmatrix}$$

with

$$R_n^\varepsilon := n \begin{pmatrix} 1 + \frac{\varepsilon}{n} & -1 \\ -1 & 1 \end{pmatrix}.$$

In particular, multiplying this inequality by the vector $(1,1)$ gives rise to

$$M_n^\varepsilon - N_n^\varepsilon \le \delta_n^\varepsilon \text{ with } \delta_n^\varepsilon := \varepsilon + \frac{\varepsilon^2}{n}. \tag{4.2.18}$$

Then, by (4.2.15) and (4.1.9), the inequality below holds

$$0 < (1-\lambda)\frac{\iota}{2} \le e^{\rho t_n^\varepsilon}\bar\gamma(\hat y_n^\varepsilon) - N_n^\varepsilon \le e^{\rho t_n^\varepsilon}\bar\gamma(\hat y_n^\varepsilon) - M_n^\varepsilon + \delta_n^\varepsilon, \tag{4.2.19}$$

in which $\hat y_n^\varepsilon \in D_\kappa^\epsilon(y_n^\varepsilon)$. In view of Remark 8, this shows that $e^{\rho t_n^\varepsilon}\bar\gamma(\hat x_n^\varepsilon) - M_n^\varepsilon > 0$ for some $\hat x_n^\varepsilon \in D_\kappa^\epsilon(x_n^\varepsilon)$, n large enough and ε small enough, recall (4.2.17). The super- and subsolution properties of $\hat V_\lambda$ and $\hat U$ then imply that we can find $u_n^\varepsilon \in [-\epsilon,\epsilon]$ together with $\hat y_n^\varepsilon$ and $\hat x_n^\varepsilon$ such that

$$\hat y_n^\varepsilon + u_n^\varepsilon\kappa(\hat y_n^\varepsilon) = y_n^\varepsilon\ ,\ \hat x_n^\varepsilon + u_n^\varepsilon\kappa(\hat x_n^\varepsilon) = x_n^\varepsilon \tag{4.2.20}$$

and

$$\rho(\hat U(t_n^\varepsilon,x_n^\varepsilon) - \hat V_\lambda(t_n^\varepsilon,y_n^\varepsilon)) \le \frac{\sigma^2(\hat x_n^\varepsilon)M_n^\varepsilon}{2(e^{\rho t_n^\varepsilon} - f(\hat x_n^\varepsilon)M_n^\varepsilon)} - \frac{\sigma^2(\hat y_n^\varepsilon)N_n^\varepsilon}{2(e^{\rho t_n^\varepsilon} - f(\hat y_n^\varepsilon)N_n^\varepsilon)}.$$

By Remark 7 and (4.2.18), this shows that

$$\begin{aligned}&\rho(\hat U(t_n^\varepsilon,x_n^\varepsilon) - \hat V_\lambda(t_n^\varepsilon,y_n^\varepsilon))\\ &\le \frac{\sigma^2(\hat x_n^\varepsilon)(N_n^\varepsilon+\delta_n^\varepsilon)}{2(e^{\rho t_n^\varepsilon} - f(\hat x_n^\varepsilon)(N_n^\varepsilon+\delta_n^\varepsilon))} - \frac{\sigma^2(\hat y_n^\varepsilon)N_n^\varepsilon}{2(e^{\rho t_n^\varepsilon} - f(\hat y_n^\varepsilon)N_n^\varepsilon)}.\end{aligned}$$

It remains to apply Proposition 5 together with (4.2.19) for n large enough and ε small enough to obtain for some $L>0$

$$\begin{aligned}&\rho(\hat U(t_n^\varepsilon,x_n^\varepsilon) - \hat V_\lambda(t_n^\varepsilon,y_n^\varepsilon))\\ &\le \frac{\sigma^2(\hat x_n^\varepsilon)N_n^\varepsilon}{2(e^{\rho t_n^\varepsilon} - f(\hat x_n^\varepsilon)N_n^\varepsilon)} - \frac{\sigma^2(\hat y_n^\varepsilon)N_n^\varepsilon}{2(e^{\rho t_n^\varepsilon} - f(\hat y_n^\varepsilon)N_n^\varepsilon)} + O_n^\varepsilon(1)\\ &\le L\,|\hat x_n^\varepsilon - \hat y_n^\varepsilon| + O_n^\varepsilon(1)\end{aligned}$$

where $O_n^\varepsilon(1)\to 0$ as $n\to\infty$ and then $\varepsilon\to 0$. By continuity and (4.2.17) combined with Remark 8 and (4.2.20), this contradicts (4.2.16) for n large enough. □

4.2.2 Supersolution property for the weak formulation

In this part, we provide a lower bound $\underline{\mathrm{v}}_{\bar\gamma}$ for $\mathrm{v}_{\bar\gamma}$ that is a supersolution of (4.1.13). It is constructed by considering a weak formulation of the stochastic target problem (4.1.10) in the spirit of [19, Section 5]. Since our methodology is slightly different, we provide the main arguments.

On $C(\mathbb{R}_+)^5$, let us now denote by $(\tilde\zeta := (\tilde a,\tilde b,\tilde\alpha,\tilde\beta),\tilde W)$ the coordinate process and let $\tilde{\mathbb{F}}^\circ = (\tilde{\mathcal{F}}_s^\circ)_{s\le T}$ be its raw filtration. We say that a probability measure $\tilde{\mathbb{P}}$ belongs to $\tilde{\mathcal{A}}_k$ if $\tilde W$ is a $\tilde{\mathbb{P}}$-Brownian motion and if for all $0\le\delta\le 1$ and $r\ge 0$ it holds $\tilde{\mathbb{P}}$-a.s. that

$$\tilde a = \tilde a_0 + \int_0^\cdot \tilde\beta_s ds + \int_0^\cdot \tilde\alpha_s d\tilde W_s \ \text{ for some } \tilde a_0\in\mathbb{R}, \tag{4.2.21}$$

$$\sup_{\mathbb{R}_+}|\tilde\zeta| \le k\ , \tag{4.2.22}$$

and

$$\mathbb{E}^{\tilde{\mathbb{P}}}\Big[\sup\Big\{|\tilde\zeta_{s'} - \tilde\zeta_s|,\ r\le s\le s'\le s+\delta\Big\}\,|\tilde{\mathcal{F}}_r^\circ\Big] \le k\delta. \tag{4.2.23}$$

For $\tilde\phi := (y, \tilde a, \tilde b)$ and $y \in \mathbb{R}$, we define $(\tilde X^{x,\tilde\phi}, \tilde Y^{\tilde\phi}, \tilde V^{x,v,\tilde\phi})$ as in (4.1.2)-(3.1.4)-(4.1.3) associated to the control $(\tilde a, \tilde b)$ with time-0 initial condition (x, y, v), and with $\tilde W$ in place of W. For $t \le T$ and $k \ge 1$, we say that $\tilde{\mathbb{P}} \in \tilde{\mathcal{G}}_{k,\bar\gamma}(t, x, v, y)$ if

$$\Big[\tilde V^{x,v,\tilde\phi}_{T-t} \ge g(\tilde X^{x,\tilde\phi}_{T-t}) \ \text{ and } \ -k \le \gamma^{\tilde a}_Y(\tilde X^{x,\tilde\phi}) \le \bar\gamma(\tilde X^{x,\tilde\phi}) \text{ on } \mathbb{R}_+\Big] \ \ \tilde{\mathbb{P}} - \text{a.s.} \tag{4.2.24}$$

We finally define

$$\underline{\mathrm{v}}^k_{\bar\gamma}(t,x) := \inf\{v = c + yx \ : \ (c,y) \in \mathbb{R}\times[-k,k] \text{ s.t. } \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y) \neq \emptyset\},$$

and

$$\underline{\mathrm{v}}_{\bar\gamma}(t,x) := \liminf_{\substack{(k,t',x') \to (\infty,t,x)\\ (t',x') \in [0,T)\times\mathbb{R}}} \underline{\mathrm{v}}^k_{\bar\gamma}(t',x'), \quad (t,x) \in [0,T]\times\mathbb{R}. \tag{4.2.25}$$

The following is an immediate consequence of our definitions.

Proposition 6. *$\mathrm{v}_{\bar\gamma} \ge \underline{\mathrm{v}}_{\bar\gamma}$ on $[0,T)\times\mathbb{R}$.*

In the rest of this section, we show that $\underline{\mathrm{v}}_{\bar\gamma}$ is a viscosity supersolution of (4.1.13). Let us start with a simple remark.

Remark 11. *The gamma constraint in (4.2.24) implies that one can find $\varepsilon > 0$ such that*

$$\frac{\varepsilon}{1+k\varepsilon^{-1}} \le \sigma^{\tilde a}_X(\tilde X^{x,\tilde\phi}) \le \varepsilon^{-1} + \varepsilon^{-2} \ \text{ and } |\tilde a| \le \varepsilon^{-1} \ \ \tilde{\mathbb{P}} - \text{a.s.},$$

for all $\tilde{\mathbb{P}} \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y)$ and $k \ge 1$. Indeed, if $\tilde a \ge -\sigma/f$ then $-k \le \gamma^{\tilde a}_Y \le \bar\gamma$ implies

$$(-\frac{k\sigma}{1+kf}) \vee (-\frac{\sigma}{f}) \le \tilde a \le \frac{\bar\gamma\sigma}{1-\bar\gamma f} \ \text{ and } \ \tilde a f + \sigma \ge \sigma/(1+kf).$$

Our claim follows from (4.1.1)-(4.1.9). On the other hand, if $\sigma + \tilde a f < 0$, then $\gamma^{\tilde a}_Y \le \bar\gamma$ implies $\tilde a \ge \bar\gamma\sigma/(1-f\bar\gamma) \ge 0$, see (4.1.9), while $\tilde a < -f/\sigma < 0$, a contradiction.

We next show that $\underline{\mathrm{v}}^k_{\bar\gamma}$ has linear growth, for k large enough.

Proposition 7. *There exists $k_o \ge 1$ such that $\{|\underline{\mathrm{v}}^k_{\bar\gamma}|, k \ge k_o\}$ is uniformly bounded from above by a continuous map with linear growth.*

Proof. a. First note that Remark 10 implies that $\{(\underline{\mathrm{v}}^k_{\bar\gamma})^+, k \ge k_o\}$ is uniformly bounded from above by a map with linear growth, for some k_o large enough.
b. Fix now $\tilde{\mathbb{P}} \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y)$. Using Remark 11 along with (4.1.1) and the condition that $(\tilde a, \tilde b, \tilde\alpha, \tilde\beta)$ is $\tilde{\mathbb{P}}$-essentially bounded, one can find $\check{\mathbb{P}} \sim \tilde{\mathbb{P}}$ under which $\int_0^\cdot \tilde Y^{\tilde\phi}_s d\tilde X^{x,\tilde\phi}_s$ is a martingale on $[0, T-t]$. Then, the condition $\tilde V^{x,v,\tilde\phi}_{T-t} \ge g(\tilde X^{x,\tilde\phi}_{T-t})$ $\tilde{\mathbb{P}}$-a.s. implies $v + \mathbb{E}^{\check{\mathbb{P}}}[\frac12\int_0^{T-t} \tilde a_s^2 f(\tilde X^{x,\tilde\phi}_s)ds] \ge \inf g > -\infty$, recall (4.1.12). By Remark 11 and (4.1.1), $v \ge \inf g - C > -\infty$, for some constant C independent of $\tilde{\mathbb{P}} \in \cup_k(\tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y))$. In consequence, $\{(\underline{\mathrm{v}}^k_{\bar\gamma})^-, k \ge k_o\}$ is bounded by a constant. □

We now prove that existence is attaint in the definition of $\underline{\mathrm{v}}^k_{\bar\gamma}$ and it is lower-semicontinuous.

Proposition 8. *For all $(t,x) \in [0,T]\times\mathbb{R}$ and $k \ge 1$ large enough, there exists $(c,y) \in \mathbb{R}\times[-k,k]$ such that $\underline{\mathrm{v}}^k_{\bar\gamma}(t,x) = c + yx$ and $\tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t, c+xy, y) \neq \emptyset$. Moreover, $\underline{\mathrm{v}}^k_{\bar\gamma}$ is lower-semicontinuous for each $k \ge 1$ large enough.*

Proof. By [44, Proposition XIII.1.5] and the condition (4.2.23) taken for $r = 0$, the set $\tilde{\mathcal{A}}_k$ is weakly relatively compact. Moreover, [33, Theorem 7.10 and Theorem 8.1] implies that any limit point $(\mathbb{P}_*, t_*, x_*, c_*, y_*)$ of a sequence $(\mathbb{P}_n, t_n, x_n, c_n, y_n)_{n\ge1}$ such that $\mathbb{P}_n \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t_n, x_n, c_n + x_n y_n, y_n)$ for each $n \ge 1$ satisfies $\mathbb{P}_* \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t_*, x_*, c_* + x_* y_*, y_*)$. Since $\underline{\mathrm{v}}^k_{\bar\gamma}$ is locally bounded, the Proposition 7 for the case $k \ge k_o$ implies the announced existence and lower-semicontinuity. □

We can finally prove the main result of this section.

Theorem 4. *The function $\underline{\mathrm{v}}_{\bar{\gamma}}$ is a viscosity supersolution of (4.1.13). It has linear growth.*

Proof. The linear growth property is an immediate consequence of the uniform linear growth of $\{|\underline{\mathrm{v}}^k_{\bar{\gamma}}|, k \geq k_o\}$ stated in Proposition 7. To prove the supersolution property, it suffices to show that it holds for each $\underline{\mathrm{v}}^k_{\bar{\gamma}}$, with $k \geq k_o$, and then to apply standard stability results, see e.g. [3].

a. We first prove the supersolution property on $[0,T)\times\mathbb{R}$. We adapt the arguments of [19] to our context. Let φ be a C^∞_b test function and $(t_0,x_0) \in [0,T)\times\mathbb{R}$ such that

$$(\text{strict}) \min_{[0,T)\times\mathbb{R}} (\underline{\mathrm{v}}^k_{\bar{\gamma}} - \varphi) = (\underline{\mathrm{v}}^k_{\bar{\gamma}} - \varphi)(t_0,x_0) = 0.$$

Recall that $\underline{\mathrm{v}}^k_{\bar{\gamma}}$ is lower-semicontinuous by Proposition 8.

Because the infimum is attaint in the definition of $\underline{\mathrm{v}}^k_{\bar{\gamma}}$, by the afore-mentioned proposition, there exists $|y_0| \leq k$ and $\tilde{\mathbb{P}} \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_k(t_0,x_0,v_0,y_0)$, such that $v_0 := c_0 + y_0x_0 = \underline{\mathrm{v}}^k_{\bar{\gamma}}(t_0,x_0)$ for some $c_0 \in \mathbb{R}$. Set $(\tilde{X},\tilde{Y},\tilde{V}) := (\tilde{X}^{x_0,\tilde{\phi}}, \tilde{Y}^{\tilde{\phi}}, \tilde{V}^{x_0,v_0,\tilde{\phi}})$ where $\tilde{\phi} = (y_0,\tilde{a},\tilde{b})$. Let θ_o be a stopping time for the augmentation of the raw filtration $\tilde{\mathbb{F}}^\circ$, and define

$$\theta := \theta_o \wedge \theta_1 \text{ with } \theta_1 := \inf\{s : |\tilde{X}_s - x_0| \geq 1\}.$$

Then, it follows from Proposition 9 below that

$$\tilde{V}_{\theta_o} \geq \underline{\mathrm{v}}^k_{\bar{\gamma}}(t_0+\theta_o, \tilde{X}_{\theta_o}) \geq \varphi(t_0+\theta_o, \tilde{X}_{\theta_o}),$$

in which here and hereafter inequalities are taken in the $\tilde{\mathbb{P}}$-a.s. sense. After applying Itô's formula twice, the above inequality reads:

$$\int_0^\theta \ell_s\, ds + \int_0^\theta \left(y_0 - \partial_x\varphi(t_0,x_0) + \int_0^s m_r dr + \int_0^s n_r d\tilde{X}_r \right) d\tilde{X}_s \geq 0. \tag{4.2.26}$$

where

$$\ell := \tfrac{1}{2}\tilde{a}^2 f(\tilde{X}) - \mathcal{L}^{\tilde{a}}\varphi(t_0+\cdot,\tilde{X}_\cdot)\ ,\ m := \mu_Y^{\tilde{a},\tilde{b}}(\tilde{X}) - \mathcal{L}^{\tilde{a}}\partial_x\varphi(t_0+\cdot,\tilde{X}_\cdot)$$
$$n := \gamma_Y^{\tilde{a}}(\tilde{X}) - \partial^2_{xx}\varphi(t_0+\cdot,\tilde{X}_\cdot),$$

with

$$\mathcal{L}^{\tilde{a}} := \partial_t + \frac{1}{2}(\sigma_X^{\tilde{a}})^2\partial^2_{xx}$$

For the rest of the proof, recall (4.2.22), together with (4.1.1) and Remark 11, these imply that $\sigma_X^{\tilde{a}}(\tilde{X})$, $\sigma_X^{\tilde{a}}(\tilde{X})^{-1}$ and $\mu_X^{\tilde{a},\tilde{b}}(\tilde{X})$ are $\tilde{\mathbb{P}}$-essentially bounded. Performing an equivalent change of measure shows one can find $\check{\mathbb{P}} \sim \tilde{\mathbb{P}}$ and a $\check{\mathbb{P}}$-Brownian motion $\check{W}$ such that:

$$\tilde{X} = \int_0^\cdot \sigma_X^{\tilde{a}_s}(\tilde{X}_s) d\check{W}_s. \tag{4.2.27}$$

Clearly, both $\check{\mathbb{P}}$ and $\check{W}$ depend on $(\tilde{a},\tilde{b},y_0)$.

