



À MA MUSE DES MATHÉMATIQUES :
LOYALE, MALGRÉ MON ESPRIT BROUILLON.

Dédicaces

Par reconnaissance et gratitude, je dédie ce travail à :

Ma famille pour son amour immense ;
Que ceci soit un témoignage de mes grands respects et remerciements.

OUSSAMA

Remerciements

Je tiens à remercier tous ceux qui ont contribué de près ou de loin au bon achèvement de ce stage.

Tout d'abord, j'adresse mes remerciements à mon professeur Monsieur **EL QALLI Yassine** pour son encadrement et son suivi durant mon projet de fin d'études. Grâce à son encouragement et les discussions que nous avons eu, j'ai pu m'accomplir totalement dans mes missions. Monsieur EL QALLI a contribué fortement à la nourriture de mon intérêt pour la Finance quantitative et les mathématiques appliquées et je lui exprime toute ma reconnaissance.

Je remercie également Monsieur **BENABADJI Mohamed Oualid**, Asset Manager à La Marocaine Vie. Je souhaite lui exprimer ma profonde gratitude pour sa disponibilité, son encadrement et les compétences qu'il a su me transmettre durant ce stage de fin d'études.

Mes remerciements vont aussi à Monsieur **LYAMMOURI Youssef** pour son accueil chaleureux et sa disponibilité. Je remercie, également, Monsieur **HOUMMANI Ayoub** pour son aide et sa disponibilité permanente.

Je remercie également Monsieur **DOGHMI Ahmed** d'avoir accepté d'examiner ce travail.

Enfin, Je tiens à exprimer toute ma reconnaissance à l'équipe Investissements de La Marocaine Vie ainsi qu'au corps professoral de l'Institut National de Statistique et d'Economie Appliquée.

Résumé

Ce travail porte sur l'analyse des méthodes du contrôle stochastique et de l'approximation stochastique ainsi que leur applications en *trading* algorithmique et haute fréquence et est composé, généralement, de deux parties.

Dans la première partie, nous présenterons la théorie du contrôle stochastique dans le cadre des diffusions continues et des diffusions à sauts ainsi que les méthodes numériques permettant la résolution des problèmes du contrôle. Ce cadre théorique abstrait sera, par la suite, appliqué à l'exécution optimale des ordres. Nous traiterons le problème de liquidation/acquisition optimal à l'aide des ordres au marché et des ordres limites. Nous proposerons un modèle mathématique qui permet de formuler ces dernières problématiques comme des problèmes du contrôle stochastique. Nous démontrerons l'existence des solutions dans les différents cadres traités ainsi que des formules explicites dans quelques cas particuliers très utilisés dans l'industrie. La théorie développée sera illustrée par différentes applications numériques.

La seconde partie est consacrée à l'application de la théorie de l'approximation stochastique à l'exécution optimale des ordres. Nous présenterons, dans un premier temps, les rudiments de la construction des algorithmes stochastiques ainsi que leurs différents résultats de convergence et leurs vitesses de convergences à l'aide d'un Théorème Central Limite des martingales. Le chapitre suivant présente un nouveau paradigme de *trading* optimal basé sur la théorie de l'approximation stochastique. Nous modéliserons le problème d'un algorithme de *market-making* qui actualise ses cotations périodiquement dans un carnet d'ordre limite électronique. Pour résoudre ce problème, nous proposerons un algorithme stochastique d'optimisation pour trouver les meilleures distances de placement d'ordres limites : il s'agit de maximiser le payoff espéré de la stratégie. Ceci mène à la construction d'un algorithme stochastique sous contraintes avec projection. Pour assurer l'existence et l'unicité de l'équilibre, des critères suffisants sur les innovations du modèle sont obtenus à l'aide d'un principe de co-monotonie fonctionnelle. Finalement, nous présenterons une application numérique illustrant la convergence de l'algorithme.

Mots clés : Contrôle stochastique optimal, équation d'Hamilton-Jacobi-Bellman, liquidation optimale, carnet d'ordre, approximation stochastique, *market-making*, co-monotonie fonctionnelle, algorithme stochastique contraint.

Table des matières

Dédicaces	2
Remerciements	3
Résumé	4
1 Introduction	10
1.1 Contexte industriel	10
1.1.1 Microstructure des marchés financiers	11
1.1.2 Mécanismes d'enchère	12
Carnet d'ordre limite	12
Dark-liquidity pools	15
Trading à haute fréquence	16
1.1.3 Trading optimal	16
Courtage (Brokerage)	16
Market-Making	17
1.2 Modèles de prix et de liquidité	18
1.2.1 Modèles mathématiques et statistiques	18
1.2.2 Modèles du Market-impact pour le trading	19
Modèles de résilience (Resilience models)	23
Modèles black-box (Black-Box models)	23
1.2.3 Le trading dans le carnet d'ordre limite	25
Market-Making à haute fréquence	25
Tactiques de la liquidation optimale	27
Trading dans les dark pools	27
1.3 Approches d'optimisation	28
1.3.1 Problème exploitation-versus-exploration	28
1.3.2 Programmation dynamique et équation HJB	29
1.3.3 Approximation stochastique et apprentissage on-line	30
1.4 Plan et motivations	31
2 Contrôle stochastique optimal	32
2.1 Contrôle stochastique des processus de diffusion	32
2.1.1 Principe de la programmation dynamique	34
2.1.2 L'équation de Hamilton-Jacobi-Bellman	35

2.1.3	Sur la régularité de la fonction valeur	36
2.1.4	Théorème de vérification	36
2.2	Contrôle stochastique des diffusions à sauts	37
2.3	Résolution numériques des problèmes de contrôle	40
2.3.1	Méthodes numériques déterministes	40
2.3.2	Méthodes numériques probabilistes	41
	Approximation des espérances conditionnelles	43
3	Exécution optimale à l'aide du contrôle stochastique	49
3.1	Exécution optimale à l'aide des ordres au marché	49
3.1.1	Contexte général	49
3.1.2	Extentions	57
3.2	Exécution optimale à l'aide des ordres limites	61
3.2.1	Cadre théorique : liquidation optimale à la Avellaneda et Stoikov	61
3.2.2	Application numérique	65
3.3	Annexe : Démonstrations des résultats	70
4	Approximation stochastique et apprentissage on-line	75
4.1	Introduction	75
4.2	Exemples d'algorithmes stochastiques	76
4.2.1	Algorithme du bandit à deux bras	76
4.2.2	Algorithme de Robbins-Monro	77
4.2.3	Recherche des extrêmes dans un cadre stochastique	78
4.3	Convergence des algorithmes stochastiques	79
4.3.1	Aperçu sur la théorie	79
4.3.2	Convergence p.s : l'approche martingale	81
4.3.3	Point de vue de l'optimisation : approche de Keifer-Wolfowitz	83
4.3.4	Convergence p.s : approche de L'EDO	83
4.3.5	Vitesse de convergence : le TCL	84
4.3.6	Principe de moyennisation (Ruppert-Polyak)	85
5	Exécution optimale à l'aide de l'approximation stochastique	87
5.1	Introduction	87
5.2	Formulation du problème	89
5.2.1	Problème d'optimisation	89
5.2.2	Problème à une période	89
5.2.3	Relation entre le prix et la liquidité	90
5.3	Analyse du problème	91
5.3.1	Résultat de convergence pour les algorithmes contraints	93
5.3.2	Convergence de la procédure récursive	94
5.4	Applications numériques	95
5.5	Annexe : Démonstrations des résultats	98
	Conclusion	104
	A Présentation de l'Organisme d'accueil	106
	Bibliographie	114

Table des figures

1.1	Organisation des marchés de l' <i>equity</i> Européens avant MiFID (2007).	11
1.2	Organisation des marchés de l' <i>equity</i> Européens après MiFID (2007).	11
1.3	Représentation graphique du carnet d'ordre limite	14
1.4	Evolution d'un carnet d'ordre limite de AMZN entre 10 :00 et 10 :06 à la date 21/06/2012	14
1.5	le <i>Signature-Plot</i> de l'action AMZN	15
1.6	À la date 06 Mai 2010, Dow Jones a expérimenté une chute de plus de 10% durant quelques minutes. L'un des facteurs amplificateurs de ce phénomène était la réaction en chaîne de différentes stratégies automatisées relatives à certaines classes d'actifs. Cet événement a soulevé plusieurs questions réglementaires. . . .	16
1.7	Fonction d'autocorrélation des signes de transactions de l'actif AMZN à la date 21/06/2012	21
1.8	Décroissance hyperbolique de la fonction d'auto-corrélation	21
1.9	Graphique des intensités des processus modélisant le nombre de LOs exécutés	26
2.1	Deux exemples de partitions de Voronoi en dimension 2	45
2.2	Quantification optimale de taille $N = 100$ de la loi normale $\mathcal{N}_2(0, I_2)$	46
3.1	graphe de ν^* et Q^{ν^*} pour $\alpha = 0.01$	54
3.2	graphe de ν^* et Q^{ν^*} pour $\alpha \rightarrow +\infty$	55
3.3	Évolution du processus du flux d'ordre net (μ) de l'action AMZN estimé sur un horizon de 3 min entre 10h00min et 10h03min à la date 21/06/2012	58
3.4	4 trajectoires du processus de l'inventaire et le taux de liquidation optimal	60
3.5	4 trajectoires du processus du <i>midprice</i> et du flux d'ordre net	60
3.6	graphiques de $t \mapsto \delta_t^{*,q}$ pour $q \in \llbracket 1, 5 \rrbracket$, $\alpha \in \{10^{-3}, 10^{-4}\}$, $\kappa = 100$, $\lambda = 50 \text{ min}^{-1}$ et $T = 1$	64
3.7	4 trajectoires du processus de l'inventaire et le taux de liquidation optimal	66
3.8	4 trajectoires du processus du <i>midprice</i> et du flux d'ordre net	66
3.9	histogramme de la différence entre le <i>price per share</i> final de l'agent et le $TWAP_T$ dans les différents scénaris (à gauche). Heatmap de l'inventaire de l'agent dans les différents scénaris (à droite)	67
3.10	Calibration de $\delta \mapsto \lambda(\delta)$ aux données de l'action AMZN à la date 21/06/2012 . .	69
4.1	Évolution des fractions gérées par les deux traders	77

5.1	Exemple d'une trajectoire de l'algorithme stochastique qui converge vers la solution	96
5.2	Procédure de Ruppert-Polyak pour différentes suites de pas	97
A.1	Logo de la marocaine vie	106
A.2	Organigramme de La Marocaine Vie	107
A.3	Répartition du département investissement	108
A.4	Répartition du département investissement	108

Liste des notations

- pour $(x, y) \in \mathbb{R}^2$, $x \wedge y = \min(x, y)$ et $x \vee y = \max(x, y)$
- $\mathbb{H}^p(\mathbb{F})$: L'ensemble des processus ν , \mathbb{F} -progressifs tel que : $\mathbb{E} \left(\int_0^T |\nu_t|^p dt \right) < \infty$ pour $p > 0$.
- $|\sigma| := \left(\text{Tr}(\sigma \sigma^T) \right)^{\frac{1}{2}}$ la norme de Frobenius de la matrice σ .
- $\langle x, y \rangle := \sum_{j=1}^d x_j y_j$ pour $x, y \in \mathbb{R}^d$. On note $|\cdot|$ la norme lui est associée.
- L'intérieur, l'adhérence et la frontière d'une ensemble \mathcal{O} sont notés respectivement $\overset{\circ}{\mathcal{O}}$, $\overline{\mathcal{O}}$ et $\partial \mathcal{O}$.
- $\mathcal{T}_{[t, T]}$ est l'ensemble des temps d'arrêts à valeurs dans $[t, T]$.
- \mathcal{S}_d est l'ensemble des matrices carrés réelles et symétriques d'ordre d .
- $\nabla_x \psi(t, x) = \left(\partial_{x_j} \psi(t, x) \right)_{j \in \llbracket 1, d \rrbracket}$ et $H_x \psi(t, x) = \left(\partial_{x_i x_j}^2 \psi(t, x) \right)_{(i, j) \in \llbracket 1, d \rrbracket^2}$ pour $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ sont respectivement le gradient et la matrice hessienne par rapport à la variable d'espace d'une fonction $\psi \in \mathcal{C}^{1,2}(\mathbb{R}_+^* \times \mathbb{R}^d)$.
- $\{f = 0\} := \left\{ x \in \mathbb{R}^d : f(x) = 0_{\mathbb{R}^n} \right\}$ pour une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}^n$.

Introduction

1.1 Contexte industriel

Le *trading* est un problème de recherche. En effet, les acheteurs et les vendeurs doivent trouver des méthodes qui leur permettent d'acheter ou de vendre la quantité souhaitée à un prix qui paraît "juste" pour eux. Les marchés financiers sont considérés comme les lieux où la demande et l'offre se confrontent. Ces lieux sont transformés aujourd'hui en des marchés électroniques. On dit que le marché est "liquide" si les participants peuvent acheter/vendre ce qu'ils veulent à tout instant et sans paiement de frais de transactions trop élevés. L'un des objectifs majeurs de l'industrie financière est de faciliter l'accès à la liquidité pour les investisseurs qui sont considérés comme les pilotes principaux des forces de l'offre et de la demande dans le marché.

Afin d'accomplir leurs objectifs, deux classes d'intermédiaires aident les investisseurs :

- *Brokers* : Ils s'agit des sociétés financières qui agissent à la place des investisseurs à travers l'exécution des ordres de ces derniers dans les échanges. Les brokers essaient de suivre un prix de référence (*benchmark price*) qui mesure leur performance. Ils font leurs profits à travers les commissions payées par les investisseurs. Dans cette classe, il existe généralement deux sous-classes de traders : les *Fundamental (Noise/Liquidity) traders* qui se basent sur des fondements économiques en dehors de l'échange et les *Informed traders* qui utilisent des informations qui ne sont pas contenues dans les prix de marché afin d'anticiper leurs mouvements.
- *Market Makers* : Considérés comme contrepartie des *Brokers*, ils sont généralement des traders professionnels qui facilitent les transactions pour un certain actif. Ils font leurs profits en vendant à un prix supérieur au prix d'achat.

On dit que les *Brokers* "consomment la liquidité", tandis que les *Markets Makers* "fournissent la liquidité".

Grâce aux avancements technologiques, la disponibilité d'une performance computationnelle faramineuse et le fait que les marchés financiers sont essentiellement électroniques, les agents financiers se basent aujourd'hui sur des stratégies automatisées (*electronic trading*) pour contrôler les coûts, gérer les risques et interagir avec le marché. De ce fait, la technologie, dont

la raison d'être est l'allocation optimale de la liquidité, est devenue l'épine dorsale de l'industrie financière.

Le présent travail se concentre sur le trading algorithmique dans les marchés de *l'equity* du point de vue d'une institution qui fournit des services aux investisseurs (*sell-side firm*). Cette section fournit le bagage nécessaire pour la compréhension du fonctionnement des marchés de *l'equity* et leurs mécanismes d'enchères (*auctions mechanisms*). En plus, on présentera les différents problèmes confrontés par les praticiens ainsi que les propriétés empiriques des marchés financiers dans le cadre du trading à haute fréquence. De ce fait, on pourra mettre nos résultats théoriques en situation en démontrant que ce domaine nécessite une approche quantitative et théorique afin de résoudre les problématiques rencontrées.

1.1.1 Microstructure des marchés financiers

La microstructure est l'éco-système où les intérêts d'achat et de ventes se rencontrent. Traditionnellement, le trading de *l'equity* était principalement centralisé dans un lieu primaire d'échange (Euronext Paris, London Stock Exchange, New York Stock Exchange, Deutsche Börse, Bolsa de Madrid, etc). Grâce aux avancements technologiques, à la globalisation et à la compétition croissante, les régulateurs ont été poussés à "libéraliser" le marché d'échange, donnant ainsi aux participants un choix entre différents échanges. Chacun d'eux a ses propres règles de *order-matching*, schémas de frais des transactions et degrés de transparence. C'est ce qu'on appelle la **fragmentation du marché** (*Market fragmentation*). En ce qui concerne l'Europe, la régulation a eu naissance en 2007 sous le nom de *Markets in Financial Instrument Directive* ou MiFID [93]. Dans les États-Unis, la régulation a été faite en 2005 sous le nom de : *Regulation National Market System* ou Reg NMS.¹

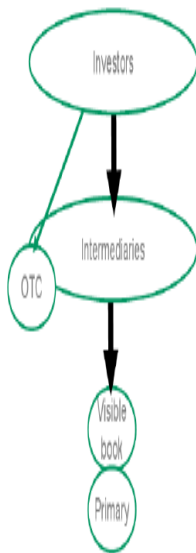


FIGURE 1.1 – Organisation des marchés de *l'equity* Européens avant MiFID (2007).

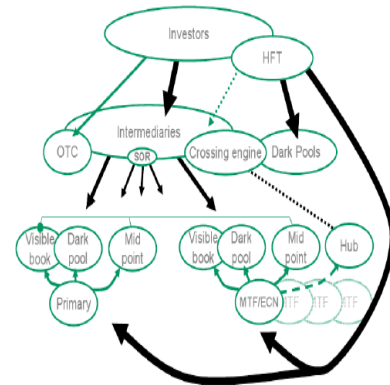


FIGURE 1.2 – Organisation des marchés de *l'equity* Européens après MiFID (2007).

Généralement, les principaux *venues* où le trading de *l'equity* est pratiqué aujourd'hui sont :

- Échanges primaires : Ils s'agissent des échanges traditionnels qui ont existé avant MiFID ou Reg NMS. Ils sont essentiellement *pre-trade transparent* ce qui veut dire que l'information sur la liquidité disponible est accessible aux participants avant le trading. Ils

1. <https://www.sec.gov/rules/final/34-51808.pdf>

restent aujourd'hui parmi les lieux les plus utilisés pour le trading comme ils constituent une référence de prix pour d'autres types d'échanges comme les *Dark Pools* (voir ci-dessous).

- *Multilateral Trading Facilities* (MTFs) : Ils s'agissent de places de trading qui sont en compétition avec les échanges primaires comme leurs services et leurs mécanismes de trading sont approximativement les mêmes. Dans le cas de l'Europe, les MTFs principaux sont BATS Chi-X (London) et Turquoise (London).
- *Dark Pools* : Ils s'agissent de places de trading qui ne fournissent pas des informations sur la liquidité disponible. L'avantage majeur de ces échanges est le fait qu'ils permettent aux investisseurs d'exécuter des ordres larges tout en évitant une fuite des informations où une manipulation par les agents opportunistes comme les traders à haute-fréquence.

M. O'Hara [99] a défini l'étude de la microstructure des marchés financiers (souvent appelée *Market Microstructure*) comme suit : "*is the study of the process and outcomes of exchanging assets under explicit trading rules*". À l'échelle de la microstructure, les marchés sont très sophistiqués. De ce fait, l'étude de la microstructure des marchés se concentre sur le processus qui permet aux investisseurs d'acheter ou de vendre des actifs tout en mettant en lumière le processus de formation du prix appelé : *Price Discovery*. Donc, ce qui différencie ce domaine des autres domaines de la finance quantitative est le fait qu'il se concentre sur deux aspects clés du trading : la liquidité et le *Price Discovery*.

1.1.2 Mécanismes d'enchère

Les forces de l'offre et de la demande à long-terme (investisseurs) ne sont pas les seuls pilotes de l'évolution du prix. En effet la manière d'interaction des *Markets Makers*, des *Brokers* et des autres agents joue également un rôle. Ces derniers influencent, de plus, le processus de formation du prix. Les mécanismes d'enchère (*Auction mechanisms*) constituent un ensemble de règles de *matching* qui définissent la manière d'interaction entre les différents agents. Pour les échanges primaires, on distingue entre deux type de phases d'enchère durant la journée : les enchères de *fixing* (*fixing auctions*) durant lesquels les ordres sont assortis après leur accumulation dans un carnet pendant une certaine période et les enchères continues (*continuous auctions*) pour lesquels les ordres sont assortis continuellement à l'instant où ils apparaissent dans le carnet. Il est à noter que la majorité du volume est échangé durant les enchères continues. De ce fait, on se concentre notamment sur la mécanique de ce type de phases d'enchère.

Carnet d'ordre limite

La place virtuelle de rencontre de l'offre et la demande durant les enchères continues est appelée : Carnet d'ordres limites (*Limit-order book*). C'est la manière de fonctionnement des échanges primaires et des MTFs. Il existe généralement deux moyens pour exprimer les intentions de trading (achat/vente) :

- Ordres au marché (*Market orders*) : Les ordres au marché (MOs) sont considérées comme ordres agressifs qui cherchent à exécuter une trade immédiatement. Dans ce type d'ordres, le trader indique qu'il veut exécuter une certaine quantité fixée d'actifs au prix le plus convenable existant dans le carnet.
- Ordres limites (*Limit orders*) : Les ordres limites (LOs) sont considérées comme ordres passives contrairement au MOs. Dans ce type d'ordres, le trader exprime son désir d'achat et de vente d'une quantité ou une quantité maximale d'actifs à un prix fixé. Comme le prix fixé par ces ordres est usuellement est plus pire que le prix du marché,

un ordre limite doit souvent attendre jusqu'à l'arrivée d'un ordre qui lui sert comme contrepartie ou être annulé.

Le carnet d'ordre limite peut être divisé en deux côtés : le côté *bid* (acheteurs) et le côté *ask* (vendeurs). le plus grand prix-*bid* est appelé le *best-bid* et le plus petit prix-*ask* est appelé le *best-ask*. Par construction, les prix-*bid* sont toujours inférieurs aux prix-*ask*. Si c'est le cas, on doit comprendre que la transaction a eu déjà lieu.

La différence entre le *best-ask* et le *best-bid* est appelée le *bid-ask spread* et il s'agit de la mesure principale de la liquidité. On définit également le *mid-price* comme la moyenne arithmétique des *best-ask* et *best-bid*. Dans la pratique, il n'existe pas une définition du "prix" d'un actif. Le *mid-price* sert comme un "prix de référence" représentant la valeur d'un actif en absence des coûts de transactions. Le *micro-price* a la même fonction que le *mid-price* et définit, à un instant t , comme suit :

$$P_t^* := \frac{V_t^b}{V_t^b + V_t^a} P_t^a + \frac{V_t^a}{V_t^b + V_t^a} P_t^b$$

tel que : V_t^a et P_t^a (resp. V_t^b et P_t^b) sont le volume-*ask* (resp. volume-*bid*) disponible au *best-ask* (resp. *best-bid*) et le *best-ask* (resp. *best-bid*). A la différence du *mid-price*, le *micro-price* mesure la tendance du prix à sauter vers le *best-bid* ou le *best-ask*. Si par exemple, V_t^b est très grand par rapport au V_t^a ($\frac{V_t^b}{V_t^b + V_t^a} \simeq 1$) alors on comprend qu'il y a plus d'acheteurs que de vendeurs et le prix doit donc augmenter pour se stabiliser vers le *best-ask*. Le lecteur doit comprendre que le *mid-price* et le *micro-price* ne sont que des mesures artificielles du "prix de référence" qui nous servent à quantifier ce dernier lors de la construction des modèles théoriques. Il s'agit de l'équivalent d'un "gaz parfait" dans la théorie cinétique des gaz en physique. Un autre concept important est le *market-depth* qui mesure la liquidité disponible dans le carnet d'ordre. Le *bid-ask spread* et le *market-depth* constituent les mesures principales du coût de la consommation de la liquidité.

Lorsqu'on traite un carnet d'ordre limite, il existe d'autres concepts qu'il faut prendre en considération :

- Les prix vivent dans une subdivision uniforme de pas appelé le *tick*. De ce fait la différence entre deux prix d'un actif ne peut pas être inférieure au *tick*. Dans les États-Unis le *tick minimum* est égal à 1 cent pour les actions ayant un prix supérieur à un dollar. Dans la bourse de Paris ou la Bolsa de Madrid le *tick* peut prendre des valeurs allant de 0.001 à 0.05 euros.
- Les ordres arrivent à des temps aléatoires. De ce fait, les changements du *mid-price* ont lieu à des instants aléatoires d'une manière discrète.
- Dans le marché de l'*equity*, les ordres sont assortis généralement en suivant une priorité prix-temps. Supposons qu'un MO demandant/offrant une certaine quantité arrive au carnet. Ce dernier sera assorti avec le LO proposant le prix le plus convenable P . Si la quantité demandée/offerte est inférieure à celle présente au prix P , alors l'algorithme de *matching* commence la liquidation par le LO le plus ancien et ainsi de suite. Si la quantité est supérieure à celle qui est présente au prix P , alors dans ce cas, l'algorithme commence l'exécution par les LO proposant le prix P et continue la liquidation par les LO proposant le deuxième plus convenable prix et ainsi de suite. Il est à noter que, dans les États-Unis, dans ce dernier cas l'ordre peut être réacheminé vers un autre échange proposant le prix P et s'il n'existe pas de liquidité disponible à ce prix dans tous les échanges, il continue comme dans le cas décrit auparavant.

Il existe d'autres concepts relatifs au carnet d'ordre et aux mécanismes d'enchères continues, tel que : le *market-latency*, *colocation*, *market-resilience*, *exchange fees schemes*, etc. Pour plus

d'informations sur les carnets d'ordres et leurs caractéristiques le lecteur est renvoyé à [119], [86] et [2].

On remarque que le processus de formation du prix est majoritairement piloté par la manière avec laquelle les agents choisissent leurs stratégies pour faire des transactions dans le carnet d'ordre limite, mais également par les différentes caractéristiques des mécanismes d'enchère. Il est à noter que les mesures classiques du risque tels que la volatilité doivent être traitées avec précaution à l'échelle microscopique à cause du bruit de la microstructure. En effet, il existe plusieurs phénomènes qui biaisent l'estimation de ces indicateurs, à titre d'exemple on cite : l'effet Epps [117], l'effet Lead-Lag [114], l'effet de signature, etc.

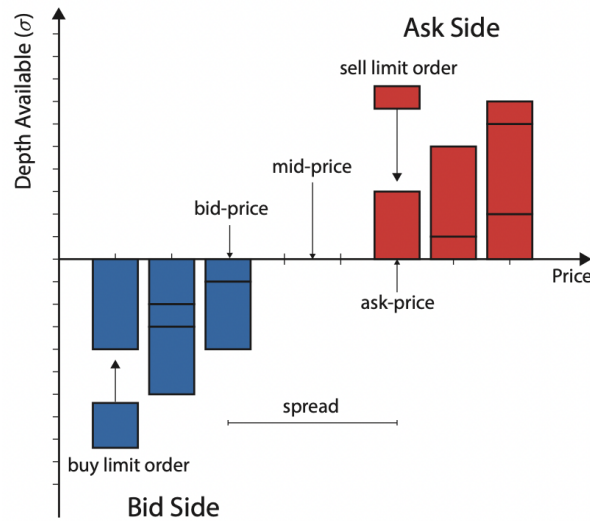


FIGURE 1.3 – Représentation graphique du carnet d'ordre limite

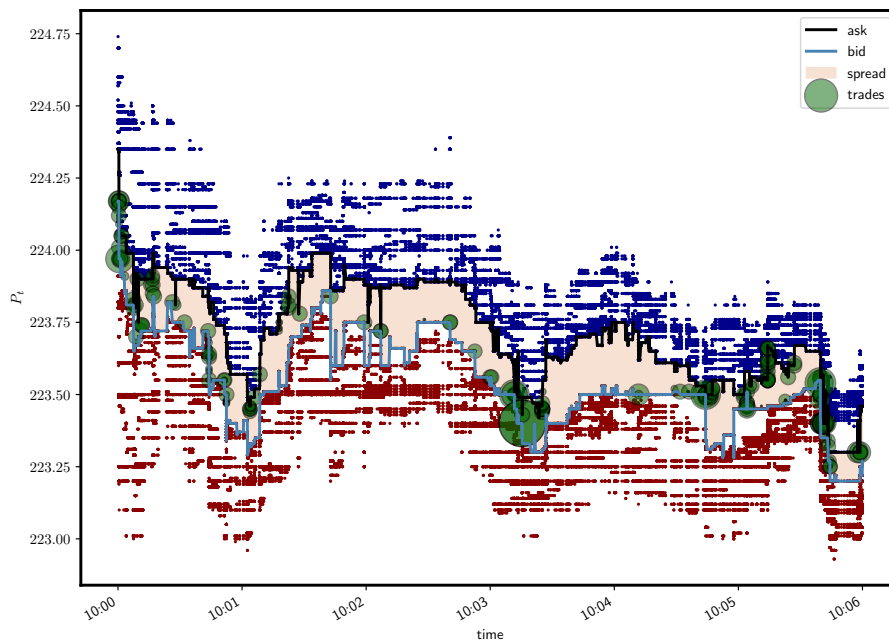


FIGURE 1.4 – Evolution d'un carnet d'ordre limite de AMZN entre 10 :00 et 10 :06 à la date 21/06/2012

La figure 1.4 représente l'évolution du carnet d'ordre limite de l'action AMZN à la date 21/06/2012 sur une durée de 6 minutes allant de 10h00 à 10h06 du matin. Les points en bleu représentent les prix-ask et ceux en rouge représentent les prix-bid. Les cercles en vert représentent les prix des transactions. Le diamètre du cercle est une fonction croissante du volume du transaction. La ligne en noir représente l'évolution du *best-ask* et celle en bleu-métallique représente le *best-bid*. L'aire hachurée entre ces deux ligne représente le *bid-ask spread*. La figure met également en évidence "la nature discontinue" du processus du prix à l'échelle microscopique.

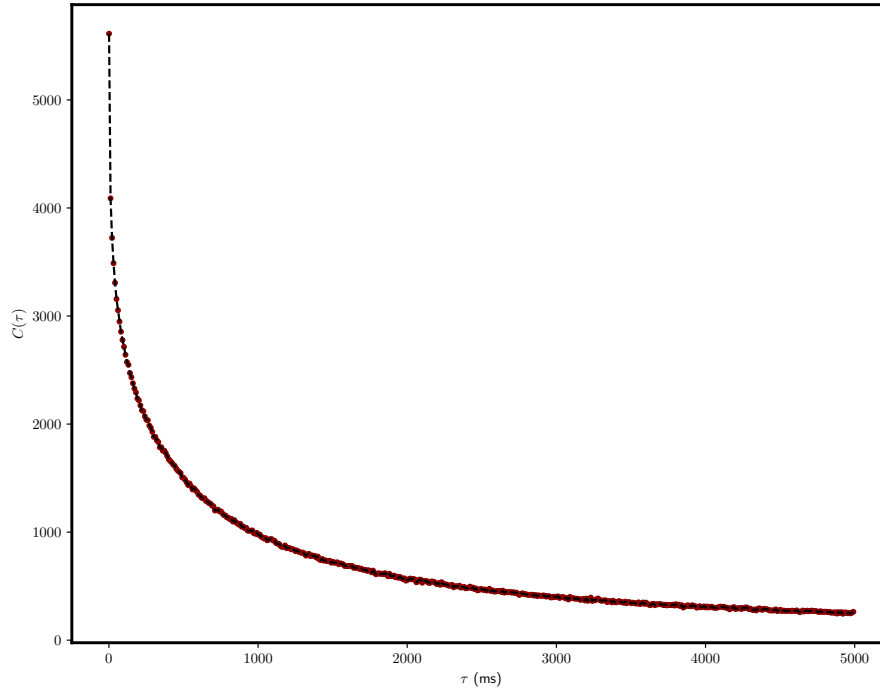


FIGURE 1.5 – le *Signature-Plot* de l'action AMZN

La figure 1.5 met en lumière l'effet de signature de la volatilité sur une échelle microscopique qui se manifeste par la croissance de celle-ci lorsque diminue la fréquence d'échantillonnage. Il s'agit du graphique de signature de l'action AMZN à la même date qu'auparavant. τ est la fréquence d'échantillonnage mesure en millisecondes, $C(\tau)$ est la variance réalisée du *mid-price* associée à la fréquence τ considérée sur un horizon T . Elle est définie par :

$$C(\tau) := \sum_{k=1}^{\lfloor T/\tau \rfloor} \left(X_{k\tau} - X_{(k-1)\tau} \right)^2$$

tel que X est le processus du *mid-price*.

Dark-liquidity pools

les autres lieux importants qui contribuent au processus de formation du prix, en dehors des carnets d'ordre limite, sont les *Dark-liquidity pool*. Le prix avec lequel les ordres sont assortis dans les *Dark pool* est le *mid-price* figurant dans les échanges primaires et les MTFs. Ainsi, ils contribuent implicitement au processus de formation du prix en attirant une part de la liquidité disponible dans les *Lit-exchanges* (tel que les échanges primaires et les MTFs).

Trading à haute fréquence

À part le trading algorithmique, un autre produit des avancées technologiques s'est apparu. Il s'agit de ce qu'on appelle le trading à haute fréquence, *High-Frequency Trading* (HFT). Le HFT est un style de trading majoritairement opportuniste dans lequel les stratégies de trading sont construites pour acquérir des profits à court terme en se basant sur des modèles statistiques [100]. Le HFT se base sur la technologie, la vitesse et le traitement rapide des données. Après le *flash crash* (voir figure 1.6), à cause de sa réputation d'introduction de liquidité toxique au marché, le HFT est devenu un sujet controversé d'un point de vue réglementaire [40], [76].

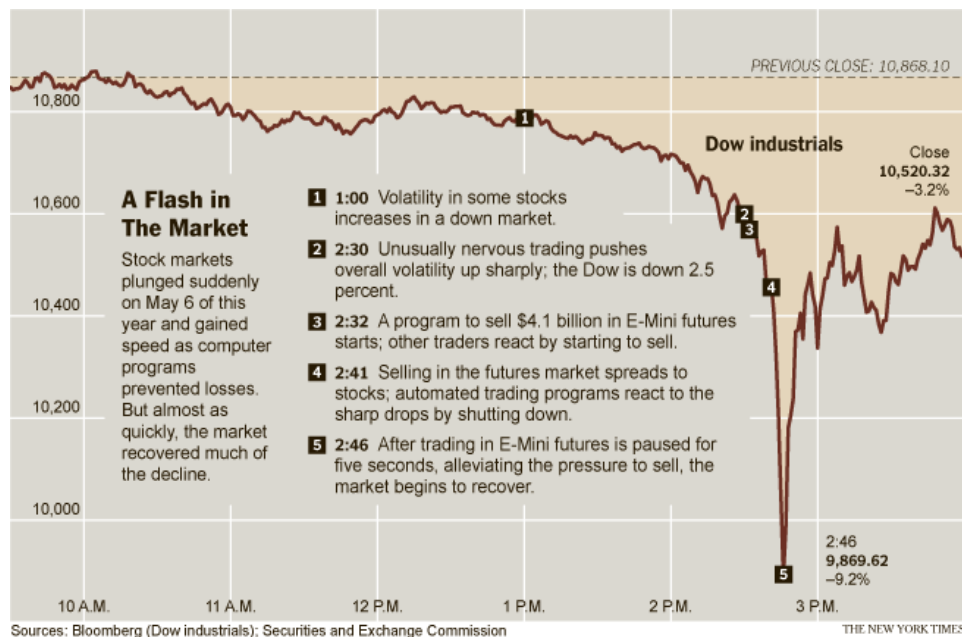


FIGURE 1.6 – À la date 06 Mai 2010, Dow Jones a expérimenté une chute de plus de 10% durant quelques minutes. L'un des facteurs amplificateurs de ce phénomène était la réaction en chaîne de différentes stratégies automatisées relatives à certaines classes d'actifs. Cet événement a soulevé plusieurs questions réglementaires.

1.1.3 Trading optimal

Les évolutions de l'automatisation du trading ont soulevé un besoin remarquable pour des approches quantitatives d'analyse statistique, de modélisation mathématique et d'optimisation. Le trading optimal (*Optimal trading*) est une ensemble de techniques quantitatives utilisées dans la construction des stratégies de trading. Il consiste à chercher une balance convenable entre la fourniture de la liquidité afin de minimiser l'impact des transaction et la consommation de la liquidité afin de minimiser l'exposition au risque du marché. Tout en créant du profit à l'aide des signaux de trading dus à l'inefficience de la liquidité. On présentera un aperçu du trading optimal dans le cadre de deux domaines qui concernent le présent travail : le courtage (*brokerage*) et le *market-making*.

Courtage (Brokerage)

Après que les investisseurs et les *brokers* se mettent d'accord sur une référence (*benchmark*) par rapport à laquelle ils mesurent leur performance, les *brokers* font face à des problèmes liés aux coûts de transactions dus à l'exécution des ordres larges. Parmi ces coûts, le *market impact* mérite une attention particulière. Il s'agit du phénomène décrivant les fluctuations du prix engendrées par l'exécution des stratégies du trading. En effet, on observe empiriquement

que l'exécution d'ordres d'achat larges engendre une hausse du prix, tandis que l'exécution d'ordres de vente larges provoque une diminution du prix. Afin de diminuer le *market impact*, *brokers* divise leurs ordres larges appelés *parent orders* en des petits ordres appelés *child/atomic orders* qui sont exécutés durant une période fixée. Cette séquence d'ordres est souvent appelé un *metaorder*. De fait, leur objectif est la recherche d'un taux de liquidation/acquisition optimal (dans un certain sens) qui leur permet de liquider/acquérir une large quantité d'un actif cible. Deux problèmes soulèvent durant ce processus : le premier est de nature stratégique qui consiste en la manière de définir le taux de trading afin de respecter les contraintes du *benchmark*. Le deuxième est un problème de nature tactique et consiste en la manière d'interaction avec le marché afin d'exécuter l'ordre avec un coût minimal.

Le contexte mathématique de résoudre ces problèmes d'optimisation sera expliqué dans les chapitres suivants. À ce stade là, il suffit de comprendre qu'une fois le *benchmark* est définie et le modèle du *market impact* est construit, la procédure d'optimisation nous donne comme résultat le processus du taux du trading optimal. Pour construire ce taux, les praticiens prennent en considération plusieurs variables tels que : la volatilité intraday, les *intraday volume patterns*, le spread, etc [77], [15], [78], [92], [94], [12].

Quant au problème d'interaction avec le marché durant l'exécution des ordres (i.e, l'aspect tactique), les praticiens font face à deux problématiques majeures qui peuvent influencer la performance du trading :

- Le placement des ordres (*Order placement*) : il s'agit du choix entre les ordres au marché (agressifs) et les ordres limites (passives). Si le *broker* choisit de consommer la liquidité à l'aide d'un ordre au marché agressif, il risque d'exécuter à de mauvais prix. Si son choix porte sur un ordre limite passif, il gagne l'avantage d'exécuter à de bon prix, toutefois il risque que son ordre ne va jamais être exécuté.
- L'acheminement des ordres (*Order routing*) : deux aspects doivent être pris en considération. Le premier est la nature du lieu d'échange qui va recevoir l'ordre, à savoir : un échange *lit* (échanges primaires, MTFs) ou un *Dark pool*. Le deuxième concerne la manière d'exécution dans chaque *venue* qui dépend des probabilités d'exécutions, les *spreads*, les schémas des frais de transactions, etc.

Market-Making

Rappelons qu'un *market maker* fournit la liquidité au marché par des propositions de vente et d'achat à travers des ordres limites. Pour que son activité soit profitable, il faut que le prix d'achat qu'il propose doit être inférieur au prix de vente proposé. La différence entre ces prix est appelé *spread* du *market maker* (qui ne doit pas être confondue avec le *bid-ask spread*). Le *market maker* ne cherche pas à acquérir un inventaire permanent. Toutefois, son activité l'expose à un risque qui consiste à avoir un inventaire déséquilibré (si, à titre d'exemple, il ne trouve pas de contrepartie pour lui vendre ce qu'il a acheté à un prix généralement plus mauvais que le *mid-price*). Afin de minimiser ce risque, les *market makers* actualisent continûment les prix qu'ils postent dans le carnet d'ordre limite. La manière avec laquelle ils choisissent leurs prix (i.e, leurs *spreads* par rapport au *mid-price* et leurs positions) est influencée par différents compromis :

- La probabilité d'exécution et le *spread* : La variation du *spread* peut engendrer des effets négatifs et positifs. En effet, le profit augmente quand le *spread* augmente. Toutefois, de larges *spreads* traduisent des prix proposés moins attrayant pour les traders agressifs, ce qui engendre une diminution de la probabilité d'exécution.
- Le risque d'inventaire, le prix et le *spread* : Le *market maker* centre ces ordres autour des prix qui maximisent la probabilité que l'exécution ait lieu pour le côté dans lequel il y a

un excès de quantité dans son inventaire. L'impact d'un inventaire déséquilibré est plus prononcé vers la fin de la journée, comme le *market maker* veut liquider son inventaire le plus rapidement possible. Le *spread* contribue également au risque d'inventaire, comme l'exécution de plusieurs ordres augmente la "variance" de l'inventaire.

- La sélection adverse : Un problème important rencontré par le *market maker* est la situation dans laquelle un trader opportuniste possédant plus d'informations sur les mouvements du prix que le premier, exécute des ordres d'achat/vente contre celles du *market maker* en sachant que le prix va augmenter (diminuer). Ce phénomène engendre un déséquilibre d'inventaire pour le *market maker* qui doit rééquilibrer son inventaire dans le future en achetant/vendant à de mauvais prix.
- Les Rabais : Les *market makers*, en tant que fournisseurs de liquidité, reçoivent des rabais ou remboursements par les organisateurs d'échange. Ils constituent la source la plus importante des profits du *market maker*.

1.2 Modèles de prix et de liquidité

Dans les sous-sections qui suivent, on soulignera les raisons qui motivent le choix des modèles que nous allons utiliser dans ce travail. On présentera la littérature existante dans le domaine du trading quantitatif.

1.2.1 Modèles mathématiques et statistiques

Comme nous avons déjà mentionné, la différence principale entre les modèles quantitatifs utilisés dans le domaine du trading algorithmique et ceux utilisés dans les domaines classiques de la finance quantitative tel que le pricing des produits dérivés, réside dans la concentration des premiers non seulement sur la modélisation du processus du prix, mais également sur la liquidité à travers la modélisation du carnet d'ordre limite et le *market impact*. De plus, ils prennent en considération les variables qui impactent le prix à court terme, tel que la taille du *tick*, le *bid-ask spread*, le *depth* du carnet d'ordre, le *price-impact*, les taux d'arrivée des ordres, etc.

Durant les années récentes, plusieurs approches ont été développées dans la littérature pour prendre en compte les aspects de la microstructure dans les modèles quantitatifs [86], [1]. Dans un premier temps, on peut classer ces modèles de prix et de liquidité en se basant sur la [granularité de la dynamique du prix](#). En ce sens, les modèles du bruit de microstructure (*microstructure-noise models*) [3], [110], [109], [111], [118], qui ont la granularité la plus grossière, modélisent le processus du prix par une diffusion d'Itô et lui ajoutent les effets du bruit de la microstructure. Ces modèles ont été originalement construits pour des raisons d'estimation statistique (estimation de la volatilité à haute fréquence, etc) en présence du bruit du microstructure (effet Epps, effet de signature ...). D'autres modèles ayant les mêmes objectifs que les premiers mais possédant une granularité plus fine ont été étudiés dans [11]. Ces derniers modèles modélisent le processus du prix comme une différence de deux processus de Hawkes [67]. De cette manière, les mécanismes statistiques locaux et asymptotiques du processus du prix peuvent être caractérisés en utilisant l'analyse spectrale.

Une autre manière de voir le prix et la liquidité est de prendre comme point de départ la [granularité la plus fine du carnet d'ordre limite](#). De cette manière, le prix et la liquidité sont considérés comme des résultats macroscopiques de l'interaction entre les ordres au marché (MOs) et les ordres limites (LOs). Dans cette direction, on peut citer les modèles à intelligence zéro (*Zero-intelligence models*) [113], [45], [36]. Ces modèles nous permettent d'obtenir des relations entre les variables macroscopiques en modélisant le carnet d'ordre limite par un sys-

tème de files d'attentes en précisant les intensités des différents événements présents dans le LOB. Même s'ils sont intéressants d'un point de vue théorique et de simulation, ces modèles présentent un inconvénient majeur qui consiste en la difficulté d'obtenir une caractérisation mathématique simple des propriétés asymptotiques du prix et du *market impact*.

Finalement, on peut citer l'approche d'écono-physique inspiré par le modèle de diffusion-réaction dans la physique théorique [30], [42], [41], [52]. Ces modèles ont été construits originalement pour résoudre des problèmes d'estimation des paramètres ou bien pour décrire simplement la relation entre le prix et la liquidité. Ces objectifs ne sont pas colinéaires avec l'optimisation des processus du trading qui est plus concentrée sur l'étude de l'impact des transactions sur le prix, les probabilités d'exécution et le risque du marché. De ce fait, ces modèles n'intègrent pas les algorithmes de trading dans leur approche.

Pour un panorama de ces modèles, le lecteur est renvoyé à [2] (partie 2) et [53].

1.2.2 Modèles du Market-impact pour le trading

Comme il a été déjà mentionné, l'un des sujets importants dans la modélisation des stratégies de trading algorithmique est la prise en considération l'impact de nos transactions sur la dynamique du prix. Ce phénomène est appelé le *market impact* et il peut être de nature mécanique (fluctuations du prix dues au volume de nos ordres) ou de nature informationnelle (mouvements du prix dues au changements des opinions des investisseurs).

La mesure la plus intéressante du *market impact* (mécanique) entre deux date t_0 et $t_0 + \Delta$ tel que $\Delta > 0$ et la corrélation conditionnelle entre le changement du prix entre ces deux date et le signe de la transaction à t_0 par rapport au volume de la transaction à t_0 . On fixe un horizon de temps $T > 0$ et on note, $(\varepsilon_t)_{t \in [0, T]}$ le processus des signes des transactions défini à un instant t par : $\varepsilon := +1$ s'il s'agit d'un achat et $\varepsilon_t := -1$ s'il s'agit d'une vente. On note également le processus du prix par $(P_t)_{t \in [0, T]}$ et le processus du volume de nos transactions (taux de participation) par $(v_t)_{t \in [0, T]}$. Alors, la courbe du *market impact* est définie pour tout $(\Delta, v) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ par :

$$\psi(\Delta, v) := \mathbb{E}(\varepsilon_{t_0} (P_{t_0+\Delta} - P_{t_0}) \mid v_{t_0} = v) - \mathbb{E}(\varepsilon_{t_0} \mid v_{t_0} = v) \mathbb{E}(P_{t_0+\Delta} - P_{t_0} \mid v_{t_0} = v)$$

Généralement, T (et donc Δ) est choisi suffisamment petit (de l'ordre de quelques secondes) de manière à ce que le processus des mouvements du prix n'a pas de drift ou bien à ce qu'il n'y a pas d'asymétrie de transactions ($\mathbb{E}(\varepsilon_{t_0} \mid v_{t_0} = v) = 0$). Ainsi, ψ est généralement égale à :

$$\psi(\Delta, v) = \mathbb{E}(\varepsilon_{t_0} (P_{t_0+\Delta} - P_{t_0}) \mid v_{t_0} = v)$$

La littérature du *market impact* a commencé avec le travail séminal de Kyle [83] durant les années 1980. Il a introduit un modèle discret pour le prix dans lequel le processus du changement du prix est donné, pour un horizon $N \in \mathbb{N}^*$, par :

$$\forall n \in \llbracket 0, N-1 \rrbracket : \quad P_{n+1} - P_n = \lambda \varepsilon_n v_n + \xi_n,$$

tel que $(\xi_n)_{n \in \llbracket 0, N \rrbracket}$ est un processus centré indépendant de $(\varepsilon_n)_{n \in \llbracket 0, N \rrbracket}$ et $\lambda > 0$ est une constante quantifiant l'intensité du *market impact*.

