



**HAL**  
open science

# Contrôle optimal, apprentissage statistique et modélisation du carnet d'ordres

Othmane Mounjid

► **To cite this version:**

Othmane Mounjid. Contrôle optimal, apprentissage statistique et modélisation du carnet d'ordres. Optimisation et contrôle [math.OC]. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLX110 . tel-02489456

**HAL Id: tel-02489456**

**<https://theses.hal.science/tel-02489456v1>**

Submitted on 24 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimal control and statistical learning for order book modelling

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'École Polytechnique

École doctorale n°574 Mathématiques Hadamard (EDMH)  
Spécialité de doctorat: Mathématiques appliquées

Thèse présentée et soutenue à Paris, le 17 Décembre 2019, par

**Othmane Mounjid**

Composition du Jury :

Bruno BOUCHARD Professeur, Université Paris-Dauphine	Président
Olivier GUEANT Professeur, Université Paris 1 Panthéon-Sorbonne	Rapporteur
Xin GUO Professeur, University of California, Berkeley	Rapporteuse
Jean-Philippe BOUCHAUD Membre académie des sciences à l'institut de France	Examineur
Sophie LARUELLE Maitre de conférence, Université Paris-Est Créteil	Examinatrice
Charles-Albert LEHALLE Head of Data Analytics, Capital Fund Management	Examineur
Eric MOULINES Professeur, École Polytechnique	Examineur
Gilles PAGES Professeur, Sorbonne Université	Examineur
Mathieu ROSENBAUM Professeur, École Polytechnique	Directeur de thèse



A mes parents...



---

## Remerciements

---

Je tiens en premier lieu à témoigner toute ma reconnaissance et ma gratitude envers mon directeur de thèse Mathieu Rosenbaum. Mathieu je te remercie pour ton soutien, ta confiance et tes conseils avisés. C'est grâce à toi que j'ai pu mener à bien ce travail. J'ai appris et découvert énormément sous ta direction. Pour finir ce paragraphe sur une note humoristique, je tenais à dire que j'ai toujours admiré ta capacité à garder ton sang froid en toute circonstance. Par exemple, Rennes, ton équipe favorite, après un début de saison historique en ligue 1 cette année s'est vue reléguée en milieu de tableau. Toutefois, tu es resté impassible face à ce petit choc émotionnel. Heureusement que la tendance va se renverser avec l'arrivée de jeunes pépites tel qu'Eyal qui porte le maillot du club dès ses 2 ans et connaît par cœur la nationalité des différents joueurs Rennais. J'espère juste qu'il n'occupera pas le poste de gardien de but comme son aïeul.

Je remercie chaleureusement Charles-Albert Lehalle pour sa disponibilité, sa bienveillance et ses critiques constructives. Charles tu es pour moi un second directeur de thèse. Pour l'anecdote, je prévoyais d'entamer une thèse en physique à la fin de mes études et c'est toi qui m'as alors proposé de poursuivre une thèse en finance avec Mathieu.

Je remercie sincèrement les différentes personnes avec lesquelles j'ai pu travailler notamment Nicolas Baradel, Bruno Bouchard, Thomas Deschatre, Pierre Gruet et Pamela Saliba, pour leurs échanges et leur disponibilité. J'ai appris plein de choses tant du point de vue humain que scientifique à vos côtés.

J'exprime mes plus sincères remerciements à Xin Guo et Olivier Guéant qui ont accepté de rapporter ma thèse. Je vous remercie pour votre lecture attentive et détaillée, l'intérêt porté à mon travail et vos commentaires pertinents et constructifs. J'adresse un remerciement tout particulier à Olivier pour le temps qu'il m'a consacré et pour m'avoir invité au séminaire qu'il organise. Merci pour ta bienveillance. Je tiens également à remercier Xin pour ses conseils avisés et l'opportunité de post-doc qu'elle m'a proposée. Toute ma reconnaissance va ensuite aux membres du jury Bruno Bouchard, Jean-Philippe Bouchaud, Sophie Laruelle, Eric Moulines et Gilles Pagès, qui me font l'honneur d'examiner ma thèse.

Je suis infiniment reconnaissant envers mes professeurs, en particulier Emmanuel Gobet, Antoine Henrot, Nicole El Karoui, Gilles Pagès, Stefano de Marco et Nizar Touzi. Comme la liste est longue, je m'excuse d'avance pour les omissions. Merci pour m'avoir transmis une fraction aussi infime soit-elle de votre connaissance, de votre passion pour les mathématiques et votre goût pour la recherche. J'ai été impressionné par vos compétences scientifiques mais encore plus touché par votre côté humain.

Je me dois de remercier les doctorants du CMAP et d'ailleurs dont la simple présence suffit à égayer le quotidien du personnel du laboratoire. Me connaissant et comme la liste est touffue, je suis sûr d'oublier quelques personnes. J'appelle donc à votre indulgence. Je commence

---

par remercier Mathieu Kholi pour sa gentillesse et sa bonne humeur ainsi qu'Hadrien de March qui peut surprendre même les plus accoutumés par son humour insolite. Cet humour se base sur la capacité légendaire de M. de March à identifier un lien tangible entre deux objets lointains. Je remercie également Martin Averseng pour sa gentillesse, son humilité et sa passion de transmettre et partager. Je ne peux oublier Massil Achab capitaine du Real Matrix qui aux côtés de notre gardien Alain Virouleau constituent une muraille que les meilleurs attaquants ont du mal à franchir. J'en profite aussi pour lancer un clin d'œil à l'ailier fort Vincent Chapellier et l'attaquant phare Wil (dont je n'ai jamais su le nom complet) du Real Matrix, deux bons vivants dont on ne se lasse jamais de leur compagnie. Je remercie également M. Heythem Farhat ce personnage au caractère exceptionnel qui renferme en son cœur une bonté inouïe. Je remercie Kevish Napal notre mauricien préféré, Remi Besson notre grand philosophe qui manie la prose aussi bien que le vers et Juliette Chevalier dont il faut éviter la provocation au risque de recevoir un coup de Krav-maga fatal. Ce n'est que très récemment que j'ai découvert l'humour assez fin de Juliette. Un grand merci à Mathilde Boissier la fille que j'apprécie le plus au laboratoire de part son caractère naturel et spontané, sa simplicité et sa grande bonté. Un merci encore plus grand aux personnes suivantes dont je ne cite que le nom par paresse: Paul Thevenin fervent supporter du PSG, Jean-Bernard Eytard notre dictionnaire ambulante, Perle Geoffroy notre rameuse d'exception, Cheikh Toure, Celine Bonnet, Florian Bourgey, Maxime Grangereau, Florian Feppon, Cyril Benezet pour son humour, son sérieux et son sens de la répartie, Ilias Ftouhi pour sa compagnie et sa bonne humeur, Brian Winchester notre premier chef cuisinier, Vincent Laurent notre second chef cuisinier, Corentin Chatelier le musicien du groupe, etc. Je dédie une pensée toute particulière aux membres du bureau 2003: Pierre Cordesse, Rubben Di Battista, Alejandro Fernandez-Montero, Thomas Lartigue, Apolline Louvet, Marc-Arthur N'Guessan, Léa Nicolas, Quentin Wargner et Alberto Remigi. Enfin, je suis reconnaissant envers nos deux informaticiens de génie Sylvain Ferrand et Pierre Straebler qui partagent une forte complicité mais dont la vision du sport les sépare.

Je remercie cordialement tout les membres de l'équipe Mathieu. Je commence par Saad le doyen de l'équipe, grand sportif au caractère réservé et extrêmement humble. Je remercie Omar pour la bonne ambiance qu'il apporte ainsi que son humour "faiblement" ironique, Pamela dont on attend toujours l'invitation au Liban, Eduardo grand esprit qui a juste besoin de réguler sa consommation de whisky, Emel pour ses discussions enflammées et ses avis tranchés, Paul le rugby-man pour son dynamisme, Bastien le corse dont je prédis un avenir prospère en affaires. Je n'oublie pas de remercier Joffrey le cadet de l'équipe qui fait preuve d'un calme olympien et d'une grande sagesse. Enfin, je remercie Mehdi pour sa sympathie, sa simplicité, sa générosité et sa tendance naturelle à évoquer des sujets qui énervent Bastien, Aditi pour sa gentillesse et sa sympathie bien que je n'ai toujours pas réussi à cerner son caractère insolite et Marcos grand maître Jedi qui suit de loin l'avancement des différents membres. On oublie souvent de te transmettre les liens vers nos derniers travaux mais tu parviens toujours à les retrouver et à les mettre sur le site. Merci pour ta disponibilité, tes conseils, ta générosité et ta bienveillance.

Il me tient à cœur de remercier M. Claisse dont j'ai toujours apprécié la simplicité, la compagnie et l'humour et M. Mastrolia pour sa sympathie et son moral d'acier. Je tiens également à remercier le duo Flore et Aline pour leur bonté, leur bienveillance et leur sérieux. Je remercie aussi Aldjia pour sa réactivité, sa perspicacité et ses conseils éclairés. Enfin, je tiens à remercier M. Bansaye pour sa gentillesse, sa spontanéité et sa générosité et M. Gerin pour son humour, sa présence et ses interventions toujours aussi pertinentes.

---

Je remercie tout le secrétariat du CMAP pour leur disponibilité, leur soutien et leur suivi notamment Maud Cadiz-Pena, Nassera Naar et Alexandra Noiret.

Avant de finir, je remercie messieurs Lacome et Lille-Palette pour m'avoir permis de travailler sur des données sportives ainsi que pour leur ouverture d'esprit, leur sympathie et leur réactivité.

Je ne remercierai jamais assez ma famille et en premier lieu mes parents pour leur présence, leur soutien et leurs encouragements. Je ne m'appesantirai pas plus sur cette partie au risque d'entamer éloges qui sera piètre quel que soit l'effort ou l'énergie investis. Je tiens juste à dire que vous occupez à mes yeux une place inégalée.





---

## Résumé

---

L'objectif principal de cette thèse est de comprendre les interactions entre les agents financiers et le carnet d'ordres. Elle se compose de six chapitres inter-connectés qui peuvent toutefois être lus indépendamment.

Nous considérons dans le premier chapitre le problème de contrôle d'un agent cherchant à prendre en compte la liquidité disponible dans le carnet d'ordres afin d'optimiser le placement d'un ordre unitaire. Pour ce faire, il peut soit rester dans le carnet d'ordres pour conserver son emplacement stratégique ou bien annuler son ordre afin de retarder son achat s'il pense pouvoir bénéficier d'un meilleur prix dans le futur. A la fin du temps imparti, si l'agent ne parvient à exécuter son ordre, il accepte de payer un coût de transaction et lance un ordre marché pour forcer l'achat. Notre stratégie permet de réduire le risque de sélection adverse: acheter alors que le prix baisse. Néanmoins, la valeur ajoutée de cette approche est affaiblie en présence de temps de latence: prédire les mouvements futurs des prix est peu utile si le temps de réaction des agents est lent.

Dans le chapitre suivant, nous étendons notre étude à un problème d'exécution plus général où les agents traitent des quantités non unitaires afin de limiter leur impact sur le prix. On considère donc un agent représentatif qui se demande quelle est la meilleure manière d'acheter une quantité donnée pendant une durée déterminée en utilisant des ordres de marché, des ordres limites ou des annulations? Pour résoudre ce problème, on introduit une formule qui permet un calcul efficace de l'impact sur le prix. On utilise ensuite cette formule pour obtenir une tactique qui produit de meilleurs résultats que les stratégies d'exécution classiques.

Dans le troisième chapitre, on s'inspire de l'approche précédente pour résoudre cette fois des problèmes de market making plutôt que des problèmes d'exécution. Le rôle des market makers est de produire de la liquidité. Pour ce faire, ils proposent des prix à l'achat en général plus bas que ceux à la vente ce qui leur permet de réaliser des profits. Ainsi, ils utilisent surtout des ordres limites et envoient des ordres de marché en principe pour gérer leur inventaire uniquement. La résolution de ce problème nous permet de proposer des stratégies pertinentes compatibles avec les actions typiques des market makers. Ensuite, nous modélisons les comportements des traders haute fréquence directionnels et des brokers institutionnels dans le but de simuler un marché où nos trois types d'agents interagissent de manière optimale les uns avec les autres.

Nous proposons dans le quatrième chapitre un modèle d'agents où la dynamique des flux dépend non seulement de l'état du carnet d'ordres mais aussi de l'historique du marché. Pour ce faire, nous utilisons des généralisations des processus de Hawkes non linéaires. En effet, le modèle pour le carnet d'ordres introduit dans les parties précédentes est réaliste et permet d'estimer les quantités d'intérêt. Toutefois, il présente quelques restrictions. Tout d'abord, les décisions prises par les agents dépendent uniquement de l'état du carnet d'ordres alors que les

---

stratégies des market makers et traders haute fréquence peuvent être plus complexes. Ensuite, les agents sont agrégés en trois groupes: market makers, traders haute fréquence directionnels et brokers institutionnels. Il est donc difficile d'analyser les différences de comportement entre les acteurs qui appartiennent au même groupe. Pour dépasser ces difficultés, on présente un modèle plus général qui permet de calculer en fonction de flux individuels plusieurs indicateurs pertinents. Il est notamment possible de classer les market makers en fonction de leur contribution à la volatilité.

Pour résoudre les problèmes de contrôle soulevés dans la première partie de la thèse, nous avons développé des schémas numériques. Une telle approche est possible lorsque la dynamique du modèle est connue. Lorsque l'environnement est inconnu, on utilise généralement les algorithmes itératifs stochastiques. Dans le cinquième chapitre, nous proposons une méthode permettant d'accélérer la convergence de tels algorithmes.

Les approches considérées dans les chapitres précédents sont adaptées pour des marchés liquides utilisant le mécanisme du carnet d'ordres. Cependant, cette méthodologie n'est plus nécessairement pertinente pour des marchés régis par des règles de fonctionnement spécifiques. Pour répondre à cette problématique, nous proposons, dans un premier temps, d'étudier le comportement des prix sur le marché très particulier de l'électricité.



---

# Abstract

---

This thesis aims at understanding the interactions between the market participants and the order book. It consists of six connected chapters which can however be read independently.

In the first chapter, we tackle the control problem of an agent who wishes to exploit the order book liquidity to optimise the placement of a unit limit order during a fixed time period. To do so, the agent can insert limit orders to keep its strategical placement or cancel its already insert limit order and delay the execution hoping to get a better execution price in the future. When the order is not executed at the end of the period, the agent sends a market order and crosses the spread to guarantee the execution. We show that our optimal tactic reduces the adverse selection risk: the execution of a buy limit order followed by a price decrease. Nonetheless, the added value of taking into account order book liquidity is eroded by latency: being able to predict future price moves is less profitable if agents reaction time is large.

In the next chapter, we extend our study to more general execution problems where agents handle non-unit quantities to mitigate their price impact. For this, we consider an agent who aims at finding the optimal way to buy a given amount of assets over some fixed time interval using limit orders, market orders and cancellations. To solve this problem, we introduce a closed-form formula which enables an efficient computation of the price impact. Using this formula, we derive the agent optimal tactic and show that it outperforms significantly standard execution strategies.

The third chapter adapts our previous approach to solve market making issues. The role of the market makers is to provide liquidity. For this, they propose bid and ask prices and make money out of the difference between these two prices. Thus, they use mainly limit orders and should somehow send market orders only for inventory management purpose. The resolution of this problem enables us to propose relevant strategies which are consistent with typical market makers behaviours. After that, we model the behaviours of directional high frequency traders and institutional brokers in order to simulate an order book driven market with our three classes of agents interacting optimally with each others.

We introduce in the fourth chapter an agent-based model where the dynamics of the flow depend not only on the order book state but also on the history of the market. For this, we use generalisations of non-linear Hawkes processes. Although we introduce in the previous chapters a realistic agent-based model which enables to estimate the quantities of interest, such approach still have some limitations. First, agents decisions depend on the market history only through the last order book state while market makers and directional high frequency traders strategies may be more intricate. Second, agents are aggregated into three groups: market makers, directional high frequency traders and institutional brokers. Consequently, it is difficult to disentangle disparities among behaviours of market participants belonging to the same group. To overcome these difficulties, we propose a general order book model which allows us to compute several relevant microstructural indicators in terms of the individual

flows. It is notably possible to rank market makers according to their own contribution to volatility.

To solve the control problems appearing in the first part of the thesis, we develop numerical schemes. This is possible when the dynamics of the model is known. To tackle control problems in an unknown environment, it is common to use stochastic iterative algorithms. In the fifth chapter, we propose a method that accelerates the convergence of such algorithms.

The approaches built in the previous chapters are appropriate for liquid markets that use an order book mechanism. However our methodologies may not be suitable for exchanges with very specific operating rules. To investigate this issue, as a first step, we study the price behaviour of the very particular intra-day electricity market.

## List of papers being part of this thesis

- Charles-Albert Lehalle and Othmane Mounjid, *Limit order strategic placement with adverse selection risk and the role of latency*, Market Microstructure and Liquidity, Vol. 03, No. 01, 1750009, 2017.
- Charles-Albert Lehalle, Othmane Mounjid and Mathieu Rosenbaum, *Optimal liquidity-based trading tactics*, submitted, 2018.
- Nicolas Baradel, Bruno Bouchard, David Evangelista and Othmane Mounjid, *Optimal inventory management and order book modeling*, ESAIM: Proceedings and Surveys, Vol. 65, pages 45 - 181, 2019.
- Othmane Mounjid, Mathieu Rosenbaum and Pamela Saliba, *From asymptotic properties of general point processes to the ranking of financial agents*, submitted, 2019.
- Charles-Albert Lehalle and Othmane Mounjid, *Improving reinforcement learning algorithms: towards optimal learning rate policies*, submitted, 2019.
- Thomas Deschatre, Pierre Gruet, Naoufal Layad, Othmane Mounjid and Mathieu Rosenbaum, *Price impact study of the electricity intra-day market*, working paper, 2019.





---

# Contents

---

<b>Contents</b>	<b>xvi</b>
<b>Introduction</b>	<b>1</b>
Motivations . . . . .	1
Outline . . . . .	4
1    Chapter I - Limit order strategic placement with adverse selection risk and the role of latency . . . . .	6
1.1    Empirical evidences . . . . .	6
1.2    Problem formulation . . . . .	8
1.3    Numerical resolution . . . . .	10
2    Chapter II - Optimal liquidity-based trading tactics . . . . .	11
2.1    Order book modelling . . . . .	12
2.2    Optimal control problem . . . . .	14
2.3    Numerical experiments . . . . .	17
3    Chapter III - Optimal inventory management and order book modeling . . . .	18
3.1    MM control problem . . . . .	19
3.2    DHFT control problem . . . . .	23
3.3    IB control problem . . . . .	24
3.4    Full market simulation . . . . .	26
4    Chapter IV - From asymptotic properties of general point processes to the ranking of financial agents . . . . .	26
5    Chapter V - Improving reinforcement learning algorithms: towards optimal learning rate policies . . . . .	31
5.1    Improvement of the asymptotic convergence rate . . . . .	32
5.2    Optimal policy for the learning rate . . . . .	34
5.3    Numerical experiments . . . . .	35
6    Chapter VI - Price impact study on the electricity intra-day market . . . . .	37
6.1    General empirical analysis . . . . .	38
6.2    Price study . . . . .	39
6.3    Price modelling . . . . .	41
 <b>I    Limit order strategic placement with adverse selection risk and the role of         latency</b>	 <b>43</b>

1	Introduction . . . . .	43
2	Main Hypothesis and Empirical Evidences . . . . .	45
2.1	Database presentation . . . . .	45
2.2	Stylized fact 1: the predictive power of the order book imbalance . . . . .	46
2.3	Stylized fact 2: Agent's decisions depend on the order book liquidity . . . . .	48
3	The Dynamic Programming Principle Applied to Limit Order Placement . . . . .	52
3.1	Formalisation of the Model . . . . .	53
3.2	Discrete time framework . . . . .	56
4	A Qualitative Understanding . . . . .	59
4.1	Numerically Solving the Control Problem . . . . .	59
4.2	The Price of Latency . . . . .	64
5	Conclusion . . . . .	64
I.A	Transition probabilities of the markov chain $U_n$ . . . . .	67
I.B	Composition of market participants groups . . . . .	70
I.C	Extreme Imbalances . . . . .	70
<b>II</b>	<b>Optimal liquidity-based trading tactics</b>	<b>73</b>
1	Introduction . . . . .	73
2	Order book modelling . . . . .	77
2.1	Preliminary: Empirical evidences . . . . .	77
2.2	Order book framework . . . . .	80
2.3	Ergodicity . . . . .	81
2.4	Computation of the long-term average price move . . . . .	83
2.5	Numerical application . . . . .	85
3	Optimal tactic control problem . . . . .	86
3.1	Presentation of the stochastic control framework . . . . .	86
4	Theoretical results . . . . .	89
4.1	Price impact computation . . . . .	89
4.2	Existence and uniqueness of the optimal strategy, regularity properties . . . . .	90
4.3	Decisions taken at fixed frequency $\Delta^{-1}$ : dynamic programming equation . . . . .	91
4.4	Second approach: Decisions taken at any time . . . . .	93
5	Numerical experiments . . . . .	95
5.1	Computation of the optimal gain: decisions taken at a fixed frequency $\Delta^{-1}$ . . . . .	95
5.2	Computation of the optimal gain: decision taken at any time . . . . .	96
II.A	Model parameters estimation . . . . .	97
II.B	Ergodicity of the process $(U_t)$ . . . . .	100
II.B.1	Outline of the proof . . . . .	100
II.B.2	Proof . . . . .	100
II.C	Proof of the computation of the long-term average price move . . . . .	102
II.C.1	Proof of Proposition 1 . . . . .	102
II.C.2	Computation methodology of $q_{i,i'}^\pm$ . . . . .	104

II.C.3	Diagonalisation of $\tilde{Q}^*$ . . . . .	105
II.D	Generalities about the state process $U_t^\mu$ and the value function . . . . .	106
II.D.1	Regularity of the regenerative process $U_t^\mu$ . . . . .	106
II.D.2	Regularity of the value function . . . . .	108
II.D.3	Execution time inequalities . . . . .	108
II.D.4	Proof of Theorem II.D.1 . . . . .	109
II.D.5	Regularity of the value function . . . . .	112
II.D.6	Proofs of Propositions II.D.2 and II.D.3 . . . . .	113
II.E	Resolution of the optimal control problem . . . . .	114
II.E.1	Proof of Theorem 1 . . . . .	114
II.E.2	Proof of Theorem 2 . . . . .	115
II.E.3	Optimal strategy . . . . .	116
II.F	Proof of Theorem 3 . . . . .	117
II.F.1	Proof of Inequality (17) . . . . .	117
II.F.2	Proof of the stationary convergence of $\mu^{Opti,\Delta}$ . . . . .	117
<b>III</b>	<b>Optimal inventory management and order book modeling</b>	<b>119</b>
1	Introduction . . . . .	119
2	General order book presentation and priors of the market participants . . . . .	122
3	Market maker's optimal control problem . . . . .	125
3.1	Market maker's strategy and the state dynamics . . . . .	125
3.2	The optimal control problem . . . . .	127
3.3	The dynamic programming equation . . . . .	129
3.4	Dimension reduction, symmetries and numerical resolution . . . . .	130
3.5	Approximate optimal controls . . . . .	132
3.6	Numerical experiments . . . . .	133
4	High frequency pair trading problem . . . . .	138
4.1	The optimal control problem . . . . .	138
4.2	The dynamic programming equation . . . . .	140
4.3	Dimension reduction, symmetries and numerical resolution . . . . .	141
4.4	Approximate optimal controls . . . . .	142
4.5	Numerical experiments . . . . .	143
5	Institutional broker strategies for portfolio liquidation . . . . .	146
5.1	Volume strategy . . . . .	146
5.2	VWAP . . . . .	148
6	Simulation of the full market . . . . .	153
7	Conclusion . . . . .	154
<b>IV</b>	<b>From asymptotic properties of general point processes to the ranking of financial agents</b>	<b>161</b>
1	Introduction . . . . .	161
2	Market modelling . . . . .	164
2.1	Introduction to the model . . . . .	164
2.2	Order book dynamics . . . . .	165

2.3	Market reconstitution . . . . .	168
2.4	Some specific models . . . . .	168
3	Ergodicity . . . . .	170
3.1	Notations and definitions . . . . .	170
3.2	Ergodicity . . . . .	170
4	Limit theorems . . . . .	174
5	Formulas . . . . .	175
5.1	Stationary probability computation . . . . .	175
5.2	Spread computation . . . . .	177
5.3	Price volatility computation . . . . .	177
5.4	An alternative measure of market stability . . . . .	178
6	Numerical experiments . . . . .	179
6.1	Database description. . . . .	179
6.2	Computation of the intensities and the stationary measure . . . . .	180
6.3	Ranking of the market makers . . . . .	181
IV.A	Market reconstitution . . . . .	183
IV.B	Proof of Remark 3.4 . . . . .	184
IV.C	Proof of Theorem 3.1 . . . . .	185
IV.C.1	Preliminary results . . . . .	185
IV.C.2	Outline of the proof . . . . .	187
IV.C.3	Proof . . . . .	187
IV.D	Proof of Theorem 3.2 . . . . .	191
IV.D.1	Preliminary result . . . . .	191
IV.D.2	Uniqueness . . . . .	193
IV.D.3	Speed of convergence . . . . .	202
IV.E	Proof of Propositions 4.1 and 4.2 . . . . .	203
IV.F	Stationary distribution computation . . . . .	204
IV.G	Proof of Proposition 5.2 . . . . .	205
IV.H	Proof of Remark 6.1 . . . . .	206
IV.I	Supplementary numerical results . . . . .	207

<b>V</b>	<b>Improving reinforcement learning algorithms: towards optimal learning rate policies</b>	<b>213</b>
1	Introduction . . . . .	213
2	Reinforcement learning . . . . .	215
3	Improvement of the asymptotic convergence rate . . . . .	217
3.1	Error decomposition . . . . .	217
3.2	Convergence rate of the estimation error . . . . .	218
3.3	Convergence rate of the optimization error . . . . .	219
3.4	Conclusion . . . . .	220
4	Optimal policy for the learning rate $\gamma$ when is countable . . . . .	220
4.1	The inner level . . . . .	221
4.2	The upper level . . . . .	225

4.3	Extension . . . . .	226
5	Some examples . . . . .	227
5.1	Methodology . . . . .	227
5.2	Drift estimation . . . . .	227
5.3	Optimal placement of a limit order . . . . .	228
5.4	Optimal execution . . . . .	230
V.A	Extension of the PASS Algorithm . . . . .	232
V.B	Implementations . . . . .	233
V.C	Additional results: optimal execution . . . . .	233
V.D	Proof of Proposition 3.1 . . . . .	237
V.E	Proof of Proposition 3.2 . . . . .	237
V.F	Proof of Proposition 4.1 . . . . .	237
V.G	Proof of Theorem 4.1 . . . . .	241
V.G.1	Preparation for the proof of Theorem 4.1 . . . . .	241
V.G.2	Propagation of the error . . . . .	244
V.G.3	Proof of Theorem 4.1 . . . . .	246
V.H	Proof of Proposition 4.2 . . . . .	248
V.I	Proof of the convergence of $(v_n^k)_{n \geq 1, k \in (\mathbb{N}^*)^2}$ . . . . .	249
<b>VI</b>	<b>Price impact study on the electricity intra-day market</b>	<b>253</b>
1	Introduction . . . . .	253
2	Descriptive analysis of the intra-day electricity market . . . . .	255
2.1	Description of the intra-day electricity market . . . . .	255
2.2	Description of the database . . . . .	256
2.3	Features of the intra-day electricity market . . . . .	256
3	Estimation methodology and definitions . . . . .	258
3.1	Estimation methodology . . . . .	259
3.2	Price impact profiles . . . . .	259
4	Price study . . . . .	261
4.1	Trade signs autocorrelation . . . . .	261
4.2	Response function autocorrelation . . . . .	262
4.3	Price diffusion . . . . .	263
5	Price modelling . . . . .	264
5.1	Description of the model . . . . .	265
5.2	Properties of the model . . . . .	265
6	Conclusion . . . . .	267
	<b>Bibliography</b>	<b>269</b>

---

# Introduction

---

The goal of this thesis is to study some control problems faced by market participants and to understand the interactions between financial processes such as order flows, liquidity, prices and volatility at different time scales (i.e. starting from the microsecond accuracy to several years). We aim at simultaneously proposing useful models and building estimation procedures for them. These models have to reproduce the most important features of the markets and enable us to design helpful tools for financial agents and regulators. Let us begin with presenting and motivating the different questions on which we want to shed some light.

## Motivations

In general, market participants can take three elementary actions to buy a given number of shares. First, they can send a market order to get immediate execution. However they have to pay a trading cost which corresponds to the bid-ask spread. Second, they can insert a limit order to avoid crossing the spread. In that case, they delay the execution. Third, they can cancel an already inserted limit order hoping to get a better acquisition price in the future. Nevertheless they postpone even more the execution. We choose to focus, as a first step, on the optimal placement problem of a unit buy order using only cancellations and insertions. A market order can only be sent at the terminal time if needed. In such case, the agents wonder how to find the right balance between fast execution and avoiding adverse selection: if the price has high chances to go down, the probability to be filled is high but it may be better to wait a little more to get a better price. Additionally, it is well known that the state of the order book, in particular the imbalance, is a key factor in the prediction of the short-term price move which is closely related to the adverse selection effect. Thus, we might ask ourselves the following question:

**Question 1.** *How to take advantage of our understanding of order flows to optimise elementary decisions in execution algorithms by reducing the adverse selection risk?*

Addressing Question 1, we highlight the disparities between the different classes of agents and illustrate how their actions depend on the imbalance. We also show that taking into account the order book dynamics, mainly through the imbalance, allows us to mitigate adverse selection effects, which improves the performance of trading strategies. Nonetheless, the added value of exploiting this order book liquidity is eroded by latency. Indeed, being able to predict

future price moves is of less use if one does not have enough time to cancel and reinsert limit orders.

In this first study, we consider a simple situation where the agent handles a unit quantity and can send a market order only at the end of the period. This framework is simple and enables us to carry on easy computations. However, it is natural to extend our model to more general optimal execution issues where agents use market orders, buy non-unit quantities and split their orders to reduce their price impact. In addition to that, we want to replace the linear approximation of the price impact used to answer Question 1, by an exact formula. Note that we choose to minimize the price impact instead of the acquisition price since it is the crucial variable for agents such as brokers who interact thousands of times per day with the order book. This pushes us to consider the following question:

**Question 2.** *How to construct, within an order book model, tractable optimal execution algorithms that reduce price impact?*

When answering this question, we propose an order book model inspired from the queue-reactive approach and prove its convergence towards a stationary regime. In this framework, we provide a closed-form formula for the endogenous price impact which allows us to compute the sequence of orders solving the execution problem of the agent. In addition, we show that our optimal tactics outperform significantly standard execution strategies.

After confirming the relevance of our approach for the optimal execution problem, we wonder how to extend our methodology to the dual problem of high frequency market making. The market makers (MMs) provide liquidity by proposing bid and ask prices and make money out of the difference between these two prices (i.e. the bid-ask spread). They use mainly limit orders and should somehow send aggressive orders only for inventory management purpose. After that we turn to the trading issue of a directional high frequency trader (DHFT). The DHFTs are agents that use their low latency technology to catch temporary arbitrage opportunities. The resolution of these control problems allows us to consider a market with three classes of agents, namely institutional brokers (IBs), MMs and DHFTs who interact with each other's in an order book driven exchange. This constitutes a first building block towards a better understanding of the interactions between the different market participants and the order book. Therefore, we raise the following question:

**Question 3.** *How to exploit our previous methodology for market making and DHFTs problems and use it to reproduce a realistic agent-based order book market?*

As an answer to Question 3, we propose reformulations of our model that are consistent with actual MMs and DHFTs issues and solve them numerically. This enables us to exhibit sophisticated optimal strategies that coincide with typical MMs and DHFTs behaviours. Then, we build a realistic market simulator where IBs, MMs and DHFTs interact optimally with each other's. This market simulator can be used to estimate market statistics for any given distribution of the agents and to back-test market participants strategies.

Although we consider a realistic agent-based model enabling us to forecast quantities of interest, the preceding approach still have some limitations. First, agents decisions depend on the market history only through the last order book state while MMs and DHFTs strategies may be more intricate. Second, agents are aggregated into three groups: IBs, MMs and DHFTs. Consequently, it is difficult to disentangle disparities among behaviours of market participants belonging a priori to the same group. This is of paramount importance for regulators who aim at understanding how each participant affects the market. To capture entirely such effect, we want to propose a more general model and to build a simple estimation procedure that quantifies the impact of each agent on market quality. This leads us to the following question:

**Question 4.** *How to build a tractable path dependent model accounting for the interactions between market participants strategies and the order book and use it to assess the contribution of each agent to market quality?*

Investigating Question 4, we design a general agent-based order book model where the order flows interact with each other through the order book state and the history of the flows. We also prove new ergodic and limit theorems for this framework that enable us to build an estimation methodology for the agents trading flows (insertion, cancellation and aggressive orders) and to derive a mathematical link between the individual behaviour of each participant and the market quality, measured for example through the volatility. Thanks to this relation, we propose a ranking methodology for high frequency market makers in term of their role in the stability of the market.

To answer Questions 1, 2 and 3, it is necessary to solve control problems. One can develop numerical schemes to solve them when the parameters governing the dynamics of the model are known or the dimension is not too large. However, when the dynamics of the model are unspecified, this approach is no longer valid. To overcome this difficulty, it is common to use stochastic iterative algorithms. These algorithms appear in many applications that involve finding zeros of a partially known function or minimising an average cost. Furthermore they require mild convergence conditions: no convexity requirements, weak regularity and partial knowledge of the cost function. In such algorithms the step size  $\gamma_k$  is crucial for the convergence. This quantity represents the maximum amount to move along a given search direction at step  $k$ . Hence, we consider the following question:

**Question 5.** *How to dynamically select the step size  $(\gamma_k)_{k \geq 0}$  to improve the convergence of stochastic iterative algorithms?*

The tools developed in the answers of the previous questions are designed for liquid markets using an order book mechanism. However it is not clear whether our approaches are suitable for markets with very specific operating rules. To tackle this issue, we propose, as a first step, to study the price behaviour of the intra-day electricity market. This market is unique since agents trade forwards contracts with a lifetime (i.e. the duration between the issue



and delivery date) smaller than one day. We aim at conducting a global study of the price variations and trading activity to determine the most relevant features of such market and build a toy model for them. Hence, the last question addressed in this thesis is the following:

**Question 6.** *What are the main properties of the price in the electricity intra-day market and how to build a model that reproduces them?*

## Outline

Each chapter of this thesis addresses one of the questions above.

In Chapter I, we answer Question 1 by studying partially labelled trades data to analyse how the decisions of each class of market participants depend on liquidity imbalance. This allows us to formulate a control problem where agents aim at buying a unit quantity of an asset and control limit orders by exploiting order book liquidity, in order to reduce adverse selection (i.e. the execution of a buy limit order followed by a price decrease). This is of great importance for the optimal placement issue of limit orders where the key point is to find the right trade-off between fast execution and avoiding adverse selection. Next, we solve our control problem and show how the obtained strategies can improve the performance of trading algorithms. Finally, we highlight that the outcome of our liquidity-based strategies is harmed by latency: the ability to predict future liquidity-consuming flows is deteriorated by the incapacity to react fast enough to order book moves. This underlines the importance of speed advantage in the reduction of adverse selection.

In Chapter II, we tackle Question 2 by first conducting an empirical study that details how market orders, limit orders, limit orders within the spread and cancellations depend on order book liquidity. This study is different from the one of Chapter I since it focuses on the order type instead of the market participant class. Based on these observations, we propose an order book model inspired from the queue-reactive dynamics, in reduced dimension, with state-dependent regeneration and non-constant spread, and show its ergodicity. Within this framework, we provide a closed-form formula for our endogenous price impact. This allows us to solve the execution problem of an agent who wants to minimize its impact and derive the associated optimal strategy. In addition, we show that our optimal tactic enables us to outperform significantly standard execution strategies.

Answers to Question 3 lie in Chapter III. We model the behaviour of three classes of agents namely institutional brokers (IBs), market makers (MMs) and directional high frequency traders (DHFTs) using an order book model. On the one hand, the MMs and DHFTs determine their optimal strategy by maximising the expected utility of their terminal wealth within a framework inspired by the one of Chapter II. In this study, the main difference between DHFTs and MMs strategies is that DHFTs exploit the correlation between two stock prices to generate profit. On the other hand, IBs have a pre-scheduled task to buy or sell many shares of the considered asset. Then, we derive the equations satisfied by the value

---

function of each agent and explain how to deduce the optimal control strategies from them. Finally, we provide an illustration of the interactions that can take place between these different market participants by simulating an order book model in which each of them applies his own (optimal) strategy.

We address Question 4 in Chapter IV. For this, we restrict ourselves to the modelling of the best bid and ask limits. We propose an approach based on the individual behaviours of market participants modelled by non-linear and state-dependent Hawkes like processes. Our model encompasses the well-known Poisson, Queue-reactive and Hawkes Queue-reactive dynamics for order books. Under mild assumptions, we prove the ergodicity and diffusivity of our model. We also derive semi-closed formulas for the spread, imbalance and market volatility in terms of the intensities of the flows of the different market participants. Thanks to these results, we are for example able to rank market makers according to their contribution to volatility. This ranking is illustrated on several CAC 40 assets. Interestingly the obtained rankings are quite homogeneous from one asset to another.

Question 5 is answered in Chapter V. This chapter investigates to what extent one can improve reinforcement learning (RL) algorithms. Our study is split into three parts. First, we show that the classical asymptotic convergence rate  $O(1/\sqrt{N})$  is pessimistic and can often be replaced by  $O((\log(N)/N)^\beta)$  with  $\frac{1}{2} \leq \beta \leq 1$  and  $N$  the number of iterations. Second, we propose a dynamic optimal policy for the choice of the learning rate  $(\gamma_k)_{k \geq 0}$  used in stochastic approximation. We decompose our policy into two interacting levels: the inner and the outer level. In the inner level, the algorithm PASS operates. This algorithm builds a new sequence  $(\gamma_k^i)_{k \geq 0}$ , based on a predefined sequence  $(\gamma_k^o)_{k \geq 0}$ , whose error decreases faster than the predefined one. In the outer level, we propose an optimal methodology for the selection of the predefined sequence  $(\gamma_k^o)_{k \geq 0}$ . These two levels are interacting in the sense that the algorithm PASS influences the construction of the sequence  $(\gamma_k^o)_{k \geq 0}$ . Third, we show numerically that our selection methodology of the learning rate outperforms significantly standard RL algorithms in the three following applications: the estimation of a drift, the optimal placement of limit orders and the optimal execution of a large number of shares.

Question 6 is discussed in Chapter VI. To do so, we investigate the price behaviour in the specific intra-day electricity market. This market is non-standard since it has unique operating rules. We first propose a descriptive analysis of the price fluctuations and exchanged volumes in such market. Then, we investigate the price impact generated by a power plant breakdown, study some price properties and finally propose a model that reproduces the different observed features. This work constitutes a first step towards the conception of optimised trading strategies in the intra-day electricity market.

# 1 Chapter I - Limit order strategic placement with adverse selection risk and the role of latency

In Chapter I, we answer Question 1: How to incorporate the dynamics of the order book in trading algorithms to reduce the adverse selection risk? We start our study by analysing the dependence between elementary decisions of the various types of agents and order book liquidity. To do so, we use partially labelled trade data that is a direct feed from NASDAQ-OMX.<sup>1</sup> Based on our observations, we propose a stochastic control framework where, in the process of buying a unit quantity, agents optimise limit orders by exploiting liquidity imbalance, to reduce adverse selection. For limit orders, we need optimal strategies essentially to identify situations when it is more suitable to stay in the order book and keep its queue priority (i.e. accelerates the execution) than to cancel its order and wait for better conditions to reinsert it. Thus, we use cancellations mainly to avoid adverse selection: having a buy limit order filled just before a price decrease. Then, we show that latency reduces the efficiency of using liquidity imbalance: it is less interesting to predict future price moves when the agent is not fast enough to catch temporary profit opportunities that he may detect.

## 1.1 Empirical evidences

We first analyse our data to emphasize the predictive power of the imbalance and then illustrate how agent's decisions depend on order book liquidity.

**Predictive power of the imbalance.** The imbalance  $\text{Imb}_t$  at time  $t$  is the easiest way to summarize the state of the order book. It is defined as follows:

$$\text{Imb}_t := \epsilon_t \frac{Q_t^{\text{Bid}} - Q_t^{\text{Ask}}}{Q_t^{\text{Bid}} + Q_t^{\text{Ask}}}, \quad (1)$$

with  $Q_t^{\text{Bid}}$  (resp.  $Q_t^{\text{Ask}}$ ) the available quantity at the best bid (resp. ask) and  $\epsilon_t$  the event sign (i.e.  $\epsilon_t = 1$  when it is a buy order and  $\epsilon_t = -1$  otherwise). Figure .1.a shows the imbalance on the  $x$ -axis and the average mid price move after 50 trades on the  $y$ -axis. We see that the imbalance is highly positively correlated to the price move after 50 trades. Figure .1.b shows the distribution of imbalance just before a change in the order book state. We remark that agents are highly active at extreme imbalance values. This is because they identify a profit opportunity to catch or on the contrary an adverse selection effect to avoid.

---

<sup>1</sup>NASDAQ-OMX is the primary market in the MiFID sense on the considered stock, see [145] for details

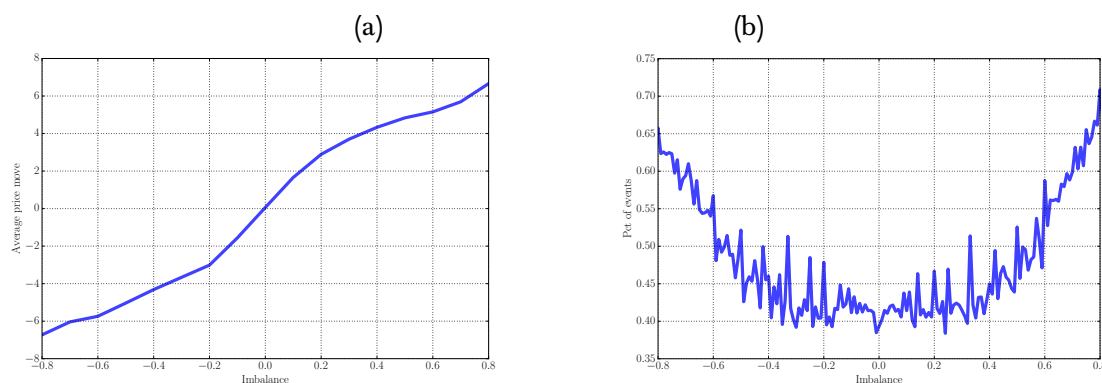


Figure .1 - (a) The predictive power of imbalance on stock Astra Zeneca: Imbalance (just before a trade) on the  $x$ -axis and the expected price move (during the next 50 trades) on the  $y$ -axis. (b) Distribution of the imbalance just before a trade from 2013-01-02 to 2013-09-30 (accounts for 376,672 trades).

**Agent’s decisions depend on order book liquidity.** On the one hand, Figure .2.a shows the average state of the imbalance for four classes of agents when they “accept” to transact via a limit order. On the other hand, we compute in Figure .2.b the average midprice move immediately before and after a class of participant accepts to transact via a limit order. In these two graphs, one can clearly see that the state of the imbalance and the price profiles are different across classes:

- Institutional brokers accept transactions when the imbalance is largely negative, i.e. they buy using a limit order when the price is going down as shown in Figure .2.b (and conversely when they sell). They suffer from a large adverse selection: had they wait a little more, the price would have been more competitive. They make this choice because they have to buy/sell fast from risk management reasons of their clients’ orders.
- High Frequency participants (here high frequency market makers and high frequency proprietary traders) trades via limit orders correspond to the less negative imbalance. They are the most opportunistic agents: look more at the order book state, have enough speed advantage to benefit from this information and their decisions are less driven by urgency considerations. Figure .2.b indeed shows that the price goes up before they buy with limit orders. This implies that they insert limit orders shortly before the trade.
- Global Investment banks are in between. Three reasons may explain such behaviour: first, their activity is a mix of client execution and proprietary trading; second, they have specific strategies to accept transactions via limit orders; third, they invest a little less than high frequency participants in low latency technology, but more than institutional brokers.

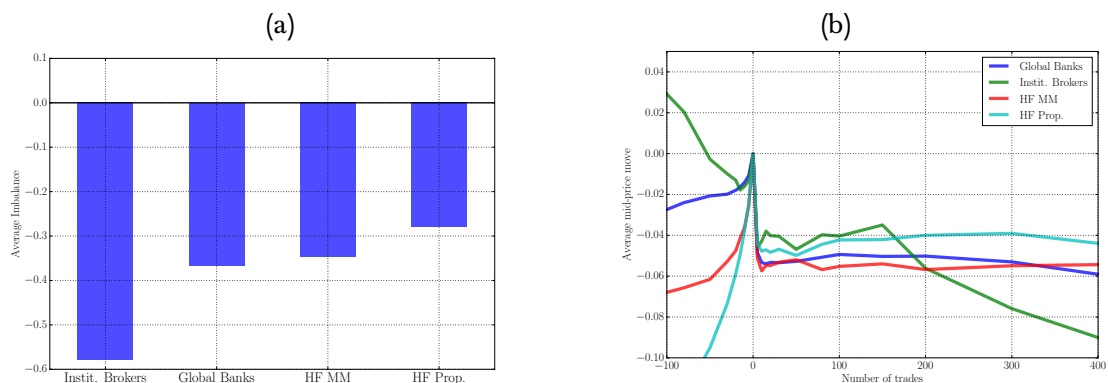


Figure .2 – (a) Neutralized order book imbalance and (b) Average mid price move when a limit order is executed for institutional brokers, global investment banks, high frequency market makers and high frequency proprietary traders. Data are from AstraZeneca (2013-01 to 2013-09).

## 1.2 Problem formulation

We place ourselves in the case of an agent who aims at buying a unit size quantity  $q$ . The order book state is modelled by  $U_n^\mu = (Q_n^{Before,\mu}, Q_n^{After,\mu}, Q_n^{Opp,\mu}, P_n^\mu)$  where  $Q_n^{Before,\mu}$  is the quantity having priority on the order  $q$ ,  $Q_n^{After,\mu}$  is the quantity posted after the order  $q$ ,  $Q_n^{Opp,\mu}$  is the first opposite limit quantity,  $P_n^\mu$  is the mid price,  $\mu$  is the control of the agent and the integer  $n$  represents the event time, see Figure .3. The agent uses two possible controls:

- $s$  (like *stay*): stay in the order book and keep its strategic placement,
- $c$  (like *cancel*): cancel the order and wait for a better order book state to reinsert it at the top of the best bid queue, see Figure .3,

to find the right balance between fast execution (i.e. control  $s$ ) and avoiding adverse selection (i.e. control  $c$ ).

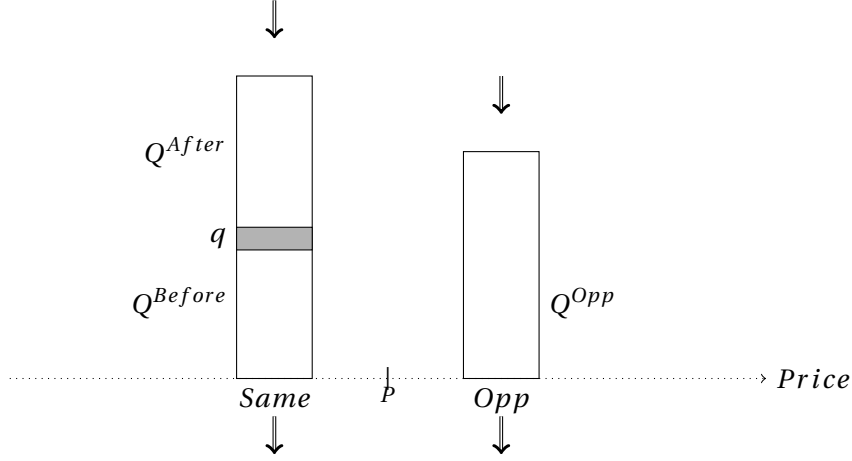


Figure .3 - Diagram of flows affecting our order book model.

**Optimal control problem.** We fix a finite horizon time  $f < \infty$  and want to compute

$$V_f(0, U) = \sup_{\mu} \mathbb{E}[\Delta P_{\infty}^{\mu}], \quad (2)$$

where:

- $U = (q^{before}, q^{after}, q^{opp}, p)$  is the initial state of the order book.
- $n_{Exec}^{\mu} = \inf\{n \geq 0, s.t. Q_n^{Before, \mu} = 0, \mu_n = s\} \wedge f$  represents the execution time.
- $\Delta P_{\infty}^{\mu} = \lim_{n \rightarrow \infty} (P_n^{\mu} - P_{n_{Exec}^{\mu}}^{Exec, \mu})$  represents the long-term underlying gain of the trader, where the execution price  $P_n^{Exec, \mu}$  satisfy  $P_n^{Exec, \mu} = P_n^{\mu} - \frac{1}{2}$  when the limit order is executed before  $f$  and  $P_n^{Exec, \mu} = P_n^{\mu} + \frac{1}{2}$  otherwise. Indeed, if at  $f$  the order has not been executed, we cross the spread to guarantee execution.

We consider the following approximation:

$$P_{\infty, n} = \mathbb{E}[\lim_{k \rightarrow \infty} P_k^{\mu} | F_n] = P_n + \frac{\alpha}{2} \cdot \frac{Q_n^{Same} - Q_n^{Opp}}{Q_n^{Opp} + Q_n^{Same}}, \quad (3)$$

where  $F_n$  denotes the filtration generated by the order book process  $U_n^{\mu}$  and  $\alpha > 0$  a parameter representing the sensitivity to the imbalance. This approximation is consistent with our empirical observations since the imbalance can predict quite accurately the price future moves, see Figure .1.a. In Chapter II, we present a more rigorous computation of this quantity.

### 1.3 Numerical resolution

**Numerical scheme.** Using Approximation (3), we have

$$\mathbb{E}[\Delta P_\infty^\mu] = \mathbb{E}[Z_{T_{Exec}^\mu}],$$

with  $Z_n = P_{\infty,n} - P_n^{Exec,\mu}$ ,  $P_{\infty,n}$  defined in (3) and  $P_n^{Exec,\mu}$  defined above. We can then compute  $V_f$  using the following result.

**Result 1.** *Using the dynamic programming algorithm we have*

$$\begin{cases} G_f = Z_f, \\ G_n = \max(P^c G_{n+1}, P^s G_{n+1}), \quad \forall n \in \{0, 1, \dots, f-1\}, \end{cases}$$

where  $V_f(0, U_0) = G_0(U_0)$ ,  $P^c$  (resp.  $P^s$ ) is the transition matrix of the Markov chain  $U_n^\mu$  when the initial decision is  $c$  (resp.  $s$ ).

**Methodology.** We assume that the order flow intensities are linear functions of the imbalance and we compare our optimal solution with another standard strategy denoted by (NC)

- (NC) corresponds to the case where no control is adopted (i.e. we always stay in the orderbook and “join the best bid” each time it changes).
- We call (OC) our optimal placement strategy: controls "c" and "s" are considered.

**Relevance of our strategy.** Let  $\mu^c$  (resp.  $\mu^*$ ) be the control under NC (resp. OC). Figure .4.a represents the variation of  $\mathbb{E}_{U_0, \mu^*}[\Delta P_\infty^\mu]$  and  $\mathbb{E}_{U_0, \mu^c}[\Delta P_\infty^\mu]$  when the initial imbalance of the order book varies. Here  $U_0$  refers to the initial order book state and  $\Delta P_\infty^\mu$  is defined in (2). As expected the strategy OC provides better results than NC. A remarkable effect is the way the optimal control anticipates adverse selection. When imbalance is highly negative, OC cancels first to take advantage of a better future opportunity.

**Latency price.** The Markov chain  $U_n^\mu$  corresponds to a market participant enabled to change his control at each period. A slower participant will not react at each limit order book move. Hence, he can be modelled by a Markov chain  $U_{\tau n}^\mu$  where  $\tau \in \mathbb{N}^*$  corresponds to a latency factor. We define  $V_f^\tau(0, U_0)$  in the same way as  $V_f(0, U_0)$ , see (2), but where the underlying order book process  $U_n^\mu$  is replaced by  $U_{\tau n}^\mu$ . In such case, the latency cost of a participant corresponds to

$$Latency_{U_0, f}(\tau) = V_f(0, U_0) - V_f^\tau(0, U_0), \quad \forall \tau \in \mathbb{N}^*.$$

Figure .4.b shows the variation of the latency cost with respect to the latency factor  $\tau$ . We observe that

- The latency cost obviously increases with the latency factor  $\tau$ .

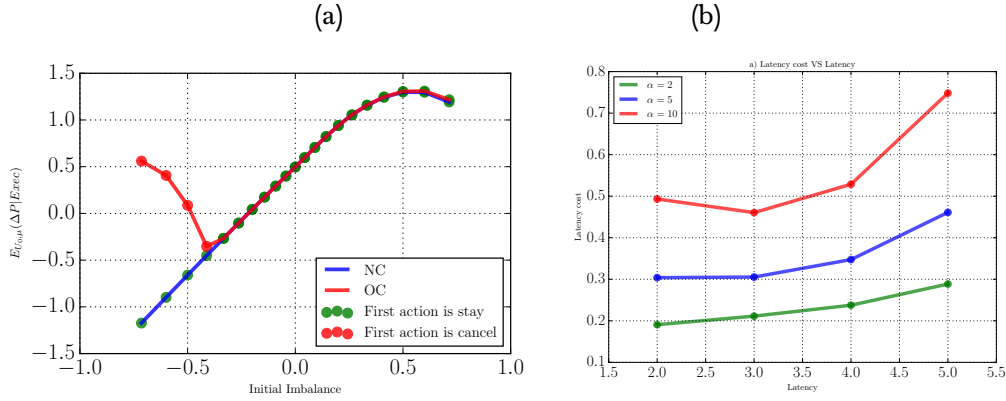


Figure .4 - (a)  $E_{U_0, \mu^*} [\Delta P_\infty^\mu]$  and  $E_{U_0, \mu^c} [\Delta P_\infty^\mu]$  when intensities are linear functions of the imbalance and (b) latency cost as a function of the latency factor  $\tau$  for different values of  $\alpha$ .

- The latency cost is higher when the sensitivity to adverse selection  $\alpha$  (see (3)) increases.

In this chapter, we show the relevance of our approach for the optimal placement problem of unit order. We wonder now how to extend our methodology to more realistic optimal execution contexts where the agent handles non-unit quantities and can send market orders at any time. These modifications generate additional difficulties since we need to take into account the splitting of the orders. In addition to that, in Chapter I, the price impact was approximated by a linear function of the imbalance. We now want to provide a proper formula for the computation of this quantity. Thus, we address in Chapter II the following question: How to extend this model to develop suitable general optimal execution algorithms that aim at reducing the price impact?

## 2 Chapter II - Optimal liquidity-based trading tactics

In this chapter, we tackle Question 2: How to construct, within a realistic order book model, optimal execution algorithms that reduce price impact? We start our study with an empirical analysis that describes how market participants decisions, namely market orders, cancellations, limit orders and limit orders within the spread are related to order book liquidity. Based on these observations, we adapt the approach of Chapter I to propose an order book model with state-dependent regeneration and non-constant spread and we show its ergodicity. We provide a closed-form formula of the endogenous price impact that allows for fast computation. Using this result, we solve the execution problem of a buying agent who aims at minimizing its impact and derive the associated optimal strategy. Finally, numerical results show that our optimal tactic enables us to outperform standard execution strategies.



## 2.1 Order book modelling

### 2.1.1 Empirical evidences

We start our analysis by investigating how market participants act when liquidity varies. To do so, we use data from Bund futures on Eurex exchange Frankfurt. This database records during one week from 1 to 5 September 2014, the state of the order book (i.e. available quantities and prices at best limits) event by event with microsecond accuracy. The database accounts for 3 407 574 events.

Let  $t$  be the current time. Figure .5 shows the average imbalance value for each event type. We give the interpretation of Figure .5 in the case of a buy limit/limit mid<sup>2</sup>/cancellation/market order, since the event sign is taken into account in the expression of the imbalance  $\text{Imb}_t$ , see (1). Note that Figure .2.a differs from Figure .5 since it focuses on the type of the orders instead of the class of the agents. We see that market participants insert limit orders when imbalance is negative (execution highly probable), cancel or send limit orders within the spread when imbalance is positive (less chance to be executed) and use market orders when imbalance is highly positive (rushing for liquidity when it is scarce).

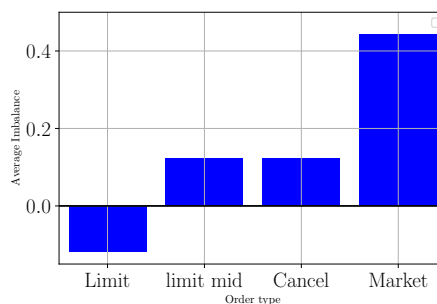


Figure .5 – Average imbalance before Limit/Limit mid/Cancel/Market order.

### 2.1.2 The model

The order book state is modelled by the Markov process  $U_t = (Q_t^1, Q_t^2, S_t) \in \mathbb{U}$  where  $Q_t^1$  (resp.  $Q_t^2$ ) is the best bid (resp. ask) quantity and  $S_t$  is the spread. The dynamics of  $U_t$  are described through a Markov jump process in the spirit of Chapter I. Within this framework, we incorporate a state-dependent regeneration which is so that the new order book state after a depletion depends on the order book state just before the depletion and the depleted side (i.e. best bid/ask). We also enforce the classical bid-ask symmetry relation that ensures no statistical arbitrage and allows us to aggregate bid and ask side data in the calibration of the model's parameters. Finally, we denote by  $P_t^1$  (resp.  $P_t^2$ ) the best bid (resp. ask) price and  $P_t$  the mid price. The mid price variations will be describe later.

<sup>2</sup>A limit mid order is a limit order inserted within the spread.

### 2.1.3 Ergodicity

Under mild conditions, the considered order book process  $U_t$  is ergodic. Indeed, we have the following result.

**Result 2** (Ergodicity). *Under mild assumptions, the process  $U_t$  is ergodic (i.e. converges towards a unique invariant distribution). Additionally, we have the following speed of convergence:*

$$\|P_u^t(\cdot) - \pi\|_{TV} \leq B(u)\rho^t,$$

with  $\|\cdot\|_{TV}$  the total variation distance,  $P_u^t(\cdot)$  the Markov kernel of the process  $U_t$  starting from the initial point  $u \in \mathbb{U}$ ,  $\pi$  the invariant distribution,  $\rho < 1$  and  $B(u)$  a constant depending on the initial state  $u$ .

This theorem is the basis for our asymptotic study of the order book dynamics, since it ensures the convergence of the order book state towards an invariant probability distribution. Thus the stylized facts observed on market data can be explained by a law of large numbers type phenomenon for this invariant distribution.

### 2.1.4 Long-term price move computation

We provide the computation methodology of the long-term price  $D(u)$  starting from the initial state  $u \in \mathbb{U}$  defined as follows:

$$D(u) = \lim_{t \rightarrow \infty} \mathbb{E}_u [P_t - P_0], \quad (4)$$

The computation of  $D(u)$  is interesting for at least two reasons. First, it allows us to predict the long-term average mid price move for any initial order book state  $u$ . Second, the quantity  $D(u)$  is crucial for the computation of the price impact.<sup>3</sup> To compute  $D(u)$  we need to specify the price dynamics.

**Price dynamics.** The mid price after the  $n$ -th order book event  $P_n$  satisfies  $P_n = P_0 + \sum_{i=1}^n \Delta P_i$  with  $\Delta P_i = P_i - P_{i-1}$ . The price jumps  $\Delta P_i$  are a deterministic function of the order book state before the jump and the order book event causing the jump. For example, we can consider the simple case where the mid price decreases (resp. increases) by one tick when the best bid (resp. ask) is depleted. We have the following result.

**Result 3** (Price impact). *Under mild assumptions, the vector  $D$  satisfies*

$$D = (I - A)^{-1}F,$$

where the matrix  $A$  encodes transition probabilities to the different order book states after the first regeneration and the vector  $F$  represents the different average mid price moves after the first regeneration. They are defined in detail in Chapter III. Figure .6 gives a numerical computation of the vector  $D$ .

---

<sup>3</sup>The connection between the computation of the price impact and Equation (4) is detailed in Chapter II.

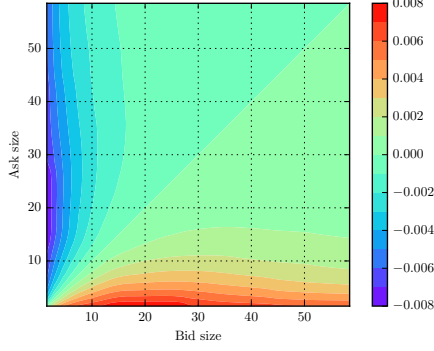


Figure .6 – Price move  $\lim_{t \rightarrow \infty} \mathbb{E}_{U_0}[P_t - P_0]$  in percentage of the tick  $\delta = 1$  cent.  $Q^1$  and  $Q^2$  are divided by the average event size.

## 2.2 Optimal control problem

### 2.2.1 Problem formulation

We express the agent control problem for a buy order of size  $q^0$ . It can be changed to a sell order in an obvious way. The agent state is modelled by the process

$$\bar{U}_t^\mu = \left( Q_t^{Bef,\mu}, Q_t^{a,\mu}, Q_t^{Aft,\mu}, Q_t^{2,\mu}, I_t^\mu, S_t^\mu, P_t^\mu, P_t^{Exec,\mu} \right),$$

where  $Q_t^{a,\mu}$  is the size of agent's limit order inserted at the best bid,  $Q_t^{Bef,\mu}$  is the quantity inserted before  $Q_t^{a,\mu}$ ,  $Q_t^{Aft,\mu}$  represents orders inserted after  $Q_t^{a,\mu}$ ,  $P_t^{Exec,\mu}$  is the acquisition price of  $q^0 - I_t^\mu$ ,  $I_t^\mu$  is the agent's inventory and  $\mu = \{\mu_t, t \leq T\}$  is the control of the agent. We recall that  $Q_t^{2,\mu}$  is the best ask limit,  $P_t^\mu$  is the mid price and  $S_t^\mu$  is the spread. Then,  $Q_t^{1,\mu} = Q_t^{Bef,\mu} + Q_t^{a,\mu} + Q_t^{Aft,\mu}$  is the total volume at the best bid. It is split into three quantities to take into account the order placement. The regeneration of the order book after the total consumption of a given best limit still depends on the depleted side (i.e. ask/bid) and the order book state before a depletion.

At every decision time, the trader can do nothing or take three possible decisions:

- *l*: He can insert a fraction of  $I^\mu$  at the top of the bid queue or within the spread if not already inserted.
- *c*: He can cancel his already existing limit order  $Q^{a,\mu}$ . By acting this way, the trader can wait for a better order book state. This control will essentially be used to avoid adverse selection, i.e. obtaining a transaction just before a price decrease.
- *m*: He can send a market order to get an immediate execution of a fraction of  $I^\mu$ .

Every decision of the agent is also characterized by a price level  $p$  and an order size  $q$ . The price level is equal to  $p \geq 0$  when the order is inserted at the limit price  $P^1 + \min(p, S)$  with  $P^1$  the best bid price and  $S$  the spread.

**Optimal control problem.** We fix a finite horizon time  $T < \infty$  and we want to compute

$$V_T(0, u) = \sup_{\mu} \mathbb{E} \left[ \underbrace{f \left( \lim_{s \rightarrow \infty} \mathbb{E}[\Delta P_s^\mu / \mathcal{F}_{T^{Exec}^\mu}] \right)}_{\text{final constraint}} - \underbrace{\gamma \int_0^{T^{Exec}^\mu} I_s^\mu ds}_{\text{running cost}} \right],$$

where

- $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec})$  is the initial agent state.
- $T_{Exec}^\mu = \inf \{ t \geq 0, s.t. I_t^\mu = 0 \} \wedge T$  represents the final execution time.
- $\Delta P_t^\mu = (q^0 P_t^\mu - P_{T_{Exec}^\mu}^{Exec, \mu})$  represents the price impact and  $q^0$  is the initial order size.
- $\gamma$  is a non-negative homogenization constant representing the waiting cost or the risk aversion of the agent and  $f: \mathbb{R} \rightarrow \mathbb{R}$  is a Lipschitz function.

We solve the agent's control problem in two situations: when decisions are taken at fixed frequency  $\Delta^{-1}$  and when they can be taken at any time. However, we present only the results for the second situation in this introduction.

### 2.2.2 Decision taken at any time

We define the value function  $V_T$  of the control problem such that

$$V_T(t, u) = \sup_{\mu} \mathbb{E} \left[ f \left( \lim_{s \rightarrow \infty} \mathbb{E}[\Delta P_s^\mu / \mathcal{F}_{T^{Exec}^\mu}] \right) - \gamma \int_t^{T^{Exec}^\mu} I_s^\mu ds \mid \bar{U}_t^\mu = u \right].$$

We provide the system of equations satisfied by  $V_T$  when decisions are taken at any time and show how to approximate it.

**Result 4.** *Let  $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec})$  be an initial state and  $t \in [0, T]$ . Then  $V(t, u)$  satisfies in the viscosity sense and almost everywhere*

- *When  $i > 0$  and  $p < \bar{P}^4$ :*

$$\max \left( \mathcal{A}V(t, \cdot) - \gamma \mathbf{I}, \sup_{e \in \mathbb{E}} V^e(t, \cdot) - V(t, \cdot) \right) = 0, \quad (5)$$

*where  $V(t, \cdot)$  and  $V^e(t, \cdot)$  are vectors such that  $V(t, \cdot)_i = V(t, u_i)^5$  and  $V^e(t, \cdot)_i = \mathbb{E}[V(t, u_i^e)]$  with  $u_i^e$  the new order book state when the decision  $e$  is taken.<sup>6</sup> Equation (5) should be understood coordinate by coordinate.*

<sup>4</sup>Here  $\bar{P}$  is a fixed threshold such that the agents liquidates its inventory for risk management reasons when the mid price gets beyond  $\bar{P}$ .

<sup>5</sup>Here  $u_i$  is the unique state associate to the index  $i$ .

<sup>6</sup>We use the expected value  $\mathbb{E}[V(t, u_i^e)]$  since the control  $e$  may lead to several states because of the regeneration.

- When  $i = 0$  or  $p \geq \bar{P}$  (execution time condition):

$$V(t, u) = \tilde{g}(u), \forall t \leq T,$$

with  $\tilde{g}(u) = f\left(\lim_{t \rightarrow \infty} \mathbb{E}_u[\Delta P_t]\right)$ .

- The terminal condition is

$$V(T, u) = g(u), \tag{6}$$

with  $g(u) = f\left(\mathbb{E}\left[\lim_{t \rightarrow \infty} \mathbb{E}_{u, m^t}[\Delta P_t]\right]\right)$  and the decision  $m^t$  represents the liquidation of the remaining inventory. We keep in mind that the control  $m$  may lead to several states because of the regeneration.

To solve numerically the preceding equations, we consider the following discrete framework.

**Discrete-time Markov chain approximation.** Let  $U_n^\Delta$  be a Markov chain with the transition matrix  $P^{*,\Delta}$  defined by

$$P_{u,u'}^{*,\Delta} = \mathbb{P}[U_\Delta = u' | U_0 = u], \quad \forall (u, u') \in \mathbb{U}^2,$$

with  $U_t$  the process defined in Section 2.1.2.<sup>7</sup> In this approximation,  $U_n^\Delta$  is viewed as the market evolution without the intervention of the agent. Associated to this new market, we introduce the controlled discrete-time Markov chain

$$\tilde{U}_n^{\Delta,\mu} = \left( Q_n^{Bef,\mu}, Q_n^{a,\mu}, Q_n^{Aft,\mu}, Q_n^{2,\mu}, I_n^\mu, S_n^\mu, P_n^\mu, P_n^{Exec,\mu} \right),$$

by using the same construction as in Section 2.2.1.<sup>8</sup> Finally, for every  $k \geq 0$ , we define the piecewise constant process  $\tilde{U}^{\Delta,\mu}$  associated to  $U_n^{\Delta,\mu}$  such that:

$$\tilde{U}_t^{\Delta,\mu} = U_k^{\Delta,\mu}, \quad \forall t \in [k\Delta, (k+1)\Delta),$$

and we denote by  $\tilde{V}^\Delta(t, U)$  the value function of the control problem 2.2.1 where the process  $U^\mu$  is replaced by  $\tilde{U}^{\Delta,\mu}$  and the agents takes its decisions only at times  $k\Delta$  with  $k \in \mathbb{N}$ .

**Solving numerically the optimal control problem in the discrete framework.** We denote by  $V^\Delta(n, u)$  the value function associated to the discrete control problem (i.e. the state process  $U_n^{\Delta,\mu}$ ), with  $n$  the period and  $u$  the order book state. The dynamic programming principle provides a numerical scheme to compute backward  $V^\Delta(0, u)$  starting from the terminal condition.<sup>9</sup> Thus, we can estimate  $\tilde{V}^\Delta(n\Delta, u)$  since  $\tilde{V}^\Delta(n\Delta, u) = V^\Delta(n, u)$ .

Let  $\mu^{Opti,\Delta}$  (resp.  $\mu^{Opti}$ ) be the optimal control associated to the process  $\tilde{U}_t^\mu$  (resp.  $U_t^\mu$ ). Then we have the following error estimate result.

<sup>7</sup>Given the infinitesimal generator  $Q$  of  $U$ , the transition matrix  $P^{*,\Delta}$  can be easily computed since  $P^{*,\Delta} = e^{\Delta Q}$ .

<sup>8</sup>In addition to that, the price impact  $PI_\infty^\Delta$  in this discrete-time approximation can be computed by following the approach of Section 2.1.4

<sup>9</sup>To compute efficiently the value function, the dynamic programming scheme can be parallelized.

**Result 5.** *Under mild assumptions,  $\tilde{V}^\Delta$  converges pointwise towards  $V$ . Additionally, we have the following error estimate:*

$$|\tilde{V}^\Delta(t, u) - V(t, u)| \leq R(T - t)\Delta, \quad \forall(t, u), \quad (7)$$

*with  $R > 0$  a positive constant. Moreover, the control  $\mu^{Opti, \Delta}$  converges towards  $\mu^{Opti}$  and for  $\Delta$  small enough we have the equality  $\mu^{Opti, \Delta} = \mu^{Opti}$ , a.s.*

The above result provides an error bound for the approximation  $\tilde{V}^\Delta(t, u)$  and ensures the convergence of the controls.

### 2.3 Numerical experiments

**Decisions taken at any time.** Figure .7 shows the value function at time zero (i.e. average gain of the optimal strategy) in red and the one of the strategy stay at the best bid in blue in percentage of the tick  $\delta = 1$  cent using the discrete approximation. When imbalance is highly negative, it is optimal to cancel the order to avoid adverse selection, when imbalance is highly positive it is optimal to send a market order or stay in the order book. In our case, stay in the order book is interesting when imbalance is highly positive since the priority value is important. When the imbalance is close to 0, it is optimal to send a market order to reduce the waiting cost. We note that the profit and loss of the optimal strategy is significantly higher than the one of the strategy stay at the bid.

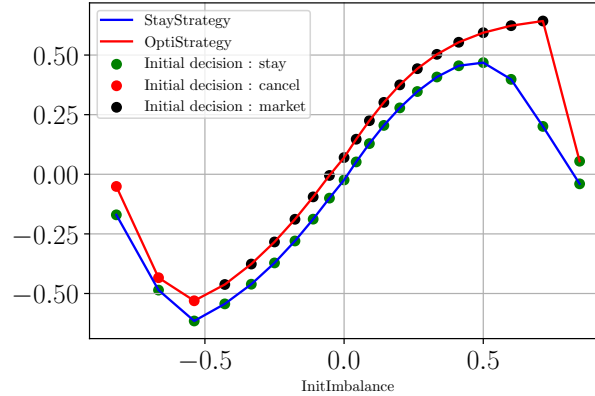


Figure .7 – The gain per tick of the optimal strategy in red and the one of the strategy join the bid in blue for different values of the initial imbalance. Initial imbalances are obtained with  $Q^{Bef} = 0$ ,  $Q^1 = 11$  and  $Q^2$  from 1 to 11. Then, we fix  $Q^2 = 11$  and vary  $Q^1$  from 10 to 1. Initial parameters are as follows: the time step is equal to  $\Delta = 1$  second, there are 10 periods, arrival rate are constant (i.e.  $\lambda^{1,+} = \lambda^{2,+} = 0.06$ ) consumption rates are constant (i.e.  $\lambda^{1,-} = \lambda^{2,-} = 0.12$ ), the new bid (resp. ask) is set to 5 and the new ask (resp. bid) to 3 after the total depletion of the bid (resp. ask) limit, the quantity  $q^0 = 1$ , the waiting cost  $c = 0.0085$ , the price increases (resp. decreases) by  $\delta = 1$  cent when the ask limit (resp. bid limit) is totally consumed and the function  $f$  is the identity.

We emphasize in this chapter the relevance of our approach for optimal execution issues. We now aim at extending our methodology for market making and high-frequency pair trading strategies. The resolution of these two problems will allow us to propose a market simulator with three classes of agents who interact with each other's in an order book driven exchange. Thus, we consider the following question: How to exploit the previous approach for market making and DHFTs problems and use it to reproduce a realistic agent-based order book driven market?

### 3 Chapter III - Optimal inventory management and order book modeling

Chapter III addresses the following question: how to exploit our previous approach to solve MMs and DHFTs problems and use it to simulate an order book driven exchange? Unlike brokers, MMs do not aim at acquiring or liquidating a given number of shares. They provide liquidity by proposing buy and sell prices and generate profit from the difference between these two prices (i.e. bid-ask spread). To do so, they are supposed to mainly use passive orders. Nonetheless, they can send aggressive orders to manage their inventory risk. In our context, the main distinction between MMs and DHFTs lies in the fact that DHFTs include the correlation between two stock prices in their decision making process. In this study, we adapt our previous approach to formulate and solve the MM and DHFT control problems.

We also use the standard Cartea-Jaimungal framework to solve the control problem of our last market participant, the IB, who needs to buy or sell a given number of shares. Numerical simulations exhibit strategies that coincide with typical MMs, DHFTs and IBs behaviours. After that, we construct a market simulator where IBs, MMs and DHFTs interact optimally with each other's and provide an illustration of the type of features that can emerge.

### 3.1 MM control problem

#### 3.1.1 Problem formulation

The MM state is modelled by the Markov process

$$Z_t^\mu = \left( Q_t^{1,\mu}, Q_t^{2,\mu}, P_t^{1,\mu}, P_t^{2,\mu}, I_t^\mu, G_t^\mu, N_t^{1,\mu}, N_t^{2,\mu}, B_t^{1,\mu}, B_t^{2,\mu}, J_t^\mu \right),$$

where  $Q_t^{1,\mu}$  (resp.  $Q_t^{2,\mu}$ ) is the best bid (resp. ask) quantity,  $P_t^{1,\mu}$  (resp.  $P_t^{2,\mu}$ ) is the best bid (resp. ask) price,  $I_t^\mu$  is the MM inventory,  $G_t^\mu$  is the MM wealth,  $N_t^{1,\mu}$  (resp.  $N_t^{2,\mu}$ ) is the size of agent's limit order inserted at the best bid (resp. ask),  $B_t^{1,\mu}$  (resp.  $B_t^{2,\mu}$ ) is the quantity inserted before the agent's limit order at the best bid (resp. ask) and  $J_t^\mu$  is the number of decisions taken by the MM.<sup>10</sup>

At any time, the MM can take the same type of actions as presented in the control problem of Chapter II. The aim of the MM is to maximise the expected utility

$$-\mathbb{E}[\exp(-\eta\{U(Z_T^\mu) - \rho j\})],$$

with

$$U(z) := g + i^+ p^1 - i^- p^2 - \kappa([i^+ - q^1]^+ + [i^- - q^2]^+),$$

for any  $z = (p^1, p^2, q^1, q^2, g, i, n^1, n^2, b^1, b^2, j) \in D_Z$  with  $D_Z$  the set of admissible values for  $z$ .<sup>11</sup> Here,  $r^+$  and  $r^-$  respectively refer to the positive and negative part of any real number  $r$ . In the above,  $\eta > 0$  is the absolute risk aversion parameter, and  $\kappa > 0$  is a penalty term. The quantity  $U(z)$  represents the Mark-to-Market value of the MM portfolio. It is the sum of two components the current wealth of the agent  $g$  and the liquidation value of its inventory  $i^+ p^1 - i^- p^2 - \kappa([i^+ - q^1]^+ + [i^- - q^2]^+)$ . We incorporate a liquidity premium through the parameter  $\kappa$ : the agent pays higher transaction costs when he executes quantities that exceed the best limits. We also add the coefficient  $\rho \geq 0$  to penalize the number of actions taken by the MM. This is the analogous of the inventory risk term in the classical Cartea-Jaimungal framework, see [52].

#### 3.1.2 Problem resolution

We provide the equations satisfied by the MM value function  $v$  and the numerical scheme used to approximate the solution. We define  $v$  such that

$$v(t, z) = \sup_{\mu} \mathbb{E}[U(Z_T^\mu) | Z_t^\mu = z],$$

<sup>10</sup>We record this quantity since it may induce a cost.

<sup>11</sup>The set  $D_Z$  is described more in detail in Chapter III



for any  $(t, z) \in [0, T] \times D_Z$  and we denote by  $v_*$  and  $v^*$  respectively the lower-semicontinuous and upper-semicontinuous envelopes of  $v$ . We consider the following notations:

$$\mathcal{I}\varphi(t, z) := \int (\mathcal{K}^e \varphi(t, z) - \varphi(t, z)) d\beta(e|p, q) \quad \text{and} \quad \mathcal{K}\varphi(t, z) := \sup_{c \in \mathbf{C}(z)} \mathcal{K}^c \varphi(t, z),$$

for any  $z = (p^1, p^2, q^1, q^2, g, i, n^1, n^2, b^1, b^2, j) \in D_Z$  where  $e \in \mathbf{C}(z)$  is an event triggered by other market participants,  $p = (p^1, p^2)$  are the prices at the best quotes,  $q = (q^1, q^2)$  are the volumes at the best limits,  $\beta(\cdot|p, q)$  is a non-negative kernel representing the instantaneous arrival rate of an event  $e$  conditional to the current state  $(p, q)$ ,  $c$  represents the MM action,  $\mathbf{C}(z)$  is the set of admissible actions and

$$\mathcal{K}^c \varphi(t, z) := \mathbb{E}[\varphi(t, z_t^c)], \tag{8}$$

with  $z_t^c$  the new agent state when he applies the control  $c$  to the current state  $Z_t = z$  at time  $t$ .<sup>12</sup> Thus, the partial differential equation characterizing  $v$  is given by

$$\begin{aligned} \min\{-\partial_t \varphi - \mathcal{I}\varphi, \varphi - \mathcal{K}\varphi\} &= 0, \quad \text{on } [0, T] \times D_Z, \\ \min\{\varphi - U, \varphi - \mathcal{K}\varphi\} &= 0, \quad \text{on } \{T\} \times D_Z. \end{aligned} \tag{9}$$

We show the following result.

**Result 6.** *Under mild assumptions, the function  $v_*$  (resp.  $v^*$ ) is a viscosity supersolution (resp. subsolution) of (9). Moreover, the value function  $v$  is the unique continuous viscosity solution of (9) in the class of (discontinuous) solutions that belongs to  $\mathcal{L}_\infty^{\text{exp}}$ .<sup>13</sup>*

We move now to the numerical scheme for (9). For this, we introduce a classical fully explicit finite difference scheme  $(v_n^k)_{k, n \geq 1}$ , with  $n$  and  $k$  are respectively time and space discretization parameters, and show its convergence.

**Result 7.** *Under mild assumptions, the sequence  $(v_n^k)_{k, n \geq 1}$  converges pointwise to  $v$  on  $[0, T] \times D_Z$  as  $k, n \rightarrow \infty$ .*

We also prove that the sequence of controls derived in a standard way from the numerical scheme is asymptotically optimal.

### 3.1.3 Numerical experiment

Figure .8 shows a simulated trajectory for the optimal strategy starting from a symmetric configuration of the order book with queue lengths equal to 6.<sup>14</sup> For this path, Figure .8 records the evolution of the MM controls, the order book state, the agent's inventory and his wealth during a trading period  $T = 59$  seconds. The time step here is 1 second. The MM's decisions are mainly driven by inventory management constraints and predictions of the next

---

<sup>12</sup>Note that we use an expectation in (8) since an action may lead to different states.

<sup>13</sup>The set of functions  $\mathcal{L}_\infty^{\text{exp}}$  is defined in Chapter III

<sup>14</sup>Our unit here is 1/2 of the ATS (i.e. mediAn Trade Size).

price moves through the imbalance variable. For example, the MM first sends limit orders of size 3 on both sides of the order book. At time 8, an aggressive buy order consumes 2 of the 3 limit orders that he posted at the best bid. He decides then to cancel the last remaining buy limit order to avoid increasing his inventory. To assess the performance of our strategy, we provide in Figure .9 the distribution of the gain of the MM for  $10^5$  simulated paths. We can see the median gain of the MM is around 0.6 which represents 6 times the tick.

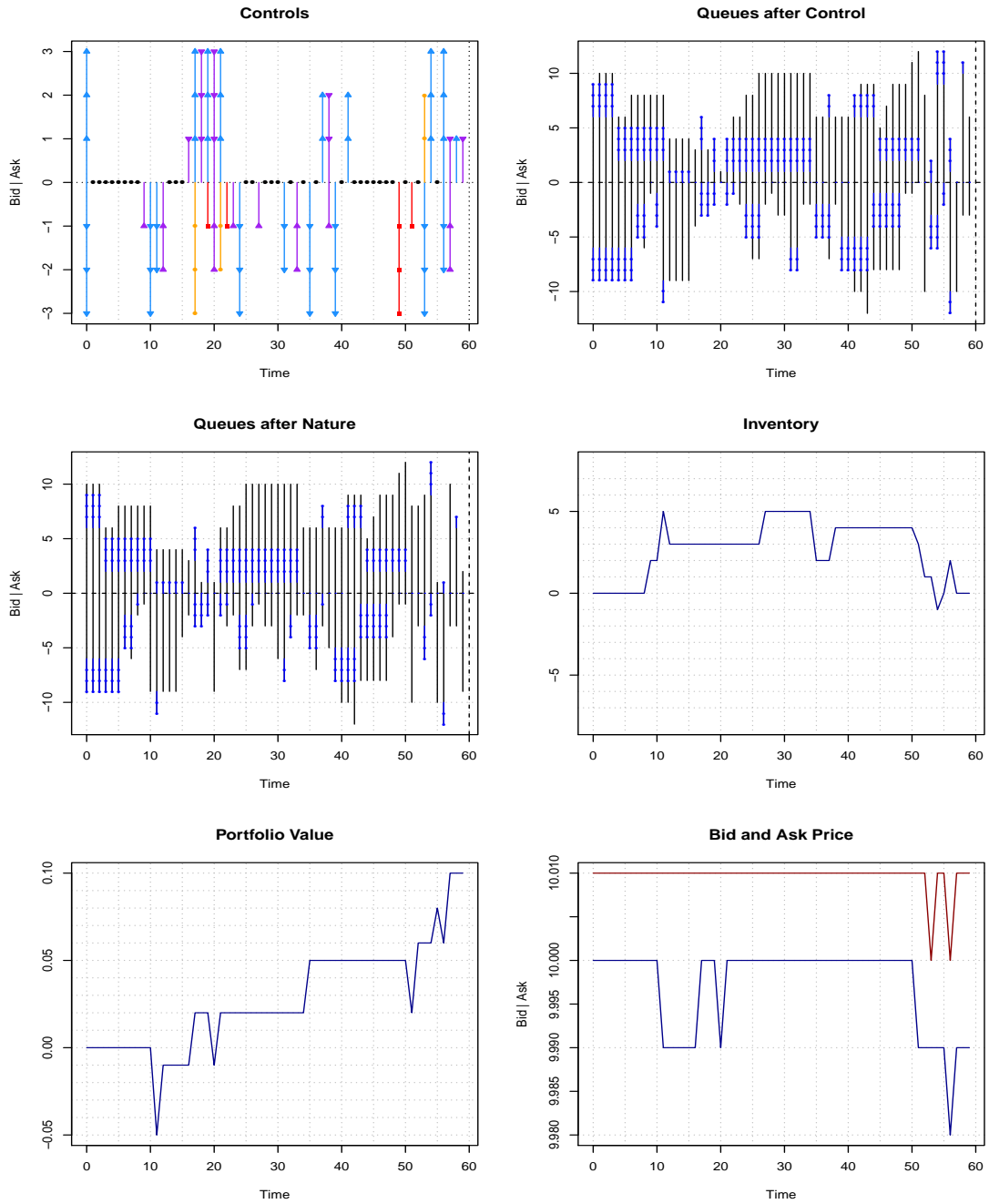


Figure .8 - Simulated path of the optimal strategy of the Market Maker. In the top left panel, the blue color refers to limit orders, the purple one corresponds to cancellations, the yellow one is associated to limit orders within the spread and the red one symbolizes market orders.