1. We first show that $y_0 = \partial_x\varphi(t_0,x_0)$, and therefore

$$\int_0^\theta \ell_s\, ds + \int_0^\theta\int_0^s m_r dr d\tilde{X}_s + \int_0^\theta\int_0^s n_r d\tilde{X}_r d\tilde{X}_s \geq 0. \tag{4.2.28}$$

Let $\check{\mathbb{P}}^\lambda \sim \check{\mathbb{P}}$ be the measure under which

$$\check{W}^\lambda := \check{W} + \int_0^\cdot \lambda[\sigma_X^{\tilde{a}_s}(\tilde{X}_s)]^{-1}(y_0 - \partial_x\varphi(t_0,x_0))ds$$

is a $\check{\mathbb{P}}^\lambda$-Brownian motion. Consider the case $\theta_o := \eta > 0$. Since all the coefficients are bounded, taking expectation under $\check{\mathbb{P}}^\lambda$ and using (4.2.26) imply that for some $C' > 0$

$$C'\eta \geq \lambda(y_0 - \partial_x\varphi(t_0,x_0))^2\mathbb{E}^{\check{\mathbb{P}}^\lambda}[\theta] + \mathbb{E}^{\check{\mathbb{P}}^\lambda}\left[\int_0^\theta \left(\int_0^s m_r dr + \int_0^s n_r d\tilde{X}_r\right)\lambda(y_0 - \partial_x\varphi(t_0,x_0))ds\right].$$

Divide both sides by η and use the fact that $(\eta \wedge \theta_1)/\eta \to 1$ $\check{\mathbb{P}}^\lambda$-a.s. as $\eta \to 0$, one obtains

$$C' \geq \lambda(y_0 - \partial_x\varphi(t_0,x_0))^2.$$

Then, we send $\lambda \to \infty$ to deduce that $y_0 = \partial_x\varphi(t_0,x_0)$.

2. We now prove that

$$\partial^2_{xx}\varphi(t_0,x_0) \leq \gamma_Y^{\tilde{a}_0}(x_0) \leq \bar{\gamma}(x_0). \tag{4.2.29}$$

First, let us consider the time change

$$h(t) = \inf\{r \geq 0 : \int_0^r \left[(\sigma_X^{\tilde{a}_s}(\tilde{X}_s))^2\mathbf{1}_{[0,\theta]}(s) + \mathbf{1}_{[0,\theta]^c}(s)\right]ds \geq t\}.$$

Again, $\sigma_X^{\tilde{a}}(\tilde{X})$ and $\sigma_X^{\tilde{a}}(\tilde{X})^{-1}$ are essentially bounded by Remark 11, so that h is absolutely continuous and its density $\mathfrak{h}$ satisfies

$$0 < \underline{\mathfrak{h}}t \leq \mathfrak{h}(t) := \left[(\sigma_X^{\tilde{a}}(\tilde{X}))^2\mathbf{1}_{[0,\theta]}(t) + \mathbf{1}_{[0,\theta]^c}(t)\right]^{-1} \leq \bar{\mathfrak{h}}t \tag{4.2.30}$$

for some constants $\underline{\mathfrak{h}}$ and $\bar{\mathfrak{h}}$, for all $t \geq 0$. Moreover, $\hat{W} := \tilde{X}_h$ is a Brownian motion in the time changed filtration. Take $\theta_o := h^{-1}(\eta)$ for some $0 < \eta < 1$, then (4.2.28) reads

$$0 \leq \int_0^{\eta\wedge h^{-1}(\theta_1)} \ell_{h(s)}\mathfrak{h}(s)\,ds + \int_0^{\eta\wedge h^{-1}(\theta_1)}\int_0^s m_{h(r)}\mathfrak{h}(r)drd\hat{W}_s + \int_0^{\eta\wedge h^{-1}(\theta_1)}\int_0^s n_{h(r)}d\hat{W}_r d\hat{W}_s. \tag{4.2.31}$$

All the involved processes are continuous and bounded, and $(\eta \wedge h^{-1}(\theta_1))/\eta \to 1$ a.s. as $\eta \to 0$, (4.2.31) combined with [19, Theorem A.1 b. and Proposition A.3] implies that

$$\gamma_Y^{\tilde{a}_0}(x_0) - \partial^2_{xx}\varphi(t_0,x_0) = \lim_{r\downarrow 0} n_{h(r)} = \lim_{r\downarrow 0} n_r \geq 0.$$

Since $\gamma_Y^{\tilde{a}}(\tilde{X}) \leq \bar{\gamma}(\tilde{X})$, this proves (4.2.29).

3. It remains to show that the first term in the definition of $F[\varphi](t_0,x_0)$ is also non-negative, recall (4.1.11). Again, let us take $\theta_o := h^{-1}(\eta)$ and recall from 2. that $\lim_{\eta\to 0}(\eta \wedge h^{-1}(\theta_1))/\eta = 1$ $\check{\mathbb{P}}$-a.s. Note that $\tilde{a}$ being of the form (4.2.21) with the condition (4.2.22), it satisfies [19, Condition (A.2)], and so does n. Using [19, Theorem A.2 and Proposition A.3] and (4.2.31), we deduce that $\ell_0\mathfrak{h}(0) - \frac{1}{2}n_0 \geq 0$. Now (4.2.30) and direct computations based on (4.1.4) imply

$$\begin{aligned}
0 &\leq \frac{1}{2}\tilde{a}_0^2 f(x_0) - \mathcal{L}^{\tilde{a}_0}\varphi(t_0,x_0) - \frac{1}{2}\left(\gamma_Y^{\tilde{a}_0}(x_0) - \partial^2_{xx}\varphi(t_0,x_0)\right)(\sigma_X^{\tilde{a}_0}(x_0))^2 \\
&= \frac{1}{2}\tilde{a}_0^2 f(x_0) - \partial_t\varphi(t_0,x_0) - \frac{1}{2}\gamma_Y^{\tilde{a}_0}(x_0)(\sigma_X^{\tilde{a}_0}(x_0))^2 \\
&= -\partial_t\varphi(t_0,x_0) - \frac{1}{2}\frac{\sigma^2(x_0)}{1 - f(x_0)\gamma_Y^{\tilde{a}_0}(x_0)}\gamma_Y^{\tilde{a}_0}(x_0) \\
&\leq -\partial_t\varphi(t_0,x_0) - \frac{1}{2}\frac{\sigma^2(x_0)}{1 - f(x_0)\partial^2_{xx}\varphi(t_0,x_0)}\partial^2_{xx}\varphi(t_0,x_0),
\end{aligned}$$

in which we use the facts that $\partial^2_{xx}\varphi(t_0,x_0) \le \gamma_Y^{\tilde a_0}(x_0) \le \bar\gamma(x_0)$ and $z \mapsto z/(1-f(x_0)z)$ in non-decreasing on $(-\infty, \bar\gamma(x_0)] \subset (-\infty, 1/f(x_0))$, for the last inequality.

b. We now consider the boundary condition at T. Since $\underline{\mathrm{v}}^k_{\bar\gamma}$ is a supersolution of $\bar\gamma - \partial^2_{xx}\varphi \ge 0$ on $[0,T)\times\mathbb{R}$, the same arguments as in [23, Lemma 5.1] imply that $\underline{\mathrm{v}}^k_{\bar\gamma} - \bar\Gamma$ is concave for any twice differentiable function $\bar\Gamma$ such that $\partial^2_{xx}\bar\Gamma = \bar\gamma$. The function $\underline{\mathrm{v}}^k_{\bar\gamma}$ being lower-semicontinuous, the map

$$x \mapsto G(x) := \liminf_{\substack{t' \to T,\, x' \to x \\ t' < T}} \underline{\mathrm{v}}^k_{\bar\gamma}(t',x')$$

is such that $G \ge g$ and $G - \bar\Gamma$ is concave. Hence, $G = (G-\bar\Gamma)^{\rm conc} + \bar\Gamma \ge (g-\bar\Gamma)^{\rm conc} + \bar\Gamma = \hat g$. □

It remains to state the dynamic programming principle used in the above proof.

Proposition 9. *Fix $(t,x,v,y) \in [0,T]\times\mathbb{R}^2\times[-k,k]$ and let θ be a stopping time for the $\tilde{\mathbb{P}}$-augmentation of $\tilde{\mathbb{F}}^\circ$ that takes $\tilde{\mathbb{P}}$-a.s. values in $[0,T-t]$. Assume that $\tilde{\mathbb{P}} \in \tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y)$. Then,*

$$\tilde V^{x,v,\tilde\phi}_\theta \ge \underline{\mathrm{v}}^k_{\bar\gamma}(t+\theta, \tilde X^{x,\tilde\phi}_\theta) \ \ \tilde{\mathbb{P}} - \text{a.s.},$$

in which $\tilde\phi := (y,\tilde a,\tilde b)$.

Proof. Since $\underline{\mathrm{v}}^k_{\bar\gamma}$ is lower-semicontinuous and all the involved processes have continuous paths, up to approximating θ by a sequence of stopping times valued in finite time grids, it suffices to prove our claim in the case $\theta \equiv r \in [0,T-t]$. Let $\tilde{\mathbb{P}}_\omega$ be a regular conditional probability given $\tilde{\mathcal{F}}^\circ_r$ for $\tilde{\mathbb{P}}$. It coincides with $\tilde{\mathbb{P}}[\cdot|\tilde{\mathcal{F}}^\circ_r](\omega)$ outside a set N of $\tilde{\mathbb{P}}$-measure zero. Then, for all $\omega \notin N$, $0 \le \delta \le 1$ and $r \ge 0$ the conditions (4.2.21)-(4.2.22)-(4.2.23) hold for $\tilde{\mathbb{P}}^r_\omega$ defined on $C(\mathbb{R}_+)^5$ by

$$\tilde{\mathbb{P}}^r_\omega[\omega' \in A] = \tilde{\mathbb{P}}_\omega[\omega'_{r+\cdot} \in A].$$

Moreover, [20, Theorem 3.3] ensures that, after possibly modifying N,

$$\tilde{\mathbb{P}}^r_\omega\left[\tilde V^{\xi_r(\omega),\vartheta_r(\omega),\hat\phi(\omega)}_{T-(t+r)} \ge g(\tilde X^{\xi_r(\omega),\hat\phi(\omega)}_{T-(t+r)})\right] = 1$$
$$\text{and } \ \tilde{\mathbb{P}}^r_\omega\left[\gamma^{\tilde a}_Y(\tilde X^{\xi_r(\omega),\hat\phi(\omega)}) \le \bar\gamma(\tilde X^{\xi_r(\omega),\hat\phi(\omega)}) \text{ on } \mathbb{R}_+\right] = 1,$$

for $\omega \notin N$, in which

$$(\xi_r,\vartheta_r,\hat\phi) := (\tilde X^{x,\tilde\phi}_r, \tilde V^{x,v,\tilde\phi}_r, (\tilde Y^{x,\tilde\phi}_r, \tilde a, \tilde b)).$$

This shows that $\vartheta_r(\omega) \ge \underline{\mathrm{v}}^k_{\bar\gamma}(t+r,\xi_r(\omega))$ outside the null set N, the required result. □

4.2.3 Conclusion of the proof and construction of almost optimal strategies

We first conclude the proof of Theorem 6.

Proof of Theorem 6. Proposition 3 and Theorem 2 imply that $\bar{\mathrm{v}}_{\bar\gamma} \ge \mathrm{v}_{\bar\gamma}$ in which $\bar{\mathrm{v}}_{\bar\gamma}$ has linear growth and is a continuous viscosity solution of (4.1.13). On the other hand, Proposition 6 and Theorem 4 imply that $\underline{\mathrm{v}}_{\bar\gamma} \le \mathrm{v}_{\bar\gamma}$ on $[0,T)\times\mathbb{R}$ in which $\underline{\mathrm{v}}_{\bar\gamma}$ has linear growth and is a viscosity supersolution of (4.1.13). By the comparison result of Theorem 3 applied with $\epsilon = 0$, $\underline{\mathrm{v}}_{\bar\gamma} \ge \bar{\mathrm{v}}_{\bar\gamma}$. Hence,

$$\mathrm{v}_{\bar\gamma} = \underline{\mathrm{v}}_{\bar\gamma} = \bar{\mathrm{v}}_{\bar\gamma} \ \text{ on } [0,T)\times\mathbb{R} \text{ and } \underline{\mathrm{v}}_{\bar\gamma} = \bar{\mathrm{v}}_{\bar\gamma} \text{ on } [0,T]\times\mathbb{R} \tag{4.2.32}$$

Along with the continuity of $\bar{\mathrm{v}}_{\bar\gamma}$, this shows that

$$\lim_{\substack{(t',x') \to (T,x) \\ t' < T}} \mathrm{v}_{\bar\gamma}(t',x') = \bar{\mathrm{v}}_{\bar\gamma}(T,x) = \underline{\mathrm{v}}_{\bar\gamma}(T,x).$$

One can conclude that $\mathrm{v}_{\bar\gamma}$ is a viscosity solution of (4.1.13), with linear growth. □

Remark 12 (Almost optimal controls)**.** *In the proof of Theorem 2, we have constructed a super-hedging strategy starting from* $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(t,x)$*. Since* $\bar{\mathrm{v}}^{\epsilon,K,\delta}_{\bar{\gamma}}(t,x) \to \bar{\mathrm{v}}_{\bar{\gamma}}(t,x) = \mathrm{v}_{\bar{\gamma}}(t,x)$ *as* $\delta, \epsilon \to 0$ *and* $K \to \infty$*, this provides a way to construct super-hedging strategies associated to any initial wealth* $v > \mathrm{v}_{\bar{\gamma}}(t,x)$*.*

4.3 Adding a resilience effect

In this section, we explain how a resilience effect can be added to our model. In the discrete rebalancement setting, we replace the dynamics (3.1.6) by

$$X^n = X_0 + \int_0^{\cdot} \mu(X^n_s)ds + \int_0^{\cdot} \sigma(X^n_s)dW_s + R^n,$$

in which R^n is defined by

$$R^n = R_0 + \sum_{i=1}^{n} \mathbf{1}_{[t^n_i,T]}\delta^n_{t^n_i} f(X^n_{t^n_i-}) - \int_0^{\cdot} \rho R^n_s ds,$$

for some $\rho > 0$ and $R_0 \in \mathbb{R}$. The process R^n models the impact of past trades on the price, the last term in its dynamics is the resilience effect. The continuous time dynamics now becomes

$$\begin{aligned}
X &= X_0 + \int_0^{\cdot} \sigma(X_s)dW_s + \int_0^{\cdot} f(X_s)dY_s + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s) - \rho R_s)ds \\
R &= R_0 + \int_0^{\cdot} f(X_s)dY_s + \int_0^{\cdot} (a_s(\sigma f')(X_s) - \rho R_s)ds \\
V &= V_0 + \int_0^{\cdot} Y_s dX_s + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X_s)ds.
\end{aligned}$$

This is obtained as a straightforward extension of Proposition 3.1.1.

Let $\mathrm{v}^R_{\bar{\gamma}}(t,x)$ be defined as the super-hedging price $\mathrm{v}_{\bar{\gamma}}(t,x)$ but for these new dynamics and for $R_t = 0$. The following states that $\mathrm{v}^R_{\bar{\gamma}} = \mathrm{v}_{\bar{\gamma}}$, i.e. adding a resilience effect does not affect the super-hedging price.

Proposition 10. $\mathrm{v}_{\bar{\gamma}} = \mathrm{v}^R_{\bar{\gamma}}$ *on* $[0,T] \times \mathbb{R}$*.*

Proof. 1. To show that $\mathrm{v}_{\bar{\gamma}} \geq \mathrm{v}^R_{\bar{\gamma}}$, it suffices to reproduce the arguments of the proof of Theorem 2 in which the drift part of the dynamics of X does not play any role. These arguments show that $\bar{\mathrm{v}}_{\bar{\gamma}} \geq \mathrm{v}^R_{\bar{\gamma}}$. Then, one uses the fact that $\mathrm{v}_{\bar{\gamma}} = \bar{\mathrm{v}}_{\bar{\gamma}}$, by (4.2.32).
2. As for the opposite inequality, we use the weak formulation of Section 4.2.2 and a simple Girsanov's transformation. For ease of notations, we restrict ourselves to $t = 0$. Fix $v > \mathrm{v}^R_{\bar{\gamma}}(0,x)$, for some $x \in \mathbb{R}$. Then, one can find $k \geq 1$, $(c,y) \in \mathbb{R} \times [-k,k]$ satisfying $v = c + yx$, and $(a,b) \in \mathcal{A}_{k,\bar{\gamma}}(0,x)$ such that $V_T \geq g(X_T)$, with (V,X,Y,R) defined by the corresponding initial data and controls. We write

$$a = a_0 + \int_0^{\cdot} \beta_s ds + \int_0^{\cdot} \alpha_s dW_s$$

as a decomposition into an Itô process. Let $\mathbb{Q}^R \sim \mathbb{P}$ be the probability measure under

which $W^R := W - \int_0^\cdot (\rho R_s/\sigma(X_s))ds$ is a $\mathbb{Q}^R$-Brownian motion, recall (4.1.1). Then,

$$\begin{aligned}
X &= X_0 + \int_0^\cdot \sigma(X_s)dW_s^R + \int_0^\cdot f(X_s)dY_s + \int_0^\cdot (\mu(X_s) + a_s(\sigma f')(X_s))ds \\
Y &= Y_0 + \int_0^\cdot (b_s + a_s\rho R_s/\sigma(X_s))ds + \int_0^\cdot a_s dW_s^R \\
a &= a_0 + \int_0^\cdot (\beta_s + \alpha_s\rho R_s/\sigma(X_s))ds + \int_0^\cdot \alpha_s dW_s^R \\
V &= V_0 + \int_0^\cdot Y_s dX_s + \frac{1}{2}\int_0^\cdot a_s^2 f(X_s)ds.
\end{aligned}$$

Upon regarding $(a, b+a\rho R/\sigma(X), \alpha, \beta+\alpha\rho R/\sigma(X), W^R)$ as a generic element of the canonical space $C([0,T])^5$ introduced in Section 4.2.2, then $\mathbb{Q}^R$ belongs to $\tilde{\mathcal{A}}_k \cap \tilde{\mathcal{G}}_{k,\bar\gamma}(t,x,v,y)$, and therefore $v > \underline{\mathrm{v}}_{\bar\gamma}(0,x)$. Hence, $\mathrm{v}^R_{\bar\gamma}(0,x) \geq \underline{\mathrm{v}}_{\bar\gamma}(0,x)$, and thus $\mathrm{v}^R_{\bar\gamma}(0,x) \geq \mathrm{v}_{\bar\gamma}(0,x)$ by (4.2.32). □

4.4 Numerical approximation and examples

In this section, we provide an example of numerical schemes that converges towards the unique continuous viscosity solution of (4.1.13) with linear growth. The scheme is then exemplified using two numerical applications in the case of constant market impact and gamma constraint.