Ainsi la variation du prix entre deux dates n et $n + k$ avec $k \in \mathbb{N}^*$ est donnée par :

$$P_{n+k} - P_n = \lambda \sum_{j=n}^{n+k-1} (\varepsilon_j v_j + \xi_j)$$

Kyle a imposé que le processus $(\varepsilon_n)_{n \in \llbracket 0, N \rrbracket}$ soit stationnaire au sens faible et non auto-corrélé. De ce fait la courbe du *market impact* sera donnée par :

$$\psi(k, v) := \lambda \sum_{j=0}^{k-1} \mathbb{E}(\varepsilon_n \varepsilon_{j+n} v_j \mid v_n = v) = \lambda v$$

Ainsi, on constate que le *market impact* est permanant et linéaire dans ce modèle.

Le modèle de Kyle n'est pas consistant avec les résultats empiriques. En effet, lorsque les institutions financières ont dessiné des graphiques de la fonction ψ , elles ont constaté qu'il existe une relation de nature concave entre le changement de prix signé et le volume. Ces observations empirique sont agrégées sous la nomination *square-root law*. Le littérature en ce sens est liée au [18] et stipule que ψ est indépendante de la duration Δ est s'écrit sous la forme :

$$\psi(\Delta, v) \propto \sigma \sqrt{\frac{v}{V}}$$

tel que V est le volume totale exécuté durant la période d'intérêt. Toutefois, plusieurs auteurs [29], [86] ont reporté une relation de la forme :

$$\psi(\Delta, v) \propto \Delta^\gamma \left(\frac{v}{V} \right)^\alpha,$$

tel que $\gamma, \alpha \in]0, 1[$.

Une des explication de la forme concave de la fonction du *market impact* est la persistance du flux d'ordres. De ce fait, l'hypothèse de la non-autocorrélation des signes des transactions émise par Kyle n'est pas vérifiée empiriquement. En effet, ce processus exhibe une mémoire longue décrite par la décroissance lente de sa fonction d'auto-corrélation. La figure 1.7 montre cette dernière propriété dans nos données pour l'actif AMZN à la date 21/06/2012.

Si on note, pour tout $k, n \in \mathbb{N}^*$:

$$\varphi(k) := \mathbb{E}(\varepsilon_n \varepsilon_{n+k})$$

alors, on constate que cette fonction prend la forme :

$$\varphi(k) \propto k^{-\gamma} \tag{1.2.1}$$

tel que $0 < \gamma < 1$.

La figure 1.8 montre un ajustement de φ de la forme 1.2.1. On constate que parmi les formes que peut prendre la fonction d'autocorrélation, une décroissance en *power-law* est un candidat potentiel.

De ce fait, en adoptant la forme 1.2.1, le raisonnement heuristique suivant nous permet d'inférer une forme plus réaliste de la courbe du *market impact* :

$$\begin{aligned} \mathbb{E}(\varepsilon_n (P_{n+k} - P_n) \mid (v_n, \dots, v_{n+k-1})) &\propto \sum_{j=0}^{k-1} v_{n+j} \mathbb{E}(\varepsilon_n \varepsilon_{j+n} \mid (v_n, \dots, v_{n+k-1})) \\ &\propto v_n + \sum_{j=1}^{k-1} v_{n+j} j^{-\gamma} \\ &\propto v_n + \sum_{j=n+1}^{n+k-1} v_j (j-n)^{-\gamma} \end{aligned}$$

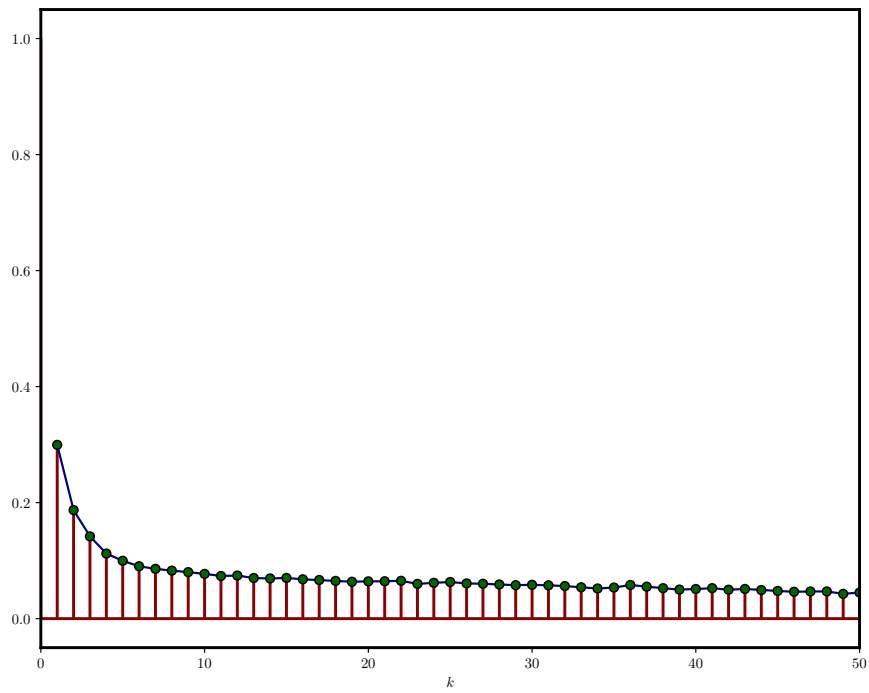


FIGURE 1.7 – Fonction d'autocorrélation des signes de transactions de l'actif AMZN à la date 21/06/2012

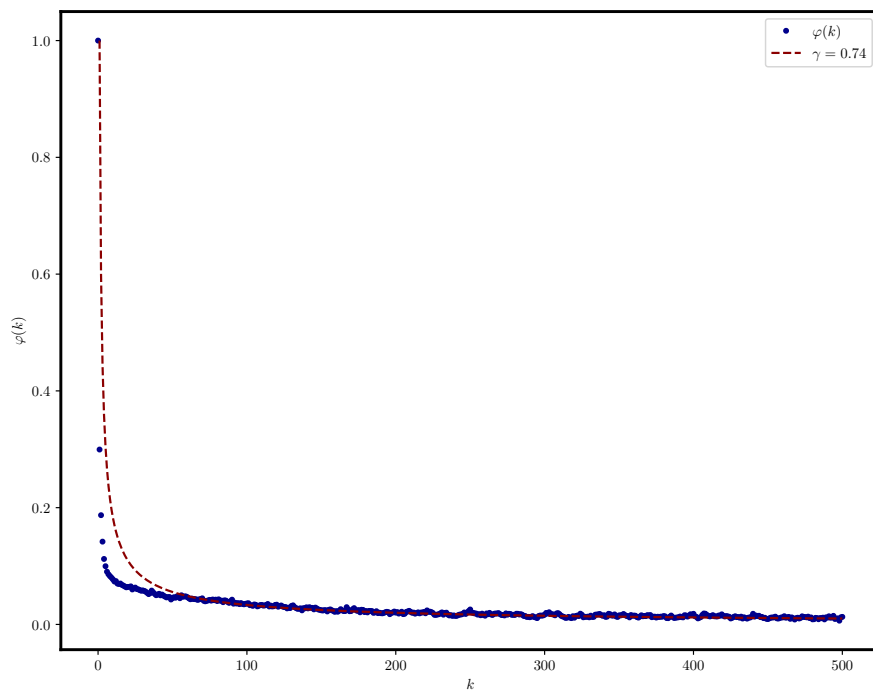


FIGURE 1.8 – Décroissance hyperbolique de la fonction d'auto-corrélation

On constate alors que le market impact n'est pas permanent comme dans le modèle de

Kyle. En effet, on constate que l'effet est "transient" et décroît durant la période d'exécution en $(j - n)^{-\gamma}$.

Pour prendre en compte le caractère transient du *market impact* et une dépendance générale entre le volume et le changement du prix, J. Bouchaud et al. [31] ont introduit le modèle du "propagateur" dans lequel le prix est un processus en temps discret défini par :

$$\forall n \in \llbracket 0, N \rrbracket : \quad P_n := P_0 + \sum_{k=0}^{n-1} G(n-k) \varepsilon_k f(v_k) + \xi_n$$

te que G et f sont des fonctions appelées respectivement le noyau de décroissance (*Decay kernel*) et la fonction du *market impact* temporaire. $N \in \mathbb{N}^*$ est un horizon de temps. $(\xi_n)_{n \in \llbracket 0, N \rrbracket}$ est un processus stationnaire, centré et indépendant de $(\varepsilon_n)_{n \in \llbracket 0, N \rrbracket}$.

Dans le cadre du modèle du propagateur le *market impact* n'est plus permanent et linéaire. En effet, il possède une composante temporaire décrite par une fonction f et décroît avec le temps. Cette décroissance temporaire traduit la composante transiente du *market impact*. On peut leur ajouter une composante permanente qui persiste après la finalisation de l'exécution du *metaorder*. Ceci a été exploré empiriquement par Bouchaud et al. [107].

Ce modèle a été un point de départ de plusieurs recherche portant sur la forme des fonction f et G . Des investigations empiriques [33], [29], [107] ont trouvé que la fonction f doit être concave. Elle prend souvent les formes suivante :

$$f(v) \propto v^{\psi(\Delta)} \quad ; \quad f(v) \propto \log(v)$$

tel que Δ est la duration du *metaorder*. Quant aux fonction G , deux formes sont utilisées généralement dans la pratique. une décroissance exponentielle [98] et hyperbolique [29].

J. Gatheral [51] a introduit une version continue du modèle du propagateur. Pour un horizon de temps $T > 0$, le processus du prix (de référence) est donnée par :

$$\forall t \in [0, T] : \quad P_t = P_0 + \int_0^t G(t-s) f(v_s) ds + \sigma W_t$$

tel que $\sigma > 0$ est la volatilité de l'actif $(W_t)_{t \in [0, T]}$ est un mouvement brownien naturel et standard. En plus, il a étudié des conditions de non-arbitrage dynamique dans le cadre de ce modèle. Un cas important est celle d'un noyau G et d'une fonction de *market impact* temporaire f ayant les formes suivantes :

$$\forall (v, t) \in \mathbb{R}_+^* \times]0, T] : \quad G(t) = t^{-\gamma} \quad , \quad f(v) = v^\alpha$$

tel que $\gamma, \alpha \in]0, 1[$. Dans ce cas, J. Gatheral a démontré qu'une condition suffisante garantissant l'absence d'arbitrage dynamique est : $\gamma + \alpha \geq 1$.

Dans la pratique, la quantité $\lim_{t \rightarrow +\infty} f(v)G(t)$ est appelé le *permanent impact* et incorpore l'effet du *metaorder* sur le prix à long terme après la fin de l'exécution. $\lim_{t \rightarrow 0^+} f(v)G(t)$ est appelé le *temporary impact* et traduit un sur-coût d'exécution sans impact visible sur le prix. $f(v) \left(\lim_{t \rightarrow 0^+} G(t) - \lim_{t \rightarrow +\infty} G(t) \right)$ est appelé le *transient impact* et s'agit de la partie progressivement absorbée par le marché au moyen terme.

Généralement, on peut classifier les modèles qui prennent en considération le *market impact* en deux classes : les modèles de résilience et les modèles *black-box*.

Modèles de résilience (Resilience models)

Cette famille de modèles a été inspirée par les travaux de Obizhaeva et Wang [98] et J. Gatheral [51] et modélise l'impact des transactions sur le prix en considérant le fait qu'une transaction engendre un mouvement du prix par une consommation de la liquidité avant que le carnet d'ordre réagit pour retourner à l'équilibre. Ce dernier phénomène est ce qu'on appelle la "résilience" du carnet d'ordre. Ce terme est inspiré de la physique qui caractérise la résistance d'un métal à un choc. Cette ligne de recherche a été largement exploré par Alfonsi et al. dans [4]. Ils modélisent la forme du carnet d'ordre à l'aide de ce qu'on appelle les *shape functions* pour en trouver les fonctions du *market impact*. En bref, supposons qu'un trader désire acheter une grande quantité d'un actif. Pour quantifier l'impact de cette transaction sur le prix de vente, on suppose que le *best ask-price* en absence de ce trader agressif est une martiale définie sur un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. On suppose de plus qu'il existe une fonction densité $f : \mathbb{R} \rightarrow \mathbb{R}_+^*$ continue appelé *shape function* tel que la quantité offerte à des prix appartenant à l'intervalle $[A_t^0, A_t^0 + \delta_t]$ à l'instant t est :

$$E_t^A = \int_0^{\delta_t} f(y) dy$$

Le cas d'une fonction f constante correspond à ce qu'on appelle *Block-shaped LOB*. Si le carnet d'ordre a cette forme particulière, alors la quantité disponible à un prix ayant une distance $\delta > 0$ du *best ask-price* est la même quelque soit la distance δ . Soit A le vrai *best ask-price* en présence du trader agressif et définissons le *extra spread* causé par les transactions du trader, pour un instant $t > 0$, par : $D_t^A := A_t - A_t^0$. De ce fait, le processus D^A quantifie l'impact des MOs sur le *best ask-price*. Maintenant, supposons que le trader exécute une quantité $x_t > 0$ à un instant t qui engendre un accroissement du prix de la valeur A_t à $A_{t+} = A_t + D_{t+}^A - D_t^A = A_t^0 + D_{t+}^A$. Alors, on a :

$$\int_{D_t^A}^{D_{t+}^A} f(x) dx = x_t \quad (1.2.2)$$

De ce fait, on remarque que si l'expression de f est spécifiée, on peut trouver une relation caractérisant le *market impact* qui lie le volume exécuté au mouvement du prix. La résilience du carnet d'ordre peut être caractérisée mathématiquement en précisant la relation qui lie E_{t+s}^A et E_t^A dans le cas où le trader agressif est absent dans l'intervalle $[t, t+s]$. Les auteurs considèrent une résilience exponentielle sous la forme :

$$E_{t+s}^A = e^{-\kappa s} E_t^A$$

tel que $\kappa > 0$ est la vitesse de la résilience.

Remarquer que dans le cas d'un *bloc-shaped LOB* dans lequel $f \equiv \beta$, on trouve un *market impact* linéaire en volume. En effet, d'après 1.2.2 :

$$\beta (D_{t+}^A - D_t^A) = \beta (A_{t+} - A_t) = x_t \implies A_{t+} - A_t = \beta^{-1} x_t$$

Malgré leur avantage théorique qui permet de caractériser l'absence d'arbitrage par des relations entre le prix et la liquidité, ces modèles sont très difficiles à calibrer aux données du marché.

Modèles black-box (Black-Box models)

Une direction de recherche plus pragmatique a été introduite par le travail séminal d'Almgren et Chriss dans [6] et les travaux d'Almgren qui en suivent [7], [5]. Ils ont introduit un modèle *black-box* en temps discret pour le *market impact* dans lequel il est facile d'obtenir une

solution optimale du problème du trading à l'aide des techniques d'optimisation classiques. L'avantage majeur de ces modèles est le fait que le *market impact* est caractérisé par un petit nombre de paramètres qui peuvent être estimés par les données du marché. De ce fait, ces modèles sont flexibles, bien adaptés aux applications numériques et peuvent être facilement généralisés. Brièvement, fixons un horizon de temps $T > 0$ et un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ sur lequel on définit les processus qui suivent. Soient $N \in \mathbb{N}^*$, $(t_k)_{k \in \llbracket 0, N \rrbracket}$ une subdivision uniforme de $[0, T]$ de pas constant Δt et $(\varepsilon_k)_{k \in \llbracket 1, N \rrbracket}$ une suite de v.a i.i.d centrés réduites. Supposons qu'un trader veut liquider une quantité large d'un actif que l'on Q en de petits lots $(Q_k)_{k \in \llbracket 0, N \rrbracket}$, tel que Q_k est la quantité qui lui reste à l'instant t_k pour tout $k \in \llbracket 0, N \rrbracket$. Pour qu'il s'agit d'une activité de liquidation on impose les conditions suivantes : $Q_0 = Q$ et $Q_N = 0$. Soit $(v_k)_{k \in \llbracket 1, N \rrbracket}$ le taux de liquidation défini par :

$$\forall k \in \llbracket 1, N \rrbracket : \quad Q_k = Q_{k-1} - v_k \Delta t$$

De ce fait $v_k \Delta t$ est la quantité d'actif vendue par le trader dans l'intervalle $]t_{k-1}, t_k]$.

Le processus du prix de référence (classiquement le *mid-price*) est défini par la relation de récurrence suivante :

$$\forall k \in \llbracket 1, N \rrbracket : \quad S_k = S_{k-1} - g(v_k) \Delta t + \sigma(\Delta t)^{1/2} \varepsilon_k$$

tel que : $g : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$ est la fonction du *permanent market impact* engendré par les actions du trader agressif, $\sigma > 0$ est la volatilité de l'actif.

Le prix d'exécution est donnée par :

$$\forall k \in \llbracket 1, N \rrbracket : \quad \tilde{S}_k = S_{k-1} - f(v_k)$$

tel que, $f : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$ est la fonction du *temporary market impact* qui est classiquement concave et croissante. Les fonctions choisies dans le travail original d'Almgren et Chriss [6] correspondent à un *market impact* (temporaire ou permanent) linéaire, i.e :

$$\exists \gamma, \eta, \alpha \in \mathbb{R}_+^*, \forall v \geq 0 : \quad f(v) := \alpha + \eta v \quad , \quad g(v) := \gamma v$$

Le *cash* obtenue par le trader à tout instant est donnée par :

$$\forall k \in \llbracket 1, N \rrbracket : \quad X_k = X_{k-1} + \tilde{S}_k v_k \Delta t$$

Finalement, le problème d'optimisation consiste à trouver le taux de liquidation optimale dans un certain sens qui nous permet de liquider la position Q . En effet, le trader peut chercher la solution qui maximise son utilité espérée ou celle qui minimise son coût (*Implementation shortfall*). Pour une fonction d'utilité de type CARA (*Constant Absolute Risk Aversion*), le problème se pose comme suit :

$$\sup_{v \in \mathcal{A}} \mathbb{E}(-\exp(-\beta X_N))$$

tel que $\beta > 0$ est l'aversion absolue au risque du trader. \mathcal{A} est l'ensemble des stratégies admissibles, i.e :

$$\mathcal{A} := \left\{ (v_k)_{k \in \llbracket 1, N \rrbracket} \text{ } \mathbb{F}\text{-prévisible et à valeurs dans } \mathbb{R}_+^* : \sum_{k=1}^N v_k = -\frac{Q}{\Delta t} \right\}$$

On pourra généralement considérer des modèles en temps continu plus sophistiqués que celui présenté ci-haut. À titre d'exemple, on peut considérer la dynamique suivante pour le processus du *mid-price* :

$$\forall t \in]0, T] : \quad dS_t = \sigma \left(dJ_t^+ - dJ_t^- \right)$$

tel que : $\sigma > 0$ est la taille du *tick* et $(J_t^\pm)_{t \in [0, T]}$ sont deux processus de Cox d'intensités stochastiques $(\kappa_t^\pm)_{t \in [0, T]}$ qui comptent le nombre des sauts positifs (+) et négatifs (-) du processus S .

1.2.3 Le trading dans le carnet d'ordre limite

Les modèles comme celui d'Almgren et Chris [6] permettent de résoudre facilement le problème de nature stratégique que l'on fait face durant la conception de l'algorithme de trading. Toutefois, la granularité de ces modèles n'est pas tellement fine afin qu'ils prennent en considération l'effet de la microstructure. Afin d'optimiser l'interaction avec le carnet d'ordre limite ou le *Dark pool* (problème tactique), on doit modéliser la manière avec laquelle la liquidité se comporte comme étant une fonction du placement des ordres (tout en prenant en considération le risque du marché mesuré par la volatilité de l'actif et le comportement asymptotique du processus du prix).

Market-Making à haute fréquence

D'un point de vue mathématique, le problème du *market-making* correspond aux choix des cotations optimales (i.e. les prix *bid/ask*) durant la session du trading afin de maximiser une utilité espérée en prenant en considération la dynamique du processus du *cash* et de l'inventaire. Pour construire un bon modèle, on doit prendre en considération deux aspects des prix :

- La probabilité qu'un ordre limite soit exécuté comme étant une fonction de la distance entre le prix proposé par le *market-maker* et le prix de référence.
- Une mesure du risque de marché qui correspond au coût de liquidation de l'inventaire à un mauvais prix à la fin de la session du trading.

L'approche qui avait eu beaucoup de succès en ce sens est celle de Avellaneda et Stoikov [9] qui ont proposé un contexte d'optimisation innovant pour aborder le problème du trading dans le carnet d'ordre limite. L'objectif le plus important de ce modèle est de considérer, dans un contexte unifié, les deux aspects importants pour un algorithme interagissant avec le carnet d'ordre : le contrôle de la probabilité qu'un ordre soit exécuté à court terme (*fill probability*) et le risque du prix durant la liquidation de l'inventaire restante, mesuré par la volatilité du son comportement à long terme. Brièvement, fixons un horizon de temps $T > 0$ et un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, tel que la filtration \mathbb{F} vérifie les conditions habituelles. On définit sur cet espace un mouvement brownien standard $(W_t)_{t \in [0, T]}$. La première brique du modèle et de préciser la dynamique du prix de référence (classiquement, le *mid-price*) que l'on note S . Avellaneda et Stoikov ont considéré la dynamique suivante :

$$\forall t \in]0, T] : \quad dS_t = \sigma dW_t \quad ; \quad S_0 = s > 0$$

avec $\sigma > 0$ est la volatilité de l'actif.

Sachant qu'un mouvement brownien peut prendre des valeurs négatives, cette dynamique ne peut être raisonnable que sur un horizon de temps court. Dans la pratique, la définition de ce qu'on appelle le prix de référence dépend des objectifs des applications. Notre objectif n'est pas la construction d'un modèle sophistiqué pour ce prix, mais de fixer une référence par rapport à laquelle on peut comparer les prix proposés par le *market-maker*. L'aspect le plus important sur lequel nous nous concentrons est le fait que ce prix ait une volatilité qui mesure le risque du à ses fluctuations.

Afin de modéliser la liquidité, l'idée est de considérer qu'un LO posté à une distance $\delta > 0$ par rapport au prix de référence est exécuté avec une probabilité donnée par l'intensité :

$$\lambda(\delta) := A \exp(-k\delta)$$

Heuristiquement, on comprends que la probabilité d'exécution d'un tel ordre dans un intervalle $[t, t + \Delta t]$ est de la forme : $A \exp(-k\delta)\Delta t$, indépendamment du passé.

Pour dériver cette forme particulière de l'intensité, Avellaneda et Stoikov se sont basé sur un argument théorique qui suppose que la relation entre le changement de prix et le *market impact* instantané est logarithmique. Toutefois, on peut justifier ce choix en se basant sur des arguments empiriques. En effet, on suppose que le nombre des LOs postés par le *market maker* et qui sont exécutés (c-à-d. les LOs ayant trouvé un MO comme contre-partie) jusqu'à l'instant t est modélisé par un processus de comptage associé à un processus ponctuel de poisson. N^+ (resp. N^-) désigne le processus comptant le nombre des LOs d'achat (resp. de vente) exécutés. Naturellement, ces processus ont des intensités $\lambda^\pm(\delta)$ qui dépendent de la distance δ entre le prix posté par le *market maker* est le prix de référence. La figure 1.9 illustre cette forme particulière de $\delta \mapsto \lambda^\pm(\delta)$. Ces graphiques sont construits en se basant sur les données du carnet d'ordre de l'action AMZN à la date 21/06/2012.

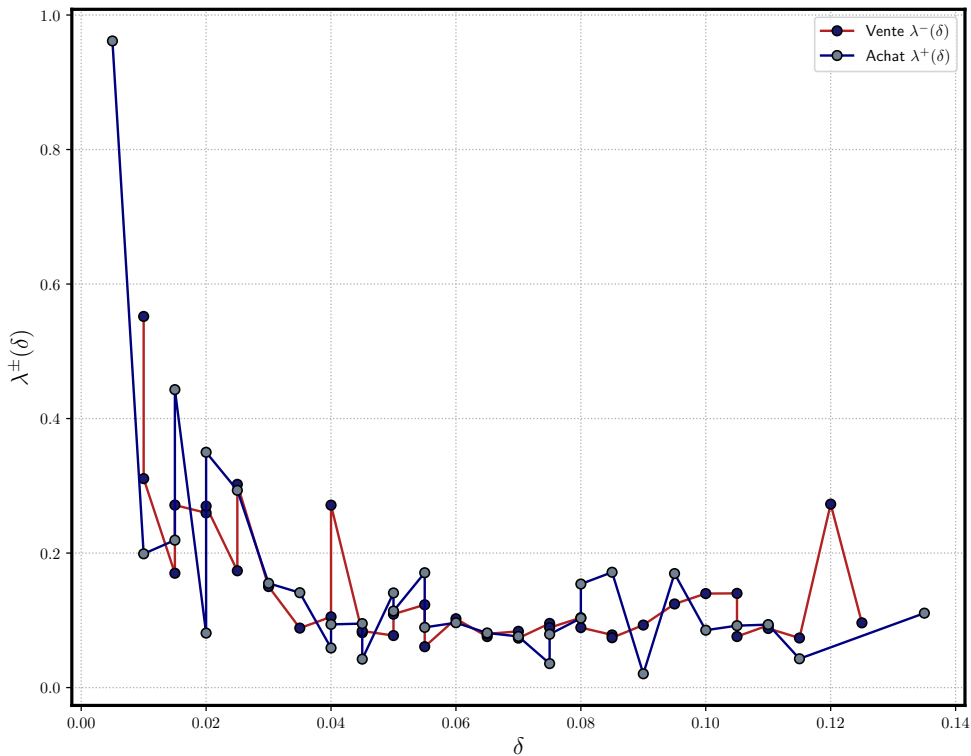


FIGURE 1.9 – Graphique des intensités des processus modélisant le nombre de LOs exécutés

Le processus d'inventaire du *market maker* est donné par :

$$\forall t \in [0, T] : \quad Q_t := N_t^+ - N_t^-$$

On note, $(\delta_t^\pm)_{t \in [0, T]}$ les processus modélisant les distances entre les prix postés par le *market maker* et le prix de référence avec du côté *bid* (+) et *ask* (-). Le processus du *cash* est donné par :

$$\forall t \in [0, T] : \quad dX_t = (S_t + \delta_t^-) dN_t^- + (S_t - \delta_t^+) dN_t^+$$

Ainsi, le problème du *market maker* se pose comme suit :

$$\sup_{(\delta^+, \delta^-) \in \mathcal{A}} \mathbb{E}(-\exp(-\gamma(X_T + Q_T S_T)))$$

tel que $\gamma > 0$ est l'aversion absolue au risque et \mathcal{A} est l'ensemble des processus \mathbb{F} - prévisibles et minorés.

Les avantages de ce modèle sont nombreux. En effet, le carnet d'ordre (la liquidité) apparaît comme un *black-box* caractérisé par quelques paramètres. Ce *black-box model* peut être calibré aux données du marché sous la conditions que les paramètres sous-jacents caractérisent la liquidité de l'actif dans le lieu d'échange d'intérêt. En plus, la dynamique à long terme du prix de référence est pilotée par un mouvement brownien, ce qui est consistant avec les modèles classiques du prix dans la finance mathématique. Finalement, les fonctions impliquées dans le modèle peuvent être exprimées analytiquement lorsqu'on aborde le problème par les techniques d'optimisation (contrôle stochastique optimal, approximation stochastique, etc). Plusieurs auteurs ont analysé les modèles du *market-making* d'une manière similaire. On cite le travail de Guéant et al. [55] dans lequel les auteurs ont complètement résolu le problème de Avellaneda et Stoikov et ont performé une analyse détaillée des solutions obtenues. Cartéa et al. [35] ont étendu le modèle à d'autres fonctions d'utilités et à des dynamiques du prix et des processus de comptage plus sophistiquées. Dans la même direction de recherche, Pham et al. [57] ont étudié le problème du trading avec les ordres limites et les ordres au marché. Les mêmes auteurs ont étudié dans [58] le problème des carnets d'ordres pro-rata.

Tactiques de la liquidation optimale

La littérature de la liquidation optimale est très proche à celle du *market-making* comme on peut la considérer, d'un point de vue tactique, comme un *market-making* qui agit d'un seul côté (i.e on propose des ordres limites que d'un seul côté du spread (*bid* ou *ask*)). Deux papiers ont, en particulier, présentés des travaux pionniers en ce sens. La première est due à Guéant et al. [56] et la deuxième à Bayraktar et al. [20]. La modélisation est approximativement la même que dans le contexte de Avellaneda et Stoikov.

Dans la même direction de recherche, plusieurs travaux ont été fait par Guéant et al. dans [62] qui ont continué l'étude de la liquidation optimale par les ordres limite en proposant des formes plus générales que celle présentées dans le papier original de Avellaneda et Stoikov. Ceci a été le point de départ de plusieurs domaines de recherche. Le premier consiste à traiter la liquidation optimale comme un "problème de pricing" [59], [60], [65] dans lequel les *brokers* garantissent à leurs clients un *benchmark* et cherchent à satisfaire des contraintes à un risque minimal. Ceci est similaire au contexte de *Black-Scholes* pour les produits dérivés. Le deuxième est l'étude du problème de la liquidation de produits plus complexes tels que les produits dérivés et les contrats ASR (*Accelerated Share Repurchase*), [63], [64], [61] afin d'étudier l'impact des coûts des transactions sur les stratégies de couverture optimale.

Une autre direction de recherche très intéressante que nous présenterons dans ce document est celle proposée par Laruelle et al. [89]. Cette approche considère comme point de départ, une version modifiée du contexte de Avellaneda et Stoikov et résout le problème non pas à l'aide des techniques de programmation dynamique (*backward*), mais en se basant sur des techniques *forward* d'approximation stochastique (*trial-and-error optimization*). Cette approche est carrément originale comme elle utilise une procédure itérative pour optimiser les tactiques du trading qui est plus flexible en termes de définition de la dynamique du prix et de l'exploitation des connaissances préalables des paramètres que celle basée sur les équations HJB.

Trading dans les dark pools

Le trading dans les *dark pools* est approximativement le même que dans le carnet d'ordre comme l'acheminement aux *venues* de liquidité est fait en prenant en considération le contexte

du trading global. Ce sujet n'a pas été trop abordé dans la littérature. Parmi les travaux qui traitent ce problème, on cite celui de Kratz et al. [79] et de Laruelle et al. [88].

1.3 Approches d'optimisation

Une fois les modèles du prix et de la liquidité sont définis, nous sommes en position de définir le problème d'optimisation qui doit être résolu. Dans ce document, nous nous concentrons sur deux situations : les tactiques du courtage algorithmique et le *market-making* à haute fréquence. Dans chacune de ces situations, la structure du problème est à peu près la même :

- Le but du tactique est de trouver le prix optimal à lequel on peut poster passivement un volume dans le carnet d'ordre limite dans le but de la consommation ou la fourniture de la liquidité.
- Les prix postés sont actualisés durant la journée d'une manière discrète ou continue en se basant sur notre connaissance croissante de la dynamique du marché et l'état de notre algorithme. De ce fait, le problème est dit : *multi-stage*.
- Mathématiquement, le but est de maximiser une utilité espérée qui dépend des paramètres de l'algorithme qui, en interagissant avec le marché, est sujet à l'influence de plusieurs facteurs aléatoires ; ils peuvent représenter un certain risque quand on connaît les lois statistiques qui gouvernent l'aléa ou une incertitude dans le cas échéant.

Comme nous allons le voir, deux types de raisonnement peuvent être adoptés quand on aborde ces problèmes d'optimisation : un raisonnement rétrograde (*backward*) en se basant sur le principe de la programmation dynamique et un raisonnement progressif (*forward*) menant à une technique d'apprentissage *on-line*.

1.3.1 Problème exploitation-versus-exploration

Lorsqu'on est entrain de résoudre un problème d'optimisation *multi-stage* en présence d'aléas, i.e lorsqu'une variable est actualisée dans chaque stage afin de contrôler un système dans l'objectif d'optimiser sa performance, deux problèmes majeures se manifestent :

- Exploitation : qui consiste à trouver une règle qui nous fournit, à chaque moment, l'action optimale du stage suivant en utilisant l'information qu'on possède entre les mains.
- Exploration : qui consiste à apprendre la nature sous-jacentes de l'aléa qui gouverne le système en explorant les différents états du monde tout en exécutant les actions optimales.

Comme nous l'avons déjà mentionné, deux raisonnements peuvent être adoptés pour s'attaquer à ce problème : un est rétrograde, l'autre est progressif. Laconiquement, l'idée de l'approche *backward* (rétrograde) est de caractériser l'aléa à l'aide d'un modèle mathématique/statistique préalablement conçu en se basant sur des informations sur la performance antérieure. De ce fait, l'espérance de la fonction impliquée dans l'optimisation peut être considérée connue ; elle peut être calculée explicitement ou obtenue par simulation. Dans ce cas, on peut définir (à l'aide du principe de la programmation dynamique) une équation déterministe de l'espérance courante/actuelle (le lecteur doit comprendre que cette espérance est une fonction du temps) du *payoff* final optimal. Ainsi, on peut directement appliquer le principe de la programmation dynamique qui nous fournit une relation rétrograde entre le *payoff* espéré optimal à la fin d'un stage et celui du stage précédent.

Il est à noter que dans cette dernière approche, la fonction qui relie l'information sur le passé au action future est **déterministe**, même si la trajectoire du système est aléatoire. La phase d'exploration est effectuée avant le lancement de l'algorithme et la méthodologie de l'optimisation s'attaque, ainsi, au problème d'exploitation seulement. Il s'agit d'une situation similaire à celle que l'on trouve dans le *pricing* des options dans laquelle le contrôle du système (i.e le Δ de l'option) peut être obtenu à l'aide d'une EDP.

Quant à l'approche *forward* (progressive), les problèmes d'exploration et d'exploitation sont résolus simultanément durant le temps d'exécution de l'algorithme, i.e il n'y a aucune spécification *à priori* sur les paramètres ou les lois qui pilotent l'aléa du système ; les données historiques ne définissent pas préalablement notre point de vue sur la dynamique du système.

De ce fait, le contrôle optimal du système est obtenu d'une manière itérative. Il est actualisé à l'aide de deux facteurs : l'état actuel du système et les innovations (le flux d'information). L'inconvénient de cette approche est le fait qu'il est difficile de constituer, *à priori*, une image globale d'ensemble du système jusqu'à la fin de la stratégie ; la fonction qui relie l'information sur le passé aux actions futures est identifiée d'une manière *on-line*. Dans cette approche, la dynamique de l'algorithme choisit les décisions futures en se basant sur le passé est **stochastique**.

Il est clair que la meilleure approche dépend de la nature du système. En effet, pour un système dont la dynamique ne change pas durant la journée, l'approche rétrograde est plus appropriée comme une phase d'exploration *on-line* est moins prioritaire. Cependant, lorsque la dynamique du système change durant la journée ou d'une journée à une autre, une approche plus adaptée au changement doit être instaurée et il est plus approprié d'utiliser les techniques d'apprentissage *on-line*. Idéalement, on veut combiner les deux approches à la fois comme c'est le cas pour quelques approches d'apprentissage par renforcement [73] ou de *dynamic policy programming* [10].

1.3.2 Programmation dynamique et équation HJB

Le principe de la programmation dynamique introduit dans le travail séminal de Bellman [21] est l'un des résultats mathématiques qui ont eu beaucoup d'influence durant le 20^{ème} siècle. Le principe est basée sur la simple observation du fait qu'une stratégie optimale de résoudre un problème *multi-stage* à $N + 1$ stages tel que $N \in \mathbb{N}$, qui commence d'un point $x \in X$ tel X est l'espace des états, peut être obtenue en deux étapes : (1) l'étape prochaine, (2) un problème *multi-stage* à N stages qui commence de $y \in X$ qui est l'état du système après l'application du contrôle durant le premier stage.

Pour illustrer ce point, supposons que nous nous intéressons à un système dont le processus des états $(x_k)_{k \in [0, N+1]}$ est déterministe à valeurs dans \mathbb{R}^d pour $d \in \mathbb{N}^*$ et évolue en temps discret selon l'équation de récurrence suivante :

$$\forall k \in [1, N + 1] : \quad x_k = x_{k-1} + g(k - 1, x_{k-1}, \varepsilon_{k-1})$$

avec $(\varepsilon_k)_{k \in [0, N]}$ est un processus de contrôle à valeurs dans \mathbb{R}^p tel que $p \in \mathbb{N}^*$. $g : \mathbb{N} \times \mathbb{R}^d \times \mathbb{R}^p \rightarrow \mathbb{R}^d$ est une fonction n'ayant aucune forme particulière.

Supposons que l'objectif du contrôle est de minimiser une fonction de coût $C : \mathbb{R}^d \times \mathbb{R}^{p \times (N+1)} \rightarrow \mathbb{R}$ de la forme :

$$C(x_0, (\varepsilon_k)_{k \in [0, N]}) := \Phi(x_N) + \sum_{k=0}^{N-1} R(k, x_k, \varepsilon_k)$$

tel que $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ et $R : \mathbb{N} \times \mathbb{R}^d \times \mathbb{R}^p \rightarrow \mathbb{R}$ sont respectivement la fonction du coût final et du coût courant.

Ce problème peut être vu comme un problème d'optimisation classique sur $\mathbb{R}^{p \times (N+1)}$. Toutefois, si p ou N sont larges, le problème va souffrir de ce qu'on appelle le *curse of dimensionality*. L'idée proposée par Bellman consiste à définir une procédure rétrograde afin de s'attaquer au problème de manière à garder une croissance linéaire de la complexité numérique en fonction du nombre des étapes. Le point clé est de représenter le problème défini sur $\mathbb{R}^{p \times (N+1)}$ par un ensemble de $N+1$ problèmes interdépendants définis sur \mathbb{R}^p . Pour ce faire, on introduit une fonction valeur appelée *optimal cost-to-go function* J définie par :

$$\forall n \in \llbracket 0, N-1 \rrbracket : \quad J(n, x_n) := \min_{(\varepsilon)_{k \in \llbracket n, N-1 \rrbracket}} \left(\Phi(x_N) + \sum_{k=n}^{N-1} R(k, x_k, \varepsilon_k) \right)$$

La quantité $J(n, x_n)$ la valeur optimale du coût si on considère que le système démarre de l'état x_n à l'instant n . On peut vérifier simplement les relations suivantes, pour tout $x \in \mathbb{R}^d$ et $n \in \llbracket 0, N-1 \rrbracket$:

$$\begin{aligned} J(N, x) &= \Phi(x) \\ J(n, x) &= \min_{\varepsilon \in \mathbb{R}^p} (R(n, x, \varepsilon) + J(n+1, x + g(n, x, \varepsilon))) \end{aligned}$$

L'algorithme de la programmation dynamique fonctionne de la manière suivante :

1. On pose $J(N, x) = \Phi(x)$.
2. On calcule (par récurrence rétrograde sur n) :

$$\begin{aligned} \tilde{\varepsilon}_n(x) &= \arg \min_{\varepsilon \in \mathbb{R}^p} (R(n, x, \varepsilon) + J(n+1, x + g(n, x, \varepsilon))) \\ J(n, x) &= (R(n, x, \tilde{\varepsilon}_n(x)) + J(n+1, x + g(n, x, \tilde{\varepsilon}_n(x)))) \end{aligned}$$

3. On utilise la condition initiale pour obtenir $\tilde{x}_{n+1} = \tilde{x}_n + g(n, \tilde{x}_n, \tilde{\varepsilon}_n(x))$ et le contrôle optimal est définie à l'instant n par $\tilde{u}_n := \tilde{u}_n(\tilde{x}_n)$.

Le dernier exemple est simpliste et ne sert que pour des raisons d'illustration qui permet au lecteur de gagner une intuition sur la structure de l'approche. Généralement, l'évolution des systèmes est purement stochastique et décrit en temps continu par un processus d'état de type : diffusion ou diffusion avec sauts. Une présentation rigoureuse de l'étude des problèmes de contrôle stochastique de ce type de processus peut être trouvée dans le chapitre 2.

1.3.3 Approximation stochastique et apprentissage on-line

L'approximation stochastique est la théorie sous-jacente de l'approche *forward*. L'idée est de représenter le contrôle optimal par un processus de Markov dont la partie stochastique est une fonction du flux d'informations (innovations). De ce fait, le contrôle actuel est actualisé afin qu'il converge au contrôle optimal qui représente le zéro d'une fonction h qui s'écrit sous la forme : $h(\theta) := \mathbb{E}(H(\theta, Y))$ tel que Y est un vecteur aléatoire de loi simulable à coût raisonnable et H est une fonction vérifiant certaines conditions que l'on spécifiera plus loin dans ce document.

D'une manière simple, si θ_n est le contrôle actuel à l'étape n , on veut concevoir un algorithme stochastique de la forme suivante :

$$\underbrace{\theta_{n+1}}_{\text{nouvelle valeur}} = \underbrace{\theta_n}_{\text{ancienne valeur}} - \underbrace{\gamma_{n+1}}_{\text{pas (poids)}} \times \underbrace{H(\theta_n, \overbrace{Y_{n+1}}^{\text{innovation}})}_{\text{correction basée sur le flux d'information entrant}}$$

De manière garantissant la convergence de la suite construite à notre valeur cible.

Il est à noter qu'on peut implémenter ce type d'algorithmes sans spécifier la dynamique stochastique du processus des innovations. Toutefois, pour des raisons garantissant la convergence de l'algorithme, les innovations doivent être i.i.d, stationnaires et ergodiques. Il existe dans la littérature plusieurs résultats généraux de convergence. On cite entre autres, les travaux de Laruelle et al. [90] (innovations moyennisantes) et de Schreck et al. [49] (innovations sous forme de chaînes de Markov contrôlées).

1.4 Plan et motivations

Le but de ce travail est l'utilisation des modèles de liquidation optimale et de *market-making* dans différentes directions liées à l'exécution optimale. Par ailleurs, on discutera les modèles utilisés de différents points de vue, tout en discutant les questions intéressantes d'un point de vue mathématique et pratique. En plus, on veut souligner à travers ce travail deux paradigmes d'optimisation intéressants (programmation dynamique et apprentissage *on-line*), l'importance des modèles *black-box* qui permettent de caractériser le comportement du marché à l'aide d'un nombre fini de paramètres qui sont estimés d'une manière *on-line* et l'importance des résultats mathématiques qui ne fournissent pas seulement des "nombres", mais également des formules explicites dans certains cas importants, ce qui nous permet d'étudier la nature des solutions, d'effectuer les statiques comparatives et de construire des approximations.

À travers cette étude, on se concentrera sur la manière d'intégrer les modèles et les techniques théoriques développées dans les applications pratiques. Par ailleurs, on testera les résultats obtenus sur des données simulées numériquement et des données réelles.

Cette étude est motivée, premièrement, par la nécessité de résoudre les problématiques de liquidation et d'acquisition de La Marocaine Vie. Ces problématiques nécessitent l'utilisation de techniques de trading qui sont basées majoritairement sur des disciplines mathématiques sophistiquées. Chose qui a motivé l'organisme d'accueil à nous recruter. De plus, pour répondre au désir de l'organisme à s'ouvrir sur le marché français, cette étude fournit à l'équipe d'investissement les différentes caractéristiques des marchés étrangers ainsi que différentes approches de trading systématique à haute fréquence qu'elle puisse utiliser dans ses activités futures. En somme, elle s'agit d'un travail de recherche qui sert à enrichir les techniques de *trading* du département d'accueil.

Contrôle stochastique optimal

2.1 Contrôle stochastique des processus de diffusion

On se donne d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, un horizon de temps T qui peut être fini ou infini. Pour $d \geq 1$, on définit sur cet espace un mouvement brownien W à valeurs dans \mathbb{R}^d . On muni cet espace de l'augmentation habituelle de la filtration canonique de W , que l'on note $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$. Pour $n \geq 1$, on définit l'ensemble $\mathbf{S} := [0, T[\times \mathbb{R}^n$, appelé **l'intérieur parabolique de l'espace d'état**. Son adhérence est : $\bar{\mathbf{S}} = [0, T] \times \mathbb{R}^n$ si T est fini et $\bar{\mathbf{S}} = \mathbf{S}$ sinon.

Pour $k \geq 1$, soit U une partie non-vide de \mathbb{R}^k . On désigne par \mathcal{U} l'ensemble des processus $(\nu_t)_{t \in [0, T]}$ \mathbb{F} -progressifs à valeurs dans U . Ces processus seront appelés les **processus de contrôle**.

On se donne de deux fonctions continues $\mu : \mathbf{S} \times U \rightarrow \mathbb{R}^n$, $\sigma : \mathbf{S} \times U \rightarrow \mathcal{M}_{n,d}(\mathbb{R})$ vérifiant les conditions de Lipschitz et de croissance sous-linéaire uniformes :

$$\exists K > 0, \forall (t, x, y, u) \in \mathbf{S} \times \mathbb{R}^n \times U : \quad |\mu(t, x, u) - \mu(t, y, u)| + |\sigma(t, x, u) - \sigma(t, y, u)| \leq K|x - y| \quad (2.1.1)$$

$$|\mu(t, x, u)| + |\sigma(t, x, u)| \leq K(1 + |x| + |u|) \quad (2.1.2)$$

Pour tout processus de contrôle, $\nu \in \mathcal{U}$, on considère l'EDS suivante :

$$dX_t = \mu(t, X_t, \nu_t) dt + \sigma(t, X_t, \nu_t) dW_t \quad ; \quad X_0 = x \in \mathbb{R}^n \quad (2.1.3)$$

On note :

$$\mathcal{U}_0 := \mathcal{U} \cap \mathbb{H}^2(\mathbb{F})$$

$$\mathcal{U}_0^t := \{\nu \in \mathcal{U}_0 : \nu \text{ est indépendant de } \mathcal{F}_t\}$$

Sous la condition 2.1.2, On a :

$$\forall x \in \mathbb{R}^n, \nu \in \mathcal{U}_0 : \quad \mathbb{E} \left(\int_0^T (|b(t, x, \nu_t)| + |\sigma(t, x, \nu_t)|^2) dt \right) < \infty \quad (2.1.4)$$

Sous les conditions 2.1.1, 2.1.2 et 2.1.4, il existe une unique solution forte de l'EDS 2.1.3 appartenant à $\mathbb{H}^2(\mathbb{F})$. De plus, elle vérifie :

$$\exists C_{K,T} > 0, \forall t \in [0, T[: \quad \mathbb{E} \left(\sup_{s \in [0, t]} |X_s^{0,x,\nu}|^2 \right) \leq C_{K,T} (1 + |x|^2) e^{C_{K,T}t}$$

Une telle solution est appelé : **diffusion contrôlée**, comme sa dynamique est pilotée par le processus de contrôle ν .