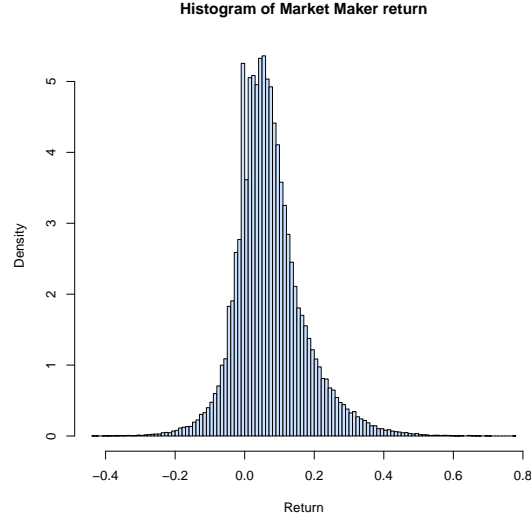


Figure .9 – Density estimation of the gain made by the Market Maker.

### 3.2 DHFT control problem

#### 3.2.1 Problem formulation

The optimisation problem of the DHFT is quite similar to that of MM. However the major difference is that the DHFT trades a second asset which corresponds to the futures of the same stock. Thus, we keep the same notations of the previous section and describe only the dynamics of the new added variable  $F_t$  associated to the future reference price

$$F_t = \frac{P_t^1 + P_t^2}{2} + S_t,$$

where  $S_t$  is a mean-reverting process

$$S_t = S_0 + \int_0^t \rho(\hat{s} - S_u) du + \int_0^t \sigma(S_u) dW_u. \quad (10)$$

Here,  $\rho$  is the strength of the mean reversion,  $\hat{s}$  is the average mean reversion level and  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  is a Lipschitz bounded function representing the volatility of the process.

The strategy of the DHFT is described by the same quantities as the one of the MM in the previous section. The key distinction lies in the fact that he constantly holds a number equal to  $-I$  units of the futures  $F$ . Additionally, we assume that buying/selling the futures leads to the payment of a proportional cost  $\kappa \geq 0$ .

The aim of the DHFT is to maximise the expected utility

$$\mathbb{E}[-\exp(-\eta\{U(S_T, Z_T^\mu) - \varrho j\})],$$

with

$$U(s, z) := g + i^+ \Delta_-^1 - i^- \Delta_+^2 - \kappa([i^+ - q^1]^+ + [i^- - q^2]^+),$$

for any  $z = (p^1, p^2, q^1, q^2, g, i, n^1, n^2, b^1, b^2, j) \in D_Z$  where the new variables  $\Delta_{\pm}^1$  and  $\Delta_{\pm}^2$  satisfy

$$\Delta_{\pm}^1 := p^1 - \frac{p^1 + p^2}{2} - s \pm \kappa, \quad \Delta_{\pm}^2 := p^2 - \frac{p^1 + p^2}{2} - s \pm \kappa.$$

The resolution of this optimization problem is very similar to that of the MM. Thus, the partial differential equation satisfied by the value function in the viscosity sense, the convergence of the numerical scheme and the one of the approximate controls still hold for the DHFT control problem. Furthermore, we repeat for the DHFT the same kind of experiments made for the MM. We find that the DHFT strategy consists in mainly making money out of the spread between the asset price and the futures price since his position in stocks is always covered by a symmetric position in the futures.

### 3.3 IB control problem

We now turn to our last market participant. The IB faces an execution problem. He needs to buy  $I_0$  shares from the stock within a time interval  $[0, T]$ . To solve the IB control problem we consider two types of strategies: a volume strategy and a Volume Weighted Average Price (VWAP) strategy. However, we only show the results for the VWAP one in this introduction.

#### 3.3.1 VWAP strategy

We now present a VWAP based strategy. The VWAP is defined as follows:

$$\text{VWAP}_T = \frac{w_T}{v(0, T)} \quad \text{where } w := \int_0^{\cdot} v_t P_t dt \quad \text{and } v(\cdot, T) := \int_{\cdot}^T v_s ds, \quad (11)$$

with  $v_t$  the traded volume on the market and  $P_t$  the stock price. We model the broker state through the process  $U_t^\mu = (P_t^\mu, I_t^\mu, G_t^\mu)$  where  $I_t$  refers to its inventory,  $G_t$  is its wealth and  $\mu_t$  corresponds to the agent's control. The IB chooses its instantaneous trading speed  $(\mu_t)_{t \geq 0}$  to maximise the expected utility

$$\mathbb{E}[-\exp[-\eta\{G_T + I_T(P_T - \tilde{\kappa} I_T) - I_0 \text{VWAP}_T\}],$$

for some  $\tilde{\kappa} > 0$  which represents a penalty in case the inventory does not match 0 at  $T$ . We have the following formula for  $v$ .

**Result 8.** *The IB value function defined such that*

$$v(t, p, i, g, w) = \sup_{\mu} \mathbb{E}[-\exp[-\eta\{G_T + I_T(P_T - \tilde{\kappa} I_T) - I_0 \text{VWAP}_T\}] | \mathcal{F}_t],$$

*satisfies the semi closed formula*

$$v(t, p, i, g, w) = -e^{-\eta\{g - \tilde{m}w - p v(t, T) \tilde{m}\} + i(p - \tilde{\kappa} i)} \times e^{h_0(t) + h_1(t)i + h_2(t)i^2},$$

*where the functions  $h_0$ ,  $h_1$  and  $h_2$  are fully characterized by ODEs.*

### 3.3.2 Numerical experiment

Figure .10 shows a simulated path for a broker with a VWAP strategy. It presents the variations of the brokers inventory, the cost of a unit share during the execution, the brokers controls and the order book state. We can see that a small number of market orders are sent. Thus, the liquidity imbalance created by this execution algorithm is the main driver of the price upward trend. Figure .11 provides an histogram of the relative error (in %) of the VWAP obtained by both the volume and the VWAP strategies with respect to the VWAP of the whole market.<sup>15</sup> One can see that actually the VWAP strategy performs better than the market on average. This is an expected result since the VWAP strategy is built in order to beat the VWAP of the market.

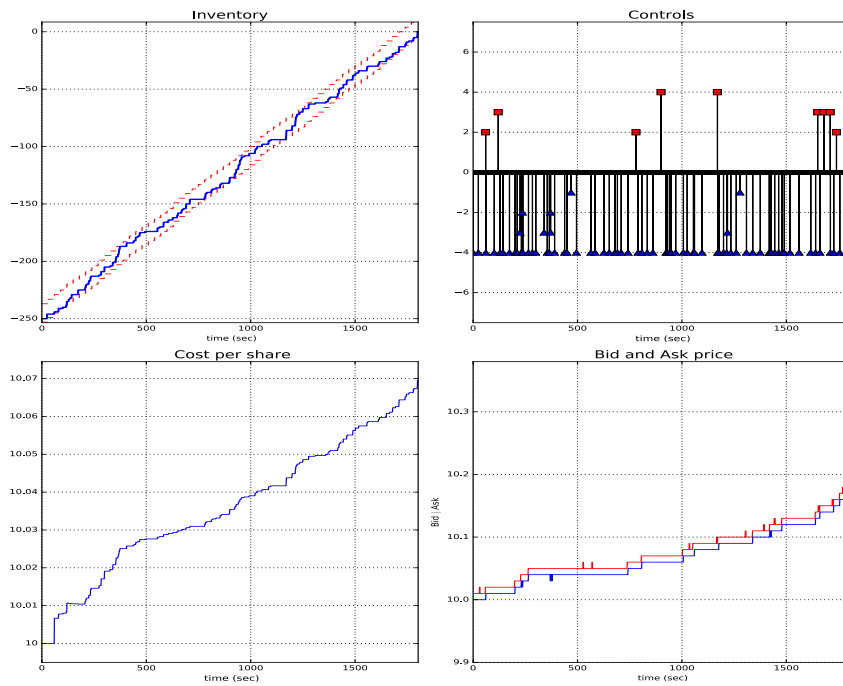


Figure .10 – Simulated path of the VWAP strategy.

<sup>15</sup>Namely,  $(VWAP_{MM} - VWAP_{Market})/VWAP_{Market}$ , in which  $VWAP_{MM}$  is the VWAP obtained by the IB by playing his optimal strategy

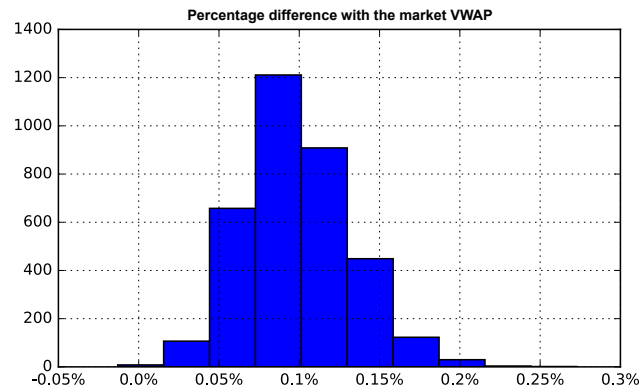


Figure .11 – Histogram of the VWAP percentage error with respect to the VWAP of the market for the VWAP strategy. Histogram based on  $10^4$  simulated paths.

### 3.4 Full market simulation

We now propose an illustration of possible interactions that may happen between several market participants acting optimally. To do so, we consider a market made of 6 agents: 1 MM, 1 DHFT, 2 IBs playing a volume strategy to execute a fixed quantity (one buys and the other sells), 2 IBs playing a VWAP strategy to execute a fixed quantity (one buys and the other sells). At each time step, each participants sends its optimal decision given the current state of the order book. The control of the DHFT is executed first, followed by the one of the MM and finally the actions of the volume and VWAP IBs. This procedure is a first building block towards a realistic market simulator that can be used to back-test strategies or estimate the quantities of interest in a simulated market with a given composition of the agents (i.e. the type of participants and their number), etc. More details are given in Chapter III.

Now we want to compare individual market participants, notably in term of their contribution to market quality. This is why we aim at proposing a model based on the individual behaviour of agents, allowing us to rank them with respect to the impact of their trading on market stability.

## 4 Chapter IV - From asymptotic properties of general point processes to the ranking of financial agents

Chapter IV answers Question 4: How to build a model for the interactions between strategies of individual market participants and use it to assess their contribution to market quality? Instead of aggregating agents into groups as in the approach of Chapter III, we want to model the specific flows of each agent taking part in the market. To do so, we restrict ourselves to the modelling of the best bid and ask dynamics. We propose a very general

framework where individual behaviours are modelled using point processes that can be seen as state-dependent and non-linear Hawkes-type processes. The well-known Poisson, Queue-reactive and Hawkes Queue-reactive dynamics for order books are particular cases of our setting. Using this approach, we establish theoretical results allowing us to assess the specific contributions of agents to market quality.

#### 4.0.1 Modelling of the best bid and ask dynamics

We use an event by event approach. Each event is characterised by  $(T_n, X_n) \in (\mathbb{R}^+, E)$  where:

- $T_n$  is the time of the  $n^{\text{th}}$  event.
- $X_n$  is a variable encoding the characteristics of the event:
  - size  $s_n$ : an integer representing the order size.
  - price  $p_n$ : equals to  $k \in \mathbb{N}$  when the order is inserted at the price best bid  $+k\tau_0$ , where  $\tau_0$  is the tick size.
  - direction  $d_n$ :  $+$  if it provides liquidity and  $-$  when liquidity is removed or consumed.
  - type  $t_n^o$ : 1 (resp. 2) when the bid (resp. ask) is modified.
  - agent  $a_n$ : the market consists in  $N$  agents.

The order book state is modelled by the process  $U_t = (Q_t^1, Q_t^2, S_t)$  where  $Q_t^1$  (resp.  $Q_t^2$ ) is the best bid (resp. ask) quantity and  $S_t$  is the spread.

**Generalised intensity and market reconstitution.** The intensity  $\lambda_t(e)$  associated to an event  $e \in E$  can be informally defined by

$$\lambda_t(e) = \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\#\{T_n \in (t, t + \delta t], X_n = e\} \geq 1 | \mathcal{F}_t]}{\delta t},$$

where  $\mathcal{F}_t$  is the sigma-algebra representing information from the history of the market. We assume it depends on the past event and current state of the market in the following way:

$$\lambda_t(e) = \psi\left(e, U_{t^-}, t, \sum_{T_i < t} \phi(e, U_{t^-}, t - T_i, X_i)\right),$$

where  $\psi$  is a possibly non-linear function,  $U_{t^-}$  is the order book state relative to the last event before  $t$  and  $\phi$  is the Hawkes-like kernel representing the influence of the past events. The functions  $\phi$  and  $\psi$  are both  $\mathbb{R}_+$ -valued. In absence of the kernel  $\phi$ , this corresponds to the Queue-reactive model (and to the zero-intelligence Poisson model when  $\psi$  does not depend on  $U_{t^-}$ ). When  $\phi$  is non-zero,  $\psi$  represents the interaction between the past events and the current order book state. Moreover, we allow  $\psi$  to have polynomial growth while in the literature it has typically at most linear growth. In this case, this model generalises the Hawkes Queue-reactive approach. Note that this intensity also encompasses Quadratic



Hawkes processes with a separable quadratic kernel. In this setting, one can recover the market intensity  $\lambda_t^M(e')$  of an anonymous event  $e'$  ( $e'$  does not contain the agent identity) using the formula below

$$\lambda_t^M(e') = \sum_{a \leq N} \lambda_t((e', a)).$$

**Ergodicity.** Our first theoretical result is relative to the ergodicity of our limit order book model.

**Result 9.** *Under suitable assumptions,  $\bar{U}_t = (Q_t^1, Q_t^2, S_t, \lambda_t)$  is ergodic: there exists a probability measure  $\mu$  such that (exponential speed of convergence)*

$$\lim_{t \rightarrow \infty} P_t(u, A) = P_t(\mu, A), \quad \forall u, \forall A,$$

where  $P_t(u, \cdot)$  is the probability distribution of  $U_t$  given the initial condition  $u \in W_0$  and  $u$  represents all the past events up to time 0. Thus,  $u$  is a sequence indexed by  $\mathbb{N}^-$  for past events and valued in  $\mathbb{R}_+ \times E$ . The measure  $\mu$  is a probability distribution defined on the space  $(W_0, \mathcal{W}_0)$ <sup>16</sup> and  $P_t(\mu, A) = \int_{W_0} P_t(x, A) \mu(dx)$  is the probability distribution of  $U_t$  starting from the random initial condition  $\mu$ .

From the ergodic property, we can derive asymptotic results for the long-term behaviour of our system.

**Scaling limits.** The reference price after  $n$  jumps  $P_n$  writes  $P_n = P_0 + \sum_{i=1}^n \Delta P_i$  where  $\Delta P_i = P_i - P_{i-1} = \eta_i$  and  $\mathbb{E}[\eta_i] = 0$ . We now assume that  $\eta_i$  is centred with  $\eta_i = f(U_i)$  for some measurable function  $f$  and consider the process

$$X_n(t) = \frac{P_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad \forall t \geq 0.$$

The next result describes the behaviour of the price at the macroscopic scale (in event time, the result in calendar time is very similar and also provided in the core of Chapter IV).

**Result 10.** *Under the stationary distribution  $\mu$ , see Result 9, the quantity  $X_n(t)$  satisfies the following convergence result:*

$$X_n(t) \xrightarrow{\mathcal{L}} \sigma W_t,$$

with  $\sigma^2 = \mathbb{E}_\mu[\eta_0^2] + 2 \sum_{k \geq 1} \mathbb{E}_\mu[\eta_0 \eta_k]$  and  $\mu$  the stationary distribution.

This result relates the individual order flow intensities and the macroscopic volatility of the asset. It enables us to rank market participants according to their contribution to volatility. We now explain how empirical computations of relevant quantities can be made.

---

<sup>16</sup>A detailed description of this space is given in Chapter IV.

### 4.0.2 Explicit computations of market quantities

In this model, we can derive semi-explicit formulas for the stationary distribution of the order book, the expected spread, the price volatility and intensities of the fluctuations of the cumulated imbalance. In the following, we provide those relative to the stationary distribution, expected spread and the price volatility.

**Stationary probability distribution of the order book.** Let  $\pi(\cdot) = P(\mu, \cdot)$  be the stationary distribution of  $U_t$ , see Result 9. Then, we can compute  $\pi$  by solving the following equation.

**Result 11.** *The distribution  $\pi$  satisfies*

$$\begin{aligned}\pi Q &= 0, \\ \pi \mathbf{1} &= 1,\end{aligned}\tag{12}$$

where the infinite dimensional matrix  $Q$  verifies

$$Q(u, u') = \sum_{e \in E(u, u')} \mathbb{E}_\mu[\lambda(e)],$$

with  $E(u, u')$  the set of events directly leading to  $u'$  from  $u$ .

In order to compute  $Q(u, u')$ , we take two states  $u$  and  $u'$  such that  $u \neq u'$  and define

$$N_t^{u, u'} = \sum_{T_i < t} \delta_{u, u'}^i, \quad t^u = \sum_{T_i < t} \Delta T_i \mathbf{1}_{U_{T_{i-1}} = u},$$

with  $\delta_{u, u'}^i = \mathbf{1}_{U_{T_{i-1}} = u, U_{T_i} = u'}$  and  $\Delta T_i = T_i - T_{i-1}$ . Therefore, one can estimate the matrix  $Q$  in the following way.

**Result 12.** *We have*

$$\hat{Q}(u, u') = \frac{N_t^{u, u'}}{t^u} \xrightarrow[t \rightarrow \infty]{} Q(u, u'), \quad a.s.\tag{13}$$

Note that the estimator of  $\hat{Q}(u, u') := N_t^{u, u'} / t^u$ , and hence  $\pi$ , does not depend on the model.

**Spread computation** We recall that the spread  $S$  is a state variable since  $U = (Q^1, Q^2, S)$ . Thus, the expected value of the spread  $\bar{S}$  under the stationary distribution satisfies

$$\bar{S} = \mathbb{E}_\pi[S] = \sum_{q^1, q^2, s} \pi(q^1, q^2, s)s.$$

We end with the computation of the volatility in practice when restricting ourselves to the Markov case (for simplicity here).

**Price volatility computation in the Markov case.** In the Markov case, the second term in the volatility formula introduced in Result 10 can be easily computed from the marginal stationary distribution  $\pi$ . Indeed, the transition matrix  $P$  of the Markov chain  $U$  (transitions after one jump) satisfies:

$$P_{u,u'} = \begin{cases} -Q(u,u')/Q(u,u), & \text{if } u \neq u' \text{ and } Q(u,u) \neq 0, \\ 0, & \text{if } u \neq u' \text{ and } Q(u,u) = 0, \end{cases} \quad \text{and} \quad P_{u,u} = \begin{cases} 0, & \text{if } Q(u,u) \neq 0, \\ 1, & \text{if } Q(u,u) = 0, \end{cases}$$

for any  $u$  and  $u'$ . In this setting, we have (recall that  $\eta_i = f(U_i)$ ):

$$\mathbb{E}_\mu[\eta_0 \eta_k] = \mathbb{E}_\pi[\eta_0 \eta_k] = \sum_u \pi(u) f(u) \mathbb{E}_u[\eta_k], \quad \mathbb{E}_u[\eta_k] = \sum_{u'} P_{u,u'}^k f(u'). \quad (14)$$

### 4.0.3 Numerical experiments

Using Result 10 and placing ourselves in the Markov case, we propose a ranking of the nine main market makers based on their impact on volatility. For this, we use four large tick assets (for which the model is very suitable): Air Liquide, EssilorLuxottica, Michelin and Orange, on Euronext, over a one year period: from January 2017 till December 2017.

For each asset, we compute first the liquidity provision and consumption intensities relative to the whole market using Equation (13). Then, we estimate the stationary measure of the order book using Equation (12). Finally we obtain the macroscopic volatility using Equation (14). We give in Figure .12 the results relative to Air Liquide.

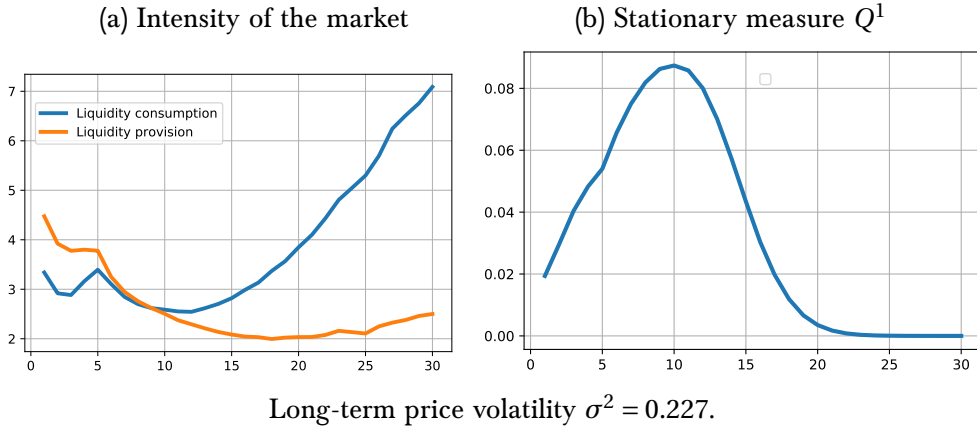


Figure .12 – (a) Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in average event size) and (b) the corresponding stationary distribution of ( $Q^1$ ) with respect to the queue size (in average event size), proper to Air Liquide.

Thereafter, for each market maker, we estimate the new market volatility  $\sigma^{2,N}$  in a situation where we suppose that he withdraws from the market and compare it with the market volatility  $\sigma^{2,M}$  when all the participants are involved: If  $\sigma^{2,N} > \sigma^{2,M}$ , this means that the market maker in question decreases the volatility (i.e. stabilizes the market) otherwise he increases

5. Chapter V - Improving reinforcement learning algorithms: towards optimal learning rate policies

the volatility (i.e. destabilises the market). Thus, the agent who increases volatility (resp. decreases) the most is ranked first (resp. last). In the following table, we establish the ranking of market makers which is actually quite stable across assets.

Market maker	Ranking Air Liquide	Market share Air Liquide	Ranking Exilor-Luxottica	Market share Exilor-Luxottica	Ranking Michelin	Market share Michelin	Ranking Orange	Market share Orange
MM1***	4	4%	3	3%	3	4%	3	3%
MM2	9	1%	9	1%	9	1%	7	1%
MM3	6	5%	6	5%	7	4%	5	4%
MM4	5	1%	4	1%	4	0%	4	1%
MM5	7	5%	8	5%	8	5%	9	5%
MM6****	1	3%	2	3%	1	3%	1	4%
MM7****	2	7%	1	12%	2	9%	2	7%
MM8*	3	9%	5	5%	5	5%	6	4%
MM9	8	2%	7	2%	6	2%	8	2%

Table .1 - Market share and ranking of markets makers. We put the symbol \* next to the name of the market maker each time he is decreasing the volatility of an asset.

In Questions 1, 2 and 3, we solve some control problems faced by different classes of financial agents using numerical schemes. Such approach is appropriate when the model and its parameters are known. However, in many situations the dynamics of the model are not clearly stated. A classical way to overcome this problem is to use stochastic iterative algorithms. These algorithms appear in many applications and require weak convergence conditions. Hence, we address in the next chapter the following question: How to dynamically choose the learning rate in order to improve the convergence of stochastic iterative algorithms?

## 5 Chapter V - Improving reinforcement learning algorithms: towards optimal learning rate policies

In this chapter, we answer Question 5: Is there an optimal policy for the selection of the learning rate in order to improve the convergence of reinforcement learning algorithms? We start our work by investigating to what extent one can improve the classical asymptotic convergence rate  $O(1/\sqrt{N})$ . Indeed, we show that this rate is pessimistic and can be replaced in many applications by  $O((\log(N)/N)^\beta)$  with  $\frac{1}{2} \leq \beta \leq 1$  and  $N$  the number of iterations. Next, we propose a dynamic optimal policy for the choice of the learning rate  $(\gamma_k)_{k \geq 0}$  used in stochastic approximation. For this, we first introduce the algorithm PASS that modifies the values of a predefined sequence  $(\gamma_k^o)_{k \geq 0}$  using the sign variations of the past observed errors. This algorithm accelerates the convergence of standard stochastic iterative algorithms with the step size  $(\gamma_k^o)_{k \geq 0}$ . Second, we propose an optimal methodology for the construction of an appropriate piecewise constant sequence  $(\gamma_k^o)_{k \geq 0}$ . Third, we combine this method and Algorithm PASS to improve the convergence of stochastic iterative algorithms. Finally, we show

that our selection methodology of the learning rate provides better convergence results than standard RL algorithms in three numerical examples: the estimation of a drift, the optimal placement of limit orders and the optimal execution of a large number of shares.

### 5.1 Improvement of the asymptotic convergence rate

Let us first formulate our problem. We aim at finding  $q^* \in \mathbb{R}^d$  solution of the equation

$$M(q, z) = \mathbb{E}[m(q, X(z), z)] = 0, \quad \forall z \in \mathcal{Z} = \{1, \dots, d\},$$

where  $X(z) \in \mathbb{R}^k$  is a random variable with an unknown distribution and  $m$  is a function from  $\mathbb{R}^d \times \mathbb{R}^k \times \mathcal{Z}$  to  $\mathbb{R}$ . Although the distribution of  $X(z)$  is unspecified, we assume that we can observe some variables  $(Z_n)_{n \geq 1}$  valued in  $\mathcal{Z}$  and  $(X_n(Z_n))_{n \geq 1}$  drawn from the same distribution of  $X(Z_n)$ . In such context, reinforcement learning (RL) addresses this problem through the following iterative procedure:

$$q_{n+1}(Z_n) = q_n(Z_n) - \gamma_n(Z_n)m(q_n, X_{n+1}(Z_n), Z_n), \quad (15)$$

with  $q_0$  a given initial condition and  $(\gamma_n)_{n \geq 0} \in \mathbb{R}^d$  a component-wise non-negative sequence. The asymptotic convergence speed of stochastic iterative algorithms is in general  $O(1/\sqrt{N})$ . We aim at investigating how one can extend and improve such asymptotic rate for (15). Note that, the function  $q^*$  is solution of the minimization problem

$$\min_{q \in \mathbb{R}^k} g(q),$$

with  $g(q) = \|M(q)\|^2$ .<sup>17</sup> Since we do not have a direct access to the distribution of  $X(z)$ , we use that at time  $n$  one keeps memory of a training sample of  $n(z)$  independent variables  $(X_i^z)_{i=1 \dots n(z)}$  drawn from  $X(z)$  distribution. Here  $n(z)$  is the number of times when the process  $Z_n$  visited  $z$ . We define  $q^n$  as a solution of

$$\min_{q \in \mathbb{R}^k} g_n(q), \quad (16)$$

with  $g_n(q) = \|M^n(q)\|^2$  and  $M^n(q, z) = \mathbb{E}^n[m(q, X(z), z)] = \frac{1}{n(z)} \sum_{j=1}^{n(z)} m(q, X_j(z), z)$  the expected value under the empirical measure  $\mu = \frac{1}{n(z)} \sum_{j=1}^{n(z)} \delta_{X_j(z)}$ . After that, we denote by  $q_k^n$  an approximate solution of the problem (16) given by an optimization algorithm after  $k$  iterations. Thus, we can decompose our total error into the estimation error and the optimization error as follows:

$$0 \leq \mathbb{E}[g(q_k^n) - g(q^*)] \leq \underbrace{\mathbb{E}[g(q^n) - g(q^*)]}_{\text{estimation error}} + \underbrace{\mathbb{E}[|g(q^n) - g(q_k^n)|]}_{\text{optimization error}}.$$

<sup>17</sup>When  $M$  is the gradient of a function  $f$  (i.e.  $\nabla f = M$ ), one can show that  $q^*$  minimises a convex and differentiable function  $g(q) = \sum_{i=1}^d \mathbb{E}[L(q, X(i), i)]$  with  $L$  a well defined operator.

### 5.1.1 Estimation error

We focus now on the first source of error.

**Slow convergence rate.** Standard uniform convergence results ensure that

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|] \leq c \frac{1}{\sqrt{n}},$$

with  $c > 0$  a positive constant. This result allows us to derive the pessimistic bound  $2c \frac{1}{\sqrt{n}}$  for the estimation error.

**Fast convergence rate.** Many authors obtain fast statistical convergence rate in the following form:

$$\mathbb{E}[\sup_q (g(q) - g_n(q))] \leq c \left(\frac{\log(n)}{n}\right)^\beta, \quad (17)$$

with  $\frac{1}{2} \leq \beta \leq 1$ .<sup>18</sup> These fast statistical estimations are established under regularity and boundedness conditions of the function  $M$ , see [19, 58] or when the data distribution satisfies some noise conditions, see for instance [19, 150]. Inequality (17) enables us to derive a bound for the estimation error proportional to  $(\frac{\log(n)}{n})^\beta$ .

### 5.1.2 Optimization error

We turn now to the optimization error. Since  $g$  is replaced by the empirical mean  $g_n$  which is known, there are many algorithms in the literature that approximate  $q^n$ . Under suitable assumptions one can expect an exponential convergence speed for such algorithms. In unfavourable situations, it is possible to ensure a convergence rate of  $O(1/N)$ .

### 5.1.3 Conclusion

We have decomposed our initial error into

- **Estimation error:** its convergence is  $O(\frac{1}{\sqrt{N}})$  in pessimistic cases. In the other situations, the convergence is faster (i.e.  $O(\frac{\log(N)}{N})^\beta$ ) with  $\frac{1}{2} \leq \beta \leq 1$ .
- **Optimization error:** the convergence is exponential under suitable conditions. In unfavourable cases, the convergence rate is  $O(\frac{1}{N})$ .

The comparison of these error sources shows that the estimation error is the dominant component. In addition to that, we can overcome the  $O(\frac{1}{\sqrt{N}})$  asymptotic speed, in some situations, by improving the estimation error.

---

<sup>18</sup>It is possible to get rid of the  $\log(n)$  factor in the above inequality using the “chaining” technique, see [40].

## 5.2 Optimal policy for the learning rate

We consider now Algorithm (15) and propose an optimal way to choose the learning rate  $\gamma_k$ . This type of algorithms appears in many contexts: stochastic iterative algorithms, gradient methods, fixed point iterative techniques, etc. It is well known that the choice of  $\gamma_k$  plays a key role for the convergence. Thus, we focus only on the step size and present a selection policy for  $(\gamma_k)_{k \geq 0}$  decomposed into two interacting levels

- The inner level: we define the algorithm PASS inspired from line search that uses a predefined sequence  $(\gamma_k^o)_{k \geq 0}$  and the sign variations of  $m(q_k, X_{k+1}(Z_k), (Z_k))$  to construct a new sequence  $(\hat{\gamma}_k)_{k \geq 0}$ . This new sequence improves the convergence of Algorithm (15).
- The upper level: we propose an optimal methodology for the selection of a piecewise constant predefined sequence  $(\gamma_k^o)_{k \geq 0}$ .

### 5.2.1 The inner level

For any predefined sequence  $(\gamma_k^o)_{k \in \mathbb{N}}$ , we introduce the following version of the algorithm PASS that is adapted to reinforcement learning (RL) issues. This is the version used for our numerical experiments.

**Algorithm 1** (PAst Sign Search (PASS) for (RL)). *We start with an arbitrary  $q^0$  and define by induction  $q_n$  and  $\hat{\gamma}_n$*

- If  $m(q_n, X_{n+1}(Z_n), Z_n) \times m(q_{r_1^n}, X_{r_1^n+1}(Z_{r_1^n}), Z_{r_1^n}) \geq 0$ , then do

$$\begin{aligned} q_{n+1}(Z_n) &= q_n(Z_n) - h(\hat{\gamma}_n(Z_n), \gamma_n(Z_n))m(q_n, X_{n+1}(Z_n), Z_n), \\ \hat{\gamma}_{n+1}(Z_n) &= h(\hat{\gamma}_n(Z_n), \gamma_n(Z_n)), \end{aligned}$$

with  $r_1^n$  is the index of the last observation when the process  $X$  visits the state  $X_n$ .

- Else, do

$$\begin{aligned} q_{n+1}(X_n) &= q_n(X_n) - l(\hat{\gamma}_n(X_n), \gamma_n(X_n))m(q_n, X_n, X_{n+1}), \\ \hat{\gamma}_{n+1}(X_n) &= l(\hat{\gamma}_n(X_n), \gamma_n(X_n)). \end{aligned}$$

We have other versions of this algorithm where the vector  $q_n$  is modified entirely and not component-wise. We define the total error  $E^k$  such that  $E^k = \sum_{z=1}^d (q_k(z) - q^*(z))^2$ . We first compare the error  $E^k$  of PASS with the one of (15) and show that it decreases faster. This enables us to derive the convergence of PASS.

**Result 13.** *Under mild assumptions, Algorithm PASS converges in the sense that*

$$\mathbb{E}[E^n] \xrightarrow{n \rightarrow \infty} 0.$$

### 5.2.2 The outer level

To select the initialisation learning rate  $(\gamma_k^o)_{k \in \mathbb{N}}$ , we construct a sequence of integers  $(k_i)_{i \in \mathbb{N}}$  such that  $\gamma_k^o$  is constant within each interval  $I_i = [k_i, k_{i+1}[$  with  $i \in \mathbb{N}$ . Then, we choose the “best” constant  $\gamma_{k_i}^o$  such that the total error  $E^{k_i}$  decreases as fast as possible by at least a factor  $\alpha$  during the time period  $I_i$  in the sense that  $\mathbb{E}[E^{k_{i+1}}] < \alpha \mathbb{E}[E^{k_i}]$ .

We place ourselves in the interval  $I_i$ . In such case, the step size  $\gamma_{k_i}^o = \gamma$  is fixed. Inside this interval, we show that

**Result 14.** *Under mild assumptions,  $\exists \bar{M} > 0, \beta \in [0, 1)$  such that the error  $E^n$  verifies*

$$\mathbb{E}[E^{n+k_i}] \leq c(\gamma)(1 - \tilde{\alpha}^n) + \tilde{\alpha}^n \bar{M} \mathbb{E}[E^{k_i}], \quad \forall \tilde{\alpha} < \max(\beta, \alpha(\gamma)), \quad (18)$$

with  $c(\gamma) = \gamma^2 R$ ,  $\alpha(\gamma) = 1 - f(\gamma)$ ,  $f(x) = 2Lx - Bx^2$  and  $(R, L, B) \in \mathbb{R}_+^3$  positive constants.

The above result ensures that the error converges exponentially towards the constant  $c(\gamma)$ . Thus, the main idea is to select  $\gamma$  such that  $c(\gamma) \leq \alpha E^{k_i}$ . This gives a set  $[0, \gamma^{\max}]$  of potential values for  $\gamma$  with  $\gamma^{\max}$  a fixed constant. Then, we select among this set the unique value of  $\gamma$  that maximises  $\alpha(\gamma)$ . After fixing the value of  $\gamma$ , we give ourselves an error precision  $r$  and define  $k_{i+1} = k_i + j_i$  with  $j_i$  the first index  $n$  such that  $\mathbb{E}[E^{n+k_i}] \leq \alpha \mathbb{E}[E^{k_i}](1 + r)$ . Inequality (18) provides an estimation for  $j_i$ .

Using the above methodology, we define recursively the sequence  $(k_i)_{i \in \mathbb{N}}$  and  $(\gamma_{k_i})_{i \in \mathbb{N}}$  as follows:

$$\begin{cases} k_{i+1} = k_i + \lceil \frac{\log(\frac{\alpha r}{2(1+r)\bar{M}})}{\log(\alpha_{k_i})} \rceil, & \gamma_{k_{i+1}} = \alpha \gamma_{k_i}, & \forall i \geq 0, \\ k_0 = 0 & \gamma_{k_0} = \tilde{\gamma}_1 \vee \tilde{\gamma}_2, \end{cases}$$

where  $\alpha_{k_i} = 1 - f(\gamma_{k_i})$ . The quantities  $\tilde{\gamma}_1$  and  $\tilde{\gamma}_2$  are two suitable initialisation constants that ensures the decrease of the error during the first iterations. We show that the convergence speed of the algorithm is exponentially fast as long as  $i \leq i^*$  with  $i^*$  a fixed threshold. However, when  $i > i^*$  we recover a slow convergence rate of  $O(\frac{1}{\sqrt{i}})$ .

### 5.3 Numerical experiments

In Chapter V we consider three numerical examples: the drift estimation, the optimal placement of limit orders and the optimal execution of a large amount of shares. As an illustration, we choose here to present only the results associated to the optimal placement problem.

**Optimal placement.** We consider an agent who aims at buying a unit quantity using limit orders and market orders. In such case, the agents wonder how to find the right balance between fast execution and avoiding trading costs associated to the bid-ask spread. The agent state at time  $t$  is modelled by  $X_t = (Q^{Before}, Q^{After}, P)$  with  $Q^{Before}$  the number of



shares placed before the agent’s order,  $Q^{After}$  the queue size after the agent’s order and  $P_t$  the mid price, see Figure .3. The agents wants to minimise the quantity

$$\mathbb{E}[F(X_{\tau \vee T^{exec}}) + \int_0^{\tau \vee T^{exec}} c ds],$$

where

- $T^{exec} = \inf\{t \geq 0, P_t = 0\}$  the first time when the limit order gets a transaction.
- $\tau$  the first time when a market order is sent.
- $X = (Q^{Before}, Q^{After}, P)$  the state of the order book.
- $F(u)$  is the price of the transaction (i.e.  $F(u) = p + \psi$  when the agents crosses the spread and  $F(u) = p$  otherwise).

Using the classical dynamic programming principle, one can solve this control problem using (15) since the  $Q$  function is a zero of a certain operator, see Chapter V. Figure V.4 shows three control maps: the x-axis reads the quantity on “same side” (i.e.  $Q^{same} = Q^{Before} + Q^{After}$ ) and the y-axis reads the position of the limit order in the queue, i.e.  $Q^{Before}$ . The color and numbers give the control associated to a pair  $(Q^{same}, Q^{Before})$ : 1 (blue) means “stay in the book”, while 0 (red) means “cross the spread” to obtain a transaction. The panel (at the left) gives the reference optimal controls obtained with a finite difference scheme, the middle panel the optimal controls obtained for the Algorithm (15) where the step-size  $(\gamma_k)_{k \geq 0}$  is derived from the upper level of our optimal policy and the right panel the optimal control obtained with our optimal policy (i.e. upper level and inner level combined). It shows that after few iterations our optimal policy already found the optimal controls. Figure V.5 compares the log of the  $L^2$  error, averaged over 1000 trajectories, between the different algorithms. We see clearly that our methodology improves basic stochastic approximation algorithms.

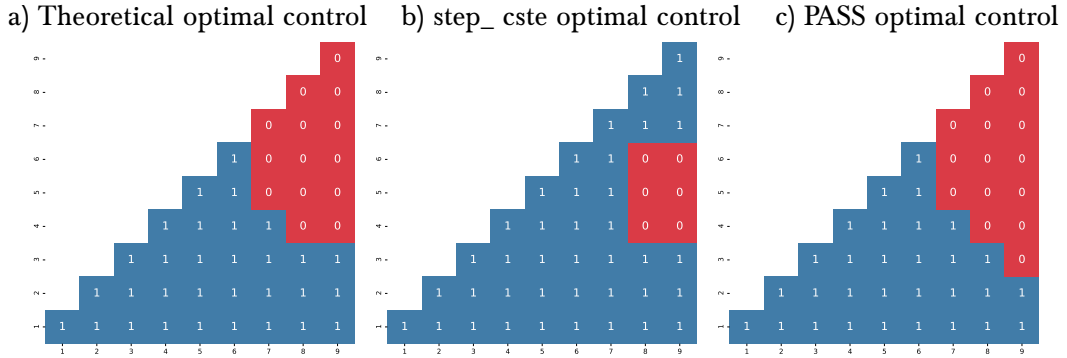


Figure .13 – Comparison optimal control after 300 iteration for different methods: left is the optimal control, middle is basic RL with a step size derived from the upper level and right is our optimal policy for the step size (i.e. upper level and inner level combined).

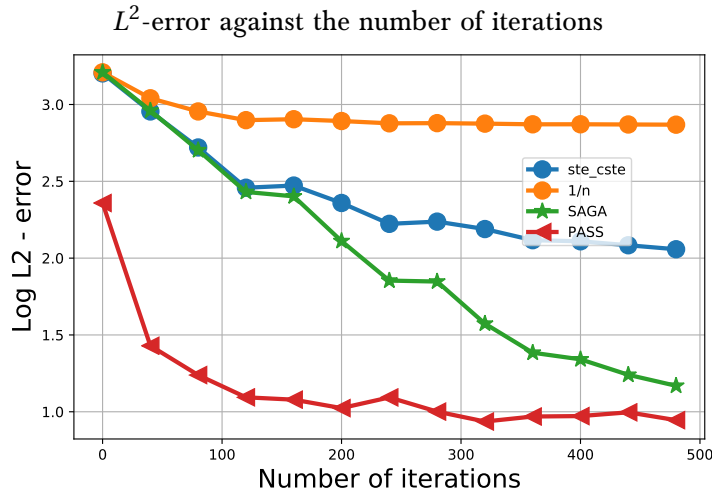


Figure .14 – The log  $L^2$ -error against the number of iterations.

The applications studied in the previous chapters are appropriate for liquid markets where the order book mechanism is well developed such as equity markets. Nonetheless, our methodology may be less adapted to markets with very specific operating rules. To address this issue, we study in the next chapter the price variations of the intra-day electricity market. Thus, we tackle our final question: What are the main price features in such market and how to design a model that reproduces them?

## 6 Chapter VI - Price impact study on the electricity intra-day market

In this last chapter, we address the following question: What are the main properties of the price in the electricity intra-day market and how to build a model that captures them? We propose a general empirical study of price fluctuations and trading activity in this market. This allows us to capture the most important characteristics of the electricity market. Then, we focus on the price impact generated by a power plant breakdown. This is of great importance for market participants that face several power plant failures per year and need to include this cost in their trading strategies. After that, we investigate some further properties of the price variations and finally propose a model consistent with our observations.

## 6.1 General empirical analysis

### 6.1.1 Description of the intra-day electricity market

Here we give a description of the electricity intra-day market runned by Epex Spot in France. Every day at 3 p.m, Epex Spot opens trading sessions<sup>19</sup> where agents trades forward contracts. The specificity of these forward contracts is that they should be delivered on the following day. Each trading session is associated to a unique time period and buying a forward contract in one given session ensures the supply of 1 MW of energy during the related time period. Each of those sessions remains open until 30 minutes before the beginning of the delivery. When a session for a given delivery period is open, trading can occur continuously and is run using an order book mechanism.

### 6.1.2 Databases description

We use two databases. The first one comes from the electricity intra-day market runned by Epex Spot in France. This database records, during two years from the 1st January 2017 to the 31th January 2019, the trades executed in the intra-day electricity market event by event with a minute accuracy for each forward contract. The second database reports all the power plants breakdowns that happened in the same period. The first database accounts for 1 146 777 events and the second 12 109 breakdowns.

### 6.1.3 Features of the intra-day electricity market

After describing the breakdowns distribution, we study the intra-day volume variations and the price seasonality.

**Intra-day volumes.** The exchanged volume is at its lowest level for early maturities when the need of energy supply is low, it attains its highest level during afternoon maturities and then stabilizes for late maturities. It is important to note that for each forward contract the market activity sharply accelerates close to the delivery hour (i.e. maturity).

**Price seasonality.** For each forward contract, we show that the weighted price at the intra-day level have the same variations of the daily exchanged volume. At the weekly scale, the weighted price decreases during the weekend since economic activity is reduced. In addition to that, at the year level, the weighted price is low during the summer since the clients consumption is weak and becomes high in the winter when the clients need more energy.

---

<sup>19</sup>There are trading sessions for the 24 hours and 48 half hours in both France and Germany. For Germany, there are also 96 sessions for the quarters of an hour

## 6.2 Price study

**Price impact.** We deal only with unexpected power plant breakdowns with a minimum level of 150 MW energy loss. Thus we keep only few thousands breakdowns. We compute, for each one of the considered breakdowns, the quantity

$$\Delta P_t = \epsilon_t \frac{P_t - P_{t_0}}{P_{t_0}}, \quad \forall t \in [t_0, t_1],$$

where  $P_t$  is the transaction price,  $\epsilon_t$  is the sign of the transaction (i.e.  $\epsilon_t = 1$  for buy market orders and  $-1$  otherwise) and the times  $t_0$  and  $t_1$  are defined as follows:

- $t_0$  is the first transaction time after the beginning of the breakdown.
- $t_1$  is the minimum time between the ending time of the breakdown, the end of the day and the closure time of the forward contract.

Finally, we define the price impact  $I_t$  as the empirical average of  $\Delta P_t$  over all the breakdowns and the rescaled time  $s$  such that  $s(t) = \frac{t-t_0}{t_1-t_0}$  for any  $t \in [t_0, t_1]$ .

**Aggregated price impact profile.** Figure .15 shows that the aggregated price impact curve (i.e. for all the forward contracts) is linear. We choose to plot the aggregated impact since we have more observations however we check that our conclusions remain valid for the most important maturities (i.e. with high exchanged volume). A linear price impact means that agents behaviours remain unchanged during the whole execution of the meta-order. This can be explained by their incapacity to detect the variations in the order book dynamics generated by the liquidation of a meta-order or by their unwillingness to react to the meta-order execution. Thus, the interactions between agents decisions are weak and their actions are mainly triggered by exogenous information like market announcements.

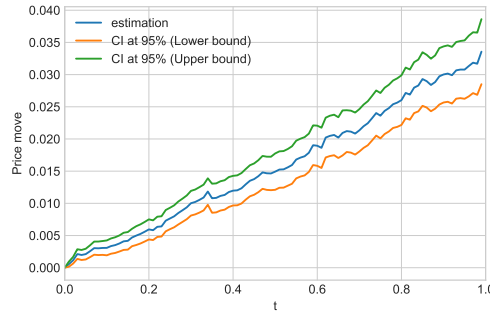


Figure .15 – Price impact profile for aggregated maturities.

**Price impact profiles for different regimes.** We investigate the connection between the price impact curve and the amplitude of the energy loss during a breakdown (i.e. size of the breakdown). We choose this variable since it is the analogous of the size of the meta-order when dealing with execution issues. Many empirical studies have already explored the

relation between these two last variables. They show that the price impact  $I$  generated by a meta-order of size  $Q$  follows the “square root law” (i.e.  $I \approx (\frac{Q}{V_D})^\delta$  with  $\delta > 0$  and  $V_D$  the daily volume). Our study consists in decomposing our initial data in three buckets with the same number of observations using quantiles. Figure .16 plots the price impact associated to each one of these buckets. We can see that the global shape of the impact curve is unchanged when the energy loss varies since the impact is linear for all the buckets. However, its magnitude increases with the breakdown size which is an expected result. Indeed, to thwart a huge energy loss, market participants buy large quantities from the intra-day market and therefore increase the price.

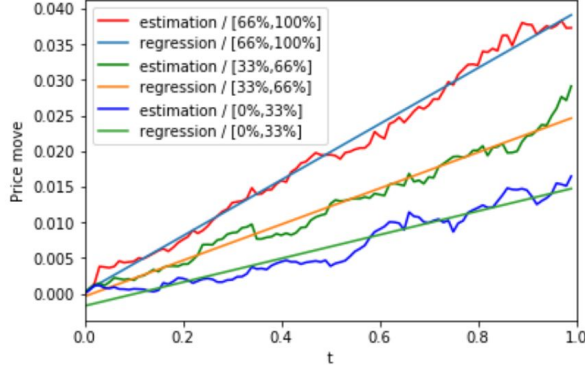


Figure .16 – Price impact for different initial price regimes.

**Trade signs autocorrelation.** Let  $l \in \mathbb{N}$ . We also study the autocorrelation function  $C(l)$  between the trade signs  $\epsilon_l$  defined such that

$$C(l) = \mathbb{E}[\epsilon_l \epsilon_0] - \mathbb{E}[\epsilon_l] \mathbb{E}[\epsilon_0],$$

with  $\epsilon_l = 1$  for buy market orders and  $-1$  otherwise. In liquid markets, one expects typically to observe a strong correlation between trade signs. However, we show the existence of a negative correlation only between the first trade and its successor. Thus it suggests the absence of memory since the long-term correlation between the trade signs is weak.

**Response function autocorrelation.** Let  $l \in \mathbb{N}$ . We study the response function  $R(l)$  between the price moves and the trade signs  $\epsilon_l$  defined such that

$$R(l) = \mathbb{E}[(\bar{p}_l - \bar{p}_0) \bar{\epsilon}_0],$$

with  $\bar{z} = z - \mathbb{E}[z]$  for any process  $z$ . We reveal a power law decrease for the response function and show that the response function remains always higher than a positive bottom level.

**Price diffusion.** We study the regularity of the price process through the Hurst exponent  $H$ . To estimate  $H$  we place ourselves in the following framework. Let  $Z_t$  be a process which satisfies

$$dZ_t = \sigma dW_t^H, \quad \forall t \geq 0,$$

where  $\sigma$  is a constant volatility and  $W_t^H$  is a fractional Brownian motion (fBm) whose Hurst exponent is  $H$ . Then, there exists a constant  $C \geq 0$  such that

$$\mathbb{E}[|Z_{t+\Delta} - Z_t|^q] = C\Delta^{Hq}, \quad \forall t \geq 0, \forall \Delta \geq 0, \forall q \geq 0.$$

This shows that

$$y_\Delta^q = A'_q \Delta + C', \quad (19)$$

with  $y_\Delta^q = \log(\mathbb{E}[|p_{t+\Delta} - p_t|^q])$ ,  $A'_q = Hq$  and  $C' = \log(C)$ . Figure .17.a plots  $y_\Delta^q$  against  $\Delta$  when the underlying process  $Z$  is the price. We see that the linear relation (19) is satisfied. Figure .17.b plots  $A'_q$  for different values of  $q$  and estimates the slope of this curve which represents the Hurst exponent. The estimated Hurst exponent is close to  $0.48 \approx 0.5$  which was not an obvious result in this very specific market.

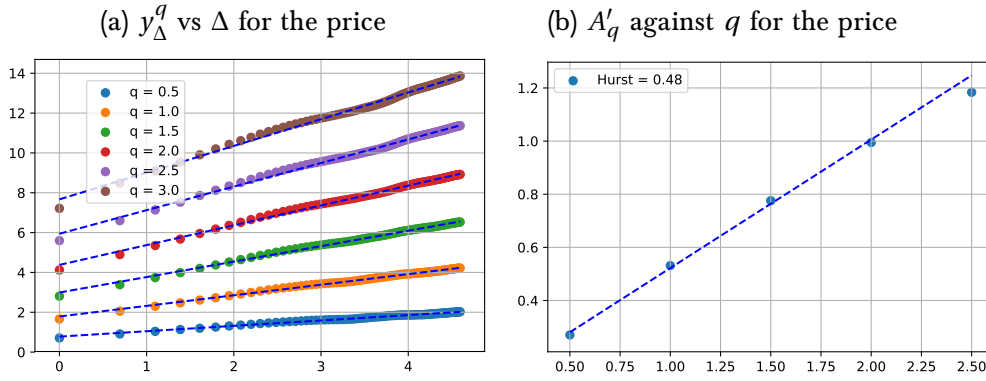


Figure .17 - (a) Plot of  $y_\Delta^q$  against  $\Delta$  during the years 2017-2018, (b) Plot of  $A'_q$  against  $q$  and estimation of the Hurst exponent for the price process associated to the maturity 22h.

### 6.3 Price modelling

We can build a price model consistent with our observations. This model is inspired from the propagator model introduced in [39]. We show how to carry computations and calibrate this model in our specific setting. Such model is useful for prediction and estimation of the quantities of interest. We consider the following dynamics for the price at time  $n$ :

$$p_n = p_0 + \mu_n n + \sum_{k \leq n-1} G(n-k) \epsilon_k,$$

with  $\mu$  a real constant representing the price trend,  $G: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  the impact function that represents the influence of a single transaction on the price,  $\epsilon_n$  is the sign of the  $n$ -th trade.

### 6.3.1 Properties of the model

When the autocorrelation between the trade signs is weak, we show that this model can reproduce a linear price impact. Moreover, we recover the power law decrease of the response function  $R$ . Finally, we study the diffusivity of the price process and link its Hurst exponent with the asymptotic behaviour of the trade signs autocorrelation function  $C(n)$  and the impact function  $G$ .

---

# Limit order strategic placement with adverse selection risk and the role of latency

---

## Abstract

This paper is split in three parts: first, we use labelled trade data to exhibit how market participants decisions depend on liquidity imbalance; then, we develop a stochastic control framework where agents monitor limit orders, by exploiting liquidity imbalance, to reduce adverse selection. For limit orders, we need optimal strategies essentially to find a balance between fast execution and avoiding adverse selection: if the price has chances to go down the probability to be filled is high but it is better to wait a little more to get a better price. In a third part, we show how the added value of exploiting liquidity imbalance is eroded by latency: being able to predict future liquidity-consuming flows is of less use if you do not have enough time to cancel and reinsert your limit orders. There is thus a rationale for market makers to be as fast as possible to reduce adverse selection. Latency costs of our limit order driven strategy can be measured numerically.

To authors' knowledge this paper is the first to make the connection between empirical evidences, a stochastic framework for limit orders including adverse selection, and the cost of latency. Our work is a first step to shed light on the role played by latency and adverse selection in optimal limit order placement.

## 1 Introduction

With the electronification, fragmentation and increase of trading frequency, order book dynamics is under scrutiny. Indeed, a deep understanding of order book dynamics provides insights on the price formation process. There are essentially two approaches for modelling the price formation process. First, *general equilibrium models* based on interactions between rational agents who take optimal decisions. General equilibrium models focus on agents behaviours and interactions. For example, investors split their metaorders into large collections of limit orders (i.e. liquidity-providing) and market orders (i.e. liquidity-consuming) (see [14], [105], [148]) while (high frequency) market makers mostly use limit orders to provide liquidity



to child orders of investors (see [31]). Second, *statistical models* where the order book is seen as a random process (see [39], [83], [90], [114] and references herein). Statistical models focus on reproducing many salient features of real markets rather than agents' behaviours and interactions. In this paper, we consider a statistical model where the arrival and cancellation flows follow size dependent Poisson processes. Using this model, we propose an optimal control for one agent targeting to obtain the "best price" by maintaining optimally a limit order in the order book .

In practice, market participants use optimal trading strategies to find a balance between at least three factors: the price variation uncertainty, the market impact and the inventory risk. For example, an asset manager who took the decision to buy or sell a large number of shares needs to adapt its execution speed to price variations. The simplest case would be to accelerate execution when price moves in its favor. He needs to consider the market impact too and in particular the price pressure of large orders: fast execution of huge quantities consumes order book liquidity and increases transaction costs. Finally, there is an inventory risk associated to the orders size: it is riskier to hold a large position than a small one during the same period of time. A fast execution reduces this inventory risk. The asset manager should then find the optimal balance between trading slow and fast. Models for these strategies are now well known (see [50, 52, 79, 80, 84, 85] ). Recent papers introduce a risk term in their optimisation problem. Moreover, some papers combine even short term anticipations of price dynamics inside these risk control frameworks. For example, in [11], authors include a Bayesian estimator of the price trend in a mean-variance optimal trading strategy. In [51], authors include an estimate of future liquidity consumption  $-\mu$  in their paper should be compared to our consuming intensities  $\lambda^{\pm}$  - in macroscopic optimal execution.

In this paper, we consider an optimal control problem where the agent faces the price variation uncertainty and the market impact but there is no inventory risk since we consider one limit order. The idea is to propose optimal strategies that can be plugged into any large scale strategy (see [145, Chapter 3] for a practitioner viewpoint on splitting the two time scales of metaorders executions) by taking profit of a short term anticipation of price moves.

After some considerations about short time price predictions and empirical evidences showing that market participants decisions depend on the imbalance (see Section 2), we show that optimal control can add value to any short term predictor (see Section 4) in the context of simple controls (cancel or insert a limit order) for a "large tick" stock. This result can thus be used by investors or market makers to include some predictive power in their optimal trading strategies. Then we show how latency influences the efficient use of such predictions. Indeed, the added value of the optimal control is eroded by latency. It allows us to link our work to regulatory questions. First of all: what is the "value" of latency? Regulators could hence rely on our results to take decisions about "slowing down" or not the market (see [44] and [67] for discussions about this topic). It sheds also light on maker-taker fees since the real value of

limit orders (including adverse selection costs), is of great importance in this debate (see [87] for a discussion).

This paper can be seen as a mix of two early works presented at the “Market Microstructure: Confronting Many Viewpoints” conference (Paris, 2014): an optimal control driven one [124] and a data-driven one focused on the predictive power of order books [146]. Our added values are first a proper combination of the two aspects (inclusion of an imbalance signal in an optimal control framework for limit orders), and then the construction of our cost function. Unlike in the second work, we do not value a transaction with respect to the mid-price at  $t = 0$ , but with respect to the microprice (i.e. the expected future price given the liquidity imbalance) at  $t = +\infty$ . We will argue that the difference is of paramount importance since it introduces an effect close to adverse selection, that is crucial in practice.

As an introduction of our framework, we will use a database of labelled transactions on NASDAQ OMX (the main Nordic European regulated markets) to show how order book imbalance is used by market participants in a way that can be seen as compatible with our theoretical results.

Hence the structure of this paper is as follows: Section 2 presents order book imbalance as a microprice and illustrates the use of imbalance by market participants thanks to the NASDAQ OMX database. Once these elements are in place, Section 3 presents our model and Section 4 shows how to numerically solve the control problem and provides main results, especially the influence of latency on the efficiency of the strategy.

## 2 Main Hypothesis and Empirical Evidences

### 2.1 Database presentation

The data used here is a direct feed from NASDAQ-OMX, that is the primary market<sup>1</sup> on the considered stock. Capital Fund Management feed recordings for AstraZeneca account for 72% of market share (in traded value) for the continuous auction on this stock over the considered period. Surprisingly, there is currently no academic paper comparing the predictive power of imbalances on different trading venues. It is outside of the scope of this paper to elaborate on this. We will hence consider that the liquidity on our primary market represents the state of the liquidity on other “large” venues (namely Chi-X, BATS and Turquoise on the considered stock). If it is not the case one can adapt our result relying on statistics on each venue, or on the aggregation of all venues. We did not aggregate venues ourselves for obvious synchronization reasons: we do not know the capability of each market participant to synchronize information coming from all venues and do not want to add noise by making more assumptions. Our idea here is to use the state of liquidity at the first limits on the primary market as a proxy of information about liquidity really used by participants.

---

<sup>1</sup>i.e. the *regulated exchange* in the MiFID sense, see [145] for details.

Venue	AstraZeneca	Vodafone
BATS Europe	7.16%	7.63%
Chi-X	19.27%	20.02%
Primary market	72.24%	61.09%
Turquoise	1.33%	11.26%

Table I.1 – Fragmentation of AstraZeneca (compared to Vodafone) from the 2013-01-02 to the 2013-09-30

We focus on NASDAQ-OMX because this European market has an interesting property: market members' identity is known. It implies that transactions are labelled by the buyer's and the seller's names. Almost all the trading activity on NASDAQ Nordic stocks was labelled this way until the end of 2014. Note that members' identity is not investors' names; it is the identity of brokers or market participants large enough to apply for a membership. High Frequency Participants (HFP) are of this kind. Of course some participants (like large asset management institutions) use multiple brokers, or a combination of brokers and their own membership. Nevertheless, one can expect to observe different behaviours when members are different enough. We will here focus on three classes of participants: High Frequency Participants (HFP), global investment banks and institutional brokers.

## 2.2 Stylized fact 1: the predictive power of the order book imbalance

**Short term price prediction utility.** Academic papers (see [39] or [90]) and brokers' research papers (see [30]) document how the sizes at first limits of the public order book<sup>2</sup> influence the next price move. It is worthwhile to underline that the identified effects are usually not strong enough to be the source of a statistical arbitrage: the expected value of buying and selling back using accurate predictions based on sizes at first limits does not beat transaction costs (bid-ask spread and fees). See [95] for a discussion. Nevertheless:

- For an investor who *already took the decision to buy or to sell*, this information can spare some basis points. For very large orders, it makes a lot of money and in any case it reduces implicit transaction costs.
- Market makers naturally use this kind of information to add value to their trading processes (see [65] for a model supporting a theoretical optimal market making framework including first limit prices dynamics).

The easiest way to summarize the state of the order book without destroying its informational content is to compute its *imbalance*: the quantity at the best bid minus the one at the best ask divided by the sum of these two quantities:

$$\text{Imb}_t := \frac{Q_t^{\text{Bid}} - Q_t^{\text{Ask}}}{Q_t^{\text{Bid}} + Q_t^{\text{Ask}}}. \quad (1)$$

---

<sup>2</sup>Limit order books are used in electronic market to store unmatched liquidity, the *bid size* is the one of passive buyers and the *ask size* the one of passive sellers, see [145] for details.

**The nature of the predictive power of the imbalance.** The predictive power of the order book imbalance is well known (see [113]). The rationale of this stylized fact (i.e. *the midprice will go in the direction of the smaller size of the order book*) is outside of the scope of this paper. We just give here some clues and intuitions to the reader:

- The future price move is positively correlated with the imbalance. In other terms

$$\mathbb{E}[(P_{t+\delta t} - P_t) \times \text{sign}(\text{Imb}_t) | \text{Imb}_t] > 0,$$

where  $P_t$  is the midprice (i.e.  $P_t = (P_t^{Bid} + P_t^{Ask})/2$ , where  $P^{Bid}$  and  $P^{Ask}$  are respectively the best bid and ask prices) at  $t$  for any  $\delta t$ . Obviously when  $\delta t$  is very large, this expected price move is very difficult to distinguish from large scale sources of uncertainty. See for instance [113] for details on the “predictive power” of such an indicator (our Figure II.2 illustrates this predictive power on real data).

- Within a model in which the arrival and cancellation flows follow independent point processes of the same intensity, the smallest queue (bid or ask) will be consumed first, and the price will be pushed in its direction. See [90] for a more sophisticated point process-driven model and associated empirical evidences.
- Another viewpoint on imbalance would be that the bid vs. ask imbalance contains information about the direction of the net value of investors’ metaorders: first, in a direct way if one is convinced that investors post limit orders; second, indirectly if one believes that investors only consume liquidity and in such case bid and ask sizes are indicators of market makers net inventory.

The focus of formula (1) on the first two limits weaken the predictive power of the bid vs. ask imbalance. For large tick assets<sup>3</sup> it may be enough to just use the first limits, but for small tick ones it certainly increases the predictive power of our imbalance indicator to take more than one tick into account. Since the predictive power of the imbalance is outside the scope of the paper, we will stop here the discussion.

**Empirical evidences on the predictive power of the imbalance.** Figure II.2.a shows the imbalance (1) on the  $x$ -axis and the midprice move after 50 trades on the  $y$ -axis. In Figure II.2.a, we recover that the imbalance is highly positively correlated to the price move after 50 trades. Figure II.2.b shows the distribution of the imbalance just before a change in the order book state. In Figure II.2.b, agents are highly active at extreme imbalance values. People become highly active at extreme imbalances because they identify a profit opportunity to catch or at the opposite an adverse selection effect to avoid. Another explanation may come from the predictive power of the imbalance (see Figure II.2.a). In fact, participants start to anticipate the next price move when the imbalance signal becomes strong while they are inactive when they have no view on the next price move (i.e. the signal imbalance is weak).

---

<sup>3</sup>For a focus on tick size, see [60].

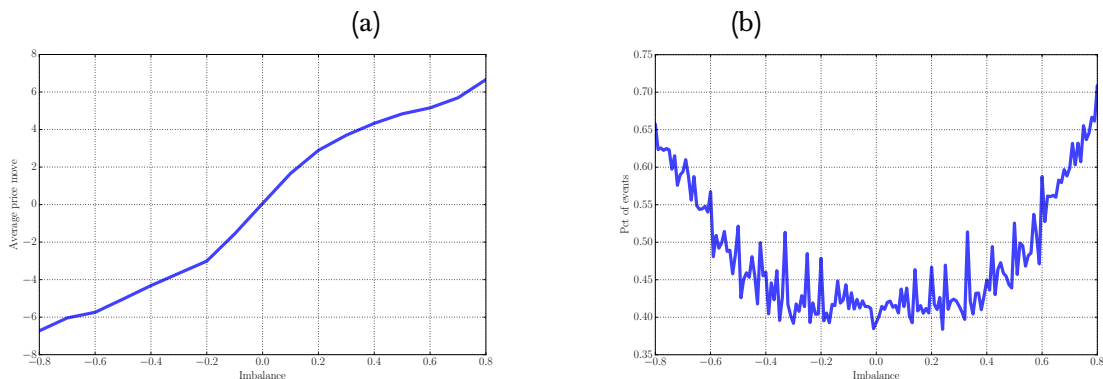


Figure I.1 – (a) The predictive power of the imbalance on stock Astra Zeneca: imbalance (just before a trade) on the  $x$ -axis and the expected price move (during the next 50 trades) on the  $y$ -axis. (b) distribution of the imbalance just before a trade. From the 2013-01-02 to the 2013-09-30 (accounts for 376,672 trades).

This paper provides a stochastic control framework to post limit orders using the information contained in the order book imbalance. In such context, we will call  $P_{+\infty}(t)$  the *microprice seen from  $t$*  and defined it such that

$$P_{+\infty}(t) = \lim_{\delta t \rightarrow +\infty} \mathbb{E}(P_{t+\delta t} | P_t, \text{Imb}_t).$$

## 2.3 Stylized fact 2: Agent’s decisions depend on the order book liquidity

### 2.3.1 Agent’s decisions depend on the order book imbalance

We expect some market participants to invest in access to data and technology to take profit of the informational content of the order book imbalance. A very simple way to test this hypothesis is to look at the order book imbalance just before a transaction with a limit order for a given class of participant. We will focus on three classes of *agents* (i.e. market participants): Global Investment Banks, High Frequency Participants (HFP), and Regional Investment Banks or Brokers. Table I.2 provides descriptive statistics on these classes of participants in the considered database.

We focus on limit orders since information processing, strategy and latency play a more important role for such orders than for market orders (market orders can be sent blindly, just to finish a small metaorder or to cope with metaorders late on schedule, see [145] for elaborations on brokers’ trading strategies).

Order type	Participant type	Order side	Avg. Imbalance	Nbe of events
Limit	Global Banks	Sell	0.35	62,111
		Buy	-0.38	63,566
	HFP	Sell	0.32	52,315
		Buy	-0.33	46,875
	Instit. Brokers	Sell	0.57	6,226
		Buy	-0.52	4,646

Table I.2 – Descriptive statistics for our three classes of agent. AstraZeneca (2013-01 to 2013-09).

For the following charts, we use labelled transactions from NASDAQ-OMX<sup>4</sup> and thanks to timestamps (and matching of prices and quantities) we synchronize them with order book data (recorded from direct feeds by Capital Fund Management). It enables us to snapshot the sizes at first limits on NASDAQ-OMX just before the transaction.

For a given participant (i.e. *agent*)  $a$  the quantity at the best bid (respectively best ask) is  $Q_\tau^{Bid}(a)$  (resp.  $Q_\tau^{Ask}(a)$ ) just before a transaction at time  $\tau$  involving a limit order owned by  $a$ . We write  $Q_\tau^{same}(a) := Q_\tau^{Bid}(a)$  (resp.  $Q_\tau^{opposite}(a) := Q_\tau^{Ask}(a)$ ) for a buy limit order and  $Q_\tau^{same}(a) := Q_\tau^{Ask}(a)$  (resp.  $Q_\tau^{opposite}(a) := Q_\tau^{Bid}(a)$ ) for a sell limit order. We normalize the quantities by the best opposite to obtain  $\rho_\tau(a) = (Q_\tau^{same}(a) - Q_\tau^{opposite}(a)) / (Q_\tau^{same}(a) + Q_\tau^{opposite}(a))$ . It is then easy to average over the transactions indexed by timestamps  $\tau$  to obtain an estimate of this expected ratio for one class of agent:

$$R(a) = \frac{1}{\text{Card}(\mathcal{F})} \sum_{\tau \in \mathcal{F}} \rho_\tau(a), \quad \lim_{\text{Card}(\mathcal{F}) \rightarrow +\infty} R(a) = \mathbb{E}_\tau \left( \frac{Q_\tau^{same}(a) - Q_\tau^{opposite}(a)}{Q_\tau^{opposite}(a) + Q_\tau^{opposite}(a)} \right).$$

It is even possible to control a potential bias by using the same number of buy and sell executed limit orders to compute this “neutralized” average:

$$R'(a) = \frac{1}{\text{Card}(\mathcal{F}(\text{buy}))} \sum_{\tau \in \mathcal{F}(\text{buy})} \rho_\tau(a) + \frac{1}{\text{Card}(\mathcal{F}(\text{sell}))} \sum_{\tau \in \mathcal{F}(\text{sell})} \rho_\tau(a).$$

Figure I.2.a shows the average state of the imbalance (via some estimates of  $R'(a)$ , on AstraZeneca from January 2013 to August 2013) for each class of agents (see Table I.3, Table I.4, Table I.5 for some lists of NASDAQ-OMX memberships used to identify agents classes). One can see the state of the imbalance is different within each class:

- Institutional brokers accept a transaction when the imbalance is highly negative, i.e. they buy using a limit order while the price is going down. It generates a large adverse selection: had they wait a little more, the price would have been cheaper. They make this choice because they have to buy fast for risk management reasons.

<sup>4</sup>For each transaction, we have a buyer ID, seller ID, a size, a price and a timestamp.

## I. Limit order strategic placement with adverse selection risk and the role of latency

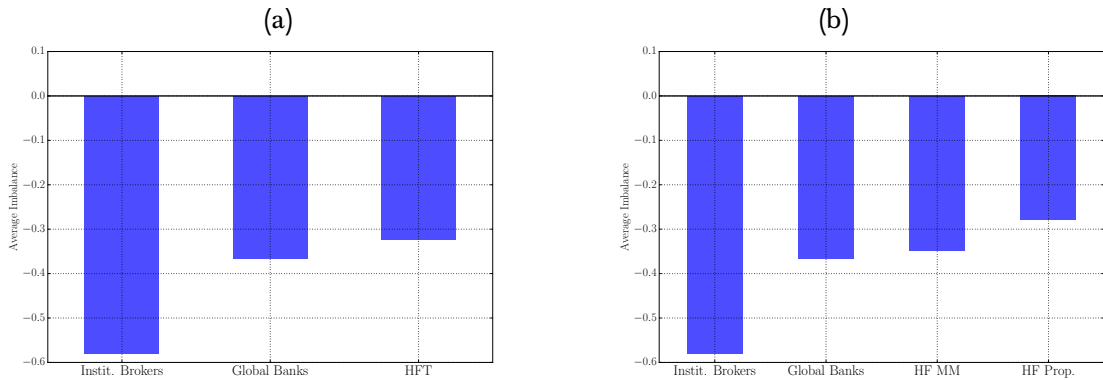


Figure I.2 – Comparison of neutralized order book Imbalance  $R'(a)$  at the time of a trade via a limit order (a) for institutional brokers, global investment banks and High Frequency Participants. (b) Shows a split of HFP between market makers and proprietary traders. Data are the ones of AstraZeneca (2013-01 to 2013-09).

- High Frequency Participants (HFP) accept a transaction when their imbalance is approximately the half of the Institutional brokers one. For sure they look more at the order book state before taking a decision. Moreover, they can be more opportunistic (i.e. ready to wait the perfect moment) instead of being lead by urgency considerations. If we split HFP between more market making-oriented ones and proprietary trading ones on Figure I.2.b we see
  - market makers (probably for inventory reasons: they have to alternate buys and sells), accept to trade when the imbalance is more negative than the average of HFP. They are probably paid back from this adverse selection by bid-ask spread gains (see [119]);
  - proprietary traders are the most opportunistic participants of our panel, leading them to have the less intense imbalance when they trade via limit orders: they seem to be the ones that suffer the less from adverse selection.
- Global Investment banks are in between. Three reasons may explain such behaviour: first, their activity is a mix of client execution and proprietary trading (hence we perceive the imbalance when they accept a trade as an average of the two categories); second, they have specific strategies to accept transactions via limit orders; third, they invest a little less than HFP in low latency technology, but more than institutional brokers.

The main effect to note is that each class of agents seems to exploit differently the state of the order book before accepting or not a transaction.

### 2.3.2 The added value of imbalance for market participants

Since we know that classes of agents take into account differently the state of the order book in their decision making process, we may ask ourselves what could be the value of such a “*high frequency market timing*”? We attempt to measure this value with a combination of NASDAQ-OMX labelled transactions and our synchronized market data. To do this, we compute the midprice move immediately before and after a class of participant  $a$  accepts to transact via a limit order as follows:

$$\Delta P_{\delta t}^{mid}(\tau, a) = \frac{P_{\tau+\delta t}^{mid} - P_{\tau}^{mid}}{\bar{\psi}} \cdot \epsilon_{\tau}(a),$$

where  $\epsilon_{\tau}(a)$  is the “sign” of the transaction (i.e. +1 for a buy and -1 for a sell) and  $\bar{\psi}$  is the average bid-ask spread on the considered stock. A “price profile<sup>5</sup>” around a trade is the averaging of this price move as a function of  $\delta t$  (between -5 minutes and +5 minutes); it is an estimate of the “expected price profile” around a trade:

$$p_a(\delta t) = \frac{1}{\text{Card}(\mathcal{F})} \sum_{\tau \in \mathcal{F}} \Delta P_{\delta t}^{mid}(\tau, a), \quad \lim_{\text{Card}(\mathcal{F}) \rightarrow +\infty} p_a(\delta t) = \mathbb{E}_{\tau} \Delta P_{\delta t}^{mid}(\tau, a).$$

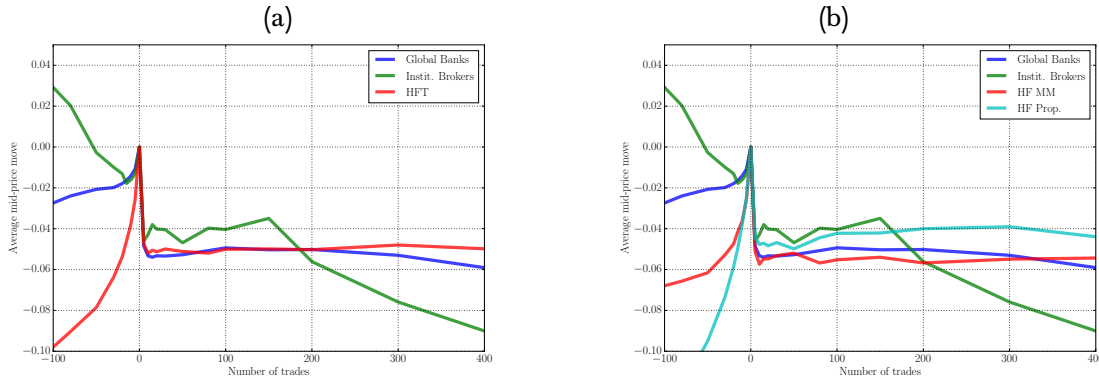


Figure I.3 - Midprice move relative to its position when a limit order is executed for (a) a High Frequency Market Maker, an investment bank, and an institutional broker; (b) makes the difference between HF market makers and HF proprietary traders. AstraZeneca (2013-01 to 2013-09).

Figure I.3.a and I.3.b show the price profiles of our three classes of participants, exhibiting real differences between them. First of all, it confirms the conclusions we draw from Figure I.2. Since it is always interesting to have a look at dynamical measures of liquidity (see [110] for a defense of the use of more dynamical measures of liquidity instead of plain averages):

<sup>5</sup>Note this “price profiles” are now used as a standard way to study the behaviour of high frequency traders in academic papers, see for instance [42] or [31].



- It is clear that Institutional brokers (green line) are buying while the price is going down. Had they bought later, they would have obtained a cheaper price. As underlined early they probably do it by purpose: they can have urgency reasons or they are using a “trading benchmark” that does pay more attention to the executed volume than to the execution price (see [145, Chapter 3] for details about brokers’ benchmarks).
- We can see that the difference between High Frequency Participants (HFP) and Global investment banks comes from the price dynamics *before the trade via a limit order*: for Investment banks the price is more or less stable before the execution and goes down when the limit order is executed. For HFP the price clearly go up *before they bought with a limit order*. It implies that they inserted their limit order shortly before the trade. In our framework we will see how cancelling and reinserting limit orders can be a way to implement an optimal strategy.
- On Figure I.3.b we see the difference between HF market makers and HF proprietary traders: the latter succeed in inserting buy (resp. sell) limit orders and obtaining transaction while the price is clearly going up (resp. down). After the trade, one can read a second difference between them and HF market makers: proprietary traders suffer from less adverse selection (the cyan curve is a little higher than the red one).

These charts show that there is a value in taking liquidity imbalance into account. In Section 4, we show the added of controlling a limit order by exploiting the order book imbalance.

**The role of latency.** Without a fast enough access to the servers of exchanges, a participant could know the best action to perform (insert or cancel a limit order), but not be able to implement it before an unexpected transaction. Since low latency has a cost, some participants may decide to ignore this information and do not access to fast market feeds, order books states, etc.

In the following sections we will not only provide a theoretical framework to “optimally” exploit the order book dynamics for limit order placement, but also study its sensitivity to latency, showing how latency can destroy the added value of understanding the order book dynamics. In our theoretical framework, we can explore situations in which the participant know the best action but cannot implement it on time.

### 3 The Dynamic Programming Principle Applied to Limit Order Placement

To set a discrete time framework, we focus on the simple case (but complex enough in terms of modelling) of one atomic quantity  $q_\epsilon$  to be executed in  $T_f$  units of time (can be order book events, trades, or seconds). It will be a buy order, but it is straightforward for the reader to transpose our results for a sell order.

From zero to  $T_f$  the trader (or software) in charge of this limit order can: cancel it (i.e. remove it from the order book) and insert it at the top of the bid queue, or do nothing. If the trader did not obtain an execution thanks to its optimal posting policy at  $T_f$ , we force him to cancel his order (if any) and send a market order to obtain a trade. For simplicity, we consider a model adapted to large tick stocks (i.e. for which the spread equals to one tick). However, our construction can be adapted to a small tick stock by enlarging the control space. For example, we can add the possibility to post a limit order in between the best bid and the best ask. Since we consider a small size order, sending a limit order in between first limits highly increases the adverse selection risk. Consequently, we neglect, as a first approximation, such a control in our model. In a more general framework, other limits should be taken into account and more controls can be considered.

### 3.1 Formalisation of the Model

Let  $q_\epsilon$  be a small limit order inserted at the first bid limit of the order book. The order book state is modelled by

$$U_t^\mu = \left( Q_t^{Before,\mu}, Q_t^{After,\mu}, Q_t^{Opp,\mu}, P_t^\mu \right),$$

where  $Q^{Before,\mu}$  is the quantity having priority on the order  $q_\epsilon$ ,  $Q^{After,\mu}$  is the quantity posted after the order  $q_\epsilon$ ,  $Q^{Opp,\mu}$  is the first opposite limit quantity,  $P_t^\mu$  is the mid price,  $\mu$  is the control of the agent and  $t$  is the current time. For simplicity, we neglect the quantity  $q_\epsilon$ :

$$Q^{Same,\mu} = Q^{Before,\mu} + Q^{After,\mu}.$$

**Limit order book dynamics.** Since we do not differentiate cancellations and market orders, the order book dynamics can be modelled by four counting processes (see Figure IV.1):

- $N_t^{Opp,+}$  (respectively  $N_t^{Same,+}$ ) with an intensity  $\lambda^{Opp,+}(Q^{Opp}, Q^{Same})$  (resp.  $\lambda^{Same,+}(Q^{Opp}, Q^{Same})$ ) representing the inserted orders in the opposite limit (resp. same limit).
- $N_t^{Opp,-}$  (resp.  $N_t^{Same,-}$ ) with an intensity  $\lambda^{Opp,-}(Q^{Opp}, Q^{Same})$  (resp.  $\lambda^{Same,-}(Q^{Opp}, Q^{Same})$ ) representing the cancelled orders in the opposite limit (resp. same limit).

In this model, these four counting processes depend only on quantities at first limits. At each event time, an atomic quantity  $\mathbf{q}$  is added or canceled from the order book. Moreover, we assume the following bid-ask symmetry relation:

$$\begin{cases} \lambda^{Opp,+}(Q^{Opp}, Q^{Same}) = \lambda^{Same,+}(Q^{Same}, Q^{Opp}), \\ \lambda^{Opp,-}(Q^{Opp}, Q^{Same}) = \lambda^{Same,-}(Q^{Same}, Q^{Opp}). \end{cases}$$

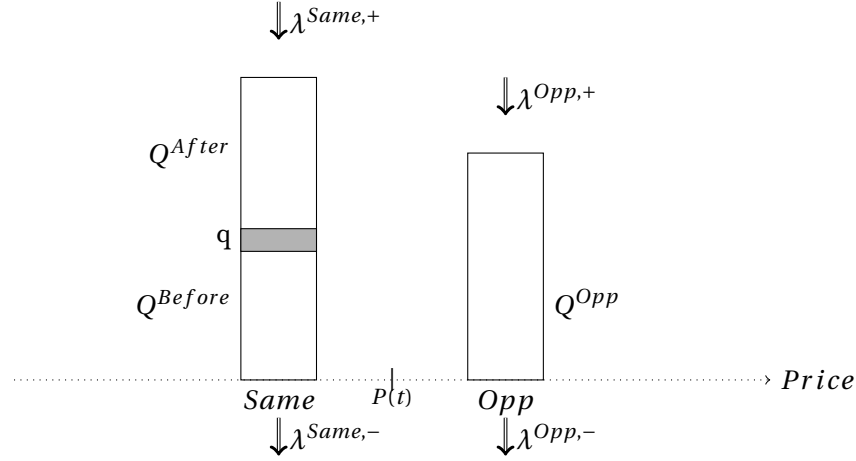


Figure I.4 – Diagram of flows affecting our order book model.

Hence, the size of the first limits can be written as follows when none of them is negative:

$$\begin{cases} Q_t^{Opp,\mu} &= Q_{t^-}^{Opp,\mu} + q\Delta N_t^{Opp,+} - q\Delta N_t^{Opp,-}, \\ Q_t^{Before,\mu} &= (Q_{t^-}^{Before,\mu} - q\Delta N_t^{Same,-})1_{Q_{t^-}^{Before,\mu} \geq q}, \\ Q_t^{After,\mu} &= Q_{t^-}^{After,\mu} + q\Delta N_t^{Same,+} - q1_{0 \leq Q_{t^-}^{Before,\mu} < q} \Delta N_t^{Same,-}. \end{cases} \quad (2)$$

**What happens when  $Q^{After,\mu}$ ,  $Q^{Before,\mu}$  or  $Q^{Opp,\mu}$  is totally consumed:** First of all, we neglect the probability that at least two of these three events happen simultaneously. When one of the two queues fully depletes, we assume that the price moves in its direction. Thus we introduce a *discovered quantity*  $Q^{Disc}$  to replace the deleted first queue and an *inserted quantity*  $Q^{Ins}$  to be put in front of  $Q^{Disc}$ . These quantities are random variables and their law are conditioned by the order book state before the depletion. In detail:

1. When  $Q_t^{Opp,\mu} = 0$ . The price increases by one tick (keep in mind for a buy order, the *opposite* is the ask side). Then, we discover a new opposite limit and a new bid quantity is inserted into the bid-ask spread (on the bid side) by other market participants (see Figure I.5). It reads

$$\begin{cases} Q_t^{Opp,\mu} &= Q^{Disc}(Q_{t^-}^{Opp,\mu}, Q_{t^-}^{Same,\mu}), \\ Q_t^{Before,\mu} &= Q^{Ins}(Q_{t^-}^{Opp,\mu}, Q_{t^-}^{Same,\mu}), \\ Q_t^{After,\mu} &= 0. \end{cases}$$

$Q^{Disc}$  is the “*discovered quantity*” and  $Q^{Ins}$  the “*inserted quantity*”.