4.4.1 Finite difference scheme

Given a map ϕ and $h := (h_t, h_x) \in (0,1)^2$, define

$$\begin{aligned}
L_1^h(t,x,y,\phi) &:= -\frac{\phi(t+h_t,x) - y}{h_t} - \frac{\sigma^2(x)G^h(t,x,y,\phi)}{2(1-f(x)G^h(t,x,y,\phi))} \\
L_2^h(t,x,y,\phi) &:= \bar\gamma(x) - G^h(t,x,y,\phi)
\end{aligned}$$

where

$$G^h(t,x,y,\phi) := \frac{\phi(t+h_t,x+h_x) + \phi(t+h_t,x-h_x) - 2y}{h_x^2}.$$

The numerical scheme is set on the grid $\pi_h := \{(t_i,x_j) = (ih_t, \underline{x}+jh_x) : i \leq n_t, j \leq n_x\}$, with $n_t h_t = T$ for some $n_t \in \mathbb{N}$, and $n_x h_x = \overline{x} - \underline{x}$, for some real numbers $\underline{x} < \overline{x}$. To paraphrase, $\mathrm{v}^h_{\bar\gamma}$ is defined on π_h as the solution of

$$\begin{aligned}
S(h,t_i,x_j,\mathrm{v}^h_{\bar\gamma}(t_i,x_j),\mathrm{v}^h_{\bar\gamma}) &= 0 \ \text{ for } \ i < n_t, 1 \leq j \leq n_x - 1 \\
\mathrm{v}^h_{\bar\gamma} &= \hat g \ \text{ on } \pi_h \cap \{(\{T\}\times\mathbb{R}) \cup ([0,T] \cap \{\underline{x},\overline{x}\})\}
\end{aligned} \tag{4.4.1}$$

where

$$S(h,t,x,y,\phi) := (\bar w - y) \vee (y - \underline{w}) \wedge \min_{l=1,2} \Big\{ L_l^h(t,x,y,\phi) \Big\}$$

with $\bar w$ and $\underline{w}$ as in Remark 9.

Theorem 5. *The equation (4.4.1) admits a unique solution $\mathrm{v}^h_{\bar\gamma}$, for all $h := (h_t,h_x) \in (0,1)^2$. Moreover, if $h_t/h_x^2 \to 0$ and $h_x^2 \to 0$, then $\mathrm{v}^h_{\bar\gamma}$ converges locally uniformly to the unique continuous viscosity solution of (4.1.13) that has linear growth.*

Proof. The existence of a solution, that is bounded by the map with linear growth $|\bar{w}|+|\underline{w}|$, is obvious. We now prove uniqueness. First observe that L_2^h is strictly increasing in its y-component, and that

$$\frac{\partial L_1^h}{\partial y}(t,x,y,\phi) = \frac{1}{h_t} + \frac{\sigma^2(x)}{h_x^2(1-f(x)G^h(t,x,y,\phi))^2} > 0$$

on the domain $\{y : L_2^h(t_i,x_j,y,\phi) \geq 0\}$. Uniqueness of the solution follows.

It is easy to see that $\phi \mapsto S(\cdot,\phi)$ is non-decreasing, so that our scheme is monotone. Consistency is clear. Moreover, it is not difficult to check that the comparison result of Theorem 3 extends to this equation (there is an equivalence of the notions of super- and subsolutions in the class of functions w such that $\underline{w} \leq w \leq \bar{w}$). It then follows from [5, Theorem 2.1] that $\mathrm{v}_{\bar{\gamma}}^h$ converges locally uniformly to the unique continuous viscosity solution with linear growth of

$$\Big[(\bar{w}-\varphi) \vee (\varphi-\underline{w}) \wedge F[\varphi]\Big]\mathbf{1}_{[0,T)} + (\varphi-\hat{g})\mathbf{1}_{\{T\}} = 0.$$

In view of (4.2.32), Remark 9 and Theorem 6, $\mathrm{v}_{\bar{\gamma}}$ is the unique viscosity solution of the above equation. □

4.4.2 Numerical examples: the fixed impact case

To illustrate the above numerical scheme, we place ourselves in the simpler case where $f \equiv \lambda > 0$ and $\bar{\gamma} > 0$ are constant. The dynamics of the stock is given by the Bachelier model

$$dX_t = \sigma\, dW_t,$$

with $\sigma := 0.2$. In the following, $T = 2$.

First, we consider a European Butterfly option with three strikes $K_1 = -1 < K_2 = 0 < K_3 = 1$, where $K_1 + 1/(2\bar{\gamma}) \leq K_2 \leq K_3 - 1/(2\bar{\gamma})$. Its pay-off is

$$g(x) = (x-K_1)^+ - 2(x-K_2)^+ + (x-K_3)^+,$$

and the corresponding face-lifted function $\hat{g}$ can be computed explicitly:

$$\begin{aligned}\hat{g}(x) \;=\; & \frac{\bar{\gamma}}{2}(x-x_1^-)^2\mathbf{1}_{[x_1^-,x_1^+)} + (x-K_1)\mathbf{1}_{[x_1^+,K_2)} \\ & +(x-K_1-2(x-K_2))\mathbf{1}_{[K_2,x_2^-)} \\ & + \Big(\frac{\bar{\gamma}}{2}(x-x_2^+)^2 + 2K_2 - (K_1+K_3)\Big)\mathbf{1}_{[x_2^-,x_2^+)} \\ & +(2K_2-(K_1+K_3))\mathbf{1}_{[x_2^+,+\infty)},\end{aligned}$$

where $x_1^\pm = K_1 \pm 1/(2\bar{\gamma})$ and $x_2^\pm = K_3 \pm 1/(2\bar{\gamma})$.

In Figure 4.1, we separately show the effect of the gamma constraint and of the market impact. As observed in Remark 6, the price is non-decreasing with respect to the impact parameter λ and bounded from below by the hedging price obtained in the model without impact nor gamma constraint. On the left and right tails of the curves, we observe the effect of the gamma constraint. It does not operate around $x = 0$ where the gamma is non-positive. The effect of the market impact operates only in areas of high convexity (around $x = -1.5$ and $x = 1.5$) or of high concavity (around $x = 0$).

In Figure 4.2, we perform similar computations but for a call spread option, where

$$g(x) = (x-K_1)^+ - (x-K_2)^+,$$

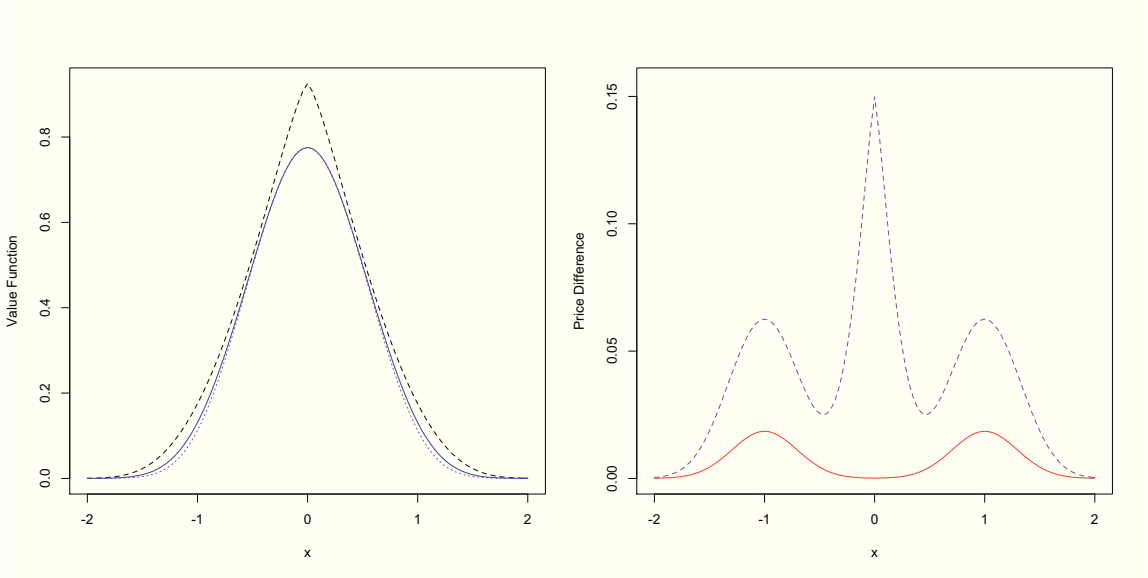

Figure 4.1: Left: Super-hedging price of the Butterfly option. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$; dotted line: $\lambda = 0$, $\bar{\gamma} = +\infty$. Right: Difference with the price associated to $\lambda = 0$, $\bar{\gamma} = +\infty$. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$.

with $K_1 = -1 < K_2 = 1$ such that $K_1 + 1/(2\bar{\gamma}) \leq K_2$. The face-lifted function $\hat{g}$ is given by

$$\hat{g}(x) = \frac{\bar{\gamma}}{2}(x - x^-)^2 \mathbf{1}_{[x^-, x^+)} + (x - K_1)\mathbf{1}_{[x^+, K_2)} + (K_2 - K_1)\mathbf{1}_{[K_2, +\infty)}$$

with $x^{\pm} = K_1 \pm 1/(2\bar{\gamma})$.

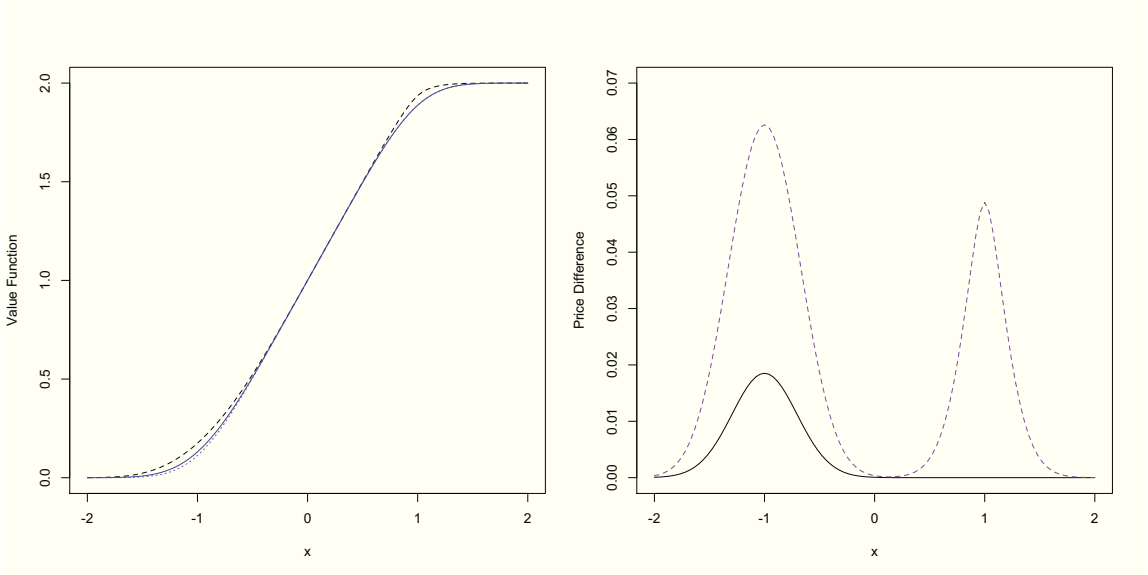

Figure 4.2: Left: Super-hedging price of the Call Spread option. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$; dotted line: $\lambda = 0$, $\bar{\gamma} = +\infty$. Right: Difference with the price associated to $\lambda = 0$, $\bar{\gamma} = +\infty$. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$.

4.5 Appendix

The following lemma is very standard, we prove it for completeness.

Lemma 2. *A upper-(resp. lower-) semicontinuous map is a viscosity sub-(resp. super-) solution of*

$$F^\epsilon_\kappa[\varphi]\mathbf{1}_{[0,T)} + (\varphi - \hat{g}^\epsilon_K)\mathbf{1}_{\{T\}} = 0$$

if and only if it is a viscosity sub-(resp. super-) solution of $F^{\epsilon,K}_{\kappa,-}[\varphi] = 0$ *(resp.* $F^{\epsilon,K}_{\kappa,+}[\varphi] = 0$*).*

Proof. The equivalence on $[0,T)$ is evident, we only consider the parabolic boundary $\{T\}\times\mathbb{R}$. Since $F^{\epsilon,K}_{\kappa,+} \geq F^\epsilon_\kappa$ and $F^{\epsilon,K}_{\kappa,-} \leq F^\epsilon_\kappa$, only one implication is not completely trivial.
a. Let v be a viscosity supersolution of $F^{\epsilon,K}_{\kappa,+}[\varphi] = 0$, and $\varphi \in C^2$ a test function such that

$$(\text{strict}) \min_{[0,T]\times\mathbb{R}} (v - \varphi) = (v - \varphi)(T, x_0) = 0,$$

for some $x_0 \in \mathbb{R}$. We define a new test function $\phi \in C^2$,

$$\phi(t,x) := \varphi(t,x) - C(T-t),$$

so that $\partial_t\phi = \partial_t\varphi + C$. For $C > 0$ large enough, we have at (T, x_0)

$$\min_{x'\in D^\epsilon_\kappa} \min\left\{-\partial_t\phi - \frac{\sigma^2(x')\partial_{xx}\phi}{2(1 - f(x')\partial_{xx}\phi)}, \bar{\gamma}(x') - \partial_{xx}\phi\right\} < 0$$

Now that the minimum of $v - \phi$, being equal to 0, is also attaint at (T, x_0), it must hold that $F^{\epsilon,K}_{\kappa,+}[\phi](T, x_0) \geq 0$, and therefore

$$v(T,x_0) - \hat{g}^\epsilon_K(x_0) = \varphi(T,x_0) - \hat{g}^\epsilon_K(x_0) = \phi(T,x_0) - \hat{g}^\epsilon_K(x_0) \geq 0.$$

b. Let v be a viscosity subsolution of $F^{\epsilon,K}_{\kappa,-}[\varphi] = 0$, and $\varphi \in C^2$ a test function such that

$$(\text{strict}) \max_{[0,T]\times\mathbb{R}} (v - \varphi) = (u - \varphi)(T, x_0),$$

for some $x_0 \in \mathbb{R}$. Then, $F^{\epsilon,K}_{\kappa,-}[\varphi](T, x_0) \leq 0$. By replacing φ by ϕ, defined for $\alpha > 0$ as

$$\phi(t,x) := \varphi(t, x_0 + \alpha(x - x_0)) + C(T-t),$$

we obtain a new test function at (T, x_0). Since $\inf\bar{\gamma} > 0$, recall (4.1.1), we can take α small enough so that

$$\min_{x'\in D^\epsilon_\kappa} \{\bar{\gamma}(x') - \partial_{xx}\phi(T, x_0)\} > 0.$$

As in the previous step, we can now choose $C > 0$ such that

$$\min_{x'\in D^\epsilon_\kappa} \left\{-\partial_t\phi - \frac{\sigma^2(x')\partial_{xx}\phi}{2(1 - f(x')\partial_{xx}\phi)}\right\} > 0$$

at (T, x_0). Since $F^{\epsilon,K}_{\kappa,-}[\phi](T, x_0) \leq 0$, we conclude that $v(T, x_0) = \phi(T, x_0) \leq \hat{g}^\epsilon_K(x_0)$. □

Chapter 5

Numerical approximation of BSDEs using local polynomial drivers and branching processes

Contents

5.1 Introduction **73**

5.2 Numerical method for a class of BSDE based on branching processes **74**

5.2.1 Local polynomial approximation of the generator 75

5.2.2 Picard iteration with doubly reflected BSDEs 76

5.2.3 A branching diffusion representation for $\bar{Y}^m$ 78

5.2.4 The numerical algorithm 79

5.3 Example of application **82**

5.4 Appendix **86**

5.4.1 Technical lemmas 86

5.4.2 More on the error analysis for the abstract numerical approximation 89

5.1 Introduction

Since the seminal paper of Pardoux and Peng [42], the theory of Backward Stochastic Differential Equations (BSDEs hereafter) has been largely developed, and has lead to many applications in optimal control, finance, etc. (see e.g. El Karoui, Peng and Quenez [24]).Different approaches have been proposed during the last decade to solve them numerically, without relying on pure PDE based resolution methods. A first family of numerical schemes, based on a time discretization technique, has been introduced by Bally and Pagès [2], Bouchard and Touzi [15] and Zhang [53], and generated a large stream of the literature. The implementation of these numerical schemes requires the estimation of a sequence of conditional expectations, which can be realized by simulations along with either non-linear regression techniques or Malliavin integration by parts based representation of conditional expectations, or by a quantization approach, see e.g. [16, 26] for references and error analysis.