Soient $f, k : \mathbf{S} \times U \rightarrow \mathbb{R}$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}$ des fonctions tels que f et k soient continues et g soit $\mathcal{B}(\mathbb{R}^n)$ - mesurable. On suppose, de plus, que $\|k^-\|_\infty < \infty$ et :

$$\exists K > 0, \forall (t, x, u) \in \mathbf{S} \times U : \quad |f(t, x, u)| + |g(x)| \leq K (1 + |x|^2 + |u|) \quad (2.1.5)$$

On définit la **fonctionnelle de gain** $J : \bar{\mathbf{S}} \times \mathcal{U} \rightarrow \mathbb{R}$, quand elle est bien définie, par :

$$\forall (t, x, \nu) \in \bar{\mathbf{S}} \times \mathcal{U} : \quad J(t, x, \nu) := \mathbb{E} \left(\int_t^T \beta(t, s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta(t, T) g(X_T^{t,x,\nu}) \mathbf{1}_{\{T < \infty\}} \right),$$

tel que pour tout $s, t \in [0, T]$ avec $s \geq t$:

$$\beta(t, s) = \exp \left(- \int_t^s k(r, X_r^{t,x,\nu}, \nu_r) dr \right)$$

Dans le cas $T < \infty$, la condition 2.1.5 et l'uniforme bornitude de k^- assurent que J est bien définie pour tout $\nu \in \mathcal{U}_0$. Dans ce cas, **l'ensemble des contrôles admissibles** est \mathcal{U}_0 . Dans le cas contraire ($T = \infty$), cet ensemble sera définie par :

$$\mathcal{U}_0 := \left\{ \nu \in \mathcal{U} : \forall x \in \mathbb{R}^n, \mathbb{E} \left(\int_0^{+\infty} \beta(0, s) \left(1 + |X_s^{0,x,\nu}|^2 + |\nu_s| \right) ds \right) < \infty \right\}$$

Ainsi, le **problème de contrôle stochastique** est définie par :

$$\forall (t, x) \in \mathbf{S} : \quad V(t, x) := \sup_{\nu \in \mathcal{U}_0} J(t, x, \nu) \quad (2.1.6)$$

V est appelée la **fonction valeur**. Notre intérêt sera l'étude du comportement local de cette fonction à l'aide du **Principe de la Programmation Dynamique** (PPD) et **l'équation de Hamilton-Jacobi-Bellman** (HJB).

Remarque 2.1.1

Pour $(t, x) \in \mathbf{S}$, s'il existe $\hat{\nu}^{t,x} \in \mathcal{U}_0$ tel que $V(t, x) = J(t, x, \hat{\nu}^{t,x})$, alors $\hat{\nu}^{t,x}$ est appelée un **contrôle optimal**. Un processus ν de \mathcal{U}_0 adapté à la filtration canonique de X notée \mathbb{F}^X est appelée un contrôle de *Feedback*. Un processus ν de \mathcal{U}_0 s'écrivant sous la forme $\nu_t := \tilde{u}(t, X_t^{0,x,\nu})$ pour $t \geq 0$, avec $\tilde{u} : \mathbf{S} \rightarrow \mathbb{R}$ est une fonction $\mathcal{B}([0, T] \times \mathbb{R}^n)$ - mesurable, est appelée un contrôle Markovien. Les processus déterministes de \mathcal{U}_0 sont appelées les contrôles *Open loop*.

Le problème de contrôle 2.1.6 peut être réduit à l'étude du problème suivant :

$$\forall (t, x) \in \mathbf{S} : \quad V(t, x) := \sup_{\nu \in \mathcal{U}_0^t} J(t, x, \nu) \quad (2.1.7)$$

On peut même se ramener par une augmentation du processus d'état, que l'on note \tilde{X} , à l'étude du problème suivant :

$$\forall (t, x) \in \mathbf{S} : \quad \tilde{V}(t, x) := \sup_{\nu \in \mathcal{U}_0^t} \mathbb{E} \left(g \left(\tilde{X}_T^{t,x,\nu} \right) \right)$$

La première formulation 2.1.7 est dite : formulation de Bolza. La deuxième est dite : formulation de Mayer.

2.1.1 Principe de la programmation dynamique

Le principe de la programmation dynamique est un principe fondamental pour la théorie du contrôle stochastique. Il s'agit de l'approche la plus populaire de l'étude de ces problèmes. Le principe est déduit de l'idée suivante de R. Bellman : *"an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."* [21].

Mathématiquement, le principe s'énonce de la manière suivante :

$$\begin{aligned} \forall (t, x) \in \mathbf{S} : \quad V(t, x) &= \inf_{\theta \in \mathcal{T}_{[t,T]}} \sup_{\nu \in \mathcal{U}_t} \mathbb{E} \left(\int_t^\theta \beta(t, s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta(t, \theta) V(\theta, X_\theta^{t,x,\nu}) \mathbb{1}_{\{T < \infty\}} \right) \\ &= \sup_{\theta \in \mathcal{T}_{[t,T]}} \sup_{\nu \in \mathcal{U}_t} \mathbb{E} \left(\int_t^\theta \beta(t, s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta(t, \theta) V(\theta, X_\theta^{t,x,\nu}) \mathbb{1}_{\{T < \infty\}} \right) \end{aligned}$$

Le problème avec cette formulation est le fait que V n'est pas nécessairement mesurable. De ce fait, le terme à droite de la dernière égalité n'est pas bien défini. Plusieurs preuves ont été proposées pour ce principe, on cite entre autres : N. El Karoui et M. Jeanblanc [43] qui se sont basées sur des théorèmes délicats de sélection mesurable.

Dans notre présentation, on suit l'approche de N. Touzi et B. Bouchard [28] qui ont établi une version faible du PPD. Pour $(t, x) \in \bar{\mathbf{S}}$, on définit l'enveloppe semi-continue supérieurement (resp. inférieurement) de V , que l'on note V^* (resp. V_*), par :

$$V_*(t, x) := \liminf_{(t', x') \rightarrow (t, x)} V(t', x') \quad ; \quad V^*(t, x) := \limsup_{(t', x') \rightarrow (t, x)} V(t', x')$$

Rappelons que V^* et V_* sont finies si V est localement bornée.

Théorème 2.1.1.1: Principe de la programmation dynamique faible

On suppose que V est localement bornée et on fixe $(t, x) \in \mathbf{S}$. Soit $(\theta^\nu)_{\nu \in \mathcal{U}_t}$ une famille de $\mathcal{T}_{[t,T]}$ indépendante de \mathcal{F}_t . Alors :

$$V(t, x) \leq \sup_{\nu \in \mathcal{U}_t} \mathbb{E} \left(\int_t^{\theta^\nu} \beta(t, s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta(t, \theta^\nu) V^*(\theta^\nu, X_{\theta^\nu}^{t,x,\nu}) \right)$$

Supposons de plus que g est semi-continue inférieurement et $(X_{s \wedge \theta^\nu}^{t,x,\nu})_{s \in [t,T]}$ est essentiellement borné dans $(\text{borné dans } \mathbb{L}^\infty)$, alors :

$$V(t, x) \geq \sup_{\nu \in \mathcal{U}_t} \mathbb{E} \left(\int_t^{\theta^\nu} \beta(t, s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta(t, \theta^\nu) V_*(\theta^\nu, X_{\theta^\nu}^{t,x,\nu}) \right)$$

Démonstration. Voir section 3.2.2 de [116].

□

Remarque 2.1.1.1

- Comme V_* (resp. V^*) est semi-continue inférieurement (resp. supérieurement), il est $\mathcal{B}([0, T] \times \mathbb{R}^n)$ - mesurable. Donc, les termes à droite des inégalités du théorème 2.1.1.1 sont bien définis.
- Si V est continue, on a $V = V_* = V^*$ et le PPD faible se réduit au PPD classique :

$$V(t, x) = \inf_{\theta \in \mathcal{T}_{[t, T]}} \sup_{\nu \in \mathcal{U}_t} \mathbb{E} \left(\int_t^\theta \beta(t, s) f(s, X_s^{t, x, \nu}, \nu_s) ds + \beta(t, \theta) V(\theta, X_\theta^{t, x, \nu}) \right)$$

2.1.2 L'équation de Hamilton-Jacobi-Bellman

L'équation de Hamilton-Jacobi-Bellman (HJB) ou l'équation de la programmation dynamique est la version infinitésimale du principe de la programmation dynamique : elle décrit le comportement local de la fonction valeur V .

L'Hamiltonien associé au problème du contrôle 2.1.6 est la fonction $\mathcal{H} : \mathbf{S} \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}_n \rightarrow \overline{\mathbb{R}}$ définie pour tout $(t, x, r, p, \gamma) \in \mathbf{S} \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}_n$ par :

$$\mathcal{H}(t, x, r, p, \gamma) := \sup_{u \in U} \left\{ -k(t, x, u)r + \langle \mu(t, x, u), p \rangle + \frac{1}{2} \text{Tr}(\sigma \sigma^\top(t, x, u) \gamma) + f(t, x, u) \right\}$$

Proposition 2.1.2.1

Supposons que $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$. Alors, V est une sur-solution classique de l'équation HJB, i.e :

$$(\forall (t, x) \in \mathbf{S}) : -\partial_t V(t, x) - \mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) \geq 0$$

Si $T < \infty$, on lui ajoute la condition terminale :

$$V(T, \cdot) \geq g$$

Démonstration. Voir proposition 3.4 de [116]. □

Proposition 2.1.2.2

Supposons que $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$ et pour tout $(t, x) \in \mathbf{S}$:

$$\mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) < \infty$$

On suppose de plus que k est bornée et que \mathcal{H} est semi-continue supérieurement. Alors, V est une sous-solution classique de l'équation HJB, i.e :

$$(\forall (t, x) \in \mathbf{S}) : -\partial_t V(t, x) - \mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) \leq 0$$

Si $T < \infty$, on lui ajoute la condition terminale :

$$V(T, \cdot) \leq g$$

Démonstration. Voir proposition 3.5 de [116]. □

Théorème 2.1.2.1

Supposons que les conditions des propositions 2.1.2.1 et 2.1.2.2 soient satisfaites. Alors V est une solution classique de l'équation HJB, i.e :

$$(\forall (t, x) \in \mathbf{S}) : -\partial_t V(t, x) - \mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) = 0 \quad (2.1.8)$$

Si $T < \infty$, on lui ajoute la condition terminale, :

$$V(T, \cdot) = g$$

2.1.3 Sur la régularité de la fonction valeur

L'objectif de cette section est de démontrer que la fonction valeur n'est pas généralement une fonction lisse. On commence par établir la continuité de la fonction valeur sous des conditions fortes sur les coefficients du problème. Pour simplifier, on adopte la formulation de Mayer ($f = k \equiv 0$) du problème 2.1.6 et on se concentre sur le cas à horizon fini ($T < \infty$).

Proposition 2.1.3.1

Supposons que g est lipschitzienne. Alors,

- V est lipschitzienne en fonction de x uniformément en t .
- Supposons de plus que U est borné. Alors, V est $\frac{1}{2}$ -höldérienne en t et :

$$\exists C > 0, \forall (t, s, x) \in [0, T]^2 \times \mathbb{R}^n : |V(t, x) - V(s, x)| \leq C(1 + |x|) \sqrt{|t - s|}$$

Démonstration. Voir proposition 3.7 de [116]. □

Pour des exemples de fonctions valeurs non différentiables, le lecteur est renvoyé aux sections 3.4.2 et 3.4.3 de [116].

2.1.4 Théorème de vérification

L'étape cruciale dans l'approche classique de la programmation dynamique consiste à montrer, étant donnée une solution régulière à l'équation d'HJB, que ce candidat, sous des conditions suffisantes, coïncide avec la fonction valeur. Ce résultat est appelé théorème de vérification et permet aussi d'obtenir un contrôle optimal. Il repose essentiellement sur la formule d'Itô.

pour $u \in U$, on note \mathcal{L}^u le générateur infinitésimal de la diffusion contrôlée $(\beta^u(0, t) X_t^{0,x,u})_{t \in [0, T]}$ défini pour tout $\varphi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$ par :

$$\forall (t, x) \in [0, T] \times \mathbb{R}^n : \mathcal{L}^u \varphi(t, x) = -k(t, x, u) \varphi(t, x) + \langle \mu(t, x, u), \nabla_x \varphi(t, x) \rangle + \frac{1}{2} \text{Tr} \left(\sigma \sigma^\top(t, x, u) H_x \varphi(t, x) \right)$$

Théorème 2.1.4.1: Vérification

Supposons que $T < \infty$. Soit $v \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n) \cap \mathcal{C}^0([0, T] \times \mathbb{R}^n)$. On suppose que f et v sont de croissance sous-quadratique, i.e :

$$\exists \kappa > 0, \forall (t, x, u) \in \mathbf{S} \times U : |f(t, x, u)| + |v(t, x)| \leq \kappa(1 + |x|^2 + |u|)$$

1. Si, v est une sur-solution du problème :

$$\begin{cases} -\partial_t V(t, x) - \mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) = 0 & (t, x) \in [0, T[\times \mathbb{R}^n \\ V(T, x) = g(x) & x \in \mathbb{R}^n \end{cases} \quad (2.1.9)$$

alors, $v \geq V$ dans $\bar{\mathbf{S}}$.

2. Supposons que v est une solution du problème 2.1.9. Pour tout $(t, x) \in \mathbf{S}$, on suppose qu'il existe un minimiseur $\tilde{u}(t, x)$ de $u \mapsto \mathcal{L}^u v(t, x) + f(t, x, u)$ sur U vérifiant :

- $\partial_t v(t, x) + \mathcal{L}^{\tilde{u}(t, x)} v(t, x) + f(t, x, \tilde{u}(t, x)) = 0$
- l'EDS suivante admet une unique solution au sens de l'unicité trajectorielle :

$$dX_s = \mu(s, x, \tilde{u}(s, X_s)) ds + \sigma(s, x, \tilde{u}(s, X_s)) dW_s \quad ; \quad X_t = x$$

- Le processus \tilde{v} défini par : $\tilde{v}_s := \tilde{u}(s, X_s^{t, x})$ pour $s \in [t, T[$ est un processus de contrôle markovien bien défini et appartient à \mathcal{U}_0 .

Alors, $v = V$ et \tilde{v} est une processus de contrôle markovien optimal.

Démonstration. Voir théorème 5.1 de [116]. □

Remarquons que si l'ensemble des contrôles U est réduit à un singleton, le problème de contrôle est dégénéré et le théorème de vérification est réduit à la fameuse représentation de *Feynman-Kac*. Remarquons de plus, que le théorème de vérification suppose l'existence d'une solution à l'équation HJB 2.1.9. Toutefois, il nous assure l'unicité d'une telle solution dans la classe des fonctions à croissance sous-quadratique (en x) uniformément (en t). On donnera sans preuve les conditions assurant l'existence d'une solution à l'équation HJB 2.1.9. On note $\mathcal{C}_b^{p, k}([0, T[\times \mathbb{R}^n)$ l'ensemble des fonctions bornées ayant des dérivées partielles par rapport à t d'ordres inférieures au égaux à p et des dérivées partielles par rapport à x d'ordres inférieures au égaux à k qui sont continues et bornées.

Théorème 2.1.4.2

On suppose que $\sigma \sigma^\top$ vérifie la condition de **parabolicité uniforme** :

$$\exists \varepsilon > 0, \forall (t, x, u) \in [0, T] \times \mathbb{R}^n \times U, \forall \xi \in \mathbb{R}^n \quad \xi^\top \sigma \sigma^\top(t, x, u) \xi \geq \varepsilon |\xi|^2$$

Supposons de plus que :

- U est compact.
- pour tout $u \in U$, $\mu(\cdot, \cdot, u)$, $\sigma(\cdot, \cdot, u)$ et $f(\cdot, \cdot, u)$ sont dans $\mathcal{C}_b^{1,2}([0, T] \times \mathbb{R}^n)$.
- $g \in \mathcal{C}_b^3(\mathbb{R}^n)$.

Alors, l'équation HJB 2.1.9 a une unique solution $V \in \mathcal{C}_b^{1,2}([0, T] \times \mathbb{R}^n)$.

2.2 Contrôle stochastique des diffusions à sauts

On se place dans le cadre présenté dans le début de la section 2.1. On s'intéresse dans cette section à l'étude des problèmes de contrôle stochastique dans le cadre des diffusions à sauts. Le lecteur est renvoyé aux livres suivants : [8], [32], [70], [71], [37] et [38] pour un rappel sur la

théorie des mesures aléatoires et les processus ponctuels.

On définit la tribu prévisible $\mathcal{P}(\mathbb{F})$ sur $[0, T] \times \Omega$ comme étant la tribu engendrée par la famille d'ensembles suivante :

$$\mathcal{R} := \{[s, t] \times F : 0 \leq s \leq t \leq T, F \in \mathcal{F}_s\} \cup \{\{0\} \times F : F \in \mathcal{F}_0\}$$

Un processus $(X_t)_{t \in [0, T]}$ est dit \mathbb{F} - prévisible si :

$$(t, \omega) \mapsto X(t, \omega) \text{ est } \mathcal{P}(\mathbb{F}) \text{ - mesurable}$$

Si (K, \mathcal{K}) est un espace mesurable et $H : ([0, T] \times \Omega \times K, \mathcal{P}(\mathbb{F}) \otimes \mathcal{K}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ est mesurable, alors le processus $(H(t, z))_{(t, z) \in [0, T] \times K}$ est dit un processus \mathbb{F} - prévisible indexé par K .

Soit N une mesure aléatoire simple et localement finie sur $[0, T] \times \mathbb{R}^d$ à base dans $[0, T]$ et marquée sur \mathbb{R}^d définie par

$$\forall [t_1, t_2] \times A \in \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^d) : N([t_1, t_2] \times A) = \sum_{n=0}^{+\infty} \mathbb{1}_{[t_1, t_2]}(T_n) \mathbb{1}_A(Z_n)$$

Tel que $(T_n, Z_n)_{n \in \mathbb{N}}$ est le processus ponctuel à valeurs dans $[0, T] \times \mathbb{R}^d$ associé à N .

Pour $t \in [0, T]$, $A \in \mathcal{B}(\mathbb{R}^d)$, on introduit la notation :

$$N(t, A) := N([0, t] \times A)$$

On suppose de plus que N admet un noyau prévisible $\tilde{\lambda}$ de la forme :

$$\forall (t, A) \in [0, T] \times \mathcal{B}(\mathbb{R}^d) : \tilde{\lambda}(t, A) = \lambda(t) \Phi(A)$$

avec Φ est une distribution de probabilité sur \mathbb{R}^d et $t \mapsto \lambda(t)$ est une fonction positive, déterministe, bornée et lipschitzienne sur $[0, T]$. On vérifie bien que λ est l'intensité de N .

On définit la mesure aléatoire compensée associée à N par :

$$\tilde{N}(dt \otimes dz) := N(dt \otimes dz) - \tilde{\lambda}(t, dz) \otimes dt$$

Soit U une partie non vide de \mathbb{R}^d . L'ensemble des contrôles admissibles \mathcal{U}_0 est constituée cette fois-ci des processus ν à valeurs dans U , \mathbb{F} - prévisibles, localement bornés et tel que $\nu \in L^2([0, T] \times \Omega)$.

Soient μ et σ les fonctions définies dans la section 2.1 vérifiant la condition de Lipschitz globale 2.1.1 et celle de la croissance sous linéaire 2.1.2. Soit $\rho : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ une fonction continue vérifiant : $\exists K > 0, \forall (t, x, y, u, v, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times U^2 \times \mathbb{R}^d$,

$$\begin{aligned} |\rho(t, x, u, z) - \rho(t, y, v, z)| &\leq K(|x - y| + |u - v|) \\ |\rho(t, x, u, z)| &\leq L(1 + |x| + |u|) \end{aligned}$$

Pour $\nu \in \mathcal{U}_0$, on considère l'EDS suivante :

$$dX_t = \mu(t, X_t, \nu_t) dt + \sigma(t, X_t, \nu_t) dW_t + \int_{\mathbb{R}^d} \rho(t, X_t, \nu_t, z) N(t, dz) \quad ; \quad X_0 = x \in \mathbb{R}^d \quad (2.2.1)$$

Sous les conditions 2.1.1, 2.1.2 et celle imposée sur ρ , l'EDS 2.2.1 admet une unique solution forte [8] vérifiant :

$$\exists C_{K,T} > 0, \forall t \in [0, T]: \quad \mathbb{E} \left(\sup_{s \in [0, t]} |X_s^{0, x, \nu}|^2 \right) \leq C_{K,T} (1 + |x|^2)$$

On s'intéresse comme dans la section précédente 2.1 aux problèmes du type :

$$\forall (t, x) \in \mathbf{S} : \quad V(t, x) := \sup_{\nu \in \mathcal{U}_0^t} J(t, x, \nu)$$

Tel que J est la fonctionnelle de gain définie dans 2.1.5 et \mathbf{S} est l'intérieur parabolique de l'espace d'état. Les fonctions f, g, k impliquées dans la définition de J vérifient les mêmes conditions de la section 2.1.

Il est à noter que les résultats généraux restent les mêmes que dans le cas des diffusions continues. En effet, le principe de la programmation dynamique s'énonce de la même manière que dans 2.1.1.1. Quant à l'équation HJB 2.1.2.1, le théorème reste valide dans ce cadre. Toutefois, l'expression de l'Hamilton associé au problème est définie dans ce cas, pour tout $v \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$, $(t, x) \in \mathbf{S}$

$$\mathcal{H}(t, x, v(t, x), \nabla_x v(t, x), H_x v(t, x)) := \sup_{u \in U} \{-k(t, x, u)v(t, x) + \mathcal{L}^u v(t, x) + f(t, x, u)\}$$

Tel que \mathcal{L}^u est l'opérateur différentiel défini, pour tout $(t, x) \in \mathbf{S}$ par :

$$\begin{aligned} \mathcal{L}^u v(t, x) &:= \langle \mu(t, x, u), \nabla_x v(t, x) \rangle + \frac{1}{2} \text{Tr} \left(\sigma \sigma^\top(t, x, u) H_x v(t, x) \right) \\ &+ \int_{\mathbb{R}^d} (v(t, x + \rho(t, x, u, z)) - v(t, x)) \lambda(t) \Phi(dz) \end{aligned}$$

Le théorème de vérification s'énonce dans ce cadre de la manière suivante présenté dans le théorème 2.2.1. Pour plus de détails sur la théorie du contrôle stochastique, les problèmes de la cible aléatoire, etc, le lecteur est renvoyé à [27] et [101]. Il est à noter que nous n'avons pas introduit le concept important de la solution de viscosité. Sur ce, on revoit le lecteur à [116] et [27] pour une présentation de ce concept.

Théorème 2.2.1: Vérification

Supposons que $T < \infty$. Soit $v \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d) \cap \mathcal{C}^0([0, T] \times \mathbb{R}^d)$ une solution de l'équation HJB :

$$-\partial_t V(t, x) - \mathcal{H}(t, x, V(t, x), \nabla_x V(t, x), H_x V(t, x)) = 0$$

Vérifiant :

$$\forall x \in \mathbb{R}^d : \quad \liminf_{(t, x') \rightarrow (T, x)} v(t, x') \geq g(x)$$

et v^- est à croissance sous-quadratique. Supposons de plus que pour tout $(t, x) \in \mathbf{S}$, il existe un minimisateur $\tilde{u}(t, x)$ de $u \mapsto -k(t, x, u)v(t, x) + \mathcal{L}^u v(t, x) + f(t, x, u)$ sur U vérifiant :

$$\bullet \quad \partial_t v(t, x) - k(t, x, \tilde{u}(t, x))v(t, x) + \mathcal{L}^{\tilde{u}(t, x)} v(t, x) + f(t, x, \tilde{u}(t, x)) = 0$$

- l'EDS suivante admet une unique solution au sens de l'unicité trajectorielle :

$$dX_s = \mu(s, x, \tilde{\nu}_s) ds + \sigma(s, x, \tilde{\nu}_s) dW_s + \int_{\mathbb{R}^d} \rho(s, X_s, \tilde{\nu}_s, z) N(t, dz) \quad ; \quad X_t = x$$

tel que : $\tilde{\nu}_s := \tilde{u}(s, X_s^{t,x})$ pour $s \in [t, T[$.

- Il existe une suite $(\theta_k)_{k \in \mathbb{N}}$ dans $\mathcal{T}_{t,T}$ croissante, vérifiant :

$$\theta_k \xrightarrow[k \rightarrow +\infty]{} T \quad \text{P.s}$$

et pour tout $k \in \mathbb{N}$, $(X_{s \wedge \theta_k}^{t,x,\tilde{\nu}})_{s \in [t,T]}$ est essentiellement borné, tel que :

$$\begin{aligned} & \mathbb{E} \left(\int_t^{\theta_k} \beta^{\tilde{\nu}}(t, s) f(s, X_s^{t,x,\tilde{\nu}}, \tilde{\nu}_s) ds + \beta^{\tilde{\nu}}(t, \theta_k) v(\theta_k, X_{\theta_k}^{t,x,\tilde{\nu}}) \right) \\ & \xrightarrow[k \rightarrow +\infty]{} \mathbb{E} \left(\int_t^T \beta^{\tilde{\nu}}(t, s) f(s, X_s^{t,x,\tilde{\nu}}, \tilde{\nu}_s) ds + \beta^{\tilde{\nu}}(t, T) g(X_T^{t,x,\tilde{\nu}}) \right) \end{aligned}$$

Alors, $v = V$ et $\tilde{\nu}$ est un processus de contrôle markovien optimal.

Démonstration. Voir théorème 2.2.4 de [27]. □

2.3 Résolution numériques des problèmes de contrôle

Trouver une solution analytique à un problème de contrôle stochastique est extrêmement difficile et rare. Sur ce, dans la majorité des cas ces problèmes sont résolus à l'aide des approximations numériques. Ceci a encouragé beaucoup de recherches en ce sens. Plusieurs méthodes numériques ont été développées qui sont notamment motivées par des problèmes rencontrés dans la finance quantitative. À titre d'exemple, on rencontre souvent des problèmes numériques très difficiles lorsqu'on traite l'optimisation du portefeuille en grande dimension sous l'information partielle ou en prenant en considération les coûts des transactions, etc.

On distingue généralement entre deux types de méthodes numériques de résolution des problèmes du contrôles, à savoir : les méthodes déterministes et probabilistes.

2.3.1 Méthodes numériques déterministes

Les méthodes purement déterministe appartiennent à l'analyse numérique et consistent en la discrétisation de l'équation non linéaire d'Hamilton-Jacobi-Bellman à l'aide des schémas de différences finies, d'éléments finies ou de volumes finis. Cette discrétisation permet, alors, d'approximer la fonction valeur sur une grille espace-temps. La convergence des schémas numériques a été démontré à l'aide des résultats de stabilité des solutions de viscosité [16]. Des résultats concernant les taux de convergence des schémas de différences finies de l'équation HJB peuvent être trouvés dans [80] et [17]. Pour quelques applications de ces méthodes à des problèmes de finance quantitative, le lecteur est renvoyé à [48] et [115].

D'un point de vue computationnel, l'inconvénient majeur de ces méthodes et qu'elles deviennent très coûteux (au sens temporel) dès que la dimension du processus d'état augmente. Dans la pratique, ces méthodes ne sont utilisées que pour de faibles dimension, e.g. 1 ou 2.

2.3.2 Méthodes numériques probabilistes

Les méthodes numériques probabilistes de la résolution des problèmes de contrôle stochastique ou d'optimisation ont connu plusieurs développements durant les dernières années. Une méthode classique basée sur le principe de la programmation dynamique a été développée par Kushner [81], à savoir : *Markov Chain Approximation Method*. Elle consiste en l'approximation du processus initial contrôlé qui est à temps continu par une chaîne de Markov contrôlée vérifiant quelques conditions de consistance locale. Ainsi, le problème du contrôle stochastique est résolu numériquement par un algorithme de récurrence *backward* obtenu par l'application du principe de la programmation dynamique à la chaîne de Markov contrôlée. Toutefois, les résultats de stabilité peuvent être très restrictives même pour le cas d'une diffusion contrôlée. De plus, la chaîne est défini à l'aide d'un arbre (*lattice*) dont la taille croît exponentiellement avec la dimension du problème. Sur ce, ces méthodes, comme les méthodes déterministes, ne sont utilisées que pour les faibles dimensions. Pour surmonter ce problème, Pham, Pagès et Printemps [105] ont développé une méthode basée sur la quantification optimale pour résoudre numériquement les problèmes de contrôle stochastique pour une dimension supérieure à 3. On présentera par la suite quelques méthodes numériques probabilistes que nous allons utiliser dans la résolution des problèmes de contrôle issus des problèmes du *trading* optimal que nous allons rencontrer dans le chapitre suivant.

On se place dans le cadre probabiliste de la section 2.1. Pour simplifier, supposons que l'horizon T est fini. Considérons une diffusion Z^α à valeurs dans \mathbb{R}^d contrôlée par un processus de contrôle α à valeurs dans $A \subset \mathbb{R}^d$ qui vérifie l'EDS suivante :

$$dZ_t^\alpha = b(Z_t^\alpha, \alpha_t) dt + \sigma(Z_t^\alpha, \alpha_t) dW_t \quad ; \quad Z_0^\alpha = z_0 \in \mathbb{R}^d \quad (2.3.1)$$

La fonctionnelle de gain est définie, pour $(t, z) \in [0, T] \times \mathbb{R}^d$ par :

$$J(t, z, \alpha) := \mathbb{E} \left(\int_t^T f(Z_s^{t,z,\alpha}, \alpha_s) ds + g(Z_T^{t,z,\alpha}) \right)$$

On note par \mathcal{A} l'ensemble des contrôles $\alpha \in \mathbb{H}^0(\mathbb{F})$ à valeurs dans A , tel que :

$$\forall (t, z) \in [0, T] \times \mathbb{R}^d : \quad \mathbb{E} \left(\int_t^T f^+(Z_s^{t,z,\alpha}, \alpha_s) ds + g^+(Z_T^{t,z,\alpha}) \right) < \infty$$

Pour alléger les notations, on omet de noter la dépendance de Z en α dans la notation Z^α .

Soient $N \in \mathbb{N}^*$, $\Delta t := T/N$ et $t_k = k\Delta t$ pour $k \in \llbracket 0, N \rrbracket$ une subdivision de $[0, T]$. On note $\mathcal{A}_{\Delta t}$ l'ensemble des processus à temps discret $(\alpha_{t_k})_{k \in \llbracket 0, N \rrbracket}$, tel que pour tout $k \in \llbracket 0, N \rrbracket$: α_{t_k} est \mathcal{F}_{t_k} -mesurable. Pour $\alpha \in \mathcal{A}_{\Delta t}$, on introduit la discrétisation de l'EDS 2.3.1 selon le schéma d'Euler-Maruyama :

$$\forall k \in \llbracket 0, N-1 \rrbracket : \quad Z_{t_{k+1}} = Z_{t_k} + b(Z_{t_k}, \alpha_{t_k}) \Delta t + \sigma(Z_{t_k}, \alpha_{t_k}) \Delta W_{t_k} \quad (2.3.2)$$

tel que $\Delta W_{t_k} \rightsquigarrow \mathcal{N}_d(0, \Delta t I_d)$ est un incrément du mouvement brownien W .

Pour $z \in \mathbb{R}^d$, $n \in \llbracket 0, N-1 \rrbracket$ et $\alpha \in \mathcal{A}_{\Delta t}$ l'approximation en temps discret de $J(t_n, z, \alpha)$ est définie par :

$$J_{\Delta t}(t_n, z, \alpha) := \mathbb{E} \left(\sum_{k=n}^{N-1} f(Z_{t_k}^{t_n, z}, \alpha_{t_k}) \Delta t + g(Z_{t_N}^{t_n, z}) \right)$$

De ce fait, l'approximation en temps discrète de $V(t_n, z)$ est :

$$V_{\Delta t}(t_n, z) := \sup_{\alpha \in \mathcal{A}_{\Delta t}} J_{\Delta t}(t_n, z, \alpha)$$

Pour $(n, z, a) \in \llbracket 0, N \rrbracket \times \mathbb{R}^d \times A$, on note $\mathbb{E}_{n,z}^a$ l'opérateur espérance conditionnelle par rapport à l'événement $\{Z_{t_n} = z, \alpha_{t_n} = a\}$. Pour $z \in \mathbb{R}^d$, le principe de la programmation dynamique affirme que $(V_{\Delta t}(t_n, z))_{n \in \llbracket 0, N \rrbracket}$ est solution de l'équation de Bellman :

$$\begin{cases} V(t_N, z) = g(z) \\ V(t_n, z) = \sup_{a \in A} \left\{ f(z, a) \Delta t + \mathbb{E}_{n,z}^a \left(V(t_{n+1}, Z_{t_{n+1}}) \right) \right\} ; \quad n \in \llbracket 0, N-1 \rrbracket \end{cases} \quad (2.3.3)$$

Remarquer que l'équation 2.3.3 fournit une procédure itérative *backward* qui permet de calculer $(V_{\Delta t}(t_n, z))_{n \in \llbracket 0, N \rrbracket}$ à condition que nous savons calculer/approximer l'espérance conditionnelle figurant dans l'équation de Bellman. Cette procédure itérative et connue dans la littérature sous le nom de : *value iteration*.

Une autre manière qui permet de calculer $(V_{\Delta t}(t_n, z))_{n \in \llbracket 0, N \rrbracket}$ consiste à remarquer, grâce au principe de la programmation dynamique, que $(V_{\Delta t}(t_n, z))_{n \in \llbracket 0, N \rrbracket}$ est solution de :

$$\begin{cases} V(t_N, z) = g(z) \\ V(t_n, z) = \sup_{a \in A} \left\{ f(z, a) \Delta t + \mathbb{E}_{n,z}^a \left(\sum_{k=n+1}^{N-1} f(Z_{t_k}^*, \alpha_{t_k}^*) \Delta t + g(Z_{t_N}^*) \right) \right\} ; \quad n \in \llbracket 0, N-1 \rrbracket \end{cases} \quad (2.3.4)$$

tel que pour tout $n \in \llbracket 0, N-1 \rrbracket$ et tout $k \in \llbracket n+1, N-1 \rrbracket$, $\alpha_{t_k}^* = \alpha^*(t_k, Z_{t_k})$ est le contrôle optimal au temps t_k avec α^* est une application mesurable définie par :

$$\alpha^*(t_k, z) = \arg \max_{a \in A} \left\{ f(z, a) \Delta t + \mathbb{E}_{n,z}^a \left(\sum_{j=k+1}^{N-1} f(Z_{t_j}^*, \alpha_{t_j}^*) \Delta t + g(Z_{t_N}^*) \right) \right\}$$

et tel que $(Z_{t_k}^*)_{k \in \llbracket n, N \rrbracket}$ est le processus démarrant à t_n de z (i.e, $Z_{t_n} = z$) contrôlé par le contrôle α définie par :

$$\alpha_{t_k} := a \mathbb{1}_{\{t_k=t_n\}} + \sum_{j=k+1}^{N-1} \alpha_{t_j}^* \mathbb{1}_{\{t_k=t_j\}}$$

Il est à noter que, comme la première équation, le système 2.3.4 fournit une procédure *backward* permettant de calculer $(V_{\Delta t}(t_n, z))_{n \in \llbracket 0, N \rrbracket}$ si on possède une expression de l'espérance conditionnelle figurant dans le système. Cette dernière procédure est appelée dans la littérature : *performance iteration*. Elle est appelée également la *policy iteration* dans la littérature du *reinforcement learning*.

À l'exception de quelques cas triviaux, une expression analytique des espérances conditionnelles figurant dans les deux systèmes 2.3.3 et 2.3.4 ne peut pas être obtenue. De ce fait, on doit les approximer numériquement, ce qui constitue la difficulté majeur que l'on fait face lorsqu'on implémente ces méthodes pour approximer la fonction valeur. Dans la section suivante, on présente différentes méthodes permettant d'approximer une espérance conditionnelle afin d'en déduire l'approximation de la fonction valeur.

Approximation des espérances conditionnelles

Dans cette sous section, on présente trois algorithmes numériques permettant de résoudre les problèmes de contrôle dans le cadre présenté auparavant. Deux de ces méthodes appartiennent à la classe des techniques *Regression Monte Carlo*, une famille d'algorithmes dont l'efficacité dépend majoritairement du choix d'une base de fonctions utilisée dans l'approximation des espérances conditionnelles. Le troisième appelé *Q-Algorithm* est basée sur la quantification optimale et a été introduit dans [69]; il consiste à approximer le processus $(Z_{t_n})_{n \in \llbracket 0, N \rrbracket}$ introduit dans 2.3.2 par une chaîne de Markov à espace d'état fini pour laquelle les espérances conditionnelles peuvent être approximées rapidement.

Régression Monte Carlo Dans le cadre simple des processus non contrôlés, la famille des algorithmes *Regression Monte Carlo* est basée sur l'idée d'approximation de l'espérance conditionnelle $\mathbb{E} \left(V_{\Delta t} (t_{n+1}, Z_{t_{n+1}}) \mid Z_{t_n} \right)$ pour $n \in \llbracket 0, N-1 \rrbracket$ par la projection orthogonale de $V_{\Delta t} (t_{n+1}, Z_{t_{n+1}})$ sur l'espace vectoriel engendré par $(\psi_k (Z_{t_n}))_{k \in \mathbb{N}^*}$, avec $(\psi_k)_{k \in \mathbb{N}^*}$ est une famille de *basis functions*, i.e. une famille de fonctions mesurables à valeurs dans \mathbb{R} tel que $(\psi_k (Z_{t_n}))_{k \in \mathbb{N}^*}$ est totale dans l'espace de Hilbert séparable $L^2 (\Omega, \sigma (Z_{t_n}), \mathbb{P})$ et pour tout $K \in \mathbb{N}^*$, $(\beta_k)_{k \in \llbracket 1, K \rrbracket}$ on a :

$$\sum_{k=1}^K \beta_k \psi_k (Z_{t_n}) = 0 \quad \text{P.s} \implies \forall k \in \llbracket 1, K \rrbracket : \quad \beta_k = 0$$

Une fois la base de fonctions et l'indice de troncation $K \geq 1$ sont fixés, l'espérance conditionnelle est approximée de la manière suivante :

$$\mathbb{E} \left(V_{\Delta t} (t_{n+1}, Z_{t_{n+1}}) \mid Z_{t_n} \right) \approx \sum_{k=1}^K \beta_k^n \psi_k (Z_{t_n})$$

tel que $\beta^n = (\beta_1^n, \dots, \beta_K^n)^\top$ est défini par :

$$\beta^n := \arg \min_{\beta \in \mathbb{R}^K} \left\{ \mathbb{E} \left(\left| V_{\Delta t} (t_{n+1}, Z_{t_{n+1}}) - \sum_{k=1}^K \beta_k^n \psi_k (Z_{t_n}) \right|^2 \right) \right\}$$

Qui peut être obtenue par de simples calculs et est égale à :

$$\beta^n = \mathbb{E} \left(\psi (Z_{t_n}) \psi (Z_{t_n})^\top \right)^\dagger \mathbb{E} \left(V_{\Delta t} (t_{n+1}, Z_{t_{n+1}}) \psi (Z_{t_n}) \right) \quad (2.3.5)$$

tel que $\psi := (\psi_1, \dots, \psi_K)^\top$ et $\mathbb{E} \left(\psi (Z_{t_n}) \psi (Z_{t_n})^\top \right)^\dagger$ est l'inverse de Moore–Penrose de la matrice en question.

β^n définie dans 2.3.5 est estimé par la méthode de Monte Carlo par :

$$\hat{\beta}^n := \left(\hat{\mathcal{A}}_n^M \right)^\dagger \frac{1}{M} \sum_{m=1}^M V_{\Delta t} (t_{n+1}, Z_{t_{n+1}}^m) \psi (Z_{t_n}^m)$$

tel que $M \in \mathbb{N}^*$ et $(Z_{t_n}^m, Z_{t_{n+1}}^m)_{m \in \llbracket 1, M \rrbracket}$ est un échantillon i.i.d ayant la même loi que $(Z_{t_n}, Z_{t_{n+1}})$ et :

$$\hat{\mathcal{A}}_n^M = \frac{1}{M} \sum_{m=1}^M \psi (Z_{t_n}^m) \psi (Z_{t_n}^m)^\top$$

Il est à noter que dans le cas ou la base $(\psi_k)_{k \in \mathbb{N}^*}$ est orthonormale, β^n définie dans 2.3.5 peut être estimée d'une manière moins coûteuse et plus efficiente. Même dans le cas ou la base n'est

pas orthonormale, on la orthonormalise par une multiplication par la l'inverse de la racine de la matrice de Gram associée, i.e :

$$(\tilde{\psi}_1, \dots, \tilde{\psi}_K)^\top := G^{-1/2} (\psi_1, \dots, \psi_K)^\top$$

tel que $G := (\langle \psi_i, \psi_j \rangle_{L^2})_{(i,j) \in \llbracket 1, K \rrbracket^2}$.¹ Pour plus de détails sur le sujet, le lecteur est renvoyé à [19].

Dans le cas des processus contrôlés, on doit prendre en considération l'effet du contrôle dans l'approximation de l'espérance conditionnelle. Ceci, peut être fait explicitement par la méthode *Control Randomization* ou implicitement par la méthode *Regress Later*.

L'idée de base de la méthode *Control Randomization*, introduite par Kharoubi et al. [74], est de remplacer la contrôle endogène figurant dans la dynamique de Z^α par un contrôle exogène $(I_{t_n})_{n \in \llbracket 0, N \rrbracket}$. Considérons un échantillon $(Z_{t_n}^m, I_{t_n}^m)_{(n,m) \in \llbracket 0, N \rrbracket \times \llbracket 1, M \rrbracket}$, tel que $M \geq 1$ et pour tout $n \in \llbracket 0, N \rrbracket$ $(I_{t_n}^m)_{m \in \llbracket 1, M \rrbracket}$ sont i.i.d de loi commune μ_n appelée *training distribution*. L'effet du contrôle est pris en compte à travers la base des fonctions et l'espérance conditionnelle est approximée par :

$$\mathbb{E} \left(V(t_{n+1}, Z_{t_{n+1}}) \mid Z_{t_n}, I_{t_n} \right) \approx \sum_{k=1}^K \hat{\beta}_k^n \psi_k(Z_{t_n}, I_{t_n})$$

Remarquer que les éléments la base de fonctions sont définie maintenant sur $\mathbb{R}^d \times A$. Pour plus de détails, le lecteur est renvoyé à [74]

La méthode *Regress Later* introduite dans [13] et étudiée dans [14], [23] et [97] approxime l'espérance conditionnelle par rapport à Z_{t_n} en deux stages. D'abord, une espérance conditionnelle par rapport à $Z_{t_{n+1}}$ est approximée par régression sur une famille de *basis functions* appliquées à $Z_{t_{n+1}}$. Ensuite, des formules analytiques sont utilisées pour conditionner cette famille de fonctions appliquées à la valeur future Z_{t_n} par rapport à Z_{t_n} . Dans cette approche, l'effet du contrôle est pris en compte implicitement dans la distribution conditionnelle de $Z_{t_{n+1}}$ par rapport à Z_{t_n} . Il est à noter que contrairement aux autres approches de *Regression Monte Carlo* tel que la méthode *Regress Now* discuté dans le début du paragraphe, la méthode *Regress Later* n'utilise pas l'information contenue dans la distribution conditionnelle durant l'étape de la régression comme celle ci est prise en compte dans le deuxième stage durant lequel les espérances conditionnelles sont calculées analytiquement. Pour plus de détails sur la méthode ainsi que des résultats de convergence et des bornes des erreurs d'approximation, le lecteur est renvoyé à [13].

Quantification La quantification optimale est une méthode issue du traitement du signal et de l'information. Elle consiste à approcher un signal (continu) à valeurs dans un espace infini par un signal (discret) à valeurs dans un espace fini. Cette technique a vu ses applications s'élargir à différents domaines, notamment, depuis quelques années, les probabilités numériques. On commence par introduire laconiquement la quantification optimale ainsi que son application à la résolution numérique des problèmes de contrôle. Plus de détails peuvent être trouvés dans [104], [103], [112], [105] et [69].

On note $|\cdot|$ la norme euclidienne sur \mathbb{R}^d . Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $X \in L^2(\Omega, \mathbb{P})$ un vecteur aléatoire à valeurs dans \mathbb{R}^d . Soient $N \in \mathbb{N}^*$ et $\Gamma_N := \{x_1, \dots, x_N\}$ une grille de points de \mathbb{R}^d . On appelle **partition de Voronoi** de Γ_N toute partition $(C_i(\Gamma_N))_{1 \leq i \leq N}$ de \mathbb{R}^d vérifiant :

1. G est bien une matrice définie-positive comme $(\psi_k)_{k \in \llbracket 1, K \rrbracket}$ est une famille libre de L^2 .

$$\forall i \in \llbracket 1, N \rrbracket : \quad C_i(\Gamma_N) \subset \left\{ u \in \mathbb{R}^d : \quad |u - x_i| = \min_{1 \leq j \leq N} |u - x_j| \right\} \quad (2.3.6)$$

Les ensembles à droite dans 2.3.6 sont appelés les cellules fermées de Voronoi engendrés par Γ_N .

On appelle **quantificateur de Voronoi** de X le vecteur aléatoire discret \hat{X}^{Γ_N} à valeurs dans Γ_N défini par :

$$\hat{X}^{\Gamma_N} := \text{Proj}_{\Gamma_N}(X) = \sum_{k=1}^N x_k \mathbf{1}_{C_k(\Gamma_N)}$$

et donc de loi de probabilité discrète $\mathbb{P}_{\hat{X}}$ caractérisée par :

$$\forall k \in \llbracket 1, N \rrbracket : \quad p_k := \mathbb{P}_{\hat{X}}(x_k) = \mathbb{P}(\hat{X}^{\Gamma_N} = x_k) = \mathbb{P}(X \in C_k(\Gamma_N))$$

Autrement dit, \hat{X}^{Γ_N} est la projection selon le plus proche voisin du vecteur aléatoire X sur la grille Γ_N .