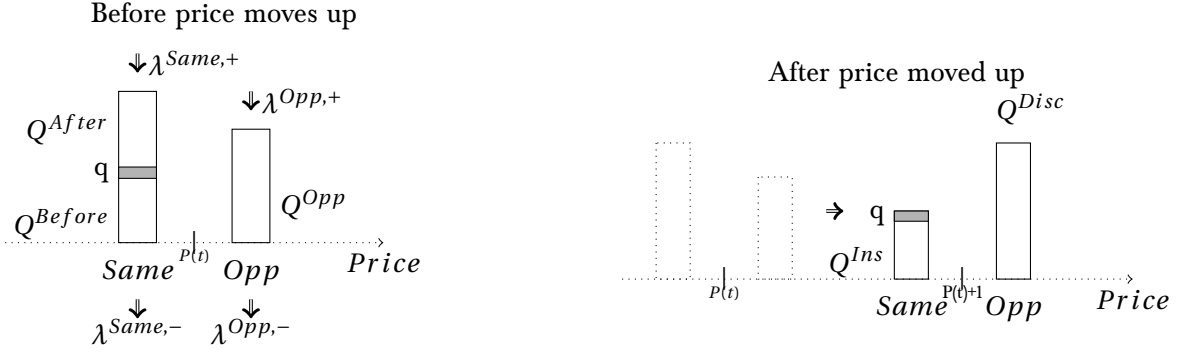


Figure I.5 – Diagram of an upward price change for our model.

2. When  $Q_t^{Before,\mu} = 0$ . The limit order is executed. This case is considered in Equation (2).
3. When moreover  $Q_t^{After,\mu} = 0$ . The price decreases by one tick. Then, we discover a new quantity on the bid side and market makers insert a new quantity on the opposite side:

$$\begin{cases} Q_t^{Opp,\mu} &= Q^{Ins}(Q_t^{Opp,\mu}, Q_t^{Same,\mu}), \\ Q_t^{Before,\mu} &= Q^{Disc}(Q_t^{Opp,\mu}, Q_t^{Same,\mu}), \\ Q_t^{After,\mu} &= 0. \end{cases}$$

If the limit order was in the order book: it has been executed. Otherwise the price moves down and the trader has the opportunity to reinsert a limit order on the top of  $Q^{Disc}$  (see Figure I.6 for a diagram).

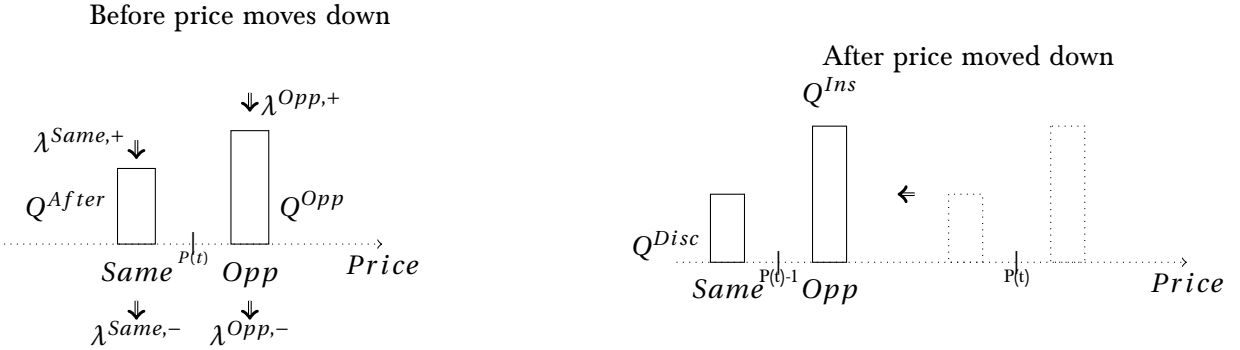


Figure I.6 – Diagram of a downward price move in our model.

**The control.** We consider two types of control  $C = \{s, c\}$ :

- $s$  (like *stay*): stay in the order book.
- $c$  (like *cancel*): cancel the order and wait for a better order book state to reinsert it at the top of  $Q^{same}$  ( $Q^{bid}$  for our buy order). This control is essentially used to avoid adverse selection, i.e. avoid to buy just before a price decrease.

**Optimal control problem.** We fix a finite horizon time  $T_f < \infty$  and we want to compute:

$$V_{T_f}(0, U) = \sup_{\mu} \mathbb{E}[\Delta P_{\infty}^{\mu}].$$

Where:

- $U = (q^{before}, q^{after}, q^{opp}, p)$  is the initial state of the order book.
- $T_{Exec}^{\mu} = \inf\{t \geq 0, s.t. Q_t^{Before, \mu} < q, \mu_t = s\} \wedge T_f$  represents the first execution time. Once the order executed, the order book is frozen.
- $\Delta P_{\infty}^{\mu} = \lim_{t \rightarrow \infty} (P_t^{\mu} - P_{T_{Exec}^{\mu}}^{Exec, \mu})$  represents the gain of the trader, where the execution price  $P_t^{Exec, \mu}$  satisfy  $P_t^{Exec, \mu} = P_t^{\mu} - \frac{1}{2}$  when the limit order is executed before  $T_f$  and  $P_t^{Exec, \mu} = P_t^{\mu} + \frac{1}{2}$  otherwise. Indeed, if at  $T_f$  the order hasn't been executed, we cross the spread to guarantee execution.

**Choice of a benchmark.** We compare the value of the obtained shares at  $t$  to *its expected value at infinity*, i.e.  $\mathbb{E}(P_{+\infty}(t)|\text{Imb}_t)$  since it is not attractive to buy at the best bid if we expect the price to continue to go down. Indeed, it is possible to expect a better future price thanks to the observed imbalance. It thus induces an *adverse selection cost* in the framework. This is not a detail since the *trader will have no incentive to put a limit order at the top of a very small queue if the opposite side of the book is large*. We will see in Section 4 using empirical evidences that this is a realistic behaviour. Such a behaviour cannot be captured by other linear frameworks like [124]. To solve the control problem, we introduce in the next section a discrete time version of the initial problem whose value function can be computed numerically. In this discrete framework,  $\lim_{t \rightarrow \infty} (P_t^{\mu} - P_{T_{Exec}^{\mu}}^{Exec, \mu})$  is computed using the imbalance.

### 3.2 Discrete time framework

Let  $\Delta_t$  be the time step and  $T_f$  be the final time. Let  $t_0 = 0 < t_1 < \dots < t_{f-1} < t_f = T_f$  different instants at which the order book is observed, such that  $t_n = n\Delta_t$  for all  $n \in \{0, 1, \dots, f\}$ . Under the assumption that between two consecutive instants  $t_n$  and  $t_{n+1}$  for all  $n \in \{0, 1, \dots, f-1\}$ , only five cases can occur:

- 1 unit quantity is added at the bid side;
- 1 unit quantity is consumed at the bid side;
- 1 unit quantity is added at the opposite side;

- 1 unit quantity is consumed at the opposite side;
- nothing happens.

We neglect the situation where at least two cases occur during the same time interval (the probability of such conjunctions are of the orders of  $\lambda^2$ , hence our approximation remains valid as far as  $(\lambda dt)^2$  is small compared to  $\lambda dt$ ).

**Framework 1** (Our setup in few words.). *In short, our main assumptions are:*

- *only one limit order of small quantity  $q_\epsilon$  is controlled, it is small enough to have no influence on order book imbalance;*
- *decrease of queue sizes at first limits is caused by transactions only (i.e. no difference between cancellation and trades);*
- *queues decrease or increase by one quantity only;*
- *the intensities of point processes (including the ones driving quantities inserted into the bid-ask spread and driving the quantity discovered when a second limit becomes a first limit) are functions of the quantities at best limits only;*
- *no notable conjunction of multiple events.*

We introduce the following Markov chain  $U_n^\mu = (Q_n^{Before,\mu}, Q_n^{After,\mu}, Q_n^{Opp,\mu}, P_n^\mu, Exec_n)$  where:

- $Q_n^{Before,\mu}$  is the  $Q^{Before,\mu}$  size at time  $t_n$  that takes value in  $\mathbb{N}$ .
- $Q_n^{After,\mu}$  is the  $Q^{After,\mu}$  size at time  $t_n$  that takes value in  $\mathbb{N}$ .
- $Q_n^{Opp,\mu}$  is the  $Q^{Opp,\mu}$  size at time  $t_n$  that takes value in  $\mathbb{N}$ .
- $P_n^\mu$  is the mid price at time  $t_n$ .
- $Exec_n$  is an additional variable taking values in  $\{-1, 0, 1\}$ .  $Exec_n$  equals to 1 when the order is executed at time  $t_n$ , 0 when the order is not executed at time  $t_n$  and -1 (a “cemetery state”) when the order has been already executed before  $t_n$ . We set  $Exec_0 = 0$ .

In the same way, we define  $N_n^{Same,+}$ ,  $N_n^{Same,-}$ ,  $N_n^{Opp,+}$  and  $N_n^{Opp,-}$  as the values of the counting processes  $N_t^{Same,+}$ ,  $N_t^{Same,-}$ ,  $N_t^{Opp,+}$  et  $N_t^{Opp,-}$  at time  $t_n$ . The transition probabilities of the Markov chain  $U_n^\mu$  are detailed in Appendix I.A.

**The terminal constraint.** The microprice  $P_{\infty,n} \approx \mathbb{E}[P_{\infty} | \mathcal{F}_n]$  is defined such as:

$$P_{\infty,n} = F(Q_n^{Opp}, Q_n^{Same}, P_n) = P_n + \frac{\alpha}{2} \cdot \frac{Q_n^{Same} - Q_n^{Opp}}{Q_n^{Opp} + Q_n^{Same}}, \quad \forall n \in \{0, 1, \dots, f\}.$$

Where  $\mathcal{F}_n$  is the filtration associated to  $U_n$  such that  $\mathcal{F}_n = \sigma(U_n, k \leq n)$  and  $\alpha$  is a parameter that represents the sensitivity of future prices to the imbalance. The execution price  $P_{Exec,n}$  is defined  $\forall n \in \{0, 1, \dots, f\}$  such that:

$$P_{Exec,n} = \begin{cases} P_n + \frac{1}{2}, & \text{when } Exec_n = 0, \\ P_n + \frac{1}{2}, & \text{when } Exec_n \in \{-1, 1\} \text{ and } P_{n+1} - P_n \neq 0, \\ P_n - \frac{1}{2}, & \text{when } Exec_n \in \{-1, 1\} \text{ and } P_{n+1} - P_n = 0. \end{cases}$$

Let  $k_0$  be the execution time:  $n_0 = \inf(n \geq 0, Exec_n = 1) \wedge f$ . Then, the terminal valuation can be written

$$Z_{k_0} = P_{\infty, n_0} - P_{Exec, n_0}. \quad (3)$$

Let  $\mathcal{U}$  be the set of all progressively measurable processes  $\mu := \{\mu_k, k < f\}$  valued in  $\{s, c\}$ . This problem can be written as a stochastic control problem:

$$V_{U_0, f} = \sup_{\mu \in \mathcal{U}^{U_0, \mu}} \mathbb{E} [Z_{n_0}] = \sup_{\mu \in \mathcal{U}^{U_0, \mu}} \mathbb{E} \left[ \sum_{i=1}^{f-1} g_i(U_i, \mu_i) + g_f(U_f) \right],$$

where  $g_i(U_i, \mu_i) = Z_i$  when  $Exec_i = 1$  and  $\mu_i = s$  and 0 otherwise for all  $i \in \{1, \dots, f-1\}$ . We write  $g_f(U_f) = Z_f$  when  $Exec_f \in \{0, 1\}$  and 0 otherwise. We want to compute  $V_{U_0, f} = \sup_{\mu \in \mathcal{U}^{U_0, \mu}} \mathbb{E} (Z_{k_0})$  using dynamic programming algorithm:

$$\begin{cases} G_f = Z_f, \\ G_n = \max(P^c G_{n+1}, P^s G_{n+1}), \quad \forall n \in \{0, 1, \dots, f-1\}, \end{cases} \quad (4)$$

where  $P^c$  and  $P^s$  are the transition probabilities of the Markov chain  $U_n^\mu$  when the initial decisions are respectively  $c$  and  $s$ . Equation (4) provides an explicit forward-backward algorithm that can be solved numerically:

- *Step 1 Forward simulation:* starting from an initial state  $u$ , we simulate all the reachable states during  $f$  periods.
- *Step 2 Backward simulation:* at the final period  $f$ , we can compute  $G_f$  for each reachable state. Then, using the backward equation (4), we can compute, recursively,  $G_i$  knowing  $G_{i+1}$  to get  $G_0$ .

For more details about the forward-backward algorithm see Appendix I.A.

## 4 A Qualitative Understanding

In this section, we present and comment the simulation results. We are going to compare two situations:

- The first one called **(NC)** corresponds to the case when no control is adopted (i.e. we always stay in the order book and “join the best bid” each time it changes).
- The second one called **(OC)** corresponds to the optimal control case: controls "c" and "s" are considered.

Moreover, our simulation results are given for two different cases:

- **Framework (CONST):** intensities of insertion and cancellation are constant:  $\lambda_k^{Same,+} = \lambda_k^{Opp,+} = 0.06$  and  $\lambda_k^{Same,-} = \lambda_k^{Opp,-} = 0.5 \forall k \in \{0, 1, \dots, f\}$ . Under (CONST), the inserted quantities  $Q^{Ins}$  and discovered quantities  $Q^{Disc}$  are constant too.
- **Framework (IMB):** intensities of cancellation and insertion are functions of the imbalance such as  $\forall k \in \{0, 1, \dots, f\}$ :

$$\begin{aligned} \lambda_k^{Opp,+} \left( Q_k^{Opp}, Q_k^{Same} \right) &= \lambda_k^{Same,+} \left( Q_k^{Same}, Q_k^{Opp} \right) = \lambda_0^+ + \beta^+ \frac{Q_k^{Opp}}{(Q_k^{Opp} + Q_k^{Same})}, \\ \lambda_k^{Opp,-} \left( Q_k^{Opp}, Q_k^{Same} \right) &= \lambda_k^{Same,-} \left( Q_k^{Same}, Q_k^{Opp} \right) = \lambda_0^- + \beta^- \frac{Q_k^{Same}}{(Q_k^{Opp} + Q_k^{Same})}, \end{aligned}$$

where  $\lambda_0^\pm$  refer to both basic insertion and cancellation intensities and  $\beta^\pm$  are predictability parameters representing the sensitivity of order flows to the imbalance. Moreover, under (IMB), inserted and discovered quantities are computed in the following way:

- **When  $Q_k^{Opp}$  is totally consumed**, we set  $Q_k^{Disc} = \lceil q_0^{disc} + \theta_{disc} \cdot Q_k^{Same} \rceil$  and  $Q^{Ins} = \lceil q_0^{ins} + \theta_{ins} \cdot Q_k^{Same} \rceil$ . Where  $\theta_{disc}$  and  $\theta_{ins}$  are coefficients associated to liquidity and  $\lceil \cdot \rceil$  is the upper rounding. The quantities  $q_0^{disc}$  and  $q_0^{ins}$  are the basic discovered and inserted quantities.
- Similarly **when  $Q_k^{Same}$  is totally consumed**, we set  $Q^{Disc} = \lceil q_0^{disc} + \theta_{disc} \cdot Q_k^{Opp} \rceil$  and  $Q^{Ins} = \lceil q_0^{ins} + \theta_{ins} \cdot Q_k^{Opp} \rceil$ .

This kind of relations is compatible with empirical findings of [90] and different from [55] in which  $Q^{Disc} = Q^{Ins}$  is independent of the liquidity imbalance.

### 4.1 Numerically Solving the Control Problem

#### 4.1.1 Anticipation of Adverse Selection

The cancellation is used by the optimal strategy to avoid adverse selection. For instance, when the quantity on the same side is extremely lower than the one on the opposite side, it is expected to cancel the order to wait for a better future opportunity. The optimal control



takes in consideration this effect and cancels the order when such a high adverse selection effect is present.

We keep the same notations of Section 3. Let  $\mu := \{\mu_k, k < f\}$  a control, we define  $\mathbb{E}_{U_0, \mu}(\Delta P | \text{Exec}) = \mathbb{E}_{U_0, \mu}(Z_{k_0})$ . The quantity  $\mathbb{E}_{U_0, \mu}(\Delta P | \text{Exec})$  depends on the control  $\mu$ , the initial state of the order book  $U_0$  and the terminal period  $f$ . This variable can be directly computed by a forward algorithm that visits all the possible states of the Markov chain  $U_n^\mu$ . For more details about the transition probabilities of the Markov chain  $U_n^\mu$  see Appendix I.A. The quantity  $\mathbb{E}_{U_0, \mu}(\Delta P | \text{Exec})$  is interesting since it corresponds to the quantity to maximize in our optimal control problem and it represents as well the profitability/trade of an agent.

Let  $\mu^s$  the control where the agent always stays in the order book (i.e. NC) and  $\mu^*$  the optimal control (i.e. OC). Figure I.7.a represents the variation of  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec})$  and  $\mathbb{E}_{U_0, \mu^s}(\Delta P | \text{Exec})$  when the initial imbalance of the order book moves under (CONST). In Figure I.7.a, blue points are initial states where it is optimal to stay in the order book at the initial time (i.e.  $t = 0$ ) while red points are initial states where it is optimal to cancel the order at  $t = 0$ . Initial parameters are fixed such that  $\lambda^{Same,+} = \lambda^{Opp,+} = 0.06$ ,  $\lambda^{Same,-} = \lambda^{Opp,-} = 0.5$ ,  $\alpha = 4$ ,  $Q^{Disc} = 6$ ,  $Q^{Ins} = 4$ ,  $f = 20$ ,  $q = 1$  and  $P_0 = 10$ . Moreover, the initial imbalance values are obtained by varying  $Q_0^{Opp}$  from 2 to 12 and  $Q_0^{After}$  from 1 to 11 while  $Q_0^{Before}$  is kept constant equal to 1. Figure I.7.b is the analogous of Figure I.7.a but under the framework (IMB). In Figure I.7.b, initial parameters are fixed such that  $\lambda_0^+ = 0.06$ ,  $\lambda_0^- = 0.5$ ,  $\beta^+ = 0.075$ ,  $\beta^- = 0.25$ ,  $q_0^{disc} = 6$ ,  $q_0^{ins} = 2$ ,  $\theta_{disc} = 3$ ,  $\theta_{disc} = 0.5$ ,  $\alpha = 4$ ,  $f = 20$  and  $P_0 = 10$ . Similarly, initial imbalance values are obtained by varying  $Q_0^{Opp}$  from 2 to 12 and  $Q_0^{After}$  from 1 to 11 while  $Q_0^{Before}$  is kept constant equal to 1.

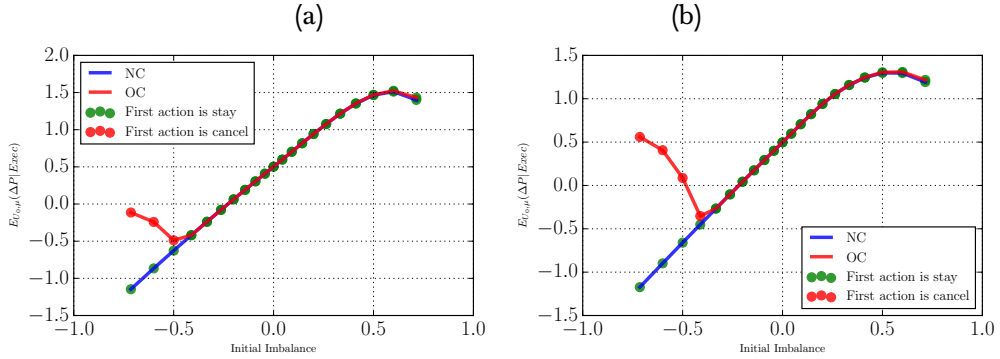


Figure I.7 – (a) (resp. (b))  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec})$  and  $\mathbb{E}_{U_0, \mu^s}(\Delta P | \text{Exec})$  when intensities are constant (CONST) (resp. (4)).

The **main effect** to note on these curves is the way the optimal control anticipates adverse selection. When imbalance is highly negative, we cancel first the order (red points) to take advantage from a better futur opportunity. We notice that under the framework (IMB), the

agent cancels earlier than under the framework (CONST) since more weights are given to cancellation events. This point is detailed in \*\*. Appendix I.C explains the downward slopes at the left of Figures I.7.a and I.7.b.

#### 4.1.2 Price Improvement comes from avoiding adverse selection

As expected, results obtained by the optimal control (OC) case are better than the ones of the non-controlled (NC) case: by cancelling and taking into account liquidity imbalance, one can be more efficient than just staying in the order book. Figure I.8.a shows the variation of the price improvement (resp.  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec}) - \mathbb{E}_{U_0, \mu^s}(\Delta P | \text{Exec})$ ) when the initial imbalance moves, under both frameworks (CONST) and (IMB). We kept the same initial parameters of Figures I.7.a and I.7.b.

Similarly, Figure I.8.b represents the variation  $\mathbb{E}_{U_0, \mu^*}(P_{k_0} | \text{Exec}) - \mathbb{E}_{U_0, \mu^s}(P_{k_0} | \text{Exec})$  when the initial imbalance moves, under both frameworks (CONST) and (IMB). The quantity  $P_{k_0}$  is the mid price at the execution time  $k_0$ . We kept the same initial parameters of Figures I.7.a and I.7.b.

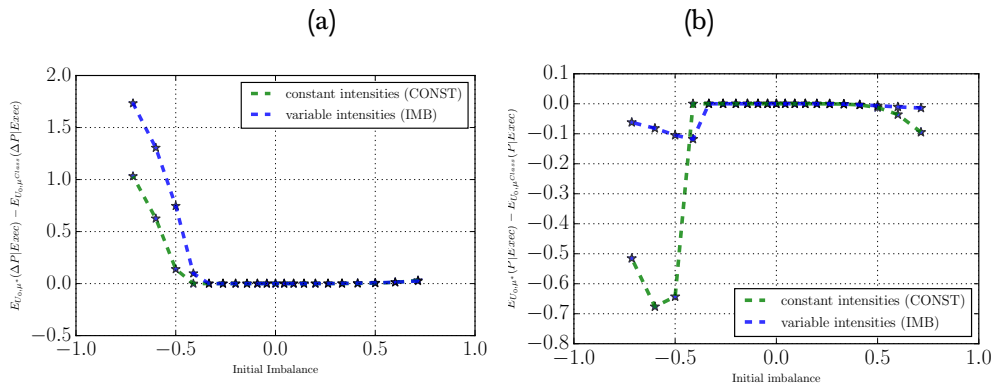


Figure I.8 – (a)  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec}) - \mathbb{E}_{U_0, \mu^s}(\Delta P | \text{Exec})$  move relative to initial imbalance under (CONST) and (IMB). (b)  $\mathbb{E}_{U_0, \mu^*}(P_{k_0} | \text{Exec}) - \mathbb{E}_{U_0, \mu^s}(P_{k_0} | \text{Exec})$  move relative to initial imbalance under (CONST) and (IMB).

Figure I.8 deserves the following comments. As expected, the optimal control provides better results than a blind “join the bid” strategy. In Figure I.8.a the price improvement is non-negative since our control maximizes  $\mathbb{E}_{U_0, \mu}(\Delta P)$ . When the initial imbalance is highly positive, the price improvement is close to 0 however when the initial imbalance is highly negative the price improvement becomes higher than 0 by avoiding adverse selection. Similarly, Figure I.8.b shows that the optimal strategy allows us to buy with a low average price when imbalance is highly negative by preventing from adverse selection.<sup>6</sup>

<sup>6</sup>Indirectly, maximizing  $\mathbb{E}_{U_0, \mu}(\Delta P | \text{Exec})$  leads to the minimization of the price.

### 4.1.3 Average Duration of Optimal Strategies

In brief, the optimal strategy aims to obtain an execution in the best market conditions (i.e. with a low adverse selection risk). It can be read on the average lifetime (i.e. "duration") of the strategy. Figure I.9.a compares the average strategy duration in both settings (NC) and (OC) when intensities are constant (CONST). We keep the same initial parameters of Figure I.7.a. Figure I.9.b is the analogous of Figure I.9.a but under the framework (IMB). Finally, Figure I.9.c shows the "stay ratio" (i.e. the proportion of stay decisions under the optimal strategy) under both settings (NC) and (OC) when intensities depend on the imbalance (IMB). Figure I.9.b and I.9.c are computed with the same initial parameters of Figure I.7.b.

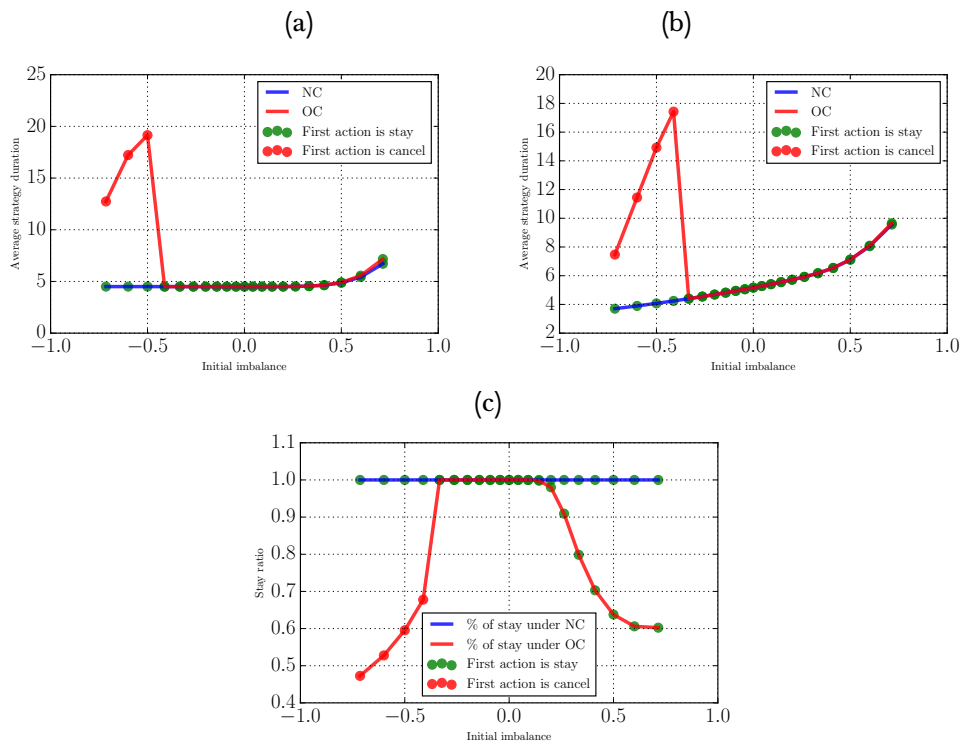


Figure I.9 – Average strategy duration as a function of (a) the initial imbalance under (CONST) and (b) the initial imbalance under (IMB). (c) Stay ratio as a function of the initial imbalance under (IMB).

In both Figures I.9.a and I.9.b, the average strategy duration of the optimal control is always higher than the non-optimal one. It is an expected result since the optimal control cancels the order and hence postpone the execution. Moreover, the algorithm cancels the order when high adverse selection is present (i.e. the imbalance is highly negative under IMB<sup>7</sup>). In such

<sup>7</sup> close to  $t = 0$ , the optimal strategy is free to cancel its limit order; but when  $T_f$  is close, it has to think about the cost of having to cross the spread in few steps.

case, the average strategy duration of the optimal control is strictly greater than the non-optimal one (see Figures I.9.a and I.9.b).

\*\* In Figure I.9.b, when intensities depend on the imbalance (IMB), the average strategy duration has an increasing trend. In fact, under (IMB), when imbalance is highly positive, more weights are given to events delaying the execution. For example, when imbalance is highly positive, the bid queue is a way larger than the opposite one. Then, the probability to obtain an execution on the bid side is low: that's why it is expected to wait more. Moreover, Figure I.9.c shows that the agent become more active when high adverse selection is present. Indeed, when the imbalance is negative (i.e. high adverse selection), the "stay ratio" decreases and consequently the "cancel ratio" increases.

#### 4.1.4 Influence of the Terminal Constraint

In this section, we want to shed light on two stylized facts:

1. the optimal strategy performs better under good market condition when there is more time left.
2. the optimal strategy becomes highly active close to the terminal time.

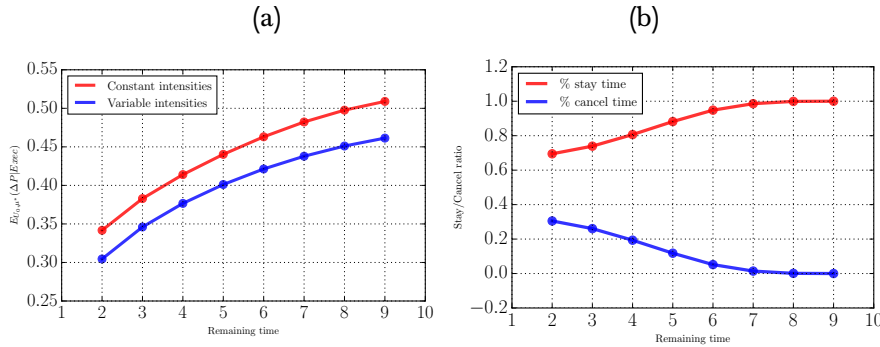


Figure I.10 – (a)  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec})$  move relative to remaining time under (CONST) and (IMB). (b) stay cancel ratio move relative to remaining time to maturity under (IMB).

Figure I.10.a compares the variation of  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec})$  as a function of the remaining time under frameworks (CONST) and (IMB). The initial imbalance is fixed equal to 0.5. Thanks to the Figure I.10.a, we can see that the more time remaining, the better for the optimal strategy. However, the concavity of the curve shows that the marginal performance  $\frac{\partial \mathbb{E}(\Delta P)_t}{\partial t}$  is decreasing. Moreover, Figure I.10.a shows also that  $\mathbb{E}_{U_0, \mu^*}(\Delta P | \text{Exec})$  may converge to a limit value when maturity time tends to infinity. Since the Markov chain  $U_n$  is ergodic (cf. [90]), we believe that this limit value is unique and independent of the initial state of the order book and could lead to an “almost ergodic” regime.

In Figure I.10.b, we present the percentage of times when the optimal strategy cancels its order and the percentage of times when it decides to stay in the order book as a function of remaining time under (CONST) and (IMB). The initial imbalance is fixed to 0.5. Thanks to Figure I.10.b, we conclude that it is optimal to be more active close to  $t = T_f$ . In Figures I.10.a and I.10.b, we keep the same initial parameters of Figures I.7.a and I.7.b

## 4.2 The Price of Latency

In Section 3, we define the Markov chain  $U_n^\mu$  that corresponds to a market participant enabled to change his control at each period. A slower participant will not react at each limit order book move. Hence, he can be modelled by the markov chain  $U_{\tau n}^\mu$  where  $\tau$  corresponds to a latency factor such as  $\tau \in \mathbb{N}^*$ . Using notations of previous sections, we define  $Z_{\tau,f}$  as the final constraint associated to the Markov chain  $U_{\tau n}^\mu$ . Thus, we define the latency cost of a participant with a latency factor  $\tau$  such as:

$$Latency_{U_0,f}(\tau) = V_{U_0,f} - V_{U_0,f,\tau}, \quad \forall \tau \in \mathbb{N}^*, \quad (5)$$

where  $V_{U_0,f,\tau} = \sup_{\mu \in \mathcal{U}} \mathbb{E}_{U_0,\mu}(Z_{\tau,f})$ . By adapting the same numerical forward-backward algorithm, the cost of latency can be computed numerically. This cost can be converted into a value: it is the value a trader should accept to pay in technology since he will be rewarded in term of performance.

Figures I.11.a and I.11.b show the variation of the latency cost with respect to the latency factor  $\tau$  under both frameworks (CONST) and (IMB) for different values of  $\alpha$ . The initial imbalance is fixed equal to 0.5 with an initial state  $Q_0^{After} = 2, Q_0^{Before} = 1$  and  $Q_0^{Opp} = 1$ . We keep the same initial parameters of Figure I.7.a and Figure I.7.b. Numerical results show:

- The latency cost increases with the latency factor  $\tau$  (cf. Figure I.11.a).
- The latency cost is higher when sensitivity to adverse selection increases (i.e.  $\alpha$  is big) (cf. Figures I.11.a and I.11.b).

Consequently, the relevance of exploiting a knowledge on liquidity is eroded by latency: being able to predict future liquidity-consuming flows is of less use if you one cannot cancel and reinsert his limit orders at each change of the order book state. For instance, when two agents act optimally according the same criterion, the faster will have more profits than the slower.

## 5 Conclusion

We have used NASDAQ-OMX labelled data to show how market participants accept or refuse transactions via limit orders as a function of liquidity imbalance. It is not an exhaustive study on this exchange from the north of Europe (we focus on AstraZeneca from January 2013 to September 2013). We first show that the order book imbalance has a predictive power on

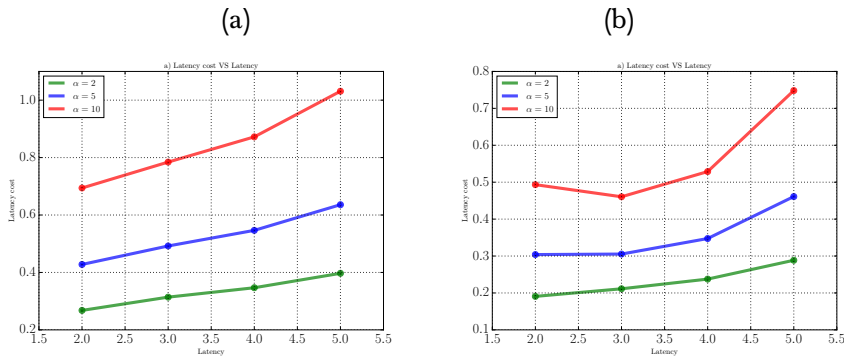


Figure I.11 – (a) Latency cost as a function of the latency factor  $\tau$  under (CONST). (b) Latency cost as a function of the latency factor  $\tau$  under (IMB) for different values of  $\alpha$ .

future mid price moves. We then focus on three types of market participants: Institutional brokers (IB), Global Investment Banks (GIB) and High Frequency Participants (HFP). Data show that IB accept to trade when the imbalance is more negative (i.e. they buy or sell while the price pressure is downward or upward) than GIB, themselves accepting a more negative imbalance than HFP. Moreover, when we split HFP between high frequency market makers and high frequency proprietary traders (HFPT), we see that HFPT achieve to buy via limit orders when the imbalance is very small. We complete this analysis with the dynamics of prices around limit orders execution, showing how strategically participants use their limit orders.

Then we propose a theoretical framework to control limit orders where the liquidity imbalance can be used to predict future price moves. Our framework includes potential adverse selection via a parameter  $\alpha$ . We use the dynamic programming principle to solve it. We show that our solutions have commonalities with our empirical findings.

In a last Section we show how the capability of exploiting the imbalance predictability using our optimal control decreases with latency: the trader has less time to put in place sophisticated strategies, hence he cannot take profit of any strategy gain.

The difficult point of using limit orders is adverse selection: if the price has chances to go down the probability to be filled is high but it is better to postpone execution to get a better price. However, when a market participant cancels his limit order (to postpone execution), he takes the risk to never obtain a transaction: the reinsertion of the order will be on the top of the bid queue. Furthermore, the price may go up again before the execution of the limit order. Our framework includes all these effects and our optimal strategy makes the choice between waiting in the queue or leaving it when the probability that the price goes down becomes too high. To do this, the position of the limit order in the queue is taken into account by our controller.

This leads to a quantitative way to understand market making and latency: if a market maker is fast enough, he will be able to play this insert, cancel and re-insert game to react to his observations of liquidity imbalance. In our framework we use the difference between the sizes of the first bid and ask queues as a proxy of liquidity imbalance. In the real world market participants can use a lot of other information (like liquidity imbalance on correlated instruments, or real time news feeds). In such context, the speed can be seen as a protection against adverse selection, potentially reducing transaction costs. Within this viewpoint, high frequency actions do not add noise to the price formation process (as opposite to the viewpoint of [44]) but allows market makers to offer better quotes. At this stage, we do not conclude speed is good for liquidity because:

- We only focus on one limit order. We should go towards a framework similar to the one of [80] to conclude on the added value of imbalance for the whole market making process, but it will be too sophisticated at this stage.
- It is not fair to draw conclusions from a knowledge of the theoretical optimal behaviour of one market participant; to go further we should model the game played by all participants, similarly to what have been done in [106]. Again it is a very sophisticated work. Nevertheless, this paper is a first step. We are convinced it is possible to obtain partially explicit formula, to enable more systematic explorations of the influence of the parameters (currently our simulations are highly memory consuming). It should allow to confront our results with observed behaviours (especially using observed values for our parameters  $\alpha, \beta, Q^{Disc}, Q^{Ins}$  and  $\lambda_s$ ).
- Last but not least, any conclusion on the added value of low latency and high frequency market making should take into account market conditions. Its value could change with the level of stress of the price formation.

This work shows that imbalance is used by participants, and provides a theoretical framework to play with limit order placement. It can be used by practitioners. More importantly, we hope other researchers will extend our work in different directions to answer to more questions, and we will ourselves continue to work further to understand better liquidity formation at the smallest time scales thanks to this new framework.

**Acknowledgements.** Authors would like to thank Sasha Stoïkov and Jean-Philippe Bouchaud for discussions about order book dynamics and optimal placement of limit orders that motivated this paper. Moreover, authors would like to underline the work of Gary Sounigo (during his Masters Thesis) and Felix Patzelt (during a post-doctoral research), who worked hard at Capital Fund Management (CFM) to understand how to align NASDAQ-OMX labelled transactions to direct datafeed of order book records. Long lasting discussions about limit order placement with Capital Fund Management execution researchers (especially with Bence Toth and Mihail Vladkov) influenced very positively our work.

## I.A Transition probabilities of the markov chain $U_n$

When first limits are totally consumed, new quantities  $Q_n^{Disc}$  and  $Q_n^{Ins}$  are inserted in the order book. We introduce then  $\phi_n$  the joint distribution of the random variables  $Q_n^{Disc}$  et  $Q_n^{Ins}$  at time  $t_n$ . We assume these two variables are independent from their past and independent from the counting processes  $N^{Same,+}$ ,  $N^{Same,-}$ ,  $N^{Opp,-}$  and  $N^{Opp,+}$ . However,  $Q_n^{Disc}$  and  $Q_n^{Ins}$  can be correlated at time  $t_n$ .

Let  $n \in \{0, 1, \dots, f\}$ ,  $p \in \mathbb{R}_+$ ,  $q^{bef} \in \mathbb{N}$ ,  $q^{aft} \in \mathbb{N}$ ,  $q^{opp} \in \mathbb{N}$ ,  $q^{disc} \in \mathbb{N}$ ,  $q^{ins} \in \mathbb{N}$  and  $e \in \{-1, 0, 1\}$

- **When the order has been executed before  $t_n$  (i.e.  $e = 1$  or  $e = -1$ ), then:**

$$\mathbb{P}\left(U_{n+1} = (p, q^{bef}, q^{aft}, q^{opp}, -1) \mid U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)\right) = 1.$$

When the order is executed a dead center is reached and both quantities and the price remain unchanged. In such case, the control has no more influence.

- **When the order isn't executed at  $t_n$  (i.e.  $e = 0$ ), we set:**

**A unit quantity is added to  $Q^{Opp}$ .** Under control "s", the transition probability is the following:

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, q^{bef}, q^{aft}, (q^{opp} + 1), e) \mid U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \mathbb{P}\left(\left\{N_{n+1}^{Opp,+} - N_n^{Opp,+} = 1\right\} \cap \left\{N_{n+1}^{Opp,-} - N_n^{Opp,-} = 0\right\}\right. \\ \left. \cap \left\{N_{n+1}^{Same,+} - N_n^{Same,+} = 0\right\} \cap \left\{N_{n+1}^{Same,-} - N_n^{Same,-} = 0\right\}\right) \\ = \lambda_n^{Opp,+} \Delta_t \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, (q^{bef} + q^{aft}), 0, (q^{opp} + 1), e) \mid U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Opp,+} \Delta_t \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

**A unit quantity is cancelled from  $Q^{Opp}$ .** We differentiate between two cases:

1. **When  $q^{opp} > 1$ , under control "s":**

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, q^{bef}, q^{aft}, (q^{opp} - 1), e) \mid U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \mathbb{P}\left(\left\{N_{n+1}^{Opp,+} - N_n^{Opp,+} = 0\right\} \cap \left\{N_{n+1}^{Opp,-} - N_n^{Opp,-} = 1\right\}\right. \\ \left. \cap \left\{N_{n+1}^{Same,+} - N_n^{Same,+} = 0\right\} \cap \left\{N_{n+1}^{Same,-} - N_n^{Same,-} = 0\right\}\right) \\ = \lambda_n^{Opp,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$



Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, (q^{bef} + q^{aft}), 0, (q^{opp} - 1), e) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Opp,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

2. **When**  $q^{opp} \leq 1$ , the price increases by one tick under control "s":

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p+1, q^{ins}, 0, q^{disc}, e) | U_n = (p, q^{bef}, q^{aft}, 1, e)) \\ = \lambda_n^{Opp,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right) \phi_{n+1}(q^{disc}, q^{ins}). \end{aligned}$$

Under control "c", the last formula does not change.

**A unit quantity is added to**  $Q^{Same}$ , under control "s" we have:

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, q^{bef}, q^{aft} + 1, q^{opp}, e) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,+} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right). \end{aligned}$$

Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, (q^{bef} + q^{aft} + 1), 0, q^{opp}, e) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,+} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,-} \Delta_t\right). \end{aligned}$$

**A unit quantity is cancelled from**  $Q^{Same}$ . We distinguish again three cases:

1. **When**  $(q^{bef} > 1 \text{ and } q^{aft} \geq 0)$  or  $(q^{bef} = 1 \text{ and } q^{aft} \geq 1)$ , under control "s":

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, q^{bef} - 1, q^{aft}, q^{opp}, e) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

We suppose that our order is executed when  $Q_n^{After}$  is consumed. Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, (q^{bef} + q^{aft} - 1), 0, q^{opp}, e) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

2. **When**  $q^{bef} = 0$  and  $q^{aft} > 1$ , under control "s":

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, 0, (q^{aft} - 1), q^{opp}, 1) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, (q^{aft} - 1), 0, q^{opp}, 0) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right). \end{aligned}$$

3. **When**  $q^{bef} + q^{aft} = 1$ , under control "s":

$$\begin{aligned} \mathbb{P}^{(c)}(U_{n+1} = (p, q^{disc}, 0, q^{ins}, 1) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right) \phi_{n+1}(q^{disc}, q^{ins}). \end{aligned}$$

Under control "c":

$$\begin{aligned} \mathbb{P}^{(s)}(U_{n+1} = (p, q^{disc}, 0, q^{ins}, 0) | U_n = (p, q^{bef}, q^{aft}, q^{opp}, e)) \\ = \lambda_n^{Same,-} \Delta_t \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right) \phi_{n+1}(q^{disc}, q^{ins}). \end{aligned}$$

**Nothing happens** in the limit order book with probability

$$\mathbb{P}^{(c)}(U_{n+1} = U_n | U_n) = \left(1 - \lambda_n^{Same,-} \Delta_t\right) \left(1 - \lambda_n^{Opp,+} \Delta_t\right) \left(1 - \lambda_n^{Opp,-} \Delta_t\right) \left(1 - \lambda_n^{Same,+} \Delta_t\right).$$

- For all the remaining cases we assume the transition probability neglectible. We hence set it to zero.

**Remark.** By taking in consideration the different cases and neglecting the terms with order strictly superior than 1 in  $\Delta_t$ , we have for any control  $i \in \{c, s\}$ :

$$\begin{aligned} \sum_{\substack{\text{states } U_n \\ \text{states } U_{n+1}}} \int_{(\mathbb{N}_+)^2} \mathbb{P}^{(i)}(U_{n+1} | U_n) \mu_{n+1}(dq^{disc}, dq^{ins}) \\ \approx 1 + \lambda_n^{Same,+} \Delta_t + \lambda_n^{Same,-} \Delta_t + \lambda_n^{Opp,+} \Delta_t + \lambda_n^{Opp,-} \Delta_t. \end{aligned}$$

Consequently, if  $\lambda_n^{Same,+} \Delta_t + \lambda_n^{Same,-} \Delta_t + \lambda_n^{Opp,+} \Delta_t + \lambda_n^{Opp,-} \Delta_t = o(1)$  ( which is true when  $\Delta_t$  is small), we end up with for any control  $i \in \{c, s\}$ :

$$\sum_{\substack{\text{states } U_n \\ \text{states } U_{n+1}}} \int_{(\mathbb{N}_+)^2} \mathbb{P}^{(i)}(U_{n+1} | U_n) \phi_{n+1}(dq^{disc}, dq^{ins}) \approx 1.$$

## I.B Composition of market participants groups

### High Frequency Traders

Name	NASADQ-OMX member code(s)	Market Maker	Prop. Trader
All Options International B.V.	AOI		
Hardcastle Trading AG	HCT		
IMC Trading B.V	IMC, IMA	Yes	
KCG Europe Limited	KEM, GEL	Yes	
MMX Trading B.V	MMX		
Nyenburgh Holding B.V.	NYE		
Optiver VOF	OPV		Yes
Spire Europe Limited	SRE, SREA, SREB		Yes
SSW-Trading GmbH	IAT		
WEBB Traders B.V	WEB		
Wolverine Trading UK Ltd	WLV		

Table I.3 – Composition of the group of HFT used for empirical examples and the composition of our “high frequency market maker” and “high frequency proprietary traders” subgroups.

### Global Investment Banks

Name	NASADQ-OMX member code(s)
Barclays Capital Securities Limited Plc	BRC
Citigroup Global Markets Limited	SAB
Commerzbank AG	CBK
Deutsche Bank AG	DBL
HSBC Bank plc	HBC
Merrill Lynch International	MLI
Nomura International plc	NIP

Table I.4 – Composition of the group of Global Investment Banks used for empirical examples.

## I.C Extreme Imbalances

**The decreasing slope at the right of the curve** in Figure I.7.a and I.7.b when imbalance is highly positive (i.e.  $Q^{Same} \gg Q^{Opp} \approx 1$ ). In this situation, the order will be executed in general before the final time  $T_f$  without being followed by a price move (1) or will be executed at  $T_f$  and followed by a price move (2). In both cases, the final constraint  $\Delta P$  is positive (see

**Institutional Brokers**

Name	NASDAQ-OMX member code(s)
ABG Sundal Collier ASA	ABC
Citadel Securities (Europe) Limited	CDG
Erik Penser Bankaktiebolag	EPB
Jefferies International Limited	JEF
Neonet Securities AB	NEO
Remium Nordic AB	REM
Timber Hill Europe AG	TMB

Table I.5 – Composition of the group of Institutional Brokers used for empirical examples.

Graph I.12). Given that  $\Delta P$  in case (2) is lower than  $\Delta P$  in case (1) and situation (2) occurs more frequently when imbalance is highly positive, it is expected to find a decreasing slope at the right of the curve.

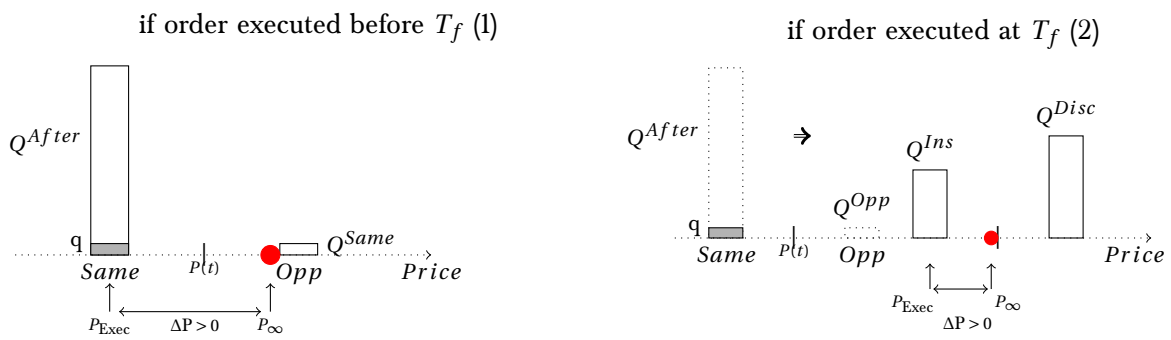


Figure I.12 –  $\Delta P$  when Imbalance highly positive



## CHAPTER II

---

# Optimal liquidity-based trading tactics

---

### Abstract

We consider an agent who needs to buy (or sell) a relatively small amount of assets over some fixed short time interval. We work at the highest frequency meaning that we wish to find the optimal tactic to execute our quantity using limit orders, market orders and cancellations. To solve the agent's control problem, we build an order book model and optimize an expected utility function based on our price impact. We derive the equations satisfied by the optimal strategy and solve them numerically. Moreover, we show that our optimal tactic enables us to outperform significantly standard execution strategies.

**Keywords:** Market microstructure; limit order book; high frequency trading; queuing model; Markov jump processes; ergodic properties; adverse selection; execution probabilities; market impact; optimal trading strategies; optimal tactics; stochastic control.

## 1 Introduction

Most electronic exchanges use an order book mechanism. In such markets, buyers and sellers send their orders to a continuous-time double auction system. These orders are then matched according to price and time priority. Each submitted order has a specific price and size and the order book is the collection of all submitted and unmatched limit orders. This is illustrated in Figure II.1, which shows a classical representation of an order book at a given time.

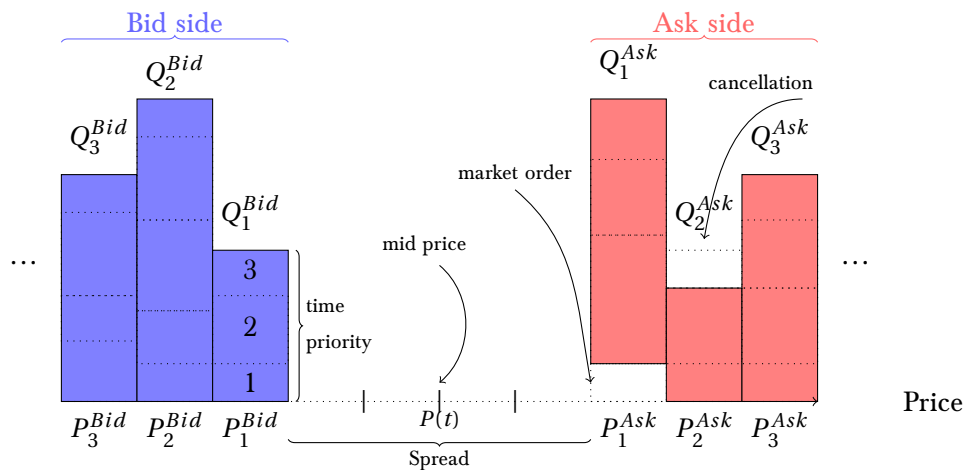


Figure II.1 – Order book representation at a given time. Here  $P_i^{Ask}$  (resp.  $P_i^{Bid}$ ) with  $i \geq 1$  are the sellers (resp. buyers) limit prices and they are increasingly (resp. decreasingly) ordered. For a given price  $P_i^{Ask}$  (resp.  $P_i^{Bid}$ ), the limit  $Q_i^{Ask}$  (resp.  $Q_i^{Bid}$ ) is the available selling (resp. buying) quantity.

In this limit order book setting, we consider the following issue: an agent has to buy or sell a given quantity of asset before a fixed time horizon. During the execution process, the agent can take four elementary decisions:

- Insert limit orders in the order book, hoping to avoid crossing the spread. We will assume that the agent does not place limit orders above the best limits however he can insert them within the spread.
- Stay in the order book with an already existing limit order, to keep his tactical placement.
- Cancel existing limit orders.
- Send market orders to get immediate execution.

Note that this is the microstructural version of the classical Almgren-Chriss optimal execution problem for the liquidation of a large quantity of asset over a time interval  $[0, T]$ , see [10, 29, 74] and [47, 48, 49, 51, 52, 76, 82, 85] for various extensions. In the setting of [10],  $[0, T]$  is split in sub time windows (typically a few minutes per window) and one derives the number of shares to be executed in each window. In our case, we want to specify how to act optimally within each window. Indeed, our buyer or seller reacts to every order book move and handles reasonably small quantities during short periods of time.

In order to solve this problem, we of course need to model the order book dynamics. There are essentially two order book modelling approaches in the literature. First, “equilibrium models”, based on interactions between rational agents who take optimal decisions, see

[66, 134, 139]. Second, “statistical models” where the order book is seen as a suitable random process, see [2, 3, 20, 55, 57, 83, 106, 107, 144]. Statistical models focus on reproducing many salient features of real markets rather than individual agents behaviours and interactions between them. In this paper, we use a statistical model. In such models, the arrival and cancellation flows often follow independent Poisson processes. The Poisson assumption allows for the derivation of simple, and often closed-form, formulas, for example for the probabilities of various order book events, see [2, 57, 83, 106, 129].

However, as clearly shown in [90], this assumption is not realistic and it is necessary to take into account accurately the local state-dependent behaviour of the order book. So in [90, 91], the authors introduce the Queue-Reactive order book model where order flows follow a Markov jump process. They also provide ergodicity conditions and model parameters calibration methodology.

**Order book model.** Here we refine the Queue-Reactive model to make it compatible with a stochastic control framework enabling us to solve important practical issues. To do so, we only consider the best bid and ask limits to work with a reasonably small state space. Furthermore, in order to get a truly good fit to real order book dynamics, we focus on the so-called regeneration process which models the order book state right after the total depletion of a limit. In our setting, when a limit is totally depleted, the order book is regenerated in a new state whose regeneration law depends on the order book state just before the depletion. In general, order book models consider several bid and ask limits and use a regeneration process independent from the order book state, see [3, 57, 90]. Here, we model the order book by a three-dimensional Markov jump process  $(Q_t^1, Q_t^2, S_t)$  where  $Q_t^1$  is the available quantity at the best bid,  $Q_t^2$  is the available quantity at the best ask and  $S_t$  is the spread. In addition, we provide a dynamics for the mid price  $P_t$  which depend on the order book state.

**The agent control problem.** Let us now introduce the agent’s control problem. We formulate it for a buy order of size  $q^a$  (it can be changed to a sell order in an obvious way). From time zero to the final time  $T$ , we assume that, at every decision time, the buyer can do nothing or use one of the three following actions: insert a fraction of the remaining quantity to buy (if not already inserted) at the top of the bid queue or within the spread (decision l), cancel the already inserted limit orders (decision c) or send a market order (decision m) for a fraction of the remaining quantity. If the agent does not obtain the total execution of  $q^a$  at time  $T$ , he cancels the remaining quantity in the order book and send a market order. Thus, the trader’s strategy is modelled by the sequence  $\mu = (\tau_i, v_i)_{i \geq 0}$  of random variables where  $(\tau_i)_{i \geq 0}$  is a sequence of increasing stopping times that represents the optimal decision times and  $v_i \in \mathbb{E} = \mathbb{T} \times \mathbb{N} \times \mathbb{P}$  refers to the optimal decision. Here, the set  $\mathbb{T} = \{l, c, m\}$  indicates the type of the order, the set  $\mathbb{N}$  specifies the order size and  $\mathbb{P}$  determines the order posting price.<sup>1</sup> The agent aims at determining the optimal sequence of decisions to reduce its price impact

<sup>1</sup>See Section 3.1 for a detailed description of the control  $\mu$  and the set  $\mathbb{P}$ .



## II. Optimal liquidity-based trading tactics

---

$PI_\infty^\mu$  which is defined such that

$$PI_\infty^\mu = -\lim_{t \rightarrow \infty} \mathbb{E}[q^a P_t^\mu - P^{Exec,\mu}], \quad (1)$$

with  $P_t^\mu$  the mid price,  $P^{Exec,\mu}$  the acquisition price of the quantity  $q^a$ ,  $t$  the current observation time and  $\mu_t$  the agent's control which remains constant when the user does nothing.<sup>2</sup>

Let  $I_t^\mu$  be the agent's inventory, that is the remaining quantity he has to buy at time  $t$ .<sup>3</sup> We denote by  $P_t^{Exec,\mu}$  the acquisition price of the quantity  $q^a - I_t^\mu$  and  $T_{Exec}^\mu$  the time where  $q^a$  is totally executed. To take into account the waiting cost, the sensitivity to the price impact and to work in a slightly more general setting, we consider the following optimisation problem:

$$\sup_{\mu} \mathbb{E} \left[ \underbrace{f \left( \lim_{t \rightarrow \infty} \mathbb{E} [q^a P_t^\mu - P_{T_{Exec}^\mu}^{Exec,\mu} / \mathcal{F}_{T_{Exec}^\mu}^\mu] \right)}_{\text{final constraint}} - \underbrace{\gamma \int_0^{T_{Exec}^\mu} I_s^\mu ds}_{\text{running cost}} \right],$$

where  $f: \mathbb{R} \rightarrow \mathbb{R}$  is a Lipschitz function and  $\gamma$  is an homogenization non-negative constant representing the waiting cost. Here we use a conditional expectation to account for the fact that agents collect information along their own trading. Note that optimal execution problems, in general, minimize the acquisition price (resp. liquidation) of the quantity to buy (resp. sell) called here  $P_{T_{Exec}^\mu}^{Exec,\mu}$ . However because of the price relaxation typically following the end of the metaorder execution, this acquisition (resp. liquidation) price is not an appropriate benchmark for agents who need to sell (resp. buy) back at least a portion of the executed shares in the future. This is the case of brokerage companies who execute a large number of clients metaorders everyday. In this work, we place ourselves in a setting where we can define a notion of the price after the relaxation (i.e. expected long term value)  $P_\infty^\mu = \lim_{t \rightarrow \infty} \mathbb{E}[P_t^\mu]$  and use it as a benchmark. Our trading has an influence on  $P_\infty^\mu$  that we will be able to compute. Note that, we send  $t$  to  $\infty$  to get the new stationary value of the price after our execution. Thus, infinity should be seen as the scale of time where the resilience of the market takes place. This notion of resilience after order executions has been studied in the literature, see [71]. In our case, its order of magnitude corresponds to few hours. We stress that the reduction of the price impact is crucial for brokers. As a matter of fact, they do not only execute one given (meta-)order, but interact thousands of times per day with the order book to execute hundreds of metaorders, which will all be subjected to the impact of previous orders. We study this problem in two cases. First, when agent's decisions are taken at fixed frequency  $\Delta^{-1}$ . This enables us to investigate latency effects and moderately high frequency trading issues. Second, when agent's decisions are taken at any time, to handle the situation where one has access to ultra high frequency trading technology.

---

<sup>2</sup>If the agent has no order in the order book and does nothing at the beginning of the period, we consider he starts with control  $c$ .

<sup>3</sup>In the latter, we slightly improperly call  $I_t^\mu$  the inventory of the agent.

**Positioning of the paper.** This paper is obviously not the first work where a stochastic control framework involving limit orders, market orders and cancellations is used to solve a high frequency trading problem. For example in [79, 108, 82] and Chapter I, the authors consider the problem of optimal posting of a limit order while market making issues are addressed in [13, 52, 80, 84] and Chapter III. The problem of optimal execution using limit and market orders is also investigated in [45, 50, 56, 75, 78]. However the interactions between market participants decisions, liquidity and behaviour of the order book are not really taken into account. In our work, the decision of the agent depends on its current position in the queue and the available liquidity on the order book. Some of the few papers considering an approach close to ours are [81, 94] and Chapter I. Beyond a slightly more general setting, compared to these papers, our main contribution is to optimize our trading tactic not only with respect to its local profit and loss but also to the endogenous price impact it generates. This is of primary importance for any market participant with intense trading activity.

**Results.** In this paper, we propose an order book model in reduced dimension, with state-dependent regeneration and non-constant spread. Within this framework, we provide a closed-form formula for the endogenous price impact. This allows us to compute the sequence of orders solving the optimal execution problem for an agent who wants to minimize its impact. When the trading frequency is fixed, an explicit expression is obtained for the value function of the agent. When decisions can be taken at any time, we explain how to design simple and efficient numerical methods and provide accurate error estimates. We even show that the optimal control in continuous time is exactly obtained when the discretization mesh of the numerical scheme is small enough. We also prove the ergodicity of our order book under mild assumptions on the intensity functions.

The paper is organized as follows. In Section 2, we introduce our order book model, prove its ergodicity and provide the formula for the price impact. In Section 3, we formulate the agent's control problem. Our main theorems including the equations satisfied by the value function and the numerical methodology to solve them are provided in Section 4. Finally, numerical experiments are given in Section 5. The proofs are relegated to an appendix.

## 2 Order book modelling

In this section, we first confirm on data that agents behaviours depend on order book liquidity, see [90, 111] and Chapter I for closely related results. Then, we describe the order book dynamics.

### 2.1 Preliminary: Empirical evidences

One specificity of our work is that we wish to carefully model the interactions between market participants and liquidity. We first show on real data that market participants act differently when facing different liquidity conditions.

**Database presentation.** Data used here are from Bund futures on Eurex exchange Frankfurt. We focus on this product since it is a good example of a very liquid and large tick asset. The database records during one week from 1 to 5 September 2014, the state of the order book (i.e. available quantities and prices at best limits) event by event with microsecond accuracy. For each day, our data cover the time period from 8 a.m to 10 p.m Frankfurt time. Each event has a type, a side (i.e. bid/ask) and a size. We consider three types of events: insertion of limit orders, cancellation of existing limit orders and market orders. The database accounts for 3 407 574 events.

Let  $t$  be the time when an event happens in the order book. We define the imbalance  $\text{Imb}_t$  and the mid price move  $\delta$  seconds after the event time  $t$ ,  $\Delta P_\delta^{\text{mid}}(t)$ , by

$$\begin{cases} \text{Imb}_t = \epsilon_t \frac{Q_t^1 - Q_t^2}{Q_t^1 + Q_t^2}, \\ \Delta P_\delta^{\text{mid}}(t) = \epsilon_t \frac{P_{\delta+t} - P_t}{s_t}, \end{cases}$$

where  $Q_t^1$  (resp.  $Q_t^2$ ) is the available quantity at the best bid (resp. ask),  $P_t$  is the mid price,  $\epsilon_t$  is the event sign (i.e.  $\epsilon_t = 1$  when it is a buy order and  $-1$  otherwise) and  $s_t$  is the spread (i.e.  $s_t = P_t^{\text{Ask}} - P_t^{\text{Bid}}$  with  $P_t^{\text{Ask}}$  the best ask price and  $P_t^{\text{Bid}}$  the best bid price).

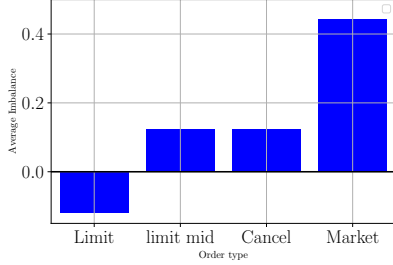
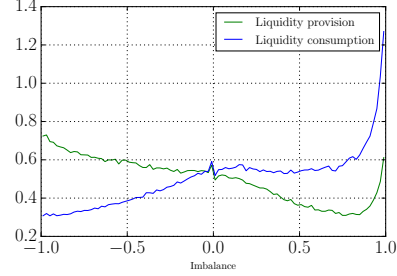
We want to confirm that agents decisions depend on the order book liquidity. A simple way to do so is to summarize the state of the order book liquidity through the imbalance. Figure II.2.a shows the average imbalance value for each event type. We give the interpretation of Figure II.2.a in the case of a buy limit/limit mid<sup>4</sup>/cancellation/market order, since the event sign is taken into account in the expression of  $\text{Imb}_t$ . We see that market participants insert limit orders when imbalance is negative (execution highly probable), cancel orders when imbalance is positive (less chance to be executed) and use market orders when imbalance is highly positive (rushing for liquidity when it is scarce).

Figure II.2.b shows the distribution of imbalance just before a liquidity provision event (i.e.. insertion of limit orders) and a liquidity consumption event (i.e.. cancellation of limit orders or market orders). We see that agents are highly active at extreme imbalance values.<sup>5</sup> Indeed, in these cases, they identify a profit opportunity to catch or on the contrary an adverse selection effect to avoid (for example buying just before a price decrease). This is related to the predictive power of the imbalance. As can be seen in Figure II.2.c,  $\Delta P_\delta^{\text{mid}}(t)$  after 2 minutes (i.e.  $\delta = 2$  min) is highly correlated to the imbalance. This means that market participants use the imbalance as a signal to anticipate next price moves.<sup>6</sup> Hence, our empirical results clearly confirm that agents decisions depend on the order book liquidity.

<sup>4</sup>Limit order inserted within the spread.

<sup>5</sup>The high rate of liquidity provision for very positive imbalance can be surprising at first sight. However it may be due to orders inserted within the spread creating a new best limit.

<sup>6</sup>Quoting Sasha Stoikov: Imbalance is the least well hidden secret of high frequency trading.

(a) Average imbalance before  
Limit/Cancel/Market order

 (b) Imbalance density before  
liquidity provision/consumption event


(c) Average price move after 2 minutes against imbalance

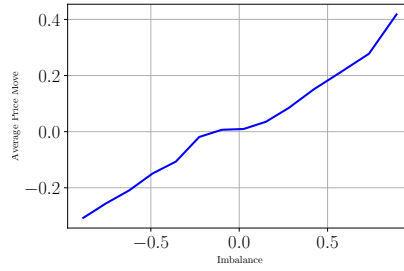


Figure II.2 – Some statistics about the imbalance.

**Bid-ask symmetry assumption.** It links buyers and sellers behaviours. To formulate the bid-ask symmetry relation, we denote by  $U_t = (Q_t^1, Q_t^2, P_t)$  the process modelling the state of the order book. We recall that  $Q_t^1$  (resp.  $Q_t^2$ ) is the available quantity at the best bid (resp. ask) and  $P_t$  is the mid price. For a buyer,  $Q^1$  represents the waiting queue before the execution and  $Q^2$  the available buying quantity. For a seller, roles of  $Q^1$  and  $Q^2$  are interchanged:  $Q^2$  is the waiting queue and  $Q^1$  the available liquidity. Additionally, buyers and sellers decisions depend on the reference price value: a buyer (resp. seller) aims to be executed at the lowest (resp. highest) price. Thus, we define  $U_t^{sym} = H(U_t) = (Q_t^2, Q_t^1, -P_t)$  the symmetric process of  $U_t$  where the function  $H$  moves us from the buyer point of view to the seller one or conversely. The symmetry assumption ensures that

$$U_t^{sym} / U_0^{sym} \stackrel{Law}{=} U_t / U_0, \quad \forall t \geq 0. \quad (2)$$

We assume the bid-ask symmetry relation since it ensures no statistical arbitrage

$$\mathbb{E}[\Delta P_t] = 0, \quad \forall t \geq 0. \quad (3)$$

with  $\Delta P_t = P_t - P_0$ . When a stationary distribution exists, we also have  $\mathbb{E}[\Delta P_\infty] = 0$ . Additionally, when  $P_t$  is a Markov process, Equation (3) proves that  $P_t$  is a martingale. Finally, when  $P_t$  is a right continuous Markov process Equation (3) ensures absence of arbitrage opportunities

## II. Optimal liquidity-based trading tactics

in the market. Finally, in our Markovian framework, the bid-ask symmetry assumption allows us to derive the optimal control for a seller directly from the buyer one.

To illustrate the bid-ask symmetry, we present, in Table II.1, basic statistics about the time between two consecutive events happening in the order book. According to Equation (2), the mean value between two consecutive events at the best bid and the best ask side should be equal. In Table II.1 column names correspond to the event type, for example in column “Global” (resp. “Bid”) we study the time between two consecutive events happening in the order book (resp. best bid) and so on. In our case, cancellation includes market orders. Table II.1 shows that the mean value between two consecutive limit/cancellation/global events at the best bid and the best ask are close.

	Global	Bid	Ask	BidCancel	AskCancel	BidLimit	AskLimit
Mean ( $10^{-2}$ sec)	1.62	2.37	2.33	2.61	2.65	4.26	4.12
Nb.Events	$3.38.10^6$	$1.67.10^6$	$1.71.10^6$	$8.66.10^5$	$8.86.10^5$	$8.08.10^5$	$8.21.10^5$

Table II.1 – Median value of the time distribution between two consecutive events. Events associated to price moves are removed.

### 2.2 Order book framework

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  be a filtered probability space with  $\mathcal{F}_0$  the trivial  $\sigma$ -algebra. The order book state is modelled by the Markov process  $U_t = (Q_t^1, Q_t^2, S_t)$  where  $Q_t^1$  (resp.  $Q_t^2$ ) is the best bid (resp. ask) quantity and  $S_t$  is the spread. We denote by  $P_t^1$  (resp.  $P_t^2$ ) the best bid (resp. ask) price,  $P_t$  the mid price and  $\tau_0$  the tick size. The mid price dynamics will be described in Section 2.4 and in this section we focus on the dynamics of  $U_t$ . For simplification, we take the state space  $\mathbb{U} = (\mathbb{N}^*)^2 \times \tau_0 \mathbb{N}^*$ .<sup>7</sup>

The Markov process  $U$  is characterized by its the infinitesimal generator  $\mathcal{Q}$ . For  $u = (q^1, q^2, s) \in \mathbb{U}$ ,  $u' = (q'^1, q'^2, s') \in \mathbb{U}$ ,  $n \in \mathbb{N}^*$ ,  $k \in \{1, \dots, \frac{s}{\tau_0} - 1\}$ ,  $e_1 = (1, 0)$ ,  $e_2 = (0, 1)$  and  $i \in \{1, 2\}$ , we consider the following form for  $U_t$ :

$$\left\{ \begin{array}{ll} \mathcal{Q}_{(q,s),(q+ne_i,s)} & = \lambda^{i,+}(u, n) + R(u, (q+ne_i, s)), \\ \mathcal{Q}_{(q,s),(q-ne_i,s)} & = \lambda^{i,-}(u, n, \mathbf{0}) + R(u, (q-ne_i, s)) & , \text{ if } q^i > n, \\ \mathcal{Q}_{(q,s),(q+(n-q^i)e_i, s-k\tau_0)} & = \lambda^{i,k}(u, n) + R(u, (q+(n-q^i)e_i, s-k\tau_0)) & , \text{ if } s > \tau_0, \\ \mathcal{Q}_{(q,s),(q',s')} & = R(u, (q', s')) & , \text{ if } s' \neq s, \end{array} \right. \quad (4)$$

with  $\mathcal{Q}_{u,u} = -\sum_{u' \neq u} \mathcal{Q}_{u,u'}$ ,  $q = (q^1, q^2)$ ,  $q' = (q'^1, q'^2)$  and

- $\lambda^{1,+}(u, n)$  (resp.  $\lambda^{2,+}(u, n)$ ) represents the arrival rate of limit orders of size  $n$  at the best bid (resp. ask) when the order book state is  $u$ .

<sup>7</sup>However we can work on the following state space  $\mathbb{U} = (\mathbb{R}_+^*)^3$  and recover similar results.

- $\lambda^{1,-}(u, n, \mathbf{0})$  (resp.  $\lambda^{2,-}(u, n, \mathbf{0})$ ) is the arrival rate of liquidity consumption orders of size  $n$  that do not deplete the best bid (resp. ask), when the order book state is  $u$ .
- $\lambda^{1,k}(u, n)$  (resp.  $\lambda^{2,k}(u, n)$ ) represents the arrival rate of buying (resp. selling) limit orders of size  $n$  within the spread at the price  $P^1 + k\tau_0$  (resp.  $P^2 - k\tau_0$ ).
- $\lambda^{1,-}(u, n, u')$  (resp.  $\lambda^{2,-}(u, n, u')$ ) represents the arrival rate of liquidity consumption orders of size  $n$  that deplete the best bid (resp. ask) and lead to a new state  $u'$  when the order book state is  $u$ .
- $R(u, u')$  verifies  $R(u, u') = \sum_{j=1}^2 \sum_{m \geq q^j} \lambda^{j,-}(u, m, u')$  and represents the order book regeneration component.

**Order book regeneration.** In our framework, when one limit is totally depleted, the order book is regenerated in a new state whose law depends on the order book state just before the depletion and the depleted side (i.e. best bid/ask). The regeneration of the process  $U$  is described through the quantity  $R(u, u')$  where  $u' \in \mathbb{U}$  is the order book state after the depletion and  $u$  is the order book state before the depletion. A simple choice is to consider the case where the spread increases by one tick when the best bid or ask are depleted and to draw new best bid and ask quantities from a fixed stationary distribution, see [55].

**Symmetry relations.** Additionally, for every  $u = (q^1, q^2, s) \in \mathbb{U}$ ,  $u' = (q'^1, q'^2, s') \in \mathbb{U}$ ,  $n \in \mathbb{N}^*$  and  $k \leq \frac{s}{\tau_0} - 1$ , we assume the following bid-ask symmetry relations:

$$\begin{cases} \lambda^{1,+}(u, n) &= \lambda^{2,+}(u^{sym}, n), \\ \lambda^{1,k}(u, n) &= \lambda^{2,k}(u^{sym}, n), \\ \lambda^{1,-}(u, n, u') &= \lambda^{2,-}(u^{sym}, n, u'^{sym}), \end{cases} \quad (5)$$

with  $u^{sym} = (q^2, q^1, s)$ . This classical bid-ask symmetry relation ensures no statistical arbitrage and allows us to aggregate bid and ask side data in the calibration of the model's parameters.

### 2.3 Ergodicity

We now provide a theoretical result on the ergodicity of the process  $U_t = (Q_t^1, Q_t^2, S_t)$  under three general assumptions. A definition of the ergodicity is given in Appendix II.B.

We denote by  $\lambda_Q^{i,+}$  (resp.  $\lambda_Q^{i,-}$ ) and  $\lambda_S^+$  (resp.  $\lambda_S^-$ ) the arrival rate of events that increase (resp. decrease) by  $n$  the size of the limit  $Q^i$  and the spread  $S$  for any  $i \in \{1, 2\}$ . For sake of completeness, we give an explicit expression for  $\lambda_Q^{i,\pm}$  and  $\lambda_S^\pm$

$$\lambda_Q^{i,\pm}(u, n) = \sum_{(q', s') \in \mathbb{U}} \mathcal{Q}_{u, (q', s')} \mathbf{1}_{q'^i = q^i \pm n}, \quad \lambda_S^\pm(u, n) = \sum_{(q', s') \in \mathbb{U}} \mathcal{Q}_{u, (q', s')} \mathbf{1}_{s' = s \pm n}, \quad \forall u = (q^1, q^2, s) \in \mathbb{U}, \forall n \geq 1.$$

## II. Optimal liquidity-based trading tactics

---

**Assumption 1** (Negative individual drift). *There exist three positive constants  $C_{bound}$ ,  $z_0 > 1$  and  $\delta > 0$  such that for any  $u = (q^1, q^2, s) \in \mathbb{U}$*

$$\begin{aligned} \sum_{n \geq 0} (z_0^n - 1) (\lambda_Q^{1,+}(u, n) - \lambda_Q^{1,-}(u, n) \frac{1}{z_0^n}) &\leq -\delta, \quad \text{when } q^1 \geq C_{bound}, \\ \sum_{k \geq 0} (z_0^{\tau_0 k} - 1) (\lambda_S^+(u, k) - \lambda_S^-(u, k) \frac{1}{z_0^{\tau_0 k}}) &\leq -\delta, \quad \text{when } s \geq C_{bound}. \end{aligned}$$

Assumption 1 ensures that the queue size and the spread value tend to decrease when they become too large. Using Equation (5), we also have

$$\sum_{n \geq 0} (z_0^n - 1) (\lambda_Q^{2,+}(u, n) - \lambda_Q^{2,-}(u, n) \frac{1}{z_0^n}) \leq -\delta, \quad \forall q^2 \geq C_{bound}.$$

**Assumption 2** (Local bound on the incoming flow). *There exists  $z_1 > 1$  such that for any  $B \geq 0$  we have*

$$\begin{aligned} \sum_{n \geq 0} z_1^n \lambda_Q^{1,+}(u, n) &\leq H^B, \quad \text{when } q^1 \leq B, \\ \sum_{k \geq 0} z_1^{\tau_0 k} \lambda_S^+(u, k) &\leq H^B, \quad \text{when } s \leq B, \end{aligned}$$

with  $u = (q^1, q^2, s) \in \mathbb{U}$  and  $H^B$  a positive constant.

Assumption 2 ensures no explosion in the system: the order arrival speed stays bounded within any bounded set of  $\mathbb{U}$ . Using the symmetry relation, we have  $\sum_{n \geq 0} z_1^n \lambda_Q^{2,+}(u, n) \leq H^B$  when  $q^2 \leq B$ . For example, Assumption 2 is satisfied in the particular case when both the size of the best limits and the the spread cannot get beyond a maximum threshold. Assumption 1 and 2 are close to those used in [91] and slightly more general. We have the following result.

**Theorem 1** (Ergodicity). *Under Assumptions 1 and 2, the process  $U_t$  is ergodic (i.e.. converges towards a unique invariant distribution). Additionally, we have the following speed of convergence:*

$$\|P_u^t(\cdot) - \pi\|_{TV} \leq B(u)\rho^t,$$

with  $\|\cdot\|_{TV}$  the total variation distance,  $P_u^t(\cdot)$  the Markov kernel of the process  $U_t$  starting from the initial point  $u \in \mathbb{U}$ ,  $\pi$  the invariant distribution,  $\rho < 1$  and  $B(u)$  a constant depending on the initial state  $u$ , see Appendix II.B.

**Remark 1.** *To prove the ergodicity we do not require the intensities to be uniformly bounded.*

This theorem is the basis for the asymptotic study of the order book dynamics in Section 2.1, since it ensures the convergence of the order book state towards an invariant probability distribution. Thus the stylized facts observed on market data can be explained by a law of large numbers type phenomenon for this invariant distribution. The proof of this result is given in Appendix II.B and it is quite inspired from [90, 91].

## 2.4 Computation of the long-term average price move

In this section, we provide the computation methodology of the following quantity:

$$D(u) = \lim_{t \rightarrow \infty} \mathbb{E}_u [P_t - P_0], \quad (6)$$

with  $u \in \mathbb{U}$ . The computation of  $D(u)$  is interesting for at least two reasons: first, it allows us to predict the long-term average mid price move for any initial order book state  $u$ . Second, the quantity  $D(u)$  is useful for the computation of the price impact  $PI_\infty^\mu$  defined in (1), see Section 4.1 for a detailed connection between the computation of  $PI_\infty^\mu$  and Equation (6). To compute numerically  $D(u)$ , we need to bound the domain  $\mathbb{U}$ . Hence, we replace Assumption 2 by slightly less general assumptions.

**Assumption 3** (Insertion Bound). *There exists a positive quantity  $Q^{max}$  such that for any  $u = (q^1, q^2, s) \in \mathbb{U}, u' \in \mathbb{U}, n \geq 0$  and  $k \leq \frac{s}{\tau_0} - 1$ ,*

$$\begin{cases} \lambda^{1,+}(u, n) = 0, & \text{when } q^1 + n > Q^{max}, \\ \lambda^{1,k}(u, n) = 0, & \text{when } q^1 + n > Q^{max}, \\ \lambda^{1,-}(u, n, u') = 0, & \text{when } q^1 > Q^{max}. \end{cases}$$

This assumption is not restrictive since quantities at the best limits remain essentially bounded. Using the symmetry relation, we have as well, for any  $u = (q^1, q^2, s) \in \mathbb{U}, u' \in \mathbb{U}, n \geq 0$  and  $k \leq \frac{s}{\tau_0} - 1$

$$\begin{cases} \lambda^{2,+ (k)}(u, n) = 0, & \text{when } q^2 + n > Q^{max}, \\ \lambda^{2,-}(u, n, u') = 0, & \text{when } q^2 > Q^{max}. \end{cases}$$

**Assumption 4** (Regeneration bound 2). *There exists a positive constant  $\tilde{Q}^{max}$  such that for all  $(u, u') \in \mathbb{U}^2$  and  $i \in \{1, 2\}$ , we have*

$$\lambda^{i,-}(u, n, u') = 0, \quad \text{when } \|u'\|_\infty > \tilde{Q}^{max},$$

with  $\|x\|_\infty = \sup_{j \leq N} |x_j|$  for any vector  $x \in \mathbb{R}^N$ .

Under the symmetry relation, Assumption (4) needs to be satisfied for only one  $i \in \{1, 2\}$ .

**Price dynamics.** The mid price after the  $n$ -th order book event  $P_n$  satisfies  $P_n = P_0 + \sum_{i=1}^n \Delta P_i$  with  $\Delta P_i = P_i - P_{i-1}$ . The price jumps  $\Delta P_i$  are a deterministic function of the order book state before the jump and the order book event causing the jump. For example, we can consider the simple case where the mid price decreases (resp. increases) by one tick when the best bid (resp. ask) is depleted.

**Assumption 5** (Mid price bound). *The process  $U_t$ , see Section 2.2, is irreducible<sup>8</sup> and there exists a state  $u \in \mathbb{U}$  such that  $D(u) < \infty$ .*

**Remark 2.** *When  $D(u) = \pm\infty$  for all  $u \in \mathbb{U}$ , there is nothing to prove. Additionally, since  $U_t$  is irreducible, Assumption 5 implies that  $D(u) < \infty$  for all  $u \in \mathbb{U}$ .*

<sup>8</sup> To ensure that  $U_t$  is irreducible we only need the intensity functions to be positive when Assumptions 3 and 4 are not satisfied.



**Computation methodology of the price impact:** A mid price move can be caused by only one of the following events:

- **event 1:** depletion of the best bid or sell limit order within the spread.
- **event 2:** depletion of the best ask or buy limit order within the spread.

Let  $t_1$  (resp.  $t_2$ ) be the first time when an event of type 1 (resp. 2) happens. At  $t_1$  (resp.  $t_2$ ) the mid price moves on average by  $\alpha_i^- = \mathbb{E}_{U_i}[\Delta P_{t_1}]$  (resp.  $\alpha_i^+ = \mathbb{E}_{U_i}[\Delta P_{t_2}]$ ) and the order book is regenerated according to a distribution  $d_i^1$  (resp.  $d_i^2$ ). Here we assimilate the initial state  $U_i$  to its unique index  $i$ . We consider the following notations:

- $q_{ii'}^- = \mathbb{P}_{U_i}[\{t_1 < t_2\} \cap \{U_{t_1}^- = U_{i'}\}]$  and  $q_{ii'}^+ = \mathbb{P}_{U_i}[\{t_2 \leq t_1\} \cap \{U_{t_2}^- = U_{i'}\}]$ . They represent respectively the probability that  $t_1 < t_2$  or conversely and the exit state is  $U_{i'}$ .
- $d_{i,k}^1$  (resp.  $d_{i,k}^2$ ) are transition probabilities from the state  $U_i$  to  $U_k$  when  $t_1 < t_2$ .
- $q_i = \sum_{i'} (q_{ii'}^+ \alpha_{i'}^+ + q_{ii'}^- \alpha_{i'}^-)$  and  $p_{i,k} = \sum_{i'} (q_{ii'}^+ d_{i',k}^2 + q_{ii'}^- d_{i',k}^1)$  represent respectively the average mid price move after the first regeneration and the probability to reach the state  $U_k$  starting from the initial point  $U_i$  right after the first regeneration.
- $U_i^{sym} = (q^2, q^1, s)$  is the symmetric state of  $U_i = (q^1, q^2, s)$  and  $i^{sym}$  is the index of the symmetric state  $U_i^{sym}$ , see (5).
- $D$  is a vector satisfying  $D_i = \lim_{t \rightarrow \infty} \mathbb{E}_{U_i}[P_t - P_0]$  for every state  $U_i = (q^1, q^2, s) \in \mathcal{D}$ . We write  $\mathcal{D}$  for the set  $\mathcal{D} = \{(q^1, q^2, s); q^1 \geq q^2\}$ .<sup>9</sup>
- $A$  is a matrix defined such that  $A_{i,k} = \frac{p_{i,k} - p_{i,k}^{sym}}{1 - (p_{i,i} - p_{i,i}^{sym})}$  when  $i \neq k$  and  $A_{i,i} = 0$  for any  $(U_i, U_k) \in \mathcal{D}^2$ .<sup>10</sup> Note that the matrix  $I - A$  is invertible.

We have the following result.

**Proposition 1** (Price impact). *Under Assumptions 3, 4 and 5, the vector  $D$  satisfies*

$$D = (I - A)^{-1} F.$$

The vector  $F$  satisfies  $F_i = \frac{q_i}{1 - (p_{i,i} - p_{i,i}^{sym})}$  for any  $U_i \in \mathcal{D}^{10}$ .