Another type of numerical algorithms is based on a pure forward simulation of branching processes, and was introduced by Henry-Labordère [27], and Henry-Labordère, Tan and Touzi [29] (see also the recent extension by Henry-Labordère et al. [28]). The main

advantage of this new algorithm is that it avoids the estimation of conditional expectations. Instead, it relies on the probabilistic representation in terms of branching processes of the so-called KPP (Kolmogorov-Petrovskii-Piskunov) equation:

$$\partial_t u(t,x) \; + \; \frac{1}{2} D^2 u(t,x) \; + \; \sum_{k\geq 0} p_k u^k(t,x) \; = \; 0, \; u(T,x) = g(x). \tag{5.1.1}$$

Here, D^2 is the Laplacian on $\mathbb{R}^d$, and $(p_k)_{k\geq 0}$ is a probability mass sequence, i.e. $p_k \geq 0$ and $\sum_{k\geq 0} p_k = 1$. This is a natural extension of the classical Feynmann-Kac formula, which has gain its popularity since the works of Skorokhod [47], Watanabe [52] and McKean [40], among others. The PDE (5.1.1) corresponds to a BSDE with a polynomial driver and terminal condition $g(W_T)$:

$$Y_\cdot = g(W_T) + \int_\cdot^T \sum_{k\geq 0} p_k (Y_t)^k dt - \int_\cdot^T Z_t dW_t,$$

in which W is a Brownian motion. Since $Y_\cdot = u(\cdot, W_\cdot)$, the Y-component of this BSDE can be estimated in terms of the branching process based Feynman-Kac representation of (5.1.1) by means of a pure forward Monte-Carlo scheme, see Section 5.2.3 below. The idea is not new. It was already proposed in Rasulov, Raimov and Mascagni [43], although no rigorous convergence analysis was provided. Extensions to more general drivers can be found in [27–29]. Similar algorithms have been studied by Bossy et al. [9] to solve non-linear Poisson-Boltzmann equations.

It would be tempting to use this representation to solve BSDEs with Lipschitz drivers, by approximating the drivers by polynomials. This is however not feasible in general. The reason being that PDEs (or BSDEs) with polynomial drivers, of degree bigger or equal to two, typically explode in finite time. They are only well posed on a small time interval. It becomes worse when the degree of the polynomial increases. Hence, no convergence can be expected for the case of general drivers.

In this paper, we instead propose to use a local polynomial approximation. The convergence of the sequence of approximating drivers to the original one can be ensured without the corresponding BSDEs exploding. Thus those BSDEs can be defined on a arbitrary time interval. This local polynomial approximation requires to be applied with the Picard iteration scheme, as the form of the polynomials depend on the space position of the solution Y itself. However, unlike classical Picard iteration schemes for BSDEs, see e.g. Bender and Denk [6], we do not need to dispose precise estimation on the whole path of the solution at each Picard iteration. Indeed, if local polynomials are defined through a partition $(A_i)_i$ of $\mathbb{R}$, then one only needs to know in which A_i the solution stays at certain branching times of the underlying process. If the A_i's are large enough, one does not need to provide high precision on the intermediate estimations. We refer to Remark 5.2.3 for more details.

All our results will be presented in a Markovian context for simplification. However, our arguments can be extended to a non-Markovian setting trivially.

5.2 Numerical method for a class of BSDE based on branching processes

Let $T > 0$, W be a standard d-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and X be the solution of the stochastic differential equation:

$$X = X_0 + \int_0^\cdot \mu(X_s)\, dt + \int_0^\cdot \sigma(X_s)\, dW_s, \tag{5.2.1}$$

where X_0 is a constant, lying in a compact subset $\mathbf{X}$ of $\mathbb{R}^d$, and $(\mu, \sigma) : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{M}^d$ is assumed to be Lipschitz continuous with support contained in $\mathbf{X}$. Our aim is to provide a numerical scheme for the resolution of the BSDE

$$Y_\cdot = g(X_T) + \int_\cdot^T f(X_s, Y_s)\, ds - \int_\cdot^T Z_s\, dW_s. \tag{5.2.2}$$

In the above, $g : \mathbb{R}^d \mapsto \mathbb{R}$ is assumed to be measurable and bounded, $f \in \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ is measurable with linear growth and Lipschitz in its second argument, uniformly in the first one. As a consequence, there exists $M \geq 1$ such that

$$|g(X_T)| \leq M \ \text{ and } |X| + |Y| \leq M \ \text{ on } [0,T]. \tag{5.2.3}$$

Remark 5.2.1. *The above conditions are imposed to easily localize the solution Y of the BSDE, which will be used in our estimates later on. One could also assume that g and f have polynomial growth in their first component and that $\mathbf{X}$ is not compact. After possibly truncating the coefficients and reducing their support, one would reduce to our conditions. Then, standard estimates and stability results for SDEs and BSDEs could be used to estimate the additional error in a very standard way. See e.g. [24].*

5.2.1 Local polynomial approximation of the generator

The first main ingredient of our algorithm consists in approximating the driver f by a driver $f_{\ell_\circ}$ that has a local polynomial structure. Namely, let

$$f_{\ell_\circ} : (x, y, y') \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \mapsto \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y'), \tag{5.2.4}$$

in which $(a_{j,\ell}, \varphi_j)_{\ell \leq \ell_\circ, j \leq j_\circ}$ is a family of continuous and bounded maps satisfying

$$|a_{j,\ell}| \leq C_{\ell_\circ}\ ,\ |\varphi_j(y_1') - \varphi_j(y_2')| \leq L_\varphi |y_1' - y_2'| \text{ and } |\varphi_j| \leq 1, \tag{5.2.5}$$

for all $y_1', y_2' \in \mathbb{R}$, $j \leq j_\circ$ and $\ell \leq \ell_\circ$, and some constants $C_{\ell_\circ}, L_\varphi \geq 0$. In the following, we shall assume that $\ell_\circ \geq 2$ without loss of generality. One can understand the $(a_{j,\ell})_{\ell \leq \ell_\circ}$ as the coefficients of a polynomial approximation of f on a subset A_j, the A_j's forming a partition of $[-M, M]$. Then, the φ_j's have to be regarded as smoothing kernels that allow one to pass in a Lipschitz way from one part of the partition to another one. We therefore assume that

$$\#\{j \in \{1, \cdots, j_\circ\} :\ \varphi_j(y) > 0\} \leq 2 \ \text{ for all } y \in \mathbb{R}, \tag{5.2.6}$$

and that $y \mapsto f_{\ell_\circ}(x, y, y)$ is globally Lipschitz. In particular,

$$\bar{Y}_\cdot = g(X_T) + \int_\cdot^T f_{\ell_\circ}(X_s, \bar{Y}_s, \bar{Y}_s)\, ds - \int_\cdot^T \bar{Z}_s\, dW_s, \tag{5.2.7}$$

admits a unique solution $(\bar{Y}, \bar{Z})$ such that $\mathbb{E}[\sup_{[0,T]} |\bar{Y}|^2] < \infty$. Moreover, by standard estimates, $(\bar{Y}, \bar{Z})$ provides a good approximation of (Y, Z) whenever $f_{\ell_\circ}$ is a good approximation of f:

$$\mathbb{E}\Big[\sup_{[0,T]} |Y - \bar{Y}|^2\Big] + \mathbb{E}\Big[\int_0^T |Z_t - \bar{Z}_t|^2 dt\Big] \leq C\mathbb{E}\Big[\int_0^T |f - f_{\ell_\circ}|^2(X_t, Y_t, Y_t) dt\Big], \tag{5.2.8}$$

for some $C > 0$ that does not depend on $f_{\ell_\circ}$, see e.g. [24].

The choice of $f_{\ell_\circ}$ will obviously depend on the application to be dealt with and needs no more comment. Let us just mention that our algorithm will be more efficient if the sets $\{y \in \mathbb{R} : \varphi_j(y) = 1\}$ are large and the intersection between the supports of the φ_j's is small, see Remark 5.2.3 below.

We also assume that

$$|\bar{Y}| \leq M. \tag{5.2.9}$$

Since we intend to keep $f_{\ell_\circ}$ with linear growth in its first component and bounded in the two other ones, uniformly in $\ell_\circ$, the above assumption does not incur any loss of generality.

5.2.2 Picard iteration with doubly reflected BSDEs

Our next step is to introduce a Picard iteration scheme to approximate the solution $\bar{Y}$ of (5.2.7). Note however that, although the map $y \mapsto f(x, y, y)$ is globally Lipschitz, the map $y \mapsto f(x, y, y')$ being a polynomial given y', is only locally Lipschitz in general. In order to reduce to a Lipschitz driver, we shall apply our Picard scheme to a doubly (discretely) reflected BSDE, with lower and upper barrier given by the bounds $-M$ and M for $\bar{Y}$, recall (5.2.9).

Let $h_\circ$ be defined by (5.4.1) in the Appendix, a lower bound for the explosion time of the BSDE of driver $y \mapsto f(x, y, y')$. Fix $h \in (0, h_\circ)$ such that $N_h := T/h \in \mathbb{N}$, and define

$$t_i^n = ih \quad \text{and} \quad \mathbb{T} := \{t_i^n,\ i = 0, \cdots, N_h\}. \tag{5.2.10}$$

We initialize our Picard scheme by setting

$$\bar{Y}_t^0 = \mathrm{y}(t, X_t) \ \text{ for } t \in [0, T], \tag{5.2.11}$$

in which y is a deterministic function, bounded by M and such that $\mathrm{y}(T, \cdot) = g$. Then, given $\bar{Y}^{m-1}$, for $m \geq 1$, we define $(\bar{Y}^m, \bar{Z}^m, \bar{K}^{m,+}, \bar{K}^{m,-})$ as the solution on $[0, T]$ of

$$\begin{aligned}
\bar{Y}_t^m \ &= \ g(X_T) + \int_t^T f_{\ell_\circ}(X_s, \bar{Y}_s^m, \bar{Y}_s^{m-1})\, ds - \int_t^T \bar{Z}_s^m\, dW_s \\
&\qquad + \int_{[t,T]\cap\mathbb{T}} d(\bar{K}^{m,+} - \bar{K}^{m,-})_s, \\
&-M \leq \bar{Y}_t^m \leq M, \ \forall t \in \mathbb{T}, \ a.s. \\
&\int_{\mathbb{T}} (\bar{Y}_s^m + M) d\bar{K}_s^{m,+} = \int_{\mathbb{T}} (\bar{Y}_s^m - M) d\bar{K}_s^{m,-} = 0,
\end{aligned} \tag{5.2.12}$$

where $\bar{K}^{m,+}$ and $\bar{K}^{m,-}$ are non-decreasing processes.

Remark 13. *Since the solution $\bar{Y}$ of (5.2.7) is bounded by M, the quadruple of processes $(\bar{Y}, \bar{Z}, \bar{K}^+, \bar{K}^-)$ (with $\bar{K}^+ \equiv \bar{K}^- \equiv 0$) is in fact the unique solution of the same reflected BSDE as in (5.2.12) but with $f_{\ell_\circ}(X, \bar{Y}, \bar{Y})$ in place of $f_{\ell_\circ}(X_s, \bar{Y}^m, \bar{Y}^{m-1})$.*

Remark 14. *One can equivalently define $\bar{Y}^m$ recursively. Let $\bar{Y}_T^m := g(X_T)$ be the terminal condition, and define on each $[t_i, t_{i+1}]$, $(Y_\cdot^m, Z_\cdot^m)$ as the solution on $[t_i, t_{i+1}]$ of*

$$Y_\cdot^m \ = \ \bar{Y}_{t_{i+1}}^m + \int_\cdot^{t_{i+1}} f_{\ell_\circ}(X_s, Y_s^m, \bar{Y}_s^{m-1}) ds - \int_\cdot^{t_{i+1}} Z_s^m dW_s. \tag{5.2.13}$$

Then, $\bar{Y}^m := Y^m$ on $(t_i, t_{i+1}]$, and $\bar{Y}_{t_i}^m := (-M) \vee Y_{t_i}^m \wedge M$.

The error due to our Picard iteration scheme is handled in a standard way. It depends on the constants

$$L_1 := 2C_{\ell_\circ} \sum_{\ell=1}^{\ell_\circ} \ell (M_h)^{\ell-1}, \quad L_2 := L_\varphi \sum_{\ell=0}^{\ell_\circ} 2C_{\ell_\circ} (M_h)^\ell,$$

where M_h is defined by (5.4.2).

Theorem 6. *The system (5.2.12) of doubly reflected BSDEs admits a unique solution $(\bar{Y}^m, \bar{Z}^m, \bar{K}^{m,+}, \bar{K}^{m,-})_{m\geq 0}$ such that $\bar{Y}^m$ is uniformly bounded for each $m \geq 0$. Moreover, for all $m \geq 0$, $|\bar{Y}^m|$ is uniformly bounded by the constant M_h, and*

$$|\bar{Y}_t^m - \bar{Y}_t|^2 \;\leq\; \frac{L_2}{\lambda^2} \Big(\frac{L_2(T-t)}{\lambda^2} \Big)^m (2M)^2 \frac{e^{\beta T}}{\beta},$$

for all $t \leq T$, and all constants $\lambda > 0$, $\beta > 2L_1 + L_2\lambda^2$.

Proof. i) First, when $\bar{Y}^m$ is uniformly bounded, $f_{\ell_\circ}(X_s, \bar{Y}_s^m, \bar{Y}_s^{m-1})$ can be considered to be uniformly Lipschitz in $\bar{Y}^m$, then (5.2.12) has at most one bounded solution. Next, in view of Lemma 5 and Remark 14, it is easy to see that (5.2.13) has a unique solution Y^m, bounded by M_h (defined by (5.4.2)) on each interval $[t_i, t_{i+1}]$. It follows the existence of the solution to (5.2.12). Moreover, $\bar{Y}^m$ is also bounded by M_h on $[0,T]$, and more precisely bounded by M on the discrete grid $\mathbb{T}$, by construction.
ii) Consequently, the driver $f_{\ell_\circ}(x, y, y')$ can be considered to be uniformly Lipschitz in y and y'. Moreover, using (5.2.5) and (5.2.6), one can identify the corresponding Lipschitz constants as L_1 and L_2.