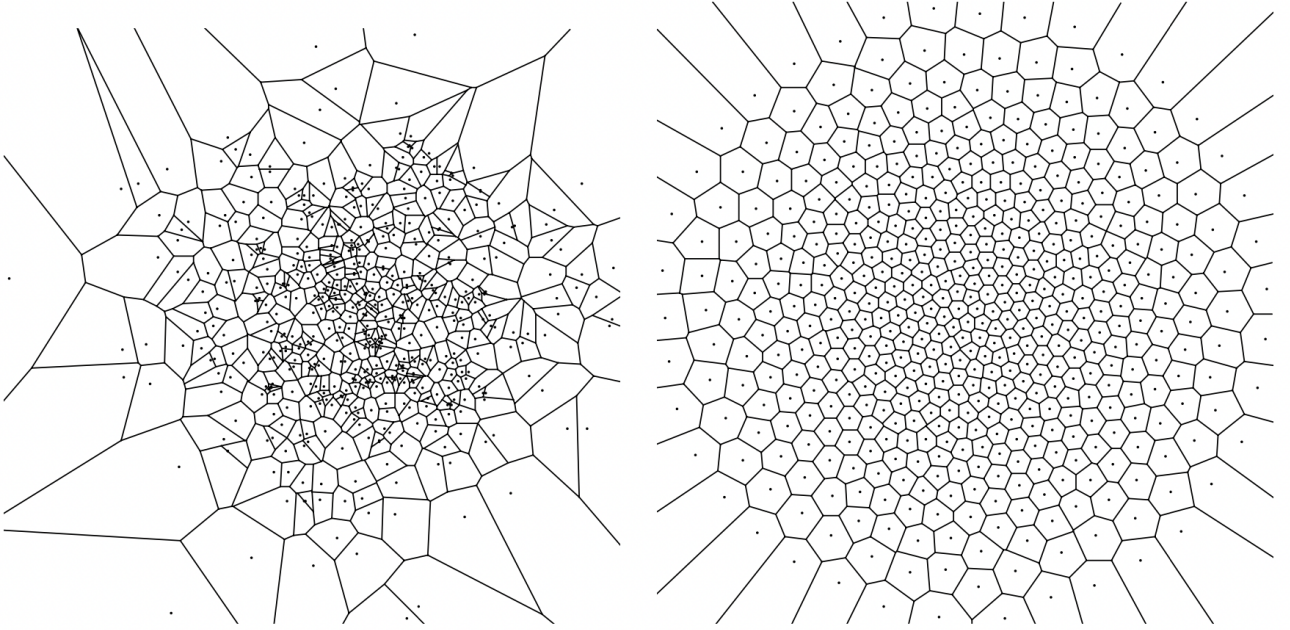


FIGURE 2.1 – Deux exemples de partitions de Voronoi en dimension 2

L'erreur résultante au carrée de quantification quadratique est appelée **distorsion** (quadratique) et s'écrit donc par définition d'une partition de Voronoi :

$$\begin{aligned} D_X^N(\Gamma_N) &:= \mathbb{E} \left(|X - \hat{X}^{\Gamma_N}|^2 \right) \\ &= \int_{\mathbb{R}^d} d_N(\Gamma_N, u) \mathbb{P}_X(du) \end{aligned}$$

où $d_N : \mathbb{R}^{d \times N} \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ est la distorsion locale définie par :

$$\forall (x, u) \in \mathbb{R}^{d \times N} \times \mathbb{R}^d : \quad d_N(x, u) := \min_{1 \leq j \leq N} |u - x_j|^2$$

Le problème de la quantification optimale, à N fixé, consiste à minimiser sur les N -grilles Γ_N la distorsion D_X^N . On dira que $\hat{X}^{\Gamma_N^*}$ est un N -quantificateur optimal de X si :

$$\mathbb{E} \left(\left| X - \hat{X}^{\Gamma_N^*} \right|^2 \right) = \inf_{\Gamma_N \in \mathbb{R}^{d \times N}} D_X^N(\Gamma_N)$$

La grille Γ_N^* qui permet de trouver le N -quantificateur optimal de X est souvent obtenue par un algorithme de gradient stochastique connue sous le nom de : algorithme de Kohonen ou *competitive learning vector quantization algorithm*. la dernière nomination décrit le fait que l'algorithme se compose à chaque itération d'une étape de compétition et une étape d'apprentissage. Plus de détails sur l'algorithme peuvent être trouvés dans [105]. On mentionne que les partitions de Voronoï optimales de la quantification de la loi normale à une dimension inférieure ou égale à 10 sont pré-calculés et peuvent être trouvés dans le site :

<http://www.quantize.maths-fi.com/>

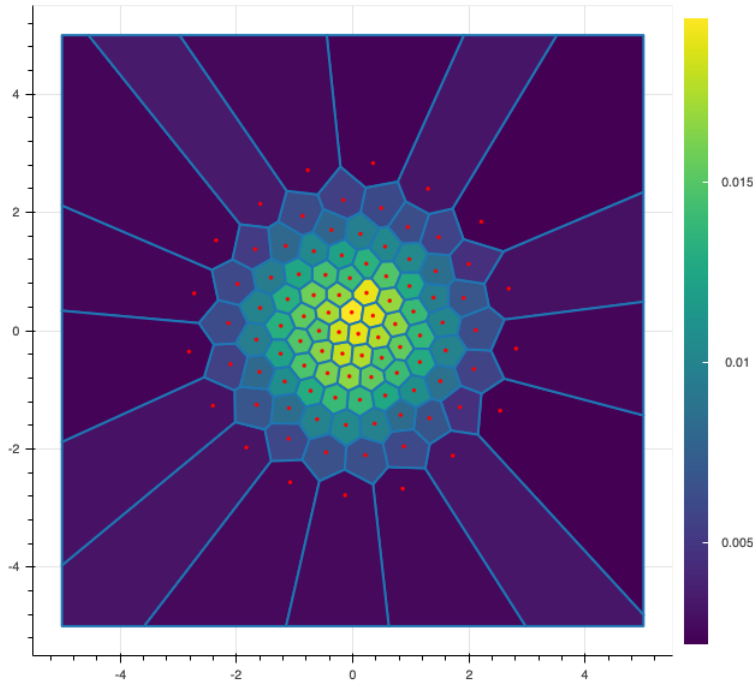


FIGURE 2.2 – Quantification optimale de taille $N = 100$ de la loi normale $\mathcal{N}_2(0, I_2)$

On présente maintenant deux méthodes basées sur la quantification optimale permettant d'approximer les espérances conditionnelles. La première appartient à la famille des approximations constantes par morceaux. La deuxième est une amélioration de la première pour laquelle la continuité de l'approximation par rapport à la variable de contrôle est préservée.

Soit $(\Gamma_n)_{0 \leq n \leq N}$ une suite d'ensembles finis tel que chacun d'eux contient des points de \mathbb{R}^d .

la méthode *Piecewise constant interpolation* repose sur l'idée suivante. Supposons qu'à l'étape n , on possède une approximation de $V_{\Delta t}(t_{n+1}, \cdot)$, la fonction valeur à l'instant t_{n+1} , que l'on note $\hat{V}_{\Delta t}^Q(t_{n+1}, \cdot)$

Alors, l'espérance conditionnelle est approximée de la manière suivante :

$$\mathbb{E}_{n,z}^a \left(\hat{V}_{\Delta t}^Q(t_{n+1}, Z_{t_{n+1}}) \right) \approx \sum_{k=1}^L p_k \hat{V}_{\Delta t}^Q(t_{n+1}, \text{Proj}_{\Gamma_{n+1}}(G_{\Delta t}(z, a, \varepsilon_k^*))) \quad \text{pour } z \in \Gamma_n \quad (2.3.7)$$

Tel que :

- $G_{\Delta t}$ est la fonction définie, en rappelant la notation de la discrétisation 2.3.2, par :

$$\forall (z, a, \varepsilon) \in \mathbb{R}^d \times A \times \mathbb{R}^d : \quad G_{\Delta t}(z, a, \varepsilon) := z + b(z, a)\Delta t + \sigma(z, a)\sqrt{\Delta t}\varepsilon$$

De manière que pour tout $n \in \llbracket 0, N-1 \rrbracket$:

$$Z_{t_{n+1}} = G(Z_{t_n}, \alpha_{t_n}, \varepsilon_{t_n})$$

tel que $(\varepsilon_n)_{0 \leq n \leq N}$ est une suite de vecteurs aléatoires i.i.d de loi commune $\mathcal{N}_d(0, I_d)$.

- $\Gamma_L^* := \{\varepsilon_1^*, \dots, \varepsilon_L^*\}$ est la grille optimale permettant d'obtenir le L -quantificateur optimal de ε_n à l'instant n et $(p_k)_{1 \leq k \leq L}$ est le vecteur des poids associée à ε^* , i.e. :

$$\forall k \in \llbracket 1, L \rrbracket : \quad p_k = \mathbb{P}(\varepsilon^* = \varepsilon_k^*)$$

Un contrôle optimal à l'instant t_n et pour $z \in \Gamma_n$ que l'on note $\hat{\alpha}_{t_n}^Q = \hat{\alpha}^Q(t_n, Z_{t_n})$ est donné par :

$$\hat{\alpha}^Q(t_n, z) := \arg \max_{a \in A} \left\{ f(z, a)\Delta t + \sum_{k=1}^L p_k \hat{V}_{\Delta t}^Q(t_{n+1}, \text{Proj}_{\Gamma_{n+1}}(G_{\Delta t}(z, a, \varepsilon_k^*))) \right\}$$

Ainsi, la fonction valeur est approximée par :

$$\hat{V}_{\Delta t}^Q(t_n, z) = \left(f(z, \hat{\alpha}_{t_n}^Q)\Delta t + \sum_{k=1}^L p_k \hat{V}_{\Delta t}^Q(t_{n+1}, \text{Proj}_{\Gamma_{n+1}}(G_{\Delta t}(z, \hat{\alpha}_{t_n}^Q, \varepsilon_k^*))) \right) \mathbb{1}_{\Gamma_n}(z)$$

La deuxième méthode appelée *semi-linear interpolation* n'est présentée que dans le cas ou le processus de contrôle est unidimensionnel ($d = 1$). Elle consiste à améliorer l'approximation 2.3.7 discontinue par rapport à la variable de contrôle par une approximation linéaire par morceaux :

$$\mathbb{E}_{n,z}^a \left(\hat{V}_{\Delta t}^Q(t_{n+1}, Z_{t_{n+1}}) \right) \approx \sum_{k=1}^L p_k \left(\lambda_{a^{\varepsilon_k^*, z}}^{\varepsilon_k^*, z} \hat{V}_{\Delta t}^Q(t_{n+1}, z_+^k) + (1 - \lambda_{a^{\varepsilon_k^*, z}}^{\varepsilon_k^*, z}) \hat{V}_{\Delta t}^Q(t_{n+1}, z_-^k) \right) \quad \text{pour } z \in \Gamma_n$$

tel que, pour tout $k \in \llbracket 1, L \rrbracket$ z_+^k et z_-^k sont définis par :

- z_+^k et z_-^k sont deux états de Γ_{n+1} les plus proches de $G_{\Delta t}(z, a, \varepsilon_k^*)$ tel que : $z_-^k < G_{\Delta t}(z, a, \varepsilon_k^*) < z_+^k$. S'ils existent, alors $\lambda_{a^{\varepsilon_k^*, z}}^{\varepsilon_k^*, z}$ est définie par :

$$\lambda_{a^{\varepsilon_k^*, z}}^{\varepsilon_k^*, z} := \frac{G_{\Delta t}(z, a, \varepsilon_k^*) - z_-^k}{z_+^k - z_-^k}$$

- Sinon, ils sont égaux et on pose : $\lambda_{a^{\varepsilon_k^*, z}}^{\varepsilon_k^*, z} = 1$.

Pour plus de détails sur la conception des ensembles d'apprentissage $(\Gamma_n)_{0 \leq n \leq N}$, la recherche du contrôle optimal ainsi que plusieurs détails pratiques, le lecteur est renvoyé à [69]. On finit la section par un pseudo-code du *Q-Algorithm* décrit ci-haut, présenté dans l'algorithme 1

Algorithme 1 : Q-Algorithm, Value iteration

Input :

- $(\Gamma_n)_{0 \leq n \leq N}$: grilles de points de \mathbb{R}^d utilisées pour l'apprentissage.
- $\Gamma_L^* := \{\varepsilon_1^*, \dots, \varepsilon_L^*\}$, $(p_k)_{1 \leq k \leq L}$: la grille optimale du L -quantificateur optimal du bruit exogène ε et la famille de poids associés.

Initialize : $\hat{V}_{\Delta t}^Q(t_N, z) = g(z)$ pour $z \in \Gamma_N$;**for** $n = N - 1$ **to** 0 **do**→ Approximer la fonction valeur à l'instant t_n , pour $z \in \Gamma_n$, par :

$$\hat{V}_{\Delta t}^Q(t_n, z) = \max_{a \in A} \left\{ f(z, a) \Delta t + \sum_{k=1}^L p_k \hat{V}_{\Delta t}^Q(t_{n+1}, \text{Proj}_{\Gamma_{n+1}}(G_{\Delta t}(z, a, \varepsilon_k^*))) \right\}$$

;

→ Collectionner les contrôles optimaux $(\hat{\alpha}^Q(t_n, z))_{z \in \Gamma_n}$;**end**→ Calculer $\tilde{V}_{\Delta t}^Q(0, z)$ pour $z \in \Gamma_0$ à l'aide des contrôles collectionnés;**Outputs :** $(\hat{\alpha}^Q(t_n, z))_{z \in \Gamma_n, n \in \llbracket 0, N-1 \rrbracket}$, $(\tilde{V}_{\Delta t}^Q(0, z))_{z \in \Gamma_0}$

Exécution optimale à l'aide du contrôle stochastique

On s'intéresse dans ce chapitre au contexte mathématique de l'exécution optimale des ordres ayant un volume très large. Comme, nous avons déjà mentionné, l'investisseur divise le *parent order* en des *child orders* qui sont exécutées durant une période fixée afin de minimiser son exposition au *market impact*. Deux problèmes seront étudiés, à savoir : la liquidation et l'acquisition optimale. Rappelons que deux problèmes soulèvent durant ce processus 1.1.3. Le problème stratégique sera étudié mathématiquement par la formulation d'un problème de contrôle stochastique dont le contrôle est le taux de liquidation/acquisition. Quant au problème tactique, on présentera l'exécution optimale dans différentes situations, c-à-d en utilisant les ordres au marché, les ordres limites ou les deux (*order placement*). On se place dans tous les cas dans un lieu d'échange primaire (*order routing*). Les modèles que nous présenterons sont des généralisations en temps continu du modèle *black-box* de Almgren et Chriss [6] présenté dans la section 1.2.2. Au lieu de se concentrer sur des coûts d'exécution ayant une forme quadratique en fonction du volume, on étudiera des formes plus générales.

3.1 Exécution optimale à l'aide des ordres au marché

3.1.1 Contexte général

Soit $T > 0$ l'horizon du temps correspondant à la période de la liquidation/acquisition de l'actif. On fixe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel on définit un mouvement brownien standard $(W_t)_{t \in [0, T]}$. On muni cet espace de l'augmentation habituelle de la filtration brownienne que l'on note $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$. On note $(\nu_t)_{t \in [0, T]}$ un processus à valeurs dans \mathbb{R}_+^* modélisant le taux de liquidation/acquisition. Le processus d'inventaire $(Q_t^\nu)_{t \in [0, T]}$ est donné par :

$$\forall t \in]0, T] : \quad dQ_t^\nu = \pm \nu_t dt \quad ; \quad Q_0^\nu = q \in \mathbb{R}_+$$

Le prix de référence (typiquement le *mid-price*) $(S_t^\nu)_{t \in [0, T]}$ est donnée par l'EDS :

$$\forall t \in]0, T] : \quad dS_t^\nu = \pm g(\nu_t) dt + \sigma dW_t \quad ; \quad S_0^\nu = S \in \mathbb{R}_+$$

tel que : $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est la fonction du *permanent market impact* modélisant l'impact permanent de l'exécution de l'investisseur sur le prix de référence. $\sigma > 0$ est la volatilité de l'actif.

Le prix d'exécution obtenu par l'investisseur $(\tilde{S}_t^\nu)_{t \in [0, T]}$ est donné par :

$$\forall t \in [0, T] : \quad \tilde{S}_t^\nu = S_t^\nu \pm \left(\frac{1}{2} \Delta + f(\nu_t) \right)$$

tel que : $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est la fonction du *temporary market impact* modélisant l'impact instantané de l'exécution de l'investisseur sur son prix d'exécution. $\Delta \geq 0$ est le *bid-ask spread* supposé constant.

Le processus du *cash* de l'investisseur est donnée naturellement par :

$$\forall t \in]0, T] : \quad dX_t^\nu = \tilde{S}_t^\nu \nu_t dt \quad ; \quad X_0^\nu = x \in \mathbb{R}_+$$

Les équations ci-haut correspondent aux deux problèmes : liquidation (-) et acquisition (+). Il est à noter que les traitements des deux problèmes sont similaires d'un point de vue théorique. Les formes particulières du processus du prix de référence et celui du prix d'exécution sont construite de manière à incorporer les phénomènes observés empiriquement que nous avons présenté dans le premier chapitre.

On introduit la fonction du coût d'exécution $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, définie pour tout $v > 0$ par : $\psi(v) = v f(v)$. On suppose que :

- $(\mathcal{H}_1) : L(0) = 0$, i.e il n'y a pas de coûts fixe.
- $(\mathcal{H}_2) : L$ est strictement convexe et croissante sur \mathbb{R}_+ .
- $(\mathcal{H}_3) : L$ est asymptotiquement sur-linéaire :

$$\lim_{v \rightarrow +\infty} \frac{L(v)}{v} = +\infty$$

L'hypothèse \mathcal{H}_1 signifie qu'il n'y a aucun coût d'exécution en absence du *trading*. La deuxième hypothèse \mathcal{H}_2 veut dire qu'il y a toujours un coût d'exécution en présence du *trading* et ce coût augmente avec le taux de ce dernier. De plus, cette relation n'est pas une simple relation linéaire. L'hypothèse \mathcal{H}_3 est une hypothèse technique assurant que la transformée de LEGENDRE-FENCHEL de L prend des valeurs finies.

D'après les résultats empiriques présentés dans le premier chapitre et dans la pratique, L prend généralement la forme : $L(v) = \eta v^{1+\phi} + \psi v$ pour $\psi, \phi > 0$. La forme particulière présentée dans le papier original de Almgren et Chriss correspond au cas : $L(v) = \eta v^2$.

On se place dans le cadre d'un investisseur possédant un large volume $Q_0 = \Xi > 0$ à l'instant initiale et qui veut le liquider pendant la période $[0, T]$. Typiquement, le terme "large" signifie que Ξ est plus grand que le volume moyen du côté *bid* présent dans le marché.

Si l'investisseur est considéré comme risque-neutre, alors il va chercher à maximiser son *payoff* final espéré ou à minimiser son coût d'exécution appelé *implementation shortfall*. Pour autoriser un univers risque-averse sans introduire explicitement une fonction d'utilité espérée, on suppose que si l'investisseur reste avec une inventaire non nulle $Q_T > 0$ à l'instant finale, alors il exécute cette quantité au *mid-price* et reçoit une pénalité paramétrisée par un réel positif α . Le revenue de cette dernière opération est donc :

$$Q_T S_T - \alpha Q_T^2 = Q_T (S_T - \alpha Q_T)$$

. Par ailleurs, on introduit une pénalité d'inventaire courue paramétrisée par un réel positif ϕ qui représente la vitesse avec laquelle l'investisseur veut exécuter les transactions. Elle prend la forme suivante :

$$-\phi \int_0^T Q_t^2 dt$$

Ainsi, l'objectif de l'investisseur est de maximiser la fonctionnelle de gain qui représente le *payoff* final espéré et pénalisé qui s'écrit sous la forme suivante pour tout $(t, x, S, q) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*$:

$$J(t, x, S, q, \nu) := \mathbb{E} \left(\underbrace{X_T^{t,x,\nu}}_{\text{cash final}} + \underbrace{Q_T^{t,q,\nu} \left(S_T^{t,x,\nu} - \alpha Q_T^{t,q,\nu} \right)}_{\text{exécution finale pénalisée}} - \underbrace{\phi \int_t^T |Q_u^{t,q,\nu}|^2 du}_{\text{pénalité d'inventaire courue}} \right)$$

Pour simplifier les notations, on n'écrira pas la dépendance des variables du contrôle ν (i.e par exemple, on écrit $X_T^{t,x}$ au lieu de $X_T^{t,x,\nu}$). De plus, on suppose que $\Delta = 0$ ce qui est équivalent à considérer S comme étant le meilleur prix opposé (le *best-bid* dans ce cas).

Le problème de l'investisseur est un problème de contrôle stochastique dont les pré-requis théoriques sont présentés dans le chapitre précédent. En effet, Le processus d'état $\psi := (X, S, Q)^T$ est une diffusion contrôlée définie par :

$$\forall t \in]0, T] : \quad d\psi_t := \mu(t, \psi_t, \nu_t) dt + \sigma(t, \psi_t, \nu_t) d\tilde{W}_t \quad (3.1.1)$$

tel que : $\mu : [0, T] \times \mathbb{R}_+^3 \times \mathbb{R}_+^* \rightarrow \mathbb{R}_+^3$, $\sigma : [0, T] \times \mathbb{R}_+^3 \times \mathbb{R}_+^* \rightarrow \mathcal{M}_3(\mathbb{R})$ sont des fonctions définies par :

$$\forall (t, \psi, u) \in [0, T] \times \mathbb{R}_+^3 \times \mathbb{R}_+^* : \quad \mu(t, \psi, u) = \begin{pmatrix} (\langle \psi, e_2 \rangle - f(u)) u \\ -g(u) \\ -u \end{pmatrix} ; \quad \sigma(t, \psi, u) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

avec $(e_i)_{1 \leq i \leq 3}$ est la base canonique de \mathbb{R}^3 . \tilde{W} est un mouvement brownien de la forme :

$$\forall t \in [0, T] : \quad \tilde{W}_t := \begin{pmatrix} 0 \\ W_t \\ 0 \end{pmatrix}$$

La fonction du gain finale $G : \mathbb{R}_+^3 \rightarrow \mathbb{R}_+$ et du gain couru $F : [0, T] \times \mathbb{R}_+^3 \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ sont définies par :

$$\forall (t, \psi, u) \in [0, T] \times \mathbb{R}_+^3 \times \mathbb{R}_+^* : \quad G(\psi) = \langle \psi, e_1 \rangle + \langle \psi, e_3 \rangle (\langle \psi, e_2 \rangle - \alpha \langle \psi, e_3 \rangle) ; \quad F(t, \psi, u) = |\langle \psi, e_3 \rangle|^2$$

Pour que J soit bien définie, l'ensemble des contrôles admissibles est :

$$\mathcal{A} := \left\{ \nu : (\omega, t) \mapsto \mathbb{R}_+^* \in \mathbb{H}^2(\mathbb{F}) : \quad -\Xi = \int_0^T \nu_t dt \in L^\infty(\Omega) \right\}$$

En adoptant les mêmes notations du chapitre 2, pour tout $t \in [0, T]$, \mathcal{A}_t désigne l'ensemble des contrôles appartenant à \mathcal{A} démarants de t et indépendants de \mathcal{F}_t . De ce fait, le problème du contrôle stochastique s'écrit sous la forme :

$$\forall (t, \psi) \in [0, T] \times \mathbb{R}_+^3 : \quad V(t, \psi) = V(t, x, S, q) = \sup_{\nu \in \mathcal{A}_t} \mathbb{E} \left(G(\psi_T^{t,\psi}) + \int_t^T F(u, \psi_u^{t,\psi}, \nu_u) du \right) \quad (3.1.2)$$

L'équation HJB associée au problème 3.1.2 est :

$$\begin{cases} \partial_t V + \frac{1}{2} \sigma^2 \partial_S^2 V - \phi q^2 + \sup_{u > 0} (u(S - f(u)) \partial_x V - g(u) \partial_S V - u \partial_q V) = 0 & ; (t, \psi) \in [0, T[\times \mathbb{R}_+^3 \\ V(T, x, S, q) = x + q(S - \alpha q) & ; x \in \mathbb{R}_+^3 \end{cases} \quad (3.1.3)$$

Définition 3.1.1: Arbitrage dynamique

Il existe une opportunité d'arbitrage dynamique dans le marché au sens de Huberman et Stanzl [68], s'il existe un processus $\nu \in \mathbb{H}^0(\mathbb{F})$ et $t_1, t_2 \in [0, T]$ tel que $t_1 < t_2$ vérifiant :

- $\int_{t_1}^{t_2} (|\nu_t| + |g(\nu_t)|) dt \in L^\infty(\Omega)$,
- $\int_{t_1}^{t_2} \nu_t dt = Q_{t_2} - Q_{t_1} = 0$,
- $\mathbb{E}(X_{t_2} | \mathcal{F}_{t_1}) > X_{t_1}$.

Autrement dit, une opportunité d'arbitrage dynamique correspond à une stratégie d'aller-retour (*round trip strategy*) qui est profitable en moyenne.

La proposition suivante stipule que l'absence d'arbitrage dynamique dans ce cadre est équivalent à un *market impact* linéaire.

Proposition 3.1.1: (cf. [54])

On suppose que $f \equiv 0$. Alors, Les assertions suivantes sont équivalentes.

- Il n'existe aucune opportunité d'arbitrage dynamique dans le cadre définie ci-haut.
- $\exists k > 0, \forall v \geq 0 : g(v) = kv$

On se place, dorénavant, dans le cadre d'un *permanent market impact* linéaire :

$$\exists k > 0, \forall v \geq 0 : g(v) = kv$$

Pour proposer une solution particulière de l'équation HJB 3.1.3, on se base sur un changement de variable basé sur le raisonnement suivant. En remarquant la forme particulière de la valeur terminale et en interprétant que V est une fonction représentant un gain. Alors, sa valeur sera heuristiquement de la forme :

$$V(t, x, S, q) = \underbrace{x}_{\text{cash initial}} + \underbrace{qS}_{\text{valeur marché de l'inventaire initiale}} + \underbrace{v(t, q)}_{\text{profit/perte du à l'exécution}} \quad (3.1.4)$$

tel que $v : [0, T] \times \mathbb{R}_+^* \rightarrow \mathbb{R}$ est une fonction vérifiant :

$$\forall q > 0 : v(T, q) = -\alpha q^2$$

On propose pour v une forme de séparation des variables t et q . Plus concrètement, on a la proposition suivante.

Proposition 3.1.2

Soit $(\varphi_q)_{q \in \mathbb{R}_+^*}$ une famille de fonction de $[0, T]$ vers \mathbb{R} solution du système des problèmes de Cauchy défini, pour tout $q > 0$, par :

$$\begin{cases} \varphi_q' - \phi + \frac{L^*(-q(k + 2\varphi_q))}{q^2} = 0 ; & t \in [0, T[\\ \varphi_q(T) = -\alpha \end{cases} \quad (3.1.5)$$

tel que L^* est la transformée de LEGENDRE-FENCHEL de L qui est fini et appartient à $\mathcal{C}^1(\mathbb{R})$.

Alors, la fonction $V : [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^* \rightarrow \mathbb{R}$ définie par :

$$\forall (t, x, S, q) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^* : \quad V(t, x, S, q) = x + qS + \varphi_q(t)q^2, \quad (3.1.6)$$

est une solution appartenant à $\mathcal{C}^{1,2}([0, T[\times \mathbb{R}_+^2 \times \mathbb{R}_+^*) \cap \mathcal{C}^0([0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*)$ de l'équation HJB 3.1.3.

Démonstration. Voir Annexe 3.3 □

Le théorème suivant montre que la fonction 3.1.6 définie dans la proposition précédente est une solution du problème de contrôle 3.1.2.

Théorème 3.1.1: Solution du problème du contrôle

Soit $(\varphi_q)_{q \in \mathbb{R}_+^*}$ la famille de fonctions définie dans la proposition 3.1.2. Alors, la fonction définie dans 3.1.6 est une solution du problème du contrôle 3.1.2. De plus, le contrôle optimale est définie, pour tout $(x, S, q) \in \mathbb{R}_+^2 \times \mathbb{R}_+^*$, par :

$$\forall t \in [0, T] : \quad \nu_t^* = \arg \max_{u > 0} [-q(k + 2\varphi_q(t))u - L(u)]$$

Démonstration. Voir Annexe 3.3 □

Pour proposer des solutions analytiques ou des approximations numériques du problème 3.1.2, on doit spécifier une forme particulière pour la fonction du coût d'exécution L . La forme la plus utilisée dans la pratique est la suivante : pour $(\beta, \eta, \kappa) \in \mathbb{R}_+^* \times \mathbb{R}_+^3$:

$$\forall v \in \mathbb{R}_+ : \quad L(v) := \underbrace{\eta v^{1+\beta}}_{\text{Market impact}} + \underbrace{\kappa v}_{\text{Bid-ask spread}} \quad (3.1.7)$$

La transformée de LEGENDRE-FENCHEL associée à L est donnée par :

$$L^*(y) := \sup_{v \in \mathbb{R}_+} (yv - L(v)) = -\eta\beta \left(\frac{y - \kappa}{\eta(1 + \beta)} \right)^{\frac{1+\beta}{\beta}} \mathbb{1}_{[\kappa, +\infty[}(y) \quad (3.1.8)$$

Traisons tout d'abord, le cas présentée dans le papier de Almgren et Chriss qui correspond à une fonction L quadratique ($\beta = 1, \kappa = 0, \eta > 0$) et donc à un *temporary market impact* linéaire.

Proposition 3.1.3

On se place dans le cas d'un fonction de coût d'exécution quadratique, i.e :

$$\exists \eta > 0, \forall v \in \mathbb{R}_+ : \quad L(v) := \eta v^2$$

On suppose que les paramètres du problème vérifient :

$$\left(\alpha - \frac{1}{2}k + \sqrt{\eta\phi} \right) \left(\alpha - \frac{1}{2}k - \sqrt{\eta\phi} \right) > 0 \quad ; \quad \sqrt{\frac{\phi}{\eta}} > \frac{1}{2T} \ln \left(\frac{\alpha - \frac{1}{2}k + \sqrt{\eta\phi}}{\alpha - \frac{1}{2}k - \sqrt{\eta\phi}} \right)$$

Alors, le taux de liquidation optimale et le processus d'inventaire sont des processus déterministes donnés par :

$$\forall t \in [0, T] : \quad \nu_t^* = \gamma \frac{\zeta e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} \Xi \quad ; \quad Q_t^{\nu^*} = \frac{\zeta e^{\gamma(T-t)} - e^{-\gamma(T-t)}}{\zeta e^{\gamma T} - e^{-\gamma T}} \Xi,$$

tel que $(\gamma, \zeta) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ sont des constantes données par :

$$\gamma := \sqrt{\frac{\phi}{\eta}} \quad ; \quad \zeta := \frac{\alpha - \frac{1}{2}k + \sqrt{\eta\phi}}{\alpha - \frac{1}{2}k - \sqrt{\eta\phi}}$$

Démonstration. Voir section 6.5 du chapitre 6 de [119]. □

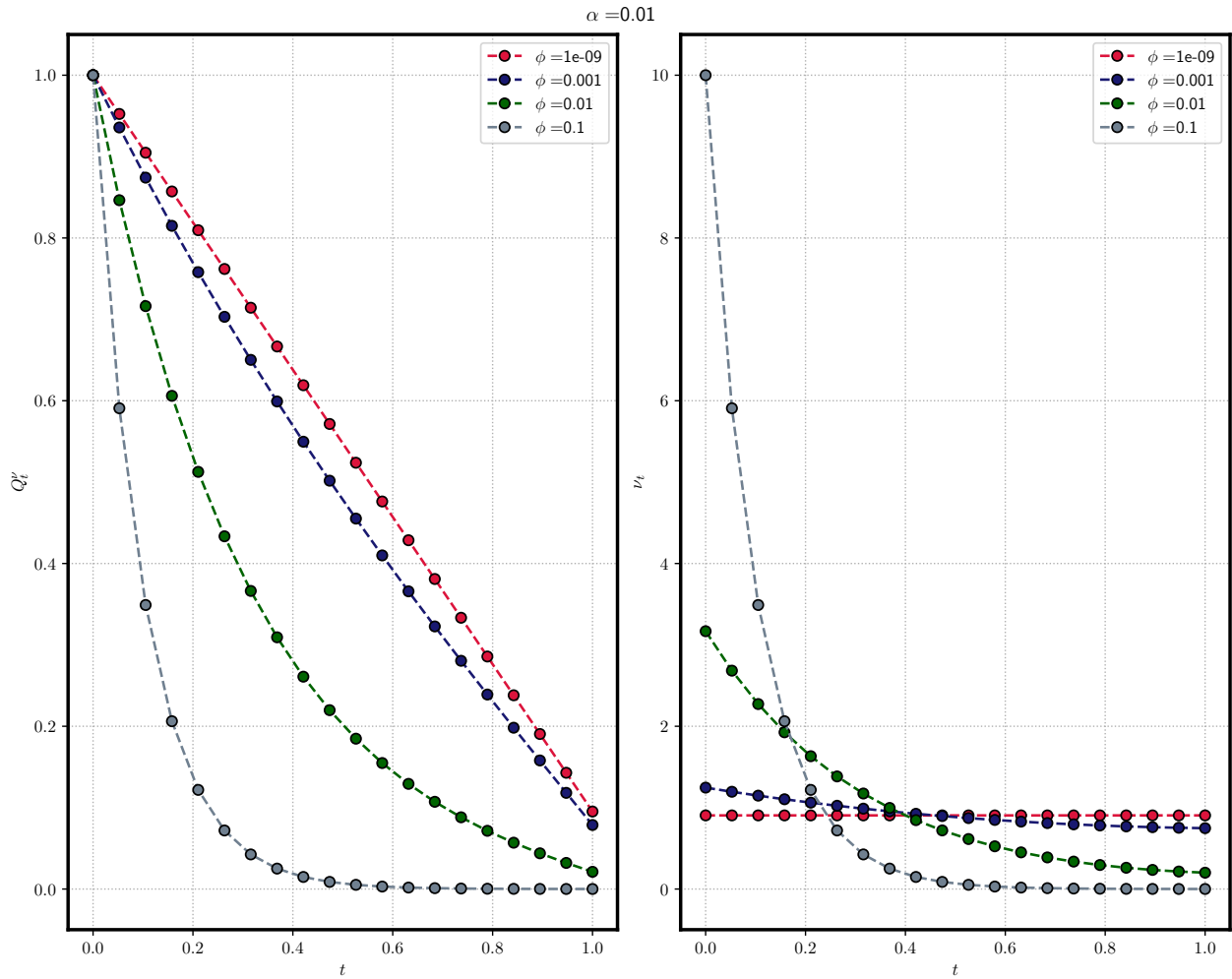


FIGURE 3.1 – graphe de ν^* et Q^{ν^*} pour $\alpha = 0.01$

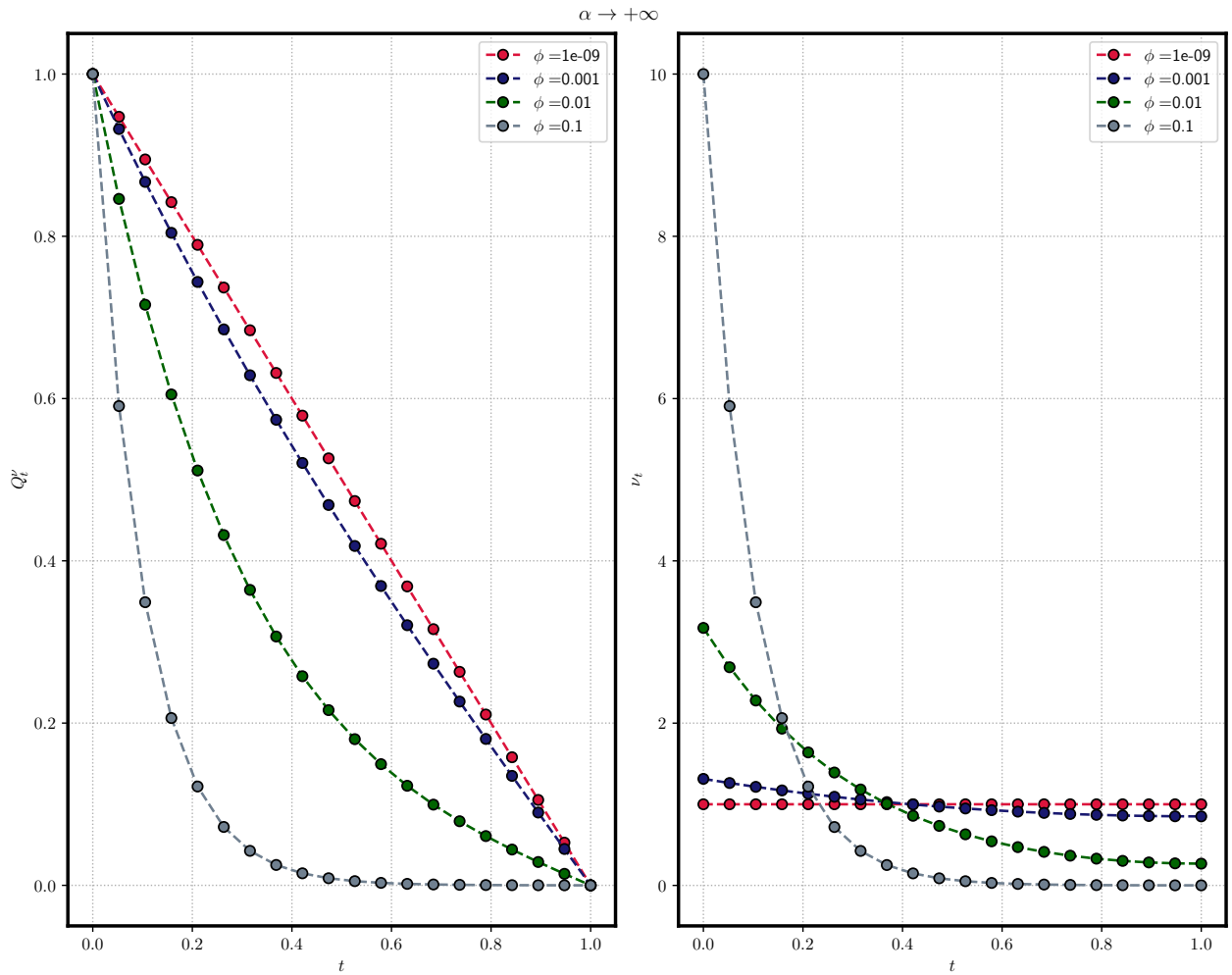


FIGURE 3.2 – graphe de ν^* et Q^{ν^*} pour $\alpha \rightarrow +\infty$

Les figures 3.1, 3.2 les graphiques de ν^* et Q^{ν^*} pour différentes valeurs de pénalités ϕ et α . Le contexte numérique des figures est :

$$(k, \eta, T, \Xi) = (10^{-3}, 10^{-3}, 1., 1.)$$

Remarque 3.1.1

- Dans le cas où l'investisseur pénalise trop une inventaire finale ($\alpha \rightarrow +\infty$), on a :

$$\nu_t^* \xrightarrow{\alpha \rightarrow +\infty} \gamma \frac{\cosh(\gamma(T-t))}{\sinh(\gamma T)} \Xi$$

Donc, la stratégie ne dépend pas du *market impact* k .

- D'après les figures 3.1 et 3.2, on constate que dans la cas où la pénalité de l'inventaire courue est nulle ($\phi = 0$), la stratégie est constante dans le temps, et correspond à une stratégie TWAP (Time Weighted Average Price) pour laquelle :

$$\nu_t^* = \frac{\Xi}{T + \frac{k}{\alpha}}$$

et dès que ϕ augmente, les courbes $t \mapsto \nu_t^*$ deviennent de plus en plus convexe et la stratégie a tendance de liquider rapidement l'inventaire. Ceci étant intuitif comme ϕ mesure la vitesse avec laquelle l'investisseur veut liquider son inventaire.

On traite maintenant le cas d'un coût d'exécution de la forme 3.1.7. D'après l'expression 3.1.8, le système différentiel 3.1.5 devient :

$$\begin{cases} \varphi'_q - \phi - \frac{\eta\beta}{q^2} \left(\frac{-q(k + 2\varphi_q) - \kappa}{\eta(1 + \beta)} \right)^{\frac{1+\beta}{\beta}} \mathbb{1}_{\{\varphi_q < -\frac{1}{2}(\frac{\kappa}{q} + k)\}} = 0 ; & t \in [0, T[\\ \varphi_q(T) = -\alpha \end{cases} \quad (3.1.9)$$

Traisons le cas des solutions vérifions :

$$\varphi_q \geq -\frac{1}{2} \left(\frac{\kappa}{q} + k \right)$$

On suppose que les paramètres du modèle vérifient :

$$\phi T + \alpha \leq \frac{1}{2} \left(\frac{\kappa}{q} + k \right)$$

Dans ce cas, les solutions ne dépendent pas de q et sont données par :

$$\varphi(t) = \phi(t - T) - \alpha < 0$$

Selon le théorème 3.1.1, le contrôle optimale peut être déduit facilement et est donné par :

$$\forall t \in [0, T] : \quad \nu_t^* = \left(\frac{-q(k + 2\phi(t - T) - 2\alpha) - \kappa}{\eta(1 + \beta)} \right)^{\frac{1}{\beta}}$$

Remarquer que la condition vérifiées par les paramètres assure la stricte positivité de la solution ν^* .

Dans le cas contraire où :

$$\varphi_q < -\frac{1}{2} \left(\frac{\kappa}{q} + k \right)$$

L'EDO est extrêmement non-linéaire et requiert une résolution numérique par une méthode de Runge-Kutta. Remarquer que tous les solutions du problème 3.1.9 sont négatives. Ce qui est un résultat intuitif puisque par la construction du modèle, le mieux que l'investisseur puisse faire est d'atteindre le *mid-price*, soit donc :

$$\forall t \in [0, T] : \quad X_t \leq x + qS \quad \text{p.s}$$

Ce qui nous rappelle l'article [25] : "To Cross or Not to Cross the Spread : That Is the Question". On pourra également résoudre numériquement l'EDP vérifiée par la fonction v définie dans 3.2.3 :

$$\begin{cases} \partial_t v - \phi q^2 + L^* (-q(k + \partial_q v)) = 0 ; & (t, q) \in [0, T[\times \mathbb{R}_+^* \\ v(T, q) = -\alpha q^2 ; & q \in \mathbb{R}_+^* \end{cases}$$

Plus généralement, on résout numériquement l'équation HJB 3.1.3 et le contrôle optimal sera construit itérativement.

3.1.2 Extentions

Incorporation du flux d'ordres : Le modèle présenté dans la dernière sous-section suppose implicitement que l'agent qui fait la liquidation est le seul qui influence le *midprice*. Autrement dit, nous avons supposé que l'action de chaque agent du marché annule celle d'un autre de manière à qu'ils contribuent au changement du *midprice* par un drift nul. Même si cette hypothèse peut apparaître raisonnable au long terme, elle n'est pas vérifiée à court terme à cause de ce qu'on appelle : *Order flow imbalance*. Voir [119] (chapitre 12) pour plus de détails.

Cartea et Jaimungal [34] ont proposée l'idée suivante. On suppose que le processus du taux du flux d'ordres d'achat $(\mu_t^+)_{t \in [0, T]}$ et celui du taux du flux d'ordres de vente $(\mu_t^-)_{t \in [0, T]}$ sont définies à l'aide des EDS suivantes :

$$d\mu_t^\pm = -\kappa\mu_t^\pm dt + \eta dJ_t^\pm \quad (3.1.10)$$

Tel que $(J_t^\pm)_{t \in [0, T]}$ sont deux processus de comptage associées à deux processus de poisson homogènes et indépendants et indépendants de tout les processus définis dans le début de la section précédente d'intensités λ^\pm . κ et η sont des constantes strictement positives. On suppose que $\lambda := \lambda^+ = \lambda^-$ pour simplifier.

Pour que le lecteur gagne un peu d'intuition sur ces processus, il doit les appercevoir de la façon suivante. On note par $(V_t^+)_{t \in [0, T]}$ (res. $(V_t^-)_{t \in [0, T]}$) le volume d'achat (resp. vente) présent au marché à tout instant $t \in [0, T]$. Alors :

$$\begin{aligned} dV_t^+ &= \mu_t^+ dt \\ dV_t^- &= \mu_t^- dt \end{aligned}$$

Ainsi, heuristiquement, $(\mu_t^\pm)_{t \in [0, T]}$ représente à $(V_t^\pm)_{t \in [0, T]}$ ce que représente le taux de liquidation/acquisition au processus d'inventaire présenté dans la début de la section précédente.

L'EDS 3.1.10 incorpore une propriété de "*mean-reversion*" observé empiriquement dans la dynamique du processus du flux d'ordres (voir figure 3.3). En effet, on suppose que les ordres d'achat et de vente arrivent indépendamment suivant deux processus ponctuels de poisson d'intensités λ et induisent un saut d'amplitude η dans la dynamiques des taux des flux d'ordres d'achat et de vente. De plus, ce saut décroît exponentiellement vers 0 (*mean-reversion*) avec une vitesse égale à κ .

Pour incorporer l'influence du flux d'ordres dans la dynamique du processus du *midprice*, on suppose que ce dernier est donné par l'EDS suivante en conservant les mêmes notation de la section précédente :

$$dS_t^\nu = (g(\mu_t^+) - g(\nu_t + \mu_t^-)) dt + \sigma dW_t$$

Remarquer qu'en choisissant la même fonction de *permanent marker impact* pour le flux d'ordres d'achat et de vente, on suppose implicitement que les action de l'agent d'intérêt et les autres agents sont traités d'un manière symétrique et influence le *mid-price* de la même façon. Dans le cadre d'un *permanent market impact* linéaire, i.e : $\exists k \in \mathbb{R}_+ : g = k\text{Id}_{\mathbb{R}_+}$, on peut réduire le nombre de variables de la manière suivante. On définit le processus du flux d'ordre net par :

$$\forall t \in [0, T] : \mu_t := \mu_t^+ - \mu_t^-$$

De ce fait, $(\mu_t)_{t \in [0, T]}$ est donné par l'EDS suivante :

$$d\mu_t = -\kappa\mu_t dt + \eta (dJ_t^+ - dJ_t^-) \quad ; \quad \mu_0 = \mu_0^+ - \mu_0^-$$

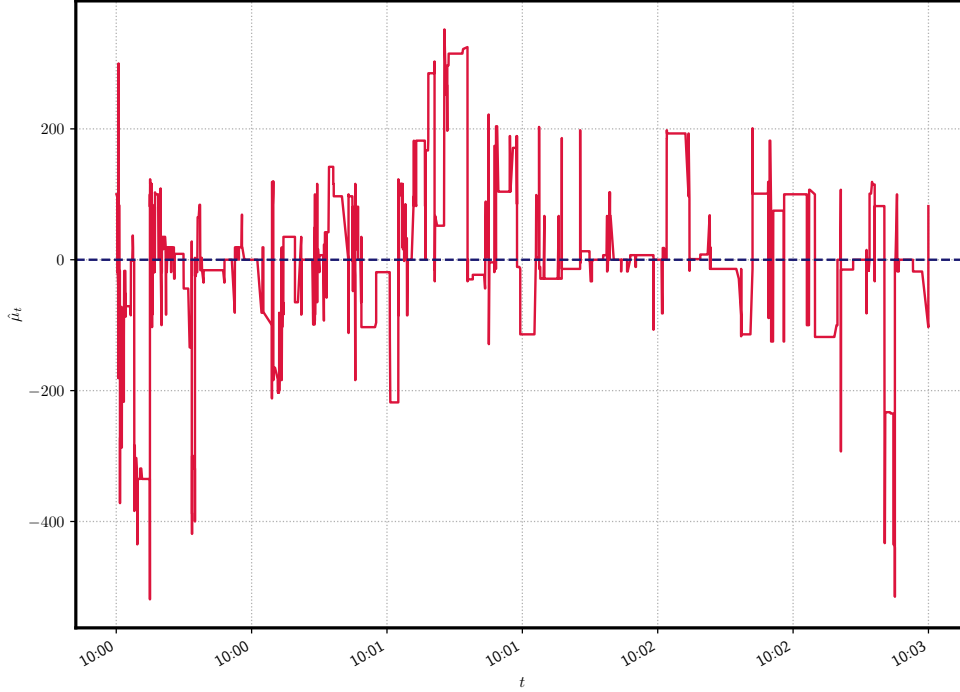


FIGURE 3.3 – Évolution du processus du flux d'ordre net (μ) de l'action AMZN estimé sur un horizon de 3 min entre 10h00min et 10h03min à la date 21/06/2012

Ainsi, $(S_t^\nu)_{t \in [0, T]}$ sera donné par :

$$dS_t^\nu = k(\mu_t - \nu_t) dt + \sigma dW_t$$

De ce fait, μ peut être ajouté au processus d'état ψ défini dans 3.1.1, i.e $\psi = (X, S, Q, \mu)^\top$, et se ramène à un problème de contrôle d'une diffusion à saut.