The proof of this result is given in Appendix II.C.1. A numerical computation of the vector  $D$  is given in Section 2.5, Figure II.3.

**Remark 3.** *To compute  $D$ , we need to estimate the regeneration distributions  $d^1$ ,  $d^2$ ,  $\alpha^\pm$  and  $q^\pm$ . The quantities  $d^1$ ,  $d^2$  and  $\alpha^\pm$  can be estimated from the order book empirical distribution after a price change. Then, we only need to estimate  $q_{ii'}^\pm$ . The estimation methodology of  $q_{ii'}^\pm$  is detailed in Lemma II.C.3 of Appendix II.C.*

<sup>9</sup>Note that  $D_i = -D_{i^{sym}}$  under the symmetry relation (5), thus it is enough to know the values  $D_i$  for any  $U_i \in \mathcal{D}$ .

<sup>10</sup> Recall that we assimilate the state  $U_i$  to its unique index  $i$ .

## 2.5 Numerical application

In this section, we compute numerically the price impact using Proposition 1. We also assess the model by comparing the theoretical and empirical distributions of  $(Q^1, Q^2)$  at long term and short term horizon.

**Approximation of  $\lim_{t \rightarrow \infty} \mathbb{E}_{U_0}[P_t - P_0]$ .** Figure II.3 shows  $\lim_{t \rightarrow \infty} \mathbb{E}_{U_0}[P_t - P_0]$  defined in Section 2.4 and computed using Proposition 1, for different values of the initial state  $U_0 = (Q^1, Q^2, \tau_0)$  where the spread is equal to one tick. Figure II.3 shows the predictive power of the imbalance: when the imbalance is positive the price increases on average and conversely. We also note that the bid-ask symmetry relation is respected.

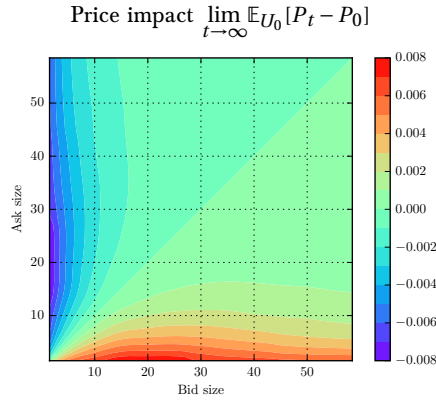
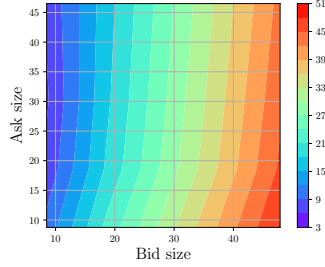
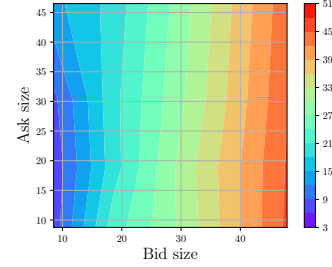


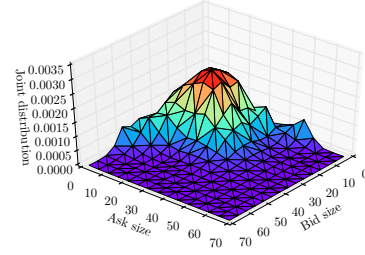
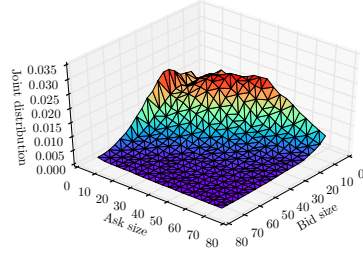
Figure II.3 – Price impact  $\lim_{t \rightarrow \infty} \mathbb{E}_{U_0}[P_t - P_0]$  in percentage of the tick size  $\delta$ .  $Q^1$  and  $Q^2$  are divided by the average event size and the tick  $\delta = 0.01$ .

**Model approximation at short-time horizon.** Figures II.4.a and II.4.b show respectively the empirical and theoretical distributions of  $Q^1$  after 20 events. We choose 20 events since it is coherent with the duration of our control. The estimation of the theoretical distribution is based on a Monte-Carlo simulation of the order book. We can see that both distributions are close and consequently that our model is consistent with the empirical order book dynamics at least during the control duration. The model is also consistent with empirical data on long term horizon, see [90].

**Model approximation at long-term horizon.** Figure II.5.a and Figure II.5.b show respectively the empirical and the theoretical distribution of  $Q^1$  and  $Q^2$ . Both distributions are close.

(a) Empirical  $Q^1$  distribution  
after 20 events

 (b) Theoretical  $Q^1$  distribution  
after 20 events

 Figure II.4 – (a) Empirical distribution of  $Q^1$  after 20 events and (b) theoretical distribution of  $Q^1$ . The quantities  $Q^1$  and  $Q^2$  are divided by the average event size.

 (a) Empirical distribution of  $(Q^1, Q^2)$ 

 (b) Stationary distribution of  $(Q^1, Q^2)$ 

 Figure II.5 – Empirical distribution of  $Q^1$  on x-axis and  $Q^2$  y-axis in (a) and theoretical one in (b).

### 3 Optimal tactic control problem

#### 3.1 Presentation of the stochastic control framework

We express the control problem for a buy order of a size  $q^0$ . It can be changed to a sell order in an obvious way.

**Order book dynamics.** The agent state is modelled by the process

$$\bar{U}_t^\mu = \left( Q_t^{Bef,\mu}, Q_t^{a,\mu}, Q_t^{Aft,\mu}, Q_t^{2,\mu}, I_t^\mu, S_t^\mu, P_t^\mu, P_t^{Exec,\mu} \right),$$

where  $Q_t^{a,\mu}$  is the size of agent's limit order inserted at the best bid,  $Q_t^{Bef,\mu}$  is the quantity inserted before  $Q_t^{a,\mu}$ ,  $Q_t^{Aft,\mu}$  represents orders inserted after  $Q_t^{a,\mu}$  (see Figure IV.1),  $P_t^{Exec,\mu}$  is the acquisition price of  $q^0 - I_t^\mu$ ,  $I_t^\mu$  is the agent's inventory and  $\mu = \{\mu_t, t \leq T\}$  is the control of

the agent. We recall that  $Q_t^{2,\mu}$  is the best ask limit,  $P_t^\mu$  is the mid price and  $S_t^\mu$  is the spread. Then,  $Q_t^{1,\mu} = Q_t^{Bef,\mu} + Q_t^{a,\mu} + Q_t^{Aft,\mu}$  is the total volume at the best bid. It is split in three quantities to take into account the order placement. We add minor changes to the order book dynamics:

- **Arrival rates:** for the best bid, we differentiate market orders consumption rate  $\lambda_m^{1,-}$  from limit orders cancellation rate  $\lambda_c^{1,-}$ . Cancellation orders consume  $Q_t^{Aft,\mu}$  first, and market orders  $Q_t^{Bef,\mu}$  first.<sup>11</sup>
- **Regeneration:** the regeneration of the process  $\bar{U}_t^\mu$  is deduced from that of  $U_t^\mu = (Q_t^{1,\mu}, Q_t^{2,\mu}, S_t^\mu)$  and  $P_t^\mu$  which are described in Section 2.2. Here, we explain the regeneration of the three variables  $Q_t^{Bef,\mu}$ ,  $Q_t^{a,\mu}$  and  $Q_t^{Aft,\mu}$  since the one of  $Q_t^{2,\mu}$ ,  $S_t^\mu$  and  $P_t^\mu$  is detailed in Section 2.2. After a regeneration  $Q_t^{a,\mu} = 0$  when the best bid is totally depleted and remains unchanged otherwise. Furthermore, the quantity  $Q_t^{Aft,\mu} + Q_t^{Bef,\mu}$  is equal to the regenerated best bid<sup>12</sup> and the position of  $Q_t^{a,\mu}$  is drawn from a distribution  $\zeta_u^i$  depending on the order book state just before the regeneration  $u$  and the depleted side  $i$  (i.e. best ask in our case). A natural choice is to set  $Q_t^{Aft,\mu} = 0$  and  $Q_t^{Bef,\mu}$  equal to the new best bid when the best bid is depleted or the price moves, and keep the quantities  $Q_t^{Bef,\mu}$ ,  $Q_t^{a,\mu}$  and  $Q_t^{Aft,\mu}$  unchanged when the best ask is depleted with no price move.

The symmetry relation (5) satisfied by  $(Q_t^{1,\mu}, Q_t^{2,\mu}, S_t^\mu)$  is unchanged.

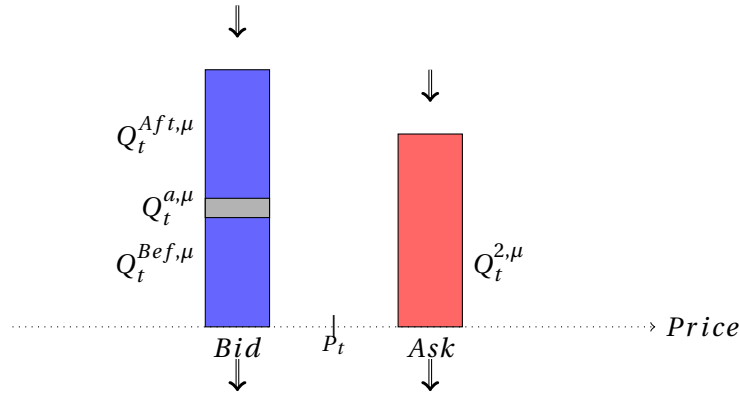


Figure II.6 – Diagram representing the position of the agent order.

**Trader's controls.** At every decision time, the trader can do nothing or take three decisions:

<sup>11</sup>This modelling is conservative since we delay the order execution as long as possible. It corresponds to the worst case scenario for the user.

<sup>12</sup>The best bid  $Q_t^{1,\mu}$  after a depletion is defined in Section 2.2.

## II. Optimal liquidity-based trading tactics

- $l$ : He can insert a fraction of  $I^\mu$  at the top of the bid queue or within the spread if not already inserted.
- $c$ : He can cancel his already existing limit order  $Q^{a,\mu}$ . By acting this way, the trader can wait for a better order book state. This control will essentially be used to avoid adverse selection, i.e.. obtaining a transaction just before a price decrease.
- $m$ : He can send a market order to get an immediate execution of a fraction of  $I^\mu$ .

Every decision of the agent is also characterized by a price level  $p$  and an order size  $q$ . Thus, the trader's control is modelled by the sequence  $\mu = (\tau_i, v_i)_{i \geq 0}$  of random variables where  $(\tau_i)_{i \geq 0}$  is a sequence of increasing stopping times (with respect to the filtration generated by  $\bar{U}^\mu$ ) that represents the optimal decision times and  $v_i \in \mathbb{E}$  is the optimal decision. We define the set  $\mathbb{E}$  such that  $\mathbb{E} = \mathbb{T} \times \mathbb{N} \times \mathbb{P}$  with  $\mathbb{T} = \{l, c, m\}$  and  $\mathbb{P} = \tau_0 \mathbb{N}$ . If the agent has no order inserted in the order book and does nothing at the beginning, the initial control is  $c$ .<sup>13</sup> The price level is equal to  $p \geq 0$  when the order is inserted at the limit price  $P^1 + \min(p, S)$ . We focus on strategies that satisfy the condition below.

**Assumption 1** (Admissible strategies). *There exists a fixed constant  $\bar{P}$  such that the agent liquidates its remaining quantity when  $P^\mu \geq \bar{P}$  (i.e.  $\mu_t = (m, S_t, I_t^\mu) \mathbf{1}_{P_t^\mu \geq \bar{P}}$ ).*

The above assumption ensures the boundedness of the execution price  $P^{Exec,\mu}$ .

**Trader's inventory and liquidation price.** Let  $e = (o, p, q) \in \mathbb{E}$  be a market event. When  $e$  is sent by another market participant, the processes  $Q_t^{a,\mu}$ ,  $I_t^\mu$  and  $P_t^{Exec,\mu}$  jump as follows:

$$\begin{cases} Q_t^{a,\mu} &= Q_{t^-}^{a,\mu} - \mathbf{1}_{o=m} \min((q - Q_{t^-}^{Bef,\mu})_+, Q_{t^-}^{a,\mu}), \\ I_t^\mu &= I_{t^-}^\mu + \mathbf{1}_{o=m} \Delta Q_t^{a,\mu}, \\ P_t^{Exec,\mu} &= P_{t^-}^{Exec,\mu} - \Delta I_t^\mu (P_{t^-}^\mu - \frac{S_{t^-}^\mu}{2}), \end{cases}$$

with  $\min((q - Q_{t^-}^{Bef,\mu})_+, Q_{t^-}^{a,\mu})$  the quantity bought at the best bid price  $P_{t^-}^\mu - \frac{S_{t^-}^\mu}{2}$ ,  $\Delta X_t = X_t - X_{t^-}$  for any càdlàg process  $X$  and  $(x)_+ = \max(x, 0)$ . When  $e$  is sent by the agent, the processes  $Q_t^{a,\mu}$ ,  $I_t^\mu$  and  $P_t^{Exec,\mu}$  are updated such that

$$\begin{cases} Q_t^{a,\mu} &= \mathbf{1}_{o=l} q + \mathbf{1}_{o=m} Q_{t^-}^{a,\mu}, \\ I_t^\mu &= I_{t^-}^\mu - \mathbf{1}_{o=m} \min(q, I_{t^-}^\mu - Q_{t^-}^{a,\mu}), \\ P_t^{Exec,\mu} &= P_{t^-}^{Exec,\mu} - \Delta I_t^\mu [P_{t^-}^\mu + \frac{S_{t^-}^\mu}{2} + \alpha(-\Delta I_t^\mu - Q_{t^-}^{2,\mu})_+], \end{cases}$$

where the parameter  $\alpha$  represents a linear temporary price impact. At the end of the execution we have the following terminal condition (i.e. liquidation of the remaining inventory):

$$\begin{cases} Q_T^{a,\mu} &= 0, \\ I_T^\mu &= 0, \\ P_T^{Exec,\mu} &= P_{T^-}^{Exec,\mu} + I_{T^-}^\mu [P_{T^-}^\mu + \frac{s}{2} + \alpha(I_{T^-}^\mu - Q_{T^-}^{2,\mu})_+]. \end{cases}$$

<sup>13</sup>To track only the first limits, we assume that the agent follows the best bid.

**Optimal control problem.** We fix a finite horizon time  $T < \infty$  and we want to compute

$$V_T(0, u) = \sup_{\mu} \mathbb{E} \left[ f \left( \underbrace{\lim_{s \rightarrow \infty} \mathbb{E}[\Delta P_s^\mu / \mathcal{F}_{T_{Exec}^\mu}]}_{\text{final constraint}} \right) - \underbrace{\gamma \int_0^{T_{Exec}^\mu} I_s^\mu ds}_{\text{running cost}} \right],$$

where

- $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec})$  is the initial agent state.
- $T_{Exec}^\mu = \inf\{t \geq 0, s.t. I_t^\mu = 0\} \wedge T$  represents the final execution time.
- $\Delta P_t^\mu = (q^0 P_t^\mu - P_{T_{Exec}^\mu}^{Exec, \mu})$  represents the price impact and  $q^0$  is the order size.<sup>14</sup>
- $\gamma$  is a non-negative homogenization constant representing the waiting cost or the risk aversion of the agent and  $f: \mathbb{R} \rightarrow \mathbb{R}$  is a Lipschitz function.

We solve the agent's control problem in two situations: when decisions are taken at fixed frequency  $\Delta^{-1}$  and when they are taken at any time.

## 4 Theoretical results

In this section, we compute  $\lim_{t \rightarrow \infty} \mathbb{E}[\Delta P_t^\mu / \mathcal{F}_{T_{Exec}^\mu}]$ , discuss the existence, uniqueness and regularity of the solution of our control problem and give equations satisfied by the value function.

### 4.1 Price impact computation

In this section Assumptions 3, 4 and 5 are satisfied. Let  $U_{T_{Exec}^\mu}^\mu$  be the order book state at the end of the execution, we split  $\Delta P_t^\mu$  in two quantities

$$\begin{aligned} \Delta P_t^\mu &= (q^0 P_t^\mu - P_{T_{Exec}^\mu}^{Exec, \mu}) = (q^0 P_t^\mu - q^0 P_{T_{Exec}^\mu}^\mu) + (q^0 P_{T_{Exec}^\mu}^\mu - P_{T_{Exec}^\mu}^{Exec, \mu}) \\ &= q^0 \Delta P_t^{\prime, \mu} + (q^0 P_{T_{Exec}^\mu}^\mu - P_{T_{Exec}^\mu}^{Exec, \mu}), \end{aligned}$$

where

- $P_{T_{Exec}^\mu}^\mu$  is the mid price at the execution (i.e.  $P_{T_{Exec}^\mu}^\mu$  and  $P_{T_{Exec}^\mu}^{Exec, \mu}$  are known at the execution).
- $\Delta P_t^{\prime, \mu} = P_t^\mu - P_{T_{Exec}^\mu}^\mu$  is the long term mid price move after the execution.

---

<sup>14</sup>We will see that  $\lim_{t \rightarrow \infty} \mathbb{E}[\Delta P_t^\mu / \mathcal{F}_{T_{Exec}^\mu}]$  is well-defined and an explicit computation of this quantity is given in Section 4.1.

Thus, we only need to compute  $\Delta P_t^{\prime,\mu}$ . Since we place ourselves after the execution, we have  $Q_{T_{Exec}^\mu}^{a,\mu} = 0$  and  $Q_{T_{Exec}^\mu}^{Aft,\mu} = 0$  which means that  $Q_{T_{Exec}^\mu}^{1,\mu} = Q_{T_{Exec}^\mu}^{Bef,\mu}$ . Since the price jumps are a deterministic function of  $U^\mu$  before each jump and  $U^\mu$  is Markov<sup>15</sup>, we have  $\lim_{t \rightarrow \infty} \mathbb{E}[\Delta P_t^{\prime,\mu} / \mathcal{F}_{T_{Exec}^\mu}^\mu] = \lim_{t \rightarrow \infty} \mathbb{E}_{U_{T_{Exec}^\mu}^\mu} [\Delta P_t^{\prime,\mu}]$ . Proposition 1 provides an explicit formula for the computation of  $\lim_{t \rightarrow \infty} \mathbb{E}_{U_{T_{Exec}^\mu}^\mu} [\Delta P_t^{\prime,\mu}]$ .

## 4.2 Existence and uniqueness of the optimal strategy, regularity properties

In the rest of the article, Assumptions 1, 3, 4 and 5 are in force. In this section, we discuss existence and uniqueness of the optimal strategy and show regularity results for the state process  $U^\mu$  and the value function. First, for a finite horizon time  $T$ , we define the value function

$$V_T(t, \bar{u}) = \sup_{\mu} \mathbb{E} \left[ f \left( \lim_{s \rightarrow \infty} \mathbb{E}[\Delta P_s^\mu / \mathcal{F}_{T_{Exec}^\mu}^\mu] \right) - \gamma \int_t^{T_{Exec}^\mu} I_s^\mu ds \mid \bar{U}_t^\mu = \bar{u} \right],$$

with  $0 \leq t \leq T$ ,  $\bar{u} \in \bar{\mathbb{U}} = \mathbb{N}^5 \times \tau_0 \mathbb{N}^* \times (\frac{\tau_0}{2} \mathbb{Z})^2$  and  $T_{Exec}^{t,\mu} = \inf\{s \geq t, s.t. I_s^\mu = 0\} \wedge T$ .<sup>16</sup>

**Existence - uniqueness of the optimal control.** The optimal strategy exists in the two frameworks (i.e. decisions taken at fixed frequency and at any time) but for different reasons. When decisions are taken at fixed frequency  $\Delta^{-1}$  the optimal strategy exists since we have a finite number of available strategies. When decisions are taken at any time, the sequence of optimal controls  $(\tau_i, v_i)_{i \geq 0}$ , where  $\tau_i$  is the optimal decision time and  $v_i$  is the optimal decision, satisfies

$$\begin{cases} \tau_0 = 0, & v_0 = \operatorname{argmax}_{r \in \mathbb{E}} \{ \mathbb{E}[V_T(0, \bar{U}_0^r)] \}, \\ \tau_{i+1} = \inf\{t > \tau_i; V_T(t, \bar{U}_t^{\hat{\mu}}) = \sup_{r \in \mathbb{E}, r \neq v_i} \mathbb{E}[V_T(t, (\bar{U}_t^{\hat{\mu}})^r)]\}, \\ v_{i+1} = \operatorname{argmax}_{r \in \mathbb{E}, v_i \neq r} \{ \mathbb{E}[V_T(\tau_{i+1}, (\bar{U}_{\tau_{i+1}}^{\hat{\mu}})^r)] \}, \end{cases} \quad (7)$$

where  $\bar{u}^r$  is the new state when the agents takes the decision  $r$  and current state is  $\bar{u}$  and  $\hat{\mu} = (\tau_j, v_j)_{j \leq i}$  for any state  $\bar{u}$  and control  $r$ . Since  $V_T$  is continuous, the optimal control is well-defined. The proof of (7) is given in Appendix II.E. However, there is a priori no uniqueness of the optimal strategy in the two frameworks.

**Regularization of the problem.** To force the uniqueness of the optimal strategy, we present a practical criterion. First, we define an order relation between trader's decisions  $c < l < m$ . The intuition behind is that  $m$  is the least risky decision because we get direct execution,  $l$  is riskier than  $m$  but less risky than  $c$  because there is no delay of the execution. Then, we order sizes (resp. prices) increasingly (resp. decreasingly) since a large size accelerates the liquidation of the inventory and a small posting price reduces instantaneously the impact. Finally, we

<sup>15</sup> $U^\mu$  does not depend on  $\mu$  after the execution since the agents leaves the market.

<sup>16</sup>The results remain valid when  $\bar{\mathbb{U}} = \mathbb{R}_+^5 \times \mathbb{R}_+^* \times (\mathbb{R})^2$ .

consider the following order relation between  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , and  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ :

$$x \leq y \text{ iff } (x_1 < y_1) \text{ or } (x_1 = y_1 \text{ and } (x_2, \dots, x_n) \leq (y_2, \dots, y_n)).$$

Hence, we can choose the least risky decision among the optimal ones in the above sense.

**Regularity of the state process and the value function.** The agent state process  $\bar{U}^\mu$  and the value function  $V_T$  are Lipschitz in space and time. The proofs of these results are given in Appendix II.D in the more general case where  $\bar{U} = \mathbb{R}_+^5 \times \mathbb{R}_+^* \times (\mathbb{R})^2$ .

### 4.3 Decisions taken at fixed frequency $\Delta^{-1}$ : dynamic programming equation

In this section, we provide and solve the system of equations satisfied by the value function  $V_T$  of the optimal control problem. The constant  $\bar{P}$  is defined in Section 3.1. We have the following result.

**Theorem 1.** *Let  $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec})$  be an initial state and  $t \in [0, T]$ . Then  $V(t, u)$  satisfies*

- When  $i > 0$  and  $p < \bar{P}$ :

- At the decision time  $t = k\Delta < T$ :

$$V(k\Delta, \cdot) = \sup_{e \in \mathbb{E}} (V^e((k\Delta)_+, \cdot)), \quad (8)$$

where  $V(t, \cdot)$  and  $V^e(t, \cdot)$  are vectors such that  $V(t, \cdot)_i = V(t, u_i)^{10}$  and  $V^e(t, \cdot)_i = \mathbb{E}[V(t, u_i^e)]$ , where  $u_i^e$  is the new order book state when the decision  $e \in \mathbb{E}$  is taken. Equation (8) should be understood coordinate by coordinate.

- At  $t \neq k\Delta < T$ :

$$0 = -\gamma \mathbf{I} + \mathcal{A}V(t, \cdot), \quad (9)$$

where  $\mathcal{A} = \partial_t + Q$  is the infinitesimal generator of the process  $U_t^\mu$ .

- When  $i = 0$  or  $p \geq \bar{P}$  (execution time condition):

$$V(t, u) = \tilde{g}(u), \quad \forall t < T, \quad (10)$$

with  $\tilde{g}(u) = f(\lim_{t \rightarrow \infty} \mathbb{E}_u[\Delta P_t])$ .

- The terminal condition is:

$$V(T, u) = g(u), \quad (11)$$

with  $g(u) = f(\mathbb{E}[\lim_{t \rightarrow \infty} \mathbb{E}_{u^{m^t}}[\Delta P_t]])$  and the decision  $m^t$  represents the liquidation of the remaining inventory. We keep in mind that a control  $m$  may lead to several states because of the regeneration.



The proof of this result is given in Appendix II.E.

**Remark 1.** *At every decision time, as long as the order is not executed, the agent compares the value function given by each control and takes the highest one, see Equation (8). When, the order is executed, the agent gain is  $\tilde{g}(u)$  with  $u$  the order book state at the end of the execution. If the order is not executed before  $T$ , the agent send a market order to obtain immediate execution and earn  $g(u)$ .*

**Remark 2.** *Without the control  $c$  and  $l$ , Equations (8), (9) and (11) are equivalent, in dimension 1, to the classical problem of finite horizon Bermudean options.*

**Explicit solution.** It is possible to exhibit a solution  $V$  for equations of Theorem 1 under a mild assumption. The solution is constructed backward and step by step within each interval  $[k\Delta, (k+1)\Delta \vee T)$  with  $k \leq k_1 = \lfloor \frac{T}{\Delta} \rfloor$  an integer.

- *Step 1 - Initialization:* We take  $k = k_1$  and place ourselves in  $[k\Delta, (k+1)\Delta \vee T)$ . Note that we can reformulate equations of Theorem 1 to obtain that  $V$  satisfies

$$\begin{cases} V(T, u) = g(u), \\ 0 = -\gamma \mathbf{I} + \tilde{\mathbf{g}} + \mathcal{A}V, \quad \forall t \in [k\Delta, (k+1)\Delta \vee T), \end{cases} \quad (12)$$

where the vector  $\tilde{\mathbf{g}}$  incorporates the execution time constraint associated to (10) and  $\mathcal{A} = \partial_t + \tilde{Q}$ . The operator  $\tilde{Q}$  is obtained by removing all the transitions to states where the inventory is zero or the price exceeds  $\tilde{P}$ . We can give the following explicit expression for  $\tilde{\mathbf{g}}$  and  $\tilde{Q}$ :

$$\begin{cases} \tilde{\mathbf{g}}(u) = \mathbf{1}_{i \neq 0, p < \tilde{P}} + \sum_{q^{bef} + q^a \leq n < q^1} \lambda_m^{1,-}(u, n, \mathbf{0}) \tilde{\mathbf{g}}(u) + \mathbf{1}_{i \neq 0, p < \tilde{P}} \sum_{u', n \geq q^1} \lambda_m^{1,-}(u, n, u') \tilde{\mathbf{g}}(u'), \\ \tilde{Q}(u, u') = Q(u, u') \mathbf{1}_{i \neq 0, p < \tilde{P}, i' = 0}, \end{cases}$$

where  $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec}) \in \bar{\mathbb{U}}$ ,  $u' = (q^{bef}, q^a, q^{aft}, q^2, i', s', p', p^{exec}) \in \bar{\mathbb{U}}$  and  $q^1 = q^{bef} + q^a + q^{aft}$ . Let  $Im(\tilde{Q})$  and  $Ker(\tilde{Q})$  be respectively the image and the kernel of  $\tilde{Q}$ . We consider the assumption  $-\gamma \mathbf{I} + \tilde{\mathbf{g}} \in Im(\tilde{Q}) + Ker(\tilde{Q})$ <sup>17</sup> which means that

$$-\gamma \mathbf{I} + \tilde{\mathbf{g}} = \tilde{\mathbf{g}}_{Im(\tilde{Q})} + \tilde{\mathbf{g}}_{Ker(\tilde{Q})},$$

with  $\tilde{\mathbf{g}}_{Im(\tilde{Q})} \in Im(\tilde{Q})$  and  $\tilde{\mathbf{g}}_{Ker(\tilde{Q})} \in Ker(\tilde{Q})$ . Since  $\tilde{\mathbf{g}}_{Im(\tilde{Q})} \in Im(\tilde{Q})$ , there exists  $\tilde{z}$  such that  $\tilde{Q}\tilde{z} = \tilde{\mathbf{g}}_{Im(\tilde{Q})}$ . Then, we can check that the following variable is solution of (12):

$$V_t^0 = e^{(T-t)\tilde{Q}} g + (T-t)\tilde{\mathbf{g}}_{Ker(\tilde{Q})} - \tilde{z}, \quad \forall t \in (k\Delta, T), \quad (13)$$

where  $g$  is the vector such that  $g_i = g(\bar{U}_i)$ . Indeed, we have

$$\begin{aligned} \mathcal{A}V_t^0 &= \partial_t V_t^0 + \tilde{Q}V_t^0 = -\tilde{Q}e^{(T-t)\tilde{Q}} g - \tilde{\mathbf{g}}_{Ker(\tilde{Q})} \\ &\quad + \tilde{Q}e^{(T-t)\tilde{Q}} g + (T-t) \underbrace{\tilde{Q}\tilde{\mathbf{g}}_{Ker(\tilde{Q})}}_{=0} - \tilde{Q}\tilde{z} \\ &= -[\tilde{\mathbf{g}}_{Ker(\tilde{Q})} + \tilde{\mathbf{g}}_{Im(\tilde{Q})}] = \gamma \mathbf{I} - \tilde{\mathbf{g}}, \end{aligned}$$

which ensures that  $V_t^0$  is solution of (12).

<sup>17</sup>This assumption verified for instance when  $\tilde{Q}$  is revertible or symmetric.

- *Step 2 - Iteration:* At time  $k\Delta$ , the agent can take a decision. So he compares expressions of Equation (8) and takes the maximum. After that, he reiterates Step 1 with new initial values and  $k = k_1 - 1$ .

#### 4.4 Second approach: Decisions taken at any time

Let us now consider the case where the agent takes a decision at any time. In this section, we provide the system of equations satisfied by the value function and we also introduce a simplified control problem whose value function can be easily computed numerically, and converges towards the one of the initial optimal control problem.

##### 4.4.1 Dynamic programming equation

We keep the same notations as in Theorem 1. We have the following result for the value function in this setting.

**Theorem 2.** *Let  $u = (q^{bef}, q^a, q^{aft}, q^2, i, s, p, p^{exec})$  be an initial state and  $t \in [0, T]$ . Then  $V(t, u)$  satisfies in the viscosity sense and almost everywhere*

- *When  $i > 0$  and  $p < \bar{P}$ :*

$$\max \left( \mathcal{A}V(t, \cdot) - \gamma \mathbf{I}, \sup_{e \in \mathbb{E}} V^e(t, \cdot) - V(t, \cdot) \right) = 0. \quad (14)$$

- *When  $i = 0$  or  $p \geq \bar{P}$  (execution time condition):*

$$V(t, u) = \tilde{g}(u), \forall t \leq T.$$

- *The terminal condition is:*

$$V(T, u) = g(u). \quad (15)$$

The proof of the result is given in Appendix II.E. Since  $\partial_t V$  is a priori not continuous, we use the notion of viscosity. However we show that  $\partial_t V$  is continuous except on the boundary of  $\{V = g\}$  and the above equations are satisfied pointwise except on this boundary, see Appendix II.E.

**Remark 3.** *When there is no control  $c$  and  $l$ , Equations (14) and (15) are equivalent, in dimension 1, to the classical problem of finite horizon American option.*

##### 4.4.2 Numerical resolution of the optimal execution problem

To solve numerically the preceding optimal control problem, we consider a discrete framework. We show here how this discrete framework can be used to approximate the solution of the continuous control problem. Furthermore, an error estimate is provided.

**Discrete-time Markov chain approximation.** Let  $U_n^\Delta$  be a Markov chain with a transition matrix  $P^\Delta$  defined by

$$P_{u,u'}^{*,\Delta} = \mathbb{P}[U_\Delta = u' | U_0 = u], \quad \forall (u, u') \in \mathbb{U}^2,$$

with  $U_t$  the process defined in Section 2.2. Given the infinitesimal generator  $\mathcal{Q}$  of  $U$ , the transition matrix  $P^{*,\Delta}$  can be easily computed since  $P^{*,\Delta} = e^{\Delta\mathcal{Q}}$ . In this approximation,  $U_n^\Delta$  is viewed as the market evolution without the intervention of the agent. Associated to this new market, we introduce the controlled discrete-time Markov chain

$$U_n^{\Delta,\mu} = \left( Q_n^{Bef,\mu}, Q_n^{a,\mu}, Q_n^{Aft,\mu}, Q_n^{2,\mu}, I_n^\mu, S_n^\mu, P_n^\mu, P_n^{Exec,\mu} \right),$$

by using the same construction as in Section 3.1. Additionally, the price impact  $PI_\infty^\Delta$  in this discrete-time approximation can be computed by following the same approach of Section 4.1. Finally, for every  $k \geq 0$ , we define the piecewise constant process  $\tilde{U}^{\Delta,\mu}$  associated to  $U_n^{\Delta,\mu}$  such that

$$\tilde{U}_t^{\Delta,\mu} = U_k^{\Delta,\mu}, \quad \forall t \in [k\Delta, (k+1)\Delta).$$

We denote by  $\tilde{V}^\Delta(t, U)$  the value function of the control problem 3.1 where the process  $U^\mu$  is replaced by  $\tilde{U}^{\Delta,\mu}$  and the agents takes its decisions only at times  $k\Delta$  with  $k \in \mathbb{N}$ .

**Solving numerically the optimal control problem in the discrete framework.** We denote by  $V^\Delta(n, u)$  the value function associated to the discrete control problem (i.e. the state process  $U_n^\Delta$ ), with  $n$  the period and  $u$  the order book state. The dynamic programming principle reads

$$V^\Delta(i, u) = \sup_{e \in \mathbb{E}} \mathbb{E} \left[ V^\Delta \left( (i+1), U_{i+1}^{\Delta,\mu} \right) - \gamma I \Delta | U_i^{\Delta,\mu} = u, \mu_i = e \right], \quad (16)$$

with the terminal constraint  $V^\Delta(n_f, u) = g(u)$ ,  $I$  the agent's inventory,  $\mu$  the control of the agent and  $n_f$  the final period. Equation (16) provides a numerical scheme to compute  $V^\Delta(0, u)$ . At the final time  $T$ , we can compute  $V^\Delta(n_f, u)$  for each reachable state. Using the backward Equation (16), we can compute  $V^\Delta(i, u)$  knowing  $V^\Delta(i+1, u)$  to get the initial value  $V^\Delta(0, u)$ . Thus, one can estimate  $\tilde{V}^\Delta(n\Delta, u)$  since  $\tilde{V}^\Delta(n\Delta, u) = V^\Delta(n, u)$ . The numerical results of simulations are presented in Section 5. To compute efficiently the value function, computations can be carried on in parallel. Let us consider the following assumption.

**Assumption 1.** *There exists a constant  $D$  such that*

$$\|P^\Delta - P^{*,\Delta}\|_F \leq D\Delta^2,$$

with  $\|\cdot\|_F$  the Frobenius norm.

**Remark 4.** *The finite difference scheme associated to the equations of Theorem 2 is the application of the discrete-time approximation with a Markov chain  $\tilde{U}_n^\Delta$  whose transition matrix  $\tilde{P}^\Delta = I + Q\Delta$ . Note that  $\tilde{P}^\Delta$  satisfies Assumption 1.*

Let  $\mu^{Opti,\Delta}$  (resp.  $\mu^{Opti}$ ) be the optimal control associated to the process  $\tilde{U}_t^{\Delta,\mu}$  (resp.  $U_t^\mu$ ) and  $H$  an upper bound for the intensity functions that exists under Assumptions 3 and 4. Note that  $g$  is bounded by  $\|g\|_\infty$  since we focus on strategies that liquidate the whole inventory when the mid price get beyond  $\bar{P}$ . Then we have the following error estimate result.

**Theorem 3.** *Under Assumption 1,  $\tilde{V}^\Delta(t, u)$  converges towards  $V(t, u)$  for every  $(t, u) \in [0, T] \times \bar{U}$ . Additionally, we have the following error estimate:*

$$|\tilde{V}^\Delta(t, u) - V(t, u)| \leq R(T - t)\Delta, \quad (17)$$

with  $R = 4cq^a H + D\|g\|_\infty$ . Moreover, the control  $\mu^{Opti,\Delta}$  converges  $\mu^{Opti}$  and for  $\Delta$  small enough we have the equality  $\mu^{Opti,\Delta} = \mu^{Opti}$ , a.s.

The proof of this result is given in Appendix II.F.

## 5 Numerical experiments

In this section, we show the relevance of the optimal strategy in both frameworks: when decisions are taken at fixed frequency  $\Delta^{-1}$  and when they are taken at any time. To do so, we compare the optimal gain given by our strategy and the one given by the standard strategy join the bid: stay in the order book at the best bid until the final time. Here, we write  $Q^1$  (resp.  $Q^2$ ) for the best bid (resp. ask) limit.

### 5.1 Computation of the optimal gain: decisions taken at a fixed frequency $\Delta^{-1}$

Figure II.7 shows for an order of size 1 the difference between the average gain (i.e. the initial value function) of the optimal strategy and the one of the strategy join the bid for different values of the initial  $Q^1$  and  $Q^2$ . The gain of the optimal strategy is obviously always higher than that of the strategy stay in the order book. However, because of the priority value, that is the advantage of a limit order compared with another limit order standing at the rear of the same queue, it is more useful to be active (i.e. cancel the order or send a market order) when imbalance is highly positive than when it is negative. Finally, note that the optimal strategy reaches the maximum value of 2.4 ticks (since the tick  $\delta = 1$  cent).

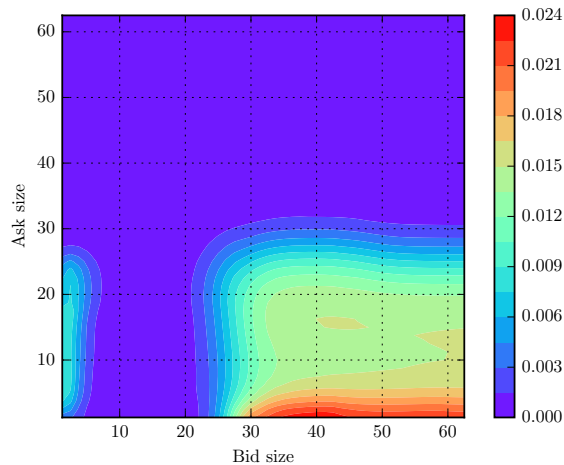


Figure II.7 - Difference between the optimal gain of the optimal strategy and the one of the strategy stay at the best bid. The initial parameters are fixed as follows: the time frequency is equal to  $\Delta = 10$  seconds, the final time  $T = 100$  seconds, arrival and consumption rates are estimated on data (see Appendix II.A), the new bid (resp. ask) is set to 5 and the new ask (resp. bid) to 3 after the total depletion of the bid (resp. ask) limit, the quantity  $q^0 = 1$ , the waiting cost  $c = 0$ , the price increases (resp. decreases) by  $\delta = 1$  cent when the ask limit (resp. bid limit) is totally consumed and the function  $f$  is equal to the identity.

## 5.2 Computation of the optimal gain: decision taken at any time

Figure II.8 shows the value function at time zero (i.e. trader's gain) of the optimal strategy in red and the one of the strategy stay at the best bid in blue in percentage of the tick  $\delta = 1$  cent using the discrete approximation. The points colors refer to the initial decision given by the strategy: green points means stay in the order book at the beginning is the best decision, red points means cancel is the best initial decision and black points means send a market order is the best initial decision. When imbalance is highly negative, it is optimal to cancel the order to avoid adverse selection, when imbalance is highly positive it is optimal to send a market order or stay in the order book. In our case, stay in the order book is interesting when imbalance is highly positive since the priority value is important (i.e.  $Q^{Bef}$  is fixed equal to 1). In mid cases (i.e. imbalance close to 0), it is optimal to send a market order to reduce the waiting cost. We note that the gain of the optimal strategy is significantly better than the one of the strategy join the bid.

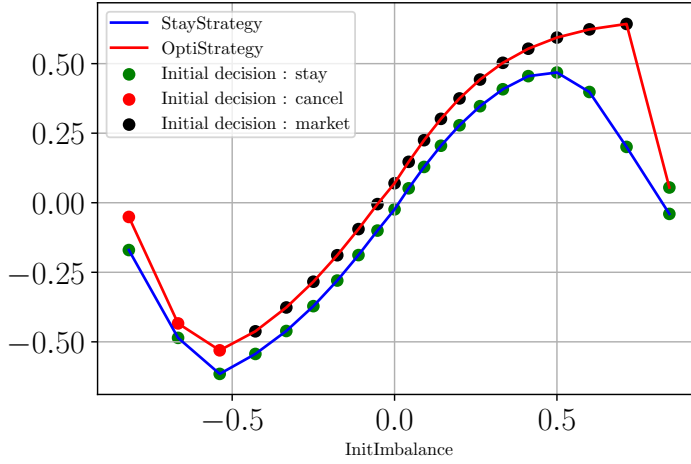


Figure II.8 – The gain per tick of the optimal strategy in red and the one of the strategy join the bid in blue for different values of the initial imbalance. Initial imbalances are obtained with  $Q^{Bef} = 0$ ,  $Q^1 = 11$  and  $Q^2$  from 1 to 11, and  $Q^2 = 11$  and  $Q^1$  from 10 to 1. Initial parameters are as follows: the time step is equal to  $\Delta = 1$  second, there are 10 periods, arrival and consumption rates are constant  $\lambda^{1,+} = \lambda^{2,+} = 0.06$  and  $\lambda^{1,-} = \lambda^{2,-} = 0.12$ , the new bid (resp. ask) is set to 5 and the new ask (resp. bid) to 3 after the total depletion of the bid (resp. ask) limit, the quantity  $q^0 = 1$ , the waiting cost  $c = 0.0085$ , the price increases (resp. decreases) by  $\delta = 1$  cent when the ask limit (resp. bid limit) is totally consumed and the function  $f$  is the identity.

The authors gratefully acknowledge the financial support of the ERC grant 679836 Staquamof and the Chair Analytics and Models for Regulation.

## II.A Model parameters estimation

The estimation methodology of the arrival and cancellation rates of limit orders is similar to that in [90]. The regeneration distribution of the order book is estimated from the empirical distribution of order book states after a depletion. In what follows, we provide the calibration results of our order book model using the database described in Section 2.1. Here, we write  $Q_t = (Q_t^1, Q_t^2)$  with  $Q_t^1$  (resp.  $Q_t^2$ ) the best bid (resp. ask) quantity and consider that intensities and regeneration distributions depend only on  $Q_t$ .

**Intensities estimation.** For every  $Q = (Q^1, Q^2)$ , we write  $\tau^{1,+}(Q) = \lambda^{1,+}/\lambda^{1,-}$  and  $\tau^{2,+}(Q) = \lambda^{1,+}/\lambda^{1,-}$  respectively for the bid and ask side growth ratios. Given the bid-ask symmetry relation, we can aggregate data and focus on the bid side only. Figures II.9.a, II.9.b, II.9.c and II.9.d show respectively  $\lambda^{1,+}$ ,  $\lambda^{1,-}$ ,  $\tau^{1,+}$  and  $\tau^{2,+}$  for different values of  $Q$ . As expected, we can see that participants insert more limit orders when the imbalance is negative (see Figure II.9.a when  $Q^2 \gg Q^1$ ) while they cancel more when the imbalance is positive (see

## II. Optimal liquidity-based trading tactics

Figure II.9.b when  $Q^1 \gg Q^2$ ). Finally, Figure II.9.c (resp. Figure II.9.d) shows that  $\tau^{1,+}$  (resp.  $\tau^{2,+}$ ) is high when imbalance is negative (resp. positive) and becomes low when imbalance is positive (resp. negative) which means that the bid limit (resp. ask limit) tends to increase (resp. decrease) when  $Q^1 \ll Q^2$  and tends to decrease (resp. increase) when  $Q^1 \gg Q^2$ .

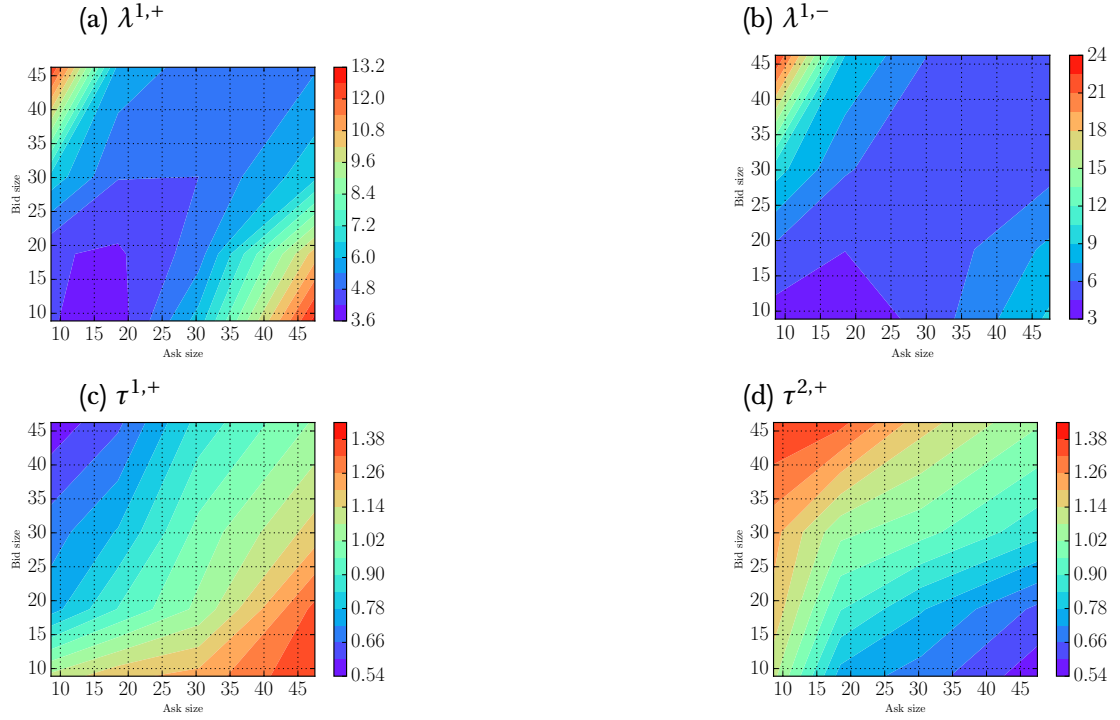


Figure II.9 – (a)  $\lambda^{1,+}$ , (b)  $\lambda^{1,-}$ , (c)  $\tau^{1,+}$  and (d)  $\tau^{2,+}$  for different values of  $(Q^1, Q^2)$ .  $Q^1$  and  $Q^2$  are divided by the average event size.

**Quantities after depletion.** When one limit is depleted, we write  $Q^{New,1}$  (resp.  $Q^{New,2}$ ) for the new best bid (resp. ask). Figures II.10.a, II.10.b and II.10.c show respectively  $Q^{New,1}$ ,  $Q^{New,2}$  and the ratio  $r^+(Q_1, Q_2) = \frac{Q^{New,1}}{Q^{New,2}}$  for different values of  $Q^1$  and  $Q^2$  before the mid price move. Since we aggregate data, the bid queue is always the depleted queue and the ask limit is the non-consumed limit. Figures II.10.a and II.10.b show that  $Q^{New,1}$  depends mainly on  $Q^2$  while  $Q^{New,2}$  depends on both  $Q^1$  and  $Q^2$ . However, the interesting point is that  $r^+$  reach its maxima in two cases, see Figure II.10.c. The first case, when the bid is low and the ask is high, can be explained by a mean reversion effect while the second one, when both queues are initially high, is due to the arrival of a large order consuming market liquidity.

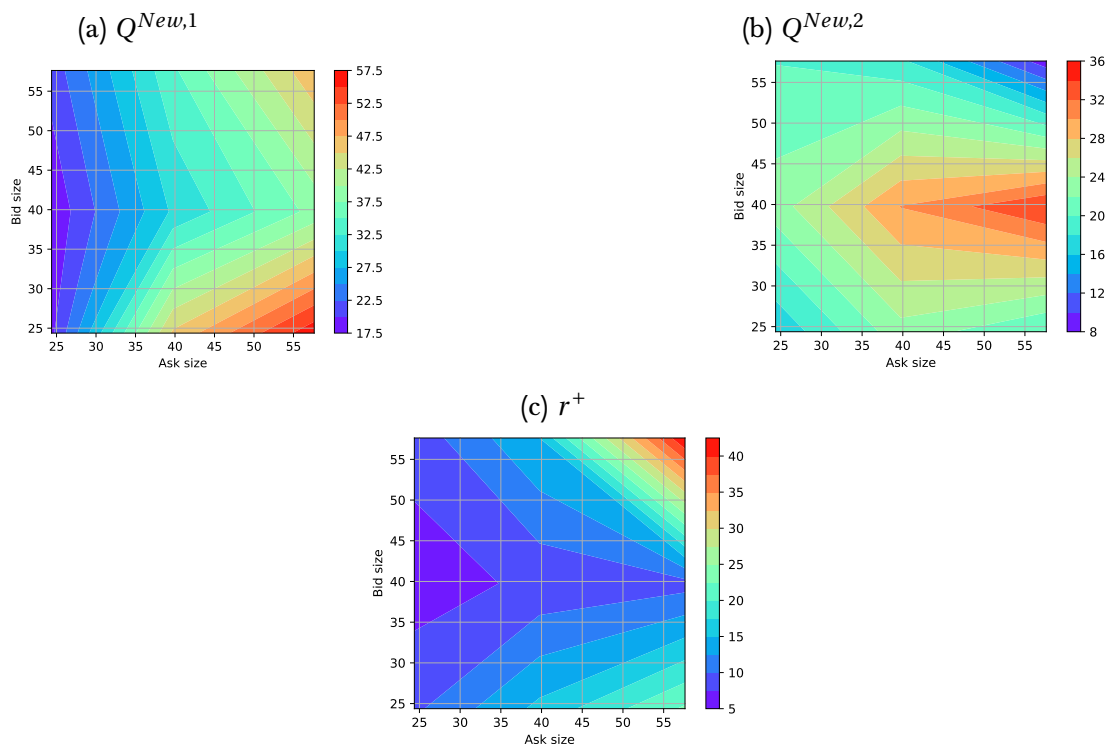


Figure II.10 - (a)  $Q^{New,1}$ , (b)  $Q^{New,2}$  and (c)  $r^+$  for different values of  $Q^1$  and  $Q^2$ .  $Q^1$  and  $Q^2$  are divided by the average event size.



## II.B Ergodicity of the process ( $U_t$ )

### II.B.1 Outline of the proof

Let  $Z_t$  be a Markov process defined on the probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  and valued in  $(W, \mathcal{W})$  and  $P_t(x, A)$  the probability transition of  $Z_t$ .

**Definition II.B.1** (Ergodicity). *The process  $Z_t$  is ergodic if there exists an invariant probability measure  $\pi$  which satisfies*

$$\lim_{t \rightarrow \infty} \|P_t(x, \cdot) - \pi(\cdot)\|_{TV} = 0,$$

where  $\|\mu - \mu'\|_{TV} = \sup_{A \in \mathcal{F}} |\mu(A) - \mu'(A)|$ .

To prove that  $U_t$  is ergodic, we design a Lyapunov function  $V : \mathbb{U} \rightarrow (0, \infty)$ , on which the following negative drift condition is satisfied for some  $c > 0$  and  $d > 0$ :

$$\mathcal{Q}V(q) \leq -cV(q) + d,$$

$\mathcal{Q}$  be the infinitesimal generator of  $U_t$ . Then, using Theorem 6.1 in [122], the Markov process  $U_t$  is non-explosive and V-uniformly ergodic. Furthermore, by Theorem 4.2 in [122], it is Harris positive recurrent.

### II.B.2 Proof

Let  $u = (q^1, q^2, s) \in \mathbb{U}$  and  $z = \min(z_0, z_1)$ , we define

$$V(u) = \sum_{i \in \{1, 2\}} z^{q^i - C_{bound}} + z^{s - C_{bound}}.$$

To simplify notations we do not write the dependence of  $\lambda_Q^{i, \pm}(n)$  and  $\lambda_S^\pm(n)$  on  $u$ . For any  $t \geq 0$ , we have

$$\mathcal{Q}_t V(u) = \sum_{u' \neq u} \mathcal{Q}_{t, u, u'} [V(u') - V(u)].$$

By rearranging the above terms, we get

$$\begin{aligned} \mathcal{Q}_t V(u) &= \sum_{i=1}^2 \sum_{1 \leq n} \left[ \lambda_Q^{i,+}(n) (z^{q^i+n-C_{bound}} - z^{q^i-C_{bound}}) + \lambda_Q^{i,-}(n) (z^{q^i-n-C_{bound}} - z^{q^i-C_{bound}}) \right] \\ &\quad + \sum_{1 \leq n} \left[ \lambda_S^+(n) (z^{s+n\tau_0-C_{bound}} - z^{s-C_{bound}}) + \lambda_S^-(n) (z^{s-n\tau_0-C_{bound}} - z^{s-C_{bound}}) \right] \\ &= \sum_{i=1}^2 z^{q^i-C_{bound}} \sum_{1 \leq n} (z^n - 1) \left[ \lambda_Q^{i,+}(n) - \lambda_Q^{i,-}(n) \frac{1}{z^n} \right] \\ &\quad + z^{s-C_{bound}} \sum_{1 \leq n} (z^{n\tau_0} - 1) \left[ \lambda_S^+(n) - \lambda_S^-(n) \frac{1}{z^{n\tau_0}} \right]. \end{aligned} \tag{18}$$

**Step (i).** When  $q^i \leq C^{bound}$ , the quantity  $z^{q^i - C^{bound}}$  is bounded and the intensities are bounded by Assumption (2). Hence, there exist  $c^1 > 0$  and  $d^1 > 0$  such that

$$\begin{aligned} z^{q^i - C^{bound}} \sum_{1 \leq n} (z^n - 1) \left[ \lambda_Q^{i,+}(n) - \lambda_Q^{i,-}(n) \frac{1}{z^n} \right] &\leq z^{q^i - C^{bound}} \left( \sum_{1 \leq n} z^n \lambda_Q^{i,+}(n) \right) \\ &\leq -c^1 z^{q^i - C^{bound}} + d^1. \end{aligned} \quad (19)$$

Similarly, when  $s \leq C^{bound}$ , we have

$$z^{s - C^{bound}} \sum_{1 \leq n} (z^{n\tau_0} - 1) \left[ \lambda_S^+(n) - \lambda_S^-(n) \frac{1}{z^{n\tau_0}} \right] \leq -c^1 z^{s - C^{bound}} + d^1. \quad (20)$$

**Step (ii).** Using Assumption (1), we deduce that

$$\begin{aligned} z^{q^i - C^{bound}} \sum_{1 \leq n} (z^n - 1) \left[ \lambda_Q^{i,+}(n) - \lambda_Q^{i,-}(n) \frac{1}{z^n} \right] &\leq -2\delta z^{q^i - C^{bound}}, \quad \text{when } q^i > C^{bound}, \\ z^{s - C^{bound}} \sum_{1 \leq n} (z^{n\tau_0} - 1) \left[ \lambda_S^+(n) - \lambda_S^-(n) \frac{1}{z^{n\tau_0}} \right] &\leq -2\delta z^{s - C^{bound}}, \quad \text{when } s > C^{bound}. \end{aligned} \quad (21)$$

**Step (iii).** By combining Inequalities (26), (20) and (27), we have

$$\mathcal{Q}_t V(u) \leq -cV(u) + d,$$

with  $c = \min(c^1, 2\delta)$  and  $d = d^1$ . This completes the proof.

## II.C Proof of the computation of the long-term average price move

### II.C.1 Proof of Proposition 1

For simplification, we fix the spread equal to one tick. Hence,  $t_1$  (resp.  $t_2$ ) are the first depletion of the best bid (resp. ask). Under Assumptions 3 and 4, the number of reachable states  $N$  is finite. For any  $i, i' \in \mathbb{N}_U$  with  $\mathbb{N}_U = \{1, \dots, N\}$  and  $(s, r) \in \mathbb{R}_+ \times \{1, 2\}$ , we denote by  $\mu_{ii'}^r(s)$  the density of  $t_{ii'}^r$  the first depletion time of the limit  $r$  when the initial state is  $U_i$  and the state before the depletion is  $U_{i'}$ . We write  $\mu_{ii'}^-$  (resp.  $\mu_{ii'}^+$ ) for the density of  $\tilde{t}_{ii'}^-$  (resp.  $\tilde{t}_{ii'}^+$ ) the first time when the bid (resp. ask) limit is **consumed before the ask** (resp. bid) one and where the initial state is  $U_i$  and the state before the depletion is  $U_{i'}$ .

**Lemma II.C.1.** *There exist  $m_1 > 0$  and  $m_2 > 0$  such that*

$$\mu_{ii'}^r(s) \leq m_1 e^{-m_2 s}, \quad \mu_{ii'}^\pm(s) \leq m_1 e^{-m_2 s}, \quad \forall s \in \mathbb{R}_+, \forall r \in \{1, 2\}, \forall (i, i') \in \mathbb{N}_U^2.$$

The proof of Lemma II.C.1 is given after the proof of Proposition 1. Let  $i \in \mathbb{N}_U$ ,  $s \geq 0$  and  $D_s^i = \mathbb{E}_{U_i}[P_s - P_0]$ . We denote by  $\mu_{ii'}^{-t}$  (resp.  $\mu_{ii'}^{+t}$ ) the density of  $\tilde{t}_{ii'}^{-t}$  (resp.  $\tilde{t}_{ii'}^{+t}$ ) the first time lower than  $t$  where the bid (resp. ask) limit is depleted before the ask (resp. bid) one and when the initial state is  $U_i$  and the state before the depletion is  $U_{i'}$ . Under Assumption 5, we have  $D_\infty^i = \lim_{t \rightarrow \infty} \mathbb{E}_{U_i}[P_t - P_0] < \infty$  at least for one  $i$ . Since  $U_t$  is irreducible it means that  $D_\infty^i < \infty$  for all  $i$ . We have the following result:

**Lemma II.C.2.**

$$\lim_{t \rightarrow \infty} \int_0^t \mu_{ii'}^{\pm t}(s) (D_t^k - D_{t-s}^k) ds = 0, \quad \forall (i, i', k) \in \mathbb{N}_U^3, \forall r \in \{1, 2\}.$$

The proof of Lemma II.C.2 is given at the end of this section. Let us prove Proposition 1.

*Proof of Proposition 1.* Let  $\Delta P_t^0 = P_t - P_0$  and  $q_{ii'}^{\pm t} = \mathbb{P}[\tilde{t}_{ii'}^\pm \leq t]$  for  $t \geq 0$ . We can write

$$\begin{aligned} \mathbb{E}_{U_i}[\Delta P_t^0] &= \mathbb{E}_{U_i}[\Delta P_t^0 \mathbf{1}_{t_2 \leq t_1 < t}] + \mathbb{E}_{U_i}[\Delta P_t^0 \mathbf{1}_{t_1 < t_2 < t}] \\ &= \mathbb{E}_{U_i}[\mathbb{E}[\Delta P_t^0 / \mathcal{F}_{t_2}] \mathbf{1}_{t_2 \leq t_1 < t}] + \mathbb{E}_{U_i}[\mathbb{E}[\Delta P_t^0 / \mathcal{F}_{t_1}] \mathbf{1}_{t_1 < t_2 < t}] \tag{22} \\ &= \sum_{i'} q_{ii'}^{+t} [\alpha_{i'}^+ + \int_0^t \mu_{ii'}^{+t}(s) (\sum_{k=1}^N d_{i',k}^2 D_{t-s}^k) ds] + \sum_{i'} q_{ii'}^{-t} [\alpha_i^- + \int_0^t \mu_{ii'}^{-t}(s) (\sum_{k=1}^N d_{i',k}^1 D_{t-s}^k) ds] \\ &= \underbrace{\sum_{i'} q_{ii'}^{+t} (\alpha_{i'}^+ + \sum_{k=1}^N d_{i',k}^2 \int_0^t \mu_{ii'}^{+t}(s) D_t^k ds) + \sum_{i'} q_{ii'}^{-t} (\alpha_i^- + \sum_{k=1}^N d_{i',k}^1 \int_0^t \mu_{ii'}^{-t}(s) D_t^k ds)}_{(1)} \\ &+ \underbrace{\sum_{i'} q_{ii'}^{+t} (\alpha_{i'}^+ + \sum_{k=1}^N d_{i',k}^2 \int_0^t \mu_{ii'}^{+t}(s) \Delta D_{t,s}^k ds) + \sum_{i'} q_{ii'}^{-t} (\alpha_i^- + \sum_{k=1}^N d_{i',k}^1 \int_0^t \mu_{ii'}^{-t}(s) \Delta D_{t,s}^k ds)}_{(2)}, \end{aligned}$$

## II.C. Proof of the computation of the long-term average price move

with  $\Delta D_{t,s}^k = D_{t-s}^k - D_t^k$ . Using Lemma II.C.2, the quantity (2) goes to 0 when  $t$  tends to infinity. Hence by sending  $t$  to infinity in (22), we find

$$D_\infty^i = \sum_{i'} q_{ii'}^+ \alpha_{i'}^+ + q_{ii'}^- \alpha_{i'}^- + \sum_{k=1}^N \sum_{i'} (q_{ii'}^+ d_{i',k}^2 + q_{ii'}^- d_{i',k}^1) D_\infty^k = q_i + \sum_{k=1}^N p_{i,k} D_\infty^k. \quad (23)$$

Additionally, using the symmetry relation, we have  $D_\infty^i = -D_\infty^{i^{sym}}$ . We write  $\mathcal{D}$  for the set  $\mathcal{D} = \{(q^1, q^2, s); p = 0, q^1 \geq q^2\}$ . Consequently, Equation (23) reads

$$D_\infty^i (1 - (p_{i,i} - p_{i,i^{sym}})) = q_i + \sum_{k \in \mathcal{D}} (p_{i,k} - p_{i,k^{sym}}) D_\infty^k.$$

Given that  $0 \leq p_{i,i} < 1$  (the price moves with a non-zero probability when one limit is totally consumed), we have  $(1 - (p_{i,i} - p_{i,i^{sym}})) > 0$ . This proves the result of Proposition 1.  $\square$

*Proof of Lemma II.C.1.* Let  $i, i' \in \mathbb{N}_U$  and  $r \in \{1, 2\}$ . Under Assumptions 3 and 4, the intensities are bounded by a constant  $\lambda_\infty$ . Let  $N_\infty$  be the Poisson process which admits  $\lambda_\infty$  as an intensity. We have

$$\mathbb{P}_{U_i}[\underline{t} \leq t_1 \leq \bar{t}, U_{t_1^-} = U_i'] \leq \mathbb{P}[N_t - N_{\underline{t}} \neq 0, \forall t \in [\underline{t}, \bar{t}]] = \int_{\underline{t}}^{\bar{t}} \lambda_\infty e^{-\lambda_\infty s} ds, \quad \forall \underline{t}, \bar{t} \in \mathbb{R}_+.$$

Consequently, we get

$$\mu_{ii'}^r(\underline{t}) = \lim_{\bar{t} \rightarrow \underline{t}} \frac{\mathbb{P}_{U_i}[\underline{t} \leq t_1 \leq \bar{t}, U_{t_1^-} = U_i']}{\bar{t} - \underline{t}} \leq \lim_{\bar{t} \rightarrow \underline{t}} \frac{\int_{\underline{t}}^{\bar{t}} \lambda_\infty e^{-\lambda_\infty s} ds}{\bar{t} - \underline{t}} = \lambda_\infty e^{-\lambda_\infty \underline{t}}.$$

By following the same methodology, we also have

$$\mu_{ii'}^\pm \leq \lambda_\infty e^{-\lambda_\infty \underline{t}}.$$

This completes the proof.  $\square$

*Proof of Lemma II.C.2.* Using Lemma II.C.1, we have

$$\mathbb{P}[\tilde{t}_{ii'}^{\pm t} \leq s] = \mathbb{P}[\tilde{t}_{ii'}^\pm \leq s | \tilde{t}_{ii'}^\pm \leq t] = \frac{\mathbb{P}[\tilde{t}_{ii'}^\pm \leq s]}{\mathbb{P}[\tilde{t}_{ii'}^\pm \leq t]} \leq \frac{m_1 e^{-m_2 s}}{q_{ii'}^{\pm t}}, \quad \forall s \leq t,$$

with  $q_{ii'}^{\pm t} = \mathbb{P}[\tilde{t}_{ii'}^\pm \leq t]$ . Under Assumptions 3 and 4 the intensities are bounded and the price jumps are also bounded since the state space is finite. Hence, there exists  $m_3 > 0$  such that

$$D_t^k \leq m_3 t, \quad \forall t \geq 0.$$

Since  $\lim_{t \rightarrow \infty} D_t^k < \infty$  is finite, we have  $\epsilon_t = \sup_{s \leq \frac{t}{2}} |D_t^k - D_{t-s}^k| \xrightarrow{t \rightarrow \infty} 0$ . Consequently, we have

$$\begin{aligned} \lim_{t \rightarrow \infty} \left| \int_0^t \mu_{ii'}^{\pm, t}(s) (D_t^k - D_{t-s}^k) ds \right| &\leq \lim_{t \rightarrow \infty} \epsilon_t \int_0^{\frac{t}{2}} \mu_{ii'}^{\pm, t}(s) ds + \int_{\frac{t}{2}}^t \mu_{ii'}^{\pm, t}(s) |D_t^k - D_{t-s}^k| ds \\ &\leq \lim_{t \rightarrow \infty} \epsilon_t + \frac{2m_3}{q_{ii'}^{\pm t_0}} t \int_{\frac{t}{2}}^t m_1 e^{-m_2 s} ds \\ &\leq \lim_{t \rightarrow \infty} \epsilon_t + \frac{2m_1 m_2 m_3}{q_{ii'}^{\pm t_0}} t (e^{-m_2 t} - e^{-m_2 \frac{t}{2}}) = 0, \end{aligned}$$

with  $t_0$  a fixed positive number such that  $q_{ii'}^{\pm t_0} > 0$ . This completes the proof.  $\square$

### II.C.2 Computation methodology of $q_{ii'}^\pm$

The reader can skip this part for a first reading. For simplification, we fix the added/cancelled quantity  $q = 1$  and the spread constant. To take into account non-unitary jumps, we can simply fill the zero values of the matrix  $\tilde{Q}^*$  with the right probabilities, see Equation (24).

Let  $R = [R^-, R^+]$  be the matrix such that  $R_{ii'}^- = q_{ii'}^-$  and  $R_{ii'}^+ = q_{ii'}^+$ . To compute the matrix  $R$ , we first fix the price  $P = 0$  since there is no price move before the total depletion of a limit and model the order book state by  $u = (q^1, q^2)$  with  $q^1$  (resp.  $q^2$ ) the best bid (resp. ask) quantity. Then, we introduce the absorbing states  $U_{0,q'}$  (resp.  $U_{q',0}$ ) with  $q' \geq 1$  associated to the cases  $u' = (0, q')$  (resp.  $u' = (q', 0)$ ) where  $Q^1$  (resp.  $Q^2$ ) is consumed before  $Q^2$  (resp.  $Q^1$ ). We want to compute the probabilities to visit  $U_{0,q'}$  and  $U_{q',0}$  with  $q' \geq 1$  starting from  $U_i'$ . To do this, we consider the infinitesimal generator  $Q^*$  of the Markov process  $(Q^1, Q^2)$  (the price  $P = 0$  is fixed)

$$Q^* = \begin{bmatrix} 0_{2Q^{max}} & \mathbf{0} \\ [Q^-, Q^+] \tilde{Q}^* \end{bmatrix},$$

where  $0_{2Q^{max}}$  is the zero square matrix of size  $2Q^{max}$ ,  $Q^-$  encodes transitions to the absorbing states  $U_{0,q'}$  and  $Q^+$  encodes transitions to the absorbing states  $U_{q',0}$  with  $1 \leq q' \leq Q^{max}$ , and  $\tilde{Q}^*$  is similar to the infinitesimal generator of the process  $U_t$  without regeneration. The matrix  $\tilde{Q}^*$  has the following form:

$$\tilde{Q}^* = \begin{bmatrix} \tilde{Q}_1^{*,(1)} & \tilde{Q}_0^{*,(1)} & 0 & 0 & \dots \\ \tilde{Q}_2^{*,(2)} & \tilde{Q}_1^{*,(2)} & \tilde{Q}_0^{*,(2)} & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \dots & 0 & 0 & \tilde{Q}_2^{*,(Q^{max})} & \tilde{Q}_1^{*,(Q^{max})} \end{bmatrix}, \quad (24)$$

where  $\tilde{Q}_0^{*,(l)}$  encodes transitions from level  $Q^1 = l$  to level  $Q^1 = l + 1$ , matrix  $\tilde{Q}_2^{*,(l)}$  encodes transition from level  $Q^1 = l$  to  $Q^1 = l - 1$  and matrix  $\tilde{Q}_1^{*,(l)}$  encodes transitions within level  $Q^1 = l$ .  $Q^{max}$  is the maximum quantity available on each limit. Within each sub-matrix  $\tilde{Q}_i^{*,(l)}$  with  $i \in \{0, 1, 2\}$ ,  $Q^1$  is equal to  $l$  and  $Q^2$  vary from 1 to  $Q^{max}$ . The sub-matrices  $\tilde{Q}_i^{*,(l)}$ , for  $i = 0, 1$ , can be written

$$\tilde{Q}_0^{*,(l)} = \begin{pmatrix} \lambda^{1,+}(l, 1) & & & & \\ & \ddots & & & \\ & & \lambda^{1,+}(l, Q^{max}) & & \end{pmatrix} \text{ and } \tilde{Q}_2^{*,(l)} = \begin{pmatrix} \lambda^{1,-}(l, 1) & & & & \\ & \ddots & & & \\ & & & & \\ & & & \lambda^{1,-}(l, Q^{max}) & \end{pmatrix}.$$

Let  $\lambda^*(l, l') = \sum_{i=1}^2 \lambda^{i,+}(l, l') + \lambda^{i,-}(l, l')$  for every  $l, l' \in \{1, \dots, Q^{max}\}$ . For  $l \leq Q^{max}$ , we have

$$\tilde{Q}_1^{*,(l)} = \begin{pmatrix} -\lambda^*(l, 1) & \lambda^{2,+}(l, 1) & 0 & 0 & \dots \\ \lambda^{2,-}(l, 2) & -\lambda^*(l, 2) & \lambda^{2,+}(l, 2) & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \dots & 0 & 0 & \lambda^{2,-}(l, Q^{max}) & -\lambda^*(l, Q^{max}) \end{pmatrix}.$$

Finally, we define the matrix  $Q^-$  such that  $Q_{ij}^- = \lambda^{1,-}(1, i)$  for  $1 \leq i \leq Q^{max}$  and 0 otherwise, and the matrix  $Q^+$  such that  $Q_{iQ^{max}+1, i+1}^+ = \lambda^{2,-}(i, 1)$  for  $0 \leq i \leq Q^{max} - 1$  and 0 otherwise.

**Lemma II.C.3** (Computation of  $q_{i'i'}^\pm$ ). *Let  $R = [R^-, R^+]$  be the matrix such that  $R_{i'i'}^- = q_{i'i'}^-$  and  $R_{i'i'}^+ = q_{i'i'}^+$ . Then,  $R$  is a solution of the equation*

$$\tilde{Q}^* R = -z^1,$$

where  $\tilde{Q}^*$  is defined in (24) and  $z^1 = [Q^-, Q^+]$  with  $Q^-$  and  $Q^+$  defined (25). The solution of this equation is unique since  $\tilde{Q}^*$  is invertible.

The proof of this lemma is given below.

**Example 1.** *The case where the best bid dynamics are independent from the one of the best ask is a classical setting that was studied in [90]. In such situation, the matrix  $\tilde{Q}^*$  is diagonalisable, see Appendix II.C.II.C.3. Additionally, when the intensities are constant, we have a closed closed-form formula for the diagonalisation of  $\tilde{Q}^*$ , see Appendix II.C.II.C.3.*

*Proof of Lemma II.C.3.* Using Theorem 3.3.1 in [131], for every absorbing state  $U_{i'}$  we have

$$\begin{cases} q_{U_{0,q}, U_{0,q}}^- = 1, & q_{U_{0,q}, U_{i'}}^+ = 0, & q_{U_{q,0}, U_{i'}}^- = 0, & q_{U_{q,0}, U_{q,0}}^+ = 1, & \text{with } q \in \{1, Q^{max}\}, \\ \sum_j Q_{i,j}^* q_{j,i'}^\pm = 0 & \forall i \in [2Q^{max} + 1, (Q^{max})^2 + 2Q^{max}]. \end{cases}$$

In the above equations, we use a slight abuse of notation and do not differentiate the state  $U_{i'}$  and the index  $i'$ . The equation above reads

$$\tilde{Q}^* R = -z^1, \tag{25}$$

with  $z^1 = [Q^-, Q^+]$  and  $R = [R^-, R^+]$  the matrix such that  $R_{i'i'}^- = q_{i'i'}^-$  and  $R_{i'i'}^+ = q_{i'i'}^+$ . The solution of this equation is unique since  $\tilde{Q}^*$  is invertible. When queues are independent  $\tilde{Q}^*$  is diagonalisable, see next sub-section. In the simple case of constant intensities,  $\tilde{Q}^*$  diagonalisation is explicetely computable.  $\square$

### II.C.3 Diagonalisation of $\tilde{Q}^*$

#### II.C.3.1 Symmetrization of $\tilde{Q}^*$ under the assumption of independent queues

The idea is to find a matrix  $P$  such that  $P^{-1}\tilde{Q}^*P$  is symmetric with  $P = LH$ . First, we consider the bloc-diagonal matrix  $H = \text{diag}\{H_1, H_2, \dots, H_{Q^{max}}\}$  where every  $H_i$  is a square matrix of size  $Q^{max}$  such that

$$\begin{cases} H_1 = I, \\ H_{i+1} = H_i \sqrt{\tilde{Q}_2^{*,(i)} \left( \tilde{Q}_0^{*,(i-1)} \right)^{-1}}, & \forall i \geq 1. \end{cases}$$

Here  $\sqrt{\cdot}$  refers to the square root of a matrix. The existence of such a matrix in this case is trivial since  $\tilde{Q}_2^{*,(i)}$  and  $\tilde{Q}_0^{*,(i-1)}$  are diagonal with strictly positive coefficients.

Next, we consider the bloc-diagonal matrix  $L = \text{diag}\{L_1, L_1, \dots, L_1\}$  where  $L_1$  is a diagonal matrix with diagonal coefficients  $L_1(1, 1) = 1$  and  $L_1(i+1, i+1) = L_1(i, i) \sqrt{\frac{\tilde{Q}_1^{*,(0)}(i+1, i)}{\tilde{Q}_1^{*,(0)}(i, i+1)}}$  for all  $i \geq 1$ . Given that queues are independent we have  $\tilde{Q}_1^{*,(0)} = \tilde{Q}_i^{*,(0)}$  for all  $i \geq 1$ . Finally, we remark that  $P^{-1}\tilde{Q}^*P$ , with  $P = LH$ , is symmetric.

### II.C.3.2 Diagonalisation of the symmetric matrix $P^{-1}\tilde{Q}^*P$ : constant coefficients

In the simple case of constant coefficients, the matrix  $P$  defined in Appendix II.C.II.C.3.II.C.3.1 satisfies

$$P^{-1}\tilde{Q}^*P = \begin{bmatrix} A(a,b) & V & 0 & 0 & 0 \\ V & A(a,b) & V & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & V & A(a,b) \end{bmatrix} \quad \text{and} \quad A(a,b) = \begin{pmatrix} a & b & 0 & 0 & \dots \\ b & a & b & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & 0 & 0 & b & a \end{pmatrix},$$

where  $V = \beta I$  with  $\beta > 0$  and  $a$  and  $b$  are some fixed constants. In such framework, the eigenvalues of  $P^{-1}\tilde{Q}^*P$  are

$$\lambda_{a,b,\beta}^{k,j} = a + 2b \cos\left(\frac{k\pi}{n+1}\right) + 2\beta \cos\left(\frac{j\pi}{n+1}\right), \quad \forall 1 \leq k, j \leq n,$$

and the associated eigenspace is generated by the eigenvector  $X^{k,j} = (v_1^j X^k, v_2^j X^k \dots, v_{Q_{max}}^j X^k)$ , where  $v_r^j$  satisfies

$$v_r^j = \sin\left(r \frac{j\pi}{n+1}\right), \quad \forall 1 \leq r, j \leq n,$$

and  $X^k$  is a vector such that

$$X_l^k = \sin\left(l \frac{k\pi}{n+1}\right), \quad \forall 1 \leq k, l \leq n.$$

## II.D Generalities about the state process $U_t^\mu$ and the value function

For simplification, we fix the added/cancelled quantity  $q = 1$  and the **spread constant equal to 1 tick**. In this section we take  $\bar{U} = \mathbb{R}_+^5 \times (\mathbb{R})^2$  and Assumptions II.D.1 and II.D.2 are in force. When  $\bar{U} = \mathbb{N}^5 \times (\frac{\mathbb{I}_0}{2}\mathbb{Z})^2$ , the results of this section remain valid but for different values of the constants and Assumptions II.D.1 and II.D.2 are not needed.

### II.D.1 Regularity of the regenerative process $U_t^\mu$

The regularity of our regenerative process is not trivial. In fact, if we consider two processes  $U^\mu$  and  $U'^\mu$  satisfying the same order book dynamics (see Section 3.1) but starting from two different initial points  $u_0$  and  $u'_0$ , as long as there is no regeneration, for every order flow trajectory, the error  $\|U^\mu - U'^\mu\|$  is equal to the initial error  $\|u_0 - u'_0\|$ . However, when one of the two processes is regenerated before the other, the regenerated one starts a new cycle from a random position and the error  $\|U^\mu - U'^\mu\|$  is no longer bounded by  $\|u_0 - u'_0\|$ . Hence, the irregularity mainly comes from the regeneration. In our case, since the regeneration law depends on the killing state, it may introduce strong irregularities. Consequently, we need an assumption to ensure that regeneration distributions are similar when exit points are close

enough. We denote by  $\tilde{d}_u^i$  the regeneration distribution of  $U_t^\mu$  when the depleted side is  $i$  and the state before the depletion is  $u$ . We give here a result on the regularity of the state process under two general assumptions.

**Assumption II.D.1** (Regeneration smoothness). *There exist positive constants  $K, q_0, q_1 \leq 1 \wedge q_0$  and  $\beta$  such that for any  $u, u' = (q^{bef}, q^a, q^{aft}, q^2, i', p', p^{exec}) \in \bar{\mathbb{U}}$ ,*

$$\begin{aligned} \|\tilde{d}_u^i - \tilde{d}_{u'}^i\|_{TV} &\leq K(\|u - u'\|_p)^p, \\ \tilde{d}_u^i(u') &= 0, \quad \text{when } \min(q^1, q^2) \leq q_0, \\ \sum_i (\lambda^{i,+}(u) + \lambda^{i,-}(u, u') \mathbf{1}_{q^i > q_1}) &\leq \beta, \end{aligned}$$

with  $q^1 = q^{bef} + q^a + q^{aft}$ ,  $\|p - p'\|_{TV} = \sup_{\mathcal{A} \in \mathcal{F}} |p(\mathcal{A}) - p'(\mathcal{A})|$  and  $\|\cdot\|_p$  the  $L_p$  norm with  $p \geq 1$ .

Assumption II.D.1 is a Lipschitz type inequality to ensure that regeneration distributions are almost similar when exit states are close enough. Furthermore, we add a support constraint to guarantee that regenerated limits have a size higher than a fixed minimum quantity  $q_0$  and consider a boundedness assumption for the intensities. We also add the following assumption.

**Assumption II.D.2** (Exit dynamics).  $\forall u = (q^1, q^2) \in \mathbb{U} = \mathbb{R}_+^2$  and  $i \in \{1, 2\}$ , *there exists a positive constants  $\beta^-$  such that  $\forall \epsilon > 0, \exists q_\epsilon > 0$ ,*

$$\sum_{u'} \lambda^{i,-}(u, u') \mathbf{1}_{q^i \leq q_\epsilon} \geq \frac{\beta^-}{\epsilon}.$$

For small size queues (i.e.  $q^i \leq q_\epsilon$ ), Assumption II.D.2 ensures that intensities of depletion are high (i.e.  $\lambda^{i,-} \geq \frac{1}{\epsilon}$ ) while other intensities are bounded. Such assumption avoids critical situations where the order book goes far away from an exit state after being too close to it. It is also consistent with empirical evidences since a limit disappears almost instantaneously when it becomes lower than a given bound  $q_\epsilon$ . We have the following result proved in Appendix II.DII.D.4.

**Theorem II.D.1** (Regularity of the state process). *Under Assumptions 2, 3, 4, 5, II.D.1 and II.D.2, the process  $U_t^\mu$  satisfies*

$$\mathbb{E}[(\|U_t^\mu - U_t'^\mu\|_p)^p] \leq K_0 e^{C_0 T} (\|U_0 - U_0'\|_p)^p, \quad \forall U_0, U_0' \in \bar{\mathbb{U}}, \forall t \in [0, T], \quad (26)$$

where  $U_t^\mu$  (resp.  $U_t'^\mu$ ) is the Markov process starting from the initial state  $U_0$  (resp.  $U_0'$ ),  $T$  is the final time,  $K_0$  and  $C_0$  are constants defined in Appendix II.DII.D.4.



### II.D.2 Regularity of the value function

In this section, we fix  $p = 1$  and write  $\|\cdot\|_p = \|\cdot\|$ . Since  $f$  is Lipschitz, the function  $g(u) = f(\mathbb{E}_u[\Delta P_\infty^\mu])$  is Lipschitz as well, see Lemma II.D.3. When the state process is valued in  $\mathbb{N}^5 \times (\frac{T_0}{2}\mathbb{N}^*)^2$  we only need  $g$  to be bounded which is always satisfied under Assumptions 3, 4 and 5. Then, we have the following regularity properties proved in Appendix II.DII.D.5.

**Proposition II.D.1.** *Let  $T$  be the final horizon, the value function  $V$  is*

- *Lipschitz in space:*

$$|V_T(t, U_1) - V_T(t, U_2)| \leq Ae^{C_0(T-t)}\|U_1 - U_2\|, \quad \forall U_1, U_2 \in \bar{U}, \forall t \in [0, T], \quad (27)$$

*with  $A$  and  $C_0$  constants respectively defined in Appendix II.DII.D.5 and Theorem II.D.1.*

- *Lipschitz in time:*

$$|V_T(t, U_1) - V_T(t', U_1)| \leq L_0|t' - t|, \quad \forall U_1 \in \bar{U}, \forall t, t' \in [0, T], \quad (28)$$

*with  $L_0 = cq^a + Ae^{C_0(T-t)}C$  and  $C$  a constant defined in Appendix II.DII.D.4.*

### II.D.3 Execution time inequalities

Here again we also fix  $p = 1$  and write  $\|\cdot\|_p = \|\cdot\|$ . We recall that  $T_{Exec}^{t, \mu}$  is defined in Section 4.2 for any control  $\mu$ . We provide here two execution time inequalities. First, when agent's decisions are taken at a fixed frequency  $\Delta^{-1}$ , we have the following inequality.

**Proposition II.D.2.** *Let  $U_1, U_2$  be two initial states and  $\mu_1^{Opti}$  (resp.  $\mu_2^{Opti}$ ) the optimal strategy for the process starting from  $U_1$  (resp.  $U_2$ ). Then, we have*

$$\mathbb{E} \left[ \left| T_{Exec}^{t, \mu_2^{Opti}} - T_{Exec}^{t, \mu_1^{Opti}} \right| \right] \leq \Delta e^{C_1(T-t)}\|U_1 - U_2\| + \Delta K_1(T-t), \quad \forall U_1, U_2 \in \bar{U}, \forall t \in [0, T],$$

*with  $C_1 = \frac{\log(K_0)}{\Delta} + C_0$ ,  $K_1 = 4\beta$  and  $C_0$  and  $K_0$  defined in Appendix II.DII.D.4.*

Proposition II.D.2 shows that both initial states and agent's latency  $\Delta$  affect the optimal execution time. We have a second inequality when decisions are taken at any time.