Let us denote $\Delta\bar{Y}^m := \bar{Y}^m - \bar{Y}$ for all $m \geq 1$. We notice that, in Remark 14, the truncation operation $\bar{Y}_{t_i}^m := (-M) \vee Y_{t_i}^m \wedge M$ can only make the value $(\Delta\bar{Y}_{t_i}^m)^2$ smaller than $(Y_{t_i}^m - \bar{Y}_{t_i})^2$, due to the fact $|\bar{Y}| \leq M$. Thus we can apply Itô's formula to $(e^{\beta t}(\Delta\bar{Y}_t^{m+1})^2)_{t\geq 0}$ on each interval $[t_i, t_{i+1}]$, and then take expectation to obtain

$$\begin{aligned} &\mathbb{E}\big[e^{\beta t}(\Delta\bar{Y}_t^{m+1})^2\big] + \beta\mathbb{E}\Big[\int_t^T e^{\beta s}|\Delta\bar{Y}_s^{m+1}|^2 ds + \int_t^T e^{\beta s}|\Delta\bar{Z}_s^{m+1}|^2 ds\Big] \\ \leq\;& 2\mathbb{E}\Big[\int_t^T e^{\beta s}\Delta\bar{Y}_s^{m+1}\big(f_{\ell_\circ}(X_s, \bar{Y}_s^{m+1}, \bar{Y}_s^m) - f_{\ell_\circ}(X_s, \bar{Y}_s, \bar{Y}_s)\big) ds\Big]. \end{aligned}$$

Using the Lipschitz property of $f_{\ell_\circ}$ and the inequality $\lambda^2 + \frac{1}{\lambda^2} \geq 2$, it follows that the r.h.s. of the above inequality is bounded by

$$(2L_1 + L_2\lambda^2)\mathbb{E}\Big[\int_t^T e^{\beta s}(\Delta\bar{Y}_s^{m+1})^2 ds\Big] + \frac{L_2}{\lambda^2}\mathbb{E}\Big[\int_t^T e^{\beta s}(\Delta\bar{Y}_s^m)^2 ds\Big].$$

By definition $\beta \geq 2L_1 + L_2\lambda^2$, the above implies

$$\mathbb{E}\Big[e^{\beta t}(\Delta\bar{Y}_t^{m+1})^2\Big] \;\leq\; \frac{L_2}{\lambda^2}\mathbb{E}\Big[\int_t^T e^{\beta s}(\Delta\bar{Y}_s^m)^2 ds\Big], \tag{5.2.14}$$

and hence

$$\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta\bar{Y}_t^{m+1})^2 dt\Big] \;\leq\; \frac{L_2}{\lambda^2} T\mathbb{E}\Big[\int_0^T e^{\beta s}(\Delta\bar{Y}_s^m)^2 ds\Big].$$

Since $|\Delta\bar{Y}^0| = |\mathrm{y}(\cdot, X) - \bar{Y}| \leq 2M$ by (5.2.9) and our assumption $|\mathrm{y}| \leq M$, this shows that

$$\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta\bar{Y}_t^m)^2 dt\Big] \;\leq\; \big(\frac{L_2}{\lambda^2}T\big)^m (2M)^2 e^{\beta T}/\beta.$$

Plugging this in (5.2.14) leads to the required result at $t = 0$. It is then clear that the above estimation does not depend on the initial condition $(0, X_0)$, so that the same result holds true for every $t \in [0,T]$. □

5.2.3 A branching diffusion representation for $\bar{Y}^m$

We now explain how the solution of (5.2.13) on $[t_i, t_{i+1})$ can be represented by means of a branching diffusion system. More precisely, consider an element $(p_\ell)_{0\le\ell\le\ell_\circ} \in \mathbb{R}_+^{\ell_\circ+1}$ such that $\sum_{\ell=0}^{\ell_0} p_\ell = 1$, set $K_n := \{(1, k_2, \dots, k_n) : (k_2, \dots, k_n) \in \{0, \dots, \ell_\circ\}^n\}$ for $n \ge 1$, and $K := \cup_{n\ge1} K_n$. Let $(W^k)_{k\in K}$ be a sequence of independent d-dimensional Brownian motions, $(\xi_k)_{k\in K}$ and $(\delta_k)_{k\in K}$ be two sequences of independent random variables, such that

$$\mathbb{P}[\xi_k = \ell] = p_\ell, \ \ \ell \le \ell_\circ, k \in K,$$

and

$$\bar{F}(t) := \mathbb{P}[\delta_k > t] = \int_t^\infty \rho(s)ds, \ t \ge 0, \ k \in K, \tag{5.2.15}$$

for some continuous strictly positive map $\rho : \mathbb{R}_+ \to \mathbb{R}_+$. We assume that

$$(W^k)_{k\in K}, \ (\xi_k)_{k\in K}, \ (\delta_k)_{k\in K} \text{ and } W \text{ are independent.} \tag{5.2.16}$$

Given the above, we construct particles $X^{(k)}$ of the dynamics (5.2.1) up to a killing time T_k at which they split into ξ_k different (conditionally) independent particles with the same dynamics (5.2.1) till their own killing time. The construction is done as follows. First, we set $T_{(1)} := \delta_1$. Given $k = (1, k_2, \dots, k_n) \in K_n$ with $n \ge 2$, we let $T_k := \delta_k + T_{k-}$ in which $k- := (1, k_2, \dots, k_{n-1}) \in K_{n-1}$. Define the Brownian particles $(W^{(k)})_{k\in K}$ by the following induction: for the first generation

$$W^{((1))} := W^1 \mathbf{1}_{[0,T_{(1)}]} \ , \ \mathcal{K}_t^1 := \{(1)\}\mathbf{1}_{[0,T_{(1)}]}(t) + \emptyset \mathbf{1}_{[0,T_{(1)}]^c}(t), \ t \ge 0,$$

then, given $n \ge 2$ and $k \in \bar{\mathcal{K}}_T^{n-1} := \cup_{t\le T} \mathcal{K}_t^{n-1}$, let

$$W^{(k\oplus j)} := \Big(W^{(k)}_{\cdot\wedge T_k} + W^{k\oplus j}_{\cdot\vee T_k} - W^{k\oplus j}_{T_k} \Big) \mathbf{1}_{[0,T_{k\oplus j}]}, \ 1 \le j \le \xi_k,$$

and

$$\bar{\mathcal{K}}_t^n := \{k \oplus j : k \in \bar{\mathcal{K}}_T^{n-1}, 1 \le j \le \xi_k \text{ s.t. } t \in (0, T_{k\oplus j}]\}, \ \ \bar{\mathcal{K}}_t := \cup_{n\ge1} \bar{\mathcal{K}}_t^n,$$

$$\mathcal{K}_t^n := \{k \oplus j : k \in \bar{\mathcal{K}}_T^{n-1}, 1 \le j \le \xi_k \text{ s.t. } t \in (T_k, T_{k\oplus j}]\}, \ \ \mathcal{K}_t := \cup_{n\ge1} \mathcal{K}_t^n,$$

in which we use the notation $(1, k_1, \dots, k_{n-1}) \oplus j = (1, k_1, \dots, k_{n-1}, j)$.

Observe that the solution X^x of (5.2.1) on $[0, T]$ of initial condition $X_0^x = x \in \mathbb{R}^d$ can be identified in law on the canonical space as a process of the form $\Phi[x](\cdot, W)$ with $(x, s, \omega) \mapsto \Phi[x](s, \omega)$ being deterministic, $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{P}$-measurable, where $\mathcal{P}$ is the predictable σ-filed on $[0, T] \times \Omega$. We then define the particles $(X^{x,(k)})_{k\in K}$ by $X^{x,(k)} := \Phi[x](\cdot, W^{(k)})$.

Based on the above construction, we can now introduce a sequence of deterministic maps associated to $(\bar{Y}^m)_{m\ge0}$. First, we set

$$v^0 := \mathrm{y} \ , \tag{5.2.17}$$

recall (5.2.11). Then, given v^{m-1}, we define

$$\begin{aligned}
V_{t,x}^m &:= \Big(\prod_{k\in\mathcal{K}_{t_{i+1}^n - t}} G_{t,x}^m(k) \Big) \Big(\prod_{k\in\bar{\mathcal{K}}_{t_{i+1}^n - t}\setminus\mathcal{K}_{t_{i+1}^n - t}} A_{t,x}^m(k) \Big), \\
G_{t,x}^m(k) &:= \frac{v^m\big(t_{i+1}, X^{x,(k)}_{t_{i+1}^n - t}\big)}{\overline{F}(t_{i+1}^n - t - T_{k-})}, \\
A_{t,x}^m(k) &:= \frac{\sum_{j=1}^{j_\circ} a_{j,\xi_k}(X^{x,(k)}_{T_k}) \varphi_j(v^{m-1}(t + T_k, X^{x,(k)}_{T_k}))}{p_{\xi_k}\, \rho(\delta_k)}, \quad \forall (t, x) \in [t_i^n, t_{i+1}^n) \times \mathbf{X}.
\end{aligned}$$

We finally set, whenever $V^m_{t,x}$ is integrable,

$$\begin{aligned} v^m(t,x) &:= \mathbb{E}\left[V^m_{t,x}\right], \ (t,x) \in (t_i, t_{i+1}) \times \mathbf{X}, \ m \geq 1, \\ v^m(t_i,x) &:= (-M) \vee \mathbb{E}\left[V^m_{t_i,x}\right] \wedge M, \ \ x \in \mathbf{X}, \ m \geq 1. \end{aligned} \tag{5.2.18}$$

Proposition 5.2.1. *For all $m \geq 1$ and $(t,x) \in [0,T] \times \mathbf{X}$, the random variable $V^m_{t,x}$ is integrable. Moreover, $\bar{Y}^m_\cdot = v^m(\cdot, X)$ on $[0,T]$.*

This follows from Proposition 12 proved in the Appendix, which is in spirit of [28]. The main use of this representation result here is that it provides a numerical scheme for the approximation of the component $\bar{Y}$ of (5.2.7), as explained in the next section.

5.2.4 The numerical algorithm

The representation result in Proposition 5.2.1 suggests to use a simple Monte-Carlo estimation of the expectation in the definition of v^m based on the simulation of the corresponding particle system. However, it requires the knowledge of v^{m-1} in the Picard scheme which is used to localize our approximating polynomials. We therefore need to approximate the corresponding (conditional) expectations at each step of the Picard iteration scheme. In practice, we shall replace the expectation operator $\mathbb{E}$ in the definition of v^m by an operator $\hat{\mathbb{E}}$ that can be computed explicitly, see Remark 5.2.2 below.

In order to perform a general (abstract) analysis, let us first recall that

$$\begin{aligned} v^m(t,x) &= \mathbb{E}[V_{t,x}(v^m(t_{i+1},\cdot), v^{m-1}(\cdot)], \ \ t \in (t_i, t_{i+1}) \\ v^m(t_i,x) &= (-M) \vee \mathbb{E}[V_{t_i,x}(v^m(t_{i+1},\cdot), v^{m-1}(\cdot)] \wedge M \end{aligned}$$

where, given two functions $\phi, \phi' : (t_i, t_{i+1}] \times \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned} V_{t,x}(\phi,\phi') &:= \Big(\prod_{k \in \mathcal{K}_{t^n_{i+1}-t}} G_{t,x}(\phi,k)\Big)\Big(\prod_{k \in \bar{\mathcal{K}}_{t^n_{i+1}-t} \setminus \mathcal{K}_{t^n_{i+1}-t}} A_{t,x}(\phi',k)\Big), \\ G_{t,x}(\phi,k) &:= \frac{\phi(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}-t})}{\overline{F}(t^n_{i+1} - t - T_{k-})}, \\ A_{t,x}(\phi',k) &:= \frac{\sum_{j=1}^{j_\circ} a_{j,\xi_k}(X^{x,(k)}_{T_k}) \varphi_j(\phi'(t+T_k, X^{x,(k)}_{T_k})}{p_{\xi_k}\, \rho(\delta_k)}. \end{aligned}$$

Let us then denote by $\mathbf{L}^\infty_{M_h}$ the class of all Borel measurable functions $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ that are bounded by M_h, and let $\mathbf{L}^\infty_{M_h,0} \subset \mathbf{L}^\infty_{M_h}$ be a subspace, generated by a finite number of basis functions. Besides, let us consider a sequence $(U_i)_{i \geq 1}$ of i.i.d. random variables of uniform distribution on $[0,1]$, independent of $(W^k)_{k \in K}$, $(\xi_k)_{k \in K}$, $(\delta_k)_{k \in K}$ and W introduced in (5.2.16). Denote $\hat{\mathcal{F}} := \sigma(U_i, i \geq 1)$.

From now on, we use the notations

$$\|\phi\|_{t^n_i} := \sup_{(t,x) \in [t^n_i, t^n_{i+1}) \times \mathbb{R}^d} |\phi(t,x)| \text{ and } \|\phi\|_\infty := \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\phi(t,x)|$$

for all functions $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$.

Assumption 3. *There exists an operator $\hat{\mathbb{E}}[\hat{V}_{t,x}(\phi,\phi')](\omega)$, defined for all $\phi, \phi' \in \mathbf{L}^\infty_{M_h,0}$, such that $(t,x,\omega) \mapsto \hat{\mathbb{E}}[\hat{V}_{t,x}(\phi,\phi')](\omega)$ is $\mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \hat{\mathcal{F}}$-measurable, and such that the function $(t,x) \in [0,T] \times \mathbb{R}^d \mapsto \hat{\mathbb{E}}[\hat{V}_{t,x}(\phi,\phi')](\omega)$ belongs to $\mathbf{L}^\infty_{M_h,0}$ for every fixed $\omega \in \Omega$. Moreover, one has*

$$\mathcal{E}(\hat{\mathbb{E}}) := \| \sup_{\phi,\phi' \in \mathbf{L}^\infty_{M_h,0}} \mathbb{E}\left[\left|\mathbb{E}\left[V_\cdot(\phi,\phi')\right] - \hat{\mathbb{E}}[\hat{V}_\cdot(\phi,\phi')]\right|\right] \|_\infty < \infty.$$

Then, one can construct a numerical algorithm by first setting $\hat{v}^0 \equiv \mathrm{y}$, $\hat{v}^m(T,\cdot) = g$, $m \geq 1$, and then by defining by induction over $m \geq 1$

$$\hat{v}^m(t,x) := (-M_h) \vee \hat{\mathbb{E}}\left[\hat{V}_{t,x}(\hat{v}^m(t^n_{i+1},\cdot),\hat{v}^{m-1})\right] \wedge M_h,\ t \in (t^n_i, t^n_{i+1}),$$

and

$$\hat{v}^m(t_i,x) := (-M) \vee \hat{\mathbb{E}}\left[\hat{V}_{t_i,x}(\hat{v}^m(t^n_{i+1},\cdot),\hat{v}^{m-1})\right] \wedge M. \tag{5.2.19}$$

In order to analyse the error due to the approximation of the expectation, let us set

$$\bar{q}_{t^n_{i+1}-t} := \#\bar{\mathcal{K}}_{t^n_{i+1}-t}\ ,\ q_{t^n_{i+1}-t} := \#\mathcal{K}_{t^n_{i+1}-t},$$

and denote

$$V^M_h := \Big(\prod_{k\in\mathcal{K}_h} \frac{M}{\bar{F}(h-T_{k-})}\Big)\Big(\prod_{k\in\bar{\mathcal{K}}_h\setminus\mathcal{K}_h} \frac{2C_{\ell_\circ}}{p_{\xi_k}\rho(\delta_k)}\Big).$$

Recall that $h < h_\circ$ that is defined by (5.4.1) in the Appendix.

Lemma 4. *The two constants*

$$M^1_h := \mathbb{E}\big[q_h V^M_h\big] \quad \textit{and} \quad M^2_h := \mathbb{E}\big[\bar{q}_h V^M_h\big]$$

are finite.

Proof. Notice that for any constant $\varepsilon > 0$, there is some constant $C_\varepsilon > 0$ such that $n \leq C_\varepsilon(1+\varepsilon)^n$ for all $n \geq 1$. Then

$$M^1_h \leq C_\varepsilon \mathbb{E}[(1+\varepsilon)^{q_h} V^M_h] \leq C_\varepsilon \mathbb{E}\Big[\prod_{k\in\mathcal{K}_t} \frac{M(1+\varepsilon)}{\bar{F}(t-T_{k-})} \prod_{k\in\bar{\mathcal{K}}_t\setminus\mathcal{K}_t} \frac{2C_{\ell_\circ}(1+\varepsilon)}{p_{\xi_k}\rho(\delta_k)}\Big],$$

where the latter expectation is finite for ε small enough. This follows from the fact that $h < h_\circ$ for $h_\circ$ defined by (5.4.1) and from the same arguments as in Lemma 5 in the Appendix. One can similarly obtain that M^2_h is also finite. □

Proposition 11. *Let Assumption 3 hold true. Then*

$$\|\mathbb{E}\left[|v^m - \hat{v}^m|\right]\|_\infty \leq \mathcal{E}(\hat{\mathbb{E}})\big(1+N_h\big)\frac{(m+N_h)^{N_h}}{N_h!}\Big((2L_\varphi M^2_h) \vee \frac{M^1_h}{M} \vee 1\Big)^{m+N_h}.$$

Before turning to its proof, we comment on the use of this numerical scheme.