L'équation HJB associée au problème 3.1.2 est donnée dans ce cas par :

$$\begin{cases} \partial_t V + \frac{1}{2} \sigma^2 \partial_{SS}^2 V - \phi q^2 + \mathcal{L}_\mu V + \sup_{u > 0} (u(S - f(u)) \partial_x V - g(u) \partial_S V - u \partial_q V) = 0 & ; (t, \psi) \in [0, T] \times \mathbb{R}_+^3 \\ V(T, x, S, q) = x + q(S - \alpha q) & ; x \in \mathbb{R}_+^3 \end{cases} \quad (3.1.11)$$

Tel que \mathcal{L}_μ est l'opérateur différentiel donné, pour tout $(t, x, S, q, \mu) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^* \times \mathbb{R}$ par :

$$\begin{aligned} \mathcal{L}_\mu V(t, x, S, q, \mu) := & -\kappa\mu \partial_\mu V(t, x, S, q, \mu) + \lambda (V(t, x, S, q, \mu + \eta) - V(t, x, S, q, \mu)) \\ & + \lambda (V(t, x, S, q, \mu - \eta) - V(t, x, S, q, \mu)) \end{aligned}$$

Un contrôle optimal a été dérivée dans [119] (chapitre 7) dans le cas d'une fonction de coût de transactions quadratique, i.e $\exists \beta \in \mathbb{R}_+ : L(v) = \beta v^2$ pour $v \in \mathbb{R}_+$ et pour $\kappa \neq \gamma$. Il est donné par :

$$\forall t \in [0, T] : \quad \nu_t^* = -\gamma \frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} Q_t^{\nu^*} - \frac{k}{2\beta} \ell(t) \mu_t \quad (3.1.12)$$

Tel que γ et ζ sont les constantes définie dans la proposition 3.1.3 (en adaptant les notations du théorème à celles du présent cadre) et ℓ est la fonction définie par :

$$\forall t \in [0, T] : \quad \ell(t) := \frac{1}{\zeta e^{\gamma T} - e^{-\gamma T}} \left(e^{\gamma t} \frac{1 - e^{-(\kappa+\gamma)t}}{\kappa + \gamma} \zeta - e^{-\gamma t} \frac{1 - e^{-(\kappa-\gamma)t}}{\kappa - \gamma} \right)$$

Remarque que dans le cadre d'Almgren et Chriss présenté dans la section précédente, le taux de liquidation optimal ν^{AC} obtenu dans la proposition 3.1.3 est lié au processus d'inventaire par la relation suivante :

$$\forall t \in [0, T] : \quad \nu_t^{\text{AC}} = -\gamma \frac{1 + \zeta e^{2\gamma(T-t)}}{1 - \zeta e^{2\gamma(T-t)}} Q_t^{\nu^{\text{AC}}}$$

Ainsi, la différence entre ce cadre et celui d'Almgren Chriss est la présence d'un terme de perturbation du taux de liquidation due à un excès du flux d'ordres. Remarquer que dans le cas où α augmente (théoriquement elle tend vers $+\infty$), on a :

$$\nu_t^{\text{AC}} \xrightarrow{\alpha \rightarrow +\infty} \nu_t^\infty \geq 0$$

$$\ell(t) \xrightarrow{\alpha \rightarrow +\infty} \ell^\infty(t) \geq 0$$

Donc lorsque, à un instant $t \in [0, T]$, le flux d'ordre est incliné vers le côté d'achat (i.e $\mu_t > 0$), l'agent diminue son taux de liquidation, ce qui est tout à fait naturel, puisqu'il anticipe que cet excès d'achat va pousser le *midprice* à la hausse. De ce fait, il attend jusqu'à l'augmentation du prix pour augmenter son taux de liquidation. Dans le cas contraire d'un excès de vente (i.e $\mu_t < 0$), l'agent augmente son taux de liquidation. Car, naturellement, cette inclination vers le côté de vente va engendrer une diminution du *midprice* et rend le trading moins profitable si il attend. En somme, l'agent essaye de cacher son activité jusqu'à ce que le flux d'ordres va dans sa direction. Une autre remarque porte sur le signe de ν^* définie dans 3.1.12. En effet, il s'agit d'un processus stochastique et il peut être négatif dans certains cas. Ainsi, on le remplace dans la pratique par la correction suivante :

$$\forall t \in [0, T] : \quad \nu_t^\dagger = \left(\nu_t^{\text{AC}} - \frac{k}{2\beta} \ell(t) \mu_t \right)_+ \mathbb{1}_{\{Q_t^{\nu^\dagger} > 0\}}$$

Les figures 3.4 et 3.5 présentent quatre trajectoires simulés des processus étudiés. Le cadre numérique de la simulation est le suivant :

$$(\kappa, k, \beta, T, \lambda, S_0, \Xi, \phi, \alpha) = (10, 10^{-4}, 10^{-3}, 1, 1000, 50, 1, 10^{-2}, 100)$$

η est pris comme étant la réalisation d'une variable aléatoire exponentielle d'intensité égale à 5. Le lecteur peut bien vérifier la relation entre ν^* et le signe de μ soulignée auparavant.

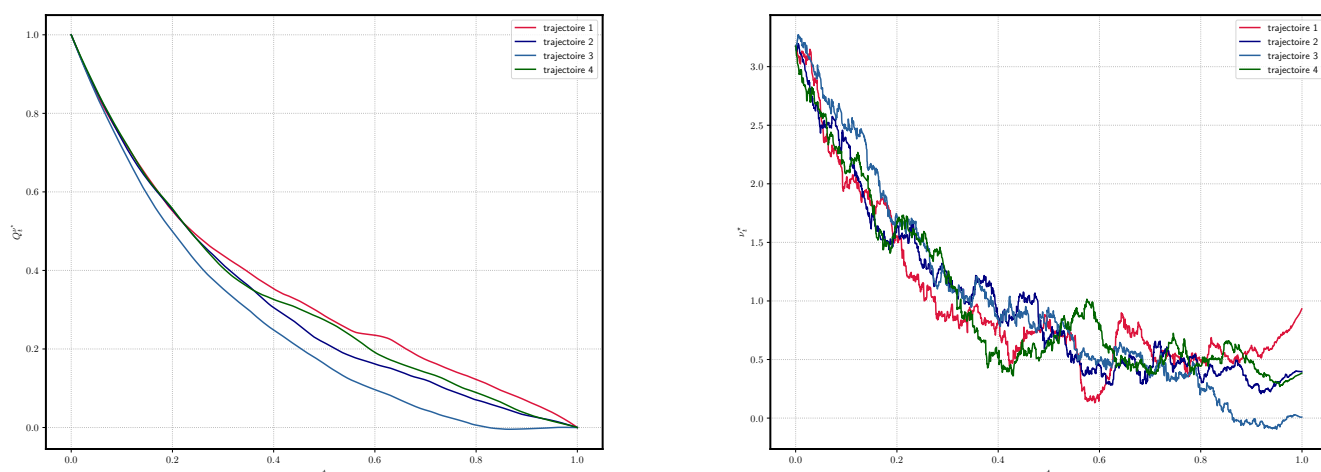


FIGURE 3.4 – 4 trajectoires du processus de l’inventaire et le taux de liquidation optimal

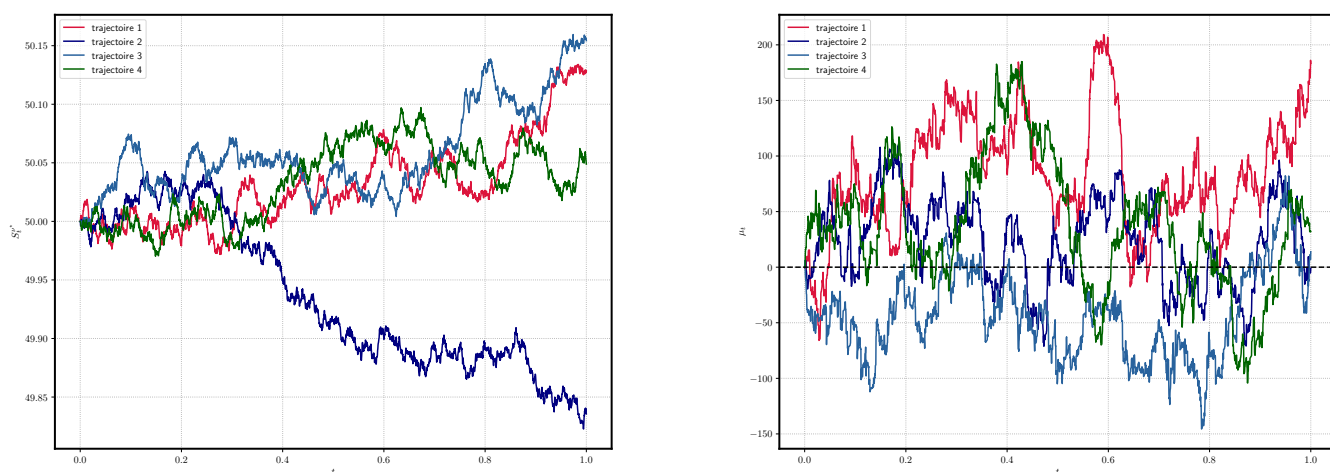


FIGURE 3.5 – 4 trajectoires du processus du *midprice* et du flux d’ordre net

Volume targeting : Pour des horizons longs (quelques minutes/heures), le cumul des ordres envoyé par l’agent peut exercer une forte pression unilatéral sur le LOB ce qui crée un fort *market impact*. Idéalement, la stratégie peut éviter cette sur-inclinaison adverse du flux d’ordres par la conception d’un algorithme qui camoufle les ordres de l’agent. Une manière d’achever cet objectif est de chercher un taux de liquidation/acquisition qui cible une fraction prédéterminée du volume totale durant la période d’exécution de la stratégie.

Généralement, deux types de stratégies peuvent être adoptés pour ce faire :

- **POV (Percentage Of Volume)** : Cibler une fraction de la vitesse avec laquelle les autres participants font leur trading,
- **PCOV (Percentage Of Cumulative Volume)** : Cibler une fraction du volume cumulé durant la période d’exécution.

Il est à noter que ces types de stratégies ne sont pas compatibles avec l’objectif de la liquidation totale de l’inventaire à l’instant T .

Les algorithmes de trading qui cible une stratégie *benchmark* sont utilisés de façon extensive dans l'industrie. Dans la section précédente, nous avons traité le cas dans laquelle l'agent maximise son P&L espérée ou, de manière équivalent, minimise son coût terminal espérée. Il s'agit de l'un de ces types de *benchmark* et est appelé : *Implementation Shortfall*. Une des stratégies les plus populaires dans l'industrie est celle qui cible ce qu'on appelle le VWAP (*Volume Weighted Average Price*). Le VWAP entre deux instants $t_1, t_2 \in [0, T]$ avec $t_1 < t_2$ est définie de la manière suivante :

$$\text{VWAP}(t_1, t_2) = \frac{\int_{t_0}^{t_1} S_t V_t dt}{\int_{t_0}^{t_1} V_t dt}$$

Tel que $(V_t)_{t \in [0, T]}$ est le processus du volume du marché que l'on suppose continue et strictement positive. $(S_t)_{t \in [0, T]}$ est le processus du *midprice*.

Le ciblage du VWAP est très difficile puisque l'agent ne connaît pas à priori la quantité future qui sera présente dans le marché durant la période d'exécution. Les investisseurs ciblent le VWAP puisqu'il désire avoir un P&L proche au P&L moyen obtenue par l'ensemble du marché durant la période d'exécution. De ce fait, un investisseur qui maximise son utilité espérée (e.g. CARA) pratiquant une activité de liquidation et qui cible le VWAP aura pour problème (en conservant les mêmes notation de la section précédente) :

$$\sup_{\nu \in \mathcal{A}} \mathbb{E} [-\exp(-(X_T - X_0 - \Xi \text{VWAP}(0, T)))]$$

Heuristiquement, une manière de cibler le VWAP et d'adopter le premier type de stratégies présenté ci-haut. En effet, si l'investisseur liquidera la totalité de sa position durant la période d'exécution qu'il fixe en ciblant la vitesse du trading des autres participants, son P&L sera proche au VWAP. Pour plus de détails, le lecteur est renvoyé au chapitre 9 de [119] ou au chapitre 4 de [54].

3.2 Exécution optimale à l'aide des ordres limites

3.2.1 Cadre théorique : liquidation optimale à la Avellaneda et Stoikov

Dans la section précédente, nous nous sommes concentrés sur des stratégies basées sur les ordres au marché (MOs). L'un des avantages de l'utilisation des MOs est le fait que l'exécution est garantie. Toutefois, le prix d'exécution est généralement plus mauvais que le *midprice* à cause de l'existence d'un spread non nul et le fait qu'un MO peut plonger dans le carnet d'ordre limite (*walking the LOB*). Dans la pratique, les agents utilisent les ordres limites (LOs) également car lorsque elles trouvent un MOs comme contrepartie, les LOs sont exécutées généralement à un prix plus profitable que le *midprice*. Toutefois, rien ne garantit l'arrivée d'un MOs qui peut s'exécuter contre un LO. À ce stade là, le lecteur doit se rappeler des différents notions et faits stylisés du trading dans le carnet d'ordre présentés dans les sous-sections 1.1.3 et 1.2.3. Dans cette section, on traitera le problème de la liquidation optimale à l'aide des LOs seulement. Il est à noter que dans ce cadre, on ne modélise pas le processus d'inventaire seulement mais également le processus comptant le nombre d'arrivée des MOs qui peuvent possiblement s'exécuter contre un LO posté par l'agent. Le modèle que nous présenterons est similaire à celui de Avellaneda et Stoikov [9].

Fixons un horizon de temps $T > 0$ et un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On définit sur cet espace un mouvement brownien naturel et standard $(W_t)_{t \in [0, T]}$. Le processus du *midprice* est

donné par :

$$\forall t \in [0, T] : S_t := S_0 + \sigma W_t$$

tel que $\sigma > 0$ est la volatilité du prix de l'actif.

Le nombre des ordres au marché (MOs) d'achat envoyés par les autres participants est modélisé par un processus de comptage associée à un processus ponctuel de poisson homogène d'intensité $\lambda > 0$ que l'on note $(M_t)_{t \in [0, T]}$. Le processus des distances entre les prix postés par l'agent et le *midprice* est noté par $(\delta_t)_{t \in [0, T]}$ et est supposé càglad (continue à gauche limitée à droite). Le nombre des MOs d'achat qui atteint le prix posté par l'agent est modélisée par un processus de comptage (contrôlé) $(N_t^\delta)_{t \in [0, T]}$ possédant une intensité qui dépend du processus δ . Le processus du *cash* est donné naturellement par :

$$dX_t^\delta = (S_t + \delta_t) dN_t^\delta ; \quad X_0^\delta := x \in \mathbb{R}$$

On suppose que l'agent possède initialement une quantité de l'actif d'intérêt égale à $\Xi \in \mathbb{N}^*$. Ainsi, le processus d'inventaire est donnée par :

$$\forall t \in [0, T] : Q_t^\delta := \Xi - N_t^\delta$$

On muni l'espace probabilisé de la filtration $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$ définie par :

$$\forall t \in [0, T] : \mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{F}_t^M \vee \mathcal{F}_t^{N^\delta}$$

On suppose de plus que δ soit \mathbb{F} - adapté.

Remarque 3.2.1

Si on se donne d'une fonction $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ convexe et décroissante sur \mathbb{R}_+ et un processus de comptage associée à un processus de poisson homogène d'intensité égale à 1 $(\tilde{N}_t)_{t \in [0, T]}$, on peut construire N^δ par le changement de temps stochastique :

$$\forall t \in [0, T] : N_t^\delta := \tilde{N}_{\int_0^t \lambda(\delta_s) ds}$$

On rappelle la forme particulière du *fill probability* (voir la figure 1.9). De ce fait, on suppose que l'intensité de N^δ soit de la forme :

$$\forall t \in [0, T] : \lambda_t := \lambda e^{-\kappa \delta_t}$$

. avec $\kappa > 0$. Remarque qu'il s'agit du choix suivant pour la fonction λ définie dans la remarque 3.2.1 :

$$\forall x \in \mathbb{R}_+ : \lambda(x) := \lambda e^{-\kappa x}$$

On définit le temps d'arrêt suivant :

$$\tau := T \wedge \left\{ t \in]0, T] : Q_t^\delta = 0 \right\}$$

. Ainsi le problème de l'agent est formulé de la même façon de la section précédente :

$$\forall (t, x, S, q) \in [0, T] \times \mathbb{R}^2 \times \mathbb{N} : V(t, (x, S, q)) := \sup_{\delta \in \mathcal{A}_t} \mathbb{E} \left(X_\tau^{t,x} + S_\tau^{t,S} Q_\tau^{t,q} + \ell(Q_T^{t,q}) \right) \quad (3.2.1)$$

tel que pour tout $t \in [0, T]$, l'ensemble des processus admissibles démarrant de l'instant t , notée \mathcal{A}_t , est constituée de tous les processus \mathbb{F} - prévisibles à valeurs dans \mathbb{R}_+ . $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_-$ est une fonction de pénalité de l'inventaire finale restante que l'on choisit de la forme suivante (pénalité quadratique) :

$$\forall x \in \mathbb{R}_+ : \ell(x) := -\alpha x^2$$

tel que $\alpha \in \mathbb{R}_+^*$.

L'équation HJB associée à ce problème est donnée par :

$$\begin{cases} \partial_t V + \frac{1}{2} \sigma^2 \partial_{SS}^2 V + \sup_{\delta \in \mathbb{R}_+} \mathcal{L}^\delta V = 0 & (t, (x, S, q)) \in]0, T[\times \mathbb{R}^2 \times \mathbb{N}^* \\ V(T, x, S, q) = x + q(S - \alpha q) & (x, S, q) \in \mathbb{R}^2 \times \mathbb{N} \\ V(t, x, S, 0) = x & (t, x, S, q) \in [0, T] \times \mathbb{R}^2 \times \mathbb{N} \end{cases} \quad (3.2.2)$$

tel que, pour $\delta \geq 0$, \mathcal{L}^δ est l'opérateur différentiel définie, pour tout $V \in \text{Dom}(\mathcal{L}^\delta)$ et $(t, (x, S, q)) \in]0, T[\times \mathbb{R}^2 \times \mathbb{N}^*$, par :

$$\mathcal{L}^\delta V(t, (x, S, q)) := \lambda e^{-\kappa \delta} (V(t, (x + S + \delta, S, q - 1)) - V(t, (x, S, q)))$$

Pour trouver une solution à l'équation HJB 3.2.2, on propose une forme similaire à celle proposée dans 3.2.3. Soit :

$$V(t, x, S, q) = \underbrace{x}_{\text{cash initial}} + \underbrace{qS}_{\text{valeur marché de l'inventaire initiale}} + \underbrace{\psi_q(t)}_{\text{profit/perte du à l'exécution}} \quad (3.2.3)$$

Plus concrètement, on a la proposition suivante :

Proposition 3.2.1

Soit $(\psi_q)_{q \in \mathbb{N}}$ une famille de fonction de $[0, T]$ vers \mathbb{R} définie de la façon suivante. Pour $q = 0$, $\psi_0 \equiv 0$ et pour $q \in \mathbb{N}^*$, ψ_q est solution du problème de Cauchy suivant :

$$\begin{cases} \psi_q' + \frac{\tilde{\lambda}}{\kappa} \exp(-\kappa(\psi_q - \psi_{q-1})) = 0 ; & t \in [0, T[\\ \psi_q(T) = -\alpha q^2 \end{cases} \quad (3.2.4)$$

avec $\tilde{\lambda} := \lambda e^{-1}$.

Alors, la fonction $V : [0, T] \times \mathbb{R}^2 \times \mathbb{N} \rightarrow \mathbb{R}$ définie par :

$$\forall (t, (x, S, q)) \in [0, T] \times \mathbb{R}^2 \times \mathbb{N} : \quad V(t, (x, S, q)) := x + qS + \psi_q(t) \quad (3.2.5)$$

est une solution de l'équation HJB 3.2.2 vérifiant :

$$\forall q \in \mathbb{N} : \quad (t, (x, S)) \mapsto V(t, (x, S, q)) \in \mathcal{C}^{1,2}([0, T[\times \mathbb{R}^2) \cap \mathcal{C}^0([0, T] \times \mathbb{R}^2)$$

Démonstration. Voir Annexe 3.3. □

Le théorème suivant affirme que la fonction V définie dans ?? est la solution du problème du contrôle 3.2.1

Théorème 3.2.1: Solution du problème du contrôle 3.2.1

Soit $(\psi_q)_{q \in \mathbb{N}_+^*}$ la famille de fonctions définie dans la proposition 3.2.1. Alors, la fonction définie dans 3.2.5 est une solution du problème du contrôle 3.2.1. De plus, le contrôle optimale est définie, pour tout $q \in \mathbb{N}$, par :

$$\forall t \in [0, T] : \quad \delta_t^{*,q} = \frac{1}{\kappa} + \psi_q(t) - \psi_{q-1}(t)$$

Démonstration. Voir Annexe 3.3. □

On présente par la suite la solution du système d'EDO couplés 3.2.4 afin d'en déduire les formules analytiques de V et δ^* .

Proposition 3.2.2

Soit $(\omega_q)_{q \in \mathbb{N}}$ la famille de fonction définie de $[0, T]$ vers \mathbb{R}_+^* par :

$$\forall t \in [0, T] : \quad \omega_q(t) := \sum_{k=0}^q \frac{\tilde{\lambda}^k}{k!} e^{-\kappa \alpha (q-k)^2} (T-t)^k$$

Alors la famille de fonctions $(\psi_q)_{q \in \mathbb{N}}$ définie par :

$$\forall t \in [0, T] : \quad \psi_q(t) := \frac{1}{\kappa} \ln (\omega_q(t))$$

est l'unique solution du système d'EDO couplés 3.2.4

Démonstration. Voir Annexe 3.3. □

D'après la proposition 3.2.2 et le théorème 3.2.1 on conclue que le *spread* de l'agent à tout instant est donné, pour tout $q \in \mathbb{N}^*$, par :

$$\forall t \in [0, T] : \delta_t^{*,q} = \frac{1}{\kappa} \left(1 + \ln \left(\frac{\sum_{k=0}^q \frac{\tilde{\lambda}^k}{k!} e^{-\kappa \alpha (q-k)^2} (T-t)^k}{\sum_{k=0}^{q-1} \frac{\tilde{\lambda}^k}{k!} e^{-\kappa \alpha (q-1-k)^2} (T-t)^k} \right) \right) \quad (3.2.6)$$

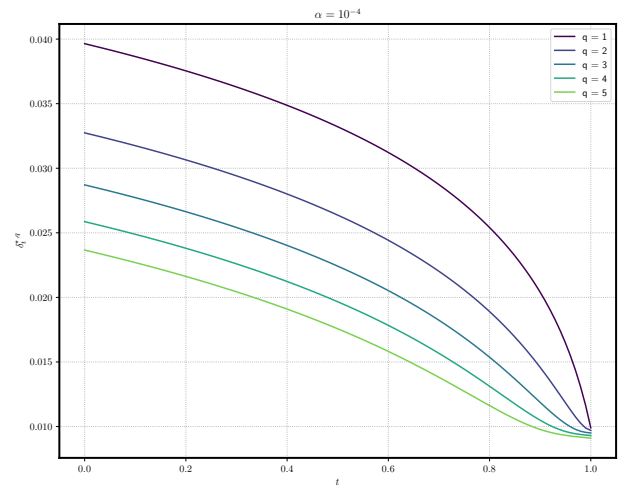
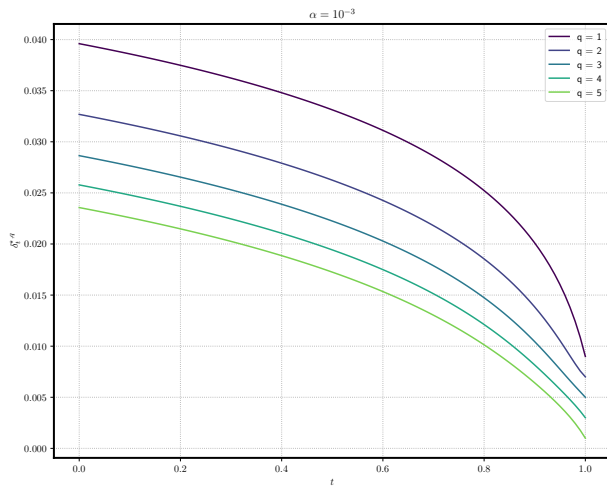


FIGURE 3.6 – graphiques de $t \mapsto \delta_t^{*,q}$ pour $q \in \llbracket 1, 5 \rrbracket$, $\alpha \in \{10^{-3}, 10^{-4}\}$, $\kappa = 100$, $\lambda = 50 \text{ min}^{-1}$ et $T = 1$.

D'après la formule de $\delta^{*,q}$ définie dans 3.2.6 et la figure 3.6, on peut déduire les remarques suivantes :

- $\delta^{*,q}$ sont décroissantes en q . Ce qui est naturel, puisque si l'agent possède un inventaire large, il est prêt à accepter une faible prime en échange de la fourniture de la liquidité au

marché, afin que d'augmenter ses *fill probabilities*. De plus, ceci assure que son inventaire sera liquidé à la date terminal ce qui lui permet de ne pas croiser le *bid-ask spread* par l'utilisation d'un MO et éviter de payer la pénalité terminale.

- pour $q \in \mathbb{N}^*$, $t \mapsto \delta^{*,q}$ est décroissante sur $[0, T]$. Ceci traduit le fait que l'agent devient de plus en plus averse à la détention d'une inventaire non nulle et recevoir une pénalité tant que le temps terminal s'approche.
- Quand le paramètre de la pénalité terminale augmente, $\delta^{*,q}$ diminue pour tout $q \in \mathbb{N}^*$. Ceci est tout à fait intuitif, comme l'augmentation de la pénalité pousse l'agent à liquider son inventaire rapidement à des prix bas.
- pour $q \in \mathbb{N}^*$, $t \mapsto \delta^{*,q}$ croît quand l'intensité de l'arrivée des MOs augmente. Là aussi la théorie est compatible avec les comportements observés sur le marché. En effet, si λ augmente, la probabilité d'arrivée d'un MO qui peut s'exécuter contre le LO posté par l'agent augmente. De ce fait, l'agent peut poster à une grande profondeur dans le carnet d'ordre ce qui lui permet de gagner plus de profits sans s'inquiéter sur le risque de non-exécution.

3.2.2 Application numérique

Simulation numérique : Afin d'explorer la stratégie développée ci-haut et illustrer la théorie étudiée, on réalise une simulation dans le cadre numérique suivant :

$$(T, \lambda, \kappa, \alpha, \Xi, S_0, \sigma) = (60 \text{ s.}, 50 \text{ min}^{-1}, 100, 10^{-3}, 5, 50, 10^{-2} \times \sqrt{60})$$

De manière à ce que l'agent possède 10% du volume moyen du marché sur $[0, T]$. On espère évaluer la performance de la stratégie en comparant le *price per share* noté $(PS_t)_{t \in [0, T]}$ et le TWAP qui utilisé souvent comme un *benchmark* de comparaison. Ils sont définis pour tout $t \in]0, T]$ par :

$$PS_t := \frac{X_t}{\Xi - Q_t}$$

$$TWAP_t := \frac{1}{t} \int_0^t S_u du$$

pour $N \in \mathbb{N}^*$, on choisit une subdivision $(t_n)_{n \in [0, N]}$ de $[0, T]$ de pas $\Delta t := T/N$. Rappelons que la probabilité qu'un LO posté à un prix $S_t + \delta$ peut être exécuté dans l'intervalle $[t_n, t_n + \Delta t]$ contre un MO avec une probabilité :

$$\lambda e^{-\kappa \delta} \Delta t$$

Les figures 3.7 et 3.8 contiennent trois trajectoires des processus simulés. On constate que dans ces trajectoires, l'agent arrive à liquider son inventaire avant la fin de la période d'exécution fixée à l'avance. Lorsqu'on combine le deux panels de la figure 3.7, on voit que dès que le LO posté par l'agent trouve une contrepartie, il augmente son spread pour générer plus de profits. Toutefois, si le LO n'est pas exécuté, il poste à prix très proche du *midprice*. Le panel droit de la figure 3.8 montre que la stratégie de l'agent sur-performe le TWAP dans les trois scénarios simulés. La raison majeure à cette sur-performance est l'utilisation des LOs qui lui permet de générer un profit supérieur à celui obtenu par le trading au *midprice*. Le panel droit de la figure 3.9 montrent une Heatmap du de l'inventaire de l'agent dans les 1000 scénarios simulés ainsi que l'inventaire moyenne et médiane. On remarque qu'en moyenne, l'agent n'arrive pas à liquider son inventaire à temps ce qui lui oblige à utiliser un MO pour terminer la liquidation. Toutefois, dans 50% des cas, il arrive à liquider son inventaire avant la fin de la période comme le montre l'inventaire médiane. Le panel gauche de la même figure montre la distribution du

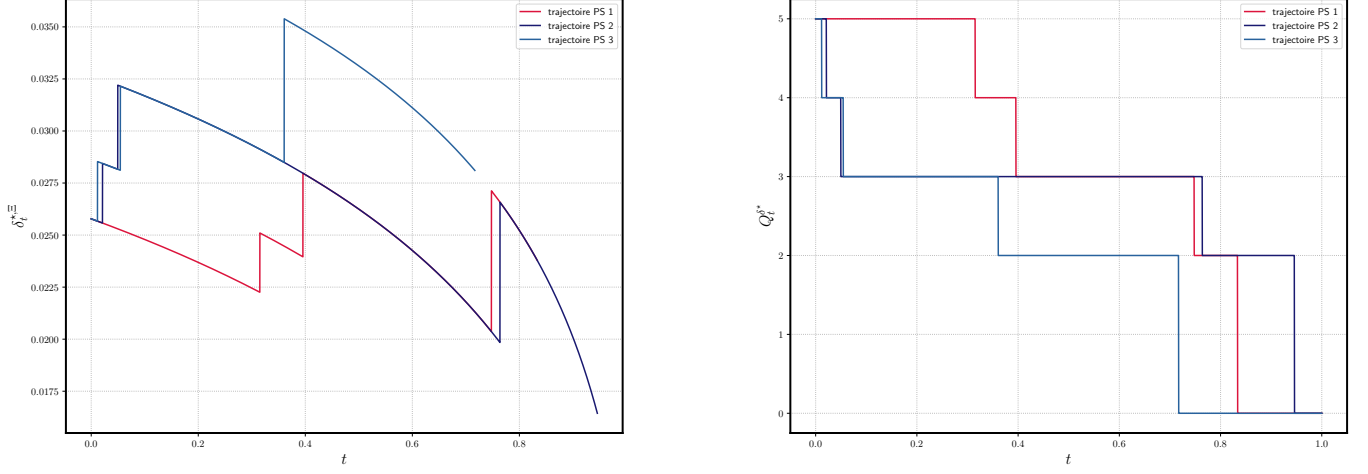


FIGURE 3.7 – 4 trajectoires du processus de l’inventaire et le taux de liquidation optimal

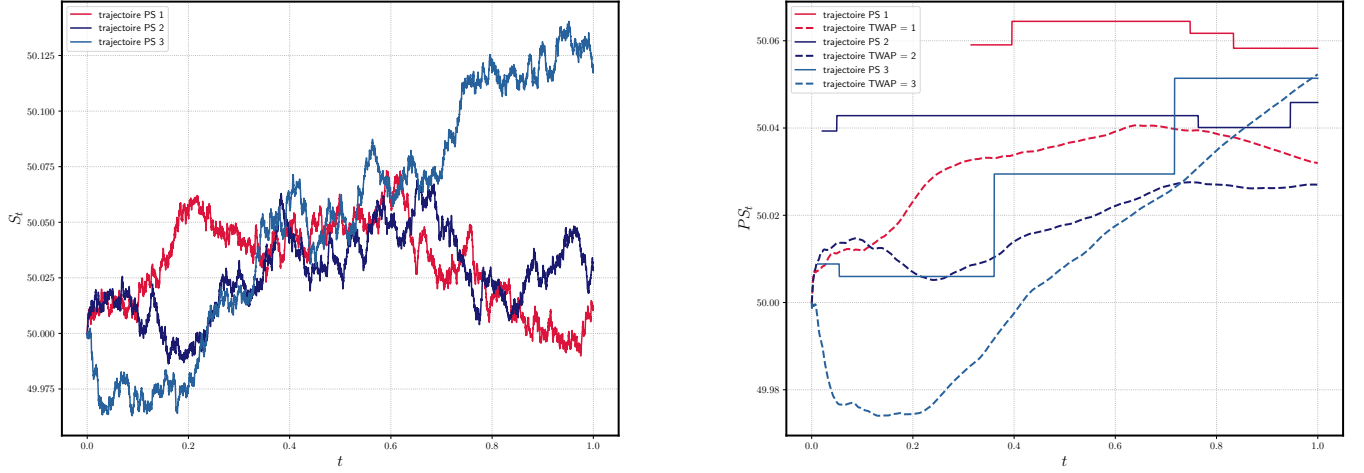


FIGURE 3.8 – 4 trajectoires du processus du *midprice* et du flux d’ordre net

P&L final de l’agent mesuré par (PS_T) et le $TWAP_T$. Là également, on conclue que la stratégie de l’agent a bien surperformé le TWAP comme la distribution est asymétrique et concentrée à droite.

Confrontation avec les données réelles : On s’intéresse dans ce paragraphe à la calibration du modèle présenté dans la section précédente. Le premier paramètre à estimer est la volatilité du prix de l’actif. Dans le cadre des données à haute fréquence, ce problème est loin d’être trivial et a été étudié extensivement dans la littérature [11], [118], [109]. Le problème majeur qui se manifeste lors de l’estimation de la volatilité à l’aide du données à haute fréquence et l’existence du bruit de la microstructure qui biaise l’estimateur classique [118] :

$$\hat{\sigma}_N = \left(\frac{1}{t_N - t_0} \sum_{k=1}^N (S_{t_k} - S_{t_{k-1}})^2 \right)^{1/2}$$

Plusieurs solutions ont été proposées dans la littérature dont on cite : le sous-échantillonnage, la modélisation du bruit de la microstructure ou la construction d’un modèle de prix comme

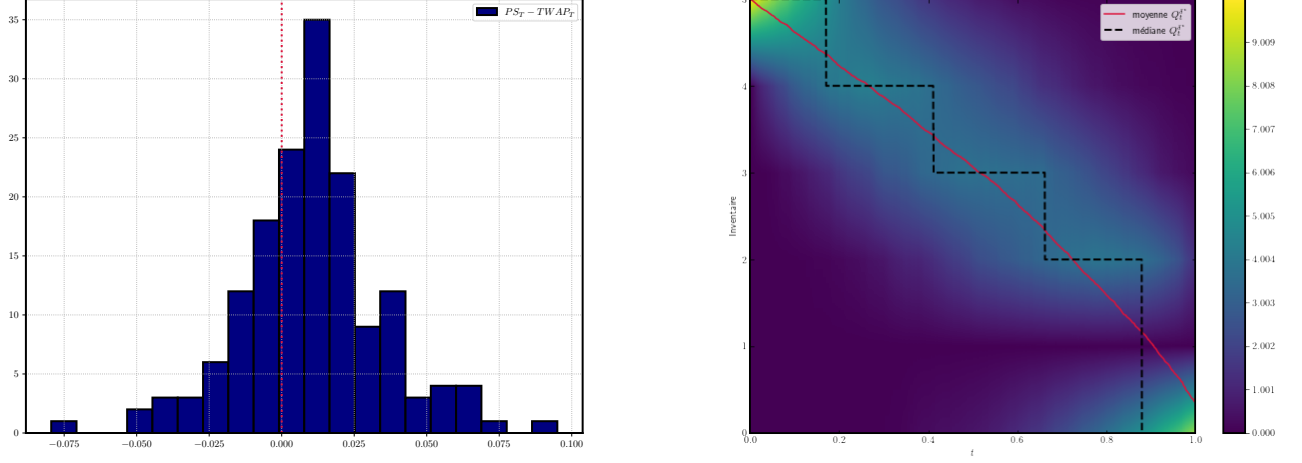


FIGURE 3.9 – histogramme de la différence entre le *price per share* final de l’agent et le $TWAP_T$ dans les différents scénaris (à gauche). Heatmap de l’inventaire de l’agent dans les différents scénaris (à droite)

étant une différence de deux processus de Hawkes [67], [11]. Dans la pratique, deux approches sont utilisées : ré-échantillonnage des données à une fréquence inférieure ou égale le pas Δt de la subdivision de $[0, T]$ ou l’utilisation d’une version adaptée au données à haute fréquence de l’estimateur de Garman & Klass [50].

Nous tournons maintenant, à la calibration du modèle de la liquidité. La calibration de ce modèle revient à estimer les paramètres $(A, \kappa) \in \mathbb{R}_+^2$, tel que :

$$\forall \delta \in \mathbb{R}_+ : \quad \lambda(\delta) := Ae^{-\kappa\delta}$$

avec $\delta \mapsto \lambda(\delta)$ est la fonction définissant l’intensité de N^δ .

La procédure de la calibration a été présenté dans le chapitre 4 de [47]. Supposons que pour un pas de temps $\Delta t > 0$ et un instant $t \in [0, T]$, l’algorithme poste un ordre au prix $S_t + \delta$ avec $\delta > 0$ et cet ordre reste inchangé durant l’intervalle $[t, t + \Delta t]$. De ce point de vue, le flux capturé ($N_{t+\Delta t}^\delta - N_t^\delta$) sera une variable aléatoire de poisson d’intensité :

$$\Lambda(\delta, t, t + \Delta t) := \int_t^{t+\Delta t} Ae^{-\kappa\delta + \kappa(S_u - S_t)} du$$

Ainsi :

$$\log(\mathbb{E}(\Lambda(\delta, t, t + \Delta t))) = \log(A) - \kappa\delta + \log\left(\int_t^{t+\Delta t} \mathbb{E}(e^{\kappa(S_u - S_t)}) du\right)$$

Supposons que $(S_t)_{t \in [0, T]}$ est un PAIS tel que la transformée de Laplace qui lui est associée existe et est finie (ce qui est le cas dans notre modèle). On note, ainsi, pour tout $u \in [0, T]$:

$$\varphi(\kappa, u) := \mathbb{E}(e^{\kappa(S_u - S_0)})$$

Pour $\delta \in \nu[[1, K]]$ tel que ν est la taille du *tick* et $\delta_{\max} := \nu K$ est la profondeur du carnet d’ordre limite, soit $\hat{\Lambda}(\delta)$ une estimateur de $\mathbb{E}(\Lambda(\delta, 0, \Delta t))$ Ainsi, les valeurs calibrées de (A, κ) correspondent à :

$$(\hat{A}, \hat{\kappa}) := \arg \min_{(A, \kappa) \in \mathbb{R}_+^2} \left\{ \sum_{j=1}^K \left(\log(\hat{\Lambda}(\delta_j)) + \kappa\delta_j - \log(A) - \log\left(\int_0^{\Delta t} \varphi(\kappa, u) du\right) \right)^2 \right\} \quad (3.2.7)$$

Dans notre modèle où S est modélisé par une diffusion brownienne, on a par une simple intégration :

$$\int_0^{\Delta t} \varphi(\kappa, u) du = \int_0^{\Delta t} \mathbb{E} \left(e^{\kappa \sigma W_u} \right) du \quad (3.2.8)$$

$$= \int_0^{\Delta t} e^{\frac{\kappa^2 \sigma^2 u}{2}} du \quad (3.2.9)$$

$$= \frac{2}{\kappa^2 \sigma^2} \left(e^{\frac{\kappa^2 \sigma^2 \Delta t}{2}} - 1 \right) \quad (3.2.10)$$

$$(3.2.11)$$

De ce fait, le problème de calibration 3.2.7 nécessite l'estimation de σ , chose que nous avons déjà traité, ainsi que l'estimation de $\Lambda(\delta) := \mathbb{E}(\Lambda(\delta, 0, \Delta t))$ pour tout $\nu \leq \delta \leq \delta_{\max}$. Sous l'hypothèse que les événements occurant dans chaque intervalle de temps sont indépendants de ceux qui arrivent dans l'intervalle suivant, le problème d'estimation de $\Lambda(\delta)$ revient à l'estimation de l'intensité d'une variable aléatoire de poisson. Nous présentons un estimateur présenté dans [?] et basé sur les temps d'attente jusqu'à l'exécution.

Pour $\nu \leq \delta \leq \delta_{\max}$ fixé, supposons que nous disposons d'un échantillon $(\tau_n)_{n \in \llbracket 1, N \rrbracket}$ de taille $N \in \mathbb{N}^*$ constitué des temps d'attente jusqu'à l'exécution de chaque agent. On définit, le temps d'attente jusqu'à l'exécution dans l'intervalle $[t, t + \Delta t]$ par :

$$\forall n \in \llbracket 1, N \rrbracket : \quad X_n = \tau_n \wedge \Delta t$$

De ce fait, l'auteur dans [47] a proposé l'estimateur suivant :

$$\hat{\Lambda}(\delta) := \frac{\sum_{n=1}^N \mathbb{1}_{\{X_n < T\}}}{\sum_{n=1}^N X_n}$$

Il a démontré, en plus, que cet estimateur est converge p.s et en L^p vers $\Lambda(\delta)$, avec biais et asymptotiquement normal [47] (section 4.4.2).

La figure 3.10 présente le graphique de $\delta \mapsto \hat{\Lambda}(\delta)$ calibrée aux données du carnet d'ordre de l'action AMZN à la date 21/06/2012 ainsi que la fonction $\delta \mapsto \hat{A}e^{-\hat{\kappa}\delta}$. Nous avons considéré un horizon de $T = 60$ s. entre 9h30 min et 9h31min.

Une fois estimés, les paramètres du modèle peuvent être passés à l'algorithme de posting afin d'exécuter la stratégie (voir le paragraphe précédent). Pour plus de détails sur la sensibilité de l'estimation des paramètres par rapport au choix de Δt ainsi qu'un traitement des différents problèmes numérique qui soulève dans ce cadre, le lecteur est renvoyé à [47] et [87].

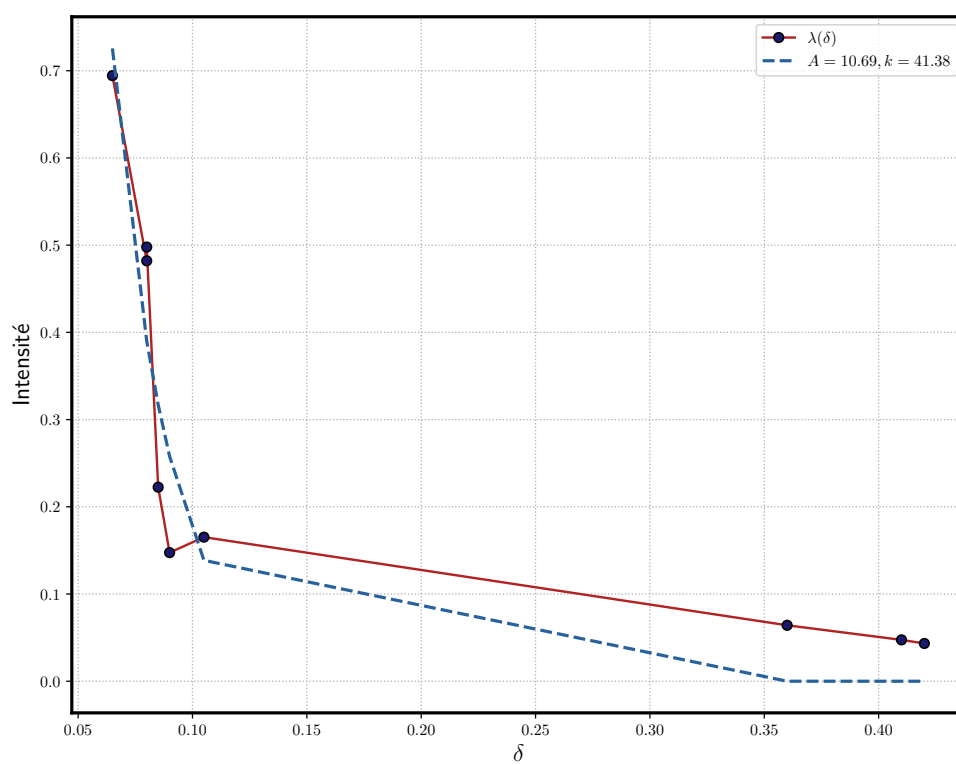


FIGURE 3.10 – Calibration de $\delta \mapsto \lambda(\delta)$ aux données de l'action AMZN à la date 21/06/2012

3.3 Annexe : Démonstrations des résultats

Proposition 3.1.2

Démonstration. Comme L est asymptotiquement sur-linéaire alors L^* est bien définie. De plus, L est strictement convexe, alors $L^* \in \mathcal{C}^1(\mathbb{R})$. De ce fait, pour tout $q > 0$, on a :

$$x \mapsto \frac{L(x)}{q^2} + \phi \in \mathcal{C}^1(\mathbb{R})$$

Ainsi, d'après le théorème de Cauchy-Lipschitz, pour tout $q > 0$, il existe une unique solution du problème 3.1.5. Soit $(\varphi_q)_{q>0}$ une famille de solutions du système 3.1.5. Montrons que la fonction V définie dans 3.1.6 est solution de l'équation HJB 3.1.3. En effet, comme pour tout $q > 0$, $\varphi_q \in \mathcal{C}^1([0, T]) \cap \mathcal{C}^0([0, T])$, alors $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*) \cap \mathcal{C}^0([0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*)$. Soit $(t, x, S, q) \in]0, T[\times \mathbb{R}_+^2 \times \mathbb{R}_+^*$. On a :

$$\partial_t V(t, x, S, q) = \varphi_q'(t)q^2 \quad ; \quad \partial_x V(t, x, S, q) = 1$$

$$\partial_S V(t, x, S, q) = q \quad ; \quad \partial_q V(t, x, S, q) = S + 2q\varphi_q(t)$$

$$\partial_{SS}^2 V(t, x, S, q) = 0$$

d'où :

$$\begin{aligned} \partial_t V(t, x, S, q) &= \phi q^2 - L^*(-q(k + 2\varphi_q)) \\ &= \phi q^2 - \sup_{u>0} (-q(k + 2\varphi_q)u - L(u)) \\ &= \phi q^2 - \sup_{u>0} (-qku - uf(u) - 2q\varphi_q(t)) \\ &= \phi q^2 - \sup_{u>0} (u(S - f(u))\partial_x V(t, x, S, q) - qg(u)\partial_S V(t, x, S, q) - u\partial_q V(t, x, S, q)) \end{aligned}$$

Or $\varphi_q(T) = -\alpha$, on a bien :

$$V(T, x, S, q) = x + qS - \alpha q^2$$

□

Théorème 3.1.1

Démonstration. On sait que V est solution de l'équation HJB 3.1.3. De ce fait, on va utiliser un argument de vérification 2.1.4.1 pour démontrer que elle s'agit bien d'une solution du problème du contrôle 3.1.2. Soient $n \in \mathbb{N}^*$ et $(t, \psi) = (t, (x, S, q)) \in [0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*$. Définissons le temps d'arrêt suivant :

$$\sigma_n := T \wedge \left\{ \tau > t : \left| X_\tau^{t,x} - x \right| \geq n, \left| S_\tau^{t,S} - S \right| \geq n, \left| Q_\tau^{t,q} - q \right| \geq n \right\}$$

On a bien :

$$\sigma_n \xrightarrow{\text{p.s.}} T$$

On procède par localisation. En effet, comme $V \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*)$ alors, par le théorème d'Itô on a :

$$V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) - V(t, \psi) = \int_t^{\sigma_n} \left(\partial_\tau V(\tau, \psi_\tau^{t,\psi}) + \left\langle \mu(\tau, \psi_\tau^{t,\psi}, \nu_\tau), \nabla_\psi V(\tau, \psi) \right\rangle \right) d\tau + \sigma \int_t^{\sigma_n} \partial_S V(\tau, \psi_\tau^{t,\psi}) dW_\tau$$

On a $\tau \mapsto Q_\tau^{t,x}$ est p.s bornée sur $[t, \sigma_n]$. φ_q est continue sur le compact $[t, \sigma_n(\omega)]$ pour tout $\omega \in \Omega$, alors elle est y bornée. De ce fait :

$$\partial_S V(\tau, \psi_\tau^{t,\psi}) = Q_\tau^{t,x} (1 + 2\varphi_q(\tau))$$

est p.s bornée sur $[t, \sigma_n]$. Alors :

$$\mathbb{E} \left(\int_t^{\sigma_n} \partial_S V(\tau, \psi_\tau^{t,\psi}) dW_\tau \right) = 0$$

ainsi :

$$\mathbb{E} \left(V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) \right) = V(t, \psi) + \mathbb{E} \left(\int_t^{\sigma_n} \left(\partial_\tau V(\tau, \psi_\tau^{t,\psi}) + \langle \mu(\tau, \psi_\tau^{t,\psi}, \nu_\tau), \nabla_\psi V(\tau, \psi) \rangle \right) d\tau \right)$$

En utilisant le fait que V résout l'équation HJB 3.1.3, On a presque sûrement :

$$\forall \tau \in [t, \sigma_n] : \left(\partial_\tau V(\tau, \psi_\tau^{t,\psi}) + \langle \mu(\tau, \psi_\tau^{t,\psi}, \nu_\tau), \nabla_\psi V(\tau, \psi) \rangle \right) \leq 0$$

soit :

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) \right) \leq V(t, \psi)$$

Maintenant, si on démontre que :

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) \right) = \mathbb{E} \left(V(T, \psi_T^{t,\psi}) \right),$$

on aura :

$$\mathbb{E} \left(V(T, \psi_T^{t,\psi}) \right) \leq V(t, \psi)$$

Avec égalité, si $\nu = \nu^*$ et par la définition de la borne supérieure, V est la solution du problème de contrôle. par la continuité de V sur $[0, T] \times \mathbb{R}_+^2 \times \mathbb{R}_+^*$ et celle de $\psi^{t,\psi}$, on a presque sûrement

$$V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) \xrightarrow{n \rightarrow +\infty} V(T, \psi_T^{t,\psi})$$

Maintenant, remarquons qu'en effectuant une intégration par partie stochastique, on a :

$$X_{\sigma_n}^{t,x} - x = - \left(Q_{\sigma_n}^{t,q} S_{\sigma_n}^{t,S} - qS \right) - k \int_t^{\sigma_n} Q_\tau^{t,q} \nu_\tau d\tau - \int_t^{\sigma_n} L(\nu_\tau) d\tau + \sigma \int_t^{\sigma_n} Q_\tau^{t,q} dW_\tau$$

De ce fait :

$$V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) = x + qS - k \int_t^{\sigma_n} Q_\tau^{t,q} \nu_\tau d\tau - \int_t^{\sigma_n} L(\nu_\tau) d\tau + \sigma \int_t^{\sigma_n} Q_\tau^{t,q} dW_\tau + \varphi_q(\sigma_n) \left| Q_{\sigma_n}^{t,q} \right|^2$$

Maintenant, remarquons que $t \mapsto Q_t$ est presque sûrement bornée par Ξ . De plus, par la définition de l'ensemble des contrôles admissibles, $\nu \in L^\infty([0, T] \times \Omega)$ ($\|\nu\|_{L^\infty} < \infty$). L étant strictement convexe sur \mathbb{R}_+ alors elle est y continue. De ce fait, comme ν prend ses valeurs presque sûrement dans $[0, \|\nu\|_{L^\infty}]$ alors $t \mapsto L(\nu_t)$ est presque sûrement bornée sur $[0, T]$ tout entier. Finalement, par la contunuité de φ_q sur $[0, T]$, on a presque sûrement :

$$\varphi_q(\sigma_n) \xrightarrow{n \rightarrow +\infty} -\alpha$$

Ainsi $(\varphi_q(\sigma_n))_{n \in \mathbb{N}^*}$ est bornée à partir d'un certain rang. D'où, presque sûrement

$$\exists K_1, K_2 > 0, \exists n_0 > 0, \forall n \geq n_0 : \left| V(\sigma_n, \psi_{\sigma_n}^{t,\psi}) \right| \leq x + qS + (\Xi k \|\nu\|_{L^\infty} + K_1) \sigma_n + \Xi \sigma |W_{\sigma_n} - W_t| + K_2 \Xi^2$$

Rappelons que $(\sigma_n)_{n \in \mathbb{N}^*}$ est bornée par T et par le théorème d'arrêt, on a : $(|W_{\sigma_n}|)_{n \in \mathbb{N}^*}$ est une sous-martingale. D'où :

$$\mathbb{E}(|W_{\sigma_n}|) \leq \mathbb{E}(|W_T|) < \infty$$

Ainsi,

$$\sup_{n \geq n_0} \mathbb{E} \left(\left| V \left(\sigma_n, \psi_{\sigma_n}^{t, \psi} \right) \right| \right) < \infty$$

Soit, par convergence dominée :

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(V \left(\sigma_n, \psi_{\sigma_n}^{t, \psi} \right) \right) = \mathbb{E} \left(V \left(T, \psi_T^{t, \psi} \right) \right)$$

Ce qui démontre que V est une solution du problème de contrôle 3.1.3 et ν^* défini dans l'énoncé du théorème est un contrôle optimal. \square

Proposition 3.2.1

Démonstration. Pour $q \in \mathbb{N}^*$, notons par S_q le problème de Cauchy 3.2.4. L'existence d'une solution de S_q pour tout $q \in \mathbb{N}^*$ se démontre par une récurrence sur q .