**Proposition II.D.3.** *Let  $U_1, U_2$  be two initial states and  $\mu_1^{Opti}$  (resp.  $\mu_2^{Opti}$ ) the optimal strategy for the process starting from  $U_1$  (resp.  $U_2$ ). Then, we have*

$$\mathbb{E} \left[ \left| T_{Exec}^{t, \mu_2^{Opti}} - T_{Exec}^{t, \mu_1^{Opti}} \right| \right] \leq K'_0\|U_1 - U_2\|, \quad \forall U_1, U_2 \in \bar{U}, \forall t \in [0, T], \quad (29)$$

*with  $K'_0 = \log(K_0)$ . The constant  $K_0$  is given in Theorem II.D.1.*

Proofs of Propositions II.D.2 and II.D.3 are given in Appendix II.D.6.

#### II.D.4 Proof of Theorem II.D.1

**Notations:** Let  $U_0, U'_0 \in \bar{U}$ ,  $U_t^\mu$  (resp.  $U_t^{\prime\mu}$ ) the process starting from  $U_0$  (resp.  $U'_0$ ),  $T$  the final time horizon,  $\bar{P}$  the maximum mid price value,  $\bar{P}^{Exec}$  the upper bound for the execution price and  $\hat{Q}^{\max} = \max(Q^{\max}, \bar{Q}^{\max}, \bar{P})$ . Let  $\tau_{1,-}$  (resp.  $\tau'_{1,-}$ ) be the first time where the best bid (resp. ask) is totally consumed,  $\tau_{1,+}$  (resp.  $\tau'_{1,+}$ ) the first time where the best ask (resp. bid) is depleted and  $\tau_1 = \tau_{1,-} \wedge \tau_{1,+}$  (resp.  $\tau'_1 = \tau'_{1,-} \wedge \tau'_{1,+}$ ) the first regeneration time of the process  $U_t^\mu$  (resp.  $U_t^{\prime\mu}$ ). Finally, we write  $\tau_2$  for  $\tau_2 = \tau_1 \vee \tau'_1$ . Let  $\epsilon = (\|U_0 - U'_0\|_p)^p > 0$ , there exists  $0 < q_\epsilon \leq q_0$  satisfying conditions of Assumption II.D.2. We fix  $p \geq 1$  and write  $\|\cdot\|_p = \|\cdot\|$ .

**Step 1:** Assume  $\|U_0 - U'_0\| > q_\epsilon$ . Let us show that

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|U_t^\mu - U_t^{\prime\mu}\|^p\right] \leq 3^{p-1} \|U_0 - U'_0\|^p e^{\tilde{C}T}, \quad (30)$$

with  $\tilde{C} = \frac{\hat{C}\beta}{q_\epsilon^p}$ ,  $\hat{C} = [6(\hat{Q}^{\max})^p + (\bar{P}^{Exec})^p]$  and  $\beta$  is defined in Assumption II.D.2.

*Proof of Equation (30).* We write  $\delta U_t^\mu = U_t^\mu - U_0$  and  $\delta U_t^{\prime\mu} = U_t^{\prime\mu} - U'_0$ . Then, we have

$$\|U_t^\mu - U_t^{\prime\mu}\|^p = 3^{p-1} \left( \|\delta U_t^\mu\|^p + \|\delta U_t^{\prime\mu}\|^p + \|U_0 - U'_0\|^p \right).$$

Since  $\mathbb{E}\left[\sup_{0 \leq t \leq T} \|\delta U_t^{\prime\mu}\|^p\right] \leq [6(\hat{Q}^{\max})^p + (\bar{P}^{Exec})^p] \beta T$ , we have

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|U_t^\mu - U_t^{\prime\mu}\|^p\right] \leq CT + B,$$

with  $C = 3^{p-1} \hat{C}\beta$  and  $B = 3^{p-1} \|U_0 - U'_0\|^p$ . Since  $\|U_0 - U'_0\| > q_\epsilon$ , we have  $B > 3^{p-1} q_\epsilon^p$ . The latter leads to the following inequality:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \|U_t^\mu - U_t^{\prime\mu}\|^p\right] \leq CT + B \leq B e^{\tilde{C}T},$$

with  $\tilde{C} = \frac{\hat{C}\beta}{q_\epsilon^p}$ . The last inequality comes from  $ax + b \leq be^{a/b_0 x}$  when  $b > b_0$ .  $\square$

**Step 2:** Assume  $\|U_0 - U'_0\| \leq q_\epsilon$ . First, let us show that

$$\mathbb{E}\|U_{\tau_2}^\mu - U_{\tau_2}^{\prime\mu}\|^p \leq \kappa \|U_0 - U'_0\|^p, \quad (31)$$

with  $\kappa = \kappa^1 \hat{C}\beta + 2\sqrt{2}\kappa^2 RK$ ,  $\kappa^1 = \frac{(1+3T\beta)}{\beta^-}$ ,  $\kappa^2 = \kappa^1 \hat{C}\beta + 1$ ,  $K$  defined in Assumption II.D.1 and  $R$  defined in Appendix II.D.4.1.

*Proof of Inequality (31).* We assume that  $\tau_1 > \tau'_1$  a.s. The general case uses the same lines of argument. We have

$$\|U_{\tau_2}^\mu - U_{\tau_2}^{\prime\mu}\|^p = \|U_{\tau_1}^\mu - U_{\tau_1}^{\prime\mu}\|^p \leq \overbrace{\|U_{\tau_1}^\mu - U_{\tau_1}^{\prime\mu}\|^p}^{(1)} + \overbrace{\|U_{\tau_1}^\mu - U_{\tau_1}^{\prime\mu}\|^p}^{(2)}.$$

## II. Optimal liquidity-based trading tactics

- Part (1) satisfies

$$\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau'_1}^\mu\|^p] \leq \hat{C}\beta\mathbb{E}[|\tau_1 - \tau'_1|] \leq \hat{C}\beta\kappa^1\|U_0 - U'_0\|^p. \quad (32)$$

In the second inequality, we use  $\mathbb{E}[|\tau_1 - \tau'_1|] \leq \kappa^1\|U_0 - U'_0\|^p$ , see Lemma II.D.1 below.

- For Part (2), let  $X = [0, \hat{Q}^{max}]^6 \times [0, \bar{P}^{Exec}]$ . There exists a Borel function  $s$  valued in  $[-1, 1]$ , see Appendix II.D.4.1 at the end of the proof, and a positive constant  $R$  such that

$$\|x - y\|^p \leq R|s(x) - s(y)|, \quad \forall (x, y) \in X^2.$$

Let us denote by  $d_x$  the regeneration distribution of  $U_t^\mu$  (i.e.  $d_x = d_x^1$  (resp.  $d_x = d_x^2$ ) when the best bid (resp. ask) is totally depleted). Finally, we denote by  $\mathcal{M}$  the set of Borel functions on  $X$  that take values in  $[-1, 1]$ . In such case, we have

$$\begin{aligned} \mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau'_1}^\mu\|^p] &\leq R\mathbb{E}[|s(U_{\tau_1}^\mu) - s(U_{\tau'_1}^\mu)|] = R\mathbb{E}[\mathbb{E}[|s(x) - s(y)| d_{U_{\tau_1}^\mu}(x) d_{U_{\tau'_1}^\mu}(y) dx dy / \mathcal{F}_{\tau_1}]] \\ &\leq \sqrt{2}R\mathbb{E}[\mathbb{E}[|\int |s(x)| d_{U_{\tau_1}^\mu}(dx) - \int |s(y)| d_{U_{\tau'_1}^\mu}(dy)| / \mathcal{F}_{\tau_1}]] \\ &\leq \sqrt{2}R\mathbb{E}[\mathbb{E}[\sup_{g \in \mathcal{M}} \left\{ \int g d_{U_{\tau_1}^\mu}(dx) - \int g d_{U_{\tau'_1}^\mu}(dx) \right\} / \mathcal{F}_{\tau_1}]] \\ &= 2\sqrt{2}R\mathbb{E}[\underbrace{\|d_{U_{\tau_1}^\mu} - d_{U_{\tau'_1}^\mu}\|_{TV}}_{(*)}] \leq 2\sqrt{2}RK\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau'_1}^\mu\|^p]. \end{aligned}$$

In (\*) we used the total variation distance property  $2\|\mu - \nu\|_{TV} = \sup_{g \in \mathcal{M}} \int g d\mu - \int g d\nu$ .

Given that  $\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau'_1}^\mu\|^p] \leq \kappa^2\|U_0 - U'_0\|^p$ , see Lemma II.D.2 below, we have

$$\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau'_1}^\mu\|^p] \leq 2\sqrt{2}RK\kappa^2\|U_0 - U'_0\|^p. \quad (33)$$

By combining Inequalities (32) and (33), we get the result.  $\square$

**Lemma II.D.1.** *We have*

$$\mathbb{E}[|\tau_1 - \tau'_1|] \leq \kappa^1\|U_0 - U'_0\|^p,$$

with  $\kappa^1 = \frac{(1+3T\beta)}{\beta^-}$ .

*Proof of Lemma II.D.1.* First, we assume that  $\tau_1 > \tau'_1$  a.s and consider the following notation. For every state  $u$ , we write  $u^{2,+}$  (resp.  $u^{2,-}$ ) for the new state of the order book when a quantity  $q = 1$  is added to (resp. cancelled from)  $Q^2$ . The same reasoning holds for  $Q^1$ . We have

$$\mathbb{E}[|\tau_1 - \tau'_1|] = \mathbb{E}[\mathbb{E}[|\tau_1 - \tau'_1| / \mathcal{F}_{\tau'_1+}]] = \mathbb{E}[\mathbb{E}[\hat{\tau}_1 / \mathcal{F}_{\tau'_1+}]] = \mathbb{E}[\mathbb{E}[\hat{\tau}_1 / U_{\tau'_1}^\mu]] = \mathbb{E}[h(U_{\tau'_1}^\mu)],$$

where  $\hat{\tau}_1 = \tau_1 - \tau'_1$  is also the first regeneration time of  $U_t^\mu$  but starting from the initial point  $U_{\tau'_1}^\mu$  and  $h(x) = \mathbb{E}[\tau / U_0^\mu = x]$  where  $\tau$  is the first regeneration time of  $U_t^\mu$  when  $U_0^\mu = x$ . Since  $U_{\tau'_1}^\mu$  is regenerated at time  $\tau'_1$ , then  $Q_{\tau'_1}^{1,\mu} \leq q$  or  $Q_{\tau'_1}^{2,\mu} \leq q$ . Let us consider the case where

$Q_{\tau'_1}^{1,\mu} \leq q$ . The case  $Q_{\tau'_1}^{2,\mu} \leq q$  is solved using the same arguments. Since  $\|U_{\tau'_1}^\mu - U_{\tau'_1}'^\mu\| = \|U_0^\mu - U_0'^\mu\|$  (i.e. the error is unchanged before the first regeneration), we have  $Q_{\tau'_1}^{1,\mu} \leq \|U_0^\mu - U_0'^\mu\| \leq q\epsilon$ . We note  $u_1 = U_{\tau'_1}^\mu$ . By considering the possible transitions of the process  $U^\mu$ , we have

$$1 + \lambda^{1,+}(u_1)h(u_1^{1,+}) + \lambda^{2,-}(u_1)h(u_1^{2,-}) + \lambda^{2,+}(u_1)h(u_1^{2,+}) - \lambda^*(u_1)h(u_1) = 0,$$

with  $\lambda^* = \sum_i \lambda^{i,+} + \lambda^{i,-}$ . Using Assumption II.D.2 and  $h(u) \leq T$ , for every initial state  $u$ , we have

$$\begin{aligned} h(u_1) &\leq \frac{1 + \lambda^{1,+}(u_1)h(u_1^{1,+}) + \lambda^{2,-}(u_1)h(u_1^{2,-}) + \lambda^{2,+}(u_1)h(u_1^{2,+})}{\lambda^*(u_1)} \leq \frac{(1+3T\beta)\epsilon}{\beta^-} \\ &= \kappa^1 \|U_0 - U_0'\|^p, \end{aligned}$$

with  $\kappa^1 = \frac{(1+3T\beta)}{\beta^-}$ . The case when the condition  $\tau_1 > \tau'_1$  a.s is not satisfied uses the same arguments. This proves the result.  $\square$

**Lemma II.D.2.** *We have*

$$\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p] \leq \kappa^2 \|U_0 - U_0'\|^p,$$

with  $\kappa^2 = \kappa^1 \hat{C}\beta + 1$ .

*Proof of Lemma II.D.2.* By following the same methodology as for Lemma II.D.1, we first note that

$$\begin{aligned} \mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p] &\leq \mathbb{E}[\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p / \mathcal{F}_{\tau_1}]] + \mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p] \\ &= \mathbb{E}[\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p / U_{\tau_1}^\mu]] + \|U_0 - U_0'\|^p \\ &\leq \hat{C}\beta \mathbb{E}[\mathbb{E}[\hat{\tau}_1 / U_{\tau_1}^\mu]] + \|U_0 - U_0'\|^p \leq \hat{C}\beta \mathbb{E}[h(U_{\tau_1}^\mu)] + \|U_0 - U_0'\|^p, \end{aligned}$$

where  $\hat{\tau}_1 = \tau_1 - \tau'_1$  is the first regeneration time of  $U_t^\mu$  but starting from the initial point  $U_{\tau'_1}^\mu$  and  $h(x) = \mathbb{E}[\tau / U_0^\mu = x]$  where  $\tau$  is the first regeneration time of  $U_t^\mu$  when  $U_0^\mu = x$ . In the proof of Lemma II.D.1, we show that  $\mathbb{E}[\mathbb{E}[h(U_{\tau_1}^\mu)]] \leq \kappa^1 \|U_0 - U_0'\|^p$ . Hence, we have  $\mathbb{E}[\|U_{\tau_1}^\mu - U_{\tau_1}'^\mu\|^p] \leq \kappa^2 \|U_0 - U_0'\|^p$ , with  $\kappa^2 = \kappa^1 \hat{C}\beta + 1$ . This proves the result.  $\square$

**Step 3:** Assume  $\|U_0 - U_0'\| \leq q\epsilon$ . Let us show that

$$\mathbb{E}\|U_t^\mu - U_t'^\mu\|^p \leq \tilde{\kappa} e^{C'T} \|U_0 - U_0'\|^p, \quad (34)$$

for any  $t < \infty$ , where  $C' = (\beta(\kappa - 1))_+$  and  $\tilde{\kappa} = 1 \vee \kappa$ .

*Proof of Inequality (34).* Let  $t < \infty$ , we denote by  $N_t^{Regen}$  (resp.  $N_t'^{Regen}$ ) the random variable representing the number of regenerations of  $U_t^\mu$  (resp.  $U_t'^\mu$ ) before  $t$ . Let  $t_1, \dots, t_{N^{Regen}}$  be the regeneration times of  $U_t^\mu$  and  $t'_1, \dots, t'_{N'^{Regen}}$  the regeneration times of  $U_t'^\mu$ . We build the sequence  $\tilde{t}_n$  such that  $\tilde{t}_1 = t_1 \vee t'_1$  and

$$\begin{cases} \tilde{t}_n^1 &= \inf\{t_i > \tilde{t}_n; i \in \{1, \dots, N_t^{Regen}\}\} \\ \tilde{t}_n^2 &= \inf\{t'_i > \tilde{t}_n; i \in \{1, \dots, N_t'^{Regen}\}\} \\ \tilde{t}_{n+1} &= \tilde{t}_n^1 \vee \tilde{t}_n^2. \end{cases}$$

## II. Optimal liquidity-based trading tactics

Let  $\tilde{N}_t^{Regen}$  be the first index  $n$  when one of the sets  $\Omega_1^n = \{t_i > \tilde{t}_n; i \in \{1, \dots, N_t^{Regen}\}\}$  or  $\Omega_2^n = \{t'_i > \tilde{t}_n; i \in \{1, \dots, N_t^{Regen}\}\}$  is empty. We adopt the convention  $\tilde{t}_n = t, \forall n > \tilde{N}_t^{Regen}$ . Thus, we have

$$\mathbb{E} \|U_t^\mu - U_t^{\prime\mu}\|^p = \mathbb{E} [\|U_{\tilde{N}_t^{Regen}}^\mu - U_{\tilde{N}_t^{Regen}}^{\prime\mu}\|^p] \leq \overbrace{\mathbb{E} [\kappa^{\tilde{N}_t^{Regen}} \|U_0^\mu - U_0^{\prime\mu}\|^p]}^{(**)} \leq \|U_0^\mu - U_0^{\prime\mu}\|^p \mathbb{E} [\kappa^{\tilde{N}_t^{Regen}}],$$

where  $(**)$  is obtained by using Inequality (31) and the conditional expectation argument. When  $\kappa > 1$ , we denote by  $N^*$  the Poisson process with constant intensity  $\beta$ . Thus, we have

$$\mathbb{E} [\kappa^{\tilde{N}_t^{Regen}}] \leq \mathbb{E} [e^{(N_t^* + 1) \log(\kappa)}] = \kappa e^{-\beta t(1-\kappa)} < \infty.$$

When  $\kappa \leq 1$ , we have  $\kappa^{\tilde{N}_t^{Regen}} \leq 1$  and consequently,  $\mathbb{E} [\kappa^{\tilde{N}_t^{Regen}}] \leq 1$ . This proves Inequality (34). By combining Inequalities (30) and (34), we complete the proof of Theorem II.D.1 with  $K_0 = \tilde{\kappa} \vee 3$  and  $C_0 = \tilde{C} \vee C'$ .  $\square$

### II.D.4.1 Existence of $s$

Let us first consider the function  $s^{-1} = -a \log(a)$  defined in  $[0, 1]$ . The function  $s^{-1}$  is continuous in  $[0, 1]$  and bijection in the subinterval  $[0, e^{-1}]$ . Moreover, it is a Hölder function satisfying  $\|s^{-1}(a) - s^{-1}(b)\| \leq R^1 \|a - b\|^{\frac{1}{p}}, \forall (a, b) \in [0, 1]^2$ . We denote by  $s^1$  the inverse of  $s^{-1}$  defined in  $[0, e^{-1}]$ . The function  $s^1$  is continuous and valued in  $[0, 1]$ . Additionally,  $s^1$  satisfies  $\|a - b\|^p \leq R^1 |s^1(a) - s^1(b)|, \forall (a, b) \in [0, e^{-1}]^2$ .

Then, we define the normalization function  $s^2$ : 
$$\begin{aligned} X &\rightarrow [0, e^{-1}]^7 \\ (x^1, \dots, x^7) &\rightarrow (s_1^2(x^1), \dots, s_1^2(x^5), s_2^2(x^6), s_3^2(x^7)) \end{aligned}$$

where  $s_1^2, s_2^2$  and  $s_3^2$  are three auxiliary functions for normalization defined such as

$$s_1^2(x) = \frac{x}{e^{-1} \hat{Q}^{\max}}, \quad s_2^2(x) = \frac{x}{e^{-1} \hat{Q}^{\max}} \quad \text{and} \quad s_3^2(x) = \frac{x}{e^{-1} \bar{p}^{Exec}}.$$

It is simple to see that  $s^2$  satisfies  $\|x - y\|^p \leq R^2 \|s^2(x) - s^2(y)\|^p, \forall (x, y) \in X^2$ , with  $R^2 = \min(\hat{Q}^{\max}, (q^a \hat{Q}^{\max} (\frac{3}{2} + a)))$ . Finally, we define  $s(x) = \sum_{i=1}^7 |s^1(s^2(x^i))|$ . Using the inequality  $\sum_{i=1}^N |x^i - y^i| \leq \sqrt{2} |\sum_{i=1}^N |x^i| - \sum_{i=1}^N |y^i||$ , we have  $\|x - y\|^p \leq R^2 R^1 \sqrt{2} |s(x) - s(y)|, \forall (x, y) \in X^2$ .

### II.D.5 Regularity of the value function

**Lemma II.D.3.** *The function  $g$  is Lipschitz.*

*Proof of Lemma II.D.3.* Since  $f$  is Lipschitz, it suffices to prove that  $r(i) = P I_\infty^i = \lim_{t \rightarrow \infty} \mathbb{E}_{U_i} [P_t - P_0]$  is Lipschitz where  $P_t$  is the mid price. Let  $U_i$  be an order book state. We denote by  $(P_t^i)_{t \geq 0}$  and  $(U_t^i)_{t \geq 0}$  respectively the mid price and the order book processes starting from  $U_i$ . Using Theorem II.D.1, we have

$$\begin{aligned} |P I_t^i - P I_t^{i'}| &\leq \mathbb{E} [|P_t^i - P_t^{i'}|] + |P_0^i - P_0^{i'}| \leq \mathbb{E} [\|U_t^i - U_t^{i'}\|] + \|U_i - U_{i'}\| \\ &\leq (K_0 e^{C_0 t} + 1) \|U_i - U_{i'}\|, \end{aligned} \quad (35)$$

with  $PI_t^i = \mathbb{E}_{U_i}[P_t - P_0]$ . Since  $PI_t^i \xrightarrow{t \rightarrow \infty} PI_\infty^i$ , there exists  $t_{\max}$  such that for any  $t \geq t_{\max}$  and state  $i$ , we have  $|PI_t^i - PI_\infty^i| \leq \|U_i - U_{i'}\|$ . Using (35), we have

$$|PI_\infty^i - PI_\infty^{i'}| \leq |PI_\infty^i - PI_{t_{\max}}^i| + |PI_{t_{\max}}^i - PI_{t_{\max}}^{i'}| + |PI_\infty^{i'} - PI_{t_{\max}}^{i'}| \leq (2 + K_0 e^{C_0 t_{\max}}) \|U_i - U_{i'}\|.$$

This completes the proof.  $\square$

*Proof of Inequality (27).* We write  $U_t^{1,\mu}$  (resp.  $U_t^{2,\mu}$ ) for the process such that  $U_t^{1,\mu} = U_1$  (resp.  $U_t^{2,\mu} = U_2$ ). Using the fact that  $g$  is Lipschitz and Inequality (26), we have

$$\begin{aligned} |V_T(t, U_1) - V_T(t, U_2)| &\leq \sup_{\mu} \mathbb{E} \left[ |g(U_{T_{Exec_1}}^{1,\mu}) - g(U_{T_{Exec_2}}^{2,\mu})| + cq^a |T_{Exec_1}^{t,\mu} - T_{Exec_2}^{t,\mu}| \right] \\ &\leq \sup_{\mu} \mathbb{E} \left[ |g(U_{T_{Exec_1}}^{1,\mu}) - g(U_{T_{Exec_1}}^{2,\mu})| + |g(U_{T_{Exec_1}}^{2,\mu}) - g(U_{T_{Exec_2}}^{2,\mu})| + cq^a |T_{Exec_1}^{t,\mu} - T_{Exec_2}^{t,\mu}| \right] \\ &\leq \sup_{\mu} \mathbb{E} \left[ g_{[Lip]} K_0 e^{C_0(T-t)} \|U_1 - U_2\| + g_{[Lip]} C |T_{Exec_2}^{t,\mu} - T_{Exec_1}^{t,\mu}| + cq^a |T_{Exec_2}^{t,\mu} - T_{Exec_1}^{t,\mu}| \right] \\ &\leq \|U_1 - U_2\| (g_{[Lip]} K_0 e^{C_0(T-t)} + K'_0 \underline{C}) \leq A e^{C_0(T-t)} \|U_1 - U_2\|, \end{aligned}$$

where  $\underline{C} = g_{[Lip]} C + cq^a$ ,  $C$  is a constant,  $K'_0 = \log(K_0)$ ,  $\mu_1^{Opti}$  (resp.  $\mu_2^{Opti}$ ) is the optimal control when  $U_1$  (resp.  $U_2$ ) is the starting point and  $A = g_{[Lip]} K_0 + K'_0 \underline{C}$ . In the penultimate inequality, we use Inequality (29) to complete the proof.  $\square$

*Proof of Inequality (28).* Inequality (28) is proved using the dynamic programming principle and Inequality (27).  $\square$

### II.D.6 Proofs of Propositions II.D.2 and II.D.3

*Proof of Proposition II.D.2.* We fix  $\Delta > 0$  and prove the result by recurrence on  $n \geq 0$  for every  $T \in [0, n\Delta]$ .

- *Initialisation:* Case  $n = 0$ , in this case  $T_{Exec}^{t,\mu_2^{Opti}} = T_{Exec}^{t,\mu_1^{Opti}} = 0$ .
- *Iteration:* Let us assume the result true for  $T \in [0, n\Delta]$ . Let  $T \in [0, (n+1)\Delta]$ . When  $T \in [0, n\Delta]$ , the result is true using the recurrence assumption. When  $T \in (n\Delta, (n+1)\Delta]$ , we can write

$$T_{Exec}^{t,\mu_2^{Opti}} = T_{Exec}^{t,\mu_2^{Opti}} \mathbf{1}_{T_{Exec}^{t,\mu_2^{Opti}} \leq \Delta} + (T_{Exec}^{t,\mu_2^{Opti}} - \Delta) \mathbf{1}_{T_{Exec}^{t,\mu_2^{Opti}} \geq \Delta} + \Delta \mathbf{1}_{T_{Exec}^{t,\mu_2^{Opti}} \geq \Delta}.$$

Let  $\tilde{U}^{\tilde{\mu}}$  be the process following the same dynamics as  $U^\mu$  but with initial value  $U^{\tilde{\mu}}$  and ending at  $T - \Delta$  with a control  $\tilde{\mu}_t = \mu_{t+\Delta}$ . Then, we have classically  $T_{Exec}^{\tilde{\mu}} = T_{Exec}^{\mu, \Delta} - \Delta$  and  $V(\Delta, u) = V^{T-\Delta}(0, u)$ . Thus, we can write

$$\begin{aligned} \mathbb{E}[|T_{Exec}^{t,\mu_2^{Opti}} - T_{Exec}^{t,\mu_1^{Opti}}|] &= \mathbb{E} \left[ \overbrace{|T_{Exec}^{t,\mu_2^{Opti}} \mathbf{1}_{T_{Exec}^{t,\mu_2^{Opti}} \leq \Delta} - T_{Exec}^{t,\mu_1^{Opti}} \mathbf{1}_{T_{Exec}^{t,\mu_1^{Opti}} \leq \Delta}|}^{(1)} \right] + \mathbb{E} \left[ \overbrace{|T_{Exec}^{t,\mu_2^{Opti}} - T_{Exec}^{t,\tilde{\mu}_1^{Opti}}|}^{(2)} \right] \\ &\quad + \Delta \mathbb{E} \left[ \overbrace{|\mathbf{1}_{T_{Exec}^{t,\mu_2^{Opti}} \geq \Delta} - \mathbf{1}_{T_{Exec}^{t,\mu_1^{Opti}} \geq \Delta}|}^{(3)} \right]. \end{aligned}$$

- For Part (1), we have

$$\begin{aligned} \mathbb{E}\left[ \left| T_{Exec}^{t, \mu_2^{Opti}} \mathbf{1}_{T_{Exec}^{t, \mu_2^{Opti}} \leq \Delta} - T_{Exec}^{t, \mu_1^{Opti}} \mathbf{1}_{T_{Exec}^{t, \mu_1^{Opti}} \leq \Delta} \right| \right] &\leq \mathbb{E}\left[ \left| T_{Exec}^{t, \mu_2^{Opti}} \mathbf{1}_{T_{Exec}^{t, \mu_2^{Opti}} \leq \Delta} \right| \right] + \mathbb{E}\left[ \left| T_{Exec}^{t, \mu_1^{Opti}} \mathbf{1}_{T_{Exec}^{t, \mu_1^{Opti}} \leq \Delta} \right| \right] \\ &\leq \Delta \left( \mathbb{E}\left[ \mathbf{1}_{T_{Exec}^{\mu_2^{Opti}} \leq \Delta} \right] + \mathbb{E}\left[ \mathbf{1}_{T_{Exec}^{\mu_1^{Opti}} \leq \Delta} \right] \right) \leq \Delta^2 2\beta. \end{aligned}$$

- For Part (2), using the recurrence assumption and Inequality (26), we have

$$\begin{aligned} \mathbb{E}\left[ \left( T_{Exec}^{\tilde{\mu}_2^{Opti}} - T_{Exec}^{\tilde{\mu}_1^{Opti}} \right) \right] &\leq \Delta e^{C_1(T-(t+\Delta))} \mathbb{E}\left[ \left| U_{\Delta}^1 - U_{\Delta}^2 \right| \right] + K_1 \Delta (T - t - \Delta) \\ &\leq \Delta K_0 e^{C_1(T-(t+\Delta)) + C_0 \Delta} \|U_1 - U_2\| + K_1 \Delta (T - t - \Delta). \end{aligned}$$

- For Part (3), using the same arguments of Part (2), we have

$$\Delta \mathbb{E}\left[ \left| \mathbf{1}_{T_{Exec}^{t, \mu_2^{Opti}} \geq \Delta} - \mathbf{1}_{T_{Exec}^{t, \mu_1^{Opti}} \geq \Delta} \right| \right] \leq \Delta^2 2\beta.$$

Finally, since  $C_1 = \frac{\log(K_0)}{\Delta} + C_0$ , we have

$$\begin{aligned} \mathbb{E}\left[ T_{Exec}^{t, \mu_2^{Opti}} - T_{Exec}^{t, \mu_1^{Opti}} \right] &\leq 4\Delta^2 \beta + K_0 \Delta e^{C_1(T-(t+\Delta)) + C_0 \Delta} \|U_1 - U_2\| + K_1 \Delta (T - t - \Delta) \\ &\leq K_1 \Delta (T - t) + K_0 \Delta e^{C_1(T-(t+\Delta)) + C_0 \Delta} \|U_1 - U_2\| \\ &= \Delta \left( K_1 (T - t) + e^{C_1(T-t)} \|U_1 - U_2\| \right). \end{aligned}$$

□

*Proof of Proposition II.D.3.* Using Proposition II.D.2, we have

$$\begin{aligned} \mathbb{E}\left[ T_{Exec}^{\mu_2^{Opti}, t} - T_{Exec}^{\mu_1^{Opti}, t} \right] &\leq \Delta (\|U_1 - U_2\| e^{C_1(T-t)} + K_1 (T - t)) \\ &\leq \lim_{\Delta \rightarrow 0} C_0 \Delta \|U_1 - U_2\| + \frac{\log(K_0)}{\Delta} \Delta \|U_1 - U_2\| + K_1 \Delta (T - t) \\ &= K'_0 \|U_1 - U_2\|. \end{aligned}$$

□

## II.E Resolution of the optimal control problem

### II.E.1 Proof of Theorem 1

First, let us assume that the time derivative  $\partial_t V$  is continuous in each sub-interval  $(k\Delta, (k+1)\Delta)$ . Then, we can show classically that  $V$  satisfies the equations of Theorem 1 by applying Itô's formula. Thus, it suffices to exhibit a solution and use a verification argument to conclude. Equation (13) in Section 4.3 provides a solution that satisfies the required regularity of  $\partial_t V$ . Thus, we conclude using a verification theorem as in Theorem 4.1 in [133].

### II.E.2 Proof of Theorem 2

Let  $G = ([0, \hat{Q}^{\max}]^7 \times [0, \bar{P}^{Exec}])$ ,  $G_0 = \left( [0, \hat{Q}^{\max}]^4 \times \{0\} \times [0, \hat{Q}^{\max}]^2 \times [0, \bar{P}^{Exec}] \right) \cup \left( [0, \hat{Q}^{\max}]^6 \times [\bar{P}, \hat{Q}^{\max}] \times [0, \bar{P}^{Exec}] \right)$ , and  $\tilde{g}$  and  $g$  are Lipschitz functions representing respectively the boundary and final constraints. We denote by  $\hat{Q}^{\max} = \max(Q^{\max}, \tilde{Q}^{\max}, \bar{P})$ ,  $\bar{P}$  the maximum mid price value and  $\bar{P}^{Exec}$  the upper bound for the execution price. Equations satisfied by  $V$  can be formally derived by assuming that  $V$  is smooth and using the dynamic programming principle

$$\begin{cases} \max\left(\partial_t v + \mathcal{Q}v, \sup_{r \in \mathbb{E}} \mathcal{K}^r v - v\right) = 0, & \text{on } [0, T) \times G, \\ v = \tilde{g}, & \text{on } [0, T) \times G_0, \\ v(T, \cdot) = g, & \text{on } G, \end{cases} \quad (36)$$

with  $\mathcal{Q}f(u) = \int f(u') - f(u) dQ(u'; u)$  the infinitesimal generator of the state process,  $\mathcal{K}^r f(u) = \int f(u') dk^r(u'; u)$  for every continuous and bounded function  $f$ , state  $u$  and control  $r \in \mathbb{E}$ . Since a control  $r$  may lead to several states, we write  $k^r(u'; u)$  for the probability to reach the state  $u'$  starting from  $u$  after taking the decision  $r$ .

**Existence, uniqueness of the solution:** Uniqueness of the solution comes from a standard comparison principle using the same arguments as in [33, Theorem 2.2]. Existence of the solution can also be derived following [33, Theorem 2.3].

**Regularity of the solution:** Let us show that  $\partial_t V$  is continuous except on the boundary of  $\{V = g\}$ . We denote by  $V$  the continuous and Lipschitz viscosity solution of (36).

Let  $r$  be the control which modifies the agent's state when it exists. Let  $O$  be the open set  $O = \{V > \max(\mathcal{K}^r V, g)\} \cup \{(t, u); V > g, k^r(u; u) = 1\}$ . On  $O$ , we have  $\partial_t V = -\mathcal{Q}V$  in the viscosity sense. Hence, by considering a sequence of smooth functions converging uniformly towards  $V$ , we have  $\partial_t V$  is continuous on  $O$ , see [37, Corollary 5.6] for a close construction.

Let  $O_1 = \{\mathcal{K}^r V = V, V > g\}$  and  $\overset{\circ}{O}_1$  its interior assumed non-empty, otherwise there is nothing to prove. Since  $V$  is Lipschitz,  $\partial_t V$  is essentially bounded. To show that  $\partial_t V$  is uniquely defined on  $\overset{\circ}{O}_1$ , we assume the opposite and consider a point  $x_0 = (t_0, u_0)$  where  $\partial_t V$  admits two possible values. We have

$$V(x_0) = \mathcal{K}^r V = \int V(t_0, u') dk^r(u'; u_0).$$

There exists at least one  $u'_0$  satisfying  $k^r(u'_0; u_0) > 0$  and  $(t_0, u'_0) \in O$ . To see this, let us take  $u'_0 = \operatorname{argmax}\{V(t_0, u'), k^r(u'; u_0) > 0\}$ . Since  $V(t_0, u'_0) \geq V(x_0) > g$ , we have  $(t_0, u'_0) \notin \{V = g\}$ . If  $(t_0, u'_0) \in O$ , it is exactly the needed result. If  $(t_0, u'_0) \notin O$ , then  $V(t_0, u'_0) = \mathcal{K}^r V = \int V(t_0, u') dk^r(u'; u'_0) \leq V(t_0, u'_0)$ . Hence, the only possibility is  $k^r(u'_0; u'_0) = 1$  which provides the needed contradiction.

Since  $u'_0 \in O$ , then  $\partial_t V(t_0, u'_0)$  is uniquely defined and  $\partial_t V$  is continuous around  $(t_0, u'_0)$ . Hence, the function  $\partial_t \tilde{V}$ , with  $\tilde{V} = V - k^r(u'_0; \cdot)V(\cdot, u'_0)$ , is not uniquely defined in  $x_0$  and  $\tilde{V}$



## II. Optimal liquidity-based trading tactics

---

satisfies

$$\tilde{V}(x_0) = \tilde{\mathcal{K}}^r V = \sum_{u', u' \neq u'_0} V(t_0, u') k^r(u'; u_0).$$

Since the sum in the above equation is finite, we can apply the same arguments as before several times to find that the null function is not uniquely defined which provides the needed contradiction. Hence  $\partial_t V$  is uniquely defined on  $\overset{\circ}{O}_1$ . Furthermore, since  $\partial_t V$  is continuous on  $O$ , we can prove by contradiction and using the same arguments that  $\partial_t V$  is continuous on  $\overset{\circ}{O}_1$  and thus on  $\bar{O}_1$ .

Let  $O_2 = \bar{O} \cap \bar{O}_1$ , where  $\bar{O}_1$  is the closure of  $O_1$  and  $x$  be a point on  $O_2$ . Thus,  $x$  is the limit point of  $(x_n)_{n \geq 0}$  and  $(x_n^1)_{n \geq 0}$ , such that  $(x_n)_n \in O$  and  $(x_n^1)_n \in O_1$ . Let  $l$  (resp.  $l_1$ ) be the limit value of  $\lim_{n \rightarrow \infty} \partial_t V(x_n)$  (resp.  $\lim_{n \rightarrow \infty} \partial_t V(x_n^1)$ ). Hence, we can check

$$l = \lim_{n \rightarrow \infty} \partial_t V(x_n) = \lim_{n \rightarrow \infty} QV(x_n) = QV(x) = \lim_{n \rightarrow \infty} QK^r V(x_n^1) = K^r \lim_{n \rightarrow \infty} QV(x_n^1) = K^r \lim_{n \rightarrow \infty} \partial_t V(x_n^1) = l_1.$$

Thus,  $\partial V$  is continuous on  $O_3 = \bar{O} \cup \bar{O}_1$ . On the set  $O_4 = \{V = g\}$ ,  $\partial_t V$  is clearly continuous since  $\partial_t V = 0$ .

Finally, we consider the set  $O_5 = \partial O_4$  and  $x$  a point on  $O_5$ . Here again,  $x$  is the limit point of  $(x_n)_{n \geq 0}$  and  $(x_n^1)_{n \geq 0}$ , such that  $(x_n)_n \in O_3$  and  $(x_n^1)_n \in O_4$ . Let  $l$  (resp.  $l_1$ ) be the limit value of  $\lim_{n \rightarrow \infty} \partial_t V(x_n)$  (resp.  $\lim_{n \rightarrow \infty} \partial_t V(x_n^1)$ ). Thus, the condition  $l = l_1$  is equivalent to  $Qg = 0$ . This relation is not necessarily satisfied.

**Conclusion:** Equation (36) is satisfied almost everywhere by  $V$ . Since  $\partial_t V$  is continuous except on the set  $O_5 = \partial\{V = g\}$ , Equation (36) is satisfied pointwise except on  $O_5$ .

### II.E.3 Optimal strategy

Let  $\tau_i^T := \tau_i \wedge T_{Exec}^\mu$ . Since  $V$  satisfies Equation (14), we have

$$\begin{aligned} \mathbb{E}[g(U_{T_{Exec}^\mu}^\mu) - cq^a T_{Exec}^\mu] &= \mathbb{E}[V(T_{Exec}^\mu, U_{T_{Exec}^\mu}^\mu)] \\ &= V(0, U_0) + \sum_{i \geq 0} \mathbb{E} \left[ \int_{\tau_i^T}^{\tau_{i+1}^T} [\mathcal{A}V(s, U_s^\mu) - cq^a] ds + [V(\tau_i, (U_{\tau_i^-}^\mu)^{\beta_i}) - V(\tau_i, U_{\tau_i^-}^\mu)] \right] \\ &= V(0, U_0). \end{aligned}$$

Since, by construction,  $\mathbb{E}[V(\tau_i, (U_{\tau_i^-}^\mu)^{\beta_i}) - V(\tau_i, U_{\tau_i^-}^\mu)] = 0$ , and  $\mathcal{A}V(\cdot, U) - cq^a = 0$ , this shows that this policy satisfies  $\mathbb{E}[g(U_{T_{Exec}^\mu}^\mu) - cq^a T_{Exec}^\mu] = V(0, U_0)$ , and is therefore optimal, by definition of  $V(0, U_0)$ .

## II.F Proof of Theorem 3

### II.F.1 Proof of Inequality (17)

Let us fix  $\Delta$  and show the result by recurrence on  $n$  for every  $T \in [0, n\Delta]$ .

**Initialisation:** in this case we have  $V' = V = g$ .

**Iteration:** let us assume the result true for  $n$ . Let  $T \in [0, (n+1)\Delta]$ .

- When  $T \in [0, n\Delta]$ : the result is true using the recurrence assumption.
- When  $T \in (n\Delta, (n+1)\Delta]$ : let  $t \in [0, T]$ . When  $t \in (\Delta, T]$ , the result is true by using  $V(t, U) = V_{T-t}(0, U)$ ,  $\tilde{V}^\Delta(t, U) = \tilde{V}_{T-t}^\Delta(0, U)$  and the recurrence hypothesis. Let us take  $t \in [0, \Delta, T)$ . Using the dynamic programming principle, we have

$$|\tilde{V}^\Delta(t, u) - V(t, u)| \leq \left| \sup_{\mu} \mathbb{E} \left[ cq^a ((\tilde{T}_{Exec}^{t, \mu} - t) \mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} \leq t+\Delta} - [T_{Exec}^{t, \mu} - t] \mathbf{1}_{T_{Exec}^{t, \mu} \leq t+\Delta}) \right. \right. \\ \left. \left. + cq^a \Delta (\mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} > t+\Delta} - \mathbf{1}_{T_{Exec}^{t, \mu} > t+\Delta}) + (\tilde{V}_{T-t}^\Delta(\Delta, \tilde{U}_\Delta^{\mu, \Delta}) - V_{T-t}(\Delta, U_\Delta^\mu)) \right] \right|.$$

- First we have

$$\mathbb{E}[|(\tilde{T}_{Exec}^{t, \mu} - t) \mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} \leq t+\Delta} - (T_{Exec}^{t, \mu} - t) \mathbf{1}_{T_{Exec}^{t, \mu} \leq t+\Delta}|] \leq \Delta \mathbb{E}[\mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} \leq t+\Delta} + \mathbf{1}_{T_{Exec}^{t, \mu} \leq t+\Delta}] \leq \Delta^2 2H.$$

- Second, using (4.4.2), we have

$$\begin{aligned} |\mathbb{E}[(\tilde{V}_{T-t}^\Delta(\Delta, \tilde{U}_\Delta^{\mu, \Delta}) - V_{T-t}(\Delta, U_\Delta^\mu))]| &= \left| \sum_{u'} [P_{u, u'}^\Delta \tilde{V}_{T-t}^\Delta(\Delta, u') - \mathbb{P}[U_\Delta = u' | U_0 = u] V_{T-t}(\Delta, u')] \right| \\ &\leq \sum_{u'} [P_{u, u'}^{*, \Delta} |(\tilde{V}_{T-t}^\Delta(\Delta, u') - V_{T-t}(\Delta, u'))|] + D \|g\|_\infty \Delta^2 \\ &\leq R(T-t-\Delta)\Delta + D \|g\|_\infty \Delta^2. \end{aligned}$$

- Finally, we have

$$cq^a \Delta \mathbb{E}[|\mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} > t+\Delta} - \mathbf{1}_{T_{Exec}^{t, \mu} > t+\Delta}|] \leq cq^a \Delta \mathbb{E}[|\mathbf{1}_{T_{Exec}^{t, \mu} \leq t+\Delta} + \mathbf{1}_{\tilde{T}_{Exec}^{t, \mu} \leq t+\Delta}|] \leq cq^a \Delta^2 2H.$$

By combining above inequalities, we conclude

$$|\tilde{V}_{T-t}^\Delta(t, u) - V_{T-t}(t, u)| \leq R(T-t-\Delta)\Delta + R\Delta^2 \leq R(T-t)\Delta.$$

### II.F.2 Proof of the stationary convergence of $\mu^{Opti, \Delta}$

Let  $\mu^{Opti, \Delta}$  be the piecewise constant optimal control associated to the process  $\tilde{U}_t^{\mu, \Delta}$ . We say that a sequence of functions  $f^n$  converges to  $f$  in a stationary way when  $\exists n_0$  such that  $\forall n \geq n_0, f^n = f$ .

## II. Optimal liquidity-based trading tactics

---

**Outline of the proof:** First, we prove the existence of a subsequence  $(\phi_n)_{n \geq 0}$  such that  $\mu^{Opti, \Delta_{\phi(n)}}(\omega) \xrightarrow{n \rightarrow \infty} \bar{\mu}(\omega)$  in a stationary way, where  $\bar{\mu}(\omega)$  is a piecewise constant function.

Then, using  $\tilde{V}^\Delta(t, U) \xrightarrow{\Delta \rightarrow 0} V(t, U)$  and the stationary convergence  $\mu^{Opti, \Delta_{\phi(n)}}$ , there exists  $n_0$

such that  $\forall n \geq n_0$ ,  $\mathbb{E}[g(\tilde{U}_{\tilde{T}_{Exec}^{\bar{\mu}}}^{\Delta_{\phi(n)}, \bar{\mu}}) - cq^a \tilde{T}_{Exec}^{\bar{\mu}}] = \mathbb{E}[g(\tilde{U}_{\tilde{T}_{Exec}^{\mu^{Opti, \Delta_{\phi(n)}}}}^{\Delta_{\phi(n)}, \mu^{Opti, \Delta_{\phi(n)}}}) - cq^a \tilde{T}_{Exec}^{\mu^{Opti, \Delta_{\phi(n)}}}] \xrightarrow{\Delta \rightarrow 0} V(t, U)$ .

Since  $U_t^\mu$  is right continuous,  $\bar{\mu}$  is optimal.

**Proof of the stationary convergence:** First, let us prove that there exists  $\epsilon > 0$  such that for every  $a \in [0, T]$ , we can find a subsequence  $\mu^{Opti, \Delta_{\phi^a(n)}}(\omega)$  which is stationary in  $[a, a + \epsilon]$ . Let  $a \in [0, T]$ , since the space  $\mathcal{C} = \{l, c, m\}$  is compact, we can extract a subsequence  $\phi^a(n)$  such that  $\mu^{Opti, \Delta_{\phi^a(n)}}(\omega)(a)$  converges towards a given limit  $\mu(\omega)(a)$ . Since  $\mathcal{C}$  is finite the sequence  $\mu^{Opti, \Delta_{\phi^a(n)}}(\omega)(a)$  is stationary. Let  $\epsilon(\omega) > 0$  be the minimum time between two consecutive jumps in  $[0, T]$ . Hence,  $\mu^{Opti, \Delta_{\phi^a(n)}}(\omega)(a)$  is constant in  $[a, a + \epsilon]$ . Consequently,  $\mu^{Opti, \Delta_{\phi^a(n)}}(\omega)(x) \xrightarrow{n \rightarrow \infty} \mu(\omega)(a)$ ,  $\forall x \in [a, a + \epsilon]$  in a stationary way.

Let  $m_\epsilon = \lfloor \frac{T}{\epsilon} \rfloor$ . For every  $i \in \{0, \dots, m_\epsilon\}$ , there exists  $\phi^{i\epsilon}$  such that

$$\mu^{Opti, \Delta_{\phi^{i\epsilon}}}(\omega)(x) \xrightarrow{n \rightarrow \infty} \mu(\omega)(i\epsilon), \quad \forall x \in [i\epsilon, (i+1)\epsilon].$$

We define the piecewise constant limit function  $\bar{\mu}(\omega)$  such that

$$\bar{\mu}(\omega)(x) = \mu(\omega)(i\epsilon), \quad \forall x \in [i\epsilon, (i+1)\epsilon].$$

By construction, there exists  $\phi(n)$  (constructed by a finite number of  $\phi^{i\epsilon}(n)$  compositions), the sequence  $\mu^{Opti, \Delta_{\phi(n)}}(\omega) \xrightarrow{n \rightarrow \infty} \bar{\mu}(\omega)$  in a stationary way.

---

# Optimal inventory management and order book modeling

---

## Abstract

We model the behaviour of three agent classes acting dynamically in a limit order book of a financial asset. Namely, we consider market makers (MM), high frequency directional trading (DHFT) firms, and institutional brokers (IB). Given a prior dynamics of the order book, similar to the one considered in the Queue-Reactive models [55, 90, 91], the MM and the HFT define their trading strategy by optimizing the expected utility of terminal wealth, while the IB has a prescheduled task to sell or buy many shares of the considered asset. We derive the variational partial differential equations that characterize the value functions of the MM and HFT and explain how almost optimal control can be deduced from them. We then provide a first illustration of the interactions that can take place between these different market participants by simulating the dynamics of an order book in which each of them acts optimally.

Key words: Optimal trading, market impact, optimal control.

MSC 2010: 49L20, 49L25.

## 1 Introduction

The comprehension of the order book dynamics have become a fundamental issue for all market participants and for regulators that try to increase the market transparency and efficiency. A deep understanding of the order book dynamics and agents behaviours enables: market makers to ensure liquidity provision at cheaper prices, high frequency traders to reduce arbitrage opportunities, investors to reduce their transaction costs, policy makers to design relevant rules, to strengthen market transparency and to reduce market manipulations. Moreover, modelling the order book provides insights on the behaviour of the price at larger time scales since the price formation process starts at the order book level, see e.g. [54] for Brownian diffusion asymptotic of rescaled price processes. Recently, the widespread market electronification has facilitated the access of high quality data describing market participants

decisions and interactions at the finest time scale, on which statistics can be based. The availability of the order book data certainly allows a better understanding of the market activity. On the other hand, the recent market fragmentation and the increase of trading frequency rise the complexity of agents actions and interactions.

The main objective of the present paper is to propose a flexible order book model close to the one of [55, 90, 91] and Chapter II, construct a first building block towards a realistic order book modelling, and try to better understand the various regimes related to the presence of different market participants. Instead of considering pure statistical dynamics as in e.g. [55], we construct a structural endogenous dynamics, see e.g. Chapter I and Chapter II, based on the optimal behaviour of agents that are assumed to be rational. For numerical tractability, we simplify the market in three classes of (most significant) participants: the market makers (MM), the directional high frequency trading (DHFT) firms, and the institutional brokers (IB). Each of them decides of his policy in an optimal way, given prior statistics, and then interact with the others given the endogenous realizations of the market.

More precisely, they assume that the market dynamics follow an order book model similar to the one suggested by the Queue-Reactive approach [90], see also [1, 55, 81], in which we restrict ourselves to the best bid and ask queues<sup>1</sup>: limit and aggressive orders arrive with certain intensities, when a queue is depleted it is regenerated according to a certain law and possibly after a price move, the spread can take two different values.<sup>2</sup> Importantly, we also take the imbalance into account in the modelling of the different transition probabilities, see e.g. Besson et al. [30].

The market participants can either put limit or aggressive orders. The aim of a market maker is to gain the spread. He should therefore essentially insert limit orders, aggressive orders being used when his inventory is too unbalanced. In our model, he can only act on the given order book. The directional high frequency trader is assumed to play on the correlation between the order book dynamics, viewed as the stock price, and the price of another asset, called futures hereafter. Indeed, he believes that the difference between the stock and the futures prices is mean-reverting. Whenever he buys/sells one unit of the stock, he sells/buys back one unit of the futures. We do not handle the order book associated to the futures but simply model the price of the futures as the mid-price of the stock to which a mean-reverting process is added. Still, we introduce a (possibly equal to 0) transaction cost proportional to the size of the transaction. As the market-makers, he seeks for a zero inventory at the end of the trading period. Finally, institutional brokers are simply assumed to play VWAP (Volume Weighted Average Price)- or Volume-based strategies (trading algorithms). Again, they essentially use limit orders and become aggressive when they are too late in their schedule.

---

<sup>1</sup>This limitation is for numerical tractability. It is already enough for most markets.

<sup>2</sup>One could consider a larger set of possible spreads in theory, but refrains to doing this for numerical tractability.

To solve the IB control problem we follow the Cartea-Jaimungal approach, see [52] and its various extensions [47, 48, 49, 50, 51].

We focus on the derivation of the optimal strategies of the market makers and the high frequency traders, and on how they can be computed numerically by solving the associated variational partial differential equations. Note that it is important to consider their strategies within a dynamical model as current actions impact the order book and therefore may modify its futures dynamics. We will actually see that, in certain situations, participants can place aggressive orders or limit orders in the spread just to try to manipulate the order book in a favorable way. Note that the dimension of our control space is higher than in [1, 81] and Chapters I and II since more complex decisions are required to tackle the market making problem.

However, the ultimate goal of this work is to provide a market simulator. In the last section, we already present simulations of the market behaviour given the pre-computed optimal strategies of the different actors. More precisely, we will only simulate the evolution of the mean-reverting process (driving the difference between the stock and the futures price) together with the reconstruction of the queues when prices move, and let the participants play their optimal strategies given the evolution of the order book due to their different actions. This should allow us to study how these different market participants may interact with each others when they act optimally. In particular, one should observe different market regimes depending on the proportion of the different participants in the total population, on their risk aversion, etc. First simulations are provided in this paper, a more throughout study is left for future research.

Our approach therefore lies in between two current streams of literature. The first one is based on “general equilibrium models”, including economic models, where the market activity is generated by interactions between rational agents who take optimal decisions that interact through the market netting process, see e.g. [66, 134, 139]. The second stream of literature considers purely statistical models where the order book is seen as a random process, see e.g. [2, 3, 20, 55, 57, 77, 83, 82, 90, 91, 106, 107, 144]. The statistical models focus on reproducing many salient features of a real market rather than agents behaviours and interactions.

We end this introduction with an outline of our paper. In Section 2, we present the general order book dynamics. The market maker control problem is studied in details in Section 3. Therefore, we present the equations satisfied by their optimal strategy and propose a numerical solution for this problem, together with numerical illustrations. In Section 4, we formulate the directional high frequency trader control problem and perform a similar analysis. The institutional broker strategy is described in Section 5, where we restrict ourselves to VWAP and Volume liquidation problems. Finally, in Section 6, we simulate a realistic market using the three agent’s optimal trading strategies.

## 2 General order book presentation and priors of the market participants

As mentioned above, we focus on a single order book and only model the best bid and ask prices, in a similar way to [55]. In this section, we describe the general market mechanisms as well as the priors on which the optimal strategies of the agents are based. We fix a terminal horizon time  $T$  and consider a probability space  $(\Omega, \mathbb{P})$ . Here,  $\Omega := \Omega_1 \times \Omega_2$  and  $\mathbb{P} := \mathbb{P}_1 \otimes \mathbb{P}_2$ , where  $\Omega_1$  is the space of  $\mathbb{R}^{11}$ -valued cadlag paths on  $[0, T]$  endowed with a probability measure  $\mathbb{P}_1$  with full support on  $\Omega_1$ , and  $\Omega_2$  is the one dimensional Wiener space endowed with the Wiener measure  $\mathbb{P}_2$ .

We denote respectively by  $(P_t^b)_{t \geq 0}$  and  $(P_t^a)_{t \geq 0}$  the best bid offer and the best ask offer processes on the market. They are valued in  $\mathfrak{d}\mathbb{Z}$  where  $\mathfrak{d} > 0$  is the tick size. We denote by  $(Q_t^b)_{t \geq 0}$  and  $(Q_t^a)_{t \geq 0}$  the sizes of the corresponding queues valued in  $\mathbb{N}^*$ . To simplify the notation, we introduce  $P := (P^b, P^a)$ ,  $Q := (Q^b, Q^a)$  and define the spread process as  $\delta P := P^a - P^b$ . Moreover, we assume<sup>3</sup> that  $\delta P_t \in \{\mathfrak{d}, 2\mathfrak{d}\}$  for all  $t \geq 0$ .

We denote by  $(\tau_i)_{i \geq 1}$  the times at which orders are sent to the market. We assume that this sequence is increasing and that  $\#\{i \in \mathbb{N} : \tau_i < T\} < \infty$  a.s. The market participants can send different types of orders at each time  $\tau_i$ :

- Aggressive orders of size  $\alpha_i^b \in \mathbb{N} \cap [0, Q^b]$  at the bid or of size  $\alpha_i^a \in \mathbb{N} \cap [0, Q^a]$  at the ask: the size of the corresponding queue,  $Q^b$  or  $Q^a$ , decreases by the size of the aggressive order,  $\alpha_i^b$  or  $\alpha_i^a$ .
- Limit orders of size  $L_i^b \in \mathbb{N}$  at the bid or of size  $L_i^a \in \mathbb{N}$  at the ask: the size of the corresponding queue,  $Q^b$  or  $Q^a$ , increases by the size of the limit order,  $L_i^b$  or  $L_i^a$ .
- When  $\delta P = 2\mathfrak{d}$ : limit orders of size  $L_i^{b, \frac{1}{2}} \in \mathbb{N}$  at the bid or of size  $L_i^{a, \frac{1}{2}} \in \mathbb{N}$  at the ask: the order is placed inside the spread, at the price  $P^b + \mathfrak{d} = P^a - \mathfrak{d}$ , which generates a new queue at the bid or at the ask, of size  $L_i^{b, \frac{1}{2}}$  or  $L_i^{a, \frac{1}{2}}$ , and a price move.
- Cancellations: cancellations of  $M_i^b \in \mathbb{N} \cap [0, Q^b]$  orders at the bid or of  $M_i^a \in \mathbb{N} \cap [0, Q^a]$  orders at the ask. The difference between cancellations and aggressive orders is that aggressive orders consume the bottom of the limit while cancellations consume the top of the limit first.

---

<sup>3</sup>The extension to more possible spread values is straightforward. We stick to this setting for notational and computational simplicity. Note that this limit is also justified by empirical evidences for many stocks, see [55].

## 2. General order book presentation and priors of the market participants

---

The sequence  $(\alpha_i^b, \alpha_i^a, L_i^b, L_i^a, L_i^{b, \frac{1}{2}}, L_i^{a, \frac{1}{2}}, M_i^b, M_i^a)_{i \geq 1}$  is a random variable leaving, with probability one, on the state space  $\mathbf{C}_\circ$  defined as the collection of elements  $(a^b, a^a, \ell^b, \ell^a, \ell^{b, \frac{1}{2}}, \ell^{a, \frac{1}{2}}, m^b, m^a) \in \mathbb{N}^8$  such that

$$\begin{cases} a^b a^a = 0, \\ \ell^b = \ell^a = 0 \text{ if } \max\{a^b, a^a\} \geq 1, \\ \ell^{b, \frac{1}{2}} = 0 \text{ if } \max\{a^b, a^a, \ell^b\} \geq 1, \\ \ell^{a, \frac{1}{2}} = 0 \text{ if } \max\{a^b, a^a, \ell^a\} \geq 1, \\ m^b = 0 \text{ if } \max\{a^b, a^a, \ell^b, \ell^{b, \frac{1}{2}}\} \geq 1, \\ m^a = 0 \text{ if } \max\{a^b, a^a, \ell^a, \ell^{a, \frac{1}{2}}\} \geq 1. \end{cases} \quad (1)$$

We interpret the above expression as follows. First, aggressive orders cannot be sent simultaneously at the bid and at the ask. Next, limit orders (at the current bid/ask prices) cannot be placed when aggressive orders are already sent. Finally, one cannot insert limit orders within the spread if limit orders at the current bid/ask prices are placed. Because we only consider the first limits, these conditions are natural whenever one presumes that orders of different market participants do not arrive exactly at the same time.

Depending on the arrival of orders, queues can be depleted. In this case, new queues can be re-generated, at the same prices or at different prices, and possibly with a change of the spread value. To model this, we introduce a sequence of random variables  $(\epsilon_i, \epsilon_i^b, \epsilon_i^a)_{i \geq 1}$  with values in  $\{0, 1\} \times (\mathbb{N}^*)^2$ . The sequence  $(\epsilon_i)_{i \geq 1}$  describes possible jumps of the bid/ask prices, while the sequence  $(\epsilon_i^b, \epsilon_i^a)_{i \geq 1}$  describes the new sizes of the queues when they are re-generated after a depletion. More precisely, we postulate that

$$\begin{aligned} P_{\tau_i}^b &= P_{\tau_{i-1}}^b + \mathfrak{d} \left[ \mathbf{1}_{\{\epsilon_i=1\}} \left( -\mathbf{1}_{\{Q_{\tau_{i-1}}^b = \hat{\alpha}_i^b\}} + \mathbf{1}_{\{\delta P_{\tau_{i-1}} = 2\mathfrak{d}\}} \mathbf{1}_{\{Q_{\tau_{i-1}}^a = \hat{\alpha}_i^a\}} \right) + \mathbf{1}_{\{L_i^{b, \frac{1}{2}} > 0\}} \right], \\ P_{\tau_i}^a &= P_{\tau_{i-1}}^a + \mathfrak{d} \left[ \mathbf{1}_{\{\epsilon_i=1\}} \left( \mathbf{1}_{\{Q_{\tau_{i-1}}^a = \hat{\alpha}_i^a\}} - \mathbf{1}_{\{\delta P_{\tau_{i-1}} = 2\mathfrak{d}\}} \mathbf{1}_{\{Q_{\tau_{i-1}}^b = \hat{\alpha}_i^b\}} \right) - \mathbf{1}_{\{L_i^{a, \frac{1}{2}} > 0\}} \right], \\ Q_{\tau_i}^b &= Q_{\tau_{i-1}}^b + L_i^b + (L_i^{b, \frac{1}{2}} - Q_{\tau_{i-1}}^b) \mathbf{1}_{\{L_i^{b, \frac{1}{2}} > 0\}} - \hat{\alpha}_i^b \mathbf{1}_{\{\Delta P_{\tau_i}^b = 0\}} + (\epsilon_i^b - Q_{\tau_{i-1}}^b) \mathbf{1}_{\{\Delta P_{\tau_i}^b \neq 0\} \cup \{\hat{\alpha}_i^b = Q_{\tau_{i-1}}^b\}}, \\ Q_{\tau_i}^a &= Q_{\tau_{i-1}}^a + L_i^a + (L_i^{a, \frac{1}{2}} - Q_{\tau_{i-1}}^a) \mathbf{1}_{\{L_i^{a, \frac{1}{2}} > 0\}} - \hat{\alpha}_i^a \mathbf{1}_{\{\Delta P_{\tau_i}^a = 0\}} + (\epsilon_i^a - Q_{\tau_{i-1}}^a) \mathbf{1}_{\{\Delta P_{\tau_i}^a \neq 0\} \cup \{\hat{\alpha}_i^a = Q_{\tau_{i-1}}^a\}}, \end{aligned} \quad (2)$$

for  $i \geq 1$ , where

$$\hat{\alpha}_i^{b/a} := \alpha_i^{b/a} + M_i^{b/a},$$

with

$$(P_0^b, P_0^a, Q_0^b, Q_0^a) \in D_{PQ} := \{(p^b, p^a, q) \in (\mathfrak{d}\mathbb{Z})^2 \times (\mathbb{N}^*)^2 : p^a - p^b \in \{\mathfrak{d}, 2\mathfrak{d}\}\},$$

and the convention  $\tau_0 = 0-$ . Note that (1) ensures that these dynamics are consistent. In particular, prices can move only if one of the queues is depleted because of the arrival of aggressive orders and cancellations or if a new limit order is inserted within the spread. These two situations cannot occur simultaneously. The ask price can move by  $\mathfrak{d}$  when the ask queue is depleted. If the spread was already  $2\mathfrak{d}$ , then the bid price moves up as well. The other way



around if the bid queue is depleted. In the following, we extend the dynamics of  $(P, Q)$  by considering it as a step constant right-continuous process on  $[0, T]$ .

We now denote by  $\mathcal{E}$  the  $\mathbb{N}^{12}$ -valued step constant right-continuous process defined by

$$\Delta \mathcal{E}_{\tau_i} := \mathcal{E}_{\tau_i} - \mathcal{E}_{\tau_{i-1}} := (\alpha_i^b, \alpha_i^a, L_i^b, L_i^a, L_i^{b, \frac{1}{2}}, L_i^{a, \frac{1}{2}}, M_i^b, M_i^a, \epsilon_i, \epsilon_i^b, \epsilon_i^a, 1), \quad i \geq 1,$$

with  $\mathcal{E}_{\tau_0} := \mathcal{E}_0 := 0$ . Later on, we shall only write

$$(P_{\tau_i}, Q_{\tau_i}) = \mathcal{T}_{P, Q}(P_{\tau_{i-1}}, Q_{\tau_{i-1}}, \Delta \mathcal{E}_{\tau_i}), \quad (3)$$

in which the map  $\mathcal{T}_{P, Q}$  is defined explicitly by (1).

The process  $\mathcal{E}$  models the flow of all the orders on the market. From the viewpoint of a market participant, it corresponds to its own orders and to the other participants' orders that we denote by  $\tilde{\mathcal{E}}$ . The process  $\tilde{\mathcal{E}}$  has jump times  $(\tilde{\tau}_i)_{i \geq 1} \subset (\tau_i)_{i \geq 1}$  of sizes

$$\Delta \tilde{\mathcal{E}}_{\tilde{\tau}_i} = \sum_{j \geq 1} \mathbf{1}_{\{\tilde{\tau}_i = \tau_j\}} \Delta \mathcal{E}_{\tau_j}, \quad \forall i \geq 1. \quad (4)$$

It induces a counting measure  $\tilde{\nu}(dt, de)$ . For a market participant, a prior on this measure is given by the compensator  $\tilde{\mu}$  of  $\tilde{\nu}$  under  $\mathbb{P}$ . We assume that it is state-dependent. More precisely, we consider a Borel kernel  $(p, q) \in D_{P, Q} \mapsto \tilde{\mu}(\cdot | p, q) \in \mathcal{M}([0, T] \times \mathbb{N}^8)$ , where  $\mathcal{M}([0, T] \times \mathbb{N}^8)$  denotes the collection of non-negative measures on  $[0, T] \times \mathbb{N}^8$ . In particular, it can depend on the order book's imbalance, as observed in e.g. [30]. To be consistent with the constraints imposed above, it satisfies

$$\tilde{\mu}(\cdot | p, q) \text{ is supported by } \mathbf{C}_\circ, \text{ for all } (p, q) \in D_{P, Q}. \quad (5)$$

Additionally, the aggressive orders and cancellations are never bigger than the corresponding queue size. This assumption is more explicit in our numerical example sections, see Section 3.4.

Next, we shall denote by  $\mathcal{E}^\phi$  the flows corresponding to the trading strategy of either a market maker, a directional high frequency trader, or an institutional broker. Thus, they face a global flow  $\mathcal{E} = \tilde{\mathcal{E}} + \mathcal{E}^\phi$ . Moreover, for simplicity, we assume that  $\tilde{\mu}$  is of the form

$$d\tilde{\mu}(c, \epsilon, \epsilon^b, \epsilon^a, dt | p, q) = d\lambda(\epsilon, \epsilon^b, \epsilon^a | p, q, c) d\beta(c | p, q) dt, \quad (6)$$

in which  $\lambda$  and  $\beta$  are bounded Borel non-negative kernels and (without loss of generality)

$$\int d\lambda(\epsilon, \epsilon^b, \epsilon^a | p, q, c) = 1, \text{ for all } (p, q, c) \in D_{P, Q} \times \mathbf{C}_\circ. \quad (7)$$

Later on, when an order  $(\alpha_i^b, \alpha_i^a, L_i^b, L_i^a, L_i^{b, \frac{1}{2}}, L_i^{a, \frac{1}{2}}, M_i^b, M_i^a)$  is sent by a MM, an HFT or an IB, we shall also assume that the conditional law of  $(\epsilon_i, \epsilon_i^b, \epsilon_i^a)$  is given by  $\lambda$ .

**Remark 2.1.** Let  $\gamma(p, q) := \int d\beta(c|p, q)$ . Then,  $\gamma$  is uniformly bounded by the above assumption. Let  $\tau$  be a stopping time and fix  $h > 0$ . Since  $\lambda$  integrates to one, it follows that  $\mathbb{P}[\#\{t \in [\tau, \tau + h] : \Delta\tilde{\mathcal{E}}_t \neq 0\} = 1 | \mathcal{F}_\tau] = h\gamma(P_\tau, Q_\tau) + o(h)$  and  $\mathbb{P}[\#\{t \in [\tau, \tau + h] : \Delta\tilde{\mathcal{E}}_t \neq 0\} > 1 | \mathcal{F}_\tau] = o(h)$ . Moreover, the process counting the number of jumps of  $\tilde{\mathcal{E}}$  is dominated by a Poisson process with intensity  $\tilde{\gamma} := \sup \gamma < \infty$ . Hence,  $\mathbb{P}[\#\{t \leq T : \Delta\tilde{\mathcal{E}}_t \neq 0\} \geq k] \leq \tilde{\gamma}T/k$ , for  $k \geq 1$ , by Markov's inequality. Similarly, if  $g$  is a non-decreasing Borel map, then  $\mathbb{E}[g(\#\{t \in [0, T] : \Delta\tilde{\mathcal{E}}_t \neq 0\})] \leq \sum_{k \geq 1} g(k) \frac{(\tilde{\gamma}T)^k}{k!} e^{-\tilde{\gamma}T}$ .

### 3 Market maker's optimal control problem

In this section, we describe the optimal control problem of the market maker, the key tools to characterize the solution and how to numerically approximate the optimal control.

#### 3.1 Market maker's strategy and the state dynamics

The market maker typically places limit orders in order to make profit of the spread but can turn aggressive when his inventory is too important. At the end of the trading period  $[0, T]$ , the later should be zero. In the following, we denote by  $G$  his gain process and by  $I$  his inventory. We also need to keep track of the sizes of the orders that he already placed at the bid queue,  $N^b$ , and at the ask queue,  $N^a$ . For simplicity, we impose that new orders cannot be inserted at the bid (respectively at the ask) if he already placed a limit order at the bid (respectively at the ask). Then, his position at the bid (resp. at the ask) is completely described by  $N^b$  (resp.  $N^a$ ) and the number of units  $B^b$  before him in the bid-queue (resp.  $B^a$  before him in the ask-queue). Later on, we only write  $N = (N^b, N^a)$  and  $B = (B^b, B^a)$ .

To define the market maker's control, we assume that he faces the exogenous process  $\tilde{\mathcal{E}}$  described in Section 2. For him, a control is a sequence of random variables  $\phi = (\tau_i^\phi, \mathbf{c}_i^\phi)_{i \geq 1}$  where  $(\tau_i^\phi)_{i \geq 1}$  is an increasing sequence of times and each

$$\mathbf{c}_i^\phi = (\alpha_i^{b,\phi}, \alpha_i^{a,\phi}, L_i^{b,\phi}, L_i^{a,\phi}, L_i^{b, \frac{1}{2}, \phi}, L_i^{a, \frac{1}{2}, \phi}, M_i^{b,\phi}, M_i^{a,\phi}),$$

is  $\mathbb{C}_0$ -valued, see below for more implicit restrictions. The quantities  $(\tau_i^\phi)_{i \geq 1}$  are the decision of the MM and  $\mathbf{c}_i^\phi$  his actions. The components of the vector  $\mathbf{c}_i^\phi$  have the same meanings as in Section 2. Given his own orders and the other participants' orders, the sequence of times at which orders are sent to the market is  $(\tau_i)_{i \geq 1}$  where  $\tau_0 := 0-$  and  $\tau_{i+1} = \min\{\tilde{\tau}_j > \tau_i, j \geq 1\} \wedge \min\{\tau_j^\phi > \tau_i, j \geq 1\}$ , see (4) and above.

We denote by  $\mathcal{E}^\phi$  the càdlàg process that jumps only at the times  $\tau_i^\phi$ s with jump size

$$\Delta \mathcal{E}_{\tau_i^\phi}^\phi = (\mathbf{c}_i^\phi, \sum_{j \geq 1} (\epsilon_j, \epsilon_j^b, \epsilon_j^a) \mathbf{1}_{\{\tau_i^\phi = \tau_j\}}), \quad \forall i \geq 1,$$

so that<sup>4</sup>

$$\mathcal{E} = \mathcal{E}^\phi + \tilde{\mathcal{E}},$$

<sup>4</sup>We keep in mind that  $\mathcal{E}$  also depends of  $\phi$  but dot not make this explicit for ease of notations.

### III. Optimal inventory management and order book modeling

from his point of view. As usual, we impose that  $\mathcal{E}^\phi$  is predictable for the (completed) filtration  $\mathbb{F}^\phi = (\mathcal{F}_t^\phi)_{t \geq 0}$  generated<sup>5</sup> by  $\mathcal{E}$ . We will also keep in mind the number of actions

$$J := \sum_{i \geq 1} \mathbf{1}_{\{\tau_i^\phi \leq \cdot\}},$$

from time 0 on, as it may induce a cost.

We now impose a minimum and maximum inventory size, denoted by  $(-I^*, I^*) \in (-\mathbb{N}) \times \mathbb{N}$ , and that

$$\#\{\tau_i^\phi \leq T, i \geq 1\} \leq k_\phi \wedge J_0 \text{ a.s., for some } k_\phi \in \mathbb{N},$$

for some  $J_0 \in \mathbb{N} \cup \{\infty\}$ . The constraint on the inventory is classic. The constraint on the number of operations can be justified by operational constraints. In the case  $J_0 = \infty$ , it just means that each control should be of essentially bounded activity, but the bound is not uniform on the set of controls and can be as large as needed. To be admissible, a control  $\phi$  should therefore be such that each  $c_i^\phi$  is  $\mathbf{C}(Z_{\tau_i^\phi-})$ -valued, where

$$Z := (P, Q, X) \text{ with } X := (G, I, N, B, J),$$

and, for any  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j)$ ,  $\mathbf{C}(z)$  is the collection of elements  $c := (a^b, a^a, \ell^b, \ell^a, \ell^{b, \frac{1}{2}}, \ell^{a, \frac{1}{2}}, m^b, m^a) \in \mathbf{C}_o \setminus \{0\}$  such that:

$$\begin{aligned} a^b &\leq \min\{i + I^* - b^a; q^b\}, \quad a^a \leq \min\{I^* - i - b^b; q^a\}, \\ \ell^b &\leq (I^* - i) \mathbf{1}_{\{n^b=0\}}, \quad \ell^a \leq (i + I^*) \mathbf{1}_{\{n^a=0\}}, \\ \ell^{b, \frac{1}{2}} &\leq (I^* - i) \mathbf{1}_{\{p^a - p^b = 2\delta\}} \mathbf{1}_{\{n^b=0\}}, \quad \ell^{a, \frac{1}{2}} \leq (i + I^*) \mathbf{1}_{\{p^a - p^b = 2\delta\}} \mathbf{1}_{\{n^a=0\}}, \\ m^b &\leq n^b, \quad m^a \leq n^a, \\ c &= 0 \text{ if } j = J_0. \end{aligned}$$

Note that the constraints on the first three lines correspond to the fact that we do not want to take a position that could lead to an inventory out of the limits  $-I^*$  and  $I^*$  if it was suddenly executed. The indicator functions correspond to additional constraints on the controls, imposed for numerical tractability: no new limit order can be sent on a side if one has not been executed or has not cancelled the position on the same side before, no limit order can be sent in the spread if it is not equal to two ticks. In this case, we write  $\phi \in \mathcal{C}(0, Z_{0-})$ . The dynamics of  $X$  are given by

$$X_{\tau_i} = \mathcal{F}_X(P_{\tau_{i-1}}, Q_{\tau_{i-1}}, X_{\tau_{i-1}}, \Delta \mathcal{E}_{\tau_i}^\phi) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi \neq 0\}} + \tilde{\mathcal{F}}_X(P_{\tau_{i-1}}, Q_{\tau_{i-1}}, X_{\tau_{i-1}}, \Delta \tilde{\mathcal{E}}_{\tau_i}) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi = 0\}}, \quad (8)$$

in which  $\mathcal{F}_X, \tilde{\mathcal{F}}_X : \mathbb{R}^{22} \mapsto \mathfrak{D}\mathbb{Z} \times \mathbb{N} \times \mathbb{N}^4 \times \mathbb{N}$ . More precisely, consider the map

$$\text{exe}(a, n, b) := \min\{(a - b)^+, n\}, \quad a, b, n \in \mathbb{N}.$$

<sup>5</sup>Note that this creates a dependence of the filtration on the control itself, which is similar to [16].

It represents the number of stocks set at a limit that are executed when an aggressive order of size  $a$  arrives, that the position in the queue is  $b$ , and the size of the posted block at this position is  $n$ . Then, having in mind the constraints encoded in  $\mathbf{C}(\cdot)$  above, see also (1), we can write

$$\mathcal{F}_X = (\mathcal{F}_G, \mathcal{F}_I, \mathcal{F}_{N^b}, \mathcal{F}_{N^a}, \mathcal{F}_{B^b}, \mathcal{F}_{B^a}, \mathcal{F}_J), \quad \tilde{\mathcal{F}}_X = (\tilde{\mathcal{F}}_G, \tilde{\mathcal{F}}_I, \tilde{\mathcal{F}}_{N^b}, \tilde{\mathcal{F}}_{N^a}, \tilde{\mathcal{F}}_{B^b}, \tilde{\mathcal{F}}_{B^a}, \tilde{\mathcal{F}}_J),$$

where

$$\begin{aligned} \mathcal{F}_G(p, q, x, \delta) &= g + (a^b - \text{exe}(a^b, n^b, b^b))p^b - (a^a - \text{exe}(a^a, n^a, b^a))p^a, \\ \mathcal{F}_I(p, q, x, \delta) &= i - (a^b - \text{exe}(a^b, n^b, b^b)) + (a^a - \text{exe}(a^a, n^a, b^a)), \\ \mathcal{F}_{N^{b/a}}(p, q, x, \delta) &= n^{b/a} + [\ell^{b/a} - n^{b/a}]^+ + \ell^{b/a, \frac{1}{2}} - m^{b/a} - \text{exe}(a^{b/a}, n^{b/a}, b^{b/a}), \\ \mathcal{F}_{B^{b/a}}(p, q, x, \delta) &= b^{b/a} + (q^{b/a} - b^{b/a})\mathbf{1}_{\{\ell^{b/a} \neq 0\}} - b^{b/a}\mathbf{1}_{\{m^{b/a} = n^{b/a}\}} - (b^{b/a} \wedge a^{b/a})\mathbf{1}_{\{a^{b/a} \neq 0\}} - b^{b/a}\mathbf{1}_{\{\ell^{b/a, \frac{1}{2}} \neq 0\}}, \\ \mathcal{F}_J(p, q, x, \delta) &= j + 1, \end{aligned}$$

and

$$\begin{aligned} \tilde{\mathcal{F}}_G(p, q, x, \delta) &= g - \text{exe}(a^b, n^b, b^b)p^b + \text{exe}(a^a, n^a, b^a)p^a, \\ \tilde{\mathcal{F}}_I(p, q, x, \delta) &= i + \text{exe}(a^b, n^b, b^b) - \text{exe}(a^a, n^a, b^a), \\ \tilde{\mathcal{F}}_{N^{b/a}}(p, q, x, \delta) &= n^{b/a} - \text{exe}(a^{b/a}, n^{b/a}, b^{b/a}), \\ \tilde{\mathcal{F}}_{B^{b/a}}(p, q, x, \delta) &= [b^{b/a} - a^{b/a}]^+ \mathbf{1}_{\{m^{b/a} = 0\}} + (b^{b/a} - [m^{b/a} - (q^{b/a} - b^{b/a} - n^{b/a})]^+)^+ \mathbf{1}_{\{m^{b/a} \neq 0\}}, \\ \tilde{\mathcal{F}}_J(p, q, x, \delta) &= 0, \end{aligned}$$

for any  $x = (g, i, n^b, n^a, b^b, b^a, j)$ ,  $\delta = (a^b, a^a, \ell^b, \ell^a, \ell^{b, \frac{1}{2}}, \ell^{a, \frac{1}{2}}, m^b, m^a, \varepsilon, \varepsilon^b, \varepsilon^a)$ ,  $p = (p^b, p^a)$  and  $q = (q^b, q^a)$ .

**Remark 3.1.** *It follows from (6) and the constraint that  $\mathcal{E}^\phi$  is predictable that the probability that  $\mathcal{E}^\phi$  and  $\tilde{\mathcal{E}}$  jump at the same time on  $[0, T]$  is zero. This justifies the formulation (8).*

For later use, note that it follows from (1) and (8) that

$$Z_{\tau_i} = \mathcal{F}(Z_{\tau_{i-1}}, \Delta \mathcal{E}_{\tau_i}^\phi) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi \neq 0\}} + \tilde{\mathcal{F}}(Z_{\tau_{i-1}}, \Delta \tilde{\mathcal{E}}_{\tau_i}) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi = 0\}}, \quad (9)$$

in which

$$\mathcal{F} = (\mathcal{F}_{P,Q}, \mathcal{F}_X) \quad \text{and} \quad \tilde{\mathcal{F}} = (\tilde{\mathcal{F}}_{P,Q}, \tilde{\mathcal{F}}_X).$$

### 3.2 The optimal control problem

The aim of the market maker is to maximize his expected utility

$$\mathbb{E}[U(Z_T)],$$

in which

$$U(z) := -\exp\left(-\eta\{g + i^+ p^b - i^- p^a - \kappa([i^+ - q^b]^+ + [i^- - q^a]^+) - \varrho j\}\right), \quad (10)$$

for any  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j)$ . In the above,  $\eta > 0$  is the absolute risk aversion parameter, and  $\kappa > 0$  is a penalty term taking into account that liquidating the current inventory may lead to a worse price than the one corresponding to the best bid or ask: the quantity  $i^+ p^b - i^- p^a$  corresponds to the liquidation value of the inventory if the bid and ask queues are big enough to absorb it, the expression starting from  $\kappa$  takes into account the number of shares that will not be liquidated at the best limit. The coefficient  $\varrho \geq 0$  penalizes the number of actions taken by the market maker. To define the corresponding value function, we now extend the definition of our state processes by writing

$$Z^{t,z,\phi} = (P^{t,z,\phi}, Q^{t,z,\phi}, X^{t,z,\phi}),$$

for the process satisfying (1)-(8) under the control  $\phi$  starting from the initial condition  $Z_{t-}^{t,z,\phi} = z \in D_Z$  where  $D_Z$  is the collection of elements  $(p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j) \in \mathbb{D}_{P,Q} \times \mathfrak{D}_Z \times \{-I^*, \dots, I^*\} \times \mathbb{N}^4 \times \{0, \dots, J_\circ\}$  such that

$$\begin{aligned} n^b + i &\leq I^*, \quad i - n^a \geq -I^*, \\ b^b + n^b &\leq q^b, \quad b^a + n^a \leq q^a. \end{aligned}$$

The corresponding set of admissible controls is  $\mathcal{C}(t, z)$ , and the filtration associated to  $\phi \in \mathcal{C}(t, z)$  is  $\mathbb{F}^{t,z,\phi}$ . We then set

$$v(t, z) := \sup_{\phi \in \mathcal{C}(t, z)} J(t, z; \phi), \quad \text{for } (t, z) \in [0, T] \times D_Z,$$

where

$$J(t, z; \phi) := \mathbb{E}[U(Z_T^{t,z,\phi})].$$

**Remark 3.2.** For later use, observe that

$$v(t, z) = e^{-\eta g} v(t, p^b, p^a, q^b, q^a, 0, i, n^b, n^a, b^b, b^a, j),$$

for all  $t \leq T$  and  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j) \in D_Z$ . Moreover, if  $J_\circ = \infty$ , we also have

$$v(t, z) = e^{-\eta(g-\varrho j)} \bar{v}(t, z) := e^{-\eta(g-\varrho j)} v(t, p^b, p^a, q^b, q^a, 0, i, n^b, n^a, b^b, b^a, 0).$$

**Remark 3.3.** Note that  $v$  is bounded from above by 0 by definition. On the other hand, for all  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j) \in D_Z$ ,

$$\begin{aligned} v(t, z) &\geq \min_{i \in [-I^*, I^*]} \mathbb{E}[U(P_T^{t,z,0}, 0, 0, g, i, 0, 0, 0, 0, j)] \\ &= e^{-\eta(g-\varrho j)} \min_{i \in [-I^*, I^*]} \mathbb{E}[U(P_T^{t,z,0}, 0, 0, 0, i, 0, 0, 0, 0, 0)], \end{aligned}$$

where  $P^{t,z,0}$  corresponds to the price dynamics when the MM does not act on the order book up to  $T$ . Moreover, it follows from (10) that

$$\mathbb{E}[U(P_T^{t,z,0}, 0, 0, 0, i, 0, 0, 0, 0, 0)] \geq -e^{\eta I^* |p^b|} \mathbb{E}[e^{\eta I^* (|P_T^{t,z,0,b} - p^b| + 2\delta + \kappa)}],$$

where

$$\sup_{p^b \in \mathfrak{d}\mathbb{Z}} \mathbb{E}[e^{\eta I^*(|P_T^{t,z,0,b} - p^b| + 2\mathfrak{d} + \kappa)}] < \infty,$$

by Remark 2.1 and the fact that the price can jump only by  $\mathfrak{d}$  when a market event occurs. Thus,  $v$  belongs to the class  $\mathfrak{L}_\infty^{exp}$  of functions  $\varphi$  such that  $\varphi/L$  is bounded, in which

$$L(p^b, g, j) := e^{-\eta(g - I^*|p^b| - \varrho j)}.$$

for any  $(p^b, g, j) \in \mathfrak{d}\mathbb{Z} \times \mathfrak{d}\mathbb{Z} \times \{0, \dots, J_\circ\}$ .