Remark 5.2.2. *In practice, the approximation of the expectation operator can be simply constructed by using pure forward simulations of the branching process. Let us explain this first in the case $h_\circ = T$. Given that $\hat{v}^m$ has already been computed, one takes it as a given function, one draws some independent copies of the branching process (independently of $\hat{v}^m$) and computes $\hat{v}^{m+1}(t,x)$ as the Monte-Carlo counterpart of $\mathbb{E}[V_{t,x}(\hat{v}^{m+1}(T,\cdot),\hat{v}^m)]$, and truncate it with the a-priori bound M_h for $(\bar{Y}^m)_{m\geq 1}$. This corresponds to the operator $\hat{\mathbb{E}}[\hat{V}_{t,x}(\hat{v}^{m+1}(T,\cdot),\hat{v}^m)]$. If $h_\circ < T$, one needs to iterate backward over the periods $[t_i,t_{i+1}]$. Obviously one cannot in practice compute the whole map $(t,x) \mapsto \hat{v}^{m+1}(t,x)$ and this requires an additional discretization on a suitable time-space grid. Then, the additional error analysis can be handled for instance by using the continuity property of v^m in Proposition 14 in the Appendix. This is in particular the case if one just computes $\hat{v}^{m+1}(t,x)$ by replacing (t,x) by its projection on a discrete time-space grid.*

Remark 5.2.3. i). *In the classical time discretization schemes of BSDEs, such as those in [15,26,53], one needs to let the time step go to* 0 *to reduce the discretization error. Here, the representation formula in Proposition 5.2.1 has no discretization error related to the BSDE itself (assuming the solution of the previous Picard iteration is known perfectly), we only need to use a fixed discrete time grid* $(t_i)_{0\le i\le N_h}$ *for* $t_i = ih$ *with* h *small enough.* ii). *Let* $A'_j := \{y \in \mathbb{R} \ : \varphi_j(y) = 1\} \subset A_j$ *for* $j \le j_\circ$*, and assume that the* A'_j*'s are disjoint. If the* A'_j *are large enough, we do not need to be very precise on* $\hat v^m$ *to obtain a good approximation of* $\mathbb{E}[V_{t,x}(g, v^m)]$ *by* $\mathbb{E}[V_{t,x}(g, \hat v^m)]$ *for* $t \in [t_{N_h-1}, t_{N_h})$*. One just needs to ensure that* $\hat v^m$ *and* v^m *belong to the same set* A'_j *at the different branching times and at the corresponding* X*-positions. We can therefore use a rather rough time-space grid on this interval (i.e.* $[t_{N_h-1}, t_{N_h}]$*). Further, only a precise value of* $\hat v^m(t_{N_h-1}, \cdot)$ *will be required for the estimation of* $\hat v^{m+1}$ *on* $[t_{N_h-2}, t_{N_h-1})$ *and this is where a fine space grid should be used. Iterating this argument, one can use rather rough time-space grid on each* (t_i, t_{i+1}) *and concentrate on each* t_i *at which a finer space grid is required. This is the main difference with the usual backward Euler schemes of [15, 26, 53] and the forward Picard schemes of [6].*

Proof of Proposition 11. Define

$$\tilde v^m(\cdot) := (-M_h) \vee \mathbb{E}\left[V_\cdot(\hat v^m(t^n_{i+1}, \cdot), \hat v^{m-1})\middle|\hat{\mathcal{F}}\right] \wedge M_h.$$

Then, Lemma 6 below with the inequality $|\varphi| \le 1$ implies that for $(t,x) \in [t_i, t_{i+1}) \times \mathbf{X}$

$$\begin{aligned}
&|\tilde v^m(t,x) - v^m(t,x)| \\
&\le \mathbb{E}\Big[\sum_{k\in\mathcal{K}_{t_{i+1}-t}} \frac{1}{M} V^M_{t^n_{i+1}-t}\big|\hat v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}}) - v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}})\big|\Big|\hat{\mathcal{F}}\Big] \\
&+ \mathbb{E}\Big[\sum_{k\in\bar{\mathcal{K}}_{t_{i+1}-t}\setminus\mathcal{K}_{t_{i+1}-t}} 2L_\varphi V^M_{t^n_{i+1}-t}\big|\hat v^{m-1}(T_k, X^{x,(k)}_{T_k}) - v^{m-1}(T_k, X^{x,(k)}_{T_k})\big|\Big|\hat{\mathcal{F}}\Big].
\end{aligned}$$

Let us compute the expectation of the first term. By denoting by $\bar{\mathcal{F}}$ the σ-field generated by the branching process, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{k\in\mathcal{K}_{t_{i+1}-t}} \frac{1}{M} V^M_{t^n_{i+1}-t}\big|\hat v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}}) - v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}})\big|\Big] \\
= \ &\mathbb{E}\Big[\sum_{k\in\mathcal{K}_{t_{i+1}-t}} \frac{1}{M} V^M_{t^n_{i+1}-t}\mathbb{E}\Big[\big|\hat v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}}) - v^m(t^n_{i+1}, X^{x,(k)}_{t^n_{i+1}})\big|\Big|\bar{\mathcal{F}}\Big]\Big] \\
\le \ &\frac{1}{M}\|\mathbb{E}[|\hat v^m - v^m|]\|_{t_{i+1}}\mathbb{E}\Big[q_{t_{i+1}-t}V^M_{t_{i+1}-t}\Big] \ \le \ \frac{M_h^1}{M}\|\mathbb{E}[|\hat v^m - v^m|]\|_{t_{i+1}}.
\end{aligned}$$

Similarly, for the second term, one has

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{k\in\bar{\mathcal{K}}_{t_{i+1}-t}\setminus\mathcal{K}_{t_{i+1}-t}} 2L_\varphi V^M_{t^n_{i+1}-t}\big|\hat v^{m-1}(T_k, X^{x,(k)}_{T_k}) - v^{m-1}(T_k, X^{x,(k)}_{T_k})\big|\Big] \\
\le \ &2L_\varphi M_h^2\|\mathbb{E}[|\hat v^{m-1} - v^{m-1}|]\|_{t_i}.
\end{aligned}$$

Notice that $\|\mathbb{E}\left[|\tilde v^m - \hat v^m|\right]\|_{t^n_i} \le \mathcal{E}(\hat{\mathbb{E}})$ by Assumption 3. All these lead to

$$\begin{aligned}
\|\mathbb{E}[|\hat v^m - v^m|]\|_{t^n_i} \ \le \ &\mathcal{E}(\hat{\mathbb{E}}) + 2L_\varphi M_h^2\|\mathbb{E}[|\hat v^{m-1} - v^{m-1}|]\|_{t^n_i} \\
&+ \frac{M_h^1}{M}\|\mathbb{E}[|\hat v^m - v^m|]\|_{t^n_{i+1}}.
\end{aligned}$$

We now appeal to Proposition 13 to obtain

$$\begin{aligned}\|\mathbb{E}[|\hat{v}^m - v^m|]\|_{t_i^n} &\leq \mathcal{E}(\hat{\mathbb{E}})\left(\sum_{i=1}^{m} C^i + \sum_{i'=2}^{N_h-i}\left(\sum_{j_1=1}^{m}\cdots\sum_{j_{i'}=1}^{j_{i'-1}} C^{m-j_{i'}}C^{i'-1}\right)\right)\\ &\leq \mathcal{E}(\hat{\mathbb{E}})(1+N_h)\frac{(m+N_h)^{N_h}}{N_h!}C^{m+N_h},\end{aligned}$$

with $C := (2L_\varphi M_h^2) \vee \frac{M_h^1}{M} \vee 1$. □

5.3 Example of application

In this section, we consider a toy example of application. Let us set $\mathbf{X} := [\underline{x}, \bar{x}]$ with $\underline{x} = \pi/8$ and $\bar{x} = 7\pi/8$, and consider the solution X of (5.2.1) with

$$\mu(x) = 0.1 \times (\frac{\pi}{2} - x) \text{ and } \sigma(x) := 0.2 \times (\bar{x} - x)(x - \underline{x}).$$

We then take

$$f(x,y) = \mu(x)\left(\sqrt{1-y^2}\mathbf{1}_{|y|\leq\bar{y}} + \sqrt{1-\bar{y}^2}\mathbf{1}_{|y|>\bar{y}}\right) + \frac{1}{2}\sigma(x)^2 y$$

with $\bar{y} := \cos(\underline{x})$. As can be seen on Figure 5.1, the Lipschitz constant of the driver is rather large. A simple application of Itô's lemma shows that the solution of (5.2.2) with $g = \cos$ is given by $Y = \cos(X)$, which will be used to assess the precision of our estimator.

The driver f is approximated by polynomials of order two weighted by localizing functions. Let $A_j^\upsilon := (y_j - \upsilon, y_{j+1} + \upsilon]$ for $j = 1, \ldots, 5$, with $\upsilon := 10^{-5}$ and

$$y_1 = -y_6 = \infty,\ y_2 = -y_5 = \bar{y},\ y_3 = -y_4 = \cos(x_{\frac{N_X+1}{4}-1}),$$

where $\{x_1, \ldots, x_{N_X}\}$ are equidistant points with $x_1 = \underline{x}$ and $x_{N_X} = \bar{x}$. Then, $f_{\ell_\circ}$ is defined as

$$f_{\ell_\circ}(x,y,y') = \sum_{j=1}^{5}\left(\mu(x)(a_{j0} + a_{j1}y + a_{j2}y^2) + \frac{1}{2}\sigma(x)^2 y\right)\varphi_j(y')$$

where

$$\varphi_j(y') = \begin{cases} \frac{y'-y_j+\upsilon}{2\upsilon} & \text{if} \quad y' \in A_j^\upsilon \cap [y_j - \upsilon, y_j + \upsilon) \\ 1 & \text{if} \quad y' \in A_j^\upsilon \cap [y_j + \upsilon, y_{j+1} - \upsilon] \\ 1 - \frac{y'-y_{j+1}+\upsilon}{2\upsilon} & \text{if} \quad y' \in A_j^\upsilon \cap [y_{j+1} - \upsilon, y_{j+1} + \upsilon) \\ 0 & \text{if} \quad y' \notin A_j^\upsilon \end{cases}$$

and

$$\begin{aligned}(a_{10}, a_{11}, a_{12}) &= (a_{50}, a_{51}, a_{52}) = ((1-\bar{y}^2)^{\frac{1}{2}}, 0, 0)\\ (a_{20}, a_{21}, a_{22}) &= (a_{40}, a_{41}, a_{42})\\ &= ((1-(y_3)^2)^{\frac{1}{2}} - a_{21}y_3, \frac{(1-(y_3)^2)^{\frac{1}{2}} - (1-\cos(x_2)^2)^{\frac{1}{2}}}{y_3 - \cos(x_2)}, 0)\\ (a_{30}, a_{31}, a_{32}) &= (1, 0, -\frac{1-(1-(y_3)^2)^{\frac{1}{2}}}{(y_3)^2}).\end{aligned}$$

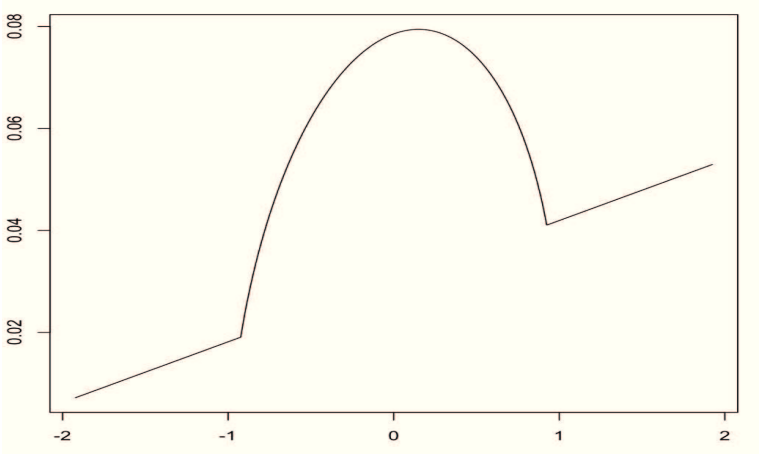

Figure 5.1: Driver $f(\pi/4, \cdot)$.

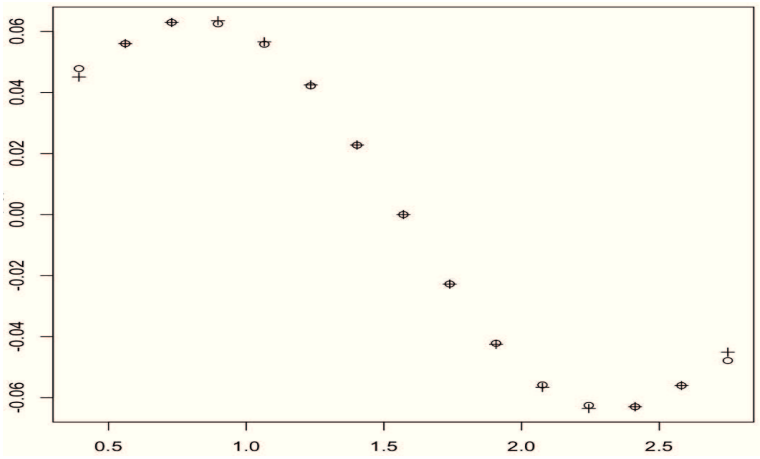

Figure 5.2: Approximation of the driver - Crosses : $f(\cdot, \cos)$. Circles: $f_{\ell_\circ}(\cdot, \cos, \cos)$.

In Figure 5.2, we plot the approximation of $x \mapsto f(x, \cos(x), \cos(x))$ by $x \mapsto f_{\ell_\circ}(x, \cos(x), \cos(x))$, that drives the driver's approximation error, recall (5.2.8). It works quite well except at the boundary points, which should not have a major impact given our mean-reverting dynamics for X.

To construct the approximation operator $\hat{\mathbb{E}}[\hat{V}]$. The time interval $[0, T)$ is divided into N_T intervals $[s_i, s_{i+1})$, $0 \le i \le N_T - 1$, of equal length, with $s_0 = 0$ and $s_{N_T} = T$. The branching density ρ is taken as the exponential law density of parameter $\lambda = 0.6$, but branching times are replaced by the next time in $(s_i)_{i \le N_T}$, if they are less than T. We draw N independent path of the Brownian particules system $(W^{(k),n}; k \in K)_{n \le N}$ (up to T) to which is associated the sequence of numbers of children, branching and birth times $(\zeta_k^n, \delta_k^n, T_k^n; k \in K)_{n \le N}$. The index sets $\mathcal{K}^n$ and $\bar{\mathcal{K}}^n$ are defined correspondingly. Let $\bar{\Phi}[x](\cdot, W)$ be the map that associates to x the Euler scheme of (5.2.1) starting from $X_0 = x$ on the grid $(s_i)_{i \le N_T}$. Then, we set $\bar{X}^{x_l,(k),n} := \bar{\Phi}[x](\cdot, W^{(k),n})$ for each $l \le N_X$. A typical path starting from $\pi/2$ is provided in Figure 5.3.

The simplest algorithm reads as follows. We fix $\mathrm{y}(t, \cdot) = (t/T)\cos$, $\hat{v}^m(T, \cdot) = \cos$, and then set, for $\kappa \ge 1$, $m \ge 0$, $i < N_T/\kappa$ with $s_{i\kappa} \in [t_{i'}, t_{i'+1})$ and $l \le N_X$,

$$\hat{\mathbb{E}}\left[\hat{V}_{s_{i\kappa},x_l}(\hat{v}^m(t_{i'+1}, \cdot), \hat{v}^{m-1})\right] := \frac{1}{N}\sum_{n=1}^{N} \hat{V}^n_{s_{i\kappa},x_l}(\hat{v}^m(t_{i'+1}, \cdot), \hat{v}^{m-1})$$

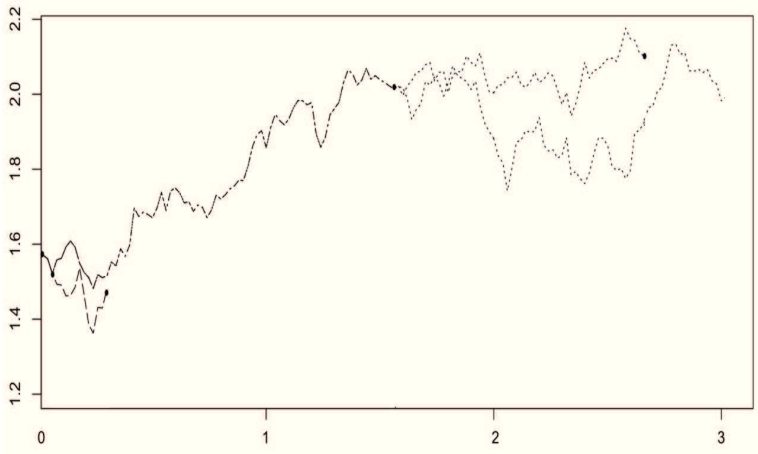

Figure 5.3: A typical simulated path of the branching diffusion starting from $\pi/2$ on $[0,3]$. Bullets denote branching or killing times.

where

$$\hat{V}^n_{s_{i\kappa},x_l}(\phi,\phi') := \Big(\prod_{k\in\mathcal{K}^n_{t_{i'+1}-s_{i\kappa}}} \hat{G}^m_{s_{i\kappa},x_l}(\phi,k)\Big)\Big(\prod_{k\in\bar{\mathcal{K}}^n_{t_{i'+1}-s_{i\kappa}}\setminus\mathcal{K}^n_{t_{i'+1}-s_{i\kappa}}} \hat{A}^n_{s_{i\kappa},x_l}(\phi',k)\Big),$$

$$\hat{G}^n_{s_{i\kappa},x_l}(\phi,k) := \frac{\phi(\bar{X}^{x_l,(k),n}_{t_{i'+1}-s_{i\kappa}})}{\overline{F}(t_{i'+1}-s_{i\kappa}-T^n_{k-})},$$

$$\hat{A}^n_{s_{i\kappa},x_l}(\phi',k) := \frac{\sum_{j=1}^{j_\circ} a_{j,\xi^n_k}(\bar{X}^{x_l,(k),n}_{T^n_k})\varphi_j(\phi'(s_{i\kappa}+T^n_k,\bar{X}^{x_l,(k),n}_{T_k}))}{p_{\xi^n_k}\,\rho(\delta^n_k)}.$$

For $m \geq 0$, $(\hat{v}^m(s_{i\kappa},\cdot))_{i<N_T/\kappa}$ is extended to $\mathbf{X}$ by a simple barycentric linearization, and $(\hat{v}^m(\cdot,x))_{x\in\mathbf{X}}$ is extended to $[0,T]$ by setting $\hat{v}^m(t,x) := \hat{v}^m(s_{(i+1)\kappa},x)$ if $t \in (s_{i\kappa},s_{(i+1)\kappa}]$. In particular, each function $\hat{v}^m$ is computed on a time grid that is κ times rougher than the one used to construct the Euler scheme of the branching system.

In practice, we proceed slightly differently. In what follows, we drop the index m for more clarity as we shall somehow mix the Picard iterations. We set $\hat{v}(s_j,\cdot) = \hat{v}(T,\cdot)$ for $j > N_T - \kappa$. Then, one can compute $\hat{v}(s_{N_T-\kappa},\cdot)$ as above, based on y, and set $\hat{v}(s_j,\cdot) = \hat{v}(s_{N_T-\kappa},\cdot)$ for $N_T-\kappa \geq j > N_T-2\kappa$. This allows to compute immediately, $\hat{v}(s_{N_T-2\kappa},\cdot)$, since it only requires the knowledge of $\hat{v}(s_j,\cdot)$ for $N_T \geq j > N_T-2\kappa$. We then set $\hat{v}(s_j,\cdot) = \hat{v}(s_{N_T-2\kappa},\cdot)$ for $N_T-2\kappa \geq j > N_T-3\kappa$, from which we can compute $\hat{v}(s_{N_T-3\kappa},\cdot)$. We go on this way. The estimation $\hat{v}(s_j,\cdot)$ corresponds to a unique Picard iteration for $N_T \geq j > N_T-2\kappa$. But, around T, we expect to be very precise with only one, as y is based on the terminal condition. The estimation $\hat{v}(s_j,\cdot)$ corresponds to a mix between a unique and two Picard iterations for $N_T-\kappa \geq j > N_T-2\kappa$, and so on. We therefore increase automatically the number of Picard iterations when we go further from the terminal horizon.