Définissons la fonction :

$$\begin{aligned} F: [0, T] \times \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (t, x, y) &\mapsto -\frac{\tilde{\lambda}}{\kappa} \exp(-\kappa(x - y)). \end{aligned}$$

De ce fait, l'EDO présente dans S_q se réécrit de la manière suivante :

$$\forall t \in [0, T]: \quad \psi'_q(t) = F(t, \psi_q(t), \psi_{q-1}(t))$$

Remarque que si pour $q \in \mathbb{N}^*$, ψ_{q-1} est une solution du système S_{q-1} , alors :

$$(t, x) \mapsto F(t, y, \psi_{q-1}(t))$$

est globalement continue est localement lipschitzienne en sa variable x sur $[0, T] \times \mathbb{R}$. De ce fait, d'après le théorème de Cauchy-Lipschitz, S_q admet une unique solution appartenant à $\mathcal{C}^1([0, T])$. Ainsi, l'existence d'une solution à S_q pour tout $q \in \mathbb{N}^*$ peut être démontrée par récurrence.

Pour démontrer que la fonction V définie dans la proposition est une solution de l'équation HJB 3.2.2, il suffit de voir que pour tout $(t, q) \in [0, T] \times \mathbb{N}^*$:

$$\arg \max_{\delta \in \mathbb{R}_+} \left\{ \lambda e^{-\kappa \delta} \underbrace{(V(t, (x + S + \delta, S, q - 1)) - V(t, (x, S, q)))}_{=\delta + h_{q-1}(t) - h_q(t)} \right\} = \left\{ \frac{1}{\kappa} + \psi_q(t) - \psi_{q-1}(t) \right\}$$

$$\sup_{\delta \in \mathbb{R}_+} \left\{ \lambda e^{-\kappa \delta} (V(t, (x + S + \delta, S, q - 1)) - V(t, (x, S, q))) \right\} = \frac{\tilde{\lambda}}{\kappa} \exp(-\kappa(\psi_q(t) - \psi_{q-1}(t)))$$

\square

Proposition 3.2.2 et théorème 3.2.1

Démonstration. Commençons par démontrer la proposition 3.2.2. En effet, $\omega_0 \equiv x \mapsto 1$ et ainsi $\psi_0 \equiv 0$. Soit $q \in \mathbb{N}^*$, on a bien $\omega_q \in \mathcal{C}^1([0, T])$ et :

$$\omega'_q + \tilde{\lambda} \omega_{q-1} = 0$$

De ce fait,

$$\psi'_q = \frac{\omega'_q}{\kappa \omega_q} = -\frac{\tilde{\lambda} \omega_{q-1}}{\kappa \omega_q} = -\frac{\tilde{\lambda}}{\kappa} \exp(-\kappa(\psi_q - \psi_{q-1}))$$

De plus :

$$\omega_q(T) = e^{-\kappa \alpha q^2} \implies \psi_q(T) = -\alpha q^2$$

Comme le système d'EDO couplées 3.2.4 admet une unique solution, $(\psi_q)_{q \in \mathbb{N}}$ définie dans la proposition est l'unique solution de ce système. \square

Démonstration. On passe maintenant à la démonstration du théorème 3.2.1. Là également, comme dans la démonstration du théorème 3.1.1, on utilise un argument de vérification (voir théorème 2.2.1). On note le processus d'état par : $Z = (X, S, Q)^\top$. On a bien :

$$dZ_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ dW_t \\ 0 \end{pmatrix} + \begin{pmatrix} S_t + \delta_t \\ 0 \\ -1 \end{pmatrix} dN_t^\delta$$

Or :

$$\forall q \in \mathbb{N} : (t, (x, S)) \mapsto V(t, (x, S, q)) \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^2) \cap \mathcal{C}^0([0, T] \times \mathbb{R}^2)$$

Alors par le théorème d'Itô pour les diffusions à sauts, on a :

pour tout $(t, z) = (t, (x, S, q)) \in]0, T] \times \mathbb{R}^2 \times \mathbb{N}^*$

$$\begin{aligned} V(T, Z_{T-}^{t,z}) - V(t, z) &= \int_t^T \left(\partial_\tau V(\tau, Z_{\tau-}^{t,z}) + \frac{\sigma^2}{2} \partial_{SS}^2 V(\tau, Z_{\tau-}^{t,z}) \right) d\tau + \sigma \int_t^T \partial_S V(\tau, Z_{\tau-}^{t,z}) dW_\tau \\ &\quad + \int_t^T \left(V(\tau, Z_{\tau-} + (S_\tau^{t,S} + \delta_\tau, 0, -1)) - V(\tau, Z_{\tau-}^{t,z}) \right) \lambda e^{-\kappa \delta_\tau} d\tau \\ &= \int_t^T \partial_\tau V(\tau, Z_{\tau-}^{t,z}) d\tau + \sigma \int_t^T \partial_S V(\tau, Z_{\tau-}^{t,z}) dW_\tau \\ &\quad + \int_t^T \left(V(\tau, (X_{\tau-}^{t,x} + S_\tau^{t,S} + \delta_\tau, S_{\tau-}^{t,S}, Q_{\tau-}^{t,q} - 1)) - V(\tau, Z_{\tau-}^{t,z}) \right) \lambda e^{-\kappa \delta_\tau} d\tau \end{aligned}$$

On démontre que le processus $(\partial_S V(\tau, Z_\tau^{t,z}))_{\tau \in [t, T]}$ est p.s borné. Or, pour $\tau \in [0, T]$:

$$\partial_S V(\tau, Z_\tau^{t,z}) = Q_\tau^{t,q} + \psi_{Q_\tau^{t,q}}(\tau)$$

et que $\tau \mapsto Q_\tau^{t,q}$ est p.s sûrement borné par Ξ sur $[0, T]$, il suffit de démontrer la bornétude du deuxième terme. En effet, pour tout $(t, q) \in]0, T] \times \mathbb{N}$, on a :

$$e^{-\kappa \alpha q^2} \left(1 + \sum_{k=1}^q \frac{(\tilde{\lambda}(T-t))^k}{k!} \right) \leq \omega_q(t) \leq \sum_{k=0}^q \frac{(\tilde{\lambda}(T-t))^k}{k!}$$

Or :

$$\sum_{k=0}^q \frac{(\tilde{\lambda}(T-t))^k}{k!} \leq \sum_{k=0}^{+\infty} \frac{(\tilde{\lambda}(T-t))^k}{k!} = e^{\tilde{\lambda}(T-t)} \leq e^{\tilde{\lambda}T} \quad ; \quad 1 + \sum_{k=1}^q \frac{(\tilde{\lambda}(T-t))^k}{k!} \geq 1,$$

on a bien :

$$-\alpha q^2 \leq \psi_q(t) \leq \frac{\tilde{\lambda}T}{\kappa}$$

Ainsi, on a presque sûrement, pour tout $\tau \in [t, T]$:

$$\left| \partial_S V \left(\tau, Z_{\tau^-}^{t,z} \right) \right| \leq \Xi + \alpha q^2 \vee \frac{\tilde{\lambda} T}{\kappa}$$

D'où :

$$\mathbb{E} \left(\int_t^T \partial_S V \left(\tau, Z_{\tau^-}^{t,z} \right) dW_{\tau} \right) = \mathbb{E} \left(\int_t^T \partial_S V \left(\tau, Z_{\tau^-}^{t,z} \right) dW_{\tau} \right) = 0$$

Maintenant, comme V est solution de l'équation HJB 3.2.2, on a presque sûrement :

$$\forall \tau \in [t, T] : \quad \partial_{\tau} V(\tau, Z_{\tau^-}^{t,z}) + \left(V(\tau, (X_{\tau^-}^{t,x} + S_{\tau}^{t,S} + \delta_{\tau}, S_{\tau^-}^{t,S}, Q_{\tau^-}^{t,q} - 1)) - V(\tau, Z_{\tau^-}^{t,z}) \right) \lambda e^{-\kappa \delta_{\tau}} \leq 0$$

Ainsi :

$$\mathbb{E} \left(V \left(T, Z_{T^-}^{t,z} \right) \right) = \mathbb{E} \left(V \left(T, Z_T^{t,z} \right) \right) \leq V(t, z) \quad (3.3.1)$$

En rappelons la démonstration 3.3, le contrôle optimal est définie par :

$$\forall t \in [0, T] : \quad \delta_t^{\star,q} = \frac{1}{\kappa} + \psi_q(t) - \psi_{q-1}(t)$$

on déduit qu'on a égalité dans 3.3.1 pour $\delta = \delta^{\star,q}$, ce qui conclue la démonstration par la définition de la borne sup. \square

Approximation stochastique et apprentissage on-line

4.1 Introduction

Dans ce chapitre, nous allons présenter les résultats essentiels de la théorie de l'approximation stochastique qui seront utilisés dans le chapitre suivant. On va, principalement, se concentrer sur le rôle de l'approximation stochastique dans la conception d'algorithmes récurrents d'optimisation en présence de l'incertitude ; ceci est particulièrement utile lorsque la source d'aléas peut être simulée numériquement ou obtenue à partir de données historiques. Il s'agit également d'un outil très puissant qui nous permet d'optimiser, d'une manière itérative et en temps réel, le comportement d'un système ; Dans ce cas, l'ensemble des techniques de l'approximation stochastique est également appelé : *on-line learning*.

La conception d'algorithmes d'optimisation récurrents à l'aide de la théorie de l'approximation stochastique n'est pas un phénomène nouveau. En effet, plusieurs méthodes ont été construites depuis la deuxième moitié du 20^{ème} siècle dont les plus célèbres sont les travaux de Robbins et Monro [108] et Keifer et Wolfowitz [75]. Comme à nos jours, nous témoignons une hausse dans la disponibilité et la vélocité des données, dans la vitesse des ordinateurs et le fait que la performance des systèmes industriels et financiers est mesuré statistiquement, ces méthodes s'adaptent parfaitement aux applications actuelles et gagnent une attention haussière de l'industrie.

En bref, étant donné une valeur d'intérêt que l'on espère approximer (point critique d'une fonction, zéro d'une application vectorielle, etc), l'idée de l'approximation stochastique est la conception d'une procédure récursive qui tente à approximer la valeur en utilisant une actualisation de "la meilleure estimation courante" par une autre en se basant sur un ensemble d'informations bruitées. Cet ensemble peut être un flux de données en temps réel, de données historique ou des données simulées.

D'une manière simple, si θ_n est "la meilleure estimation courante" à l'étape n , on veut concevoir un algorithme stochastique de la forme suivante :

$$\underbrace{\theta_{n+1}}_{\text{nouvelle valeur}} = \underbrace{\theta_n}_{\text{ancienne valeur}} - \underbrace{\gamma_{n+1}}_{\text{pas (poids)}} \times \underbrace{H(\theta_n, \overbrace{Y_{n+1}}^{\text{innovation}})}_{\text{correction basée sur le flux d'information entrant}}$$

de manière garantissant la convergence de la suite construite à notre valeur cible. Mathématiquement, ceci signifie que si $(Y_n)_{n \in \mathbb{N}}$ satisfait une propriété de stationnarité ergodique, $(\theta_n)_{n \in \mathbb{N}}$ converge presque sûrement vers la vraie valeur θ^* . Le cadre le plus simple, consiste en des innovation $(Y_n)_{n \in \mathbb{N}}$ i.i.d et peut être vu comme une sorte de généralisation de la loi forte des grands nombres (pour $\gamma_n := n^{-1}$, $H(\theta_n, Y_{n+1}) = \theta_n - Y_{n+1}$).

D'un point de vue mathématique, l'analyse de la convergence de ce type d'algorithmes se base fortement sur la théorie des martingales et la théorie de la stabilité des EDO. Plusieurs références de la théorie mathématique de l'approximation stochastique existent dans la littérature dont on cite entre autres : les livres de Duflo [39], Benveniste et al. [24], Kushner et al. [82] et Pagès [104]. On commence par présenter quelques exemples classiques d'algorithmes stochastiques avant de présenter la théorie mathématique de l'approximation stochastique.

4.2 Exemples d'algorithmes stochastiques

4.2.1 Algorithme du bandit à deux bras

Un exemple très intéressant de l'utilisation de l'approximation stochastique dans l'allocation optimale d'un budget entre deux agents est l'algorithme appelé : *Two-armed bandit*. Cette procédure récursive est très connue dans le domaine de la psychologie mathématique et l'apprentissage par renforcement.

Supposons que les actionnaires d'un fond dirigé par deux traders A et B veulent allouer le budget géré entre ces deux traders de manière à donner la fraction (du budget) la plus grande au trader le plus performant. Pour ce faire, ils décident de créer une procédure de ré-allocation périodique du fond entre les deux traders basée sur leurs performances (journalières, mensuelle, etc.). Cette procédure ne doit pas être "décourageante" pour les traders afin de garder leur confiance en soi. Une manière de ce faire est la suivante : Soit, pour $n \in \mathbb{N}$, $X_n \in [0, 1]$ la fraction du fond géré par le trader A durant la période n , ce qui laisse au trader B, la fraction $1 - X_n$. Chaque période, un trader entre les deux est choisi au hasard pour évaluer sa performance. Supposons qu'il s'agit du trader A. Si sa performance à la date n est considérée comme bonne, il sera récompensé à la date $n + 1$ par une allocation supplémentaire égale à $\gamma_{n+1} \in]0, 1[$ multipliée par la fraction gérée par le trader B durant la date n quelque soit la performance de ce dernier comme elle n'a pas été évaluée. De ce fait, le trader A va gérer une fraction $X_n + \gamma_{n+1}(1 - X_n)$ à la date $n + 1$. La même procédure s'applique au trader B s'il est choisi. Dans ce dernier cas, le trader A va gérer une fraction $X_n - \gamma_{n+1}X_n$ à la date $n + 1$.

Plus concrètement, fixons un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et modélisons les performances périodiques des deux traders par deux suites d'événement $(A_n)_{n \in \mathbb{N}}$ et $(B_n)_{n \in \mathbb{N}}$ définies sur cet espace, tel que : pour $n \in \mathbb{N}$, A_n (resp. B_n) : "la performance du trader A (resp. B) est bonne". Soit $(U_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires i.i.d de loi commune $\mathcal{U}_{[0,1]}$, tel que, si l'événement $\{U_{n+1} \leq X_n\}$ désigne le fait que le trader A est choisi à la fin de la date n pour être évalué. De ce fait, la dynamique de $(X_n)_{n \in \mathbb{N}}$ est décrite par :

$$\forall n \in \mathbb{N}^* : X_{n+1} = X_n - \gamma_{n+1} \left(X_n \mathbb{1}_{\{U_{n+1} > X_n\} \cap B_n} - (1 - X_n) \mathbb{1}_{\{U_{n+1} \leq X_n\} \cap A_n} \right) \quad ; \quad X_0 \in [0, 1]$$

tel que $(\gamma_n)_{n \in \mathbb{N}^*}$ est une suite de pas à valeurs dans $]0, 1[$ vérifiant :

$$\sum_{n=1}^{+\infty} \gamma_n = +\infty$$

La figure 4.1 montrent l'évolution de l'algorithme du bandit dans le cadre théorique suivant :

- $(\mathbb{1}_{A_n}, \mathbb{1}_{B_n})_{n \in \mathbb{N}}$ sont i.i.d tel que $p_A := \mathbb{P}(A_1)$ et $p_B := \mathbb{P}(B_1)$,
- $(U_n)_{n \in \mathbb{N}}$ et $(\mathbb{1}_{A_n}, \mathbb{1}_{B_n})_{n \in \mathbb{N}}$ sont indépendantes,
- $\forall n \in \mathbb{N}^* : \quad \gamma_n := \frac{1}{\sqrt{n}}$

les valeur des paramètres sont :

$$(X_0, p_A, p_B) = (0.7, 0.4, 0.7)$$

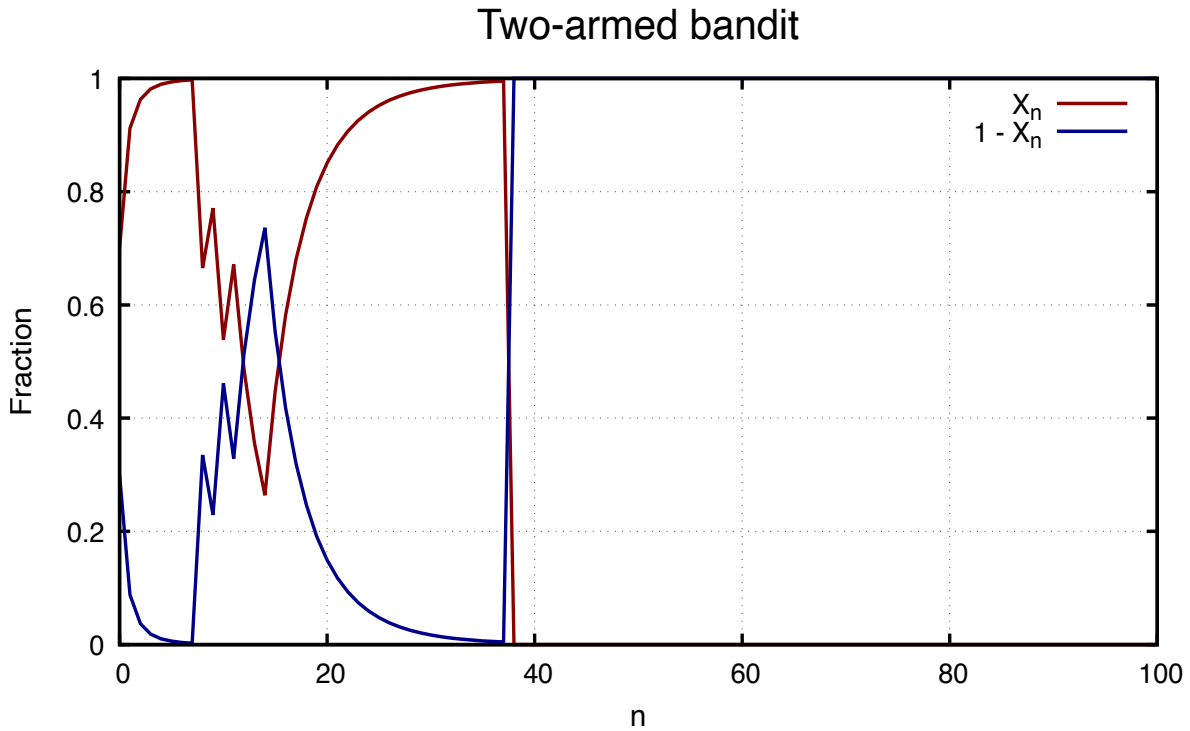


FIGURE 4.1 – Évolution des fractions gérées par les deux traders

Cet algorithme a été étudié d'une manière extensive dans la littérature. Pour une étude détaillé de sa convergence, le lecteur est renvoyé à [84] et [85].

Comme le lecteur a pu remarquer, la fonction de mise à jour H peut être déduite d'une manière heuristique de la construction de l'algorithme et une analyse mathématique ne sera utilisée que pour démontrer la convergence de l'algorithme. Toutefois, ceci n'est pas toujours comme dans de nombreuses applications, l'identification de la fonction H est loin d'être triviale.

4.2.2 Algorithme de Robbins-Monro

Soient $H : \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^d$ une fonction borélienne et Y est une variable aléatoire à valeurs dans \mathbb{R}^q tel que :

$$\forall \theta \in \mathbb{R}^d : \quad \mathbb{E}|H(\theta, Y)| < \infty$$

On s'intéresse à la recherche des zéros de la fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ définie par :

$$\forall \theta \in \mathbb{R}^d : \quad h(\theta) := \mathbb{E}(H(\theta, Y)) \quad (4.2.1)$$

Notons, $\theta^* \in \mathbb{R}^d$ s'il existe, une des zéros de h . Si l'expression de h est connue et suffisamment régulière, on peut achever notre objectif par la méthode de Newton-Raphson :

$$\theta_{n+1} = \theta_n - (Jh(\theta_n))^{-1} h(\theta_n)$$

avec $Jh(\theta_n)$ est la matrice jacobienne de h au point θ_n qui est supposée inversible.

L'idée de Robbins et Monro consiste à appliquer la même procédure précédente, non pas sur la fonction h , mais sur la fonction H en appliquant une randomisation de la procédure, soit :

$$\theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})$$

tel que, $(Y_n)_{n \in \mathbb{N}^*}$ est suite de variable aléatoires (innovations) i.i.d de même loi que Y et $(\gamma_n)_{n \in \mathbb{N}^*}$ est une suite numérique de pas à valeurs dans \mathbb{R}_+ vérifiant certaines conditions garantissant la convergence de l'algorithme.

Il est à noter que cette procédure peut être compris par deux perspectives :

- On ne sait pas l'expression de la fonction H , mais on peut observer la valeur $H(\theta_n, Y_{n+1})$ lorsqu'on contrôle l'entrée,
- On connaît l'expression de H , mais on ne connaît pas la loi gouvernant la suite des innovations à part quelques propriétés de la suite tel que le fait qu'elle soit : i.i.d, moyennisante, stationnaire ou une chaîne de Markov contrôlée. Ce cas est le plus répandu dans les applications en finance quantitative.

4.2.3 Recherche des extrémis dans un cadre stochastique

Le dernier exemple concerne la recherche des points critiques d'une fonction que l'on cherche à minimiser/maximiser et est le plus important pour nous dans le présent travail. On peut généralement distinguer deux situations. La première concerne une fonction qui admet une représentation sous forme d'une espérance d'une application différentiable (algorithme du gradient stochastique). La deuxième concerne la recherche des extrémis d'une fonction qui n'est observable qu'à travers une procédure de *trial-and-error* (algorithme de Kiefer-Wolfowitz).

Algorithme de descente du gradient stochastique : Supposons que l'on cherche à minimiser une fonction h de la forme introduite dans 4.2.1. Supposons de plus que :

$$\forall y \in \mathbb{R}^q : \quad \theta \mapsto H(\theta, y) \in \mathcal{C}^1(\mathbb{R}^d)$$

et que l'on a accès à l'expression de $\theta \mapsto \nabla_{\theta} H(\theta, \cdot)$ (gradient local). L'idée de cet algorithme est d'appliquer une randomisation de la méthode de la descente du gradient, qui est très connue dans la littérature de l'analyse numérique, de la manière suivante :

$$\theta_{n+1} = \theta_n - \gamma_{n+1} \nabla_{\theta} H(\theta_n, Y_{n+1})$$

Comme nous allons le voir, sous certaines conditions, cet algorithme converge vers la valeur recherchée.

Algorithme de Kiefer-Wolfowitz : On se place dans le même cadre précédant de la minimisation d'une fonction ayant une représentation sous forme d'une espérance, à l'exception que dans ce cadre l'expression du gradient local de H n'est pas disponible et que h est à valeurs dans \mathbb{R}_+ .

L'idée de Keifer et Wolfowitz est simplement d'approcher le gradient de h par des schémas de différence finie :

$$\forall \theta \in \mathbb{R}^d, \forall i \in \llbracket 1, d \rrbracket : \quad \partial_{\theta_i} h(\theta) \approx \frac{h(\theta + \varepsilon_i e_i) - h(\theta - \varepsilon_i e_i)}{2\varepsilon_i}$$

avec $(e_i)_{i \in \llbracket 1, d \rrbracket}$ est la base canonique de \mathbb{R}^d et $(\varepsilon_i)_{i \in \llbracket 1, d \rrbracket}$ est un vecteur de pas strictement positifs.

Ce terme de différence finie admet une représentation sous forme d'espérance donnée par :

$$\forall \theta \in \mathbb{R}^d, \forall i \in \llbracket 1, d \rrbracket : \quad \frac{h(\theta + \varepsilon_i e_i) - h(\theta - \varepsilon_i e_i)}{2\varepsilon_i} = \mathbb{E} \left(\frac{H(\theta + \varepsilon_i e_i, Y) - H(\theta - \varepsilon_i e_i, Y)}{2\varepsilon_i} \right)$$

Partant de cette représentation, on peut déduire une procédure récursive stochastique pour trouver $\arg \min_{\theta \in \mathbb{R}^d} h(\theta)$ comme suit :

$$\forall i \in \llbracket 1, d \rrbracket : \quad \theta_{n+1}^i = \theta_n^i - \gamma_{n+1} \left(\frac{H(\theta_n + \varepsilon_n^i e_i, Y_{n+1}) - H(\theta_n - \varepsilon_n^i e_i, Y_{n+1})}{2\varepsilon_n^i} \right)$$

La convergence de cet algorithme a été étudiée dans [75].

L'objectif de la présentation des exemples précédents est de familiariser le lecteur avec la construction des algorithmes stochastique ainsi que l'utilité de la théorie de l'approximation stochastique. On va, par la suite, donner un cadre mathématique rigoureux qui nous permettra de présenter les différents résultats de convergence de ces algorithmes.

4.3 Convergence des algorithmes stochastiques

4.3.1 Aperçu sur la théorie

Deux problèmes courants rencontrés en finance mathématique ou plus généralement en analyse numérique sont la recherche de l'ensemble des points où une fonction traverse un niveau donné et la recherche des points où une fonction atteint son minimum. Ces deux problèmes sont liés puisqu'on peut les ramener à des recherche de zéros d'une fonction. Dans un cadre déterministe, la recherche des zéros peut être achevée par une procédure récursive similaire à celle de Newton-Raphson :

$$\forall n \in \mathbb{N} : \quad \theta_{n+1} = \theta_n - \gamma_{n+1} h(\theta_n) \tag{4.3.1}$$

tel que $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction vectorielle continue vérifiant une propriété de croissance sous-linéaire au voisinage de $+\infty$. $(\gamma_n)_{n \geq 1}$ est une suite numérique à valeurs dans $]0, \gamma_0]$ tel que $\gamma_0 > 0$. Sous certaines conditions de "*mean-reversion*", on montre qu'une telle procédure est bornée et éventuellement converge vers un zéro θ^* de h .

L'approximation stochastique est une extension de cette méthode dans le cas où la fonction dont on cherche le zéro admet une représentation sous forme d'espérance, i.e : $h(\theta) = \mathbb{E}(H(\theta, Y))$ pour $\theta \in \mathbb{R}^d$, où Y est une variable aléatoire à valeurs dans \mathbb{R}^q de loi simulable (cf. 4.2.1). On construit alors une procédure récursive stochastique de la forme :

$$\forall n \in \mathbb{N} : \quad \theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})$$

où, $(Y_n)_{n \geq 1}$ sont i.i.d de même loi que Y . Cette procédure converge également vers un zéro de h sous des hypothèses appropriées, spécifiées par la suite, sur H et la suite de pas/gain $(\gamma_n)_{n \geq 1}$.

La propriété de *mean-reversion* citée auparavant peut être obtenue par une hypothèse de croissance faite sur h , en supposant que :

$$\forall \theta \in \mathbb{R}^d \setminus \{\theta^*\} : \quad \langle h(\theta), \theta - \theta^* \rangle > 0$$

Plus généralement, le *mean-reversion* peut découler de l'existence de ce qu'on appelle : **une fonction de Lyapunov**. Pour introduire cette notion, faisons tout d'abord une légère connexion avec les EDO et les système différentiels. En effet, lorsque $\gamma_n = \gamma > 0$ pour tout $n \geq 1$, l'équation 4.3.1 n'est que le schéma d'Euler associée à l'EDO :

$$\text{ODE}_h : y' = -h(y)$$

Une fonction de Lyapunov pour ODE_h est une fonction $L : \mathbb{R}^d \rightarrow \mathbb{R}_+$ tel que pour toute solution $t \mapsto x(t)$ de ODE_h , la fonction $L \circ x$ est décroissante. Dans le cas où L est différentiable, cette dernière condition est équivalente à :

$$\langle \nabla L, h \rangle \geq 0$$

Si une telle fonction de Lyapunov existe (ce qui n'est souvent pas le cas) le système différentiel est dit "dissipatif".

En ce qui concerne le rôle des fonction de Lyapunov dans l'analyse des algorithmes stochastiques, on peut généralement citer deux situations :

- La fonction L est identifiée à priori et est l'objet d'intérêt pour des raisons d'optimisation (e.g dans une procédure de gradient stochastique). Dans ce cadre, on peut concevoir une fonction h basée sur L , e.g : en posant $h = \nabla L$ ou plus généralement $h = \rho \nabla L$ avec ρ est une fonction presque partout positive.
- la fonction d'intérêt qui est impliquée naturellement dans le problème est h et on doit chercher une fonction de Lyapunov (qui peut ne pas exister). Ce cas requiert souvent une compréhension profonde de la nature dynamique du système étudié.

Deux approches sont généralement utilisées pour démontrer la convergence des algorithmes stochastique. **L'approche martingale** qui consiste en l'étude de la convergence de suite $L(\theta_n)_{n \geq 0}$ où, L est une fonction analogue à la fonction de Lyapunov. En utilisant, une propriété de convergence des surmartingales, l'hypothèse de *mean-reversion* et la divergence de la série de terme générale $(\gamma_n)_{n \geq 1}$ on peut démontrer la convergence presque sûre et en L^2 de l'algorithme. La deuxième approche, appelée : **approche de l'EDO** et considère l'équation définissant l'algorithme stochastique comme une version "perturbée" du schéma d'Euler de l'EDO $_h$. L'idée est de démontrer, via des arguments de topologie générale et d'analyse fonctionnelle, que les valeurs d'adhérence de la suite $(\theta_n)_{n \geq 0}$ est presque sûrement stable par le flot EDO $_h$.

Maintenant, supposons qu'aucun accès direct aux valeurs numériques de la fonction h n'est possible mais que h a une représentation sous forme d'une espérance donnée comme suit : fixons un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et soient $H : \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^d$ une fonction borélienne, Y une variable aléatoire définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{R}^q tel que :

$$\forall \theta \in \mathbb{R}^d : \quad \mathbb{E}|H(\theta, Y)| < \infty$$

et h la fonction définie par :

$$\forall \theta \in \mathbb{R}^d : \quad h(\theta) := \mathbb{E}(H(\theta, Y)) \tag{4.3.2}$$

On s'intéresse à la procédure récursive (algorithme stochastique) de recherche des zéros de h donnée par :

$$\forall n \in \mathbb{N} : \quad \theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1}) \quad (4.3.3)$$

où, $(\gamma_n)_{n \in \mathbb{N}^*}$ est une suite numérique à valeurs positives et $(Y_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{R}^q , i.i.d et de même loi que Y .

On espère que la suite de variables aléatoires $(\theta_n)_{n \in \mathbb{N}}$ converge (dans un certain sens) vers un des zéros de h .

4.3.2 Convergence p.s : l'approche martingale

L'approximation stochastique fournit des théorèmes variés qui garantissent la convergence p.s. et/ou dans L^p des procédures d'approximation stochastique. Nous donnons ci-dessous un résultat préliminaire général (multidimensionnel) connu comme le Lemme de Robbins-Siegmund, considéré comme un "métathéorème" de l'approximation stochastique, à partir duquel les principaux résultats de convergence seront facilement déduits (bien qu'il n'en contienne aucun stricto sensu).

Théorème 4.3.2.1: Robbins-Siegmund

Soit $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ la fonction définie par 4.3.2. Supposons qu'il existe une constante $K > 0$ et une fonction $L : \mathbb{R}^d \rightarrow \mathbb{R}_+$ continûment différentielle satisfaisant :

$$\nabla L \text{ est Lipchitz continue et } |\nabla L|^2 \leq K(1 + L) \quad (4.3.4)$$

tel que h satisfasse l'hypothèse de *mean-reversion* :

$$\langle \nabla L, h \rangle \geq 0$$

On suppose, de plus, que H satisfasse l'hypothèse de croissance peuso-linéaire suivante :

$$\forall \theta \in \mathbb{R}^d : \quad \|H(\theta, Y)\|_{L^2(\Omega, \mathbb{P})} \leq K\sqrt{1 + L(\theta)}$$

Soit $(\gamma_n)_{n \in \mathbb{N}^*}$ une suite de pas à valeurs dans \mathbb{R}_+ satisfaisant :

$$\sum_{n=1}^{+\infty} \gamma_n = +\infty \quad ; \quad \sum_{n=1}^{+\infty} \gamma_n^2 < +\infty \quad (4.3.5)$$

Finalement, supposons que θ_0 est indépendante de $(Y_n)_{n \geq 1}$ et que $\mathbb{E}(L(\theta_0)) < +\infty$.

Alors la procédure récursive définie par 4.3.3 satisfait :

- $\theta_n - \theta_{n-1} \xrightarrow{\mathbb{P}\text{-p.s. \& } L^2(\Omega, \mathbb{P})} 0$,
- La suite $(L(\theta_n))_{n \geq 0}$ est $L^1(\Omega, \mathbb{P})$ - bornée,
- $L(\theta_n) \xrightarrow{\mathbb{P}\text{-p.s.}} L_\infty \in L^1(\Omega, \mathbb{P})$
- $\sum_{n=1}^{+\infty} \gamma_n \langle \nabla L, h \rangle(\theta_{n-1}) < +\infty$

Démonstration. Voir théorème 6.1 de [104]

□

Remarque 4.3.2.1

- Si la fonction L satisfait aussi la propriété de coercivité : $\lim_{|\theta| \rightarrow +\infty} L(\theta) = +\infty$, alors L est souvent appelée fonction de Lyapunov du système comme dans la Théorie des Équations Différentielles Ordinaires (EDO).
- Notons que l'hypothèse 4.3.4 implique que $\nabla \sqrt{1+L}$ est bornée donc que \sqrt{L} a au plus une croissance linéaire (ce qui implique que L a au plus une croissance quadratique)
- Une lecture attentive de la preuve présentée dans [104] montre que l'hypothèse $\sum_{n=1}^{+\infty} \gamma_n = +\infty$ n'est pas nécessaire. Cependant, nous la gardons puisqu'elle est indispensable pour n'importe quelle application de ce Lemme. Heuristiquement, cette hypothèse assure que la valeur initiale de l'algorithme sera oubliée. Ces d'hypothèses sont connues dans la littérature comme "hypothèses de Robbins-Siegmund".

On présente par la suite deux corollaires importants du théorème de Robbins-Siegmund 4.3.2.1.

Corollaire 4.3.2.1

- **ALGORITHME DE ROBBINS-MONRO** : Supposons que la fonction moyenne h soit continue et séparante en un point $\theta^* \in \mathbb{R}^d$, i.e :

$$\forall \theta \in \mathbb{R}^d \setminus \{\theta^*\} : \langle h(\theta), \theta - \theta^* \rangle > 0$$

ce qui implique que $\{h = 0\} = \{\theta^*\}$. Supposons de plus que $\theta_0 \in L^2(\Omega, \mathbb{P})$ et qu'il existe $K > 0$ tel que H vérifie :

$$\forall \theta \in \mathbb{R}^d : \mathbb{E}(|H(\theta, Y)|^2) \leq K(1 + |\theta|^2)$$

Supposons que la suite de pas vérifie 4.3.5, alors :

$$\theta_n \xrightarrow{\text{p.s.} \& L^p(\Omega, \mathbb{P})} \theta^*$$

pour tout $p \in]0, 2[$.

- **GRADIENT STOCHASTIQUE** : Soit $L : \mathbb{R}^d \rightarrow \mathbb{R}_+$ une fonction différentiable vérifiant 4.3.4 et la propriété de coercivité : $\lim_{|\theta| \rightarrow +\infty} L(\theta) = +\infty$, ce qui implique que :

$$\arg \min_{\theta \in \mathbb{R}^d} L(\theta) \subseteq \{\nabla L = 0\}$$

On suppose que $\{\nabla L = 0\} = \{\theta^*\}$ et la fonction moyenne h soit donnée par $h := \nabla L$. On suppose, de plus, que H vérifie :

$$\forall \theta \in \mathbb{R}^d : \mathbb{E}(|H(\theta, Y)|^2) \leq K(1 + |\theta|^2),$$

que $\theta_0 \in L^1(\Omega, \mathbb{P})$ et que la suite de pas $(\gamma_n)_{n \in \mathbb{N}^*}$ vérifie 4.3.5. Alors, $L(\theta^*) = \min_{\theta \in \mathbb{R}^d} L(\theta)$,

$$\theta_n \xrightarrow{\text{p.s.}} \theta^* \quad \text{et} \quad \nabla L(\theta_n) \xrightarrow{L^p(\Omega, \mathbb{P})} 0$$

pour tout $p \in]0, 2[$.

Démonstration. Voir corollaire 6.1 de [104]. □

4.3.3 Point de vue de l'optimisation : approche de Keifer-Wolfowitz

On revient au problème d'optimisation décrit dans 4.2.3. Nous donnons ci-dessous le résultat de convergence pour les procédures de Kiefer-Wolfowitz (qui est la contrepartie naturelle du Lemme de Robbins-Siegmund dans le cadre du gradient stochastique)

Théorème 4.3.3.1: Keifer-Wolfowitz

Supposons que la fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}_+$ soit deux fois différentiable sur \mathbb{R}^d de Hessienne Lipschitzienne. Supposons, de plus, que $\theta \mapsto \|H(\theta, Y)\|_{L^2(\Omega, \mathbb{P})}$ a au plus une croissance sous-linéaire et que les suites de pas vérifient :

$$\forall i \in \llbracket 1, d \rrbracket : \quad \sum_{n=1}^{+\infty} \eta_n^i = \sum_{n=1}^{+\infty} \gamma_n = +\infty, \quad \sum_{n=1}^{+\infty} \gamma_n < +\infty, \quad \eta_n \xrightarrow{n \rightarrow +\infty} 0_{\mathbb{R}^d}, \quad \text{et} \quad \sum_{n=1}^{+\infty} \left(\frac{\gamma_n}{\eta_n^i} \right)^2 < +\infty$$

Alors, $(\theta_n)_{n \geq 0}$ converge p.s vers une composante connexe de $\{h = \alpha\} \cap \{\nabla h = 0\}$ pour un niveau $\alpha \in \mathbb{R}_+$.

Démonstration. Voir [75]. □

Il est à noter que cette procédure requiert un soin particulier lors du choix des suites $(\gamma_n)_{n \geq 1}$ et $(\eta_n)_{n \geq 1}$. Le problème des pièges (convergence vers un minima local de h , un point selle, etc) rencontré lors de l'utilisation de l'approche de Keifer-Wolfowitz a été étudié dans [91]. Plus de détails sur l'approche, le lecteur est renvoyé à [24].

4.3.4 Convergence p.s : approche de L'EDO

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ un espace probabilisé filtré sur lequel on définit la procédure récursive suivante, pour $n_0 \in \mathbb{N}$:

$$\forall n \geq n_0 : \quad \theta_{n+1} = \theta_n - \gamma_{n+1} h(\theta_n) + \gamma_{n+1} (\Delta M_{n+1} + r_{n+1}) \quad (4.3.6)$$

où $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction continue et localement lipschitzienne, θ_{n_0} est un vecteur aléatoire \mathcal{F}_{n_0} mesurable et, pour tout $n \geq n_0$, ΔM_{n+1} est un \mathcal{F}_n -accroissement de martingale et $(r_n)_{n \geq n_0}$ est un terme de reste (\mathcal{F}_n) -adapté.

Le théorème de convergence p.s suivant découle des liens subtils entre le comportement asymptotique des algorithmes stochastiques et celui des EDO (cf. [82] (chapitre 5) ou [24] (chapitre 2)).

Théorème 4.3.4.1: Méthode de l'EDO

Supposons que h soit localement Lipschitzienne, que :

$$r_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} 0 \quad \text{et} \quad \sup_{n \geq n_0} \mathbb{E}(|\Delta M_{n+1}|^2 | \mathcal{F}_n) < +\infty \quad \text{p.s.}$$

et que $(\gamma_n)_{n \in \mathbb{N}^*}$ suite de pas à valeurs dans \mathbb{R}_+ satisfaisant :

$$\sum_{n=1}^{+\infty} \gamma_n = +\infty \quad ; \quad \sum_{n=1}^{+\infty} \gamma_n^2 < +\infty$$

Alors l'ensemble Θ^∞ des valeurs d'adhérence de la suite $(\theta_n)_{n \geq n_0}$ est p.s un ensemble compact connexe, stable par le flot de :

$$\text{EDO}_h : y' = -h(y)$$

et par :

$$\text{EDO}_h^* : y' = h(y)$$

De plus, si $\theta^* \in \Theta^\infty$ est l'unique équilibre uniformément stable sur Θ^∞ de EDO_h , alors :

$$\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \theta^*$$

COMMENTAIRE : Par uniformément stable, on entend :

$$\sup_{\theta \in \Theta^\infty} |\theta(\theta_0, t) - \theta^*| \xrightarrow[t \rightarrow +\infty]{} 0$$

où $(\theta(\theta_0, t))_{(t, \theta_0) \in \mathbb{R}_+ \times \Theta^\infty}$ est le flot de EDO_h sur Θ^∞ .

Démonstration. Voir [22] ou [104] (chapitre 6, section 6.4). □

La convergence de la dernière procédure récursive peut être obtenue sous des conditions plus faibles (cf. [96], [22]).

4.3.5 Vitesse de convergence : le TCL

Dans des cadre standards, sous des conditions "naturelles", un algorithme stochastique converge vers sa cible à une vitesse $\sqrt{\gamma_n}$ (ce qui suggère d'utiliser des pas de la forme $\gamma_n = \frac{a}{b+n}$ où $a, b > 0$ pour obtenir une grande vitesse de convergence). Plus précisément, on peut démontrer que $\left(\frac{\theta_n - \theta^*}{\sqrt{\gamma_n}}\right)_{n \geq n_0}$ converge en loi vers un vecteur gaussien qui dépend de la matrice jacobienne $Jh(\theta^*)$. Nous nous plaçons ici dans le cadre d'algorithme stochastique avec terme de reste, i.e. du type 4.3.6.