### 3.3 The dynamic programming equation

The derivation of the dynamic programming equation is standard, and is based on the dynamic programming principle. We state below the weak version of Bouchard and Touzi [37], and let  $v_*$  and  $v^*$  denote the lower- and upper-semicontinuous envelopes of  $v$ .

**Proposition 3.1.** *Fix  $(t, z) \in [0, T] \times D_Z$  and a family  $\{\theta^\phi, \phi \in \mathcal{C}(t, z)\}$  such that each  $\theta^\phi$  is a  $[t, T]$ -valued  $\mathbb{F}^{t,z,\phi}$ -stopping time and  $\|Z_{\theta^\phi}^{t,z,\phi}\|_{\mathbb{L}_\infty} < \infty$ . Then,*

$$\sup_{\phi \in \mathcal{C}(t,z)} \mathbb{E}\left[v_*(\theta^\phi, Z_{\theta^\phi}^{t,z,\phi})\right] \leq v(t, z) \leq \sup_{\phi \in \mathcal{C}(t,z)} \mathbb{E}\left[v^*(\theta^\phi, Z_{\theta^\phi}^{t,z,\phi})\right].$$

*Proof.* The right-hand side inequality follows from a conditioning argument, see [37]. The left-hand side is more delicate because the set of admissible controls depends on the initial data. However, it can be easily proved along the lines of Baradel et al. [16] when  $\{\theta^\phi, \phi \in \mathcal{C}(t, z)\}$  is  $[t, T] \cap (\mathbb{N} \cup \{t, T\})$ -valued. Then, the general case is obtained by approximating  $[t, T]$ -valued stopping times from the right (recall that  $Z$  is right-continuous).  $\square$

One can then derive the corresponding dynamic programming equation. For any  $z = (p, q, x) \in D_Z$ ,  $c \in \mathbf{C}(z)$ ,  $t \leq T$ , and a continuous and bounded function  $\varphi$ , we set

$$\mathcal{I}\varphi(t, z) := \int (\mathcal{K}^c \varphi(t, z) - \varphi(t, z)) d\beta(c|p, q) \text{ and } \mathcal{K}\varphi(t, z) := \sup_{c \in \mathbf{C}(z)} \mathcal{K}^c \varphi(t, z),$$

where we use the convention that  $\mathcal{K}^0 = \sup \emptyset = -\infty$ , and

$$\mathcal{K}^c \varphi(t, z) := \int \varphi(t, \mathcal{F}(z, c, \varepsilon, \varepsilon^b, \varepsilon^a)) d\lambda(\varepsilon, \varepsilon^b, \varepsilon^a | p, q, c),$$

recall (6), (7) and (9). The partial differential equation characterization of  $v$  is then at least formally given by

$$\begin{aligned} \min\{-\partial_t \varphi - \mathcal{I}\varphi, \varphi - \mathcal{K}\varphi\} &= 0, \text{ on } [0, T] \times D_Z, \\ \min\{\varphi - U, \varphi - \mathcal{K}\varphi\} &= 0, \text{ on } \{T\} \times D_Z. \end{aligned} \tag{11}$$

In order to ensure that the above is correct, we need two additional conditions.

**Assumption 3.1.** *For all upper-semicontinuous (resp. lower-semicontinuous)  $\varphi \in \mathfrak{L}_\infty^{exp}$ , the map  $(t, z) \in [0, T] \times D_Z \mapsto (\mathcal{I}, \mathcal{K})\varphi(t, z)$  is upper-semicontinuous (resp. lower-semicontinuous) and belongs to  $\mathfrak{L}_\infty^{exp}$ .*

**Assumption 3.2.** *There exists a Borel function  $\psi$  that is continuously differentiable in time and such that*

(i)  $0 \geq \partial_t \psi + \mathcal{I} \psi$ , on  $[0, T] \times D_Z$ ,

(i)  $\psi - \mathcal{K} \psi \geq \iota$ , on  $[0, T] \times D_Z$ , for some  $\iota > 0$ ,

(i)  $\psi \geq U$ , on  $\{T\} \times D_Z$ ,

(i)  $\liminf_{n \rightarrow \infty} (\psi/L)(t_n, z_n) = \infty$  if  $|z_n| \rightarrow \infty$  as  $n \rightarrow \infty$ , for all  $(t_n, z_n)_{n \geq 1} \subset [0, T] \times D_Z$ .

Then, one can actually prove that  $v$  is the unique solution of (11) in the class  $\mathfrak{L}_{\infty}^{\text{exp}}$  defined in Remark 3.3.

**Theorem 3.1.** *Let Assumption 3.1 hold. Then,  $v_*$  (resp.  $v^*$ ) is a viscosity supersolution (resp. subsolution) of (11). If moreover Assumption 3.2 holds, then  $v$  is continuous on  $[0, T] \times D_Z$  and is the unique viscosity solution of (11), in the class of (discontinuous) solutions in  $\mathfrak{L}_{\infty}^{\text{exp}}$ .*

*Proof.* In view of Proposition 3.1, the derivation of the viscosity super- and subsolution properties is very standard under Assumption 3.1, see e.g. [16, 37]. As for uniqueness, let us assume that  $v$  and  $w$  are respectively a super- and a subsolution. Let  $\psi$  be as in Assumption 3.2. Then,  $(v - w - \psi)(t_n, z_n)$  converges to  $-\infty$  if  $|z_n| \rightarrow \infty$  as  $n \rightarrow \infty$ , for any sequence  $(t_n, z_n)_{n \geq 1} \subset [0, T] \times D_Z$ , and showing that  $v \geq w$  on  $[0, T] \times D_Z$  can be done by, e.g., following the line of arguments of [16, Proposition 5.1]. Finally,  $v^*, v_* \in \mathfrak{L}_{\infty}^{\text{exp}}$  by Remark 3.3.  $\square$

**Remark 3.4.** *If  $J_{\circ} < \infty$  and the supports of  $\lambda(\cdot|p, q, c)$  and  $\gamma(\cdot|p, q)$  are bounded, uniformly in  $(p, q, c) \in (\partial\mathbb{Z})^2 \times \mathbb{N}^2 \times \mathbf{C}_{\circ}$ , then it is not difficult to see that the function defined by*

$$\psi(t, z) := e^{2\eta(1+I^*)|z|} e^{-r(j+t)}, \text{ for } z = (p, q, g, i, n, b, j) \in D_Z \text{ and } t \leq T,$$

*satisfies the requirements of Assumption 3.2, for  $r$  large enough. Verifying Assumption 3.2 in the case  $J_{\circ} = \infty$  seems much more difficult. On the other hand, the sequence of value functions associated to a sequence  $(J_{\circ}^n)_{n \geq 1}$  increases to the value function associated to  $J_{\circ} = \infty$  as  $J_{\circ}^n \rightarrow \infty$ , which provides a natural way to construct a convergent numerical scheme for the computation of  $v$  and the optimal control policy, see Sections 3.4 and 3.5 below. Standard arguments based on this approximation would also imply that  $v_* = v$  and that  $v$  is the smallest supersolution of (11), in the class of (discontinuous) solutions in  $\mathfrak{L}_{\infty}^{\text{exp}}$ .*

### 3.4 Dimension reduction, symmetries and numerical resolution

Before providing a converging numerical scheme for (11), let us first recall that the variables  $g$  (and  $j$  if  $J_{\circ} = \infty$ ) can be omitted, see Remark 3.2. If moreover, the transition kernels depend on prices only through the spread (which is a natural assumption at least on a rather short horizon time), then one more dimension can be eliminated.

**Assumption 3.3.** *The kernel  $(p, q, c) \in D_{BQ} \times \mathbf{C} \mapsto (\lambda(\cdot|p, q, c), \beta(\cdot|p, q))$  depends on  $p = (p^b, p^a)$  only through the value of the half-spread  $\delta p := (p^a - p^b)/2$ .*

Indeed, if Assumption 3.3 holds, then one easily checks that

$$\begin{aligned} e^{-\eta i p^\circ} \bar{v}(t, -\delta p, \delta p, q^b, q^a, 0, i, \cdot) &= \bar{v}(t, p^\circ - \delta p, p^\circ + \delta p, q^b, q^a, 0, i, \cdot) \\ &= \bar{v}(t, p^b, p^a, q^b, q^a, 0, i, \cdot), \end{aligned}$$

with  $p^\circ := (p^a + p^b)/2$ , so that

$$e^{\eta i (p^b + \delta p)} \bar{v}(t, p^b, p^b + 2\delta p, q^b, q^a, 0, i, \cdot).$$

does not depend on  $p^b$  but only on  $\delta p$ . The resolution of the equation can also be simplified by using potential symmetries, in the sense of the following assumption.

**Assumption 3.4.** For all  $(p, q) \in D_{RQ}$ ,  $c := (a^b, a^a, \ell^b, \ell^a, \ell^{b, \frac{1}{2}}, \ell^{a, \frac{1}{2}}, m^b, m^a) \in \mathbf{C}$  and all Borel sets  $O \subset \{0, 1\}$ ,  $O^b, O^a \subset \mathbb{N}$ ,

$$\int_{O \times O^b \times O^a} d\lambda(\varepsilon, \varepsilon^b, \varepsilon^a | p, q, c) = \int_{O \times O^a \times O^b} d\lambda(\varepsilon, \varepsilon^a, \varepsilon^b | \bar{p}, \bar{q}, \bar{c}),$$

where  $\bar{p} = (-p^a, -p^b)$ ,  $\bar{q} = (q^a, q^b)$  and  $\bar{c} = (a^a, a^b, \ell^a, \ell^b, \ell^{a, \frac{1}{2}}, \ell^{b, \frac{1}{2}}, m^a, m^b)$ . Moreover, for all Borel sets  $O = O_a^b \times O_a^a \times O_\ell^b \times O_\ell^a \times O_{\ell^{\frac{1}{2}}}^b \times O_{\ell^{\frac{1}{2}}}^a \times O_m^b \times O_m^a \subset \mathbb{N}^8$ ,

$$\int_O d\beta(c | p, q) = \int_{\bar{O}} d\beta(c | \bar{p}, \bar{q}),$$

where  $\bar{O} := O_a^a \times O_a^b \times O_\ell^a \times O_\ell^b \times O_{\ell^{\frac{1}{2}}}^a \times O_{\ell^{\frac{1}{2}}}^b \times O_m^a \times O_m^b$ .

The above assumption implies that the transition probabilities of the order book are symmetric at the bid and at the ask, whenever the configurations are. Then,  $v$  admits a symmetry which can be exploited to reduce the complexity of the numerical resolution of (11). Namely, under Assumption 3.4, we have

$$\bar{v}(t, p^b, p^a, q^b, q^a, 0, i, n^b, n^a, b^b, b^a, j) = \bar{v}(t, -p^a, -p^b, q^a, q^b, 0, -i, n^a, n^b, b^a, b^b, j).$$

Let us now turn to the definition of a numerical scheme for (11). We now make the additional assumption that the supports of  $\lambda$  and  $\beta$  are bounded (not that they are already discrete, by nature).

**Assumption 3.5.**  $J_\circ < \infty$  and there exists finite Borel sets  $O_1 \subset \mathbb{N}^8$  and  $O_2 \subset \mathbb{N}^3$  such that  $\beta(\cdot | p, q)$  is supported by  $O_1$  and  $\lambda(\cdot | p, q, c)$  is supported by  $O_2$ , for all  $(p, q, c) \in D_{RQ} \times \mathbf{C}$ .

Then, the operators  $\mathcal{F}$  and  $\mathcal{K}$  are explicit. Hence, the only required discretization is in time. For a time step  $T/n > 0$ , we define a time grid  $\pi_n := \{t_i^n, i \leq n\}$  where  $t_i^n = iT/n$  for  $i \leq n$ . We next consider the sequence of space domains  $D_Z^k := D_Z \cap [-k, k]$ <sup>11</sup> for  $k \geq 1$ , and we let  $v_n^k$  be the solution of

$$\begin{aligned} v_n^k(t_i^n, \cdot) &= \max \left\{ v_n^k(t_{i+1}^n, \cdot) + \frac{T}{n} \mathcal{F} v_n^k(t_{i+1}^n, \cdot), \max_{c \in \mathbf{C}(\cdot)} \mathcal{K}^{c, n} v_n^k(t_{i+1}^n, \cdot) \right\} = 0, \text{ on } D_Z^k, i \leq n-1, \\ v_n^k - U &= 0, \text{ on } (\{T\} \times D_Z^k) \cup (\pi_n \times (D_Z \setminus D_Z^k)), \end{aligned} \quad (12)$$



where

$$\mathcal{K}^{c,n} = \mathcal{K}^c + \frac{T}{n} \mathcal{F} \circ \mathcal{K}^c.$$

This fully explicit scheme is convergent.

**Proposition 3.2.** *Let Assumption 3.5 hold, then the sequence  $(v_n^k)_{k,n \geq 1}$  converges pointwise to  $v$  on  $[0, T) \times D_Z$  as  $k, n \rightarrow \infty$ .*

*Proof.* First note that  $(v_n^k/L)_{k,n \geq 1}$  is uniformly bounded, where  $L$  is defined in Remark 3.3. This follows from Assumption 3.5 and a simple induction argument, compare with Remark 3.3. Then, standard stability results, see e.g. [17] and [15, Section 3.2], imply that the relaxed upper-limit  $v^\infty$  and lower-limit of  $v_\infty$  of  $(v_n^k)_{k,n \geq 1}$  are respectively sub- and supersolution of (11) and belong to the class  $\mathcal{L}_\infty^{\text{exp}}$ . The comparison result mentioned in the proof of Theorem 3.1, see Remark 3.4, thus implies that  $v_\infty \geq v \geq v^\infty$  while  $v_\infty \leq v^\infty$  by definition.  $\square$

### 3.5 Approximate optimal controls

In the following, we estimate the optimal control in a classical way. For each  $i < n$ , we choose a measurable map  $\hat{c}_n^k(t_i^n, \cdot)$  such that

$$\begin{aligned} \hat{c}_n^k(t_i^n, \cdot) &\in \arg \max\{\mathcal{K}^{c,n} v_n^k(t_{i+1}^n, \cdot), c \in \mathbf{C}(\cdot)\}, \text{ on } D_Z^k, \\ \hat{c}_n^k(t_i^n, \cdot) &= 0, \text{ on } D_Z \setminus D_Z^k, \end{aligned}$$

and define the sequence of stopping times

$$\hat{t}_{j+1}^{n,k} := \min\{t_i^n : i \geq 0, t_i^n > \hat{t}_j^{n,k}, (v_n^k - \mathcal{K}^{\hat{c}_n^k} v_n^k(t_{i+1}^n, \cdot))(t_i^n, \hat{Z}_{t_i^n}^{n,k}) = 0\}, j \geq 0,$$

with  $\hat{t}_0^{n,k} := 0-$ , and in which  $\hat{Z}^{n,k} = (\hat{P}^{n,k}, \hat{Q}^{n,k}, \hat{X}^{n,k})$  is defined as in (9) for the initial condition  $Z_{0-}$  at 0 and the control associated to  $\hat{\phi}_n^k := (\hat{t}_i^{n,k}, \hat{c}_n^k(\hat{t}_i^{n,k}, \hat{Z}_{\hat{t}_i^{n,k}}^{n,k}))_{i \geq 1}$  in a Markovian way. This provides a sequence of controls that is asymptotically optimal.

**Proposition 3.3.** *Let the conditions of Proposition 3.2 hold. Then,*

$$\lim_{n,k \rightarrow \infty} \mathbb{E}[U(\hat{Z}_T^{n,k})] = v(0, Z_{0-}).$$

*Proof.* Let  $\gamma(p, q) := \int d\beta(c|p, q)$  and recall that  $\gamma$  is uniformly bounded by assumption, as well as the sequence  $(v_n^k)_{k,n \geq 1}$ . Then, it follows from Remark 2.1 that

$$\begin{aligned} v_n^k(t_{i+1}^n, \hat{Z}_{t_i^n}^{n,k}) + \frac{T}{n} \mathcal{F} v_n^k(t_{i+1}^n, \hat{Z}_{t_i^n}^{n,k}) &= v_n^k(t_{i+1}^n, \hat{Z}_{t_i^n}^{n,k}) \left(1 - \frac{T}{n} \gamma(\hat{P}_{t_i^n}^{n,k}, \hat{Q}_{t_i^n}^{n,k})\right) \\ &\quad + \frac{T}{n} \int \mathcal{K}^c v_n^k(t_{i+1}^n, \hat{Z}_{t_i^n}^{n,k}) d\beta(c|\hat{P}_{t_i^n}^{n,k}, \hat{Q}_{t_i^n}^{n,k}) \\ &= \mathbb{E}[v_n^k(t_{i+1}^n, \tilde{\mathcal{F}}(\hat{Z}_{t_i^n}^{n,k}, \Delta \tilde{\mathcal{E}}_{t_{i+1}^n}^k)) | \mathcal{F}_{t_i^n}^{\hat{\phi}_n^k}] + o(n^{-1}), \end{aligned}$$

and, similarly,

$$\mathcal{K}_{\hat{c}_n^k(\hat{\tau}_i^{n,k}, \hat{Z}_{\hat{\tau}_i^{n,k}-}^{n,k}), n} v_n^k(\hat{\tau}_i^{n,k} + \frac{T}{n}, \hat{Z}_{\hat{\tau}_i^{n,k}-}^{n,k}) = \mathbb{E}[v_n^k(\hat{\tau}_i^{n,k} + \frac{T}{n}, \tilde{\mathcal{F}}(\hat{Z}_{\hat{\tau}_i^{n,k} + \frac{T}{n}-}^{n,k}, \Delta \tilde{\mathcal{E}}_{\hat{\tau}_i^{n,k} + \frac{T}{n}})) | \mathcal{F}_{\hat{\tau}_i^{n,k}-}^{\hat{\phi}_n^k}] + o(n^{-1}),$$

with the convention that  $\tilde{\mathcal{F}}(\cdot, 0)$  is the identity. Let  $\theta_k^n$  be the first time when  $\hat{Z}^{n,k}$  exists the domain  $D_Z^k$ . In view of (12), Assumption 3.5 (in particular that  $J_0 < \infty$ ) and Remark 2.1 (see the arguments at the end of Remark 3.3), an induction implies that

$$v_n^k(0, Z_{0-}) = \mathbb{E}[U(\hat{Z}_{\theta_k^n}^{n,k})] + o_n(1) = \mathbb{E}[U(\hat{Z}_T^{n,k})] + o_n(1) + o_k(1),$$

in which  $o_n(1)$  and  $o_k(1)$  go to 0 as  $n \rightarrow \infty$  and  $k \rightarrow \infty$ . It remains to appeal to Proposition 3.2.  $\square$

### 3.6 Numerical experiments

We now turn to a numerical experiment. We compute an approximation of the optimal control as described in Section 3.5, using the simplifications detailed in Section 3.4.

Let us first describe a realistic prior distribution for the evolution of  $\tilde{\mathcal{E}}$ . The coefficients we use are inspired from the behaviour of the stock Société Générale (GLE FP)<sup>6</sup> and from [30]. As for the prior on the dynamics of the market. We simply consider that both market and limit orders arrive according to a Poisson process. Both limit and aggressive orders arrive with an intensity of 0.6 per second. When a limit order arrives, it is assigned to the bid or the ask with probability 1/2. When a market order arrives, we assign it to the bid or the ask according to the statistics described for big caps in [30, Chart 8 p.10]. Namely, if

$$\text{Imb}_{\tilde{\tau}_i} := \frac{Q_{\tilde{\tau}_i}^b - Q_{\tilde{\tau}_i}^a}{Q_{\tilde{\tau}_i}^b + Q_{\tilde{\tau}_i}^a},$$

is the order book imbalance at the time  $\tilde{\tau}_i$  at which the market order arrives, then it arrives at the ask with probability  $0.5 + 0.35 \times \text{Imb}_{\tilde{\tau}_i}$ . To describe the size of the orders and of the inventory, we take 1/2 of the ATS (mediAn<sup>7</sup> Trade Size) as the unit.

The size of the market orders is also assumed to be dependent on the order book imbalance. We use [30, Chart 16 p.10] to estimate that the size of the trade arriving at the ask (if it arrives at the ask), represents a percentage of the queue (that we round to an integer number, by above). Namely, we set  $\hat{f}_{\tilde{\tau}_i}^a := 0.7 + 0.3 \times \text{Imb}_{\tilde{\tau}_i}$ , and assign a (conditional) probability of 60% that the order is of size  $\text{round}[\hat{f}_{\tilde{\tau}_i}^a \times Q_{\tilde{\tau}_i}^a]$  and the same probability that the executed volume deviates from the latter by one unit (with equal probability to be by one more and one less unit - if quantities are negative or bigger than the size of the queue, we obviously set them to 0 or to the size of the queue).

<sup>6</sup>We thank Kepler-Cheuvreux for providing us these data.

<sup>7</sup>Following [30, Page 23].

We use a simpler modelling approach for the limit orders. With 55% probability a limit order (if it arrives) is of size of 2 (recall that the unit is 1/2 of an ATS). It is of size 3 with probability 10% and of size 1 with probability 35%. Again, it is based on [30].

When a queue is depleted, the probability of a price move is set to  $\mathbb{P}[\epsilon_i = 1 | \mathcal{F}_{\bar{\tau}_i}] = 75\%$ . If the bid is depleted but the bid price does not go down, the size of the new bid queue is set to 2 units with probability 60%, 1 unit with probability 25% and 3 units with probability 15%, otherwise it is set to 10 units with probability 60%, 5 units with probability 25% and 12 units with probability 15%. The same applies to the ask price if this is the ask queue that is depleted. If both ask and bid prices move, recall (3), we take the same distribution for both queues and consider them as being (conditionally) independent. The distribution corresponds to the one of a price move, as described above.

When the spread is equal to two ticks, the next limit order arrives in the spread with probability 90%. This models the fact that a spread of two ticks is not common, see e.g. [55]. To be consistent with the probability of arrival of market orders at the bid or at the ask, we assume that a new bid limit is created in the spread with probability  $0.5 + 0.35 \times \text{Imb}_{\bar{\tau}_i}$  (otherwise, this is a new ask limit). The size of this new limit is 2 with 55%, of size 3 with probability 10% and of size 1 with probability 35%. This corresponds to the behaviour of the stock Société Générale (GLE FP). We do not change the dynamics of aggressive orders with the size of the spread.

Let us now describe the other parameters of the Market Maker's optimal control problem. We set the bid price to 10 at time  $0-$ , the spread is one tick, and the tick equal to 0.01, recall from Section 3.4 that only the spread size matters (because we have here the required symmetry) and observe that the latter can be rescaled together with the level of risk aversion, which is here fixed to  $\eta = 1$ . We take  $\kappa = 0.02$  and  $\rho = 10^{-20}$ . The horizon time is  $T = 59$  seconds, and the time step is 1 second. We keep a small horizon time for a better visibility of the evolution of the order book. We do not consider cancellations from the rest of the market for simplicity. Moreover, in order to reduce the computation time, we add the additional constraint that the MM cannot cancel only part of position in a queue, he can only cancel the whole position. We also fix a maximal queue size of 12 and fix the maximal absolute value of the inventory to  $I^* = 7$ . The size of the limit and market orders sent by the MM are constrained to be less than 3. This corresponds to adding an additional constraint in the definition of  $\mathcal{C}(\cdot)$  which does not change the above analysis.

In Figure III.2, we provide a simulated path of the optimal strategy, starting from a symmetric configuration of the order book, with queue lengths equal to 6. In this simulation, the MM always play before the other (random) players.<sup>8</sup> The top left graphic describes the control played by the MM. Triangles pointing outward (with respect to the zero line) correspond to limit orders, the number of triangles giving their size. Arrows with triangles pointing inward are cancellations, again the number of triangles gives the size. Aggressive orders are symbol-

---

<sup>8</sup>This is just a convention, since the transition probabilities do not depend on time.

ized by lines with squares, while limit orders within the spread correspond to the lines with dots. The top right graphic gives the state of the order book just after the MM action, and before the nature (i.e. the other players) plays. The size of the lines gives the total length of each queue, while the dots symbolize the position of the MM in the queue. The middle left graphic describes the state of the order book after the nature decision. The bottom left graphic gives the value of the portfolio of the MM if he had to liquidate (by sending aggressive orders) his stock holding at the best bid/ask price. The final value gives the "true" liquidation value of the book. In the case that the final inventory cannot be liquidated at the best bid/ask price, we liquidate the remaining part at the best bid/ask price minus/plus one tick. The middle right graphic is the inventory just before he plays and the bottom right one is the bid and ask prices just after he plays. Figure III.1 provides the distribution of the gain made by the MM, it uses  $10^5$  simulated paths.

His strategy can seem difficult to interpret at first sight. But, we have to remember that he believes that the imbalance plays an important role in the book order dynamics, and that he not only should take care of it but that he can actually use it: some limit orders are sent not to be executed but to influence the evolution of the price in a favorable direction. Also note that he should avoid a price decrease/increase when his inventory is positive/negative. Having this in mind, it is not surprising that he can be sometimes at the limit of a price manipulation strategy, if he is big enough with respect to the market, which is the case in these simulations. Finally, we have to keep in mind that his strategy is constrained. He sometimes cancels positions to be free to react more quickly to a more favorable market configuration later on. Not surprisingly, the MM first sends limit orders of equal sizes on both sides of the order book. Nothing happens until time  $t = 9$ . At this time his inventory is 2 and his position at the ask is still far from the beginning of the queue. In order to avoid increasing again his inventory, he cancels his remaining position of one unit at the bid. He puts a new position at the bid at time 10 after the queue has slightly increased, because of an exogenous limit order. Unfortunately for him, this new position is immediately executed, and his inventory jumps to 5. After the queue is regenerated, he immediately puts a limit order of size 2 at the bid. His reasoning is the following: 1. this position has little chance to be executed immediately, he does not take a risk of again increasing his inventory; 2. by doing this he increases the imbalance and therefore the probability of being executed at the ask so has to gain the spread and reduce his inventory. This strategy is successful since immediately 2 units of his positions at the ask are executed. The imbalance is still good even if he cancels his last unit at the bid, so as to be free to play the control he wants once his last unit at the ask will be executed. In fact, it does not work and he decides to refresh his global position at time 16 by cancelling his final unit at the ask and putting limit orders again at time 17 in a symmetric way. This is a limit spread order at the bid. This makes sense since he has to avoid the stock price to go down, because he has a positive inventory. The fact that just after he alternates between putting and cancelling limit orders at the ask is probably a numerical artifact: on the one hand, he wants to have a limit position at the ask because his inventory is large, on the other hand he does not want to increase the imbalance to avoid having a too big probability of being executed at the bid and of seeing the price go down. Similarly, he wants to keep a position at the bid to

### III. Optimal inventory management and order book modeling

---

avoid a downward price move, while he really does not want to be executed on this side. The aggressive order at the bid at time 19 is just a partial cancellation, that is completed at time 20. He just does not want to be executed. By doing so, he unfortunately causes a downward jump of the bid, which is not good for him. Just after he sends a limit spread order at the bid to fight against this downward pressure, and then cancels it once the bid is back at a distance one tick of the ask. This strategy is successful. The rest of his orders can be interpreted in a similar manner. Just note that he starts to be aggressive at the bid side at time 49 because his inventory is too big and the maturity starts to be quite close, in particular the first market order of size 3 compensates the execution of a position of size 3 at the bid just before (so that the graphic of the inventory does not move, although the inventory temporary jumps to the upper limit 7).

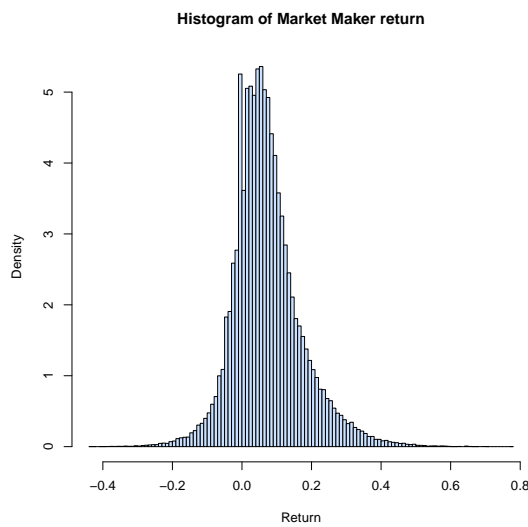


Figure III.1 – Density estimation of the gain made by the Market Maker.

### 3. Market maker's optimal control problem

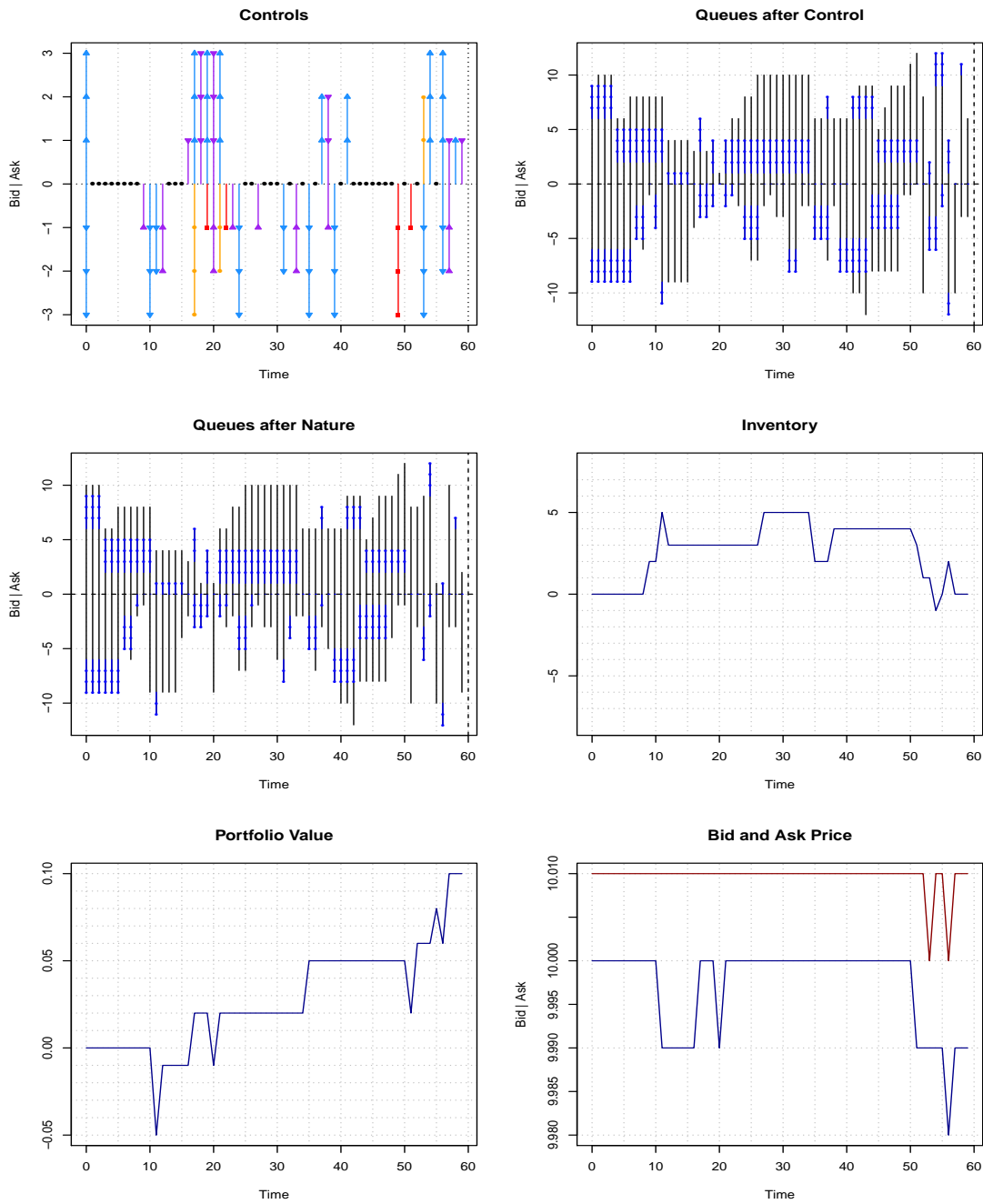


Figure III.2 - Simulated path of the optimal strategy of the Market Maker.

## 4 High frequency pair trading problem

We now turn to the DHFT strategy. We consider a pair trading strategy where the trader invests in the difference of two highly correlated assets. Here, we choose the futures price of the stock to be the second asset. The order book defined in Section 2 represents the dynamics of the stock. Whenever the trader buys/sells  $n$  units of stocks (being by an aggressive order or by the execution of a limit order), he sells/buys immediately  $n$  units of the futures. His inventory should be fully liquidated at  $T$ .

### 4.1 The optimal control problem

We assume that the reference price  $F$  of the futures is given by

$$F = \frac{P^b + P^a}{2} + S, \quad (13)$$

where  $S$  is a mean-reverting process

$$S = S_0 + \int_0^\cdot \rho(\hat{s} - S_t) dt + \int_0^\cdot \sigma(S_t) dW_t. \quad (14)$$

Here,  $\rho$  is the strength of mean reversion,  $\hat{s}$  is the average of mean reversion and  $\sigma : \mathbb{R} \mapsto \mathbb{R}$  is a Lipschitz bounded function representing the volatility of the process. In the following, we shall assume that

$$\text{the support of } \sigma \text{ is bounded.} \quad (15)$$

**Remark 4.1.** *The above implies in particular that  $S$  lies in a certain compact set  $D_S$  as soon as  $S_0$  does. This could clearly be relaxed to the price of a finer analysis.*

The strategy of the DHFT is described by the same quantities used for the MM case, see Section 3. The only difference lies in the fact that he constantly holds a number equal to  $-I$  units of the futures  $F$ . We assume that buying/selling the futures leads to the payment of a proportional cost  $\kappa \geq 0$ . Then, the dynamics of the corresponding state process  $X$  are given by

$$X_{\tau_i} = \mathcal{T}_X(S_{\tau_i}, P_{\tau_{i-1}}, Q_{\tau_{i-1}}, X_{\tau_{i-1}}, \Delta \mathcal{E}_{\tau_i}^\phi) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi \neq 0\}} + \tilde{\mathcal{T}}_X(S_{\tau_i}, P_{\tau_{i-1}}, Q_{\tau_{i-1}}, X_{\tau_{i-1}}, \Delta \tilde{\mathcal{E}}_{\tau_i}) \mathbf{1}_{\{\Delta \mathcal{E}_{\tau_i}^\phi = 0\}}, \quad (16)$$

with  $\mathcal{T}_X$  and  $\tilde{\mathcal{T}}_X$  now defined such that

$$\begin{aligned} \mathcal{T}_G(s, p, q, x, \delta) &= g + (a^b - \text{exe}(a^b, n^b, b^b)) \Delta_-^b - (a^a - \text{exe}(a^a, n^a, b^a)) \Delta_+^a, \\ \mathcal{T}_I(s, p, q, x, \delta) &= i - (a^b - \text{exe}(a^b, n^b, b^b)) + (a^a - \text{exe}(a^a, n^a, b^a)), \\ \mathcal{T}_{N^{b/a}}(s, p, q, x, \delta) &= n^{b/a} + [\ell^{b/a} - n^{b/a}]^+ + \ell^{b/a, \frac{1}{2}} - m^{b/a} - \text{exe}(a^{b/a}, n^{b/a}, b^{b/a}), \\ \mathcal{T}_{B^{b/a}}(p, q, x, \delta) &= b^{b/a} + (q^{b/a} - b^{b/a}) \mathbf{1}_{\{\ell^{b/a} \neq 0\}} - b^{b/a} \mathbf{1}_{\{m^{b/a} = n^{b/a}\}} - (b^{b/a} \wedge a^{b/a}) \mathbf{1}_{\{a^{b/a} \neq 0\}} - b^{b/a} \mathbf{1}_{\{\ell^{b/a} \neq 0\}}, \\ \mathcal{T}_J(s, p, q, x, \delta) &= j + 1, \end{aligned}$$

and

$$\begin{aligned}
 \tilde{\mathcal{T}}_G(s, p, q, x, \delta) &= g - \text{exe}(a^b, n^b, b^b)\Delta_+^b + \text{exe}(a^a, n^a, b^a)\Delta_-^a, \\
 \tilde{\mathcal{T}}_I(s, p, q, x, \delta) &= i + \text{exe}(a^b, n^b, b^b) - \text{exe}(a^a, n^a, b^a), \\
 \tilde{\mathcal{T}}_{N^{b/a}}(s, p, q, x, \delta) &= n^{b/a} - \text{exe}(a^{b/a}, n^{b/a}, b^{b/a}), \\
 \tilde{\mathcal{T}}_{B^{b/a}}(p, q, x, \delta) &= [b^{b/a} - a^{b/a}]^+ \mathbf{1}_{\{m^{b/a}=0\}} + (b^{b/a} - [m^{b/a} - (q^{b/a} - b^{b/a} - n^{b/a})]^+)^+ \mathbf{1}_{\{m^{b/a} \neq 0\}}, \\
 \tilde{\mathcal{T}}_J(s, p, q, x, \delta) &= 0,
 \end{aligned}$$

in which

$$\Delta_{\pm}^b := p^b - \frac{p^b + p^a}{2} - s \pm \kappa, \quad \Delta_{\pm}^a := p^a - \frac{p^b + p^a}{2} - s \pm \kappa.$$

In the above, we use the notations  $x = (g, i, n^b, n^a, b^b, b^a, j)$ ,  $\delta = (a^b, a^a, \ell^b, \ell^a, \ell^{b, \frac{1}{2}}, \ell^{a, \frac{1}{2}}, m^b, m^a, \varepsilon, \varepsilon^b, \varepsilon^a)$ ,  $p = (p^b, p^a)$  and  $q = (q^b, q^a)$ . The set of admissible controls  $\mathcal{C}(0, S_0, Z_{0-})$  is defined as in Section 3 but with respect to the (completed) filtration  $\mathbb{F}^\phi = (\mathcal{F}_t^\phi)_{t \geq 0}$  generated by  $(S, \mathcal{E})$ .

We also assume that DHFT has an exponential type utility function, with risk aversion parameter  $\eta > 0$ . Then, she wants to maximize over  $\phi \in \mathcal{C}(0, S_0, Z_{0-})$  the expected utility

$$\mathbb{E}[U(S_T, Z_T^\phi)],$$

where

$$U(s, z) := -\exp\left(-\eta\{g + i^+ \Delta_-^b - i^- \Delta_+^a - \kappa([i^+ - q^b]^+ + [i^- - q^a]^+)\} - \rho j\right), \quad (17)$$

for any  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j)$ , and where  $\Delta_{\pm}^b$  and  $\Delta_{\pm}^a$  are defined as above.

As in Section 3, we next extend the definition of our state processes by writing

$$(S^{t,s}, Z^{t,s,z,\phi}) = (S^{t,s}, P^{t,s,z,\phi}, Q^{t,s,z,\phi}, X^{t,s,z,\phi}),$$

for the process satisfying (1)-(8)-(14) under the control  $\phi$  and the initial condition  $(S_t^{t,s}, Z_{t-}^{t,s,z,\phi}) = (s, z) \in D_S \times D_Z$ . The corresponding set of admissible controls is  $\mathcal{C}(t, s, z)$ , and the filtration associated to  $\phi \in \mathcal{C}(t, s, z)$  is  $\mathbb{F}^{t,s,z,\phi} = (\mathcal{F}_s^{t,s,z,\phi})_{s \in [t, T]}$ . We finally define the value function

$$v(t, s, z) := \sup_{\phi \in \mathcal{C}(t,s,z)} J(t, s, z; \phi) \text{ for } (t, s, z) \in [0, T] \times D_S \times D_Z,$$

where

$$J(t, s, z; \phi) := \mathbb{E}[U(S_T^{t,s}, Z_T^{t,s,z,\phi})].$$

We close this section with Remarks that are the counterparts of Remarks 3.2 and 3.3.

**Remark 4.2.** *For later use, observe that*

$$v(t, s, z) = e^{-\eta(g-\rho j)} \bar{v}(t, z) := e^{-\eta(g-\rho j)} v(t, s, p^b, p^a, q^b, q^a, 0, i, n^b, n^a, b^b, b^a, 0),$$

for all  $t \leq T$ ,  $s \in D_S$  and  $z = (p^b, p^a, q^b, q^a, g, i, n^b, n^a, b^b, b^a, j) \in D_Z$ .

**Remark 4.3.** *Note that  $v$  is bounded from above by 0 by definition. It also follows from (15) that  $S^{t,s}$  takes values in the compact set  $D_S$  so that  $v \in \mathbb{L}_{\infty}^{exp}$  by the same reasoning as in Remark 3.3.*



## 4.2 The dynamic programming equation

As in Section 3.3, we first provide a dynamic programming principle. Again, we let  $v_*$  and  $v^*$  denote the lower- and upper-semicontinuous envelopes of  $v$ .

**Proposition 4.1.** *Fix  $(t, s, z) \in [0, T] \times D_S \times D_Z$  and a family  $\{\theta^\phi, \phi \in \mathcal{C}(t, s, z)\}$  such that each  $\theta^\phi$  is a  $[t, T]$ -valued  $\mathbb{F}^{t, s, z, \phi}$ -stopping time and  $\|(S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi})\|_\infty < \infty$ . Then,*

$$\sup_{\phi \in \mathcal{C}(t, x, q)} \mathbb{E} \left[ v_*(\theta^\phi, S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi}) \right] \leq v(t, s, z) \leq \sup_{\phi \in \mathcal{C}(t, x, q)} \mathbb{E} \left[ v^*(\theta^\phi, S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi}) \right].$$

*Proof.* Let  $\mathcal{C}_k(t, x, z)$  be the set of controls  $\phi$  satisfying the additional constraint  $\#\{\tau_i^\phi, i \geq 1\} \leq k$  a.s., and let  $v_k$  be the corresponding value function, for  $k \geq 1$ . Then, it is not difficult to see that  $v_k$  is continuous, and the arguments of [37] imply that

$$\begin{aligned} \sup_{\phi \in \mathcal{C}_k(t, x, z)} \mathbb{E} \left[ v_k(\theta^\phi, S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi}) \right] &\leq v_k(t, s, z) \\ &\leq \sup_{\phi \in \mathcal{C}_k(t, x, z)} \mathbb{E} \left[ v_k(\theta^\phi, S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi}) \right] \\ &\leq \sup_{\phi \in \mathcal{C}(t, x, z)} \mathbb{E} \left[ v(\theta^\phi, S_{\theta^\phi}^{t, s}, Z_{\theta^\phi}^{t, s, z, \phi}) \right]. \end{aligned}$$

Since by definition  $v = \lim_{k \rightarrow \infty} \uparrow v_k$  and  $\mathcal{C}(t, x, z) = \cup_{k \geq 1} \mathcal{C}_k(t, x, z)$ , sending  $k \rightarrow \infty$  in the above leads to the required result, recall Remark 4.3.  $\square$

The partial differential equation associated to  $v$  is then at least formally given by

$$\begin{aligned} \min \{-\mathcal{L}\varphi - \mathcal{I}\varphi, \varphi - \mathcal{K}\varphi\} &= 0, \text{ on } [0, T] \times D_S \times D_Z, \\ \min \{\varphi - U, \varphi - \mathcal{K}\varphi\} &= 0, \text{ on } \{T\} \times D_S \times D_Z, \end{aligned} \tag{18}$$

in which  $\mathcal{L}$  is the Dynkin operator associated to (14):

$$\mathcal{L}\varphi = \partial_t \varphi + \rho(\hat{s} - s)\partial_s \varphi + \frac{1}{2}\sigma^2 \partial_{ss}^2 \varphi.$$

To fully characterize the value function, we need the additional assumption, similar to Assumption 3.2.

**Assumption 4.1.** *There exists a Borel function  $\psi \in C^{1,2}([0, T] \times D_S \times D_Z)$  such that*

- (i)  $0 \geq \mathcal{L}\psi + \mathcal{I}\psi$ , on  $[0, T] \times D_S \times D_Z$ ,
- (i)  $\psi - \mathcal{K}\psi \geq \iota$ , on  $[0, T] \times D_S \times D_Z$  for some  $\iota > 0$ ,
- (i)  $\psi \geq U$ , on  $\{T\} \times D_S \times D_Z$ ,
- (i)  $\liminf_{n \rightarrow \infty} (\psi/L)(t_n, s_n, z_n) = \infty$  if  $|z_n| \rightarrow \infty$  as  $n \rightarrow \infty$ , for all  $(t_n, s_n, z_n)_{n \geq 1} \subset [0, T] \times D_S \times D_Z$ .

**Theorem 4.1.** *Let Assumption 3.1 hold. Then,  $v_*$  (resp.  $v^*$ ) is a viscosity supersolution (resp. subsolution) of (18). If moreover Assumption 4.1 holds, then  $v$  is continuous on  $[0, T] \times D_Z$  and is the unique viscosity solution of (18), in the class  $\mathbb{L}_\infty^{exp}$ .*

*Proof.* In view of Proposition 4.1, the derivation of the viscosity super- and subsolution properties is very standard under Assumption 3.1, see e.g. [37]. As for uniqueness, this follows from a comparison principle that can be proved in the class  $\mathcal{L}_\infty^{\text{exp}}$  by arguing as in the proof of Theorem 3.1.  $\square$

As for the MM problem, Assumption 4.1 is easily checked when  $J_\circ < \infty$ .

**Remark 4.4.** *If  $J_\circ < \infty$  and the supports of  $\lambda(\cdot|p, q, c)$  and  $\gamma(\cdot|p, q)$  are bounded, uniformly in  $(p, q, c) \in (\partial\mathbb{Z})^2 \times \mathbb{N}^2 \times \mathbf{C}_\circ$ , then the function  $\psi$  defined in Remark 4.4 also satisfies the requirements of Assumption 4.1, for  $r$  large enough.*

### 4.3 Dimension reduction, symmetries and numerical resolution

As in Section 3.4, one can use specificities of the value function to reduce the complexity of the resolution of (18). First, the variable  $(g, j)$  can be omitted, see Remark 4.2. Moreover, if Assumption 3.3 holds, then one easily checks that

$$\bar{v}(t, s, p^b, p^b + 2\delta p, q^b, q^a, 0, i, n^b, n^a, b^b, b^a, 0),$$

does not depend on  $p^b$ . The difference with the relation obtained in Section 3.4 is due to (13) and the fact that the HFT always holds a symmetric position in the stock and the futures (he is protected against evolutions of the mid-price). The other symmetry relations described in Section 3.4 do not hold because of the dependence on the process  $S$ .

Let us now turn to the definition of a numerical scheme for (18). Recall that, under Assumption 3.5, the operators  $\mathcal{S}$  and  $\mathcal{K}$  are explicit. Hence, the only required discretization is in time and in the variable  $s$ . We shall consider separately the diffusion part and the obstacle part of the PDE. More precisely, we fix a time and a space grid  $\pi_t^n := \{t_i^n, i \leq n_t\}$  and  $\pi_s^n := \{s_i^n, i \leq n_s\}$  where  $t_i^n = iT/n_t$  for  $i \leq n_t$  and  $s_i^n = \underline{s} + i(\bar{s} - \underline{s})/n_s$ ,  $i \leq n_s$ . Here,  $\underline{s}$  and  $\bar{s}$  are such that  $D_S = [\underline{s}, \bar{s}]$ , recall Remark 4.1, and  $n := (n_t, n_s) \in \mathbb{N}^2$ . We next define the sequence of space domains  $D_Z^k := D_Z \cap [-k, k]^{11}$  for  $k \geq 1$ , and we let  $v_n^k$  be defined by

$$\begin{aligned} v_n^k &= \max \left\{ \check{v}_n^k, \mathcal{K} \check{v}_n^k \right\} = 0, \text{ on } \pi_t^n \times \pi_s^n \times D_Z^k, \\ v_n^k - U &= 0, \text{ on } (\{T\} \times \pi_s^n \times D_Z^k) \cup (\pi_t^n \times \pi_s^n \times (D_Z \setminus D_Z^k)). \end{aligned} \quad (19)$$

Here, for  $i \leq n_t - 1$  and  $i' \leq n_s$ ,

$$\check{v}_n^k(t_i^n, s_{i'}^n, \cdot) = \mathbb{E}[v_n^k(t_{i+1}^n, p_n(S_{t_{i+1}^n}^{t_i^n, s_{i'}^n}), \cdot)] + \frac{T}{n} \mathcal{S} v_n^k(t_{i+1}^n, s_{i'}^n, \cdot), \text{ on } D_Z^k, \quad (20)$$

where  $p_n$  is the left-hand side projection operator on  $\pi_s^n$ . Note that the above numerical scheme is not fully explicit as it requires to compute conditional expectations. This can however be easily performed either by a finite difference scheme or by Monte-Carlo techniques in a very classical way. Note in particular that the randomness in these conditional expectations only comes from a one dimensional factor.

**Proposition 4.2.** *Let Assumptions 4.1 and 3.5 hold, then the sequence  $(v_n^k)_{n \geq 1}$  converges pointwise to  $v$  as  $n_t, n_s, k \rightarrow \infty$ .*

*Proof.* Note that Assumption 4.1 ensures that comparison holds for (18) in the class  $\mathbf{L}_\infty^{\text{exp}}$ , see e.g., [16, Proposition 5.1]. Thus, as for Proposition 3.2, the result is an easy consequence of the stability result of [17].  $\square$

#### 4.4 Approximate optimal controls

The optimal control can be numerically estimated as in Section 3.4. We first extend  $(v_n^k, \check{v}_n^k)$  to  $\pi_t^n \times D_S \times D_Z$  by setting  $(v_n^k, \check{v}_n^k)(\cdot, s, \cdot) := (v_n^k, \check{v}_n^k)(\cdot, p_n(s), \cdot)$ . Then, we choose a measurable map  $\hat{c}_n^k(t_i^n, \cdot)$  such that

$$\begin{aligned} \hat{c}_n^k(t_i^n, \cdot) &\in \arg \max\{\mathcal{K}^c \check{v}_n^k(t_i^n, s, \cdot), c \in \mathbf{C}(\cdot)\}, \text{ on } D_S \times D_Z^k, \\ \hat{c}_n^k(t_i^n, \cdot) &= 0, \text{ on } D_S \times (D_Z \setminus D_Z^k), \end{aligned}$$

and define the sequence of stopping times

$$\hat{t}_{j+1}^{n,k} := \min\{t_i^n : i \geq 0, t_i^n > \hat{t}_j^{n,k}, (v_n^k - \mathcal{K}^{\hat{c}_n^k} \check{v}_n^k(t_i^n, \cdot))(t_i^n, S_{t_i^n}, \hat{Z}_{t_i^n-}) = 0\}, \quad j \geq 0,$$

with  $\hat{t}_0^{n,k} := 0-$ , and in which  $\hat{Z}^{n,k} = (\hat{P}^{n,k}, \hat{Q}^{n,k}, \hat{X}^{n,k})$  is defined as in (9)-(14) for the initial condition  $Z_{0-}$  and the control associated to  $\hat{\phi}_n^k := (\hat{t}_i^{n,k}, \hat{c}_n^k(\hat{t}_i^{n,k}, S_{\hat{t}_i^{n,k}}, \hat{Z}_{\hat{t}_i^{n,k}-}))_{i \geq 1}$  in a Markovian way. Again, this provides a sequence of controls that is asymptotically optimal.

**Proposition 4.3.** *Let the conditions of Proposition 4.2 hold. Then,*

$$\lim_{k \rightarrow \infty} \lim_{n, n_s \rightarrow \infty} \mathbb{E}[U(S_T, \hat{Z}_T^{n,k})] = v(0, S_0, Z_{0-}),$$

in which the limit is taken along sequences  $n$  such that  $n_t^2 n_s^{-1} \rightarrow 0$ .

*Proof.* Let  $\gamma(p, q) := \int d\beta(c|p, q)$  and recall that  $\gamma$  is uniformly bounded by assumption. The family  $\{(v_n^k, \check{v}_n^k)/L\}_{k, n \geq 1}$  is uniformly bounded, compare with Remark 4.3. Also note that  $(s, z) \in D_S \times D_Z^k \mapsto U$  is  $C_k$ -Lipschitz, for some  $C_k > 0$  that only depends on  $k$ . By induction (recall that the component  $s$  is projected on  $\pi_s^n$ ),

$$|\check{v}_n^k(\cdot, s, \cdot) - \check{v}_n^k(\cdot, s', \cdot)| \leq C'_k[|s - s'| + n_t O(n_s^{-1})], \text{ for all } s, s' \in D_S, \text{ on } D_Z^k,$$

for some  $C'_k > 0$  that does not depend on  $n \in \mathbb{N}^2$ . Since  $S$  is 1/2-Hölder in time in  $\mathbf{L}_2$ , together with Assumption 3.5, this implies that

$$\begin{aligned} \mathcal{K}^c \mathcal{I} \check{v}_n^k(t_{i+1}^n, s_{i+1}^n, z) &= \mathcal{K}^c \mathbb{E}[\mathcal{I} \check{v}_n^k(t_{i+1}^n, S_{t_{i+1}^n}^{t_i^n, s_{i+1}^n}, z)] + O_k(n_t^{-\frac{1}{2}}) + n_t O_k(n_s^{-1}), \\ &= \mathbb{E}[\check{v}_n^k(t_{i+1}^n, S_{t_{i+1}^n}^{t_i^n, s_{i+1}^n}, Z_{t_{i+1}^n}^{t_i^n, z, c})] + O_k(n_t^{-\frac{1}{2}}) + n_t O_k(n_s^{-1}) \text{ for } z \in D_Z^k, c \in \mathbf{C}(z), \end{aligned}$$

in which the exponent  $c$  in  $Z_{t_{i+1}^n}^{t_i^n, z, c}$  means that the impulse  $c$  is given at  $t_i^n$ , and we use Remark 2.1 again. In the above,  $O_k(\xi)$  is a function, that may depend on  $k$ , but such that  $\xi \in \mathbb{R} \setminus \{0\} \mapsto |O_k(\xi)|/|\xi|$  is bounded in a neighborhood of 0. By the arguments already used in

the proof of Proposition 3.3, it follows that

$$\begin{aligned} v_n^k(t_i^n, s_{i'}^n, z) = & \mathbb{E}[v_n^k(t_{i+1}^n, S_{t_{i+1}^n}^{t_i^n, s_{i'}^n}, Z_{t_{i+1}^n}^{t_i^n, z, 0})] \vee \max_{c \in \mathbf{C}(z)} \mathbb{E}[v_n^k(t_{i+1}^n, S_{t_{i+1}^n}^{t_i^n, s_{i'}^n}, Z_{t_{i+1}^n}^{t_i^n, z, c})] \\ & + o_k(n_t^{-1}) + n_t O_k(n_s^{-1}), \end{aligned}$$

and therefore

$$\begin{aligned} v_n^k(0, S_0, Z_{0-}) &= \mathbb{E}[U(S_{\theta_k^n}, \hat{Z}_{\theta_k^n}^{n,k})] + o_n^k(1) + O_k(n_t^2 n_s^{-1}) \\ &= \mathbb{E}[U(S_T, \hat{Z}_T^{n,k})] + o_n^k(1) + O_k(n_t^2 n_s^{-1}) + o_k(1), \end{aligned}$$

by induction, in which  $o_n^k(1)$  goes to 0 as  $n \rightarrow \infty$ ,  $o_k(1)$  goes to 0 as  $k \rightarrow \infty$ , and  $\theta_k^n$  is as in the proof of Proposition 3.3. It remains to appeal to Proposition 4.2.  $\square$

#### 4.5 Numerical experiments

We use the same model as the one described in Section 3.6. As for the new parameters, we take  $\hat{s} = 0$ ,  $\rho = 50$  and  $\sigma = 0.2$ , in particular the mean reversion parameter is taken to be large.

### III. Optimal inventory management and order book modeling

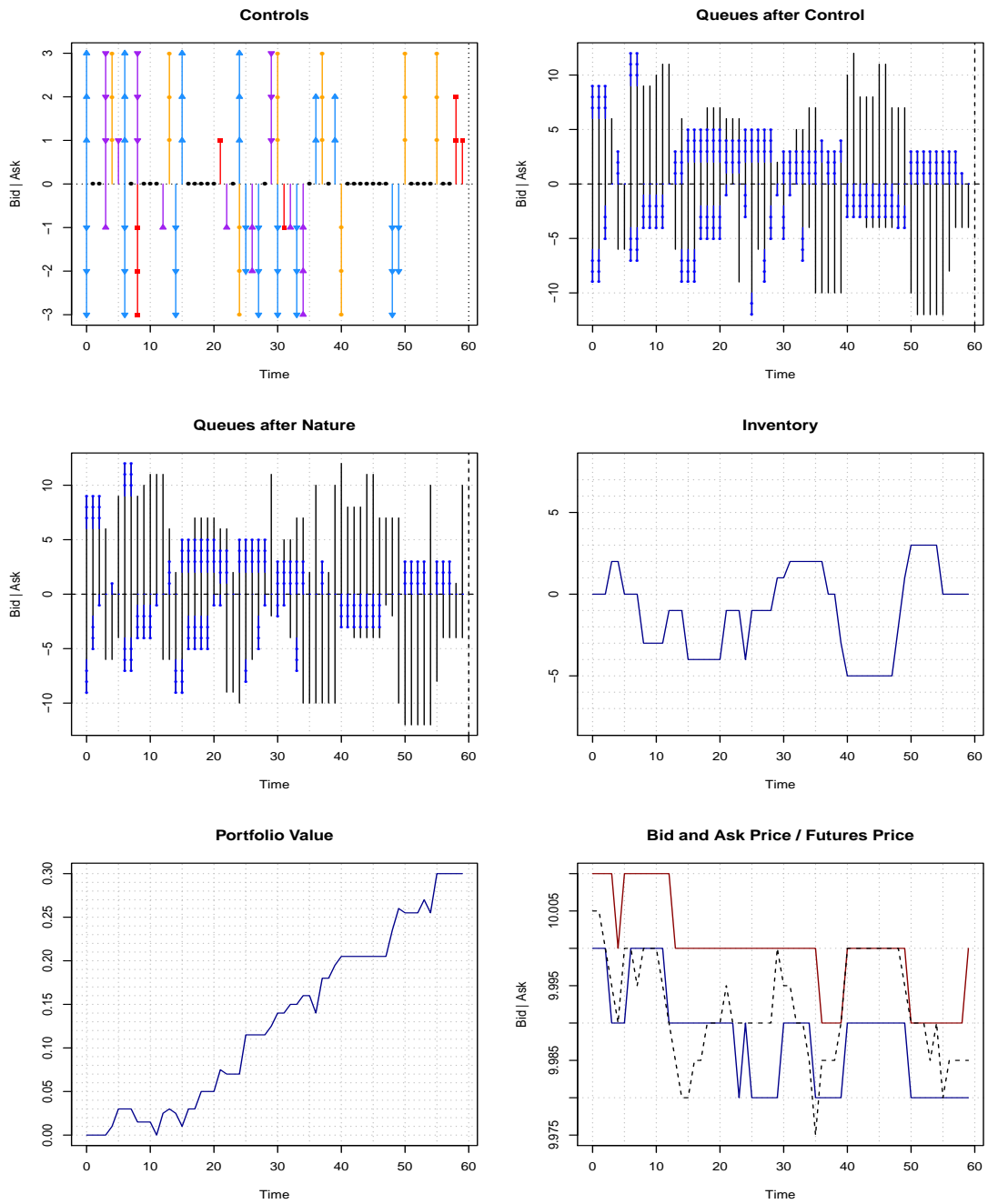


Figure III.3 - Simulated path of the optimal strategy of the High Frequency Trader.

We approximate the behaviour of the spread process  $S$  by a trinomial tree based on the transition probabilities associated to the diffusion (14), so that the expectation in (20) can be computed explicitly. More precisely, we consider a centered 6 points grid with mesh equal to half a tick (i.e. 0.005).

The graphics in Figure III.3 have the same interpretation as in Section 3.6 except that we now also provide the evolution of the Futures process  $F$ , this is the dashed line in the bottom right graphic. Remember that the HFT does not gain from the evolution of the mid-price, as his position in stocks is always covered by a symmetric position in the futures. He only gains from the evolution of the spread process  $S$  or from the bid-ask spread of the stock if  $S$  does not move. Not surprisingly, his behaviour is quite different from the one of the MM described in Section 3.4.

As the MM, the DHFT starts with a symmetric position because the spread with the future is 0. Then he tries selling at time 4 by sending a limit sell order in the spread to buy the futures whose price (compared to the mid-price of the stock) decreases. He is starting to play on the stock-futures spread and follows this strategy until time 14. He plays in a more symmetric way after this until time 35, with a slight tendency to resume his inventory. At time 35, he decides to clearly sell stocks again and buy the futures whose price is again very low. From time 40 on, he inverts his position on the book to try buying the stock and thus sells the futures whose price went up. He finally inverts his position again at time 50 when the futures price goes back to the mid price, to liquidate his position on the pair. Compared with the MM in Figure III.2, DHFT positions are clearly driven by the stock-futures spread. The DHFT also does not seem to try to control the stock's mid-price as the MM did. In Figure III.4, we provide an estimation of the density of the gain of the HFT based on  $10^5$  simulated paths.

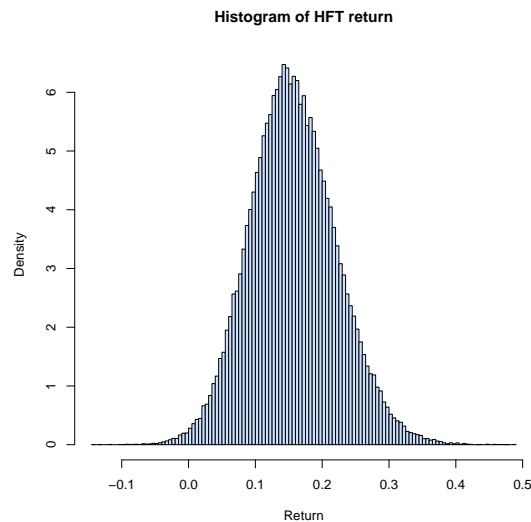


Figure III.4 – Density estimation of the gain made by the High Frequency Trader.

## 5 Institutional broker strategies for portfolio liquidation

We now turn to the Institutional Broker problem. We consider in the following two common strategies for buying/selling a block of stocks. We focus on the buying side, selling being performed in a symmetrical way.

### 5.1 Volume strategy

We first consider a simple volume strategy. We fix a participation rate  $f \in (0, 1)$ . It corresponds to the percentage of the total traded volume executed by the broker between 0 and the end of the execution. This means that the agents should buy the percentage  $f/(1-f)$  of the other market participants trades volume. To acquire  $I_0$ , the IB splits  $[0, T]$  into subintervals  $[t_i, t_{i+1}]$  of  $\mathbb{R}_+$ . At the beginning of each subinterval, he estimates the conditional probability  $p_{i,i+1}$  that a trade arrives at the best bid, given that a trade arrives and given the order book state at  $t_i$ . Then, he inserts a buy limit bid order of size  $(f/(1-f))Q_{t_i}^1/p_{i,i+1}$  at  $t_i$ . Here, the quantity  $Q_{t_i}^1/p_{i,i+1}$  represents the average number of trades needed to deplete the best bid and  $(f/(1-f))$  is the percentage of this traded volume that the IB want to buy. In the meantime, the IB compares the realized volume  $\Delta I := I - I_0$  with the total traded volume on the market  $\tilde{v}(0, \cdot) + \Delta I$ . If  $\Delta I(1-f) > f\tilde{v}(0, \cdot) + \delta_I$  with  $\delta_I > 0$  a given threshold, the broker stops adding limit orders and cancels the already inserted limit orders. If  $\Delta I(1-f) < f\tilde{v}(0, \cdot) - \delta_I$ , the agents sends an aggressive order so as to turn to a position  $\Delta I(1-f) \geq f\tilde{v}(0, \cdot) - \delta_I$  as soon as possible. The IB leaves the market when  $I \geq 0$ .

For our numerical experiment, we take the same model as in Section 3.6, we consider a participation rate  $f = 0.2$  and simply take  $p_{i,i+1} = 1/2$ . The time intervals  $[t_i, t_{i+1}]$  have a length of 60 seconds and the time step is 1/2 second. We take  $\delta_I = 4$  (i.e. 2 ATS). The number of units to buy is  $I_0 = 250$  units. A simulated path of the strategy is provided in Figure III.5. In the top left graphic, the dashed lines correspond to the target volume  $\pm\delta_I$ , while the solid curve is the volume traded by the IB. The top right graphic gives the control of the IB: lines with inward pointing arrows are limit orders, lines with squares are market orders. The bottom left graphic provides the evolution of the average price at which stocks have been bought by the IB from time 0 on. The bottom right graphic gives the evolution of the bid and ask prices. One can see that this very simple strategy is quite efficient in the sense that only a limited number of market orders are sent. On the other hand, the systematic position of the IB at the bid limit creates an important imbalance that contributes to push up the price.

Figure III.6 provides an histogram of the relative error (in %) of the VWAP obtained by this strategy with respect to the VWAP realized at the level of the whole market.<sup>9</sup> It is based on  $10^4$  simulated paths. One can see that this average price is typically slightly higher than the VWAP of the market.

---

<sup>9</sup>Namely,  $(\text{VWAP}_{\text{Vol}} - \text{VWAP}_{\text{Market}})/\text{VWAP}_{\text{Market}}$  in which  $\text{VWAP}_{\text{Vol}}$  is the VWAP obtained by the IB by playing his volume strategy.

## 5. Institutional broker strategies for portfolio liquidation

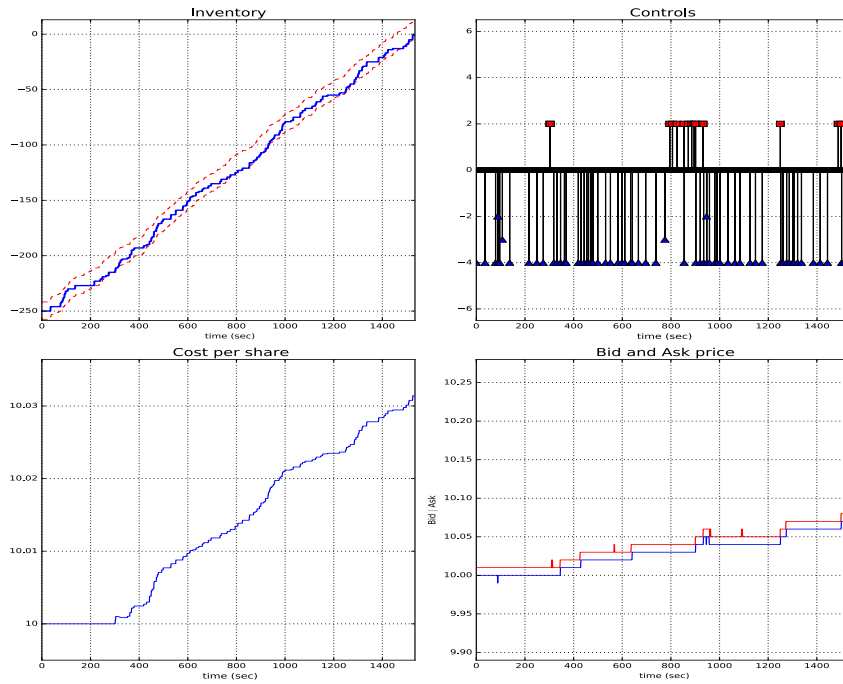


Figure III.5 – Simulated path of the volume strategy.

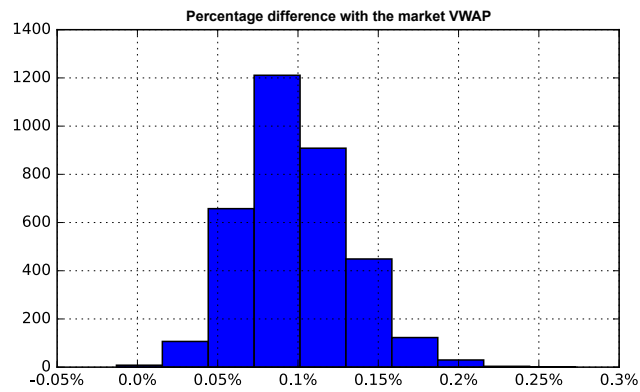


Figure III.6 – Histogram of the VWAP percentage error with respect to the VWAP of the market, for the Volume strategy.



## 5.2 VWAP

We now present a VWAP (Volume Weighted Average Price) based trading strategy frequently practiced by institutional brokers. Suppose that, at initial time 0, an institutional broker decides to buy a quantity  $I_0 \in \mathbb{N}$  of a tradable asset using a VWAP based strategy, i.e. by trying to obtain an average execution cost not more than

$$\text{VWAP}_T = \frac{w_T}{v(0, T)} \quad \text{where } w := \int_0^\cdot v_t P_t dt \text{ and } v(\cdot, T) := \int_\cdot^T v_s ds, \quad (21)$$

in which  $v$  is a deterministic non-negative continuous process such that  $\int_0^T v_t dt > 0$ , which represents the trading volume of the market, and  $P$  models the stock price.<sup>10</sup>

### 5.2.1 Abstract continuous time resolution

Following the seminal work of Almgren and Chriss [9], see also [76, Section 4.4], we first consider the idealized world in which trading is done continuously at a bounded intensity, the trading speed  $\vartheta$ , taken as a process in the class  $\mathcal{C}$  of non-negative processes adapted to the (completed) filtration generated by  $W$ . He assumes that the dynamics of the asset reference price  $P$  have a permanent linear price impact generated by the agent's trading activity, see [9]. More precisely, he assumes that the stock price evolves according to

$$dP_t = \beta \vartheta_t dt + \sigma dW_t, \quad (22)$$

with  $\beta, \sigma > 0$ , while his inventory  $I$  follows the dynamics

$$dI_t = \vartheta_t dt,$$

with the initial condition  $I_0 = -I_0$ , meaning that he is short at 0 of  $I_0$  stocks that he has to buy on  $[0, T]$ . He also assumes that his wealth  $G$  is affected by a temporary linear market impact  $\kappa \vartheta$ , for some  $\kappa > 0$ , i.e.

$$dG_t = -\vartheta_t [P_t + \kappa \vartheta_t] dt. \quad (23)$$

The goal of the agent is to maximise over  $\vartheta \in \mathcal{C}$  the expected utility

$$\mathbb{E} \left[ -\exp \left\{ -\eta \left\{ G_T + I_T (P_T - \tilde{\kappa} I_T) - I_0 \frac{w_T}{v(0, T)} \right\} \right\} \right], \quad (24)$$

for some  $\tilde{\kappa} > 0$  which represents a penalty in case the inventory does not match 0 at  $T$ . This type of problems has been widely studied, see the books [76, 82] for references. In the present form, it can be solved explicitly by using a verification argument. We have the following result.

---

<sup>10</sup>We take it deterministic for simplicity. In practice, it generally corresponds to a market volume curve estimated by the broker.

**Theorem 5.1.** *The IB value function defined such that*

$$v(t, p, i, g, w) = \sup_{\mu} \mathbb{E}[-\exp[-\eta\{G_T + I_T(P_T - \tilde{\kappa}I_T) - I_0 VWAP_T\}] | \mathcal{F}_t],$$

*satisfies the semi closed formula*

$$v(t, p, i, g, w) = -e^{-\eta\{g - \bar{m}w - p v(t, T) \bar{m} + i(p - \tilde{\kappa}i)\}} \times e^{h_0(t) + h_1(t)i + h_2(t)i^2},$$

*where the functions  $h_0$  and  $h_1$  are fully characterized by ODEs*

$$\begin{aligned} \partial_t h_0 &= -\frac{1}{2}\sigma^2(\eta v(t, T) \bar{m})^2 + \frac{(h_1 + \beta\eta v(t, T) \bar{m})^2}{4\kappa\eta}, \\ \partial_t h_1 &= \sigma^2\eta^2 v(t, T) \bar{m} + (h_1 + \beta\eta v(t, T) \bar{m}) \frac{-\beta\eta + 2\eta\tilde{\kappa} + 2h_2}{2\kappa\eta}, \\ \partial_t h_2 &= -\frac{\eta^2\sigma^2}{2} + \frac{[-\beta\eta + 2\eta\tilde{\kappa} + 2h_2]^2}{4\kappa\eta}. \end{aligned}$$

The above result shows that the optimal control policy is given by

$$\hat{\theta} = \hat{v}(\cdot, I),$$

where

$$\begin{aligned} \hat{v}(t, i) &:= \frac{\beta\partial_p v(t, p, i, g, w) + \partial_i v(t, p, i, g, w) - p\partial_g v(t, p, i, g, w)}{2\kappa\partial_g v(t, p, i, g, w)} \\ &= \frac{\beta[i - v(t, T) \bar{m}] - 2\tilde{\kappa}i - [h_1(t) + 2h_2(t)i]/\eta}{2\kappa}. \end{aligned}$$

One can check that the function  $h_2$  has the following form:

$$h_2(t) = \frac{1}{c_1 + c_2 e^{c_3(T-t)}} - c_4,$$

in which the constants  $c_1, \dots, c_4$  can be computed explicitly by using the above differential equation and the terminal condition  $h_2(T) = 0$ . Namely, set

$$a_0 := -\frac{\eta^2\sigma^2}{2} + \frac{\eta(2\tilde{\kappa} - \beta)^2}{4\kappa}, \quad a_1 := \frac{2\tilde{\kappa} - \beta}{\kappa}, \quad a_2 = \frac{1}{\kappa\eta},$$

and let  $y_\circ$  be a root<sup>11</sup> of

$$(4a_0a_2 - a_1^2)y^2 + (a_1^2 - 4a_0a_2)y + a_0a_2 = 0,$$

then

$$c_3 = \frac{a_1}{1 - 2y_\circ}, \quad c_4 = \frac{a_0}{(1 - y_\circ)c_3}, \quad c_1 = \frac{y_\circ}{c_4}, \quad c_2 = \frac{1}{c_4} - c_1.$$

Existence of  $y_\circ$  is guaranteed since

$$a_1^2 - 4a_0a_2 = \frac{\eta^2\sigma^2}{2} > 0.$$

<sup>11</sup>The function  $h_2$  does not depend on the choice of the root.

### III. Optimal inventory management and order book modeling

*Proof of Theorem 5.1.* The value function  $v$  satisfies the following Hamilton-Jacobi-Bellman equation:

$$0 = \sup_{u \geq 0} \left( \partial_t \varphi + \frac{1}{2} \sigma^2 \partial_{pp}^2 \varphi + p v_t \partial_w \varphi + u (\beta \partial_p \varphi + \partial_i \varphi - p \partial_g \varphi) - u^2 \kappa \partial_g \varphi \right), \text{ on } [0, T) \times \mathbb{R}^4$$

with terminal condition

$$\varphi(T, p, i, g, w) = -e^{-\eta\{g+i(p-\bar{\kappa}i)-\bar{m}_0 w\}}, \text{ for } (p, i, g, w) \in \mathbb{R}^4,$$

where

$$\bar{m} := I_0 / v(0, T).$$

To simplify the above, we first write  $v$  in the form

$$v(t, p, i, g, w) = e^{-\eta(g-\bar{m}w-pv(t,T)\bar{m})} \bar{v}(t, p, i) \text{ with } \bar{v}(t, p, i) := v(t, p, i, 0, 0),$$

so that  $\bar{v}$  formally solves

$$0 = \partial_t \varphi + \frac{1}{2} \sigma^2 (\partial_{pp}^2 \varphi + 2\eta v(t, T) \bar{m} \partial_p \varphi + \eta^2 v(t, T)^2 \bar{m}^2 \varphi) - \frac{(\beta \partial_p \varphi + \beta \eta v(t, T) \bar{m} \varphi + \partial_i \varphi + \eta p \varphi)^2}{4\kappa \eta \varphi},$$

on  $[0, T) \times \mathbb{R}^4$ , with terminal condition

$$\varphi(T, p, i, g, w) = -e^{-\eta\{i(p-\bar{\kappa}i)\}} \text{ for } (p, i, g, w) \in \mathbb{R}^4.$$

One possible solution is given by

$$\bar{v}(t, p, i) = -e^{-\eta\{i(p-\bar{\kappa}i)\}} e^{h_0(t)+h_1(t)i+h_2(t)i^2},$$

in which  $h_0, h_1$  and  $h_2$  solve<sup>12</sup> on  $[0, T)$

$$\begin{aligned} \partial_t h_0 &= -\frac{1}{2} \sigma^2 (\eta v(t, T) \bar{m})^2 + \frac{(h_1 + \beta \eta v(t, T) \bar{m})^2}{4\kappa \eta}, \\ \partial_t h_1 &= \sigma^2 \eta^2 v(t, T) \bar{m} + (h_1 + \beta \eta v(t, T) \bar{m}) \frac{-\beta \eta + 2\eta \bar{\kappa} + 2h_2}{2\kappa \eta}, \\ \partial_t h_2 &= -\frac{\eta^2 \sigma^2}{2} + \frac{[-\beta \eta + 2\eta \bar{\kappa} + 2h_2]^2}{4\kappa \eta}, \end{aligned}$$

with  $h_0(T) = h_1(T) = h_2(T) = 0$ . Then,  $h_0, h_1$  and  $h_2$  are fully characterized. We complete the proof with an easy verification argument.  $\square$

<sup>12</sup>Just insert the above in the PDE of  $\bar{v}$  and match the orders in the  $i$  variable.

### 5.2.2 Strategy in practice and simulations

In practice, this optimal strategy cannot be implemented within an order book. We therefore consider a “discrete” version. In this version, we assume that the IB tries to keep his inventory  $I$  equal to the optimal inventory  $\hat{I} := \int_0^T \hat{\vartheta}_t dt - I_0$ . To do this, he fixes a time grid  $\{t_i, i \leq n\}$  with  $t_0 = 0$  and  $t_n = T$ . At time  $t_i$ , he estimates that she has to execute a volume of  $V_{i,i+1} := \int_{t_i}^{t_{i+1}} \hat{\vartheta}(t, I_{t_i}) dt$  on  $[t_i, t_{i+1}]$  while the volume of the market will be  $\tilde{V}_{i,i+1} = \int_{t_i}^{t_{i+1}} v_t dt$ . Then, he follows a volume strategy with threshold  $\delta_I = +\infty$ , see Section 5.1 above, on  $[t_i, t_{i+1}]$  with a target participation rate of  $f = V_{i,i+1} / (\tilde{V}_{i,i+1} + V_{i,i+1})$ . When  $\Delta I := I - I_0 > \hat{I} + \bar{\delta}_I$ , for some  $\bar{\delta}_I > 0$ , orders are cancelled and he waits until  $\Delta I \leq \hat{I} + \bar{\delta}_I$ . In the case where  $\Delta I < \hat{I} - \bar{\delta}_I$ , he send market orders to reduce to the difference  $\hat{I} - \bar{\delta}_I - \Delta I$  as fast as possible.

For our numerical experiment, we take the same configuration as in Section 5.1, with  $\bar{\delta}_I = 4$ , i.e. 2 ATS. The optimization of the VWAP strategy is done with a time horizon of 30 minutes and a flat volume curve (so that the control does in fact not depend on it). The number of units to buy is  $I_0 = 250$  units. The additional parameters<sup>13</sup> are set to  $\eta = 1$ ,  $\sigma = 0.2$ ,  $\beta = 0,0004$ ,  $\kappa = 0.003$  and  $\tilde{\kappa} = \kappa * 60$ . The latter corresponds to the cost incurred when buying the remaining shares  $I_T$  in 1 additional minute after  $T$ , at a flat intensity in the theoretical continuous time model of Section 5.2.1. The volume intensity  $v$  corresponds to 0.6 ATS per second, i.e. 1.2 units per second.

The interpretation of the different graphics in Figure III.7 is the same as in Section 5.1, except that the dashed lines in the top left graphic correspond now to the optimal VWAP trading curve  $\pm 2$  ATS. Again, we see that only a limited number of market orders were needed to be sent, but that the imbalance created by the VWAP trading algorithm drives the price up. Figure III.8 provides an histogram of the relative error (in %) of the VWAP obtained by this strategy with respect to the VWAP of the whole market.<sup>14</sup> It is based on  $10^4$  simulated paths. One can see that he actually typically performs better than the market. Not surprisingly, this strategy performs better than the volume strategy in terms of VWAP.

<sup>13</sup>The coefficients  $\beta$  and  $\kappa$  are estimated for our book dynamics. Given priors with simulate a bunch of paths and estimated them by a moment matching approach based on (22)-(23). We then use the updated values to simulate a new bunch of paths and we re-estimate them. And so on, until convergence.

<sup>14</sup>Namely,  $(\text{VWAP}_{\text{MM}} - \text{VWAP}_{\text{Market}}) / \text{VWAP}_{\text{Market}}$ , in which  $\text{VWAP}_{\text{MM}}$  is the VWAP obtained by the IB by playing his optimal VWAP strategy

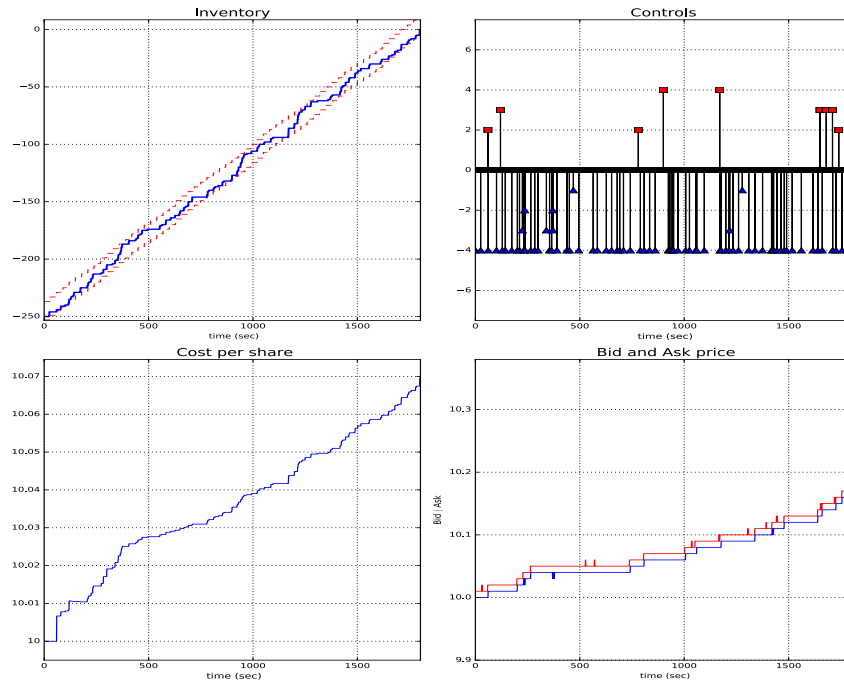


Figure III.7 – Simulated path of the VWAP strategy.

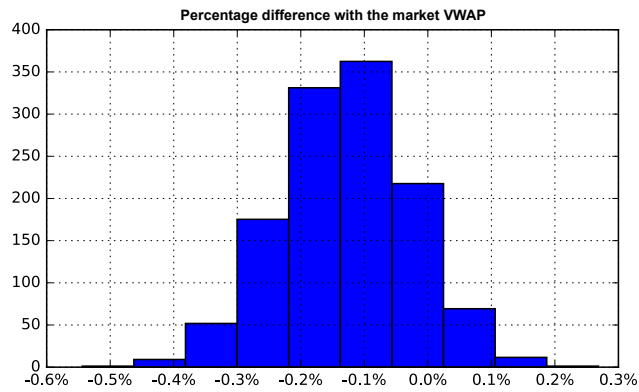


Figure III.8 – Histogram of the percentage error with respect to the VWAP of the market, for the VWAP strategy.

## 6 Simulation of the full market

We now provide an illustration of possible interactions of several market participants. Again, a more realistic and general study is left for future research.

The prior and parameters of the participants are the same as the one that have been used in the above numerical experiments. More precisely, we consider one MM, one DHFT and four IB. Among the IB, two are playing volume strategies (a buy and a sell) and two are playing VWAP strategies (a buy and a sell). At each time step, each participant decides of his optimal control given the current state of the order book. They send their orders at the same time. The control of the DHFT is executed first, then the one of the MM, and finally the controls of the Volume and VWAP trading algorithms are executed (we choose the order among them randomly, according to a uniform distribution).

The exogenous randomness only comes from the simulation of the stock-futures spread process  $S$  and from the new queues created when one queue is depleted. If the bid queue is depleted, then the bid price moves down. If the bid-ask spread is equal to two ticks, then the ask price moves down as well. The other way round if the ask queue is depleted. If the bid (resp. the ask) price moves down (resp. up), we consider that this is a discovered limit and its size is chosen as in Section 3.4: 10 units with probability 60%, 5 units with probability 25% and 12 units with probability 15%. If the bid (resp. the ask) price move up (resp. down), which can happen if the bid-ask spread is already of two ticks, we consider that this is a created limit and its size is chosen again as in Section 3.4: 2 units with probability 60%, 1 unit with probability 25% and 3 units with probability 15%. There is no other randomness. The rest of the dynamics is generated by the HFT, the MM and the four other trading algorithms.