In Figure 5.4, we plot the solution $x \mapsto \cos(x)$ and the confidence interval obtained by computing the mean estimated value over 100 independent estimations $\pm$ twice the standard deviation computed over these 100 estimations, for $N = 10^3$, $N_X = 31$, $N_T = 50$, $\kappa = 10$ and $N_h = 1$. As can be seen, the algorithm is already quite efficient with only a rather small number of simulations. Figure 5.5 provides the same curves in the case $N = 10.10^3$.

In Figure 5.6 and Figure 5.7, we consider the case $T = 2$ with $N_h = 2$, $N_T = 140$ and $N_X = 47$. We use $N = 10^3$ and $N = 10.10^3$ simulations, respectively.

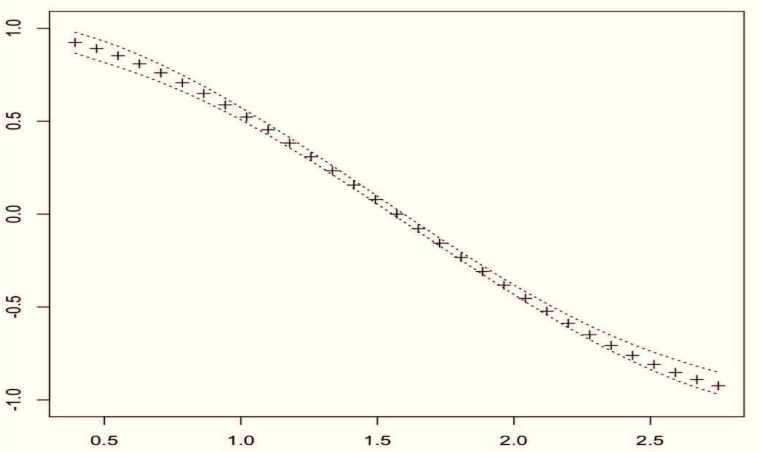

Figure 5.4: $T = N_h = 1$ with $N = 10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values.

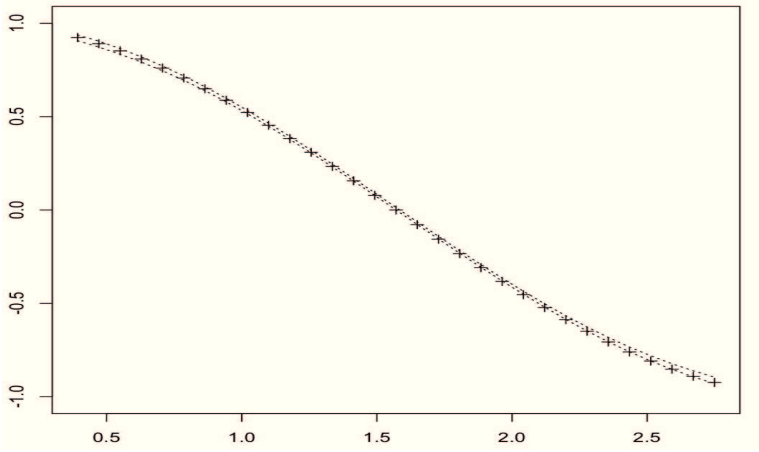

Figure 5.5: $T = N_h = 1$ with $N = 10.10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values.

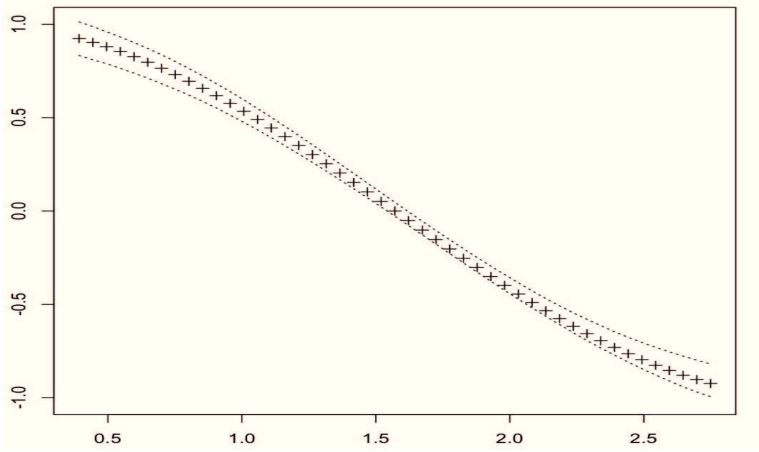

Figure 5.6: $T = N_h = 2$ with $N = 10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values.

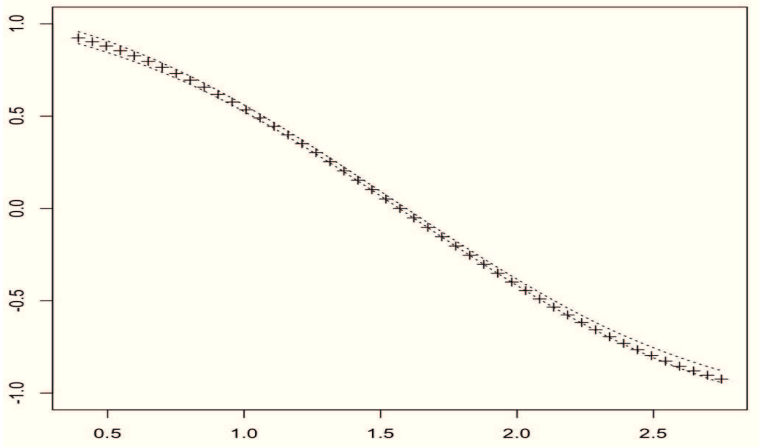

Figure 5.7: $T = N_h = 2$ with $N = 10.10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values.

5.4 Appendix

5.4.1 Technical lemmas

Lemma 5. *The ordinary differential equation* $\eta'(t) = \sum_{\ell=0}^{\ell_\circ} 2C_{\ell_\circ}\eta(t)^\ell$ *with initial condition* $\eta(0) = M > 0$ *has a unique solution on* $[0, h_\circ]$ *for*

$$h_\circ := \frac{(\ell_\circ - 1)(1-M)_+ + (1 \vee M)^{-(\ell_\circ - 1)}}{(\ell_\circ + 1)(\ell_\circ - 1)2C_{\ell_\circ}}. \tag{5.4.1}$$

Moreover, it is bounded on $[0, h_\circ]$ *by*

$$M_h := \max\Big(1, \big((1 \vee M)^{1-\ell_\circ} + (\ell_\circ - 1)(1-M)^+ - h_\circ \ell_\circ(\ell_\circ - 1)2C_{\ell_\circ}\big)^{(1-\ell_\circ)^{-1}}\Big). \tag{5.4.2}$$

Consequently, one has, for all $t \in [0, h_\circ]$,

$$\mathbb{E}\Big[\Big(\prod_{k\in\mathcal{K}_t} \frac{M}{\bar{F}(t - T_{k-})}\Big)\Big(\prod_{k\in\bar{\mathcal{K}}_t\setminus\mathcal{K}_t} \frac{2C_{\ell_\circ}}{p_{\xi_k}\rho(\delta_k)}\Big)\Big] \;\le\; \eta(t) \;\le\; M_h. \tag{5.4.3}$$

Proof. i) We first claim that

$$\int_M^{M_h} \frac{dy}{2C_{\ell_\circ}(1 + y + \cdots + y^{\ell_\circ})} \ge h_\circ. \tag{5.4.4}$$

Then, for every $t \in [0, h_\circ]$, there is some constant $M(t) \le M_h < \infty$ such that

$$\int_M^{M(t)} \frac{dy}{2C_{\ell_\circ}(1 + y + \cdots + y^{\ell_\circ})} \;=\; t = \int_0^t ds.$$

This indicates that $(M(t))_{t\in[0,h_\circ]}$ is a bounded solution (hence the unique one) of $\eta'(t) = \sum_{\ell=0}^{\ell_\circ} 2C_{\ell_\circ}\eta(t)^\ell$ of initial condition $\eta(0) = M > 0$. In particular, it is bounded by M_h.
ii) Let us now prove (5.4.4). Notice that $y^k \le 1 \vee y^{\ell_\circ}$ for any $y \ge 0$ and $k = 0, \cdots, \ell_\circ$. Then, it is enough to prove that

$$\int_M^{M_h} \Big(1 \wedge \frac{1}{y^{\ell_\circ}}\Big) dy \ge h_\circ(\ell_\circ + 1)2C_{\ell_\circ}. \tag{5.4.5}$$

By direct computation, the l.h.s. of (5.4.5) equals

$$(M_h - M)\mathbf{1}_{\{M_h \le 1\}} + \Big((1-M)^+ + \frac{1}{\ell_\circ - 1}\big((1 \vee M)^{1-\ell_\circ} - M_h^{1-\ell_\circ}\big)\Big)\mathbf{1}_{\{M_h > 1\}}.$$

When $h_\circ$ satisfies (5.4.1), it is easy to check that (5.4.5) holds true.
iii) We now prove (5.4.3). Recall that $\bar{\mathcal{K}}_t^n$ denotes the collection of all particles in $\bar{\mathcal{K}}_t$ of generation n. Set

$$\chi_t^n := \Big(\prod_{k \in \cup_{j=1}^n \mathcal{K}_t^j} \frac{M}{\bar{F}(t - T_{k-})}\Big)\Big(\prod_{k \in \cup_{j=1}^n (\bar{\mathcal{K}}_t^j \setminus \mathcal{K}_t^j)} \frac{2C_{\ell_\circ}}{p_{\xi_k}\rho(\delta_k)}\Big)\Big(\prod_{k \in \bar{\mathcal{K}}_t^{n+1}} \eta(t - T_{k-})\Big).$$

Since $\bar{\mathcal{K}}_t^n$ has only finite number of particles, the random variable χ_t^n is uniformly bounded. Then by exactly the same arguments as in (5.4.6) and (5.4.7) below, and by repeating this argument over n, one has

$$\eta(t) = M + \int_0^t \sum_{\ell=0}^{\ell_\circ} 2C_{\ell_\circ}\eta(s)^\ell ds = \mathbb{E}\big[\chi_t^1\big] = \mathbb{E}\big[\chi_t^n\big],\ \forall n \ge 1.$$

It follows by Fatou Lemma that

$$\mathbb{E}\Big[\Big(\prod_{k \in \mathcal{K}_t} \frac{M}{\bar{F}(t - T_{k-})}\Big)\Big(\prod_{k \in \bar{\mathcal{K}}_t \setminus \mathcal{K}_t} \frac{2C_{\ell_\circ}}{p_{\xi_k}\rho(\delta_k)}\Big)\Big] = \mathbb{E}\big[\lim_{n\to\infty} \chi_t^n\big] \ \le\ \lim_{n\to\infty} \mathbb{E}[\chi_t^n] = \eta(t).$$

□

For completeness, we provide here the proof the representation formula of Proposition 5.2.1 and of the technical lemma that was used in the proof of Proposition 11.

Proposition 12. *The representation formula of Proposition 5.2.1 holds.*

Proof. We only provide the proof on $[t_{N_h - 1}, T]$, the general result is obtained by induction. The representation holds true by construction for $m = 0$. Let us now fix $m \ge 1$.

First notice that Lemma 5 shows that the random variable $V_{t,x}^m$ is integrable. Next, Set $(1)+ := \{(1,j), j \le \ell_\circ\} \cap \bar{\mathcal{K}}_T$ and define $\mathcal{K}_t(1) := \mathcal{K}_t \cap (1)+$ and $\bar{\mathcal{K}}_t(1) := \bar{\mathcal{K}}_t \cap (1)+$. For ease of notations, we write $X^x := X^{x,((1))}$. Then, for all $(t,x) \in [t_{N_h-1}, T] \times \mathbb{R}^d$,

$$\begin{aligned}\mathbb{E}[V_{t,x}^m] = \mathbb{E}\left[\frac{g(X_{T-t}^x)}{\overline{F}(T-t)}\mathbf{1}_{\{T_{(1)} \ge T-t\}}\right] \\ + \mathbb{E}\left[\mathbf{1}_{\{T_{(1)} < T-t\}} \frac{\sum_{j=1}^{j_\circ} a_{j,\xi_{(1)}}(X_{T_{(1)}}^x)\varphi_j(v^{m-1}(t + T_{(1)}, X_{T_{(1)}}^x))}{p_{\xi_{(1)}}\rho(\delta_{(1)})} R_{t,x}^m\right]\end{aligned}$$

where

$$R_{t,x}^m := \Big(\prod_{k \in \mathcal{K}_{T-t}(1)} G_{t,x}(k)\Big)\Big(\prod_{k \in \bar{\mathcal{K}}_{T-t}(1) \setminus \mathcal{K}_{T-t}(1)} A_{t,x}^m(k)\Big)$$

satisfies

$$\mathbb{E}[R_{t,x}^m | \mathcal{F}_{T_{(1)}}] = \prod_{k \in (1)+} v^m\big(T_{(1)}, X_{T_{(1)}}^{t,x}\big) = \big[v^m(T_{(1)}, X_{T_{(1)}}^x)\big]^{\xi_{(1)}},$$

by (5.2.16). On the other hand, (5.2.15) and (5.2.16) imply

$$\mathbb{E}\left[\frac{g(X_{T-t}^x)}{\overline{F}(T-t)}\mathbf{1}_{\{T_{(1)} \ge T-t\}}\right] = \mathbb{E}[g(X_{T-t}^x)] \tag{5.4.6}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\{T_{(1)}<T-t\}}\frac{\sum_{j=1}^{j_\circ} a_{j,\xi_{(1)}}(X^x_{T_{(1)}})\varphi_j(v^{1,m-1}(t+T_{(1)},X^x_{T_{(1)}}))}{p_{\xi_{(1)}}\,\rho(\delta_{(1)})}[v^m(T_{(1)},X^x_{T_{(1)}})]^{\xi_{(1)}}\right]\\
&=\mathbb{E}\left[\int_0^{T-t}\frac{\sum_{j=1}^{j_\circ} a_{j,\xi_{(1)}}(X^x_s)\varphi_j(v^{1,m-1}(t+s,X^x_s))}{p_{\xi_{(1)}}}[v^m(s,X^x_s)]^{\xi_{(1)}}ds\right]\\
&=\mathbb{E}\left[\int_0^{T-t}\sum_{j=1}^{j_\circ}\sum_{\ell\le\ell_\circ} a_{j,\ell}(X^x_s)\varphi_j(v^{1,m-1}(t+s,X^x_s))[v^m(s,X^x_s)]^{\ell}ds\right]\\
&=\mathbb{E}\left[\int_0^{T-t} f_{\ell_\circ}(X^x_s,v^m(t+s,X^x_s),v^{m-1}(t+s,X^x_s))ds\right]. \qquad (5.4.7)
\end{aligned}$$

Combining the above implies that

$$v^m(t,X_t)=\mathbb{E}\left[g(X_T)+\int_t^T f_{\ell_\circ}(X_s,v^m(s,X_s),v^{m-1}(s,X_s))ds\Big|\mathcal{F}_t\right],$$

and the required result follows by induction. □

Lemma 6. *Let $(x^i,y^i)_{i\le I}$ be a sequence of real numbers. Then,*

$$\left|\prod_{i=1}^I x^i-\prod_{i=1}^I y^i\right|\le\sum_{i\in I}\Big(|x^i-y^i|\prod_{j\neq i}\max(|x^j|,|y^j|)\Big).$$

Proof. It suffices to observe that

$$\prod_{i=1}^I x^i-\prod_{i=1}^I y^i \;=\; (x^1-y^1)\prod_{i=2}^I x^i+y^1\Big(\prod_{i=2}^I x^i-\prod_{i=2}^I y^i\Big),$$

and proceed by induction arguments. □

Proposition 13. *Let $c_1,c_2,c_3\ge 0$, and let $(u^i_m)_{m\ge0,i\ge0}$ be a sequence such that*

$$u^i_m\le c_1u^i_{m-1}+c_2u^{i+1}_m+c_3\;\; \textit{for } m\ge1, i<N_h.$$

Then

$$\begin{aligned}
u^i_m\le& c_1^m u^i_0+\sum_{i'=1}^{N_h-i}\left(\sum_{j_1=1}^{m}\sum_{j_2=1}^{j_1}\cdots\sum_{j_{i'}=1}^{j_{i'-1}}c_1^m c_2^{i'}u_0^{i+i'}\right)\\
&+c_3\left(\sum_{i=1}^m c^i+\sum_{i'=2}^{N_h-i}\Big(\sum_{j_1=1}^{m}\sum_{j_2=1}^{j_1}\cdots\sum_{j_{i'}=1}^{j_{i'-1}}c_1^{m-j_{i'}}c_2^{i'-1}\Big)\right).
\end{aligned}$$