Théorème 4.3.5.1: TCL (cf. [104], [24], [39], [82])

Soit $\theta^* \in \{h = 0\}$. Supposons que la fonction h soit différentiable au point θ^* et que toutes les valeurs propres de $Jh(\theta^*)$ aient une partie réelle positive. Supposons que, pour un certain $\delta > 0$:

$$\sup_{n \geq n_0} \mathbb{E} \left(|\Delta M_{n+1}|^{2+\delta} \mid \mathcal{F}_n \right) < +\infty \quad \text{p.s.}, \quad \mathbb{E} \left(\Delta M_{n+1} \Delta M_{n+1}^\top \mid \mathcal{F}_n \right) \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \Gamma$$

où $\Gamma \in \mathcal{S}_d(\mathbb{R})$ (déterministe et symétrique) est définie-positive et pour $\varepsilon > 0$:

$$\mathbb{E} \left((n+1) |r_{n+1}|^2 \mathbb{1}_{\{|\theta_n - \theta^*| \leq \varepsilon\}} \right) \xrightarrow[n \rightarrow +\infty]{} 0$$

Définissons la suite de pas comme suit :

$$\forall n \in \mathbb{N}^* : \quad \gamma_n := \frac{\alpha}{n}, \quad \alpha > \frac{1}{2\text{Re}(\lambda_{\min})}$$

où $\lambda_{\min} := \arg \min_{\lambda \in \text{Sp}_{\mathbb{C}}(Jh(\theta^*))} \text{Re}(\lambda)$, avec $\text{Re}(z)$ est la partie réelle du nombre complexe z et $\text{Sp}_{\mathbb{C}}(A)$ est le spectre complexe de la matrice A . Autrement dit, λ_{\min} est la valeur propre

complexe de $Jh(\theta^*)$ ayant la plus petite partie réelle. Alors, la convergence p.s est régie sur l'ensemble de convergence $\{\theta_n \xrightarrow[n \rightarrow +\infty]{p.s} \theta^*\}$ par le Théorème Central Limite suivant :

$$\sqrt{n}(\theta_n - \theta^*) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}_d(0, \alpha \Sigma)$$

où :

$$\Sigma := \int_0^{+\infty} \left[\exp \left(- \left(Jh(\theta^*) - \frac{1}{2\alpha} I_d \right) u \right) \right]^\top \Gamma \exp \left(- \left(Jh(\theta^*) - \frac{1}{2\alpha} I_d \right) u \right) du$$

4.3.6 Principe de moyennisation (Ruppert-Polyak)

L'implémentation pratique des algorithmes stochastiques récurrents montre que même si la convergence est régie par un TCL, elle reste "chaotique" même dans les dernières étapes de l'algorithme, sauf si la suite des pas est choisie de manière que $(\theta_n)_{n \geq 0}$ ait asymptotiquement une variance minimale.

L'idée originale de l'introduction du principe de moyennisation était de "lisser" le comportement d'un algorithme stochastique convergeant en considérant la moyenne arithmétique des valeurs passées plutôt que la valeur calculée à la $n^{\text{ème}}$ itération. De façon surprenante, si on combine cette procédure de moyennisation et une suite de pas qui "décroît lentement", on atteint gratuitement la meilleure vitesse de convergence possible.

Plus précisément, soit $(\gamma_n)_{n \geq 1}$ la suite de pas définie par :

$$\forall n \in \mathbb{N}^* : \quad \gamma_n := \left(\frac{\alpha}{n + \beta} \right)^\vartheta, \quad \vartheta \in]1/2, 1[, \alpha > 0, \beta \geq 0$$

On se place dans le cadre de la procédure récursive usuelle 4.3.3 et on pose :

$$\forall n \in \mathbb{N}^* : \quad \bar{\theta}_n := \frac{1}{n} \sum_{k=0}^n \theta_k$$

Remarquer que $(\bar{\theta}_n)_{n \geq 1}$ vérifie la procédure récursive suivante :

$$\forall n \in \mathbb{N}^* : \quad \bar{\theta}_{n+1} = \bar{\theta}_n - \frac{1}{n+1} (\bar{\theta}_n - \theta_n) \quad ; \quad \bar{\theta}_0 = 0$$

Sous des hypothèses naturelles (cf. [106]), on montre dans différents cas étudiés précédemment (gradient stochastique, Robbins-Monro) que :

$$\bar{\theta}_n \xrightarrow[n \rightarrow +\infty]{p.s} \theta^*$$

où θ^* est la cible de l'algorithme et de plus :

$$\sqrt{n}(\bar{\theta}_n - \theta^*) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}_d(0, \Sigma_{\min}^*)$$

où Σ_{\min}^* est la matrice des variances-covariances asymptotiquement minimale parmi toutes les matrices des variances-covariances que l'on peut obtenir. En particulier, dans le cas où $d =$, on a :

$$\Sigma_{\min}^* = \frac{\text{Var}(H(\theta^*, Y))}{h'(\theta^*)^2}$$

Théorème 4.3.6.1: Ruppert & Polyak (cf. [39])

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ un espace probabilisé filtré sur lequel on définit la procédure réursive suivante :

$$\forall n \in \mathbb{N} : \quad \theta_{n+1} = \theta_n - \gamma_{n+1} (h(\theta_n) + \Delta M_{n+1})$$

où, $(\Delta M_n)_{n \geq 1}$ est un processus d'accroissements de martingale et h est une fonction borélienne continue en son unique zéro θ^* , vérifiant :

$$\forall \theta \in \mathbb{R}^d : \quad h(\theta) = Jh(\theta^*) (\theta - \theta^*) + \mathcal{O}(|\theta - \theta^*|^2)$$

où toutes les valeurs propres de $Jh(\theta^*)$ ont une partie réelle strictement positive. Supposons de plus, qu'il existe une constante $C > 0$ tel que :

$$\forall n \in \mathbb{N} : \quad \mathbb{E}(\Delta M_{n+1} \mid \mathcal{F}_n) \mathbf{1}_{\{|\theta_n - \theta^*| \leq C\}} = 0 \quad \text{p.s}$$

et qu'il existe un exposant $\delta > 0$ tel que :

$$\begin{aligned} \sup_{n \in \mathbb{N}} \mathbb{E}(|\Delta M_{n+1}|^{2+\delta} \mid \mathcal{F}_n) \mathbf{1}_{\{|\theta_n - \theta^*| \leq C\}} &< +\infty \quad \text{p.s} \\ \mathbb{E}(\Delta M_{n+1} \Delta M_{n+1}^\top \mid \mathcal{F}_n) &\xrightarrow[n \rightarrow +\infty]{\text{p.s}} \mathbf{1}_{\{|\theta_n - \theta^*| \leq C\}} \Gamma \in \mathcal{S}_d(\mathbb{R}) \end{aligned}$$

Alors, si pour tout $n \geq 1$, $\gamma_n := \alpha n^{-1}$, où $\alpha \in]1/2, 1[$, la suite des moyennes empiriques $(\bar{\theta}_n)_{n \geq 0}$ définie ci-haut, vérifie sur l'ensemble (évènement) de convergence $\{\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s}} \theta^*\}$, le TCL avec la variance minimale suivant :

$$\sqrt{n} (\bar{\theta}_n - \theta^*) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}_d(0, (Jh(\theta^*))^{-1} \Gamma Jh(\theta^*)) \quad \text{sur } \{\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s}} \theta^*\}$$

Autrement dit,

$$\forall g \in \mathcal{C}_b^0(\mathbb{R}^d, \mathbb{R}), A \in \mathcal{F} : \quad \mathbb{E} \left(\mathbf{1}_{A \cap \{\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s}} \theta^*\}} g(\sqrt{n} (\bar{\theta}_n - \theta^*)) \right) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E} \left(\mathbf{1}_{A \cap \{\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s}} \theta^*\}} g(Z) \right)$$

où :

$$Z \rightsquigarrow \mathcal{N}_d(0, (Jh(\theta^*))^{-1} \Gamma Jh(\theta^*))$$

Exécution optimale à l'aide de l'approximation stochastique

5.1 Introduction

Dans ce chapitre, on présente un cadre d'optimisation pour le *Market-Making* dans le carnet d'ordre limite comme celui présenté par J. Fernandez-Tapia dans [46]. On considère une version en temps discret du modèle d'Avellaneda et Stoikov [9] présenté dans la section 3.2.1 similaire à celle développée dans le travail de Laruelle, Lehalle et Pagès dans [89] dans le contexte des tactiques de la liquidation optimale. L'idée est de profiter de la nature itérative de la procédure d'actualisation des prix postés afin de concevoir un algorithme qui optimise sa stratégie d'une manière "on-line" de *trial-and-error* (i.e, apprentissage on-line). Un des avantages de cette approche est le fait que l'exploration du système est effectuée en temps-réel (cf. 1.3.1). De ce fait, la spécification d'un modèle pour le prix n'est pas nécessaire contrairement au cadre du contrôle stochastique présenté dans le chapitre 3.

On considère dans ce chapitre un *market maker* qui interagit avec un carnet d'ordre limite électronique. Rappelons que le rôle d'un *market maker* est de fournir la liquidité au marché en postant des cotations au côté *bid* et *ask*. Il crée son profit en achetant une quantité et la revendant à un prix supérieur au prix d'achat. Plus concrètement, son profit est dû à la différence des deux derniers prix appelée *spread* du *market maker*. Ainsi, son objectif est la maximisation du nombre de ce type de transactions (i.e, les paires (achat, revente)) qui ont, de plus, des *spread* larges, tout en gardant une inventaire proche de 0 à la fin de la session du trading. De ce fait, le *market maker* fait face aux arbitrages suivants : Premièrement, s'il espère avoir un large *spread*, la probabilité d'exécution de son LO diminue. De plus, un *spread* relativement étroit n'engendre que peu de profits. Ceci d'une part, d'autre part, si son algorithme poste les prix que d'un seul côté (*bid/ask*) (e.g. à cause des mouvements du prix), son inventaire s'éloigne de zéro ce qui lui force à exécuter cet excès d'inventaire à l'aide d'un MO à la fin de la session du trading. Ceci motive son algorithme à centrer les prix postés autour du prix de référence (e.g le *mid-price*) afin de garder son inventaire proche de 0.

REMARQUE : Remarquer que le problème de la liquidation/acquisition optimale à l'aide des LOs n'est, *mutatis mutandis*, qu'une sorte de *market-making* unilatéral. De ce fait, le traitement

des trois problèmes est similaire d'un point de vue théorique.

D'un point de vue de modélisation, une itération d'une tactique de *market-making* peut être vue comme une interaction avec un *black-box* à lequel on donne, comme *input*, les contrôles δ_a et δ_b qui représentent les distances entre les prix postés dans les côté *ask* et *bid* et le prix de référence. Ainsi, le *black-box* donne, comme *output*, deux variables N_a et N_b représentant les flux (les liquidités) capturés dans les côtés du *spread* durant une fenêtre de longueur $\Delta t > 0$ (représentant la durée d'une itération de l'algorithme) et qui dépendent des contrôles ainsi que d'autres variables exogènes.

À la fin de chaque itération, le payoff obtenu par l'algorithme est une variable aléatoire :

$$\Xi(\delta_a, \delta_b) := \Pi(N_a(\delta_a, \xi), N_b(\delta_b, \xi), \xi)$$

tel que ξ représente les variables aléatoires qui influencent le payoff (e.g. prix, *spread*, etc).

Dans la pratique, ξ peut être un vecteur aléatoire ou un processus stochastique et il est modélisé de manière à représenter des "innovations" qui peuvent être i.i.d, stationnaires et ergodiques ou faiblement dépendantes. Comme le lecteur va remarquer ci-dessous, ξ sera prix comme le rendement du prix de référence de la période d'intérêt.

Notre objectif est de maximiser le payoff espéré $\mathbb{E}(\Xi(\delta_a, \delta_b))$ sur l'ensemble des valeurs prises par le couple (δ_a, δ_b) . La question qui se pose à ce stade là est de savoir comment exploiter l'aspect itératif de l'algorithme afin de trouver le maximisateur du payoff espéré, tout en proposant un cadre adaptatif et flexible qui ne requiert pas une modélisation détaillée de la dynamique du marché (e.g. prix, liquidité, etc).

Comme le lecteur a pu (probablement) déjà remarquer, notre modélisation rejoint naturellement le cadre de l'approximation stochastique présenté dans le chapitre précédent, dans lequel on cherche à concevoir une procédure itérative qui converge vers les points critiques/zéros d'une fonction h admettant une représentation intégrale de la forme, $h(\delta) := \mathbb{E}(H(\delta, \xi))$ pour $\delta \in \mathbb{R}^d$, dans le cas où on ne peut pas calculer h explicitement mais, une expression du gradient de H par rapport à δ est disponible et ξ peut être simulée à un coût raisonnable ou peut être obtenue en temps réel à l'aide d'un flux de données (dans ce dernier cas, la procédure est connue sous le nom de *on-line learning*).

Comme nous avons déjà mentionné, l'avantage de cette approche est sa flexibilité et son aptitude à approcher des situations dans lesquelles le prix de référence n'est pas modélisé nécessairement par une diffusion brownienne comme dans l'approche du contrôle stochastique. En fait, on peut traiter des situations beaucoup plus générales où l'algorithme collecte continûment les informations de son environnement sans devoir spécifier un modèle décrivant la dynamique du prix. De plus, la nature récursive de la procédure la rend adaptative, facilement implémentable et peut être généralisée pour traiter d'autres cadres comme le trading dans les *dark-pools*. Voir e.g les travaux de Laruelle et al. [89], [88].

5.2 Formulation du problème

5.2.1 Problème d'optimisation

On fixe un horizon de temps $T > 0$ et on se place dans le cadre d'un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ dont la filtration vérifie les conditions habituelles. On divise l'intervalle de temps $[0, T]$ en $N \in \mathbb{N}^*$ périodes de trading de longueur $\Delta T > 0$ secondes, i.e :

$$[0, T] = \bigcup_{n=1}^N [(n-1)\Delta T, n\Delta T]$$

On s'intéresse au problème d'un algorithme de *market making* qui actualise ses prix postés dans les deux côtés du carnet d'ordre limite (i.e. *bid/ask*) au début de chaque période. Au début de la période $n \in \llbracket 1, N \rrbracket$, i.e dans l'intervalle $[(n-1)\Delta T, n\Delta T]$, l'algorithme poste ses cotations à des distances $\delta_a^{(n-1)}$ (pour le côté *ask*) et $\delta_b^{(n-1)}$ (pour le côté *bid*) d'un prix de référence $S_{(n-1)\Delta T}$ (e.g. le *mid-price* à l'instant $(n-1)\Delta T$). Les ordres restent inchangés jusqu'à la fin de la période.

Pour $n \in \llbracket 1, N \rrbracket$, soit $N_a^{(n)}$ (resp. $N_b^{(n)}$) le nombre des ordres postés par l'algorithme et qui sont exécutés durant la $n^{\text{ème}}$ période dans le côté *ask* (resp. *bid*). Le payoff réalisé par le *market maker* durant cette période est donné par :

$$\Xi_n := \underbrace{N_a^{(n)} (S_{(n-1)\Delta T} + \delta_a^{(n-1)})}_{\text{Vente}} - \underbrace{N_b^{(n)} (S_{(n-1)\Delta T} - \delta_b^{(n-1)})}_{\text{Achat}} + \underbrace{(N_b^{(n)} - N_a^{(n)}) S_{n\Delta T}}_{\text{Revalorisation de l'inventaire}}$$

On introduit le rendement du prix de référence en tout instant de la $n^{\text{ème}}$ période qui est défini par :

$$\forall t \in [0, \Delta T] : Y_t^{(n)} := S_{(n-1)\Delta T + t} - S_{(n-1)\Delta T}$$

On introduit également les variables suivantes :

$$\begin{aligned} N^{(n)} &:= N_a^{(n)} + N_b^{(n)} \quad ; \quad Q^{(n)} = N_b^{(n)} - N_a^{(n)} \\ \psi^{(n)} &:= \frac{\delta_a^{(n)} + \delta_b^{(n)}}{2} \quad ; \quad \theta^{(n)} := \frac{\delta_a^{(n)} - \delta_b^{(n)}}{2} \end{aligned}$$

$N^{(n)}$ et $Q^{(n)}$ représentent respectivement le nombre de transactions exécutées par le *market maker* et l'accroissement de son inventaire durant la période $[(n-1)\Delta T, n\Delta T]$. $\psi^{(n)} \in \mathbb{R}_+$ est le *demi-spread* du *marker maker* et $\theta^{(n)} \in \mathbb{R}$ est le facteur *off-center*.

Après ce changement de variables, le payoff réalisée durant la $n^{\text{ème}}$ période se réécrit de la façon suivante :

$$\Xi_n = \psi^{(n-1)} N^{(n)} - (\theta^{(n-1)} - Y_{\Delta T}^{(n)}) Q^{(n)}$$

Ainsi, le but du *marker maker* est de maximiser l'espérance de cette quantité au début de chaque période.

5.2.2 Problème à une période

Notons $\mathbb{D}([0, T], \mathbb{R})$ l'espace des fonctions càdlàg définie de $[0, T]$ à valeurs dans \mathbb{R} muni de la topologie de la convergence uniforme qui est plus grossière que la topologie J_1 de Skorokhod usuelle (voir [72]). Soit $H : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{D}([0, \Delta T], \mathbb{R}) \rightarrow \mathbb{R}$ une fonctionnelle borélienne définie de la manière suivante :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R}, \forall n \in \llbracket 1, N \rrbracket : H(\psi, \theta, (Y_t^{(n)})_{t \in [0, \Delta T]}) := \mathbb{E}(\Xi_n^{\psi, \theta} | \mathcal{F}_{n\Delta T}^Y) \quad (5.2.1)$$

où $\mathcal{F}_{n\Delta T}^Y$ est la tribu engendrée par $Y^{(n)}$, i.e :

$$\mathcal{F}_{n\Delta T}^Y := \sigma \left(Y_t^{(n)}, t \in [0, \Delta T] \right)$$

et :

$$\Xi_n^{\psi, \theta} := \psi N^{(n)} - \left(\theta - Y_{\Delta T}^{(n)} \right) Q^{(n)}$$

Remarquer que si, pour tout $n \in \llbracket 1, N \rrbracket$, $Q^{(n)}$ et $N^{(n)}$ sont indépendantes de $Y^{(n)}$, on a :

$$H \left(\psi, \theta, \left(Y_t^{(n)} \right)_{t \in [0, \Delta T]} \right) = \psi \mathbb{E} \left(N^{(n)} \right) - \left(\theta - Y_{\Delta T}^{(n)} \right) \mathbb{E} \left(Q^{(n)} \right) \quad (5.2.2)$$

On suppose que le processus du prix de référence est à accroissements indépendants et stationnaires ce qui implique que :

$$\mathbb{E} \left(H \left(\psi, \theta, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) \right) = \dots = \mathbb{E} \left(H \left(\psi, \theta, \left(Y_t^{(N)} \right)_{t \in [0, \Delta T]} \right) \right)$$

De ce fait, le problème à plusieurs période se réduit à un problème à une période défini de la manière suivante :

$$\sup_{(\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R}} \pi(\psi, \theta) \quad (5.2.3)$$

tel que $\pi : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ est la fonction du payoff espéré, définie par :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : \quad \pi(\psi, \theta) := \mathbb{E} \left[H \left(\psi, \theta, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) \right] \quad (5.2.4)$$

5.2.3 Relation entre le prix et la liquidité

Nous avons jusqu'à présent introduit le problème d'optimisation sans relier le prix à la liquidité. Pour ce faire, on considère la version modifiée du modèle d'Avellaneda et Stoikov que nous avons présenté dans la section 3.2.1 et qui est similaire à celle introduite par Laruelle et al. dans [88].

On rappelle que dans ce modèle, la probabilité qu'un LO posté à une distance δ du prix de référence soit exécuté dans l'intervalle $[t, t + \delta t]$, pour $t \in [0, \Delta T]$ avec $\delta t > 0$, est donnée par :

$$Ae^{-k\delta\delta t}$$

Ainsi, les flux capturés par des ordres placés à des prix $S_0 + \delta_a$ et $S_0 - \delta_b$ avec $\delta_a, \delta_b > 0$ durant la période $[0, \Delta T]$ sont des variables aléatoires de poisson d'intensités :

$$\begin{aligned} \lambda^+ \left(\delta_b, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) &:= \int_0^{\Delta T} Ae^{-k(S_t - (S_0 - \delta_b))} dt = \int_0^{\Delta T} Ae^{-k(Y_t^{(1)} + \delta_b)} dt \\ \lambda^- \left(\delta_a, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) &:= \int_0^{\Delta T} Ae^{-k((S_0 + \delta_a) - S_t)} dt = \int_0^{\Delta T} Ae^{-k(\delta_a - Y_t^{(1)})} dt \end{aligned}$$

où λ^+ (resp. λ^-) concerne les transactions d'achat (resp. de vente). Pour proposer une modélisation cohérente avec le problème 5.2.4, on introduit le changement de variables suivant :

$$\begin{aligned} \lambda \left(\psi, \theta, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) &:= \lambda^- \left(\delta_b, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) + \lambda^+ \left(\delta_b, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) \\ \mu \left(\psi, \theta, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) &:= \lambda^- \left(\delta_b, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) - \lambda^+ \left(\delta_b, \left(Y_t^{(1)} \right)_{t \in [0, \Delta T]} \right) \end{aligned}$$

D'où :

$$\begin{aligned}\lambda\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right) &= 2Ae^{-k\psi} \int_0^{\Delta T} \cosh\left(k\left(\theta - Y_t^{(1)}\right)\right) dt \\ \mu\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right) &= 2Ae^{-k\psi} \int_0^{\Delta T} \sinh\left(k\left(\theta - Y_t^{(1)}\right)\right) dt\end{aligned}$$

De ce fait, on a :

$$\begin{aligned}\mathbb{E}\left(N^{(1)} \mid \mathcal{F}_{\Delta T}^Y\right) &= \lambda\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right) \\ \mathbb{E}\left(Q^{(1)} \mid \mathcal{F}_{\Delta T}^Y\right) &= \mu\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right)\end{aligned}$$

Finalement, la fonctionnelle du payoff conditionnel au trajectoire $Y^{(1)}$ définie dans 5.2.1 se réécrit de la manière suivante :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : H\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right) = \psi \lambda\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right) - (\theta - Y_{\Delta T}^{(1)}) \mu\left(\psi, \theta, \left(Y_t^{(1)}\right)_{t \in [0, \Delta T]}\right)$$

Pour alléger la notation, on omet de noter l'indice de la période dans $Y^{(1)}$, i.e on écrit, dorénavant, Y au lieu de $Y^{(1)}$.

5.3 Analyse du problème

On s'intéresse dans cette section à l'analyse du problème d'optimisation à une période introduit 5.2.3. On espère concevoir une procédure récursive "à la Robbins-Monro" qui permet de résoudre le problème. On démontre l'existence et l'unicité du maximum et on se base sur la théorie de l'approximation stochastique développée dans le chapitre 4 afin de démontrer la convergence de la procédure.

Remarquons, tout d'abord, que le fonction du payoff espéré 5.2.4 se réécrit de la manière suivante :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : \pi(\psi, \theta) = 2Ae^{-k\psi} (\psi h(\theta) - g(\theta)) \quad (5.3.1)$$

où, $h, g : \mathbb{R} \rightarrow \mathbb{R}$ sont les fonctions définies, pour tout $\theta \in \mathbb{R}$, par :

$$h(\theta) = \mathbb{E} \left(\int_0^{\Delta T} \cosh(k(\theta - Y_t)) dt \right) \quad (5.3.2)$$

$$g(\theta) = \mathbb{E} \left((\theta - Y_{\Delta T}) \int_0^{\Delta T} \sinh(k(\theta - Y_t)) dt \right) \quad (5.3.3)$$

Rappelons qu'une fonction càdlàg sur $[0, \Delta T]$ est bornée (voir e.g. [26], chapitre 3). De plus, on a $\cosh' = \sinh$, $\sinh' = \cosh$, $\cosh \leq \exp \circ |\cdot|$ et pour tout $\theta \in \mathbb{R}$, les fonctions $(t, \omega) \mapsto \sinh\left(k\left(\theta - \inf_{s \in [0, \Delta T]} Y_s(\omega)\right)\right)$ et $(t, \omega) \mapsto \sup_{s \in [0, \Delta T]} |Y_s(\omega)| \exp\left(k\left(|\theta| + \sup_{s \in [0, \Delta T]} |Y_s(\omega)|\right)\right)$ sont dans $L^1([0, \Delta T] \times \Omega)$. Donc, d'après le théorème de Lebesgue de la dérivation sous le signe intégrale (voir e.g. [44], chapitre 8), les fonctions h et g sont dérivables sur \mathbb{R} . Il s'ensuit que π est différentiable sur $\mathbb{R}_+^* \times \mathbb{R}$ et on a bien le droit de calculer le gradient de π dans tout ce qui suit.

Le théorème suivant affirme l'existence et l'unicité d'un maximisateur de π ce qui implique, en particulier, que $\{\nabla \pi = 0\} = \{(\psi^*, \theta^*)\}$.

Théorème 5.3.1: Existence et unicité du maximum

La fonction π définie par 5.3.1 admet un unique maximisateur sur $\mathbb{R}_+ \times \mathbb{R}$.

Démonstration. Voir Annexe 5.5. □

On espère concevoir un algorithme de Robbins-Monro pour chercher les zéros de $\nabla \pi$. Pour ce faire, on calcule, tout d'abord, le gradient local de π , i.e on calcule le gradient de $(\psi, \theta) \mapsto H(\psi, \theta, Y)$ pour toute fonction $t \mapsto Y_t$ de $\mathbb{D}([0, \Delta T], \mathbb{R})$.

Proposition 5.3.1

Soit $Y : t \mapsto Y_t \in \mathbb{D}([0, \Delta T], \mathbb{R})$, alors $(\psi, \theta) \mapsto H(\psi, \theta, Y) \in \mathcal{C}^1(\mathbb{R}_+ \times \mathbb{R})$ et son gradient est donné par :

$$\nabla_{(\psi, \theta)} H(\psi, \theta, Y) = \begin{pmatrix} \partial_\psi H(\psi, \theta, Y) \\ \partial_\theta H(\psi, \theta, Y) \end{pmatrix} = \begin{pmatrix} 1 - k\psi & k(\theta - Y_{\Delta T}) \\ -k(\theta - Y_{\Delta T}) & -(1 - k\psi) \end{pmatrix} \begin{pmatrix} \lambda(\psi, \theta, Y) \\ \mu(\psi, \theta, Y) \end{pmatrix}$$

Démonstration. Voir Annexe 5.5. □

Comme on verra dans les démonstrations, la fonction $\nabla_{(\psi, \theta)} H$ ne vérifie pas la condition de croissance pseudo-linéaire de l'algorithme de

$$\mathbb{E} \left(\left| \nabla_{(\psi, \theta)} H(\psi, \theta, (Y_t)_{t \in [0, \Delta T]}) \right|^2 \right) \leq K e^{2k(|\theta| - \psi)} (1 + \psi^2 + \theta^2)$$

pour une constante positive K .

Ainsi, pour garantir la convergence de la procédure itérative de la recherche du zéro de $\nabla \pi$, on doit contrôler les valeurs de H pour les grandes valeurs θ et on est obligé à modifier la procédure récursive usuelle, i.e, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} \psi_{n+1} &= \psi_n - \gamma_{n+1} \partial_\psi H(\psi_n, \theta_n, Y_{n+1}) \\ \theta_{n+1} &= \theta_n - \gamma_{n+1} \partial_\theta H(\psi_n, \theta_n, Y_{n+1}) \end{aligned}$$

Par la procédure récursive suivante :

$$\psi_{n+1} = \psi_n - \gamma_{n+1} \rho(\psi_n, \theta_n) \partial_\psi H(\psi_n, \theta_n, Y_{n+1}) \quad (5.3.4)$$

$$\theta_{n+1} = \theta_n - \gamma_{n+1} \rho(\psi_n, \theta_n) \partial_\theta H(\psi_n, \theta_n, Y_{n+1}) \quad (5.3.5)$$

où $\rho : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}_+^*$ est la fonction définie par :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : \quad \rho(\psi, \theta) := e^{-k(|\theta| - \psi)}$$

Pagès et al. ont démontré dans [95] que la convergence de la procédure n'est pas modifiée et la deuxième procédure modifiée converge vers le zéro de la fonction d'intérêt pour une fonction ρ strictement positive (ce qui est le cas pour la fonction ρ définie ci-haut).

Avant de présenter le résultat de convergence de cette dernière procédure récursive, on présente tout d'abord quelques contraintes que les paramètres doivent vérifier dans les applications numériques. En effet, le fait d'avoir des cotations de type *bid* dans le côté *ask* (et vice versa) n'a pas de sens dans la pratique. Sur ce, on impose la contrainte suivante :

$$|\theta| < C + \psi$$

tel que $C > 0$ est une borne supérieure du demi *bid-ask spread*. De plus, on veut que l'algorithme cherche la solution dans la direction où le payoff soit positif. Ainsi, on espère que la solution appartient à l'ensemble suivant :

$$\Pi^+ := \{(\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : |\theta| < C + \psi, \pi(\psi, \theta) \geq 0\}$$

Sur ce, on va considérer un algorithme stochastique avec projection de la forme :

$$\forall n \in \mathbb{N} : X_{n+1} = \text{Proj}_{\Pi^+} (X_n - \gamma_{n+1} H(X_n, Y_{n+1}))$$

où, Proj_{Π^+} est la projection euclidienne sur Π^+ . On cite par la suite, un théorème de convergence pour les algorithmes stochastiques avec projection.

5.3.1 Résultat de convergence pour les algorithmes contraints

On se place dans le cadre de 4.2.1 du chapitre précédent 4. Le but est de déterminer les éléments de l'ensemble $\{\theta \in \Theta : h(\theta) = \mathbb{E}(H(\theta, Y)) = 0\}$ tel que $\Theta \subset \mathbb{R}^d$ est un ensemble convexe fermé. Soit $(\theta_n)_{n \in \mathbb{N}}$ la suite de variables aléatoires à valeurs dans \mathbb{R}^d définie sur un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ par :

$$\forall n \in \mathbb{N} : \theta_{n+1} = \text{Proj}_{\Theta} (\theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})) \quad ; \quad \theta_0 \in \Theta \quad (5.3.6)$$

où $(Y_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires à valeurs dans \mathbb{R}^q , i.i.d et de même loi que Y . $(\gamma_n)_{n \geq 1}$ est une suite de pas à valeurs dans \mathbb{R}_+^* et Proj_{Θ} est la projection euclidienne sur Θ . La procédure récursive 5.3.6 peut se réécrire sous la forme :

$$\forall n \in \mathbb{N} : \theta_{n+1} = \theta_n - \gamma_{n+1} h(\theta_n) - \gamma_{n+1} \Delta M_{n+1} + \gamma_{n+1} p_{n+1} \quad (5.3.7)$$

où, pour $n \in \mathbb{N}$, $\Delta M_{n+1} := H(\theta_n, Y_{n+1}) - h(\theta_n)$ est un \mathcal{F}_n -accroissement de martingale et $(p_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires à valeurs dans \mathbb{R}^d définie par :

$$\forall n \in \mathbb{N}^* : p_{n+1} := \frac{1}{\gamma_{n+1}} (\text{Proj}_{\Theta} (\theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})) - \theta_n) + H(\theta_n, Y_{n+1})$$

Théorème 5.3.1.1: Convergence p.s des algorithmes contraints (c.f [82], [66])

Soit $(\theta_n)_{n \in \mathbb{N}}$ la suite définie par 5.3.7. Supposons que $\{h = 0\} \cap \Theta = \{\theta^*\}$ et que la fonction moyenne h vérifie l'hypothèse de *mean-reversion* suivante :

$$\forall \theta \in \Theta \setminus \{\theta^*\} : \langle h(\theta), \theta - \theta^* \rangle > 0$$

Supposons que la suite de pas vérifie :

$$\sum_{n=1}^{+\infty} \gamma_n = +\infty \quad \text{et} \quad \sum_{n=1}^{+\infty} \gamma_n^2 < +\infty$$

De plus, s'il existe $K > 0$ tel que la fonction H vérifie :

$$\forall \theta \in \Theta : \mathbb{E}(|H(\theta, Y)|^2) \leq K(1 + |\theta|^2)$$

alors :

$$\theta_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \theta^*$$

5.3.2 Convergence de la procédure récursive

Comme elle a été déjà mentionnée, la fonction H ne vérifie pas une condition du théorème de convergence de l'algorithme de Robbins-Monro ce qui nous a obligé à modifier la procédure récursive usuelle par celle présentée dans (5.3.8, 5.3.9). Heureusement, les auteurs dans [95] ont démontré que cette modification ne change pas la convergence de l'algorithme pour une classe large de fonctions, à savoir : les fonctions (presque partout) strictement positives. De ce fait, il suffit de vérifier les conditions du théorème de convergence de l'algorithme de Robbins-Monro présenté dans 4.3.2.1. Nous nous plaçons dans le cadre d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Théorème 5.3.2.1: Convergence p.s

Soit $\rho : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}_+^*$ la fonction définie par :

$$\forall (\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R} : \quad \rho(\psi, \theta) := e^{-k(|\theta| - \psi)}$$

et $\left((Y_t^{(n)})_{t \in [0, \Delta T]} \right)_{n \geq 1}$ une suite de trajectoires i.i.d de même loi que $(Y_t)_{t \in [0, \Delta T]}$. On note $\{(\psi^*, \theta^*)\} := \{\nabla \pi = 0\} = \arg \max_{(\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R}} \pi(\psi, \theta)$ et on définit la procédure récursive suivante :

$$\psi_{n+1} = \psi_n - \gamma_{n+1} \rho(\psi_n, \theta_n) \partial_\psi H(\psi_n, \theta_n, Y_{n+1}) \quad (5.3.8)$$

$$\theta_{n+1} = \theta_n - \gamma_{n+1} \rho(\psi_n, \theta_n) \partial_\theta H(\psi_n, \theta_n, Y_{n+1}) \quad (5.3.9)$$

Supposons que :

$$\begin{aligned} \mathbb{E} \left(Y_{\Delta T}^2 \int_0^{\Delta T} \exp(2k|Y_t|) dt \right) &< +\infty \\ \mathbb{E} \left(\int_0^{\Delta T} \exp(2k|Y_t|) dt \right) &< +\infty \end{aligned}$$

et que :

$$\psi_0 \in L^1(\Omega, \mathbb{P}) \quad ; \quad \theta_0 \in L^1(\Omega, \mathbb{P})$$

Soit $(\gamma_n)_{n \geq 1}$ une suite de pas à valeurs dans \mathbb{R}_+^* vérifiant :

$$\sum_{n=1}^{+\infty} \gamma_n = +\infty \quad \text{et} \quad \sum_{n=1}^{+\infty} \gamma_n^2 < +\infty$$

Alors :

$$(\psi_n, \theta_n) \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} (\psi^*, \theta^*)$$

Démonstration. Voir Annexe 5.5. □

Un problème délicat qui n'a pas été traité auparavant est le fait que les innovations dans la procédure précédente sont infinie-dimensionnelles, i.e des processus stochastiques (processus du rendement du prix de référence), tandis que dans le théorème de Robbins-Monro 4.3.2.1, nous avons démontré la convergence de la procédure pour des innovations finie-dimensionnelles. Comme l'affirme le théorème suivant, après un changement de variable, la procédure récursive définie dans 5.3.2.1 ne dépend de la trajectoire de Y qu'à travers trois variables aléatoires ce qui nous ramène au cas usuel des innovations fini-dimensionnelles.

Théorème 5.3.2.2: Problème des innovations

La fonctionnelle H définie dans 5.2.1 ne dépend de $(Y_t)_{t \in [0, \Delta T]} \in \mathbb{D}([0, \Delta T], \mathbb{R})$ qu'à travers la variable $Y_{\Delta T}$ et

$$\alpha_k := \frac{1}{2} \ln \left(\int_0^{\Delta T} e^{kY_t} dt \int_0^{\Delta T} e^{-kY_t} dt \right)$$

$$\beta_k := \frac{1}{2} \ln \left(\frac{\int_0^{\Delta T} e^{-kY_t} dt}{\int_0^{\Delta T} e^{kY_t} dt} \right)$$

De plus, on a les formules suivantes (par un abus des notations), pour $(\psi, \theta) \in \mathbb{R}_+ \times \mathbb{R}$:

$$\lambda(\psi, \theta) = 2Ae^{-k\psi + \alpha_k} \cosh(k\theta + \beta_k)$$

$$\mu(\psi, \theta) = 2Ae^{-k\psi + \alpha_k} \sinh(k\theta + \beta_k)$$

Démonstration. Voir Annexe 5.5. □

5.4 Applications numériques

On présente dans cette section une application numérique des résultats théoriques développés dans les sections précédentes. Nous avons prouvé la convergence de la procédure itérative pour un processus de prix de référence càdlàg dans un cadre théorique abstrait. Pour qu'on puisse faire des illustrations numériques, on suppose que le dernier processus se comporte comme une diffusion brownienne, i.e :

$$dS_t = \sigma dW_t \quad ; \quad S_0 > 0$$

où W est un mouvement brownien standard et $\sigma > 0$ est la volatilité du prix.

Dans ce cadre, le processus du rendement du prix sur la période $[0, \Delta T]$ est donné par :

$$\forall t \in [0, \Delta T] : \quad Y_t = \sigma W_t$$

Cette étude ne sert pas pour des raisons d'illustrations seulement, mais elle permet, également, de comprendre les différentes relations entre les paramètres et de trouver des valeurs initiales des algorithmes stochastiques lorsqu'ils sont utilisés dans une procédure d'apprentissage *on-line*. Un autre avantage de ce choix particulier est l'existence d'une formule explicite pour la solution optimale du problème. En effet, pour $\theta \in \mathbb{R}$:

$$h(\theta) = \mathbb{E} \left(\int_0^{\Delta T} \cosh(k(\theta - Y_t)) dt \right)$$

$$= \frac{1}{2} \left(e^{k\theta} I_{\Delta T}(-k\sigma) + e^{-k\theta} I_{\Delta T}(k\sigma) \right)$$

où,

$$I_{\Delta T}(\pm k\sigma) := \int_0^{\Delta T} \mathbb{E} \left(e^{\pm k\sigma W_t} \right) dt$$

Sachant, que $-W$ est également un mouvement brownien, on a : $I_{\Delta T}(-k\sigma) = I_{\Delta T}(k\sigma)$. Cette dernière intégrale a été déjà calculée dans la dernière section du chapitre précédent 3.2.10 et est donnée par :

$$I_{\Delta T}(k\sigma) = \frac{2}{k^2 \sigma^2} \left(e^{\frac{k^2 \sigma^2 \Delta T}{2}} - 1 \right)$$

$\theta \mapsto g(\theta)$ peut être calculée d'une manière similaire pour obtenir, pour tout $\theta \in \mathbb{R}$:

$$\begin{aligned}
h(\theta) &= \frac{2 \cosh(k\theta)}{k^2 \sigma^2} \left(e^{\frac{k^2 \sigma^2 \Delta T}{2}} - 1 \right) \\
g(\theta) &= \frac{2\theta \sinh(k\theta)}{k^2 \sigma^2} \left(e^{\frac{k^2 \sigma^2 \Delta T}{2}} - 1 \right) \\
&\quad + \frac{2 \cosh(k\theta)}{k^3 \sigma^2} \left(e^{\frac{k^2 \sigma^2 \Delta T}{2}} (k^2 \sigma^2 \Delta T - 2) + 2 \right)
\end{aligned}$$

En rappelant l'expression du payoff espéré 5.3.1, de simple calculs permettent de conclure que le maximum de la fonction est obtenu pour :

$$\psi^* = \frac{1}{k} \left(\frac{k^2 \sigma^2 \Delta T - 1 + e^{-\frac{k^2 \sigma^2 \Delta T}{2}}}{1 - e^{-\frac{k^2 \sigma^2 \Delta T}{2}}} \right) ; \quad \theta^* = 0$$

Dans la pratique, le processus Y est simulé par un schéma d'Euler-Maruyama de pas $\delta t := \Delta T/m$:

$$\forall j \in \llbracket 0, m-1 \rrbracket : \quad \bar{Y}_{(j+1)\delta t} = \bar{Y}_{j\delta t} + \sqrt{\frac{\Delta T}{m}} Z_{j+1}$$

où $(Z_j)_{j \in \llbracket 0, m \rrbracket}$ sont i.i.d de loi $\mathcal{N}(0, 1)$.

De ce fait, les variables α_k et β_k sont simulées par :

$$\bar{\alpha}_m := \frac{X_m^- + X_m^+}{2} ; \quad \bar{\beta}_m := \frac{X_m^- - X_m^+}{2}$$

où :

$$X_m^\pm := \ln \left(\frac{\Delta T}{m} \sum_{j=0}^m e^{\pm k Y_{j\delta t}} \right)$$

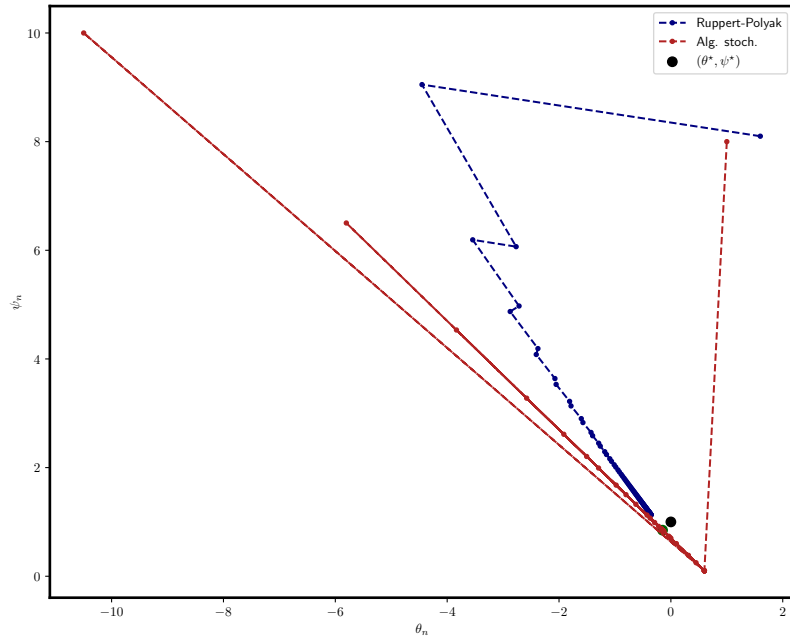


FIGURE 5.1 – Exemple d'une trajectoire de l'algorithme stochastique qui converge vers la solution

La figure 5.1 montre la convergence de l'algorithme stochastique ainsi que la version de moyennisation de Ruppert-Polyak pour un nombre d'itération égale à 100. Le cerle en noir représente la vraie solution théorique obtenue auparavant. Le cadre numérique de la simulation est le suivant :

$$\Delta T = 5s. ; m = 20 ; \sigma = 10^{-2} ; A = 5 ; k = 1$$

La suite de pas utilisée est définie par :

$$\forall n \in \mathbb{N}^* : \gamma_n := \frac{1}{2n}$$

Les valeurs initiales sont : $(\psi_0, \theta_0) := (8, 1)$. On constate que la procédure de Ruppert-Polyak est moins instable que la procédure "Batch" et stagne dans la dernière étape de la convergence.

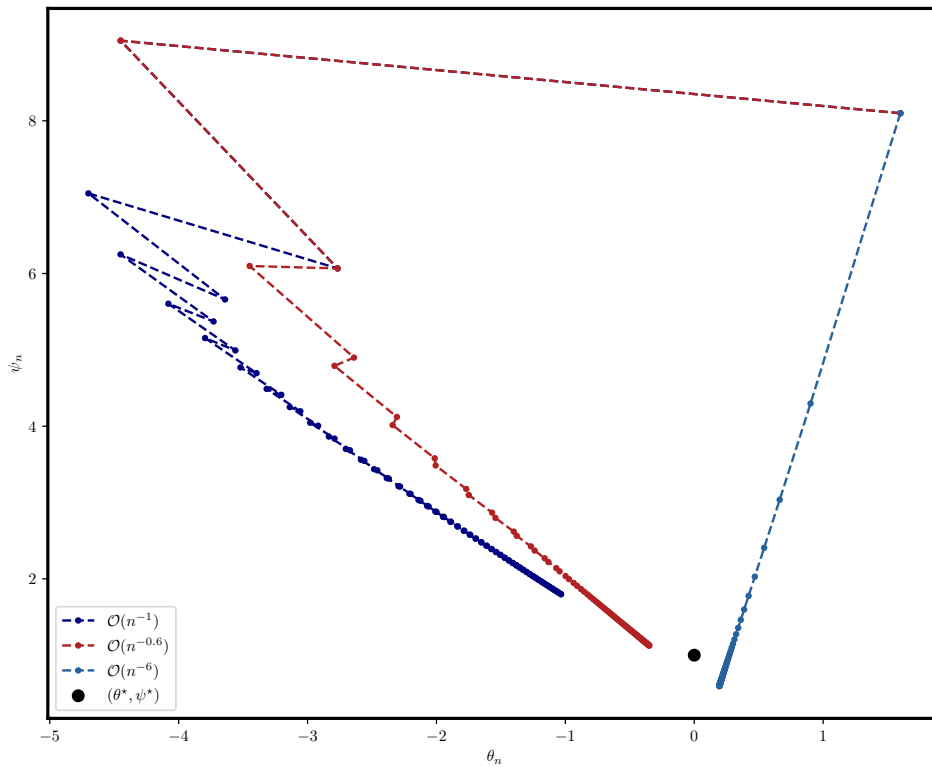


FIGURE 5.2 – Procédure de Ruppert-Polyak pour différentes suites de pas

La figure 5.2 montre la convergence de la procédure de Ruppert-Polyak pour différentes suites de pas. On remarque que, conformément aux résultats théoriques, la procédure converge rapidement pour les pas ayant un exposant dans $]1/2, 1[$.

Pour un traitement d'un cadre où on introduit une pénalisation de l'inventaire dans l'expression de la fonction objective, le lecteur est renvoyé au travail de Fernandez-Tapia [46] et Laruelle et al. [88]. L'auteur de [47] a proposé un cadre théorique abstrait s'appuyant sur la théorie de l'approximation stochastique pour traiter ce type de problèmes.

Plusieurs extensions peuvent être traitées dans le cadre de ce paradigme de trading optimal s'appuyant sur l'approximation stochastique. À titre d'exemple, on peut modéliser le processus du prix de référence par une différence de deux processus de Hawkes ou modéliser le processus du log-rendement par un mouvement brownien fractionnaire.

5.5 Annexe : Démonstrations des résultats

Théorème 5.3.1

Avant de démontrer l'existence et l'unicité du maximum de π , on introduit tout d'abord un concept qui nous sera très utile dans la démonstration.

Co-monotonie fonctionnelle : Le principe de co-monotonie fonctionnelle introduit par Pagès dans [102] est l'extension du principe de co-monotonie classique pour les variables aléatoires à des processus stochastiques comme les diffusions browniennes, les processus à accroissements indépendants, les processus de Liouville et les mouvements browniens fractionnaires. On fixe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et on rappelle le principe de co-monotonie classique :

Proposition 5.5.1: (c.f e.g [104])

Soient X une variable aléatoire réelle et $f, g : \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions co-monotones sur \mathbb{R} (i.e elle ont la même monotonie sur \mathbb{R}) vérifiant :

$$f(X) \in L^1(\mathbb{P}), \quad g(X) \in L^1(\mathbb{P}), \quad f(X)g(X) \in L^1(\mathbb{P})$$

Alors :

$$\text{Cov}(f(X), g(X)) \geq 0$$

Afin d'étendre ce dernier principe au cas des fonctionnelles des processus stochastiques, il faut tout d'abord introduire une relation d'ordre sur l'ensemble fonctionnel $\mathbb{D}([0, \Delta T], \mathbb{R})$ et considérer les processus stochastiques comme des variables aléatoires à valeurs dans $\mathbb{D}([0, \Delta T], \mathbb{R})$, i.e on identifie le processus stochastique $(X_t)_{t \in [0, \Delta T]}$ et la variable aléatoire :

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{D}([0, \Delta T], \mathbb{R}) \\ \omega &\mapsto (t \mapsto X_t(\omega)) \end{aligned}$$

Pour que la dernière définition ait un sens, on doit munir $\mathbb{D}([0, \Delta T], \mathbb{R})$ d'une tribu rendant l'application X mesurable. On définit, à cet égard, la tribu suivante :

$$\mathcal{D}_{\Delta T} := \sigma(\alpha \mapsto \alpha(t), t \in [0, \Delta T])$$

On définit sur $\mathbb{D}([0, \Delta T], \mathbb{R})$ la relation d'ordre partielle suivante :

$$\forall \alpha, \beta \in \mathbb{D}([0, \Delta T], \mathbb{R}) : \quad \alpha \leq \beta \implies \forall t \in [0, \Delta T] : \alpha(t) \leq \beta(t)$$

On dit qu'une fonctionnelle définie sur $\mathbb{D}([0, \Delta T], \mathbb{R})$ est monotone si elle est croissante par rapport à la relation d'ordre \leq . Deux fonctionnelles sont dites co-monotones si elles partagent la même monotonie. Elles sont dites anti-comotones dans le cas contraire.