Note that only the VWAP trading algorithms are forced to trade, when they reach the (upper or lower) limit of their prescribed inventory path. In principle, the MM plays aggressive orders only when he needs to adjust quickly his inventory. For a typical path of the stock-futures spread, the HFT also has no incentive to send aggressive orders, except to adjust his inventory. Without trades, the Volume trading algorithms do not act as well (because the market volume does not move). Therefore, if the MM and the HFT have a zero initial inventory, we expect to have to wait for the VWAP trading algorithms to initiate first aggressive orders, and starts the whole dynamics. In our illustration, we start with an initial state in which the HFT and the MM have a zero initial inventory. Each VWAP trading algorithm has to buy/sell 75 stocks within 5 minutes of trading. This corresponds to an average of 0.5 trades per second, which is consistent with the priors of the MM and the HFT.

The Figures III.9, III.10, III.11, III.12, III.13 and III.14 have the same interpretation as in Section 3.4 except that now the position of the agent in the queue is in blue in the top right and middle left graphics (the black part corresponding to the other participants). The top right graphic is the state of the book just after the control of the agent is executed, the middle left is the state of the order book after the controls of all the participants have been executed. Note

that the trades are surprisingly due to the VWAP trading algorithms, to which the Volume trading algorithms need to adjust. The aggressive orders sent by the HFT and the MM are either cancellations or corrections of their inventory near the terminal time, except around time 35 at which the HFT clearly wants to take a position on the stock-futures spread because it is very low.

## 7 Conclusion

We have proposed a simplified but still realistic modelling of an order book, whose dynamics depend on the current imbalance. We have derived optimal strategies for three key actors: Market Makers, High Frequency Traders and Institutional Brokers (volume and VWAP trading algorithms). In this model, optimal strategies can be estimated numerically. Simulations show how complex optimal strategies are when actors believe that the imbalance actually has an impact on the dynamics of order flows. Our numerical simulations show that the construction of a realistic market simulator is feasible. A desirable next step would be to consider realistic proportions of market participants for the simulations proposed in Section 6 and to see whether the market statistics used to calibrate the different strategies can be retrieved through the simulated interactions of the various players. If this is the case, one can imagine trying to exhibit different market patterns depending on the nature and proportions of actors currently acting.

**Acknowledgment:** This research is part of a Cemracs 2017 project and benefited from the support of the Initiative de Recherche from Kepler-Cheuvreux and Collège de France. We are very grateful for the support of P. Besson and his team.

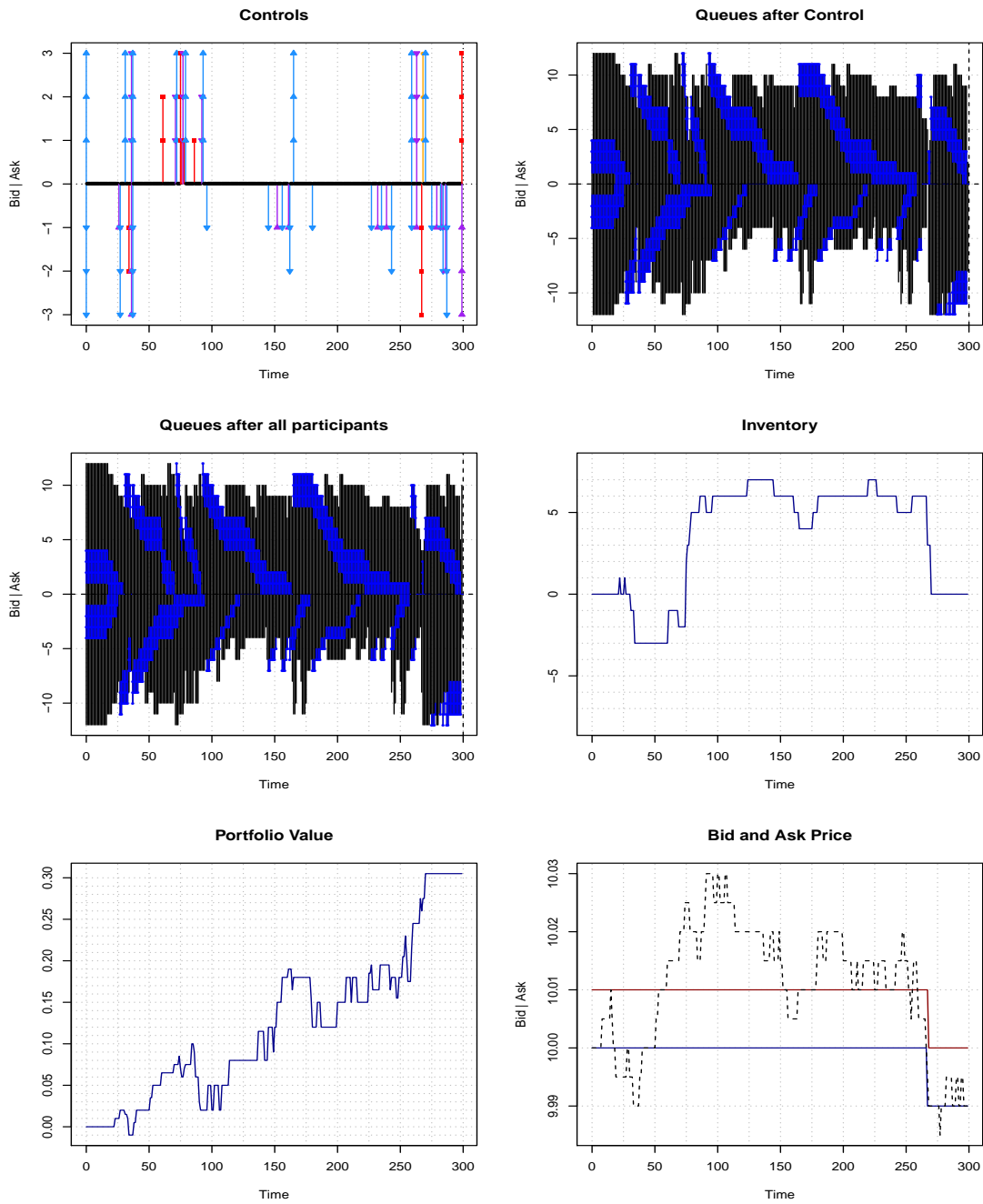


Figure III.9 - Optimal strategy of the High Frequency Trader when agents play together.



### III. Optimal inventory management and order book modeling

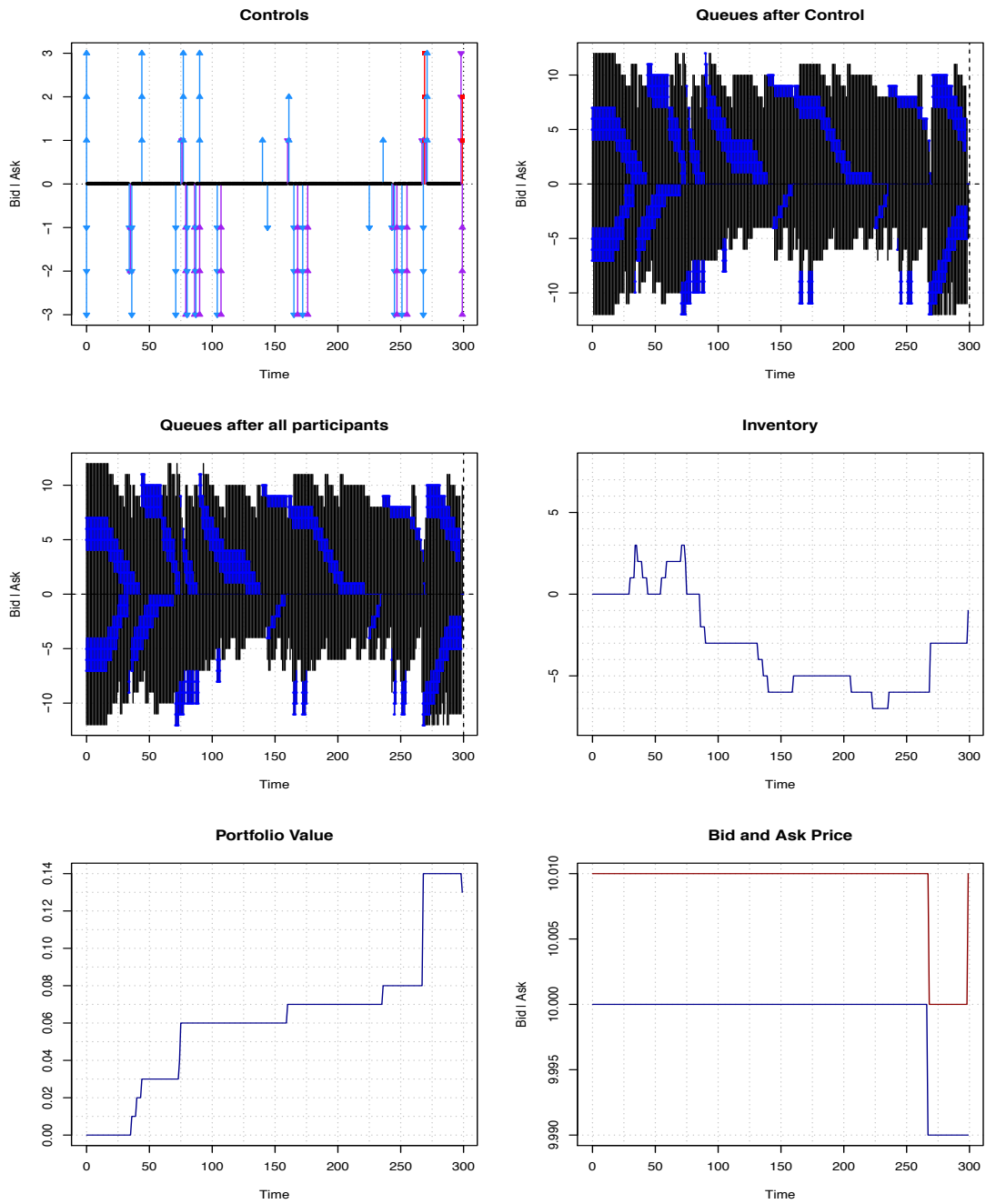


Figure III.10 - Optimal strategy of the Market Maker when agents play together.

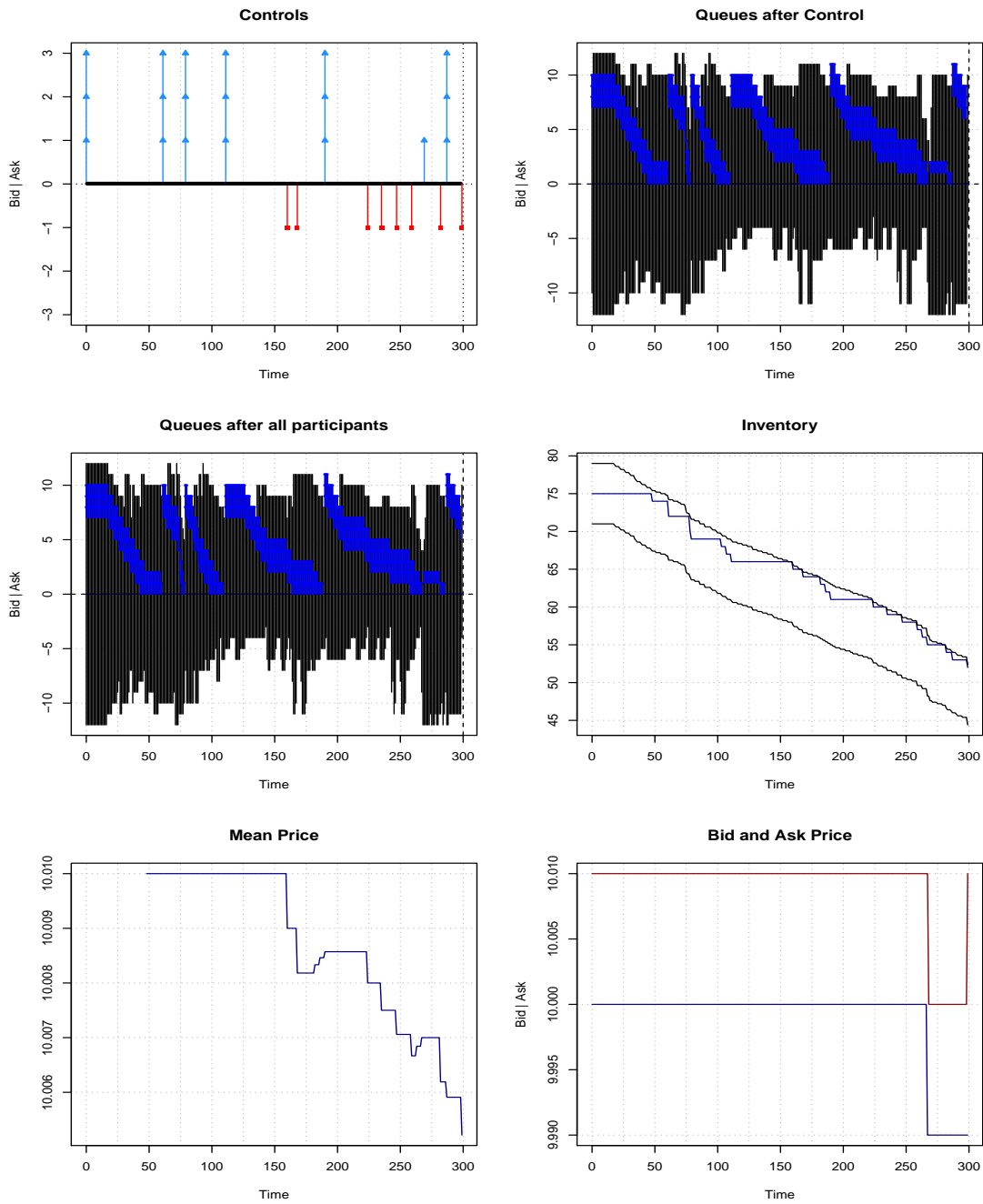


Figure III.11 - Optimal strategy of the Volume trading algorithm (seller) when agents play together.

### III. Optimal inventory management and order book modeling

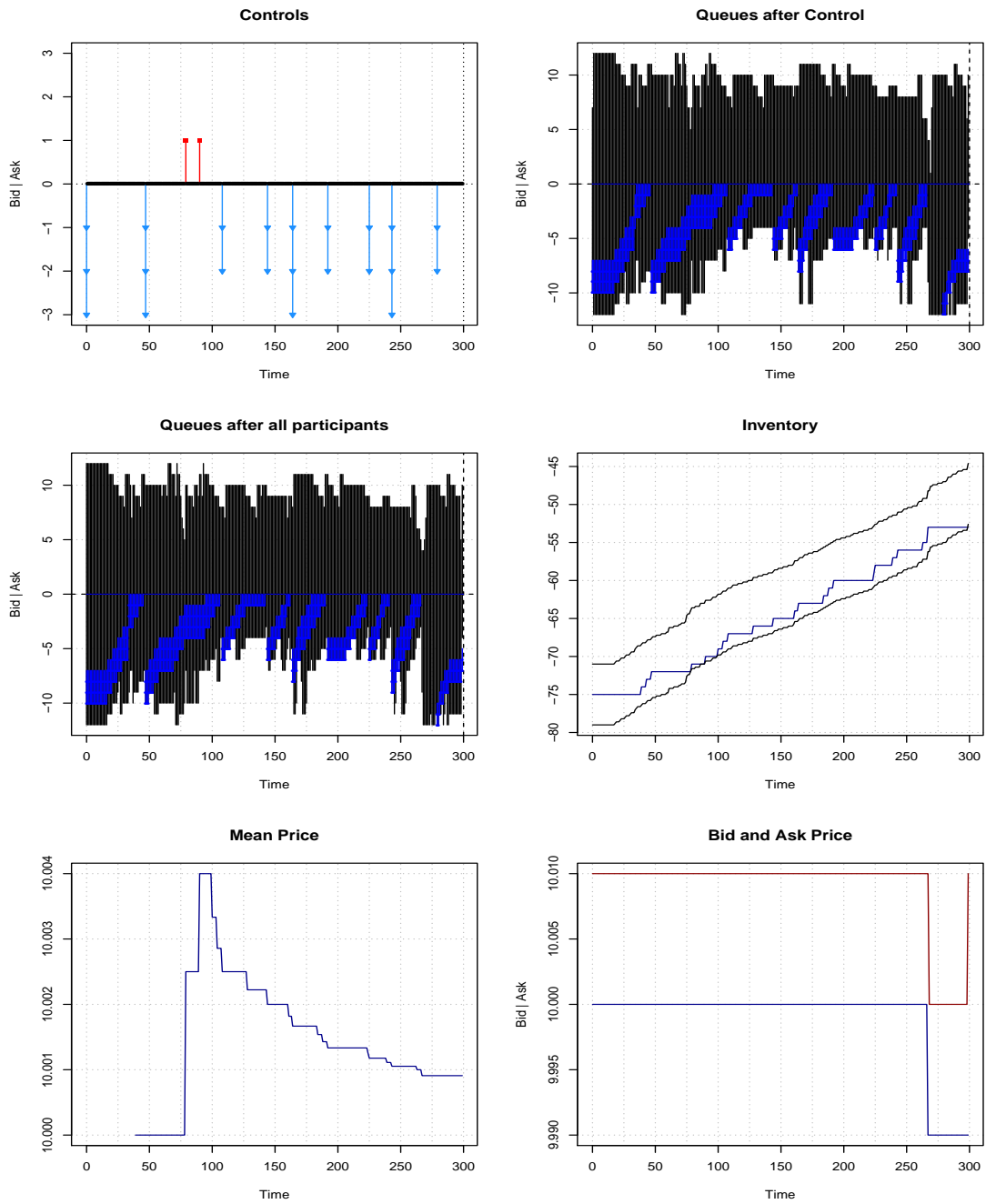


Figure III.12 - Optimal strategy of the Volume trading algorithm (buyer) when agents play together.

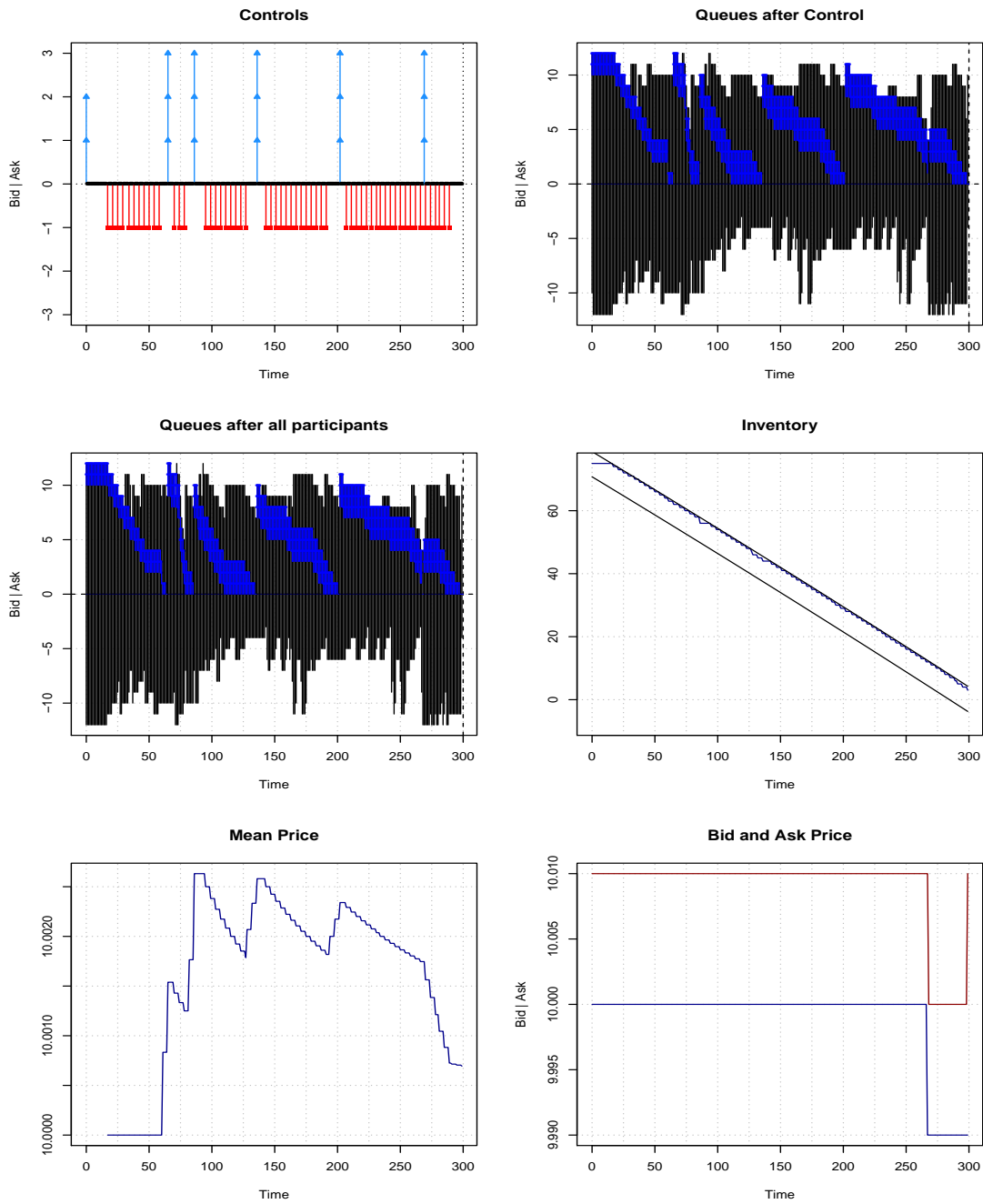


Figure III.13 - Optimal strategy of the VWAP trading algorithm (seller) when agents play together.

### III. Optimal inventory management and order book modeling

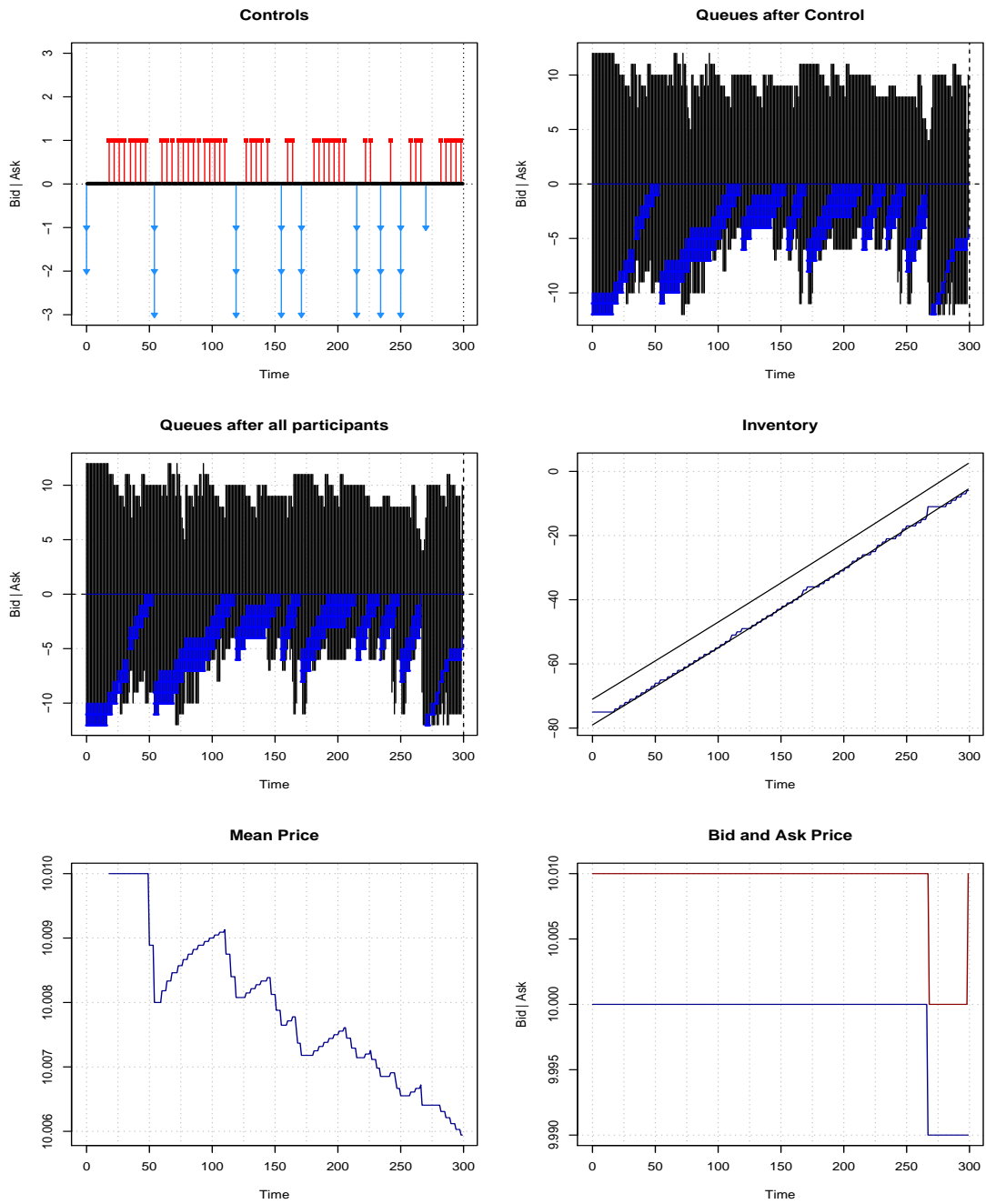


Figure III.14 - Optimal strategy of the VWAP trading algorithm (buyer) when agents play together.

---

# From asymptotic properties of general point processes to the ranking of financial agents

---

## Abstract

We propose a general non-linear order book model that is built from the individual behaviours of the agents. Our framework encompasses Markovian and Hawkes based models. Under mild assumptions, we prove original results on the ergodicity and diffusivity of such system. Then we provide closed-form formulas for various quantities of interest: stationary distribution of the best bid and ask quantities, spread, liquidity fluctuations and price volatility. These formulas are expressed in terms of individual order flows of market participants. Our approach enables us to establish a ranking methodology for the market makers with respect to the quality of their trading.

**Keywords:** Market microstructure, limit order book, high-frequency trading, market making, queuing model, Hawkes processes, ergodic properties, volatility, regulation.

## 1 Introduction

In the last two decades, the development of electronic and fragmented markets has lead to a deep disruption in the landscape of market participants. In particular, traditional market making institutions have been largely replaced by high-frequency market makers. Market makers are intermediaries between buyers and sellers. In an electronic limit order book, they provide liquidity to market participants willing to trade immediately by simultaneously posting limit orders on both sides of the book. Market makers undergo different types of risk, mainly adverse selection and inventory risks. To avoid adverse selection risk, they must be able to update very frequently their quotes in response to other order submissions or cancellations. To minimise their inventory risk, they need to use smart algorithms enabling them to hold positions for very short time periods only, see for example [115].

High-frequency traders (HFTs) are now the only market participants that are indeed able to play the role of market makers on liquid stocks, see [98]. This is achieved thanks to an intense use of speed (co-location) and technology. They are supposedly capable to maintain a strong presence at best price limits and control adverse selection at the same time, see [99], while operating efficient inventory management in an increasingly fast-moving market, see [5, 18]. This is to the extent that HFTs are described as the new market makers in [120].

Since the arrival of these new market makers, academics, regulators and practitioners aim at understanding whether their activity is harmful or beneficial for markets. On the one hand, some argue that HFTs have a positive impact on markets: the competition between market makers leads to an increase in market depth, to narrower bid-ask spreads which is equivalent to reduced trading costs for other investors, see [88, 99] and to better price discovery, see [88, 137]. On the other hand, others assert that high-frequency market makers have toxic consequences. For example, they worsen market volatility during flash crashes by aggressively liquidating their long positions, see [101, 116].

One important common point in most studies analysing the behaviour of HFTs is that they try to measure how HFTs impact the market as a group, without investigating individual behavioural disparities among them. Chapters I and II shed light on the fact that all HFTs do not behave similarly, showing for example that they have very different levels of aggressiveness and liquidity provision. In this paper, we wish to participate to the debate about the role of HFTs on market quality by bringing some new quantitative elements enabling regulators and exchanges to assess the individual effects of each high-frequency market maker operating on the market. In particular, we want to be able to rank market makers according to the quality of their trading.

We use several metrics for market quality such as spread and liquidity fluctuations, but a particular focus is given to the price volatility. This idea of disentangling market participants contribution to volatility is used in [136]. In this work, the authors nicely model the interactions between the various orders of the different market participants using linear Hawkes processes. This model is very interpretable: an order of type  $A$  of Agent  $i$  raises the likelihood of an order of type  $B$  of Agent  $j$  by a certain amount. Consequently, the authors naturally define the contribution of Agent  $A$  to the volatility by the weighted sum over all possible types of orders of Agent  $A$  of the squared mean price jump triggered by each of these orders, the associated weight being the intensity of the corresponding order type.

Our focus here is on market makers. Thus one crucial element to take into account is the well-known fact that the main market driver of any market making strategy is the state of the limit order book (and not single individual orders of other market participants), see [90, 125] and Chapter I. Therefore, in the spirit of the Queue-reactive model of [90], we assume that the state of the order book, which is a common component, affects the interactions between our high-frequency market participants. However, to get a really accurate modelling of the behaviour of the agents, we also let their individual actions depend on their own past ones

and on those of other participants, in the spirit of [136]. We allow for strong non-linearities in the dependences with the past, leading to a much generalised version of Hawkes-Queue-reactive type order book models, see [126, 153].

In this extended and non-Markovian framework, we are able to prove the ergodicity and diffusivity of our system, see [91] for inspiring ideas. Furthermore, we provide asymptotic expressions for market quantities such as spread, liquidity fluctuations or price volatility in terms of the individual order flows of market participants. This notably enables us to forecast the dynamics of the market in case one market maker leaves it. The idea is that we consider that market makers interact with the market through their algorithms which are specified for example in term of average event size or in term of relative quantities such as the imbalance. If we remove one market participant while the others do not modify their algorithms, we can for instance compute a new volatility. If it is larger (smaller) than the actual one, we can say that the considered market maker has a stabilising (destabilising) effect on the market. This eventually leads us to a ranking of market makers with respect to the quality of their trading.

Let us now give a brief description of our model. Let  $n$  be a positive integer representing the index of the  $n$ -th order book event  $e_n$ . Each event  $e_n$  happens at time  $T_n$  and is characterised by a variable  $X_n$  that encodes all the needed information to describe  $e_n$ . For example,  $X_n$  contains the order size, the type of the order (limit order, liquidity-consuming order such as market order or cancellation), the order posting price and the identity of the agent. A detailed description of the sequences  $(T_n)_{n \geq 1}$  and  $(X_n)_{n \geq 1}$  is given in Section 2.2. The order book state is modelled by the process  $U_n = (Q_n^1, Q_n^2, S_n)$  with  $Q_n^1$  the available quantity at the best bid,  $Q_n^2$  the available quantity at the best ask and  $S_n$  the spread at time  $T_n$ . For a detailed description of the dynamics of  $U_n$ , see Equation (1). Here we focus on the first limits to reduce the dimension of the state space and keep a tractable model.<sup>1</sup> Finally, we use a general approach to infer the behaviour of the price process from that of  $(U_n)$ , in the spirit of [91] and Chapter II, see Section 4 for the detailed formulation. We define the non-linear Hawkes-Markovian arrival rate  $\lambda_t(e)$  of an order book event  $e$  ( $e$  containing the identity of the involved agent) at time  $t \in \mathbb{R}_+$  as follows:

$$\lambda_t(e) = \psi\left(e, U_{t-}, t, \sum_{T_i < t} \phi(e, U_{t-}, t - T_i, X_i)\right),$$

where  $\psi$  is a non-linear function,  $U_{t-}$  is the order book state relative to the last event before  $t$  and  $\phi$  is the Hawkes kernel representing the influence of past events. The functions  $\phi$  and  $\psi$  are both  $\mathbb{R}_+$ -valued. In absence of the kernel  $\phi$ , the function  $\psi$  leads to a classical Markovian approach since the arrival rate of an event  $e$  depends essentially on the order book state  $U_{t-}$ . When  $\phi$  is non-zero,  $\psi$  controls the interactions between the past events and the current order book state. Note that we allow  $\psi$  to have a polynomial growth while in the literature, it has at most a linear growth, see [41]. Additionally, we do not impose  $\psi$  and  $\phi$  to be continuous, which means that a sudden change of regime in the order book dynamics are

<sup>1</sup>However, we can model deeper limits by enlarging the dimension of the state space.



also incorporated in our modelling. Finally, we propose an agent-based model since market participant identities are contained in the order book events  $e$  through the variables  $(X_i)_{i \geq 1}$ .

Our framework is a generalised order book model where the arrival rate of the events follows a non-linear Hawkes-type dynamics that depend on the order book state. This approach covers most existing bid-ask order book models. It is a natural extension of the Poisson intensity models, see [2, 83, 144], the Markovian Queue-reactive model introduced in [90] and the Hawkes based models such as [3, 126, 136]. In this setting, under mild assumptions, we provide new ergodic results and limit theorems, expressing all the limiting quantities in terms of the individual flows of market participants. Furthermore, we build an estimation methodology for the intensity functions which turns out to be similar to the one used in the Queue-reactive case, see [91], although the model here is much more general and non-Markovian. These theoretical results for our point processes, which largely extend classical ergodicity properties limited to the Markov case, are the basis for the assessment of the role of the different market participants on market quality as explained above.

The paper is organised as follows. First, we introduce in Section 2 our order book model and describe how to recover the market dynamics from the individual behaviours of each agent. Then, we prove the ergodicity of our system in Section 3 and its diffusivity in Section 4. In Section 5, we provide the needed formulas to compute the order book stationary distribution, the price volatility and the liquidity fluctuations. Finally, numerical results and ranking of market makers on several assets are provided in Section 6. Proofs and additional results are relegated to an appendix.

## 2 Market modelling

In this section, we describe the order book model and show how to recover the market dynamics given the agents individual behaviours.

### 2.1 Introduction to the model

In the order book mechanism buyers and sellers send their orders to a continuous-time double auction system. Market participants orders have a specific size that is measured in average event size (AES)<sup>2</sup> and the orders can be sent to different price levels that are separated by a minimum distance which is the tick size. In our model, we only consider the price levels between the best bid and ask prices to reduce the dimension of the state space. Additionally, we assume that the agents can take three elementary decisions:

- Insert a limit order of a specific size at the best bid or ask price, hoping to get an execution.
- Insert a buying or selling limit order of a specific size within the spread.

---

<sup>2</sup>AES is the average size of events observed in the limit order book.

- Send a liquidity-consuming order of a specific size at the best bid or ask price. Cancellation and market orders have the same effect on liquidity. Thus, they are aggregated to constitute the liquidity consumption orders.

The size of the orders is not constant in the model. Finally, the mid price is valued on the set  $\tau_0\mathbb{Z}$  where  $\tau_0$  is the tick size. A simple example is to consider the case where the mid price decreases (resp. increases) by one tick when the best bid (resp. ask) is totally depleted. Here, the mid price jumps size may be larger than one tick. In the rest of the article, we take the mid price as our reference price for simplification. The dynamics of the model are illustrated in Figure IV.1.

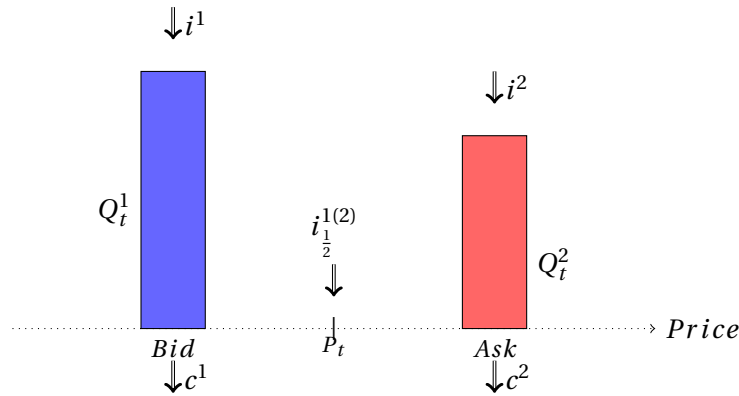


Figure IV.1 – Diagram of flows affecting our order book model. The quantity  $i^1$  (resp.  $i^2$ ) represents the insertion of limit orders at the best bid (resp. ask). The quantity  $i_{\frac{1}{2}}^1$  (resp.  $i_{\frac{1}{2}}^2$ ) is associated to buying (resp. selling) limit orders within the spread. The quantities  $c^1$  and  $c^2$  refer to the orders that consume respectively the liquidity at the best bid and ask.

**Notations.** We consider the following notations:

- The current physical time is  $t$ .
- The mid price is  $P_t$ , the best bid price is  $P_t^1$  and the best ask price is  $P_t^2$ .
- The spread is  $S_t = P_t^2 - P_t^1$  and  $\alpha_0$  is the tick size.
- The available quantity at the best bid (resp. ask) is  $Q_t^1$  (resp.  $Q_t^2$ ).

## 2.2 Order book dynamics

Let  $(\Omega, \mathcal{F})$  be a measurable space and  $(T_n)_{n \geq 1}$  a non-decreasing sequence of random variables such that  $T_n < T_{n+1}$  on the event  $\{T_n < \infty\}$ . We associate to each  $T_n$  a random variable  $X_n$  taking its value on a measurable space  $(E, \mathcal{E})$ . In our case,  $T_n$  is the time when events

#### IV. From asymptotic properties of general point processes to the ranking of financial agents

happen in the order book and  $X_n$  are variables describing each event. We endow  $\Omega$  with the filtration  $(\mathcal{F}_t)_{t \geq 0}$  defined as  $\mathcal{F}_t = \sigma(\{T_n \in C\} \times \{X_n \in B\}, C \in \mathcal{B}(\mathbb{R}) \cap (-\infty, t], B \in \mathcal{E})$ . Each event is characterised by:

- **The size of the order:** is an integer representing the minimum quantity that can be inserted in the order book.<sup>3</sup>
- **The price of the order:** equals to  $k \in \mathbb{N}$  when the order is inserted at the price  $P^1 + k\alpha_0$ .
- **The direction of the order:** equals to  $+1$  if it provides liquidity and  $-1$  when liquidity is removed.
- **The type of the order:** equals to  $1$  (resp.  $2$ ) when it modifies the bid (resp. ask)<sup>4</sup> side.
- **The identity of the agent:** is valued in  $\{1, \dots, N\}$  since the market consists in  $N$  agents.

Since we track only the first limits, we add the following variables to describe the new order book state when one of these limits is depleted:  $\tilde{Q}^1$  (resp.  $\tilde{Q}^2$ ) the new bid (resp. ask) queue and  $\tilde{S}$  the new spread after a depletion. Note that when there is no depletion, the random vector  $(\tilde{Q}^1, \tilde{Q}^2, \tilde{S})$  is arbitrary<sup>5</sup> and its values are not used. Finally, we record the order book state after an event to add a dependence between the arrival rate of the events and the past order book states. The order book dynamics are described below. Hence, we consider the following form for  $E = \bar{\mathbb{N}} \times \mathbb{T} \times \mathbb{S} \times \mathbb{B} \times \tilde{\mathbb{U}} \times \mathbb{U} \times \mathbb{A}$  with:

- $\bar{\mathbb{N}} = \mathbb{N}^*$ : the set where the orders size is valued.
- $\mathbb{T} = \mathbb{N}$ : the set where the price levels are valued.
- $\mathbb{S} = \{+1, -1\}$ : the set where the orders direction is valued.
- $\mathbb{B} = \{1, 2\}$ : the set where the orders type is valued.
- $\tilde{\mathbb{U}} = \{\mathbb{N}^2 \times \alpha_0 \mathbb{N}\} \setminus \mathbb{U}^0$ : the set where the order book states after a depletion are valued.<sup>6</sup>
- $\mathbb{U} = \{\mathbb{N}^2 \times \alpha_0 \mathbb{N}\} \setminus \mathbb{U}^0$ : the set where the order book states after an event are valued.
- $\mathbb{A} = \{1, \dots, N\}$ : the set where the agents identity is valued.
- $\mathbb{U}^0 = \{0\}^2 \times \alpha_0 \mathbb{N}$ : the set of unreachable order book states.

**Example 2.** *We place ourselves in the case where the minimum order size is a quarter of the AES and  $(\tilde{Q}^1, \tilde{Q}^2, \tilde{S}) = c$  when there is no depletion with  $c$  is a fixed constant. Thus, a buy limit order of size  $0.5$  AES inserted at the best bid price  $+1$  tick by the agent  $5$  when the best bid size is  $Q_i^1 = 1$  AES, the best ask size is  $Q_i^2 = 3$  AES and the spread  $S = 2$  ticks is represented by the event  $e = (2, 1, +1, 1, c, u, 5)$  with  $u = (2, 12, 1)$ .*

<sup>3</sup>In practice, the minimum quantity can be taken as a quarter of the the average event size (AES).

<sup>4</sup>A buy (sell) limit order within the spread, a liquidity consumption at the bid (resp. ask) or a limit order at the bid (resp. ask) modify the bid side first.

<sup>5</sup>To fix the ideas we can take  $(\tilde{Q}^1, \tilde{Q}^2, \tilde{S}) = c$  with  $c$  a fixed constant when there is no depletion.

<sup>6</sup>The state where the best bid or ask size is zero is fictitious state that allow us to model the price changes, see Remark 5.8.

**Order book dynamics.** The order book state is modelled by the process  $U_t = (Q_t^1, Q_t^2, S_t)$  where  $Q_t^1$  (resp.  $Q_t^2$ ) is the best bid (resp. ask) quantity and  $S_t$  is the spread. The dynamics of the reference price are going to be deduced from the one of the process  $(U_t)_{t \geq 0}$ , see Section 4. The process  $U_t$  is defined in the following way:

$$U_t = \sum_{T_i < t} \Delta U_i, \quad \Delta U_i = U_i - U_{i-1},$$

with  $U_i = (Q_i^1, Q_i^2, S_i) \in \mathbb{U}$  the order book state after the  $i$ -th event (we write  $U_i$  for  $U_{T_i}$  when no confusion is possible). Thus, we only need to describe the variables  $(U_i)_{i \geq 1}$ . Let  $i \geq 1$  and  $X_i = (n_i, t_i, s_i, b_i, \tilde{U}_i, U_i, a_i) \in E$  with  $n_i \in \mathbb{N}$ ,  $t_i \in \mathbb{T}$ ,  $s_i \in \mathbb{S}$ ,  $b_i \in \mathbb{B}$ ,  $\tilde{U}_i = (\tilde{Q}_i^1, \tilde{Q}_i^2, \tilde{S}_i) \in \mathbb{U}$ ,  $U_i = (Q_i^1, Q_i^2, S_i) \in \mathbb{U}$  and  $a_i \in \mathbb{A}$ . The variable  $U_i$  satisfies

$$\begin{aligned} S_i &= \mathbf{1}_{\epsilon_i=0} S_{i-1} - (t_i^1 + t_i^2) + \mathbf{1}_{\epsilon_i=1} \tilde{S}_i, \\ Q_i^1 &= \mathbf{1}_{\epsilon_i=0} Q_{i-1}^1 + (n_i^{1,+} - n_i^{1,-} + n_i^{1,\frac{1}{2}}) + \mathbf{1}_{\epsilon_i=1} \tilde{Q}_i^1, \\ Q_i^2 &= \mathbf{1}_{\epsilon_i=0} Q_{i-1}^2 + (n_i^{2,+} - n_i^{2,-} + n_i^{2,\frac{1}{2}}) + \mathbf{1}_{\epsilon_i=1} \tilde{Q}_i^2, \end{aligned} \quad (1)$$

where  $\epsilon_i$  is a price move indicator (i.e.  $\epsilon = 0$  when there is no depletion and  $\epsilon = 1$  otherwise), the variable  $t_i^1$  (resp.  $t_i^2$ ) is the spread variation when a buy (resp. sell) limit order is inserted within the spread. The variables  $n_i^{1,+}$  (resp.  $n_i^{2,+}$ ),  $n_i^{1,-}$  (resp.  $n_i^{2,-}$ ) and  $n_i^{1,\frac{1}{2}}$  (resp.  $n_i^{2,\frac{1}{2}}$ ) are respectively the best bid (resp. ask) increments when a buy limit order is inserted at the best bid (resp. ask), when a consumption order is sent at the best bid (resp. ask) and when a buy (resp. sell) limit order is inserted within the spread. We now explain how the previous quantities can be written in terms of the state variables:

$$\begin{aligned} \epsilon_i &= \mathbf{1}_{\{s_i=-1\} \cap (\{b_i=1, n_i \geq Q_{i-1}^1\} \cup \{b_i=2, n_i \geq Q_{i-1}^2\})}, \\ t_i^1 &= \min(t_i \alpha_0, S_{i-1} - \alpha_0) \mathbf{1}_{\{b_i=1, t_i \neq 0\}}, \\ t_i^2 &= (S_{i-1} - t_i \alpha_0) + \mathbf{1}_{\{b_i=2, t_i \neq \frac{S_{i-1}}{\alpha_0}\}}, \\ n_i^{1(2),+} &= n_i \mathbf{1}_{\{s_i=+1, t_i=0(\frac{S_{i-1}}{\alpha_0}), b_i=1(2)\}}, \\ n_i^{1(2),-} &= n_i \mathbf{1}_{\{s_i=-1, t_i=0(\frac{S_{i-1}}{\alpha_0}), b_i=1(2), n_i < Q_{i-1}^{1(2)}\}}, \\ n_i^{1(2),1/2} &= n_i \mathbf{1}_{\{s_i=+1, t_i \notin \{0, \frac{S_{i-1}}{\alpha_0}\}, b_i=1(2)\}}. \end{aligned}$$

We denote by  $\lambda_t$  the intensity of the point process  $(T_n, X_n)$ . For  $e \in E$ ,  $\lambda_t(e)$  corresponds to the arrival rate of an event of type  $e$  conditional on the past history of the process and it is defined informally as

$$\lambda_t(e) = \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\#\{T_n \in (t, t + \delta t], X_n = e\} \geq 1 | \mathcal{F}_t]}{\delta t},$$

with  $\#A$  is the cardinality of the set  $A$ . We consider the following expression for the intensity:

$$\lambda_t(e) = \psi(e, U_{t-}, t, \sum_{T_i < t} \phi(e, U_{t-}, t - T_i, X_i)), \quad (2)$$

#### IV. From asymptotic properties of general point processes to the ranking of financial agents

where  $\psi$  and  $\phi$  are  $\mathbb{R}_+$ -valued functions. The individual behaviour of each agent is encoded in the functions  $\psi$  and  $\phi$  through  $e$  and  $(X_i)_{i \geq 1}$ , see Equation (2).

Note that we can recover the full definition of the intensity of the process  $N = (T_n, X_n)$  using the following proposition:

**Proposition 2.1.** *For any  $B \in \mathcal{E}$  and  $t \in \mathbb{R}_+$ , we have*

$$\lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\#\{T_n \in (t, t + \delta t], X_n \in B\} \geq 1 | \mathcal{F}_t]}{\delta t} = \sum_{e \in B} \lambda_t(e). \quad (3)$$

The proof of Proposition 2.1 is given in Appendix IV.A. The existence and the uniqueness of a probability measure  $\mathbb{P}$  on the filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t)$  such that (3) is satisfied and  $\lambda_t$  verifies Equation (2) is ensured as soon as  $\sum_{e \in E} \lambda_t(e)$  is locally integrable, see [93]. We prove that  $\sum_{e \in E} \lambda_t(e)$  is locally integrable in Appendix IV.C.

### 2.3 Market reconstitution

We can recover the market intensity  $\lambda_t^M$  using the corollary below.

**Corollary 2.1.** *When  $\lambda_t$  verifies Equation (2), the market intensity  $\lambda_t^M(e')$  of an event  $e'$  ( $e'$  does not contain the identity of the agent) in the exchange is given by*

$$\lambda_t^M(e') = \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\#\{T_n \in (t, t + \delta t], X_n \in (e', \mathbb{A})\} \geq 1 | \mathcal{F}_t]}{\delta t} = \sum_{a \in \mathbb{A}} \lambda_t((e', a)), \quad (4)$$

for any  $e' \in E' = \bar{\mathbb{N}} \times \mathbb{T} \times \mathbb{S} \times \mathbb{B} \times \tilde{\mathbb{U}} \times \mathbb{U}$ .

The proof of Corollary 2.1 is a consequence of Proposition 2.1.

### 2.4 Some specific models

**Poisson intensity.** We introduce here a simple version of the Poisson intensity model where the variable  $X_n = (n_n, t_n^o, s_n, b_n, \tilde{U}_n, U_n, a_n)$  with  $U_n = (Q_n^1, Q_n^2, S_n)$  satisfies

- the order size  $n_n = 1$ : all the events have the same size 1 AES.
- the price level  $t_n^o \in \{0, \frac{S_n}{\alpha_0}\}$ : orders are inserted at the best bid or ask.
- the law of  $\tilde{U}_n$  is unchanged: when one limit is depleted, the new state is drawn from the stationary distribution of the order book.

For any  $e = (n, t^o, s, b, \tilde{u}, u, a) \in E$  with  $u = (Q^1, Q^2, S)$ , we can recover Poisson models by taking the following choice of the parameters:

$$\psi(e, u, t, z) = \tilde{h}(s, b, a) \mathbf{1}_{n=1, t^o \in \{0, \frac{S}{\alpha_0}\}}, \quad \forall z, t \in \mathbb{R}_+,$$

with  $\tilde{h}$  a deterministic function valued on  $\mathbb{R}_+$ . Thus, the expression of the intensity becomes

$$\lambda_t(e) = \tilde{h}(s, b, a) \mathbf{1}_{n=1, t^o \in [0, \frac{s}{a_0}]}$$

Such modelling was introduced in [2, 55, 144].

**Queue-reactive intensity.** In the Queue-reactive model, the arrival rate of the events depends only on the current order book state. For any  $e \in E$  and  $u \in U$ , we take

$$\psi(e, u, t, z) = \tilde{h}(e, u), \quad \forall z, t \in \mathbb{R}_+,$$

to reproduce the Queue-reactive dynamics with  $\tilde{h}$  a deterministic function valued on  $\mathbb{R}_+$ . Hence, the intensity reads

$$\lambda_t(e) = \tilde{h}(e, u).$$

Such modelling was studied in [90, 91].

**Hawkes Queue-reactive intensity.** In the Hawkes framework, the arrival rate of each event depends fully on all the past market events. For any  $e \in E$  and  $u \in U$ , we generate the Hawkes Queue-reactive dynamics by taking

$$\psi(e, u, t, z) = h(e, u, t) + z, \quad \forall z, t \in \mathbb{R}_+.$$

Thus intensity has the following expression

$$\lambda_t(e) = h(e, U_{t-}, t) + \sum_{T_i < t} \phi(e, U_{t-}, t - T_i, X_i).$$

Close modelling was used [3, 14, 95, 126, 136].

**Quadratic Hawkes process.** The quadratic Hawkes processes generalise the linear Hawkes processes by adding an interaction term between the pairs of past events. In the classical one-dimensional case, the intensity function of a quadratic Hawkes process reads

$$\lambda_t(e) = h(t) + \sum_{T_i < t} \phi(t - T_i) + \sum_{T_i, T'_i < t} K(t - T_i, t - T'_i),$$

with  $K : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$  the quadratic kernel. We can recover a simple case of the quadratic Hawkes models when  $K$  is separable (i.e.  $K(t, s) = k(t)k(s)$  with  $k$  a non negative function) by taking  $\psi$  of the following form:

$$\psi(e, u, t, z) = h(e, u, t) + z^2, \quad \forall z, t \in \mathbb{R}_+.$$

Hence, the expression of the intensity becomes

$$\lambda_t(e) = h(e, U_{t-}, t) + \sum_{T_i < t} \phi^2(e, U_{t-}, t - T_i, X_i) + \sum_{T_i, T'_i < t} \phi(e, U_{t-}, t - T_i, X_i) \phi(e, U_{t-}, t - T'_i, X'_i).$$

Quadratic Hawkes models were introduced in [34, 132].

**Remark 2.1.** *In our modelling, the linear term is necessarily  $\phi^2$ . However, to overcome this limitation we can add a new argument to the function  $\psi$  which differentiates the linear kernel from the quadratic one. This will not modify the proofs.*

### 3 Ergodicity

#### 3.1 Notations and definitions

Let  $Z_t$  be a process defined on the probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  and valued in  $(W_0, \mathcal{W}_0)$ . We consider another process  $V_t$  defined on  $(W_0, \mathcal{W}_0)$  and valued in  $(X, \mathcal{X})$  and we denote by  $P_t(x, \cdot)$  the probability distribution of  $V_t^{0,x}$  starting at 0 with the initial condition  $x \in W_0$ . For any measure  $\mu$  defined on  $(W_0, \mathcal{W}_0)$  viewed as a random starting condition, we denote by  $P_t(\mu, \cdot) = \int_{W_0} P_t(x, \cdot) \mu(dx)$ .

**Definition 3.1** (Invariant distribution). *The measure  $\mu$  is invariant if the probability distribution  $P_t(\mu, \cdot)$  does not depend on the time  $t$ .*

This definition is consistent with the one given in [41, 86, 123]. The process  $V_t$  starting with the initial distribution  $\mu$  is stationary if and only if  $\mu$  is invariant. We define the total variation distance between two measures  $\pi$  and  $\pi'$  such that  $\|\pi - \pi'\|_{TV} = \sup_{A \in \mathcal{X}} |\pi(A) - \pi'(A)|$ .

**Definition 3.2** (Ergodicity). *Let  $C \in \mathcal{W}_0$ . The process  $V_t$  is  $C$ -ergodic if there exists an invariant measure  $\mu$  such that for any  $x \in C$ , we have  $P_t(x, \cdot) \xrightarrow[t \rightarrow \infty]{} P_0(\mu, \cdot)$  in total variation.*

**Remark 3.1.** *This definition is consistent with the one given in [123]. Ergodicity is interesting since it ensures the convergence of the order book process  $U_t$  towards an invariant probability distribution. Thus the stylized facts observed on market data can be explained by a law of large numbers type phenomenon for this invariant distribution.*

**Remark 3.2.** *In this section, we work with a continuous time processes  $Z_t$  and  $V_t$  with  $t \in \mathbb{R}_+$ . However, all the definitions are similar for a discrete time processes  $Z_n$  and  $V_n$  with  $n \in \mathbb{N}$ . We just have to replace  $t$  by  $n$  in the definitions above.*

The space  $\Omega$  and the filtration  $\mathcal{F}_t$  considered here are defined in Section 2.2,  $\mathcal{F} = \mathcal{F}_\infty$ , the filtered space  $W_0$  is the space of sequences indexed by  $\mathbb{N}^-$  and valued on  $\mathbb{R}_+ \times E$ ,  $X = \cup \times (\mathbb{R}_+)^E$  and  $\mathcal{X} = \mathcal{U} \times \mathcal{B}(\mathbb{R}_+)^{\otimes E}$  with  $\mathcal{U}$  the  $\sigma$ -algebra generated by the discrete topology on  $\cup$ ,  $\mathcal{B}(\mathbb{R}_+)^{\otimes E}$  the cylinder  $\sigma$ -algebra for  $(\mathbb{R}_+)^E$ ,  $\mathcal{B}(\mathbb{R}_+)$  the borel  $\sigma$ -algebra of  $\mathbb{R}_+$  and  $\mathcal{W}_0 = (\mathcal{B}(\mathbb{R}_+) \times \mathcal{E})^{\otimes \mathbb{N}^-}$  with  $\mathcal{E}$  the  $\sigma$ -algebra generated by the discrete topology on  $E$ . We need to work on the space  $W_0$  since the dynamics of the process depend on its whole past.

#### 3.2 Ergodicity

In this section, we assume that the process  $\bar{U}_t$  is irreducible and provide under general assumptions a theoretical result on the ergodicity of the process  $\bar{U}_t = (Q_t^1, Q_t^2, S_t, \lambda_t)$  with  $\lambda_t$  the intensity defined by (2).

We denote by  $\lambda_Q^{i,+}$  (resp.  $\lambda_Q^{i,-}$ ) and  $\lambda_S^+$  (resp.  $\lambda_S^-$ ) the arrival rate of the events that respectively increase (resp. decrease) the limit  $Q^i$  and the spread  $S$  for any  $i \in \mathbb{B}$ . Let  $U_t = (Q_t^1, Q_t^2, S_t)$  be the order book process and  $e \in E$  be a market event, the quantities  $\lambda_Q^{i,\pm}$  and  $\lambda_S^\pm$  are defined

by the following formulas:

$$\lambda_Q^{i,\pm}(U_{t^-}, n) = \sum_{e \in E_Q^{i,\pm}(U_{t^-}, n)} \lambda_t(e), \quad \lambda_S^\pm(U_{t^-}, k) = \sum_{e' \in E_S^\pm(U_{t^-}, k)} \lambda_t(e), \quad (5)$$

with  $n \in \mathbb{N}$ ,  $k \in \mathbb{N}$  and

$$\begin{aligned} E_Q^{i,\pm}(U_{t^-}, n) &= \{e \in E; \text{s.t. } \Delta Q_t^i = \pm n\}, \\ E_S^\pm(U_{t^-}, k) &= \{e \in E; \text{s.t. } \Delta S_t = \pm k\}, \end{aligned} \quad (6)$$

with  $\Delta X_t = X_t - X_{t^-}$  for any process  $X_t$ . For simplicity and since there is no ambiguity, we do not write the dependence of  $\lambda_Q^{i,\pm}$  and  $\lambda_S^\pm$  on the current time  $t$ . For any  $n \in \mathbb{N}^*$ , we write

$$\mathcal{P}(n) = \{\mathbf{k}_m = \{k_1, \dots, k_m\} \in (\mathbb{N}^*)^m; \text{ s.t. } k_1 + \dots + k_m = n, \quad m \in \mathbb{N}^*\},$$

for the set containing all the partitions of  $n$ .

**Assumption 3.1** ( $\psi$  growth). *We assume that there exist  $c \geq 0$ ,  $d \geq 0$  and  $n_\psi \in \mathbb{N}$  such that*

$$\begin{aligned} \tilde{\psi}(e, z) &\leq c(e) + d(e)z^{n_\psi}, \\ \sup_{e \in E} \left\{ d(e) \sum_{\mathbf{k}_m \in \mathcal{P}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_m} \int_{\mathbb{R}_+^m} \prod_{i=1}^m \phi^{*k_i}(e, s_i) ds_i \right\} &< 1, \end{aligned}$$

with  $\tilde{\psi}(e, z) = \sup_{(u,t) \in \mathbb{U} \times \mathbb{R}_+} \psi(e, u, t, z)$ ,  $\phi^*(e, s) = \sup_{u \in \mathbb{U}} \sum_{x \in E} \phi(e, u, s, x)$  and  $\binom{n_\psi}{\mathbf{k}_m} = \binom{n_\psi}{k_1, \dots, k_m} = \frac{n_\psi!}{k_1! \dots k_m!}$ .

Assumption 3.1 is natural. To see this, we take a 1-d stationary non-linear Hawkes process  $N_t$  with an intensity  $\lambda_t$  that verifies

$$\lambda_t = c + d \left( \sum_{T_i < t} \phi(t - T_i) \right)^{n_\psi} = c + d \left( \int_{-\infty}^t \phi(t - s) dN_s \right)^{n_\psi}, \quad \forall t \in \mathbb{R}_+.$$

By stationarity, we have

$$\begin{aligned} \bar{\lambda} = \mathbb{E}[\lambda_t] &= c + d \mathbb{E} \left[ \left( \int_{-\infty}^t \phi(t - s) dN_s \right)^{n_\psi} \right] \\ &= c + d \left\{ \sum_{\mathbf{k}_m \in \mathcal{P}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_m} \int_{(-\infty, t)^m} \prod_{i=1}^m \phi^{k_i}(t - s_i) \mathbb{E}[dN_{s_1} \dots dN_{s_m}] \right\}, \end{aligned}$$

with  $\binom{n_\psi}{\mathbf{k}_m}$  an enumeration factor. In fact, if we have  $n_\psi$  possible events divided in  $m$  groups such that the  $j$ -th group is composed of  $k_j$  events, then the quantity  $\binom{n_\psi}{\mathbf{k}_m}$  counts the number of possible groups. Here each group represents the jumps that happen at the same time. Since the jumps have a unit size, the Brascamp-Lieb inequality ensures that  $\mathbb{E}[dN_{s_1} \dots dN_{s_m}] \leq \prod_{i=1}^m \mathbb{E}[dN_{s_i}]^{1/m} = \prod_{i=1}^m \mathbb{E}[dN_{s_i}]^{1/m} = \prod_{i=1}^m \mathbb{E}[\lambda_{s_i}]^{1/m} = \bar{\lambda}$  which leads to

$$\bar{\lambda} \leq c + q\bar{\lambda},$$

with  $q = d/m! \sum_{\mathbf{k}_m \in \mathcal{P}(n_\psi)} \binom{n_\psi}{\mathbf{k}_m} \int_{(\mathbb{R}_+)^m} \prod_{i=1}^m \phi^{k_i}(e, s_i) ds_i$ . The condition  $q < 1$  of Assumption 3.1 guarantees that  $\bar{\lambda}$  is finite.



**Remark 3.3.** *Non linear Hawkes process are studied mainly when the function  $\psi$  admits at most a linear growth (i.e.  $n_\psi \leq 1$ ). When  $n_\psi = 1$ , we recover the classical condition*

$$\sup_{e \in E} d(e) \left\{ \int_{\mathbb{R}_+} \phi^*(e, s) ds \right\} < 1.$$

When  $n_\psi = 2$ , Assumption 3.1 becomes

$$\sup_{e \in E} d(e) \left\{ \left( \int_{\mathbb{R}_+} \phi^*(e, s) ds \right)^2 + \int_{\mathbb{R}_+} \phi^*(e, s)^2 ds \right\} < 1.$$

**Assumption 3.2** (Negative drift). *There exist positive constants  $C_{bound}$ ,  $z_0 > 1$  and  $\delta$  such that*

$$\begin{aligned} \sum_{n \geq 0} (z_0^n - 1) (\lambda_Q^{i,+}(U_{t^-}, n) - \lambda_Q^{i,-}(U_{t^-}, n) \frac{1}{z_0^n}) &\leq -\delta, & \text{a.s. when } Q_{t^-}^i \geq C_{bound}, \\ \sum_{k \geq 0} (z_0^{\alpha_0 k} - 1) (\lambda_S^+(U_{t^-}, k) - \lambda_S^-(U_{t^-}, k) \frac{1}{z_0^{\alpha_0 k}}) &\leq -\delta, & \text{a.s. when } S_{t^-} \geq C_{bound}, \end{aligned} \quad (7)$$

for any  $i \in \mathbb{B}$  and  $U_t = (Q_t^1, Q_t^2, S_t) \in \mathbb{U}$  where  $\alpha_0$  is the tick size.

Assumption 3.2 ensures that both the size of the first limits and the spread tend to decrease when they become too large. Same kind of hypothesis are used in [90] and Chapter II but when the order book dynamics are Markov.

**Remark 3.4.** *In practice, Assumption 3.2 is verified when the following conditions are satisfied:*

$$\begin{aligned} \sum_{n \geq 0} (z_0^n - 1) (\psi_Q^{i,+}(u, n, t, z) - \psi_Q^{i,-}(u, n, t, z) \frac{1}{z_0^n}) &\leq -\delta, & \text{when } q^i \geq C_{bound}, \\ \sum_{n \geq 0} (z_0^{\alpha_0 k} - 1) (\psi_S^+(u, k, t, z) - \psi_S^-(u, k, t, z) \frac{1}{z_0^{\alpha_0 k}}) &\leq -\delta, & \text{when } s^i \geq C_{bound}, \\ \phi_Q^{i,+}(u, n, t, x) &\leq \phi_Q^{i,-}(u, n, t, x), & \text{when } q^i \geq C_{bound}, \\ \phi_S^+(u, k, t, x) &\leq \phi_S^-(u, k, t, x), & \text{when } s^i \geq C_{bound}, \\ \psi(e, u, t, z) &\text{ is non-decreasing in } z, & \text{when } q^i \geq C_{bound}, \\ \psi(e, u, t, z) &\text{ is non-decreasing in } z, & \text{when } s^i \geq C_{bound}, \end{aligned} \quad (8)$$

where  $u = (q^1, q^2, s) \in \mathbb{U}$ ,  $i \in \mathbb{B}$  and  $\psi_Q^{i,\pm}$ ,  $\psi_S^\pm$ ,  $\phi_Q^{i,\pm}$  and  $\phi_S^\pm$  are functions defined such that

$$\begin{aligned} \psi_Q^{i,\pm}(u, n, t, z) &= \sum_{e \in E_Q^{i,\pm}(u, n)} \psi(e, u, t, z), & \phi_{Q(S)}^{i,+}(u, n, t, x) &= \sup_{e \in E_Q^{i,+}(u, n)} \phi(e, u, t, x), \\ \psi_S^\pm(u, k, t, z) &= \sum_{e \in E_S^\pm(u, k)} \psi(e, u, t, z), & \phi_{Q(S)}^-(u, k, t, x) &= \inf_{e \in E_{Q(S)}^-(u, k)} \phi(e, u, t, x), \end{aligned}$$

with  $(n, k, t, z) \in \mathbb{N}^2 \times \mathbb{R}_+^2$ . Although Inequalities (7) and (8) are not equivalent, there is a large panel of functions that satisfy (8). A proof of this result is given Appendix IV.B.

**Assumption 3.3** (Bound on the overall flow). *We assume that there exist  $z_1 > 1$ ,  $M$  and  $\underline{\psi} > 0$  satisfying*

$$\begin{aligned} c^* &= \sum_{e \in E} c(e) < \infty, \\ \lambda^* &= \sum_{e \in E, \mathbf{k}_m \in \mathcal{D}(n_\psi)} d(e) / m! \binom{n_\psi}{\mathbf{k}_m} \int_{\mathbb{R}_+^{n_\psi}} \prod_{j=1}^m \phi^{*k_j}(e, s_j) ds_j < \infty, \\ Q_\infty^i &= \sum_{n \in \mathbb{N}} (z_1^n - 1) \mathbb{E}_x \left[ \lambda_Q^{i,+}(u, n) - \frac{\lambda_Q^{i,-}(u, n)}{z_1^n} \right] < M, & \text{when } q^i \leq C_{bound}, \\ S_\infty &= \sum_{k \in \mathbb{N}} (z_1^k - 1) \mathbb{E}_x \left[ \lambda_S^+(u) - \frac{\lambda_S^-(u, n)}{z_1^k} \right] < M, & \text{when } s \leq C_{bound}, \\ \underline{\lambda}_t(e) &= \sum_{e \in E} \lambda_t(e) \geq \underline{\psi}, & \text{a.s.} \end{aligned}$$

with  $c(e)$ ,  $d(e)$  and  $\phi^*$  defined in Assumption 3.1,  $i \in \mathbb{B}$ ,  $\mathbf{x} \in W_0$  and  $C_{\text{bound}}$  defined in Assumption (3.2). Similar assumptions are considered in [90] and Chapter II in the Markov case.

Assumption 3.3 ensures no explosion in the system since it forces the arrival rate of orders, the size of the limits and the spread to stay bounded.

**Remark 3.5.** *In practice, we can find pathwise conditions similar to those used in Remark 3.4 such that the inequalities  $Q_\infty^i < M$ ,  $S_\infty < M$  and  $\underline{\lambda}_t(e) \geq \bar{\psi}$ , a.s are satisfied.*

**Theorem 3.1** (Existence). *Under Assumptions 3.1, 3.2 and 3.3, the process  $\bar{U}_t = (Q_t^1, Q_t^2, S_t, \lambda_t)$  admits an invariant distribution.*

The proof of this result is given in Appendix IV.C.

**Assumption 3.4** (Regularity). *We assume that  $\psi$  is a càdlàg function continuous with respect to  $z$ ,  $\phi$  is a positive càdlàg function and there exist  $\bar{\psi} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and  $n_1 \in \mathbb{N}$  such that*

$$|\psi(e, u, s, x) - \psi(e, u, s, y)| \leq |\bar{\psi}(x) - \bar{\psi}(y)|, \quad \forall (e, u, s, x, y) \in E \times \mathbb{U} \times \mathbb{R}_+^3,$$

and

$$|\bar{\psi}(x) - \bar{\psi}(y)| \leq K|x - y|1 + x^{n_1} + y^{n_1}, \quad \forall (x, y) \in \mathbb{R}_+^2,$$

with  $K$  a positive constant.

**Remark 3.6.** *Assumption 3.4 is satisfied in the special case where  $\bar{\psi}$  is a polynomial.*

We have the following result.

**Theorem 3.2** (Ergodicity). *Under Assumptions 3.1, 3.2, 3.3 and 3.4, the process  $\bar{U}_t$  is  $\mathcal{W}_0$ -ergodic, which means that there exists an invariant measure  $\mu$ , see Definition 3.1, that satisfies*

$$\lim_{t \rightarrow \infty} P_t(x, A) = P_0(\mu, A), \quad \forall x \in W_0, A \in \mathcal{X},$$

where  $P_t(x, A)$  is the probability that  $\bar{U}_t \in A$  starting from the initial condition  $x$ . Additionally, we have the following speed of convergence:

$$\|P_t(x, \cdot) - P_0(\mu, \cdot)\|_{TV} \leq K_1 e^{-K_2 t}, \quad \forall x \in W_0,$$

with  $K_1, K_2$  are positive constants and  $\|\cdot\|_{TV}$  the total variation distance.

The proof of this result is given in Appendix IV.D. We can construct pathwise the point process  $N = (T_n, X_n)$  defined in Section 2 using the following algorithm.

**Remark 3.7** (Pathwise construction of  $N$ ). *Using the thinning algorithm proposed by Lewis in [112] and Ogata in [132], the point process  $N = (T_n, X_n)$  defined in Section 2 satisfies  $N = \lim_{m \rightarrow \infty} N^m$  where  $N^m$  is defined as follows*

$$\begin{aligned} \lambda_t^{m+1}(e) &= \psi(e, U_t^m, t, \sum_{T^m < t} \phi(e, U_{t^-}^m, t - T^m, X^m)) \mathbf{1}_{T^m \leq t < T^{m+1}} + \lambda_t^m(e) \mathbf{1}_{t < T^m}, \\ N^{m+1}((0, t] \times B) &= \int_{(T^m, T^{m+1}] \times B} N^*(dt \times (0, \lambda_t^{m+1}(e)] \times de) \mathbf{1}_{t > T^m} + N^m((0, t \wedge T^m] \times B), \\ T^{m+1} &= \sup\{t > T^m; \int_{(T^m, t] \times \mathcal{E}} N^*(dt \times (0, \lambda_t^m(e)] \times de) = 0\}, \end{aligned}$$

with  $U^m$  the order book process generated by  $N^m$  and described in (1),  $N^* = (T_n^*, R_n^*, X_n^*)$  a Poisson process valued on  $\mathbb{R}_+^2 \times E$  which admits  $\text{dtdz}\nu(\text{de})$  as an  $\mathcal{F}_t^{N^*}$  intensity and  $\nu = \sum_{e \in E} \delta_e$ .

This is a well known result that were used in many contexts, see [41, 59, 109, 112, 132]. The proof of Theorem 3.1 ensures that the above algorithm is well defined.

## 4 Limit theorems

Let  $n$  be the index of the  $n$ -th jump,  $(\eta_n)_{n \geq 0}$  be a process satisfying  $\eta_n = f((U_i)_{i \leq n}, (Y_i)_{i \leq n})$  with  $f$  a measurable function valued on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ,  $(Y_i)_{i \geq n}$  is a geometrically ergodic sequence, see 15.7 in [121], independent of  $(U_i)_{i \geq n}$ . Here, we write  $\bar{\mu}$  for the invariant measure of the joint process  $(\bar{U}, Y)^7$ ,  $V_n = \sum_{k=1}^n \eta_k$  and  $S_n = \sum_{k=1}^n (\eta_k - \mathbb{E}_{\bar{\mu}}[\eta_k])$ . We denote by

$$X_n(t) = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad \forall t \geq 0.$$

**Assumption 4.1.** *Under the invariant measure  $\bar{\mu}$ , the sequence  $(\eta_i)_{i \geq 0}$  is stationary and  $\mathbb{E}_{\bar{\mu}}[|\eta_0|] < \infty$ .*

**Assumption 4.2.** *Under the invariant measure  $\bar{\mu}$ , we have  $\mathbb{E}_{\bar{\mu}}[(\eta_0 - \mathbb{E}_{\bar{\mu}}[\eta_0])^2] < 1$ .*

**Proposition 4.1.** *Under Assumption 4.1, we have*

$$\frac{V_n}{n} \xrightarrow[n \rightarrow \infty]{} \mathbb{E}_{\bar{\mu}}[\eta_0], \quad a.s. \quad (9)$$

Moreover when both Assumptions 4.1 and 4.2 are verified, the quantity  $X_n(t)$  satisfies

$$X_n(t) \xrightarrow{\mathcal{L}} \sigma W_t, \quad (10)$$

with  $\sigma^2 = \mathbb{E}_{\bar{\mu}}[\eta_0^2] + 2 \sum_{k \geq 1} \mathbb{E}_{\bar{\mu}}[\eta_0 \eta_k]$  and  $\bar{\mu}$  the invariant measure of  $(U_i, Y_i)$  and  $W_t$  a standard brownian motion.

Note that  $\sigma^2 < \infty$  under Assumption 4.2. The proof of this result is given in Appendix IV.E.

**Remark 4.1.** *The leading term in the expression of  $\sigma^2$  is  $\mathbb{E}_{\bar{\mu}}[\eta_0^2]$ . Numerically, it can be computed as soon as we have an estimate of the stationary distribution of  $\eta_0$ , see Proposition 5.1.*

Proposition 4.1 ensures that the large scale limit of  $S$  in event time is a brownian motion. However, it is more relevant to study the large scale limit of the process  $S$  in calendar time. Thus we now consider the process

$$\tilde{X}_n(t) = \frac{S_{N(nt)}}{\sqrt{n}}, \quad \forall t \geq 0.$$

The following proposition provides the large scale limit of the process  $S_{N(nt)}$ .

<sup>7</sup>Recall that  $\bar{U}$  is the process  $\bar{U} = (Q^1, Q^2, S, \lambda)$ .

**Proposition 4.2.** *Under Assumption 4.1, we have*

$$\frac{V_{N(nt)}}{n} \xrightarrow[n \rightarrow \infty]{} \frac{\mathbb{E}_{\bar{\mu}}[\eta_0]}{\mathbb{E}_{\bar{\mu}}[\Delta T_1]}, \quad a.s. \quad (11)$$

Moreover when both Assumptions 4.1 and 4.2 are verified, the quantity  $\tilde{X}_n(t)$  satisfies

$$\tilde{X}_n(t) \xrightarrow{\mathcal{L}} \frac{\sigma}{\sqrt{\mathbb{E}_{\bar{\mu}}[\Delta T_1]}} W_t, \quad (12)$$

with  $\sigma^2 = \mathbb{E}_{\bar{\mu}}[\eta_0^2] + 2 \sum_{k \geq 1} \mathbb{E}_{\bar{\mu}}[\eta_0 \eta_k]$ ,  $\bar{\mu}$  the invariant measure of  $(U_i, Y_i)$ ,  $\Delta T_n = T_n - T_{n-1}$  the inter-arrival time between the  $n$ -th and  $(n-1)$ -th jump and  $W_t$  a standard brownian motion.

The proof of this result is given in Appendix IV.E.

**Remark 4.2.** *The mid price after  $n$  jumps  $P_n$  satisfies  $P_n = P_0 + \sum_{i=1}^n \Delta P_i$  with  $\Delta P_i = (P_i - P_{i-1}) = \eta_i$ . When  $(\eta_i)_{i \geq 0}$  verifies Assumptions 4.1 and 4.2, the rescaled price process  $\tilde{P}_n(t) = \frac{P_{N(nt)}}{\sqrt{n}}$  converges towards a Brownian diffusion.*

## 5 Formulas

In this section, we provide a calibration methodology for the intensities and computation formulas for the quantities of interest: the stationary distribution of the order book, the price volatility and the fluctuations of liquidity.

### 5.1 Stationary probability computation

In this section, we denote by  $\mu$  the invariant measure of  $\bar{U} = (Q^1, Q^2, S, \lambda)$  defined on  $(W_0, \mathcal{W}_0)$ , see Theorem 3.2. Let  $\zeta_t = f((U_i)_{T_i \leq t})$  be a stationary process under  $\mu$  with  $f$  a measurable function valued in  $(Z, \mathcal{Z})$ ,  $Z$  a countable space and  $\pi$  the stationary distribution of  $\zeta_t$ . The proposition below provides a fixed point formula satisfied by  $\pi$ .

**Proposition 5.1.** *The stationary distribution  $\pi$  satisfies*

$$\begin{aligned} \pi Q &= 0, \\ \pi \mathbf{1} &= 1. \end{aligned} \quad (13)$$

where the infinite dimensional matrix  $Q$  verifies

$$Q(z, z') = \sum_{e \in E(z, z')} \mathbb{E}_{\mu}[\lambda(e) | \zeta_0 = z], \quad (14)$$

with  $E(z, z')$  the set of events directly leading to  $z'$  from  $z$ .

The proof of this result is provided in Appendix IV.F.

**Remark 5.1.** *When  $\zeta_t = U_t = (Q_t^1, Q_t^2, S_t)$ , Proposition 5.1 provides a fixed point equation for the computation of the stationary distribution  $\pi$  of the order book.*

**Remark 5.2.** *The operator  $Q$  is the infinitesimal generator of the process  $\zeta$  defined such that  $Q(z, z') = \lim_{\delta \rightarrow 0} \frac{\mathbb{P}_\mu[\zeta_\delta = z' | \zeta_0 = z]}{\delta}$  for any  $z \neq z'$ . The proof of this result is given in Equation (61) of Appendix IV.F.*

### 5.1.1 Markov framework

In the Markov case, it is a well known result that  $Q$  satisfies (13), see [131]. In this case, the coefficients of  $Q$  are parameters of the model and can be estimated using (15).

### 5.1.2 General case

Let us take two states  $z$  and  $z'$  such that  $z \neq z'$ . We define  $N_t^{z, z'}$  and  $t^z$  such that

$$N_t^{z, z'} = \sum_{T_i < t} \delta_{z, z'}^i, \quad t^z = \sum_{T_i < t} \Delta T_i \mathbf{1}_{\{\zeta_{T_{i-1}} = z\}},$$

with  $\delta_{z, z'}^i = \mathbf{1}_{\{\zeta_{T_{i-1}} = z, \zeta_{T_i} = z'\}}$  and  $\Delta T_i = T_i - T_{i-1}$ . We have the following result.

**Proposition 5.2.** *When  $(\delta_{z, z'}^i)_{i \geq 1}$  satisfies Assumption 4.1, we have*

$$\hat{Q}(z, z') = \frac{N_t^{z, z'}}{t^z} \xrightarrow[t \rightarrow \infty]{} Q(z, z'), \quad a.s. \quad (15)$$

The proof of this result is given in Appendix IV.G.

**Remark 5.3** (Confidence interval). *We can compute a confidence interval for the estimator  $\hat{Q}(z, z')$ , see Appendix IV.G for the details.*

**Remark 5.4.** *When  $\zeta_t = U_t = (Q_t^1, Q_t^2, S_t)$ , Proposition 5.2 provides an estimator for the operator  $Q(u, u')$  with  $u, u' \in \mathbb{U}$  and  $u \neq u'$ .*

**Remark 5.5.** *In the Markov case and when  $\zeta_t = U_t$ , see [91], the authors used the estimator presented in Proposition 5.2 to evaluate  $Q(u, u')$ .*

**Remark 5.6.** *Let  $(z, z') \in \mathbb{U}^2$  such that  $z \neq z'$  and  $a \in \mathbb{A}$ , we consider the quantity  $Q(z, z', a) = \sum_{e \in E(z, z') \cap E(a)} \mathbb{E}[\lambda(e) | \zeta_0 = z]$  with  $E(a)$  the set of events generated by the agent  $a$ . This quantity represents the infinitesimal probability that the agent  $a$  sends an order that moves  $\zeta$  from  $z$  to  $z'$ . It can be estimated by  $\hat{Q}(z, z', a) = N_t^{z, z', a} / t^z$  which satisfies*

$$\hat{Q}(z, z', a) = \frac{N_t^{z, z', a}}{t^z} \xrightarrow[t \rightarrow \infty]{} Q(z, z', a), \quad a.s., \quad (16)$$

with  $N_t^{z, z', a} = \sum_{T_i < t} \delta_{z, z', a}^i$  and  $\delta_{z, z', a}^i = \mathbf{1}_{\{\zeta_{T_{i-1}} = z, \zeta_{T_i} = z', A_i = a\}}$  where  $A_i$  is the identity of the agent causing the  $i$ -th event. The quantity  $Q(z, z', a)$  allows us to infer the market dynamics (i.e. the operator  $Q$ ) for a specific combination of the agents, see Equation (14).

## 5.2 Spread computation

Since the process  $U_t$  is ergodic the spread  $S_t$  has a stationary distribution. Then, we can compute  $\mathbb{E}_\pi[S_\infty]$  where  $\pi$  is the stationary distribution of  $U$ . The computation formula for  $\pi$  is detailed in Proposition 5.1 and the estimation methodology of  $Q$  is described in Proposition 5.2.

## 5.3 Price volatility computation

We place ourselves in the case of Remark 4.2 and assume that the mid price moves  $(\eta_i)_{i \geq 0}$  are valued in  $\zeta = \alpha_0 \mathbb{Z}$  with  $\alpha_0$  the tick size. In such situation, the limit theorem of Section 4 ensures the convergence of  $\bar{P}_n(t)$  towards

$$\bar{P}_n(t) \xrightarrow{\mathcal{L}} \sigma W_t,$$

with  $\sigma^2 = \mathbb{E}_{\bar{\mu}}[\eta_0^2] + 2 \sum_{k \geq 1} \mathbb{E}_{\bar{\mu}}[\eta_0 \eta_k]$  and  $\bar{\mu}$  the invariant measure of  $(\bar{U}, Y)$ , see Section 4. The quantity of interest is  $\sigma^2$ . To compute  $\sigma^2$ , we need to evaluate  $\mathbb{E}_{\bar{\mu}}[\eta_0 \eta_k]$  for all  $k \geq 0$ . We have

$$\begin{aligned} \mathbb{E}_{\bar{\mu}}[\eta_0^2] &= \sum_{\eta \in \zeta} \pi_{\eta_0}(\eta) \eta^2, \\ \mathbb{E}_{\bar{\mu}}[\eta_0 \eta_k] &= \sum_{\eta \in \zeta} \pi_{\eta_0}(\eta) \eta \mathbb{E}_{\bar{\mu}}[\eta_k | \eta_0 = \eta], \quad \forall k \geq 1, \end{aligned} \tag{17}$$

with  $\pi_{\eta_0}(\eta) = \mathbb{P}_{\bar{\mu}}[\eta_0 = \eta]$ . Thus we need to estimate  $\pi_{\eta_0}$  and  $\mathbb{E}_{\bar{\mu}}[\eta_k | \eta_0 = \eta]$  to evaluate  $\sigma^2$ . The computation of the leading term  $\mathbb{E}_{\bar{\mu}}[\eta_0^2]$  requires only the knowledge of the stationary distribution  $\pi_{\eta_0}$ . The latter is evaluated using Proposition 5.1. To estimate  $\mathbb{E}_{\bar{\mu}}[\eta_k | \eta_0 = \eta]$  with  $k \geq 1$ , we use the following proposition.

**Proposition 5.3.** *Let us take  $k \geq 1$ ,  $\eta \in \zeta$  and*

$$N_n^{\eta, (k)} = \sum_{j \leq n} \eta_j \delta_\eta^{j(k)},$$

with  $\delta_\eta^{j(k)} = \mathbf{1}_{\{\eta_{j-k} = \eta\}}$  and  $n^\eta = \sum_{j \leq n} \delta_\eta^{j(k)}$ . When both  $(\eta_i \delta_\eta^{i(k)})_{i \geq 1}$  and  $(\delta_\eta^{i(k)})_{i \geq 1}$  satisfy Assumption 4.1, we have

$$\hat{E}(\eta_0, k) = \frac{N_n^{\eta_0, (k)}}{n^\eta} \xrightarrow{n \rightarrow \infty} \mathbb{E}_{\bar{\mu}}[\eta_k | \eta_0 = \eta], \quad a.s. \tag{18}$$

The proof of this result is similar to the one of Proposition 5.2.