Proof. We have

$$u^i_m\le(c_1)^m u^i_0+\sum_{j=1}^m(c_1)^{m-j}(c_2u^{i+1}_m+c_3).$$

The required result then follows from a simple induction. □

5.4.2 More on the error analysis for the abstract numerical approximation

The regression error will depend on the regularity of v^m. Here we prove that $v^m(t,x)$ is Hölder in t and Lipschitz in x under additional conditions, and provide some estimates on the corresponding coefficients. Given $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, denote

$$[\phi]_{t_i^n} := \sup_{(t,x)\neq(t',x')\in[t_i^n,t_{i+1}^n]\times\mathbf{X}} \frac{|\phi(t,x)-\phi(t',x')|}{|t-t'|^{\frac{1}{2}}+|x-x'|}.$$

Since (μ,σ) is assumed to be Lipschitz, it is clear that there exists $L_X > 0$ such that for all $(t,x),(t',x') \in [0,T]\times\mathbf{X}$,

$$\|X_t^x - X_{t'}^{x'}\|_{\mathbf{L}^2} \le L_X\Big(\sqrt{|t'-t|}+|x'-x|\Big). \tag{5.4.8}$$

Proposition 14. *Suppose that $x \mapsto g(x)$ and $x \mapsto f_{\ell_\circ}(x,y,y')$ are uniformly Lipschitz with Lipschitz constants L_g and L_f. Let β and $\lambda_1,\lambda_2 > 0$ such that $\frac{L_2}{\lambda_2^2}T < 1$ and $\beta \ge 2L_1 + L_f\lambda_1^2 + L_2\lambda_2^2$, then for all $m \ge 1$ and $i \le N_h$,*

$$\begin{aligned}[v^m]_{t_i^n} \le L_v := {} & (1+L_X)L_X\sqrt{\big(L_g^2+\frac{L_f}{\beta\lambda_1^2}\big)Te^{\beta T}/\big(1-\frac{L_2}{\lambda_2^2}T\big)} \\ & + 2(1+\ell_\circ)C_\ell(1\vee M_h^{\ell_\circ})\sqrt{h_\circ}.\end{aligned}$$

Proof. For ease of notations, we provide the proof for $t = 0$ only.
i) Let $x_1,x_2 \in \mathbb{R}^d$ and $Y^{m,1} := v^m(\cdot,X^{x_1})$, $Y^{m,2} := v^m(\cdot,X^{x_2})$,and denote $\Delta Y^m := Y^{m,1} - Y^{m,2}$, $\Delta X := X^{x_1} - X^{x_2}$, where X^{x_1} (resp. X^{x_2}) denotes the solution of SDE (5.2.1) with initial condition $X_0 = x_1$ (resp. $X_0 = x_2$). Using the same arguments as in the proof of Theorem 6, it follows that, for any $\beta \ge 2L_1 + L_f\lambda_1^2 + L_2\lambda_2^2$, one has

$$\begin{aligned}\mathbb{E}[e^{\beta t}(\Delta Y_t^{m+1})^2] \le {} & \mathbb{E}[e^{\beta T}(\Delta Y_T^{m+1})^2] + \frac{L_f}{\lambda_1^2}\mathbb{E}\Big[\int_t^T e^{\beta s}|\Delta X_s|^2 ds\Big] \\ & + \frac{L_2}{\lambda_2^2}\mathbb{E}\Big[\int_t^T e^{\beta s}(\Delta Y_s^m)^2 ds\Big]\end{aligned} \tag{5.4.9}$$

and then

$$\begin{aligned}\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta Y_t^{m+1})^2 dt\Big] &\le T\mathbb{E}[e^{\beta T}(\Delta Y_T^{m+1})^2] + T\frac{L_f}{\lambda_1^2}\mathbb{E}\Big[\int_0^T e^{\beta s}|\Delta X_s|^2 ds\Big] \\ &\quad + T\frac{L_2}{\lambda_2^2}\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta Y_t^m)^2 dt\Big] \\ &\le Te^{\beta T}\Big(L_g^2+\frac{L_f}{\beta\lambda_1^2}\Big)L_X^2|x_1-x_2|^2 \\ &\quad + T\frac{L_2}{\lambda_2^2}\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta Y_t^m)^2 dt\Big].\end{aligned}$$

Since $\frac{L_2}{\lambda_2^2}T < 1$, this induces that

$$\mathbb{E}\Big[\int_0^T e^{\beta t}(\Delta Y_t^{m+1})^2 dt\Big] \le \frac{Te^{\beta T}\big(L_g^2+\frac{L_f}{\beta\lambda_1^2}\big)L_X^2|x_1-x_2|^2}{1-\frac{L_2}{\lambda_2^2}T}.$$

Plugging the above estimates into (5.4.9), it follows that

$$(\Delta Y_0^m)^2 \;\leq\; \hat{L}_v^2|x_1 - x_2|^2, \text{ with } \hat{L}_v^2 := \frac{\big(L_g^2 + \frac{L_f}{\beta\lambda_1^2}\big)L_X^2 T e^{\beta T}}{1 - \frac{L_2}{\lambda_2^2}T}.$$

ii) For the Hölder property of v^m, it is enough to notice that for $t \leq h_\circ$,

$$\begin{aligned} |v^m(0,x) - v^m(t,x)| &\leq \mathbb{E}\Big[|v^m(t, X_t^x) - v^m(t,x)| + \int_0^t |f(X_s^x, Y_s^m, Y_s^{m-1})|ds\Big] \\ &\leq \hat{L}_v L_X \sqrt{t} + 2(1+\ell_\circ)C_\ell(1 \vee M_h^{\ell_\circ})t, \end{aligned}$$

where the last inequality follows from the Lipschitz property of v^m in x and the fact that Y^m is uniformly bounded by M_h. We hence conclude the proof. □

List of Figures

4.1 Left: Super-hedging price of the Butterfly option. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$; dotted line: $\lambda = 0$, $\bar{\gamma} = +\infty$. Right: Difference with the price associated to $\lambda = 0$, $\bar{\gamma} = +\infty$. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$. 71

4.2 Left: Super-hedging price of the Call Spread option. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$; dotted line: $\lambda = 0$, $\bar{\gamma} = +\infty$. Right: Difference with the price associated to $\lambda = 0$, $\bar{\gamma} = +\infty$. Dashed line: $\lambda = 0.5$, $\bar{\gamma} = 1.75$; solid line: $\lambda = 0$, $\bar{\gamma} = 1.75$. 71

5.1 Driver $f(\pi/4, \cdot)$. 83

5.2 Approximation of the driver - Crosses : $f(\cdot, \cos)$. Circles: $f_{\ell_\circ}(\cdot, \cos, \cos)$. 83

5.3 A typical simulated path of the branching diffusion starting from $\pi/2$ on $[0, 3]$. Bullets denote branching or killing times. 84

5.4 $T = N_h = 1$ with $N = 10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values. 85

5.5 $T = N_h = 1$ with $N = 10.10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values. 85

5.6 $T = N_h = 2$ with $N = 10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values. 85

5.7 $T = N_h = 2$ with $N = 10.10^3$ - Crosses: cos function. Dotted lines: mean of estimations $\pm$ 2 standard deviation computed over the estimated values. 86

Bibliography

[1] F. Abergel and G. Loeper. Pricing and hedging contingent claims with liquidity costs and market impact. SSRN.

[2] Vlad Bally and Gilles Pages. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic processes and their applications*, 106(1):1–40, 2003.

[3] G. Barles. *Solution de viscosités des équations d'Hamilton Jacobi*, volume 17 of *Mathématiques et Applications*. Springer Verlag, 1994.

[4] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*. Springer-Verlag, Berlin, 1994.

[5] Guy Barles and Panagiotis E Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.

[6] Christian Bender and Robert Denk. A forward scheme for backward sdes. *Stochastic processes and their applications*, 117(12):1793–1812, 2007.

[7] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control. The Discrete-Time Case*. Academic Press, New York, 1978.

[8] V. I. Bogachev and M. A. S. Ruas. *Measure theory*, volume 1. Springer, 2007.

[9] Mireille Bossy, Nicolas Champagnat, Hélène Leman, Sylvain Maire, Laurent Violeau, and Mariette Yvinec. Monte carlo methods for linear and non-linear poisson-boltzmann equation. *ESAIM: Proceedings and Surveys*, 48:420–446, 2015.

[10] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.

[11] B Bouchard, G Loeper, and Y Zou. Hedging of covered options with linear market impact and gamma constraint. *arXiv preprint arXiv:1512.07087*, 2015. (en révision mineur pour SIAM SICON).

[12] Bruno Bouchard, Grégoire Loeper, and Yiyi Zou. Almost-sure hedging with permanent price impact. *Finance and Stochastics*, 20(3):741–771, 2016.

[13] Bruno Bouchard and Marcel Nutz. Stochastic target games and dynamic programming via regularized viscosity solutions. *arXiv preprint arXiv:1307.5606*, 2013.

[14] Bruno Bouchard, Xiaolu Tan, and Yiyi Zou. Numerical approximation of bsdes using local polynomial drivers and branching processes. *arXiv preprint arXiv:1612.06790*, 2016. (soumis).

[15] Bruno Bouchard and Nizar Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[16] Bruno Bouchard and Xavier Warin. Monte-carlo valuation of american options: facts and new algorithms to improve existing methods. In *Numerical methods in finance*, pages 215–255. Springer, 2012.

[17] U. Cetin, R. A. Jarrow, and P. Protter. Liquidity risk and arbitrage pricing theory. *Finance and stochastics*, 8(3):311–341, 2004.

[18] U. Çetin, H. M. Soner, and N. Touzi. Option hedging for small investors under liquidity costs. *Finance Stoch.*, 14(3):317–341, 2010.

[19] Patrick Cheridito, H Mete Soner, and Nizar Touzi. The multi-dimensional super-replication problem under gamma constraints. In *Annales de l'IHP Analyse non linéaire*, volume 22, pages 633–666, 2005.

[20] J. Claisse, D. Talay, and x. Tan. A note on solutions to controlled martingale problems and their conditioning. 2014.

[21] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[22] H. Crauel. *Random probability measures on Polish spaces*, volume 11. CRC Press, 2003.

[23] Jakša Cvitanić, Huyen Pham, and Nizar Touzi. Super-replication in stochastic volatility models under portfolio constraints. *Journal of Applied Probability*, 36(02):523–545, 1999.

[24] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[25] R. Frey. Perfect option hedging for a large trader. *Finance and Stochastics*, 2:115–141, 1998.

[26] Emmanuel Gobet, Jean-Philippe Lemor, Xavier Warin, et al. A regression-based monte carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005.

[27] Pierre Henry-Labordère. Cutting cva's complexity. *Risk*, 25(7):67, 2012.

[28] Pierre Henry-Labordere, Nadia Oudjane, Xiaolu Tan, Nizar Touzi, and Xavier Warin. Branching diffusion representation of semilinear pdes and monte carlo approximation. *arXiv preprint arXiv:1603.01727*, 2016.

[29] Pierre Henry-Labordere, Xiaolu Tan, and Nizar Touzi. A numerical algorithm for a class of bsdes via the branching process. *Stochastic Processes and their Applications*, 124(2):1112–1140, 2014.

[30] Hitoshi Ishii. On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions. *Funkcial. Ekvac*, 38(1):101–120, 1995.

[31] Robert Jensen. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Archive for Rational Mechanics and Analysis*, 101(1):1–27, 1988.

[32] NV Krylov. On the rate of convergence of finite-difference approximations for bellmans equations with variable coefficients. *Probability theory and related fields*, 117(1):1–16, 2000.

[33] Thomas G Kurtz and Philip E Protter. Weak convergence of stochastic integrals and differential equations. In *Probabilistic models for nonlinear partial differential equations*, pages 1–41. Springer, 1996.

[34] O. A. Ladyzhenskaia, V. Solonnikov, and N. N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[35] G. M. Lieberman. *Second order parabolic differential equations*. World scientific, 1996.

[36] H. Liu and J. M. Yong. Option pricing with an illiquid underlying asset market. *Journal of Economic Dynamics and Control*, 29:2125–2156, 2005.

[37] G. Loeper. Option pricing with market impact and non-linear black and scholes pde's. ArxivN.

[38] Grégoire Loeper. Solution of a fully non-linear black and scholes equation coming from a linear market impact mode. *SSRN*.

[39] Grégoire Loeper. Option pricing with market impact and non-linear black and scholes pdes. 2013.

[40] Henry P McKean. Application of brownian motion to the equation of kolmogorov-petrovskii-piskunov. *Communications on Pure and Applied Mathematics*, 28(3):323–331, 1975.

[41] H. M. Soner P. Cheridito and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de l'Institut Henri Poincaré, Série C: Analyse Non-Linéaire*, 22:633–666, 2005.

[42] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[43] Abdujabor Rasulov, Gulnora Raimova, and Michael Mascagni. Monte carlo solution of cauchy problem for a nonlinear parabolic equation. *Mathematics and Computers in Simulation*, 80(6):1118–1123, 2010.

[44] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 1999.

[45] P.J. Schönbucher and P. Wilmott. The feedback effects of hedging in illiquid markets. *SIAM Journal on Applied Mathematics*, 61:232–272.

[46] K.R. Sircar and G. Papanicolaou. Generalized black-scholes models accounting for increased market volatility from hedging strategies. *Applied Mathematical Finance*, 5(1):45–82, 1998.

[47] Anatolii Volodimirovich Skorokhod. Branching diffusion processes. *Theory of Probability & Its Applications*, 9(3):445–449, 1964.

[48] H. M. Soner and N. Touzi. Superreplication under gamma constraints. *SIAM J. Control Optim.*, 39:73–96, 2000.

[49] H. M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4(3):201–236, 2002.

[50] H. M. Soner and N. Touzi. Hedging under gamma constraints by optimal stopping and face-lifting. *Mathematical finance*, 17:59–80, 2007.

[51] M. Soner and N. Touzi. The dynamic programming equation for second order stochastic target problems. *SIAM Journal on Control and Optimization*, pages 2344–2365.

[52] Shinzo Watanabe et al. On the branching process for brownian particles with an absorbing boundary. *Journal of Mathematics of Kyoto University*, 4(2):385–398, 1965.

[53] J Zhang. A numerical scheme for backward stochastic differential equations: approximation by step processes. *Annals of Applied Probability,*, 14(1):459–488, 2004.

Résumé

La théorie classique de la valorisation des produits dérivés se repose sur l'absence de coûts de transaction et une liquidité infinie. Ces hypothèses sont toutefois ne plus satisfaites dans le marché réel, en particulier lorsque la transaction est grande et les actifs non-liquides. Dans ce marché imparfait, on parle du prix de sur-réplication puisque la couverture parfaite est devenue parfois infaisable.

La première partie de cette thèse se concentre sur la proposition d'un modèle qui intègre à la fois le coût de transaction et l'impact sur le prix du sous-jacent. Nous commençons par déduire la dynamique de l'actif en temps continu en tant que la limite de la dynamique en temps discret. Sous la contrainte d'une position nulle sur l'actif au début et à la maturité, nous obtenons une équation quasi-linéaire pour le prix du dérivé, au sens de viscosité. Nous offrons la stratégie de couverture parfaite lorsque l'équation admet une solution régulière. Quant à la couverture d'une option européenne "covered" sous la contrainte gamma, le principe de la programme dynamique utilisé précédemment n'est plus valide. En suivant les techniques du cible stochastique et de l'équation différentielle partielle, nous démontrons que le prix de la sur-réplication est devenue une solution de viscosité d'une équation non linéaire de type parabolique. Nous construisons également la stratégie ε-optimale, et proposons un schéma numérique.

La deuxième partie de cette thèse est consacrée aux études sur un nouveau schéma numérique d'EDSR, basé sur le processus de branchement. Nous rapprochons tout d'abord le générateur Lipschitzien par une suite de polynômes locaux, puis appliquons l'itération de Picard. Chaque itération de Picard peut être représentée en termes de processus de branchement. Nous démontrons la convergence de notre schéma sur l'horizon temporel infini.

Mots Clés

Réplication dynamique, Impact permanent, Cible stochastique, EDSR, Méthodes Monte-Carlo, Processus de branchement

Abstract

Classical derivatives pricing theory assumes frictionless market and infinite liquidity. These assumptions are however easily violated in real market, especially for large trades and illiquid assets. In this imperfect market, one has to consider the super-replication price as perfect hedging becomes infeasible sometimes.

The first part of this dissertation focuses on proposing a model incorporating both liquidity cost and price impact. We start by deriving continuous time trading dynamics as the limit of discrete rebalancing policies. Under the constraint of holding zero underlying stock at the inception and the maturity, we obtain a quasi-linear pricing equation in the viscosity sense. A perfect hedging strategy is provided as soon as the equation admits a smooth solution. When it comes to hedging a covered European option under gamma constraint, the dynamic programming principle employed previously is no longer valid. Using stochastic target and partial differential equation smoothing techniques, we prove the super-replication price now becomes the viscosity solution of a fully non-linear parabolic equation. We also show how ε-optimal strategies can be constructed, and propose a numerical resolution scheme.

The second part is dedicated to the numerical resolution of the Backward Stochastic Differential Equation (BSDE). We propose a purely forward numerical scheme, which first approximates an arbitrary Lipschitz driver by local polynomials and then applies the Picard iteration to converge to the original solution. Each Picard iteration can be represented in terms of branching diffusion systems, thus avoiding the usual estimation of conditional expectation. We also prove the convergence on an unlimited time horizon. Numerical simulation is also provided to illustrate the performance of the algorithm.

Keywords

Dynamic hedging, Price impact, Stochastic target, BSDE, Monte-Carlo methods, Branching process.