On muni l'espace $\mathbb{D}([0, \Delta T], \mathbb{R})$ par la topologie de la convergence uniforme au lieu de la topologie usuelle J_1 de Skorokhod qui est plus fine que la première (c.f [102]).

On dit que la fonctionnelle $F : \mathbb{D}([0, \Delta T], \mathbb{R}) \rightarrow \mathbb{R}$ est continue en $\alpha \in \mathcal{C}^0([0, \Delta T], \mathbb{R})$ si :

$$\forall (\alpha_n)_{n \leq 0} \in \mathbb{D}([0, \Delta T], \mathbb{R})^{\mathbb{N}} : \|\alpha_n - \alpha\|_{\infty} \xrightarrow{n \rightarrow +\infty} 0 \implies |F(\alpha_n) - F(\alpha)| \xrightarrow{n \rightarrow +\infty} 0$$

Une fonctionnelle est dit \mathcal{C}^0 -continue si elle est continue en tout point de $\mathcal{C}^0([0, \Delta T], \mathbb{R})$. Il est clair qu'une fonctionnelle \mathcal{C}^0 -continue et \mathbb{P}_X - p.s continue pour tout processus X de trajectoires p.s continues.

Une famille $(P(x, dy))_{x \in \mathbb{R}}$ est une famille de probabilités de transitions si :

- $\forall x \in \mathbb{R} : P(x, dy)$ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
- $\forall B \in \mathcal{B}(\mathbb{R}) \ x \mapsto P(x, dy)$ est une fonction borélienne sur \mathbb{R} .

On dit qu'une famille de probabilités de transitions $(P(x, dy))_{x \in \mathbb{R}}$ "conserve la monotonie" si, pour tout fonction $f : \mathbb{R} \rightarrow \mathbb{R}_+$ bornée et monotone, la fonction Pf définie par :

$$\forall x \in \mathbb{R} : Pf(x) := \int_{\mathbb{R}} f(y) P(x, dy)$$

est monotone avec la même monotonie de f .

Nous sommes maintenant en mesure de citer le principe de co-monotonie fonctionnelle :

Théorème 5.5.1: Co-monotonie fonctionnelle (c.f [102])

Soit $X := (X_t)_{t \in [0, \Delta T]}$ un processus de Markov de trajectoires p.s continues et de probabilités de transitions conservant la monotonie. Si, $F, G : \mathbb{D}([0, \Delta T], \mathbb{R}) \rightarrow \mathbb{R}$ sont deux fonctionnelles co-monotones, \mathcal{C}^0 -continues vérifiant^a :

$$F(X) \in L^1(\mathbb{P}), \quad G(X) \in L^1(\mathbb{P}), \quad F(X)G(X) \in L^1(\mathbb{P})$$

alors,

$$\text{Cov}\left(F\left((X_t)_{t \in [0, \Delta T]}\right), G\left((X_t)_{t \in [0, \Delta T]}\right)\right) \geq 0$$

a. Cette condition est vérifiée si on émet une hypothèse de croissance sous-polynômiale sur F et G

On applique le dernier théorème dans notre cadre pour tirer une inégalité qui nous sera utile dans la démonstration par la suite.

Proposition 5.5.2

Supposons que Y vérifie les conditions du théorème 5.5.1 précédent. Alors, pour tout $(\theta, \psi) \in \mathbb{R}_+ \times \mathbb{R}$:

$$\mathbb{E}\left((\theta - Y_{\Delta T}) \int_0^{\Delta T} \sinh(k(\theta - Y_t)) dt\right) \geq (\theta - \mathbb{E}(Y_{\Delta T})) \mathbb{E}\left(\int_0^{\Delta T} \sinh(k(\theta - Y_t)) dt\right)$$

Démonstration. Ceci découle facilement de l'application du principe de co-monotonie fonctionnelle 5.5.1 aux fonctionnelles F et G définies par :

$$\forall X \in \mathbb{D}([0, \Delta T], \mathbb{R}) : F(X) := \theta - X_{\Delta T}, \quad G(X) := \int_0^{\Delta T} \sinh(k(\theta - X_t)) dt$$

Le lecteur pourra facilement vérifier que ces fonctionnelles sont décroissantes, \mathcal{C}^0 -continues. L'inégalité découle finalement du fait que pour toutes variables aléatoires X, Y réelles de carré intégrables, on a :

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

□

Existence et unicité du maximum : La preuve s'appuie sur une série de propositions. Définissons à cet égard, les fonctions suivantes définies pour tout $\theta \in \mathbb{R}$ par :

$$\begin{aligned} a(\theta) &:= \mathbb{E} \left(\int_0^{\Delta T} \cosh(k(\theta - Y_t)) dt \right) \\ b(\theta) &:= \mathbb{E} \left((\theta - Y_{\Delta T}) \int_0^{\Delta T} \sinh(k(\theta - Y_t)) dt \right) \\ c(\theta) &:= \mathbb{E} \left(\int_0^{\Delta T} \sinh(k(\theta - Y_t)) dt \right) \\ d(\theta) &:= \mathbb{E} \left((\theta - Y_{\Delta T}) \int_0^{\Delta T} \cosh(k(\theta - Y_t)) dt \right) \end{aligned}$$

En utilisant les mêmes arguments de domination développés dans le début de la section 5.3, le lecteur pourra facilement remarquer que toutes ces fonctions sont au moins de classe \mathcal{C}^2 sur \mathbb{R} .

Proposition 5.5.1

1. $\forall \theta \in \mathbb{R} : a(\theta) > |c(\theta)|$

2. c est une fonction strictement croissante sur \mathbb{R} vérifiant :

$$\lim_{\theta \rightarrow -\infty} c(\theta) = -\infty \quad ; \quad \lim_{\theta \rightarrow +\infty} c(\theta) = +\infty$$

3. Il existe une constante $C > 0$ tel que :

$$\forall \theta \in]-\infty, -C[\cap]C, +\infty[: \quad b(\theta) > 0$$

4. il existe deux constantes $C_-, C_+ \in \mathbb{R}$ tel que :

$$\begin{aligned} \forall \theta \geq C_+ : \quad c(\theta)b(\theta) &\leq a(\theta)b(\theta) \\ \forall \theta \leq C_- : \quad c(\theta)b(\theta) &\geq -a(\theta)b(\theta) \end{aligned}$$

Démonstration. 1. Ceci découle facilement de l'inégalité de Jensen appliquée à la fonction convexe $x \mapsto |x|$ et de l'inégalité classique $\cosh \geq |\sinh|$.

2. Or $\cosh \geq 1$, on a pour tout $\theta \in \mathbb{R} : c'(\theta) = k \mathbb{E} \left(\int_0^{\Delta T} \cosh(k(\theta - Y_t)) dt \right) \geq k\Delta T > 0$ Les deux limites découlent facilement du fait qu'un processus càdlàg est p.s borné et :

$$\lim_{x \rightarrow -\infty} \sinh(x) = -\infty \quad ; \quad \lim_{x \rightarrow +\infty} \sinh(x) = +\infty$$

3. D'après le principe de co-monotonie fonctionnelle 5.5.2, on a :

$$\forall \theta \in \mathbb{R} : \quad b(\theta) \geq (\theta - \mathbb{E}(Y_{\Delta T})) c(\theta) \tag{5.5.1}$$

En utilisant la définition des deux limites dans 2., on a :

$$\begin{aligned} \exists A > 0, \forall \theta \geq A : \quad c(\theta) &> 0 \\ \exists B < 0, \forall \theta \leq B : \quad c(\theta) &< 0 \end{aligned}$$

On pose $C' := A \vee (\mathbb{E}(Y_{\Delta T}) + 1)$ et $C'' = A \wedge (\mathbb{E}(Y_{\Delta T}) - 1)$. Il est clair que si $\theta \geq C'$ les facteurs à droite dans l'inégalité 5.5.1 sont strictement positifs. De même, si $\theta \leq C''$ les facteurs à droite dans l'inégalité 5.5.1 sont strictement négatifs. Poser $C = C' \vee C''$ conclue la démonstration.

4. Par un argument similaire à celui utilisé dans 3., il existe $C_+ \in \mathbb{R}$ tel que pour tout $\theta \geq C_+$, $b(\theta)$ et $c(\theta)$ soient positives. De ce fait, la première inégalité découle de 1. De même, il existe $C_- \in \mathbb{R}$ tel que pour tout $\theta \leq C_-$, $b(\theta)$ soit positive et $c(\theta)$ soit négative. En utilisant une deuxième fois l'inégalité dans 1., on obtient la deuxième inégalité. \square

Proposition 5.5.2

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ la fonction définie par :

$$\forall \theta \in \mathbb{R} : f(\theta) := b(\theta)c(\theta) - a(\theta)d(\theta)$$

Alors, f est continue, strictement décroissante sur \mathbb{R} et vérifie :

$$\lim_{\theta \rightarrow -\infty} f(\theta) > 0 \quad ; \quad \lim_{\theta \rightarrow +\infty} f(\theta) < 0$$

Ceci implique, en particulier, que f est bijective sur \mathbb{R} et $0 \in f(\mathbb{R})$ ce qui nous garantit l'existence et l'unicité de la solution de $f = 0$.

Démonstration. Il est clair que f est dérivable sur \mathbb{R} (et donc continue). Or, pour tout $\theta \in \mathbb{R}$:

$$\begin{aligned} a'(\theta) &= kc(\theta) & c'(\theta) &= ka(\theta) \\ b'(\theta) &= c(\theta) + kd(\theta) & d'(\theta) &= a(\theta) + kb(\theta) \end{aligned}$$

on a :

$$\forall \theta \in \mathbb{R} : f'(\theta) = c^2(\theta) - a^2(\theta)$$

Donc f' est strictement négative sur \mathbb{R} d'après le premier résultat de la proposition précédente. D'après le dernier résultat de la dernière proposition, il existe $C_+ \in \mathbb{R}$, tel que :

$$\forall \theta \geq C_+ : f(\theta) \leq -a(\theta)(d(\theta) - b(\theta))$$

Ainsi :

$$\forall \theta \geq C_+ : f(\theta) \leq -a(\theta) \mathbb{E} \left((\theta - Y_{\Delta T}) \int_0^{\Delta T} e^{-k(\theta - Y_t)} dt \right) \leq -a(\theta)e^{-k\theta} \mathbb{E} \left((\theta - Y_{\Delta T}) \int_0^{\Delta T} e^{kY_t} dt \right)$$

En utilisant le fait qu'un processus càdlàg est p.s borné, on peut trouver une constante $A < 0$ qui dépend de C_+ tel que pour tout $\theta > C_+ \vee 0$, $f(\theta) < A$. Ce qui implique que $\lim_{\theta \rightarrow +\infty} f(\theta) < 0$ par la définition de la limite. On peut, *mutattis mutandis*, démontrer l'autre résultat en s'appuyant sur la deuxième inégalité du dernier résultat de la proposition précédente. \square

Proposition 5.5.3

Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ la fonction définie par :

$$\forall \theta \in \mathbb{R} : g(\theta) := a(\theta) \exp \left(-k \frac{b(\theta)}{a(\theta)} \right)$$

Alors, g est strictement concave et admet un unique maximum sur \mathbb{R} .

Démonstration. Il est clair que g soit dérivable sur \mathbb{R} . Alors, pour tout $\theta \in \mathbb{R}$, on a :

$$\begin{aligned} g'(\theta) &= k \exp \left(-k \frac{b(\theta)}{a(\theta)} \right) \left(a'(\theta) \left(\frac{1}{k} + \frac{b(\theta)}{a(\theta)} \right) - b'(\theta) \right) \\ &= k^2 \exp \left(-k \frac{b(\theta)}{a(\theta)} \right) \left(\frac{b(\theta)c(\theta) - a(\theta)d(\theta)}{a(\theta)} \right) \\ &= k^2 \exp \left(-k \frac{b(\theta)}{a(\theta)} \right) \frac{f(\theta)}{a(\theta)} \end{aligned}$$

Donc le signe de g' est le signe de f définie dans la proposition précédente. On sait que :

$$\exists ! \theta^* \in \mathbb{R} : f(\theta^*) = 0$$

et

$$\forall \theta \leq \theta^* : f(\theta) \geq 0 \quad ; \quad \forall \theta \geq \theta^* : f(\theta) \leq 0$$

d'où, g est strictement croissante sur $] -\infty, \theta^* [$ et strictement décroissante sur $] \theta^*, +\infty [$ et admet son unique maximum en θ^* (ce qui implique que g est strictement concave sur \mathbb{R}) \square

Revenons maintenant à notre théorème. Soit $(\psi, \theta) \in \mathbb{R}_+^* \times \mathbb{R}$ et rappelons que $\pi(\psi, \theta) \propto e^{-k\psi} (\psi a(\theta) - b(\theta))$. Par des calculs simples, si on fixe $\theta \in \mathbb{R}$, on peut démontrer que les points critiques ψ_θ^* de $\psi \mapsto \pi(\psi, \theta)$ vérifient :

$$\psi_\theta^* = \frac{1}{k} + \frac{b(\theta)}{a(\theta)}$$

De plus,

$$\pi(\psi_\theta^*, \theta) \propto k^{-1} e^{-1} g(\theta)$$

D'après la proposition précédente, g admet un unique maximum sur \mathbb{R} en un point $\theta^* \in \mathbb{R}$. Ainsi, $(\psi_{\theta^*}^*, \theta^*)$ est l'unique maximum de π sur $\mathbb{R}_+ \times \mathbb{R}$. QED.

Proposition 5.3.1

Démonstration. Par les mêmes arguments développés dans 5.3, on démontre que, pour tout $Y \in \mathbb{D}([0, \Delta T], \mathbb{R})$, la fonction $(\psi, \theta) \mapsto H(\psi, \theta, Y)$ est de classe \mathcal{C}^2 sur $\mathbb{R}_+^* \times \mathbb{R}$. De plus, des calculs directs montrent que pour tout $(\psi, \theta, Y) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{D}([0, \Delta T], \mathbb{R})$, on a :

$$\begin{aligned} \partial_\psi \lambda(\psi, \theta, Y) &= -k \lambda(\psi, \theta, Y) \\ \partial_\theta \lambda(\psi, \theta, Y) &= k \mu(\psi, \theta, Y) \\ \partial_\psi \mu(\psi, \theta, Y) &= -k \mu(\psi, \theta, Y) \\ \partial_\theta \mu(\psi, \theta, Y) &= k \lambda(\psi, \theta, Y) \end{aligned}$$

Or :

$$\begin{aligned} \partial_\psi H(\psi, \theta, Y) &= \lambda(\psi, \theta, Y) + \psi \partial_\psi \lambda(\psi, \theta, Y) - (\theta - Y_{\Delta T}) \partial_\psi \mu(\psi, \theta, Y) \\ \partial_\psi H(\psi, \theta, Y) &= \psi \partial_\theta \lambda(\psi, \theta, Y) - \mu(\psi, \theta, Y) - (\theta - Y_{\Delta T}) \partial_\theta \mu(\psi, \theta, Y) \end{aligned}$$

on a bien les formules présentées dans la proposition. \square

Théorème 5.3.2.1

Démonstration. En comparant les conditions du théorèmes avec celles de l'algorithme de Robbins-Monro, on remarque qu'il suffit de démontrer la condition de croissance pseudo-linéaire de $\nabla_{(\psi, \theta)} H$.

Soit $(\psi, \theta, Y) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{D}([0, \Delta T], \mathbb{R})$. Pour alléger l'écriture, on omet d'écrire les variables des fonctions impliquées dans la suite d'égalités suivante.

$$\begin{aligned}
 \left| \nabla_{(\psi, \theta)} H \left(\psi, \theta, (Y_t)_{t \in [0, \Delta T]} \right) \right|^2 &= |(\theta - Y_{\Delta T}) \nabla \mu - \psi \nabla \lambda|^2 + \lambda^2 + \mu^2 \\
 &\quad + 2\psi (\lambda \partial_\psi \lambda - \mu \partial_\theta \lambda) + 2(\theta - Y_{\Delta T}) (\mu \partial_\theta \mu - \lambda \partial_\psi \mu) \\
 &= |(\theta - Y_{\Delta T}) \nabla \mu - \psi \nabla \lambda|^2 + (\lambda^2 + \mu^2) (1 - 2k\psi) + 4k\lambda\mu (\theta - Y_{\Delta T}) \\
 &\leq k^2 (|\theta - Y_{\Delta T}|^2 + \psi^2) (\lambda^2 + \mu^2) + (\lambda^2 + \mu^2) (1 - 2k\psi) + 4k\lambda\mu (\theta - Y_{\Delta T}) \\
 &= (\lambda^2 + \mu^2) (k^2 |\theta - Y_{\Delta T}|^2 + (1 - k\psi)^2) + 4k\lambda\mu (\theta - Y_{\Delta T}) \\
 &\leq 2\lambda^2 (k^2 |\theta - Y_{\Delta T}|^2 + (1 - k\psi)^2 + 4k(\theta - Y_{\Delta T})) \\
 &\leq 2\lambda^2 (3k^2 |\theta - Y_{\Delta T}|^2 + (1 - k\psi)^2 + 1) \\
 &\leq 12k^2 \lambda^2 (\theta^2 + \psi^2 + k^{-2} + Y_{\Delta T}^2)
 \end{aligned}$$

Par l'inégalité de Jensen appliquée à la fonction convexe $x \mapsto x^2$, on a :

$$\lambda^2 \leq 4A^2 e^{-2k\psi} \int_0^{\Delta T} \cosh^2(k(\theta - Y_t)) dt \leq 4A^2 e^{-2k\psi} \int_0^{\Delta T} e^{2k|\theta|} e^{2k|Y_t|} dt$$

ce qui implique :

$$\left| \nabla_{(\psi, \theta)} H \left(\psi, \theta, (Y_t)_{t \in [0, \Delta T]} \right) \right|^2 \leq 48k^2 A^2 e^{2k(|\theta| - \psi)} (\theta^2 + \psi^2 + k^{-2} + Y_{\Delta T}^2) \int_0^{\Delta T} e^{2k|Y_t|} dt$$

D'après les hypothèses d'intégrabilité émises sur Y dans le théorème, le terme à droite dans la dernière inégalité est intégrable. Donc, on posant :

$$K := 48k^2 A^2 \mathbb{E} \left((1 + k^{-2} + Y_{\Delta T}^2) \int_0^{\Delta T} e^{2k|Y_t|} dt \right) < +\infty$$

on a bien :

$$\mathbb{E} \left(\left| \rho(\psi, \theta) \nabla_{(\psi, \theta)} H \left(\psi, \theta, (Y_t)_{t \in [0, \Delta T]} \right) \right|^2 \right) \leq K (1 + \psi^2 + \theta^2)$$

De ce fait, le théorème de Robbins-Monro s'applique et garantit la convergence p.s de la procédure itérative. QED. \square

Théorème 5.3.2.2

Démonstration. Remarquer que, à part $Y_{\Delta T}$, la dépendance de H en Y vient de la dépendance des fonctions λ et μ en Y . Notre résultat, découle de la définition des fonctions \sinh et \cosh et de la remarque suivante :

$$\int_0^{\Delta T} e^{k(\theta - Y_t)} dt = e^{k\theta} \sqrt{\frac{\int_0^{\Delta T} e^{-kY_t} dt}{\int_0^{\Delta T} e^{kY_t} dt}} \sqrt{\int_0^{\Delta T} e^{-kY_t} dt \int_0^{\Delta T} e^{kY_t} dt} = e^{k\theta + \beta_k} e^{\alpha_k}$$

De même :

$$\int_0^{\Delta T} e^{-k(\theta - Y_t)} dt = e^{-k\theta} \sqrt{\frac{\int_0^{\Delta T} e^{kY_t} dt}{\int_0^{\Delta T} e^{-kY_t} dt}} \sqrt{\int_0^{\Delta T} e^{-kY_t} dt \int_0^{\Delta T} e^{kY_t} dt} = e^{-k\theta - \beta_k} e^{\alpha_k}.$$

\square

Conclusion

Dans ce dernier chapitre de conclusion, nous citons les idées générales des chapitres précédents en surlignant nos contributions scientifiques.

La première idée concerne la décomposition des algorithmes de *trading* en stratégies et tactiques. Alors que les stratégies ont une logique financière dominante et un horizon temporel plus large, la logique d'optimisation (et le *payoff*) d'une tactique est basée sur des aspects microstructurels de l'interaction dans le marché, et même pour des stratégies dont la logique financière n'est pas la même, les tactiques associées peuvent être similaires (c'est par exemple le cas du *market-making* et de la liquidation optimale, qui obéissent au même type de logique au niveau tactique - c'est-à-dire un algorithme qui capture de la liquidité et interagit avec le carnet d'ordres afin de maximiser une fonction d'utilité espérée).

Dans cette étude, nous avons également voulu mettre en avant l'utilisation de modèles mathématiques et statistiques de type *black-box*, c'est-à-dire qu'au lieu de modéliser le carnet d'ordres de manière exacte, nous caractérisons la liquidité par un ensemble de paramètres de faible dimension qui sont estimés à partir de données ; cela rend le modèle plus flexible. C'est le cas du modèle de Avellaneda-Stoikov qui permet un traitement mathématique direct des problèmes de *trading* optimal (via la programmation dynamique ou les techniques d'apprentissage *on-line*).

Enfin, nous avons comparé deux approches d'optimisation pour s'attaquer aux problèmes *multi-stage* que nous traitons tout au long de cette étude : le principe de programmation dynamique (équation de Bellman) qui résout les problèmes par un raisonnement *backward*, et, d'autre part, l'apprentissage *on-line* qui tend à être plus "*model-free*" et permet à l'algorithme d'apprendre de son interaction avec l'environnement. Les deux approches présentent des avantages et des inconvénients (par exemple, dans le cas du *market-making*, la programmation dynamique est plus naturelle pour le contrôle de l'inventaire, tandis que l'apprentissage *on-line* représente une manière de suivre les tendances pour s'adapter aux différentes dynamiques du marché).

Concernant nos contributions dans ce travail, nous avons prouvé l'existence de la solution du problème du contrôle stochastique modélisant la problématique de la liquidation optimale dans le cadre des ordres au marché comme dans celui des ordres limites. Nous avons, également, trouvé des formules explicites des solutions dans des cas particuliers très utilisés dans l'industrie.

Tout au long de ce travail, nous n'étudions pas seulement les problèmes mathématiques, mais nous les plaçons dans leur contexte industriel afin de comprendre une série de questions, de la modélisation mathématique à l'implémentation des méthodes numériques. La calibration sur des données réelles a également été présentée.

Plusieurs problèmes peuvent être dérivés de notre étude. Tout d'abord, les plus ordinaires consistent à étendre nos résultats à des cas où le volume des transactions est stochastique, à donner des formes générales à la probabilité d'exécution ou à appliquer les mêmes idées à d'autres types de modèles (e.g, dans lesquels le processus du prix de référence est modélisé par une différence de deux processus de Hawkes, etc). Il est à noter que le paradigme du *trading* optimal basé sur l'approximation stochastique rejoint d'une manière naturelle le cadre de l'apprentissage par renforcement profond (*deep Q-learning*) que nous pouvons traiter, probablement, dans le futur.

Présentation de l'Organisme d'accueil

Dans le cadre de notre stage de fin d'Études, nous avons effectué notre stage d'une durée de quatre mois au sein de
La Marocaine Vie.



FIGURE A.1 – Logo de la marocaine vie

La Marocaine-Vie, seule compagnie spécialisée dans les branches d'assurance de personnes, est présente sur le marché marocain depuis 1978. Acquisée en 2001 par le Groupe Société Générale (via Sogécap, compagnie d'assurance vie du Groupe, et Société Générale Maroc), elle a depuis lors réorienté son activité vers le développement de la bancassurance avec les réseaux et filiales de la Société Générale au Maroc.

Parmi les valeurs essentielles fixées par la Marocaine Vie nous trouvons :

- Le professionnalisme : Cette caractéristique est renforcée par le développement du savoir faire et de l'échange de l'expérience.
- L'esprit d'équipe : Le partage du savoir et du savoir faire est l'une des caractéristiques du travail en groupe qui permet d'aboutir à un travail efficace et par suite une amélioration remarquable des prestations.
- L'innovation : La marocaine vie essaye d'inventer de nouveaux produits et de nouveaux services qui répondent aux besoins de sa clientèle et attirent des nouveaux adhérents.

La répartition du capital de la Marocaine Vie est présentée dans le tableau suivant :

Actionnaire	%
Sogécap	74%
Société Générale Maroc	26%

TABLE A.1 – La répartition du capital de la Marocaine Vie

Avec :

Sogécap : Compagnie d'assurance vie et de capitalisation du Groupe Société Générale, est en France l'un des leaders de la bancassurance. En fin 2009, Sogécap gère plus de 8 millions de contrats pour un encours supérieur à 70 milliards d'euros.

La société Générale : Banque universelle créée en 1913, Société Générale Maroc est l'une des principales filiales du Groupe Société Générale France. Groupe Société Générale Maroc est le premier groupe financier international du pays.

Les différentes divisions composant la Marocaine Vie sont présentées dans l'organigramme suivant :

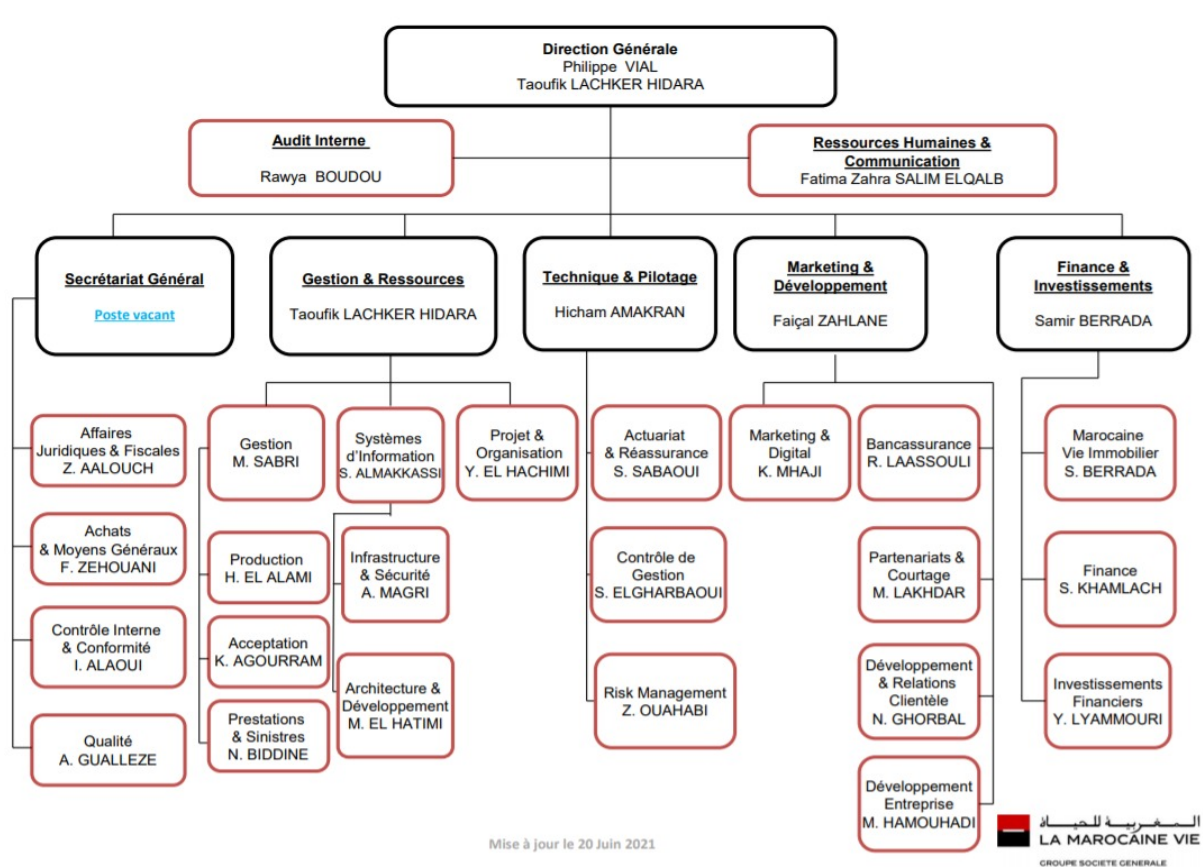


FIGURE A.2 – Organigramme de La Marocaine Vie

Le présent projet s'est déroulé au sein du département Investissement de La Marocaine Vie. Ce département se divise en 3 entités, comme expliqué ci-dessous :



FIGURE A.3 – Répartition du département investissement

Le rôle du département Investissements Financiers peut se résumer par la schématisation ci-dessous :

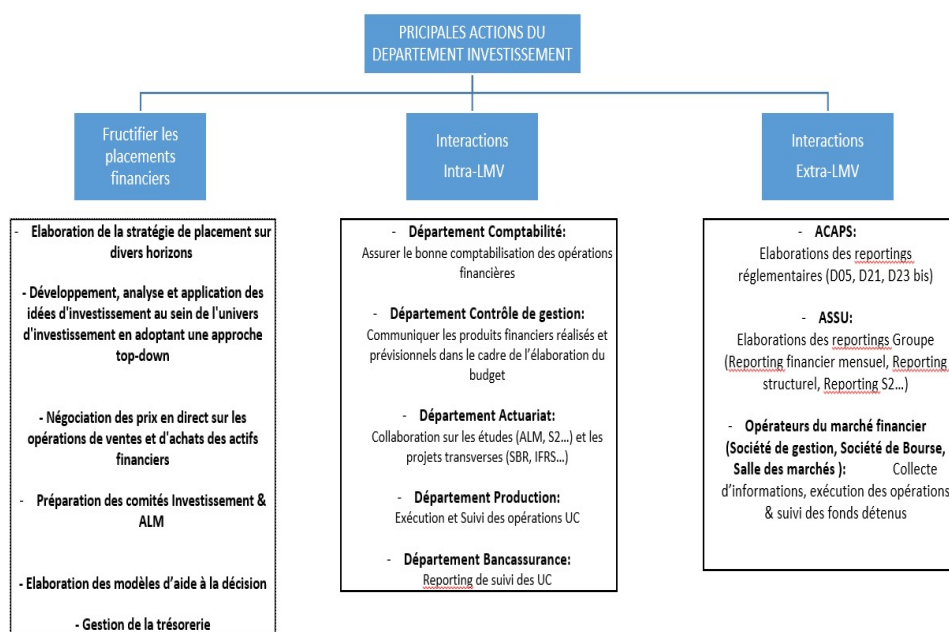


FIGURE A.4 – Répartition du département investissement

Bibliographie

- [1] Frédéric Abergel, Jean-Philippe Bouchaud, Thierry Foucault, Mathieu Rosenbaum, and Charles-Albert Lehalle. *Market microstructure : confronting many viewpoints*. Wiley, 2012.
- [2] Frédéric Abergel, Marouane Anane, Anirban Chakraborti, Aymen Jedidi, and Ioane Muni Toke. *Limit Order Books*. Physics of Society : Econophysics and Sociophysics. Cambridge University Press, 2016.
- [3] Yacine Ait-Sahalia, Per A. Mykland, and Lan Zhang. Ultra high frequency volatility estimation with dependent microstructure noise. (2005,30), 2005.
- [4] Aurélien Alfonsi, Antje Fruth, and Alexander Schied. Optimal execution strategies in limit order books with general shape functions. 2007.
- [5] Robert Almgren. Optimal trading in a dynamic market. Technical report, 2008.
- [6] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, pages 5–39.
- [7] Robert F. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10(1) :1–18, 2003.
- [8] David Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2009.
- [9] Marco Avellaneda and Sasha Stoikov. High frequency trading in a limit order book. *Quantitative Finance*, 8 :217–224, 04 2008.
- [10] Mohammad Gheshlaghi Azar, Vicenç Gómez, and Hilbert J. Kappen. Dynamic policy programming. *J. Mach. Learn. Res.*, 13(1) :3207–3245, nov 2012.
- [11] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Modeling microstructure noise with mutually exciting point processes, 2011.
- [12] Emmanuel Bacry, Adrian Iuga, Matthieu Lasnier, and Charles-Albert Lehalle. Market impacts and the life cycle of investors orders. 2014.
- [13] Alessandro Balata and Jan Palczewski. Regress-later monte carlo for optimal control of markov processes, 2017.
- [14] Alessandro Balata and Jan Palczewski. Regress-later monte carlo for optimal inventory control with applications in energy, 2018.
- [15] Peter Bank and Antje Fruth. Optimal Order Scheduling for Deterministic Liquidity Patterns. (1310.3077), October 2013.

- [16] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully non-linear second order equations. In *29th IEEE Conference on Decision and Control*, pages 2347–2349 vol.4, 1990.
- [17] Guy Barles and Espen Jakobsen. Error bounds for monotone approximation schemes for hamilton-jacobi-bellman equations. *Siam Journal on Numerical Analysis - SIAM J NUMER ANAL*, 43 :540–558, 01 2005.
- [18] BARRA. *Market impact model handbook*. Berkeley, California, Barra,, 1997.
- [19] Christian Bayer, Martin Redmann, and John Schoenmakers. Dynamic programming for optimal stopping via pseudo-regression, 2018.
- [20] Erhan Bayraktar and Michael Ludkovski. Liquidation in limit order books with controlled intensity, 2011.
- [21] Richard Bellman. *Dynamic Programming*. Dover Publications, 1957.
- [22] Michel Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de probabilités de Strasbourg*, 33 :1–68, 1999.
- [23] Christian Bender and Jessica Steiner. Least-squares monte carlo for backward sdes. 2012.
- [24] Albert Benveniste, Pierre Priouret, and Michel Métivier. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, Berlin, Heidelberg, 1990.
- [25] Paul Besson, Stéphanie Pelin, and Matthieu Lasnier. To cross or not to cross the spread : That is the question. *The Journal of Trading*, 11(4) :77–91, 2016.
- [26] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons Inc., second edition, 1999. A Wiley-Interscience Publication.
- [27] Bruno Bouchard. Introduction to stochastic control of mixed diffusion processes, viscosity solutions and applications in finance and insurance. 01 2007.
- [28] Bruno Bouchard and Nizar Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49, 01 2011.
- [29] J. P. Bouchaud. Price impact, 2009.
- [30] Jean-Philippe Bouchaud, J. Dooyne Farmer, and Fabrizio Lillo. How markets slowly digest changes in supply and demand, 2008.
- [31] Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart. Fluctuations and response in financial markets : the subtle nature of random price changes. *Quantitative Finance*, 4(2) :176–190, apr 2004.
- [32] Pierre Bremaud. *Point Process Calculus in Time and Space An Introduction with Applications*. Springer, 1 edition, 2021.
- [33] Frédéric Bucci, Michael Benzaquen, Fabrizio Lillo, and Jean-Philippe Bouchaud. Slow decay of impact in equity markets : insights from the ancerno database, 2019.
- [34] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *SSRN Electronic Journal*, 01 2015.
- [35] Álvaro Cartea, Sebastian Jaimungal, and Jason Ricci. Buy low, sell high : A high frequency trading perspective. *SIAM Journal on Financial Mathematics*, 5(1) :415–444, 2014.
- [36] Rama Cont, Sasha Stoikov, and Rishi Talreja. A stochastic model for order book dynamics. *Operations Research*, 58(3) :549–563, 2010.
- [37] D. Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. I : Elementary theory and methods. 2nd ed*, volume Vol. 1. 01 2003.
- [38] Daryl Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. II : General theory and structure. 2nd revised and extended ed*, volume 2. 03 1990.

- [39] Marie Duflo and Stephen S. Wilson. *Random Iterative Models*. Springer-Verlag, Berlin, Heidelberg, 1st edition, 1997.
- [40] David A. Easley, Marcos M. López de Prado, and Maureen O'Hara. The microstructure of the "flash crash" : Flow toxicity, liquidity crashes, and the probability of informed trading. *The Journal of Portfolio Management*, 37 :118 – 128, 2011.
- [41] Zoltan Eisler, Jean-Philippe Bouchaud, and Julien Kockelkoren. The price impact of order book events : market orders, limit orders and cancellations, 2009.
- [42] Zoltan Eisler, János Kertész, and Fabrizio Lillo. The limit order book on different time scales. *Proceedings of SPIE - The International Society for Optical Engineering*, 6601, 06 2007.
- [43] Nicole El Karoui and Monique Jeanblanc Picque. Contrôle de processus de Markov. *Séminaire de probabilités de Strasbourg*, 22 :508–541, 1988.
- [44] Briane Marc et Pagès Gilles. *Théorie de l'intégration : Cours et exercices*. Vuibert, 4ed. edition, 2006.
- [45] J. Doynne Farmer, Paolo Patelli, and Ilija I. Zovko. The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Sciences*, 102(6) :2254–2259, 2005.
- [46] Joaquin Fernandez-Tapia. High-frequency trading meets reinforcement learning : Exploiting the iterative nature of trading algorithms. *Capital Markets : Market Microstructure eJournal*, 2015.
- [47] Joaquin Fernandez-Tapia. *Modeling, optimization and estimation for the on-line control of trading algorithms in limit-order markets*. PhD thesis, 09 2015.
- [48] Ben G. Fitzpatrick and Wendell H. Fleming. Numerical methods for an optimal investment-consumption model. *Math. Oper. Res.*, 16(4) :823–841, nov 1991.
- [49] Gersende Fort, Eric Moulines, Amandine Schreck, and Matti Vihola. Convergence of markovian stochastic approximation with discontinuous dynamics, 2014.
- [50] Mark B. Garman and Michael J. Klass. On the estimation of security price volatilities from historical data. *The Journal of Business*, 53(1) :67–78, 1980.
- [51] Jim Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7) :749–759, 2010.
- [52] Laszlo Gillemot, J. Farmer, and Fabrizio Lillo. There's more to volatility than volume. *Quantitative Finance*, 6 :371–384, 09 2005.
- [53] Martin D. Gould, Mason A. Porter, Stacy Williams, Mark McDonald, Daniel J. Fenn, and Sam D. Howison. Limit order books, 2010.
- [54] Olivier Guéant. *The Financial Mathematics of Market Liquidity : From Optimal Execution to Market Making*. CRC Press - Taylor and Francis, 2016.
- [55] Olivier Guéant, Charles-Albert Lehalle, and Joaquin Fernandez-Tapia. Dealing with the inventory risk : a solution to the market making problem. *Mathematics and Financial Economics*, 7(4) :477–507, sep 2012.
- [56] Olivier Guéant, Charles-Albert Lehalle, and Joaquin Fernandez-Tapia. Optimal portfolio liquidation with limit orders. *SIAM Journal on Financial Mathematics*, 3(1) :740–764, jan 2012.
- [57] Fabien Guilbaud and Huyen Pham. Optimal high frequency trading with limit and market orders, 2011.
- [58] Fabien Guilbaud and Huyên Pham. Optimal high frequency trading in a pro-rata microstructure with predictive information. *Mathematical Finance*, 25, 05 2012.

- [59] Olivier Guéant. Execution and block trade pricing with optimal constant rate of participation, 2012.
- [60] Olivier Guéant. Optimal execution and block trade pricing : a general framework, 2012.
- [61] Olivier Guéant. Optimal execution of asr contracts with fixed notional, 2014.
- [62] Olivier Guéant and Charles-Albert Lehalle. General intensity shapes in optimal liquidation, 2012.
- [63] Olivier Guéant and Jiang Pu. Option pricing and hedging with execution costs and market impact, 2013.
- [64] Olivier Guéant, Jiang Pu, and Guillaume Royer. Accelerated share repurchase : pricing and execution strategy, 2013.
- [65] Olivier Guéant and Guillaume Royer. Vwap execution and guaranteed vwap. *SIAM Journal on Financial Mathematics*, 5(1) :445–471, 2014.
- [66] Dean S. Clark (auth.) Harold J. Kushner. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Applied Mathematical Sciences (Switzerland) N.26. Springer, 1 edition, 1978.
- [67] Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1) :83–90, 1971.
- [68] Gur Huberman and Werner Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4) :1247–1275, 2004.
- [69] Côme Huré. *Numerical Methods and Deep Learning for Stochastic Control Problems and Partial Differential Equations*. Theses, Université Paris Diderot (Paris 7), Sorbonne Paris Cité, June 2019.
- [70] Ken ito Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1st edition, 1999.
- [71] Ken ito Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1st edition, 1999.
- [72] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. 1987.
- [73] Hilbert J. Kappen. An introduction to stochastic control theory, path integrals and reinforcement learning. In Joaquín Marro, Pedro L. Garrido, and Joaquín J. Torres, editors, *Cooperative Behavior in Neural Systems*, volume 887 of *American Institute of Physics Conference Series*, pages 149–181, February 2007.
- [74] Idris Kharroubi, Nicolas Langrené, and Huyên Pham. A numerical algorithm for fully nonlinear HJB equations : An approach by control randomization. *Monte Carlo Methods and Applications*, 20(2), jan 2014.
- [75] J. Kiefer and J. Wolfowitz. Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, 23(3) :462 – 466, 1952.
- [76] Andrei Kirilenko, Albert Kyle, Mehrdad Samadi, and Tugkan Tuzun. The flash crash : The impact of high frequency trading on an electronic market. *SSRN Electronic Journal*, 05 2011.
- [77] Robert Kissell. Statistical methods to compare algorithmic performance. *The Journal of Trading*, 2(2) :53–62, 2007.
- [78] Robert Kissell. Transaction cost analysis. *The Journal of Trading*, 3 :29–37, 03 2008.
- [79] Peter Kratz and Torsten Schöneborn. Optimal liquidation in dark pools. *Quantitative Finance*, 14(9) :1519–1539, 2014.

- [80] N. V. Krylov. On the rate of convergence of finite-difference approximations for bellman's equations with variable coefficients. *Probability Theory and Related Fields*, 117(1) :1–16, May 2000.
- [81] Harold J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM Journal on Control and Optimization*, 28(5) :999–1048, 1990.
- [82] Harold J. Kushner and George Yin. *Stochastic Approximation and Recursive Algorithms with Applications*, 2nd edn. Springer-Verlag, 2003.
- [83] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6) :1315–1335, 1985.
- [84] Damien Lamberton and Gilles Pagès. How fast is the bandit? *Stochastic Analysis and Applications*, 26(3) :603–623, 2008.
- [85] Damien Lamberton, Gilles Pagès, and Pierre Tarrès. When can the two-armed bandit algorithm be trusted? *The Annals of Applied Probability*, 14(3), aug 2004.
- [86] Charles-Albert Lehalle; Sophie Laruelle. *Market Microstructure in Practice*. World Scientific, 2nd edition, 2018.
- [87] Sophie Laruelle. Faisabilité de l'apprentissage des paramètres d'un algorithme de trading sur des données réelles. *Cahier de la Chaire Finance and Sustainable Development, Hors-Série Microstructure des marchés*, 2013.
- [88] Sophie Laruelle, Charles-Albert Lehalle, and Gilles Pagès. Optimal split of orders across liquidity pools : a stochastic algorithm approach, 2009.
- [89] Sophie Laruelle, Charles-Albert Lehalle, and Gilles Pagès. Optimal posting price of limit orders : learning by trading. 2011.
- [90] Sophie Laruelle and Gilles Pagès. Stochastic approximation with averaging innovation applied to finance, 2010.
- [91] V.A. Lazarev. Convergence of stochastic-approximation procedures in the case of a regression equation with several roots. 28(3) :66–78, 1992.
- [92] Charles-Albert Lehalle. The impact of liquidity fragmentation on optimal trading. *Trading*, 2009(1) :80–87, 2009.
- [93] Charles-Albert Lehalle. Market microstructure knowledge needed for controlling an intra-day trading process. 2013.
- [94] Charles-Albert Lehalle and Ngoc Minh Dang. Rigorous post-trade market impact measurement and the price formation process. *Trading*, 2010(1) :108–114, 2010.
- [95] Vincent Lemaire and Gilles Pagès. Unconstrained recursive importance sampling. *The Annals of Applied Probability*, 20(3), jun 2010.
- [96] M. Métivier and P. Priouret. Théorèmes de convergence presque sure pour une classe d'algorithmes stochastiques à pas décroissant. *Probability Theory and Related Fields*, 74(3) :403–428, 1987.
- [97] Selvaprabu Nadarajah, François Margot, and Nicola Secomandi. Comparison of least squares monte carlo methods with applications to energy real options. *European Journal of Operational Research*, 256(1) :196–204, 2017.
- [98] Anna Obizhaeva and Jiang Wang. Optimal trading strategy and supply /demand dynamics. *Journal of Financial Markets*, 16(1) :1–32, 2013.
- [99] Maureen O'Hara. *Market Microstructure Theory*. Wiley, 1995.
- [100] Maureen O'Hara. High frequency market microstructure. *Journal of Financial Economics*, 116(2) :257–270, 2015.

- [101] Bernt Øksendal and Agnès Sulem. *Applied Stochastic Control of Jump Diffusions*. Springer, Universitext, 2019. 3rd edition.
- [102] Gilles Pagès. Functional co-monotony of processes with an application to peacocks. Technical report, September 2012.
- [103] Gilles Pagès and Jacques Printems. Optimal quantization for finance : from random vectors to stochastic processes. *Handbook of Numerical Analysis*, 15, 2008.
- [104] Gilles Pagès. *Introduction to Numerical Probability applied to finance*. 09 2014.
- [105] Gilles Pagès, Huyên Pham, Modeles Aleatoires, Jacques Printems, and Centre Mathématiques. Optimal quantization methods and applications to numerical problems in finance. 06 2003.
- [106] Mariane Pelletier. Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing. *The Annals of Applied Probability*, 8(1):10–44, 1998.
- [107] Marc Potters and Jean-Philippe Bouchaud. More statistical properties of order books and price impact. *Physica A : Statistical Mechanics and its Applications*, 324(1-2) :133–140, jun 2003.
- [108] Herbert Robbins and Sutton Monro. A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22(3) :400 – 407, 1951.
- [109] Christian Robert. A new approach for the dynamics of ultra-high-frequency data : The model with uncertainty zones. *Journal of Financial Econometrics*, 9 :344–366, 06 2010.
- [110] Christian Robert and Mathieu Rosenbaum. Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Mathematical Finance*, 22 :133 – 164, 01 2012.
- [111] Mathieu Rosenbaum. A new microstructure noise index. *Quantitative Finance*, 11(6) :883–899, 2011.
- [112] Afef Sellami. Méthodes de quantification optimale pour le filtrage et applications à la finance. Master’s thesis, Université Paris Dauphine, 2005.
- [113] Eric Smith, J Doyne Farmer, László Gillemot, and Supriya Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative Finance*, 3(6) :481–514, dec 2003.
- [114] Julyerme Tonin, João Tonin, Marina Cunha, and José Bornia. Lead-lag effect in the stock market of brics. volume 5, pages 54–66, 12 2013.
- [115] Agnès Tourin and Thaleia Zariphopoulou. Numerical schemes for investment models with singular transactions. *Comput. Econ.*, 7(4) :287–307, nov 1994.
- [116] Nizar Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*, volume 29. 01 2013.
- [117] Bence Tóth and János Kertész. The epps effect revisited. *Quantitative Finance*, 9 :793–802, 10 2009.
- [118] Lan Zhang, Per A Mykland, and Yacine Aït-Sahalia. A tale of two time scales. *Journal of the American Statistical Association*, 100(472) :1394–1411, 2005.
- [119] José Penalva Álvaro Cartea, Sebastian Jaimungal. *Algorithmic and High-Frequency Trading*. Mathematics, Finance and Risk. Cambridge University Press, 1 edition, 2015.