**Remark 5.7** (Markov case). *When the dynamics of  $U$  are Markov and  $\eta_i = f_0(U_i)$  for any  $i \geq 0$  with  $f_0$  a deterministic function, see Remark 5.8. We have*

$$\mathbb{E}_\pi[\eta_0 \eta_k] = \sum_{u \in \mathbb{U}} \pi(u) \eta_0(u) \mathbb{E}_u[\eta_k], \tag{19}$$

where  $\pi$  is the stationary distribution of  $U$  that can be computed using Proposition 5.1 and  $\mathbb{E}_u[\eta_k] = (P^k * \eta_0)_u = \sum_{u' \in \mathbb{U}} P_{u,u'}^k \eta_0(u')$  with  $P^k$  the  $k$ -th power of the Markov chain  $P$  associated to the process  $U$  and which satisfies

$$\begin{aligned} P_{u,u'} &= \begin{cases} -Q_{u,u'} / Q_{u,u}, & \text{if } u \neq u' \text{ and } Q_{u,u} \neq 0, \\ 0, & \text{if } u \neq u' \text{ and } Q_{u,u} = 0, \end{cases} \\ P_{u,u} &= \begin{cases} 0, & \text{if } Q_{u,u} \neq 0, \\ 1, & \text{if } Q_{u,u} = 0, \end{cases} \end{aligned} \quad (20)$$

where the quantity  $P_{u,u'}$  represents  $P_{u,u'} = \mathbb{P}[U_1 = u' | U_0 = u]$  with  $U_1$  the state of the order book after one jump.

**Remark 5.8.** In Section 6, for any  $u = (q^1, q^2, s)$ , we consider the following function:

$$f_0(u) = \begin{cases} -1, & \text{if } q^1 = 0 \text{ and } q^2 > 0, \\ +1, & \text{if } q^2 = 0 \text{ and } q^1 > 0, \\ 0, & \text{otherwise,} \end{cases}$$

for the numerical simulations. Note that the states where  $q^1 = 0$  or  $q^2 = 0$  are fictitious states that are not observable in practice. These states are introduced to handle the price changes. Indeed, the states where  $q^1 = 0$  (resp.  $q^2 = 0$ ) correspond to a price decrease (resp. increase) by one tick and the states where both  $q^1 = 0$  and  $q^2 = 0$  are unreachable.

#### 5.4 An alternative measure of market stability

Another way to look at market stability is to investigate the behaviour of the disequilibrium between offer and demand. This equilibrium can be for example measured through the cumulative imbalance  $N_t = V_t^b - V_t^a$  where  $V_t^b$  (resp.  $V_t^a$ ) is the net number of inserted limit orders at the bid (resp. ask). From no arbitrage argument, we know that the dynamics of  $N_t$  are closely related to that of the price [95, 100]. Consequently, it is natural to view the long term volatility of this object as an alternative measure of market stability.

In this section, we follow the same methodology used in Section 5.3. The cumulative imbalance after  $n$  jumps  $N_n$  satisfies  $N_n = N_0 + \sum_{i=1}^n \Delta N_i$  where  $\Delta N_i = N_i - N_{i-1} = n_i$ . Hence, when  $(n_i)_{i \geq 0}$  satisfies Assumptions 4.1 and 4.2, we have the following convergence result:

$$X_n^N = \frac{\sum_{k=1}^n (n_k - \mathbb{E}_u[n_k])}{\sqrt{n}} \xrightarrow{\mathcal{L}} \tilde{\sigma} W_t,$$

with  $\tilde{\sigma}^2 = \mathbb{E}_\mu[n_0^2] + 2 \sum_{k \geq 1} \mathbb{E}_\mu[n_0 n_k]$  and  $\mu$  the stationary distribution of  $\bar{U}$  given by Theorem 3.1. The quantity  $\mathbb{E}_\mu[n_0 n_k]$  can be computed using the same methodology of Section 5.3.

## 6 Numerical experiments

In this section, we propose a ranking of the market makers for four different assets, based on their impact on volatility. For each asset, we compute first the liquidity provision and consumption intensities relative to the whole market using Equation (15).<sup>8</sup> Then, we estimate the stationary measure of the order book, see Equation (13), and use it to compute the two following estimators of the market volatility:

$$\begin{aligned}\sigma^{2,G} &= \mathbb{E}_\mu[\eta_0^2], \\ \sigma_k^{2,M} &= \mathbb{E}_\pi[\eta_0^2] + 2 \sum_{j=1}^k \mathbb{E}_\pi[\eta_0 \eta_j],\end{aligned}$$

where  $\mu$  is the invariant measure of  $\bar{U}$  given by Theorem 3.2,  $\pi$  is the stationary distribution of  $U$  when the order book dynamics are Markov and  $\eta_i = f_0(U_i)$  with  $f_0$  defined in Remark 5.8. The estimator  $\sigma^{2,G}$  is computed by applying Equation (17) and  $\sigma_k^{2,M}$  is evaluated using Remark 5.7. Thereafter, for each market maker, we compute its own intensities using Equation (16). After that, we estimate the new market intensities in a situation where we suppose that he withdraws from the exchange by subtracting the agent intensity from the market one, see Corollary 2.1. We finally compute the new market volatility estimators  $\sigma^{2,G}$  and  $\sigma_k^{2,M}$  corresponding to this new scenario using Equation (17) and Remark 5.7 again.

**Remark 6.1.** *In the simple case where the order book dynamics are Markov and the queues are independent, see Section 2.3.3 in [90], minimizing the first order approximation of the price volatility  $\sigma^2 \sim \mathbb{E}_\pi[\eta_0^2]$  is similar to selecting the agent with the highest ratio insertion/consumption  $\frac{\lambda_Q^{1(2),+}}{\lambda_Q^{1(2),-}}$ . This condition is a well-known result which means that the new agent needs to have an insertion/consumption ratio greater than the one of the market. The proof of this result is given in Section IV.H.*

**Remark 6.2.** *The reconstruction methodology of the market assumes that other participants will not modify their behaviours when an agent leaves the market. In practice, this assumption is satisfied since agents react to global variables such as the imbalance and not to a specific agent-based information. Additionally, when an agent leaves the market, the other participants do not have enough order flow history to calibrate all the parameters of their models.*

**Remark 6.3.** *The reconstruction methodology of the market takes into account the volume exchanged by each agent since this information is included in the estimated intensities. Indeed, the intensity of an agent who trades a large volume is high because he either interacts frequently with the market or generates significant changes in the order book state.*

### 6.1 Database description.

We study four large tick European stocks: Air Liquide, EssilorLuxottica, Michelin and Orange, on Euronext, over a year period: from January 2017 till December 2017. The data under study are provided by the French Regulator Autorité des marchés financiers. For each

<sup>8</sup>A liquidity provision (resp. consumption) event is assimilated to an increase (resp. decrease) of the best bid or ask size by 1 unit. To fix the ideas, one (AES) is our unit here.



#### IV. From asymptotic properties of general point processes to the ranking of financial agents

of these assets, we have access to the trades and orders data. Using both data, we rebuild the Limit Order book (LOB) up to the first limit of both sides, whenever an event (an order insertion, an order cancellation or an aggressive order) happens on one of these limits. Note that we remove market data corresponding to the first and last hour of trading, as these periods have usually specific features because of the opening/closing auction phases. We present in Table IV.1 some preliminary statistics on the different considered assets.

Asset	Number of insertion orders (in millions of orders)	Number of cancellation orders (in millions of orders)	Number of aggressive orders (in millions of orders)	Ratio of cancellation orders number over aggressive orders number	Average spread (in ticks)
Air Liquide	2.36	2.40	0.21	11.4	1.07
EssilorLuxottica	3.90	3.96	0.34	11.6	1.11
Michelin	3.81	4.01	0.32	12.5	1.14
Orange	6.60	6.66	0.47	14.1	1.14

Table IV.1 – Preliminary statistics on the assets.

Table IV.1 shows that the number of insertion orders is lower than that of cancellation orders. A priori, this seems contradictory, but what happens in practice is that some agents insert orders that they cancel partially and progressively at a later stage by sending multiple cancellation orders, which leads to a number of cancellation orders higher than that of insertion orders.

The considered market makers, that we aim at ranking, are the Supplemental Liquidity Providers (SLP) members. The SLP programme imposes a market making activity on programme members, including order book presence time at competitive prices. In return, they get favorable pricing and rebates in the form of a maker-taker fees model directly comparable to those of the major competing platforms. This programme includes 9 members. Some of them have at the same time SLP activity and other activities, such like proprietary or agency activity. In our analysis, we only analyse the SLP flow of these members. We denote the market makers by MM1 to MM9.

## 6.2 Computation of the intensities and the stationary measure

We compute the liquidity consumption and provision intensities at the first limit relative to the whole market according to the queue size, the corresponding stationary measure and the long term volatility for Air Liquide. Results relative to EssilorLuxottica, Michelin and Orange are relegated to Appendix IV.I. The estimation methodology of the intensities is based on Proposition 5.2. To apply this proposition, we record, for every event occurring in the LOB at the best limits (best ask and bid), the type of this order (insertion or consumption), the waiting time (in number of seconds) between this event and the preceding one occurring at the same limit and the queue size before the event. The queue size is then approximated by

the smaller integer that is larger than or equal to the volume available at the queue, divided by the stock average event size (AES) computed for each limit on a daily basis. In practice, the spread cannot be equal to one tick all the time. This is why we exclude from our analysis all the events that occur when the spread is higher than one tick.

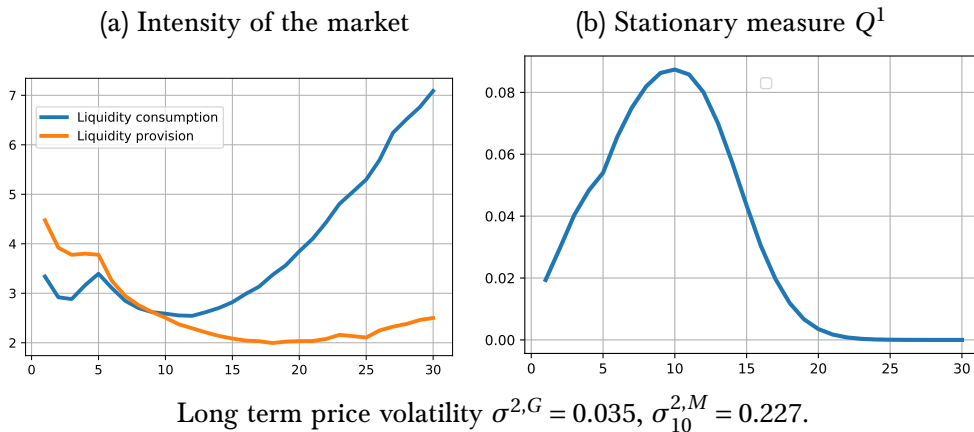


Figure IV.2 - (a) Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and (b) the corresponding stationary distribution of  $Q^1$  with respect to the queue size (in AES), proper to Air Liquide.

We can see that for all these assets, the liquidity provision intensity is approximately a decreasing function of the queue size. This result reveals a quite common strategy used in practice: posting orders when the queue is small to seize priority (for further details about the priority value, see Chapter III). For all assets, the consumption intensity is an increasing function when the queue size is large. For small queue sizes, we notice a slight decrease of this intensity, see Figure IV.2. Indeed, the increasing aspect corresponding to large queue sizes is explained by market participants waiting for better price when liquidity is abundant. The decreasing aspect associated to small queue sizes is due to aggressive orders sent by agents to get the last remaining quantities available at the first limits: market participants rushing for liquidity when it is rare. The lower the ratio of cancellation orders number over aggressive orders number is, the clearer the decreasing shape for small queue sizes stands out, see Table IV.1 and Figures IV.2, IV.4, IV.5 and IV.6.

### 6.3 Ranking of the market makers

For each of the assets and for each one of the market makers, we compute the liquidity consumption and provision intensities, and the corresponding price volatility  $\sigma_{10}^{2,M}$  that we would obtain in a situation where the studied market maker withdraws from the market. Since the estimators  $\sigma^{2,G}$  and  $\sigma_{10}^{2,M}$  give the same ranking, we choose to show the values for  $\sigma_{10}^{2,M}$  alone. We show next the results relative to Air Liquide; those of EssilorLuxottica, Michelin and Orange are relegated to Appendix IV.I.

IV. From asymptotic properties of general point processes to the ranking of financial agents

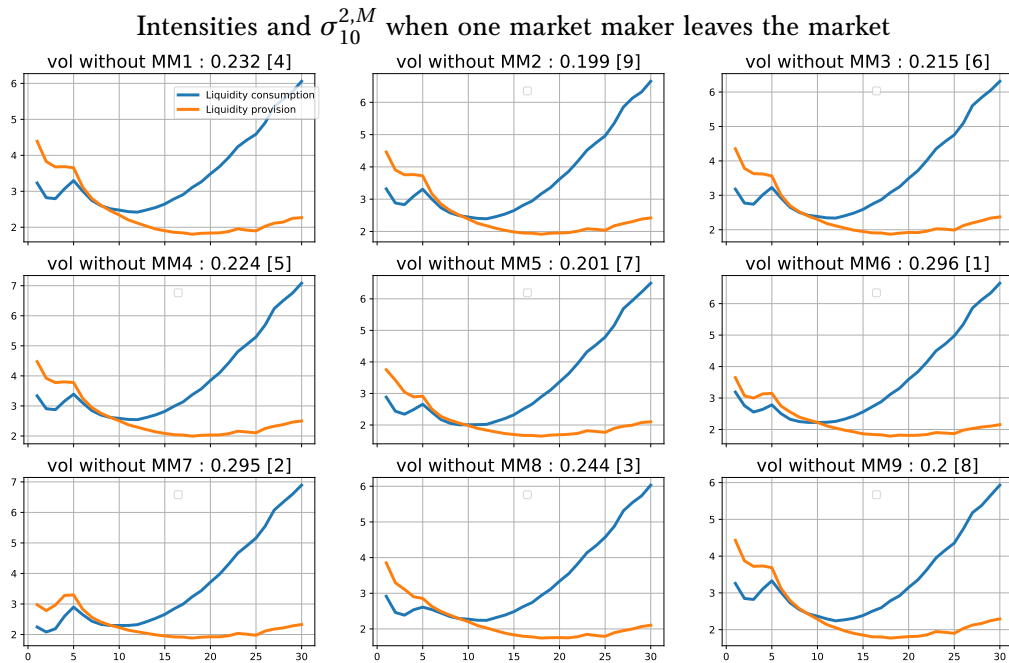


Figure IV.3 – Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and  $\sigma_{10}^{2,M}$  when one market maker is ejected from the market for the stock Air Liquide.

Based on the previous results, we carry out for each asset the ranking of the different market makers according to their contribution to volatility. To do so, we compare the expected volatility when removing each market maker from the market to the actual one when all the market makers in the market: if the expected volatility is higher (resp. lower) than the actual one, this means that the market maker into question decreases (resp. increases) market volatility. The market maker who decreases<sup>9</sup> (resp. increases<sup>10</sup>) volatility the most is ranked first (resp. last). In the following table, we add a star next to market makers decreasing volatility: a zero star (resp. a four stars) means that the market maker increases (resp. decreases) the market volatility of the 4 studied assets.

<sup>9</sup>The expected volatility of the new market without this market maker is the highest.

<sup>10</sup>The expected volatility of the new market without this market maker is the lowest.

Market maker	Ranking Air Liquide	Market share Air Liquide	Ranking Exilor-Luxottica	Market share Exilor-Luxottica	Ranking Michelin	Market share Michelin	Ranking Orange	Market share Orange
MM1***	4	4%	3	3%	3	4%	3	3%
MM2	9	1%	9	1%	9	1%	7	1%
MM3	6	5%	6	5%	7	4%	5	4%
MM4	5	1%	4	1%	4	0%	4	1%
MM5	7	5%	8	5%	8	5%	9	5%
MM6****	1	3%	2	3%	1	3%	1	4%
MM7****	2	7%	1	12%	2	9%	2	7%
MM8*	3	9%	5	5%	5	5%	6	4%
MM9	8	2%	7	2%	6	2%	8	2%

Table IV.2 – Market share and ranking of markets makers

## IV.A Market reconstitution

*Proof of Proposition 2.1.* Let  $t \geq 0$  be the current time. For any  $B \in \mathcal{E}$ , we denote by  $T^{t,e}$  the first time greater than  $t$  when an event  $e \in B$  happens given  $\mathcal{F}_t$  and  $T^{t,B} = \min_{e \in B} T^{t,e}$  the time of the next market event. Thus, we have

$$\begin{aligned} \lambda_t(B) &= \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\#\{T_n \in (t, t + \delta t], X_n \in B\} \geq 1 | \mathcal{F}_t]}{\delta t} \\ &= \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\{T^{t,B} \in (t, t + \delta t)\} | \mathcal{F}_t]}{\delta t}. \end{aligned}$$

We write  $f^{t,e}$  for the density function of  $T^{t,e}$  and  $F_B^{t,e}(s) = \mathbb{P}[(\min_{\bar{e} \in B \setminus \{e\}} T^{t,\bar{e}}) \geq s | T^{t,e} \leq s]$  for any  $s \geq 0$ . Using the monotone convergence theorem, we have

$$\begin{aligned} \lim_{\delta t \rightarrow 0} \frac{\mathbb{P}[\{T^{t,B} \in (t, t + \delta t)\} | \mathcal{F}_t]}{\delta t} &= \lim_{\delta t \rightarrow 0} \frac{\sum_{e \in B} \int_t^{t+\delta t} f^{t,e}(s) F_B^{t,e}(s) ds}{\delta t} \\ &= \sum_{e \in B} \lim_{\delta t \rightarrow 0} \frac{\int_t^{t+\delta t} f^{t,e}(s) F_B^{t,e}(s) ds}{\delta t} \\ &= \sum_{e \in B} f^{t,e}(t) F_B^{t,e}(t) = \sum_{e \in B} \lambda_t(e), \end{aligned}$$

since  $f_a^{t,e'}(t) = \lambda_t((e', a))$  using Equation (2) and  $F_a^{t,e'}(t) = 1$  by definition. This completes the proof.  $\square$

### IV.B Proof of Remark 3.4

*Proof of Remark 3.4.* Let  $N = (T_n, X_n)$  be the point process defined in Section 2 and  $i \in \mathbb{B} = \{1, 2\}$ . We define  $\phi_Q^{i,\pm,n}$  in the following way:

$$\begin{aligned}\phi_Q^{i,+n} &= \sup_{e \in E_Q^{i,+}(u,n)} \sum_{T_i < t} \phi(e, U_{t^-}, n, t - T_i, X_i), \\ \phi_Q^{i,-n} &= \inf_{e \in E_Q^{i,-}(u,n)} \sum_{T_i < t} \phi(e, U_{t^-}, n, t - T_i, X_i),\end{aligned}$$

with  $U_t = (Q_t^1, Q_t^2, S_t)$ . When  $Q_{t^-}^i \geq C_{bound}$ , using that  $\psi$  is non-decreasing in  $z$ , we have

$$\begin{aligned}\sum_{n \geq 0} (z_0^n - 1) (\lambda_Q^{i,+}(U_{t^-}, n) - \lambda_Q^{i,-}(U_{t^-}, n) \frac{1}{z_0^n}) &\leq \sum_{n \geq 0} (z_0^n - 1) (\psi_Q^{i,+}(U_{t^-}, n, t, \phi_Q^{i,+n}) - \lambda_Q^{i,-}(U_{t^-}, n) \frac{1}{z_0^n}) \\ &= \sum_{n \geq 0} (z_0^n - 1) (\psi_Q^{i,+}(U_{t^-}, n, t, \phi_Q^{i,+n}) - \psi_Q^{i,-}(U_{t^-}, n, t, \phi_Q^{i,+n}) \frac{1}{z_0^n}) \\ &\quad + \sum_{n \geq 0} (1 - \frac{1}{z_0^n}) (\psi_Q^{i,-}(U_{t^-}, n, t, \phi_Q^{i,+n}) - \lambda_Q^{i,-}(U_{t^-}, n)) = (i) + (ii).\end{aligned}$$

Using Equation (8), we have

$$(i) = \sum_{n \geq 0} (z_0^n - 1) (\psi_Q^{i,+}(U_{t^-}, n, t, \phi_Q^{i,+n}) - \psi_Q^{i,-}(U_{t^-}, n, t, \phi_Q^{i,+n}) \frac{1}{z_0^n}) \leq -\delta, \quad a.s., \quad (21)$$

when  $Q_{t^-}^i \geq C_{bound}$ . Moreover, using that  $\psi$  is non-decreasing in  $z$ , we have

$$\begin{aligned}(ii) &= \sum_{n \geq 0} (1 - \frac{1}{z_0^n}) \sum_{e \in E_Q^{i,-}(u,n)} (\psi(e, U_{t^-}, n, t, \phi^{i,+n}) - \psi(e, U_{t^-}, n, t, \sum_{T_i < t} \phi(e, U_{t^-}, n, t - T_i, X_i))) \\ &\leq \sum_{n \geq 0} (1 - \frac{1}{z_0^n}) \sum_{e \in E_Q^{i,-}(u,n)} (\psi(e, U_{t^-}, n, t, \phi^{i,+n}) - \psi(e, U_{t^-}, n, t, \phi^{i,-n})), \quad a.s.,\end{aligned}$$

when  $Q_{t^-}^i \geq C_{bound}$ . Since Equation (8) ensures that  $\phi^{i,+n} \leq \phi^{i,-n}$ , *a.s.* and  $\psi$  is non-decreasing in  $z$ , we deduce that

$$(ii) = \sum_{n \geq 0} (1 - \frac{1}{z_0^n}) (\psi_Q^{i,-}(U_{t^-}, n, t, \phi_Q^{i,+n}) - \lambda_Q^{i,-}(U_{t^-}, n)) \leq 0, \quad a.s., \quad (22)$$

when  $Q_{t^-}^i \geq C_{bound}$ . Using Equations (21) and (22), we get

$$\sum_{n \geq 0} (z_0^n - 1) (\lambda_Q^{i,+}(U_{t^-}, n) - \lambda_Q^{i,-}(U_{t^-}, n) \frac{1}{z_0^n}) \leq -\delta, \quad a.s.,$$

when  $Q_{t^-}^i \geq C_{bound}$ . By following the same methodology, we also get

$$\sum_{n \geq 0} (z_0^n - 1) (\lambda_S^{i,+}(U_{t^-}, n) - \lambda_S^{i,-}(U_{t^-}, n) \frac{1}{z_0^n}) \leq -\delta, \quad a.s.,$$

when  $S_{t^-} \geq C_{bound}$ . This completes the proof.  $\square$

## IV.C Proof of Theorem 3.1

### IV.C.1 Preliminary results

For any  $k \geq 1$ , we denote by  $T_{n+1}(e)$ ,  $T_{Q_{n+1}}^{i\pm}(k)$  and  $T_{S_{n+1}}^{i\pm}(k)$  respectively the arrival time of the first event  $e$ ,  $e_Q^{i\pm}(k) \in E_Q^{i\pm}$  and  $e_S^\pm(k) \in E_S^\pm$  greater than  $T_n$ . The sets  $E_Q^{i,\pm}$  and  $E_S^\pm$  are defined in Equation (6). They contain the events that increase or decrease the best bid, best ask and spread by  $k$ .

**Lemma IV.C.1.** *Let  $n \geq 0$  and  $i \in \mathbb{B}$ . The order book increments satisfy the following formulas:*

$$\begin{aligned}\mathbb{P}[\Delta Q_{n+1}^i = \pm k] &= \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{Q_n}^{i,\pm}(t, k) Z_n(t) dt\right], \\ \mathbb{P}[\Delta S_{n+1} = \pm k] &= \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{S_n}^\pm(t, k) Z_n(t) dt\right],\end{aligned}$$

with  $\Delta Q_{n+1}^i = Q_{n+1}^i - Q_n^i$ ,  $\Delta S_{n+1} = S_{n+1} - S_n$  and

$$\begin{aligned}Z_n(t) &= e^{-[\sum_e \int_0^t \lambda_n(e, s+T_n) ds]}, & \lambda_{Q_n}^{i,\pm}(t, k) &= \sum_{e \in E_Q^{i,\pm}(k)} \lambda_n(e, t+T_n), \\ \lambda_{S_n}^\pm(t, k) &= \sum_{e \in E_S^{i,\pm}(k)} \lambda_n(e, t+T_n), & \lambda_n(e, t) &= \psi(e, U_{T_n}, t, \sum_{T_i \leq T_n} \phi(e, U_{T_n}, t - T_i, X_i)), \quad \forall t \geq 0.\end{aligned}$$

*Proof of Lemma IV.C.1.* We write  $\Delta T_{n+1}(e) = T_{n+1}(e) - T_n$  for any event  $e \in E$  and  $\Delta T_{Q_{n+1}}^{i\pm}(k) = T_{Q_{n+1}}^{i\pm}(k) - T_n$ . Using Remark 3.7, the increments  $(\Delta T_{n+1})_{n \geq 0}$  are independent given  $\mathcal{F}_n$  and  $\Delta T_{n+1}(e) | \mathcal{F}_n$  follows a non homogeneous exponential distribution with an intensity  $\lambda_n(e, \cdot)$ . Thus, we have

$$\begin{aligned}\mathbb{P}[\Delta Q_{n+1}^i = \pm k] &= \mathbb{E}[\mathbb{P}[\Delta T_{Q_{n+1}}^{i\pm}(k) < \Delta T_{n+1}(e), \quad \forall e \notin E_Q^{i\pm}(k) | \mathcal{F}_n]] \\ &= \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{Q_n}^{i,\pm}(t, k) e^{-\int_0^t \lambda_{Q_n}^{i,\pm}(s, k) ds} dt \prod_{e \notin E_Q^{i\pm}(k)} \left(\int_{\mathbb{R}_+} \mathbf{1}_{t < t_e} \lambda_n(e, t_e) e^{-\int_0^{t_e} \lambda_n(e, s) ds} dt_e\right)\right] \\ &= \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{Q_n}^{i,\pm}(t, k) e^{-[\sum_e \int_0^t \lambda_n(e, s) ds]} dt\right] = \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{Q_n}^{i,\pm}(t, k) Z_n(t) dt\right].\end{aligned}\quad (23)$$

By following the same methodology used in Equation (23), we get

$$\mathbb{P}[\Delta S_{n+1} = \pm k] = \mathbb{E}\left[\int_{\mathbf{R}_+} \lambda_{S_n}^\pm(t, k) Z_n(t) dt\right],$$

which completes the proof.  $\square$

Let  $\tau_{\mathcal{O}}$  be the first entrance period of  $N^i = (T_{i+j}, X_{i+j})_{j \leq 0}$  to the set  $\mathcal{O} \in \mathcal{W}_0$ ,  $C_{bound}$  defined in Assumption 3.2 and  $1 < z \leq \min(z_0, z_1)$  with  $z_0$  and  $z_1$  are respectively defined in Assumptions 3.2 and 3.3.

**Lemma IV.C.2** (Drift condition). *Under Assumptions 3.2 and 3.3, the process  $U_n = (Q_n^1, Q_n^2, S_n)$  satisfies the following drift condition:*

$$\begin{aligned} \mathbb{E}[z^{Q_{n+1}^i - C_{\text{bound}}} \mathbf{1}_{\tau_\theta \geq n+1}] &\leq \lambda \mathbb{E}[z^{Q_n^i - C_{\text{bound}}} \mathbf{1}_{\tau_\theta \geq n+1}] + B \mathbb{E}[\mathbf{1}_{\tau_\theta \geq n+1}], \\ \mathbb{E}[z^{S_{n+1} - C_{\text{bound}}} \mathbf{1}_{\tau_\theta \geq n+1}] &\leq \lambda \mathbb{E}[z^{S_n - C_{\text{bound}}} \mathbf{1}_{\tau_\theta \geq n+1}] + B \mathbb{E}[\mathbf{1}_{\tau_\theta \geq n+1}], \end{aligned} \quad \forall n \in \mathbb{N}, \forall i \in \mathbb{B},$$

with  $\lambda < 1$  and  $B$  two constants.

**Remark IV.C.1.** We define

$$V_{C_{\text{bound}}}(u) = \sum_{i \in \{1,2\}} z^{q^i - C_{\text{bound}}} + z^{s - C_{\text{bound}}}, \quad \forall u \in \mathbb{U}. \quad (24)$$

Using Lemma IV.C.2, we deduce that

$$\mathbb{E}[V_{C_{\text{bound}}}(U_{n+1}) \mathbf{1}_{\tau_\theta \geq n+1}] \leq \lambda \mathbb{E}[V_{C_{\text{bound}}}(U_n) \mathbf{1}_{\tau_\theta \geq n+1}] + 3B \mathbb{E}[\mathbf{1}_{\tau_\theta \geq n+1}], \quad \forall n \in \mathbb{N},$$

*Proof of Lemma IV.C.2.* We write  $\tilde{\mathbb{E}}[X] = \mathbb{E}[X \mathbf{1}_{\tau_\theta \geq n+1}]$  for any random variable  $X$  to simplify the notations and  $V$  instead of  $V_{C_{\text{bound}}}$  since there is no possible confusion. We have

$$\tilde{\mathbb{E}}[z^{Q_{n+1}^i} | \mathcal{F}_n] = \tilde{\mathbb{E}}[z^{Q_n^i} | \mathcal{F}_n] + \sum_{u' \neq U_n} \tilde{\mathbb{P}}[Q_{n+1}^i = q' | \mathcal{F}_n] [z^{q'} - z^{Q_n^i}].$$

Using Lemma IV.C.1, we get

$$\mathbb{P}[\Delta Q_{n+1}^i = \pm k] = \mathbb{E}\left[\int_{\mathbb{R}_+} \lambda_{Q_n^i}^{i,\pm}(t, k) Z_n(t) dt\right],$$

which leads to

$$\begin{aligned} \tilde{\mathbb{E}}[z^{Q_{n+1}^i}] &= \tilde{\mathbb{E}}[z^{Q_n^i}] + \tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \left\{ \sum_{k \geq 1} \lambda_{Q_n^i}^{i,+}(t, k) [z^{Q_n^i+k} - z^{Q_n^i}] + \sum_{k \geq 1} \lambda_{Q_n^i}^{i,-}(t, k) [z^{Q_n^i-k} - z^{Q_n^i}] \right\} dt\right], \\ &= \tilde{\mathbb{E}}[z^{Q_n^i}] + \tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \{\mathcal{Q}_u(t, U_n)\} dt\right], \end{aligned} \quad (25)$$

with  $\mathcal{Q}_u(t, U_n) = \sum_{k \geq 1} \lambda_{Q_n^i}^{i,+}(t, k) [z^{Q_n^i+k} - z^{Q_n^i}] + \sum_{k \geq 1} \lambda_{Q_n^i}^{i,-}(t, k) [z^{Q_n^i-k} - z^{Q_n^i}]$ . By rearranging the above terms, we get

$$\mathcal{Q}_u(t, U_n) = z^{Q_n^i - C_{\text{bound}}} \sum_{1 \leq k} (z^k - 1) \left[ \lambda_{Q_n^i}^{i,+}(t, k) - \lambda_{Q_n^i}^{i,-}(t, k) \frac{1}{z^k} \right].$$

We write  $\tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \{\mathcal{Q}_u(t, U_n)\} dt\right] = T_1 + T_2$  with

$$\begin{aligned} T_1 &= \tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \mathbf{1}_{Q_n^i \leq C_{\text{bound}}} \{\mathcal{Q}_u(t, U_n)\} dt\right], \\ T_2 &= \tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \mathbf{1}_{Q_n^i > C_{\text{bound}}} \{\mathcal{Q}_u(t, U_n)\} dt\right]. \end{aligned}$$

We first handle the term  $T_1$ . When  $Q_n^i \leq C^{bound}$ , the quantity  $z^{Q_n^i - C^{bound}} < 1$  is bounded. Additionally, we have  $\sum_{e \in E} \lambda_n(e, s + T_n) \geq \underline{\psi} > 0$  under Assumption 3.3. This ensures that  $Z_n(t) \leq e^{-\psi t}$ , *a.s.* Thus, there exist  $c^1 > 0$  and  $d^1 > 0$  such that

$$T_1 \leq \int_{\mathbb{R}_+} e^{-\psi t} \tilde{\mathbb{E}}[\{\mathcal{Q}_u(t, U_n)\} dt] \leq -c^1 \tilde{\mathbb{E}}[z^{Q_n^i - C^{bound}} \mathbf{1}_{Q_n^i \leq C^{bound}}] + d^1. \quad (26)$$

In the last inequality we used Assumption 3.3 again. For the term  $T_2$ , we use Assumption 3.2 and  $Z_n(t) \leq e^{-\psi t}$ , *a.s.*, to deduce that

$$T_2 \leq -\frac{\delta}{\psi} \tilde{\mathbb{E}}[z^{Q_n^i - C^{bound}} \mathbf{1}_{Q_n^i > C^{bound}}]. \quad (27)$$

By combining Inequalities (26) and (27), we have

$$\tilde{\mathbb{E}}\left[\int_{\mathbb{R}_+} Z_n(t) \{\mathcal{Q}_u(t, U_n)\} dt\right] \leq -c \tilde{\mathbb{E}}[z^{Q_n^i - C^{bound}}] + \tilde{\mathbb{E}}[d],$$

with  $c = \min(c^1, \frac{\delta}{\psi})$  and  $d = d^1$  which proves the first inequality of Lemma IV.C.2. By following the same steps, we also prove the second inequality. This completes the proof.  $\square$

### IV.C.2 Outline of the proof

To prove the existence of an invariant distribution, we first construct  $N$  as a limiting process of the sequence  $N^m$  defined in Remark 3.7. This construction is based on the thinning algorithm. After that, we show, in Steps (ii) and (iii), that  $N$  is well defined. Then, we introduce the process  $\bar{U}^\infty = \text{ess sup}_{t \geq 0} \bar{U}_t$  which dominates  $\bar{U}_t$  and prove that it does not explode in Step (iv). This ensures the tightness of the family  $\cup_{t \geq 0} \bar{U}_t$ . Additionally, the process  $\bar{U}$  satisfies the Feller property since  $E$  is a countable space and  $\mathbb{E}[\|\bar{U}_t\|]$  is uniformly bounded. Thus, we deduce that  $\bar{U}$  admits an invariant distribution and complete the proof.

### IV.C.3 Proof

*Proof of Theorem 3.1.* Let us take  $N^*$  and  $U^*$  the processes described in Remark 3.7 with  $\nu = \sum_{e \in E} \delta_e$ . For clarity, we forget the dependence of  $\mathbb{E}_x[\cdot]$  on the initial condition  $\mathbf{x} \in W_0$ .

**Step (i):** In this step, we prove that the process  $N$ , defined by Equation (3), exists as a limiting process of the sequence  $N^m$ . To do so, we first introduce some notations. We define recursively the processes  $\lambda^m$  and  $N^m$  as in Remark 3.7. Note that  $U^m = (Q^{m1}, Q^{m2}, S^m)$  can be decomposed in the following way:

$$Q_t^{mi} = Q_t^{mi,+} - Q_t^{mi,-}, \quad S_t^m = S_t^{m+} - S_t^{m-}, \quad (28)$$

with

$$\begin{aligned} Q_t^{mi,+} &= \sum_{T^m < t} \Delta Q_t^{mi} \mathbf{1}_{\Delta Q_t^{mi} > 0}, & Q_t^{mi,-} &= \sum_{T^m < t} \Delta Q_t^{mi} \mathbf{1}_{\Delta Q_t^{mi} < 0}, \\ S_t^{m+} &= \sum_{T^m < t} \Delta S_t^m \mathbf{1}_{\Delta S_t^m > 0}, & S_t^{m-} &= \sum_{T^m < t} \Delta S_t^m \mathbf{1}_{\Delta S_t^m < 0}, \end{aligned}$$



IV. From asymptotic properties of general point processes to the ranking of financial agents

with  $i \in \mathbb{B}$  and  $\Delta Z_t = Z_t - Z_{t-}$  for any process  $Z$ . For all  $\omega \in \Omega$ , each one of the processes  $N^m$ ,  $\lambda^m$ ,  $Q^{m,i,\pm}$  and  $S^{m,\pm}$  is non decreasing with  $m$  by induction. Hence, they admit limiting processes  $N$ ,  $\lambda$ ,  $Q^{1(2),\pm}$  and  $S^\pm$ . This implies that  $U^m$  converges towards  $U$ . To ensure that  $N$  admits  $\lambda$  as an intensity, we need to prove that  $\sum_{e \in E} \lambda_t(e)$  and  $U$  are both finite *a.s.*, see Steps (ii)-(iii).

**Step (ii):** In this step, we prove by induction on  $m$  that  $\sup_t \mathbb{E}[\sum_{e \in E} \lambda_t^m(e)]$  is uniformly bounded which ensures that  $\sup_t \mathbb{E}[\sum_{e \in E} \lambda_t(e)]$  is finite and that  $\sum_{e \in E} \lambda_t(e)$  does not explode. We write  $\lambda_n^m(e, t) = \lambda_t^m(e) \mathbf{1}_{T_n^m < t \leq T_{n+1}^m}$ . For  $m = 0$ , we have  $\mathbb{E}[\lambda_n^m(t, e)] = 0$  since  $\lambda_t^m(e) = 0$  for any  $t \geq 0$ . We have by construction

$$\begin{aligned} \mathbb{E}[\lambda_n^{m+1}(e, t)] &= \mathbb{E}[\lambda_n^m(e, t)], & \text{when } n \leq m, \\ \mathbb{E}[\lambda_n^{m+1}(e, t)] &= 0, & \text{when } n > m + 1. \end{aligned}$$

for any  $t \geq 0$ . Thus, we only need to study the case  $n = m + 1$ . Using Remark 3.7 and Assumption 3.1, we have

$$\begin{aligned} \sup_t \mathbb{E}[\lambda_{m+1}^{m+1}(e, t)] &\leq c(e) + d(e) \sup_t \mathbb{E}[\left( \sum_{T_i^m < t} \bar{\phi}(e, t - T_i^m, X_i^m) \right)^{n_\psi}] \\ &= c(e) + d(e) \sup_t \sum_{\{\mathbf{k}_m\} \in \mathcal{D}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_k} \sum_{\mathbf{x} \in E^k} \int_{(-\infty, t)^k} \prod_{i=1}^k \bar{\phi}^{k_i}(e, t - s_i, x_i) \mathbb{E}[dN_{s_1}^m \dots dN_{s_k}^m], \end{aligned}$$

with  $\bar{\phi}(e, t, x) = \sup_{u \in \mathbb{U}} \phi(e, u, t, x)$  and  $\binom{n_\psi}{\mathbf{k}_k} = \frac{n_\psi!}{k_1! \dots k_k!}$ . Using the above equation and the Brascamp-Lieb inequality, we have

$$\begin{aligned} \sup_t \mathbb{E}[\lambda_{m+1}^{m+1}(e, t)] &\leq c(e) + d(e) \sup_t \sum_{\mathbf{k}_k \in \mathcal{D}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_k} \sum_{\mathbf{x} \in E^k} \int_{(-\infty, t)^k} \prod_{i=1}^k \bar{\phi}^{k_i}(e, t - s_i, x_i) (\sup_{t,n} \mathbb{E}[\lambda_n^m(x_i, t)])^{1/k} ds_i \\ &= c(e) + \bar{\lambda}^m d(e) \sum_{\mathbf{k}_k \in \mathcal{D}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_k} \sum_{\mathbf{x} \in E^k} \int_{\mathbb{R}_+^k} \prod_{i=1}^k \bar{\phi}^{k_i}(e, s_i, x_i) ds_i \\ &\leq c(e) + q \bar{\lambda}^m, \end{aligned} \tag{29}$$

with  $\bar{\lambda}^m = \sup_{t,e,n} \mathbb{E}[\lambda_n^m(e, t)]$  and  $q = \sup_e \{d(e) \sum_{\mathbf{k}_k \in \mathcal{D}(n_\psi)} 1/m! \binom{n_\psi}{\mathbf{k}_k} \int_{\mathbb{R}_+^k} \prod_{i=1}^k \phi^{*k_i}(e, s_i) ds_i\}$  where  $\phi^*$  is defined in Assumption 3.1. Using (29), we deduce that

$$\bar{\lambda}^{m+1} \leq \frac{c}{1-q} + q^{m+1} \bar{\lambda}^0 = \bar{x}.$$

Since  $q < 1$  under Assumption 3.1, it ensures that  $\bar{\lambda} = \sup_m \bar{\lambda}^m$  is finite. To complete the proof, we use (29) and Assumption 3.3, to get the following inequality:

$$\sup_t \mathbb{E}[\sum_{e \in E} \lambda_t^m(e)] \leq c^* + \bar{\lambda} \sum_{e \in E, \mathbf{k}_k \in \mathcal{D}(n_\psi)} d(e)/m! \binom{n_\psi}{\mathbf{k}_k} \sum_{\mathbf{x} \in E^k} \int_{\mathbb{R}_+^k} \prod_{i=1}^k \bar{\phi}^{k_i}(e, t - s_i, x_i) ds_i < \infty.$$

**Step (iii):** We write  $U_n^m = (Q_n^{m1}, Q_n^{m2}, S_n^m) = U_{T_n}^m$ . We prove here that  $\mathbb{E}[Q_n^{mi}]$  and  $\mathbb{E}[S_n^m]$  are uniformly bounded for all  $m \geq 0$  and  $n \geq 0$  to ensure that  $S$  and  $Q^i$  do not explode. Let us prove that

$$\mathbb{E}[z^{Q_{n+1}^{mi}}] \leq \lambda \mathbb{E}[z^{Q_n^{mi}}] + B, \quad \forall n \leq m-1, m \geq 1. \quad (30)$$

with  $z \leq \min(z_0, z_1)$  and  $z_0$  and  $z_1$  are respectively defined in Assumption 3.2 and 3.3,  $\lambda < 1$  and  $B \geq 0$ . Let  $m \geq 1$ , we have by construction

$$\mathbb{E}[z^{Q_{n+1}^{mi}}] = \mathbb{E}[z^{Q_n^{mi}}], \quad \text{when } n \leq m-1.$$

Thus, we only need to investigate the case  $n = m$ . This is proved in Lemma IV.C.2. Using Inequality (30), we get

$$\mathbb{E}[z^{Q_n^{mi}}] \leq \frac{B}{1-\lambda} + \lambda^n z^{Q_0^{mi}}, \quad \forall n \leq m, \quad (31)$$

with  $z^{Q_0^{mi}}$  fixed. Thus,  $\mathbb{E}[Q_n^{mi}]$  is uniformly bounded. Using similar lines of argument, we also have  $\mathbb{E}[S_n^m]$  uniformly bounded. Hence, the limiting processes  $U$  does not explode.

**Step (iv):** First, note that the process  $N$  is well defined since  $\lambda_t$  is locally integrable, see Step (ii)-(iii) and [93]. Additionally, we can construct it pathwise using the thinning algorithm, see Remark 3.7.

Let  $\bar{U}_s$  be the process described in Theorem 3.1 and for which we just proved the existence. This process is dominated by the process  $\bar{U}^\infty = (U^\infty, \lambda^\infty) = \text{esssup}_{s \geq 0} \bar{U}_s$ . In this part, we prove that both  $\mathbb{E}[U^\infty]$  and  $\mathbb{E}[\lambda^\infty]$  are finite.

First, we prove that  $\mathbb{E}[U^\infty] < \infty$ . Let  $\lambda < \rho < \rho^1 < 1$ ,  $C > 0$ ,  $\mathcal{S}$  the set  $\mathcal{S} = \{u \in \mathbb{U}; u > C, c.w.\}$  where *c.w.* means component-wise and  $\mathcal{S}$  a set  $S \in \mathcal{U} \subset \mathcal{S}$ . Since  $U_n \mathbf{1}_{U_n \in \mathcal{S}^c}$  is bounded *a.s.*, we only need to show  $\mathbb{E}[U^{\infty, \mathcal{S}}]$  is finite with  $U^{\infty, \mathcal{S}} = \text{esssup}_{n \in \mathbb{N}} U_n^{\mathcal{S}}$  and  $U_n^{\mathcal{S}} = U_n \mathbf{1}_{U_n \in \mathcal{S}}$ . Using the Doob's decomposition, we have  $U_n^{\mathcal{S}} = M_n^{\mathcal{S}} + A_n^{\mathcal{S}}$  with  $M_n^{\mathcal{S}}$  a martingale and  $A_n^{\mathcal{S}} = \sum_{k=1}^n (\mathbb{E}[U_k^{\mathcal{S}} | \mathcal{F}_{k-1}] - U_{k-1}^{\mathcal{S}})$  a predictable process. Thus, we get

$$\mathbb{E}[U^{\infty, \mathcal{S}}] \leq \mathbb{E}[\text{esssup}_{n \geq 0} M_n^{\mathcal{S}}] + \mathbb{E}[\text{esssup}_{n \geq 0} A_n^{\mathcal{S}}], \quad c.w.$$

The Doob's inequality and Fatou's Lemma ensure that  $\mathbb{E}[\sup_{n \geq 0} M_n^{\mathcal{S}}] \leq 2 \lim_{n \rightarrow \infty} \mathbb{E}[M_n^{\mathcal{S}^2}]^{\frac{1}{2}}$ , *c.w.* Using the martingale property of  $M_n^{\mathcal{S}}$  and the Doob's decomposition of  $U_n^{\mathcal{S}}$ , we find

$$\mathbb{E}[(M_n^{\mathcal{S}})^2] - \mathbb{E}[(M_0^{\mathcal{S}})^2] = \sum_{k=1}^n \mathbb{E}[(M_k^{\mathcal{S}} - M_{k-1}^{\mathcal{S}})^2], \quad M_k^{\mathcal{S}} - M_{k-1}^{\mathcal{S}} = U_k^{\mathcal{S}} - \mathbb{E}[U_k^{\mathcal{S}} | \mathcal{F}_{k-1}], \quad c.w.$$

We have

$$\mathbb{E}[(M_k^{\mathcal{S}} - M_{k-1}^{\mathcal{S}})^2] = \mathbb{E}[(U_k^{\mathcal{S}} - \mathbb{E}[U_k^{\mathcal{S}} | \mathcal{F}_{k-1}])^2] \leq 2(\mathbb{E}[(U_k^{\mathcal{S}})^2] + \mathbb{E}[\mathbb{E}[(U_k^{\mathcal{S}})^2 | \mathcal{F}_{k-1}]]) \leq 4\mathbb{E}[(U_k^{\mathcal{S}})^2], \quad c.w.$$

#### IV. From asymptotic properties of general point processes to the ranking of financial agents

Let us prove that  $\sum_{k \geq 0} \mathbb{E}[(U_k^S)^2] < \infty$ . Using Lemma IV.C.2 and by taking  $\mathcal{O} = \{(T_j, X_j)_{j \leq 0} \in W_0; X_j = (n_j, t_j, b_j, \tilde{u}_j, u_j, a_j) \in E \text{ and } u_0 \geq C, c.w.\}$ , we have

$$\mathbb{E}[V_C(U_{n+1})\mathbf{1}_{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}}] \leq \mathbb{E}[V_C(U_{n+1})\mathbf{1}_{U_n \in \mathcal{S}}] \leq \lambda \mathbb{E}[V_C(U_n)\mathbf{1}_{U_n \in \mathcal{S}}] + \mathbb{E}[B\mathbf{1}_{U_n \in \mathcal{S}}]. \quad (32)$$

By following the same lines of arguments used to prove (25) in Lemma IV.C.2 and basic approximations, we have the following inequality:

$$\mathbb{E}[z^{Q_{n+1}^i - C}\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}] \leq \mathbb{E}[z^{Q_n^i - C}\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}] + \mathbb{E}\left[\int_{\mathbb{R}_+} Z_n(t)\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}\{\mathcal{Q}_u(t, U_n)\} dt\right].$$

In the set  $\{U_n \in \mathcal{S}^c\}$ , we have  $Q_n^i \leq C$  which implies  $z^{Q_n^i - C} < 1$ . Additionally, we have  $\sum_{e \in E} \lambda_n(e, s + T_n) \geq \underline{\psi} > 0$  under Assumption 3.3. This ensures that  $Z_n(t) \leq e^{-\psi t}$ , *a.s.* Thus, using Assumption 3.3, there exists  $B^1$  such that

$$\mathbb{E}[z^{Q_n^i - C}\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}] + \mathbb{E}\left[\int_{\mathbb{R}_+} Z_n(t)\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}\{\mathcal{Q}_u(t, U_n)\} dt\right] \leq \mathbb{E}[B^1\mathbf{1}_{\{U_{n+1} \in \mathcal{S}, U_n \in \mathcal{S}^c\}}]. \quad (33)$$

We take  $C \geq C^* = \max(\log(\frac{2B}{\rho - \lambda} + 1), \log(\frac{B^1}{1 - \rho^1}), C_{bound})$  to ensure that

$$\begin{cases} [B - (\rho - \lambda)V_C(U_n)]\mathbf{1}_{U_n \in \mathcal{S}} < 0, & a.s. \\ [B^1 - (1 - \rho^1)V_C(U_n + 1)]\mathbf{1}_{U_{n+1} \in \mathcal{S}} < 0, & a.s. \end{cases}$$

By combining Inequalities (32) and (33) and taking  $C \geq C^*$ , we deduce that

$$\begin{aligned} \rho^1 \mathbb{E}[V_C(U_{n+1})\mathbf{1}_{U_{n+1} \in \mathcal{S}}] &\leq \rho \mathbb{E}[V_C(U_n)\mathbf{1}_{U_n \in \mathcal{S}}] + \mathbb{E}[(B - (\rho - \lambda)V_C(U_n))\mathbf{1}_{U_n \in \mathcal{S}}] \\ &\quad + \mathbb{E}[(B^1 - (1 - \rho^1)V_C(U_{n+1}))\mathbf{1}_{U_{n+1} \in \mathcal{S}}], \\ &\leq \rho \mathbb{E}[V_C(U_n)\mathbf{1}_{U_n \in \mathcal{S}}], \end{aligned}$$

which ensures that  $\mathbb{E}[V_C(U_{n+1})\mathbf{1}_{U_{n+1} \in \mathcal{S}}] \leq r \mathbb{E}[V_C(U_n)\mathbf{1}_{U_n \in \mathcal{S}}]$  with  $r = \frac{\rho}{\rho^1} < 1$ . Since  $(U_k^S)^2 \leq c_1 V_C(U_k^S)$ , this proves that  $\sum_{k \geq 0} \mathbb{E}[(U_k^S)^2] < c_1 \sum_{k \geq 0} \mathbb{E}[V_C(U_k^S)] \leq \frac{c_1}{1 - \rho} < \infty$ . Hence, we get  $\mathbb{E}[\text{esssup}_{n \geq 0} M_n^S] \leq (\frac{c_1}{1 - \rho})^{\frac{1}{2}}$ , *c.w.* We also have

$$A_n^S \leq \tilde{A}_n^S = \sum_{k=1}^n |\mathbb{E}[U_k^S | \mathcal{F}_{k-1}] - U_{k-1}^S| \leq 2 \sum_{k=1}^n \mathbb{E}[|U_k^S|], \text{ c.w.}$$

with  $\tilde{A}_n^S$  a component-wise non-decreasing process. Since  $\mathbb{E}[|U_k^S|] \leq (\mathbb{E}[(U_k^S)^2])^{\frac{1}{2}}$ , we get  $\mathbb{E}[\tilde{A}_n^S] \leq (\frac{c_1}{1 - \rho})^{\frac{1}{2}}$ . Hence, we deduce that  $\mathbb{E}[\text{esssup}_{n \geq 0} A_n^S] \leq (\frac{c_1}{1 - \rho})^{\frac{1}{2}}$ , *c.w.* which ensures that  $\mathbb{E}[U^{\infty, S}] < \infty$ .

Second, we prove that  $\mathbb{E}[\lambda^\infty]$  is finite. Let  $t \geq 0$  and  $\mathcal{T} = \{t_0 = 0 < t_1 < \dots < t_n = t\}$  be a partition of  $[0, t]$ . Using the monotone convergence theorem, we have

$$\mathbb{E}\left[\sum_{k=1}^n |\lambda_{t_k} - \lambda_{t_{k-1}}|\right] \leq \mathbb{E}\left[\frac{\sum_{k=1}^n (t_k - t_{k-1})}{t} |\tilde{\lambda}_{t_k} - \tilde{\lambda}_{t_{k-1}}|\right] = \mathbb{E}\left[\frac{\int_0^t f^{\mathcal{T}} ds}{t}\right] \leq \frac{\int_0^t \mathbb{E}[f^{\mathcal{T}}] ds}{t},$$

with  $f^{\mathcal{F}} = \sum_{k=1}^n |\lambda_{t_k} - \lambda_{t_{k-1}}| \mathbf{1}_{t_{k-1} \leq t < t_k}$ . Since  $\mathbb{E}[|\lambda_{t_k} - \lambda_{t_{k-1}}|] \leq 2 \sup_t \mathbb{E}[|\lambda_t|] \leq \frac{c}{1-q} < \infty$ , we get

$$\mathbb{E}\left[\sum_{k=1}^n |\lambda_{t_k} - \lambda_{t_{k-1}}|\right] \leq \frac{c}{1-q} < \infty.$$

We can then apply Bichteler-Dellacherie theorem to write  $\lambda_t = M_s + A_s$  with  $M_s$  a martingale and  $A_s$  a predictable process with almost surely finite variation over finite time intervals such that

$$\mathbb{E}[\mathbf{var}_t(\lambda)] = \mathbb{E}[\mathbf{var}_t(M)] + \mathbb{E}[\mathbf{var}_t(A)],$$

where  $\mathbf{var}_t(Z)$  is the variation of any process  $Z$  over the interval  $[0, t]$ . Since

$$\mathbb{E}[\lambda^\infty] \leq \mathbb{E}[\text{ess sup}_s M_s] + \mathbb{E}[\text{ess sup}_s A_s], \quad \text{ess sup}_{s \leq t} M_s \leq \mathbf{var}_t(M), \quad \text{ess sup}_{s \leq t} A_s \leq \mathbf{var}_t(A),$$

and  $\sup_t \mathbb{E}[\mathbf{var}_t(\lambda)] < \infty$ , we deduce that  $\mathbb{E}[\lambda^\infty] < \infty$ . Finally, we have  $\mathbb{E}[\|\tilde{U}_t\|] \leq \mathbb{E}[\|\tilde{U}^\infty\|] < \infty$ , for all  $t \geq 0$ . Thus, the family  $\cup_{t \geq 0} \tilde{U}_t$  is tight. Moreover, the process  $\tilde{U}_t$  satisfies the Feller property since  $\cup$  and  $E$  are countable states and  $\mathbb{E}[\|\tilde{U}_t\|]$  is uniformly bounded. Thus the process  $\tilde{U}$  admits an invariant distribution  $\mu$  which completes the proof.  $\square$

## IV.D Proof of Theorem 3.2

### IV.D.1 Preliminary result

**Lemma IV.D.1.** *Let  $(\mathcal{F}_n)_{n \geq 0}$  be a sequence of  $\sigma$ -algebras such that  $\mathcal{F}_n \xrightarrow{n \rightarrow \infty} \mathcal{F}_\infty$  with  $\mathcal{F}_\infty$  a  $\sigma$ -algebra and  $(X_n)_{n \geq 0}$  be a sequence of random variables valued in  $\mathbb{R}$  such that  $X_n \xrightarrow{n \rightarrow \infty} X$ , a.s.,  $X_n$  is  $\mathcal{F}_\infty$ -measurable,  $X$  is  $\mathcal{F}_\infty$ -measurable and  $\sup_n \mathbb{E}[X_n^2] < \infty$ . Then, we have*

$$\mathbb{E}[X_n | \mathcal{F}_n] \xrightarrow{n \rightarrow \infty} X, \quad a.s.$$

**Remark IV.D.1.** *In the above Lemma IV.D.1, we can replace the condition  $\sup_n \mathbb{E}[X_n^2] < \infty$  by the condition  $\mathbb{E}[\sup_n X_n] < \infty$  and recover the same result.*

*Proof of Lemma IV.D.1.* Let  $m$  and  $n$  be two positive integers. We write  $X_n^m = \mathbb{E}[X_m | \mathcal{F}_n]$ .

**Step (i):** Since  $\sup_n \mathbb{E}[X_n^2] < \infty$ , we can apply a conditional dominated convergence theorem to show that  $X_n^m \xrightarrow{m \rightarrow \infty} X_n = \mathbb{E}[X | \mathcal{F}_n]$ , a.s.

IV. From asymptotic properties of general point processes to the ranking of financial agents

**Step (ii):** Since  $\mathcal{F}_\infty = \lim_{n \rightarrow \infty} \mathcal{F}_n$ , there exists a sequence  $(A_n)_{n \geq 0}$  such that  $A_n \in \mathcal{F}_n$  and  $A_n \xrightarrow[n \rightarrow \infty]{} A$ . By definition, we have

$$\mathbb{E}[X_n \mathbf{1}_{A_n}] = \mathbb{E}[X \mathbf{1}_{A_n}].$$

Note that the family  $(X_n)_{n \geq 0}$  is tight. Indeed, using Doob's and Jensen's inequalities, we have

$$\mathbb{E}\left[\sup_{i \leq n} |X_i|\right] \leq \mathbb{E}\left[\left(\sup_{i \leq n} |X_i|\right)^2\right]^{\frac{1}{2}} = \mathbb{E}\left[\sup_{i \leq n} X_i^2\right]^{\frac{1}{2}} \leq 2\mathbb{E}[X_n^2]^{\frac{1}{2}}.$$

Then, using Fatou's Lemma, we get  $\mathbb{E}\left[\sup_{i \leq n} X_i\right] \leq 2(\sup_n \mathbb{E}[X_n^2])^{\frac{1}{2}} < \infty$  which ensures that  $(X_n)_{n \geq 0}$  is tight. Thus, we can extract a sub sequence  $(X_{n_k})_{k \geq 0}$  such that  $X_{n_k} \xrightarrow[k \rightarrow \infty]{} Z$  *a.s.*

Since  $\sup_n \mathbb{E}[X_n^2] < \infty$ , we can use the dominated convergence theorem to get

$$\mathbb{E}[Z \mathbf{1}_A] = \lim_{k \rightarrow \infty} \mathbb{E}[X_{n_k} \mathbf{1}_{A_{n_k}}] = \lim_{k \rightarrow \infty} \mathbb{E}[X \mathbf{1}_{A_{n_k}}] = \mathbb{E}[X \mathbf{1}_A].$$

Thus, we have  $Z = X$ ,  $\mathcal{F}_\infty$ -*a.s.* Since all the variables  $X_k$  are  $\mathcal{F}_\infty$ -measurable, the variable  $Z$  is also  $\mathcal{F}_\infty$ -measurable for any  $n \geq 0$ . Given that  $Z$  and  $X$  are both  $\mathcal{F}_\infty$ -measurable, we deduce that every accumulation point  $Z$  of  $(X_n)_{n \geq 0}$  satisfies  $Z = X$ , *a.s.* Finally, we get  $\lim_{n \rightarrow \infty} X_n^m = X$ , *a.s.* and we can use a composition argument, to deduce that  $\mathbb{E}[X_n | \mathcal{F}_n] \xrightarrow[n \rightarrow \infty]{} X$ , *a.s.*  $\square$

We borrow the following definition from [41].

**Definition IV.D.1 (Coupling).** *Two point processes  $N$  and  $N'$  couple if and only if*

$$\lim_{t \rightarrow \infty} \mathbb{P}[N_s = N'_s, \quad \forall s \in (t, \infty)] = 1.$$

**Lemma IV.D.2.** *Let  $N$  be a point process and  $\lambda$  its intensity. We have*

$$\mathbb{P}[N_s - N_t = 0, \quad \forall s \in (t, \infty) | \mathcal{F}_t] = \mathbb{E}[e^{-\int_t^\infty \lambda_u \mathbf{1}_{A_u} ds} | \mathcal{F}_t],$$

with  $A_u = \{N_u - N_t = 0\}$  for all  $u \geq t$ .

*Proof.* See Lemma 1 in [41].  $\square$

**Lemma IV.D.3.** *Two point processes  $N$  and  $N'$  which admit respectively  $\lambda$  and  $\lambda'$  as intensities couple if and only if*

$$\int_0^\infty \sup_{e \in E} \mathbb{E}[|\lambda_s(e) - \lambda'_s(e)|] ds < \infty.$$

*Proof.* Let  $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^{N'}$ . Using the canonical coupling, the point process  $|N - N'|$  admits  $|\lambda_t - \lambda'_t|$  as an  $\mathcal{F}_t$ -intensity. Using Lemma (IV.D.2) and Jensen's Inequality, we have

$$\mathbb{P}\left[\sup_e |N_s(e) - N'_s(e)| = 0, \quad \forall s \in (t, \infty)\right] \geq \mathbb{E}\left[e^{-\int_t^\infty \sup_e |\lambda_s(e) - \lambda'_s(e)| ds}\right] \geq e^{-\int_t^\infty \sup_e \mathbb{E}[|\lambda_s(e) - \lambda'_s(e)|] ds}.$$

Since  $\int_0^\infty \sup_e \mathbb{E}[|\lambda_s(e) - \lambda'_s(e)|] ds < \infty$ , we have  $\int_t^\infty \sup_e \mathbb{E}[|\lambda_s(e) - \lambda'_s(e)|] ds \xrightarrow[t \rightarrow \infty]{} 0$  which implies that

$$\mathbb{P}\left[\sup_e |N_s(e) - N'_s(e)| = 0, \quad \forall s \in (t, \infty)\right] \xrightarrow[t \rightarrow \infty]{} 1.$$

This completes the proof.  $\square$

## IV.D.2 Uniqueness

### IV.D.2.1 Outline of the proof

Let  $N^\infty = (T_i^\infty, X_i^\infty)$  be the stationary process constructed in Theorem 3.1 and  $N = (T_i, X_i)$  be a point process whose intensity satisfies (2). We write  $\lambda$  (resp.  $\lambda^\infty$ ) for the intensity of  $N$  (resp.  $N^\infty$ ). To prove the uniqueness of the invariant distribution, we only need to show that  $\int_0^\infty \sup_{e \in E} \mathbb{E}[|\lambda_s(e) - \lambda_s^\infty(e)|] ds < \infty$ , see Lemma IV.D.3. To do so, we first show that  $(U_n)_{n \geq 0}$  is  $f$ -geometrically ergodic, see Lemma IV.D.6. The proof of this result requires Lemmas IV.D.4 and IV.D.5. Using this result, we prove, in Lemma IV.D.7, that  $f(t) = \sup_e \mathbb{E}[|\lambda_t(e) - \lambda_t^\infty(e)|]$  satisfies the following inequality:

$$f(t) \leq u(t) + c_3 G \left( \int_0^t \bar{h}(t-s) f(s) ds \right),$$

with  $u(t) = c_2 \mathbb{E}[\|U_t - U_t^\infty\|] + c_1 \mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^{\frac{1}{\beta p}}$ ,  $G(t) = t^{\frac{1}{\beta}}$  and  $\bar{h}(t) = \sup_{e, u, x} \phi(e, u, t, x)$  with  $c_1, c_2, c_3, \beta > 1$  and  $p > 1$  positive constants. Then, we use Theorem 3 in [21] and the above inequality, to show that  $\int_{\mathbb{R}_+} f(t) dt < \infty$  which ensures the uniqueness.

### IV.D.2.2 Proof

Let  $\lambda < 1$  given by Lemma IV.C.2 and  $\lambda < \rho < 1$ . We denote by  $s = \{(T_j, X_j)_{j \leq 0} \in W_0; X_j = (n_j, t_j, b_j, \tilde{u}_j, u_j, a_j) \in E \text{ and } V(u_0) \leq \frac{2B}{\rho - \lambda} + 1\}$  and by  $\alpha$  a set  $\alpha \in \mathcal{W}_0 \subset s$ . We have the following lemma.

**Lemma IV.D.4.** *Under Assumptions 3.2 and 3.3, the function  $f = V + 1$  with  $V$  defined in Equation (24) and  $r > 1$  such that*

$$\sup_{x \in W} \mathbb{E}_x \left[ \sum_{n=1}^{\tau_\alpha} f(U_n) r^n \right] < \infty.$$

*Proof.* The proof is similar to Theorem 6.3 in [121]. □

Let  $\mathcal{F}_n$  and  $\mathcal{F}_{l \leq j \leq n}$  be respectively defined in the following way  $\mathcal{F}_n = \sigma(T_j \times X_j, \forall j \leq n)$ ,  $\mathcal{F}_{l \leq j \leq n} = \sigma(T_j \times X_j, \forall l \leq j \leq n)$ . We also write  $p_k^n$  as follows:

$$p_k^n(u) = |\mathbb{P}[U_n = u | \mathcal{F}_{k \leq j \leq n-1}] - \mathbb{P}[U_n = u | \mathcal{F}_{j \leq n-1}]|, \quad \forall n \in \mathbb{N}, \forall k \leq n-1, \forall u \in \mathbb{U}.$$

**Lemma IV.D.5.** *Under Assumptions 3.1, 3.3 and 3.4, we have*

$$p_k^n = \sup_{u \in \mathbb{U}} p_k^n(u) \xrightarrow{k \rightarrow \infty} 0, \quad a.s. \tag{34}$$

IV. From asymptotic properties of general point processes to the ranking of financial agents

*Proof.* Using Lemma IV.C.1, we have

$$\begin{aligned} p_k^n(u) &= |\mathbb{E}[\int_{\mathbb{R}_+} Z_n(t)\lambda_n(u, t) dt | \mathcal{F}_{k \leq j \leq n-1}] - \mathbb{E}[\int_{\mathbb{R}_+} Z_n(t)\lambda_n(u, t) dt | \mathcal{F}_{j \leq n-1}]| \\ &= |\mathbb{E}[\int_{\mathbb{R}_+} Z_n(t)\lambda_n(u, t) dt | \mathcal{F}_{k \leq j \leq n-1}] - \int_{\mathbb{R}_+} Z_n(t)\lambda_n(u, t) dt|, \end{aligned}$$

with  $\lambda_n(u, t) = \sum_{e \in E(U_{n-1}, u)} \lambda_n(e, t)$ ,  $\lambda_n(e, t) = \psi(e, U_{n-1}, t + T_{n-1}, r_n(t))$ ,  $r_n(t) = \sum_{j \leq n-1} \phi(e, U_{n-1}, t + T_{n-1} - T_j, X_j)$  and  $Z_n(t) = e^{-\int_0^t \sum_e \lambda_n(e, s) ds}$ .

Since  $p_k = \sup_{u \in \mathbb{U}} p_k^n(u)$ , we can construct a sequence  $(u_j)_{j \geq 0}$  such that  $p_k^n(u_j) \xrightarrow{j \rightarrow \infty} p_k^n$ , *a.s.*

We write  $u_j = (q_j^1, q_j^2, s_j)$ . Without loss of generality, we can consider that  $(q_j^1)_{j \geq 0}$  is monotonic by taking a sub-sequence of  $(q_j^1)_{j \geq 0}$ . Hence, there exists a limiting process  $q_\infty^1$  such that  $q_j^1 \xrightarrow{j \rightarrow \infty} q_\infty^1$ , *a.s.* By repeating this argument several times, we can always construct  $(u_j)$  such that

$$p_k^n(u_j) \xrightarrow{j \rightarrow \infty} p_k, \quad u_j \xrightarrow{j \rightarrow \infty} u_\infty, \quad a.s.$$

Let us prove that  $\lambda_n(u_j, t) \xrightarrow{j \rightarrow \infty} \lambda_n(u_\infty, t)$ , *a.s.* To do so, we distinguish two sets  $A_1 = \{w \in \Omega; u_\infty(w) < \infty\}$  and  $A_2 = \{w \in \Omega; u_\infty(w) = \infty\}$ . When  $u_\infty < \infty$ , we have  $u_j = u_\infty$  for  $j$  large enough since  $\mathbb{U}$  is countable. This ensures that  $E(U_{n-1}, u_j) = E(U_{n-1}, u_\infty)$ , *a.s.* for  $j$  large enough. Thus, we get

$$\lambda_n(u_j, t) \mathbf{1}_{A_1} \xrightarrow{j \rightarrow \infty} \sum_{e \in E(U_{n-1}, u_\infty)} \psi(e, U_{n-1}, t + T_{n-1}, r_n(t)) \mathbf{1}_{A_1}, \quad a.s.$$

When  $u_\infty = \infty$ , we have  $\sum_{e \in E(U_{n-1}, u_\infty)} \lambda_n(e, t) = 0$  since  $E(U_{n-1}, u_\infty) = \emptyset$ . Using  $\sum_{e \in E} \lambda_{n_j}(e, t) < \infty$ , *a.s.*, see Step (ii) in the proof of Theorem 3.1, we deduce that  $\sum_{e \in E(U_{n_j-1}, C^c)} \lambda_{n_j}(e, t) \xrightarrow{c \rightarrow \infty} 0$ , *a.s.* with  $C^c = \{u \in \mathbb{U}; u > c, c.w\}$ ,  $c > 0$  and  $c.w$  means component-wise. Since  $E(U_{n_j-1}, u_j) \subset E(U_{n_j-1}, C^c)$  for  $j$  large enough, we get  $\sum_{e \in E(U_{n_j-1}, u_j)} \lambda_{n_j}(e, t) \xrightarrow{j \rightarrow \infty} 0$ , *a.s.* which means that

$$\lambda_n(u_j, t) \mathbf{1}_{A_2} \xrightarrow{j \rightarrow \infty} \sum_{e \in E(U_{n-1}, u_\infty)} \psi(e, U_{n-1}, t + T_{n-1}, r_n(t)) \mathbf{1}_{A_2} = 0, \quad a.s.,$$

and proves  $\lambda_n(u_j, t) \xrightarrow{j \rightarrow \infty} \lambda_n(u_\infty, t)$ , *a.s.*

Additionally, we have  $\mathbb{E}[\sup_{n,s} \sum_e \lambda_n(e, s)] < \infty$ , see Step (iv) in the proof of Theorem 3.1. Thus, we get  $\mathbb{E}[\sup_{n,u,s} \lambda_n(u, s)] < \infty$ . Since  $\sum_e \lambda_n(e, s) \geq \underline{\psi}$  under Assumption 3.3, we have  $Z_n(t) \leq e^{-\underline{\psi}t}$ , *a.s.* Then, we can apply the dominated convergence theorem to show that

$$\int_{\mathbb{R}_+} Z_n(t)\lambda_n(t, u_j) dt \xrightarrow{j \rightarrow \infty} \int_{\mathbb{R}_+} Z_n(t)\lambda_n(t, u_\infty) dt, \quad a.s.$$

Furthermore, we have

$$\mathbb{E}[\sup_j \int_{\mathbb{R}_+} Z_n(t)\lambda_n(u_j, t) dt] \leq \mathbb{E}[\int_{\mathbb{R}_+} \sup_j e^{-\underline{\psi}t} \lambda_n(u_j, t) dt] \stackrel{\text{Fubini}}{=} \int_{\mathbb{R}_+} e^{-\underline{\psi}t} \mathbb{E}[\sup_j \lambda_n(u_j, t)] dt,$$

with  $\mathbb{E}[\sup_j \lambda_n(u_j, t)] < \infty$ . Hence, we can use the conditional dominated convergence to show

$$\mathbb{E}\left[\int_{\mathbb{R}_+} Z_n(t) \lambda_n(u_j, t) dt \mid \mathcal{F}_{k \leq r \leq n-1}\right] \xrightarrow{j \rightarrow \infty} \mathbb{E}\left[\int_{\mathbb{R}_+} Z_n(t) \lambda_n(u_\infty, t) dt \mid \mathcal{F}_{k \leq r \leq n-1}\right].$$

Finally, since  $\mathcal{F}_{k \leq r \leq n-1} \xrightarrow{k \rightarrow \infty} \mathcal{F}_{r \leq n-1}$ , we can apply Lemma IV.D.1 to deduce that

$$p_k^n \xrightarrow{k \rightarrow \infty} 0, \text{ a.s.}$$

This completes the proof.  $\square$

Let  $\Delta T_n = T_n - T_{n-1}$  be the inter-arrival time between  $n$ -th jump and the  $n-1$ -th jump with  $T_n$  the time of the  $n$ -th event. Let  $N^\infty = (T_i^\infty, X_i^\infty)$  be the stationary process constructed in Lemma 3.1 and  $N = (T_i, X_i)$  be a point process whose intensity satisfies (2). We write  $U^\infty = (Q^{1^\infty}, Q^{2^\infty}, S^\infty)$  (resp.  $U = (Q^1, Q^2, S)$ ) for the order book state associated to  $N^\infty$  (resp.  $N$ ). We denote by  $\lambda^\infty$  (resp.  $\lambda$ ) the intensity of  $N^\infty$  (resp.  $N$ ). We have the following result.

**Lemma IV.D.6.** *Under Assumptions 3.1, 3.2, 3.3 and 3.4, the process  $(U_n)_{n \geq 0}$  is  $f$ -geometrically ergodic, see 15.7 in [123], in the sense that there exists  $r > 1$  such that*

$$\sup_{\mathbf{x} \in W_0} \sum_{n \geq 1} \mathbb{E}_{\mathbf{x}}[\|f(U_n) - f(U_n^\infty)\| r^n] < \infty.$$

*Proof.* Let  $P^n(\mathbf{x}, A)$  be the probability of being in the set  $A = \{(t_k, x_k)_{k \leq 0} \in W_0; x_k = (n_k, t_k, b_k, s_k, \tilde{u}_k, u_k, a_k), u_0 \in a\}$ ,  $a \in \mathcal{U}$ , with  $\mathcal{U}$  the  $\sigma$ -algebra generated by the discrete topology on  $\mathbb{U}$ , after  $n$  jumps conditional on  $\mathbf{x} = (t_k, x_k)_{k \leq 0} \in W_0 = (\mathbb{R}_+ \times E)^{\mathbb{N}^-}$ . Let  $\mathbf{y} \in W_0$ . We write  $\pi$  for the stationary distribution of the process  $U_n^\infty = (Q_n^{1^\infty}, Q_n^{2^\infty}, S_n^\infty)$  and  $\tau_{\alpha^k}$  for the first entrance time of  $U$  to the set  $\alpha^k = \{\mathbf{z} \in W_0; \mathbf{z}_{-k+1 \leq j \leq 0} = \mathbf{y}_{-k+1 \leq j \leq 0}\}$ . Using the first-entrance last-exit decomposition of  $P^n(\mathbf{x}, A)$ , see Section 8.2 in [123], we have

$$\begin{aligned} P^n(\mathbf{x}, A) &= \alpha^k P^n(\mathbf{x}, A) + \sum_{j=1}^n \sum_{i=1}^j \left[ \int_{\mathbb{U}_{\mathbf{u}, \alpha^k}^{j-i}} \int_{\mathbb{U}_{\mathbf{x}, \alpha^k}^i} \alpha^k P^i(\mathbf{x}, d\mathbf{u}) P^{j-i}(\mathbf{u}, d\mathbf{v}) \alpha^k P^{n-j}(\mathbf{v}, A) \right] \\ &= \alpha^k P^n(\mathbf{x}, A) + \sum_{j=1}^n \sum_{i=1}^j \left[ \int_{\mathbb{U}_{\mathbf{u}, \alpha^k}^{j-i}} \int_{\mathbb{U}_{\mathbf{x}, \alpha^k}^i} \alpha^k P^i(\mathbf{x}, d\mathbf{u}) P^{j-i}(\mathbf{u}, d\mathbf{v}) \alpha^k P^{n-j}(\mathbf{y}, A) \right] \\ &\quad + \sum_{j=1}^n \sum_{i=1}^j \left[ \int_{\mathbb{U}_{\mathbf{u}, \alpha^k}^{j-i}} \int_{\mathbb{U}_{\mathbf{x}, \alpha^k}^i} \alpha^k P^i(\mathbf{x}, d\mathbf{u}) P^{j-i}(\mathbf{u}, d\mathbf{v}) | \alpha^k P^{n-j}(\mathbf{v}, A) - \alpha^k P^{n-j}(\mathbf{y}, A) | \right]. \end{aligned} \quad (35)$$

with  $\alpha^k P^n(\mathbf{x}, A) = \mathbb{P}[(T_k, X_k)_{k \leq 0} = \mathbf{x}, U_n \in A, \tau_{\alpha^k} \geq n]$  and  $\mathbb{U}_{\mathbf{x}, \alpha^k}^i = \{\mathbf{z} \in \alpha^k; (\mathbf{z}_k)_{k \leq i} = \mathbf{x}\}$ . Using  $\mathbb{E}_{\mathbf{x}}[\tau_{\alpha^k}] < \infty$  for all  $\mathbf{x} \in S$  and the arguments used in the proof of Theorem 10.2.1 in [123], we deduce that the stationary distribution admits the following representation:

$$\pi(A) = \mathbb{E}_{\mathbf{y}}[\tau_{\alpha^k}]^{-1} \mathbb{E}_{\mathbf{y}}\left[\sum_{j=1}^{\tau_{\alpha^k}} \mathbf{1}_{\tilde{U}_n \in A}\right] = \pi(\alpha^k) \sum_{j=1}^{\infty} \alpha^k P^j(\mathbf{y}, A). \quad (36)$$



#### IV. From asymptotic properties of general point processes to the ranking of financial agents

By combining (35) and (36), we get

$$\begin{aligned} P(\mathbf{x}, A) - \pi(A) &= \alpha^k P^n(\mathbf{x}, A) + \left[ (\alpha^k P(\mathbf{x}) * P(\alpha^k) - \pi(\alpha^k)) * \alpha^k P(\mathbf{y}) \right]_n(A) + \pi(\alpha^k) \sum_{j \geq n+1} \alpha^k P^j(\mathbf{y}, A) \\ &\quad + (\alpha^k P(\mathbf{x}) * P(\alpha^k)) * (\alpha^k P(\alpha^k) - \alpha^k P(\mathbf{y}))_n(A). \end{aligned} \quad (37)$$

with  $*$  the integrated Cauchy product between two sequences which is defined as follows:

$$[u(B) * v(C)]_n(A) = \sum_{i=1}^n \int_{\cup_{B,C}^i} u_i(B, du) v_{n-i}(\mathbf{u}, A), \quad \forall (B, C, A) \in (\mathcal{W}_0)^3,$$

with  $(u_n)_{n \geq 0}$  and  $(v_n)_{n \geq 0}$  two sequences such that  $u_n, v_n : (\mathcal{W}_0)^2 \rightarrow \mathbb{R}$ . Let  $f$  be the function defined in Lemma IV.D.4,  $\pi(f) = \int_{\mathcal{W}_0} \pi(du) f(u) < \infty$ ,  $\mathbb{E}_{\mathbf{x}}[f(U_n)] = \int_{\mathcal{W}_0} P^n(\mathbf{x}, du) f(u)$  and  $|P^n(\mathbf{x}, \cdot) - \pi|_f = |\mathbb{E}_{\mathbf{x}}[f(U_n)] - \pi(f)|$ . Using (37), we have

$$\begin{aligned} |P^n(\mathbf{x}, \cdot) - \pi|_f &\leq \mathbb{E}_{\mathbf{x}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}] + [(\alpha^k P(\mathbf{x}) * P(\alpha^k) - \pi(\alpha^k)) * t_n^f] \\ &\quad + \pi(\alpha^k) \sum_{j \geq n+1} t_j^f + |(\alpha^k P(\mathbf{x}) * P(\alpha^k)) * \Delta t_n^f|, \end{aligned} \quad (38)$$

with  $t_n^f = \mathbb{E}_{\mathbf{y}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}]$  and  $\Delta t_n^f(\mathbf{v}) = (\mathbb{E}_{\mathbf{v}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}] - t_n^f)$ . To prove geometric ergodicity we have to show

$$\sup_{\mathbf{x}} \sum_{n \geq 1} |P^n(\mathbf{x}, \cdot) - \pi|_f r^n < \infty, \quad (39)$$

with  $r > 1$ . Let us take  $\bar{n} \in \mathbb{N}^*$  and the delay  $k(\bar{n}) \in \mathbb{N}$  associated to  $\alpha^k$  depending on  $\bar{n}$ . Using (38), we have

$$\begin{aligned} \sum_{n \geq 1} |P^n(\mathbf{x}, \cdot) - \pi|_f r^n &\leq \sum_{n \geq 1} \mathbb{E}_{\mathbf{x}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}] r^n + \sum_{n \geq 1} |(\alpha^k P(\mathbf{x}) * P(\alpha^k) - \pi(\alpha^k)) * t_n^f r^n| \\ &\quad + \pi(\alpha^k) \sum_{n \geq 1} \sum_{j \geq n+1} t_j^f r^n + \sum_{n \geq 1} [(\alpha^k P(\mathbf{x}) * P(\alpha^k)) * \Delta t_n^f] r^n = \text{(i)} + \text{(ii)} + \text{(iii)} + \text{(iv)}. \end{aligned}$$

The error term (i) can be dominated by

$$\sum_{n \geq 1} \mathbb{E}_{\mathbf{x}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}] r^n \leq \sum_{n \geq 1} \mathbb{E}_{\mathbf{x}}[f(U_n) \mathbf{1}_{\tau_{\alpha^k} \geq n}] r^n = \mathbb{E}_{\mathbf{x}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right]. \quad (40)$$

The error term (iii) can be bounded by

$$\pi(\alpha^k) \sum_{n \geq 1} \sum_{j \geq n+1} t_j^f r^n \leq \pi(\alpha^k) \sum_{n \geq 1} \sum_{j \geq n+1} t_j^f r^n \leq \frac{\pi(\alpha^k)}{r-1} \sup_{\mathbf{v}} \mathbb{E}_{\mathbf{v}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right]. \quad (41)$$

Now we move to the error term (iv). We have

$$(\alpha^k P(\mathbf{x}) * P(\alpha^k)) * \Delta t_n^f \leq \sum_{j \leq n, i \leq j} \left[ \int_{\cup_{\mathbf{u}, \alpha^k}^{j-i} \times \cup_{\mathbf{x}, \alpha^k}^i \times W_0} \alpha^k P^i(\mathbf{x}, du) P^{j-i}(\mathbf{u}, dv) \right] \Delta_{\alpha^k} P^{n-j}(dw) f(w),$$

with  $\Delta_{\alpha^k} P^{n-j}(dw) = |\alpha^k P^{n-j}(\mathbf{v}, dw) - \alpha^k P^{n-j}(\mathbf{y}, dw)|$ . Using Equations (35) and (36), we get

$$\left\{ \begin{array}{l} \Delta_{\alpha^k} P^{n-j}(dw) \leq \alpha^k P^{n-j}(\mathbf{v}, dw) + \alpha^k P^{n-j}(\mathbf{y}, dw), \\ \sum_{j \leq n, i \leq j} \left[ \int_{\cup_{\mathbf{u}, \alpha^k}^{j-i} \times \cup_{\mathbf{x}, \alpha^k}^i \times \mathcal{W}} \alpha^k P^i(\mathbf{x}, du) P^{j-i}(\mathbf{u}, dv) \right] \alpha^k P^{n-j}(\mathbf{v}, dw) f(w) \leq \mathbb{E}_{\mathbf{x}}[f(U_n) r^n] < \infty, \\ \sum_{j \leq n, i \leq j} \left[ \int_{\cup_{\mathbf{u}, \alpha^k}^{j-i} \times \cup_{\mathbf{x}, \alpha^k}^i \times \mathcal{W}} \alpha^k P^i(\mathbf{x}, du) P^{j-i}(\mathbf{u}, dv) \right] \alpha^k P^{n-j}(\mathbf{y}, dw) f(w) \leq \mathbb{E}_{\mathbf{y}}[f(U_{n-j}) r^n] < \infty, \end{array} \right.$$

Since  $\Delta_{\alpha^k} P^{n-j} \xrightarrow[k \rightarrow \infty]{} 0$ , see Lemma IV.D.5, the dominated convergence theorem ensures that

$$(\alpha^k P(\mathbf{x}) * P(\alpha^k)) * \Delta t_n^f \xrightarrow[k \rightarrow \infty]{} 0.$$

Thus, there exists  $\bar{k}(\bar{n})$  such that  $(\alpha^k P(\mathbf{x}) * P(\alpha^k)) * \Delta t_n^f \leq \epsilon(\bar{n})$  for any  $k \geq \bar{k}(\bar{n})$ . Hence the error term (iv) can be majorated by

$$\sum_{n \geq 1}^{\bar{n}} [\alpha^k P(\mathbf{x}) * P(\alpha^k)] * \Delta t_n^f r^n \leq \epsilon(\bar{n}) \frac{r^{\bar{n}+1} - 1}{r - 1}, \quad (42)$$

which means that we have to choose  $\epsilon(\bar{n}) < c_1 \frac{r-1}{r^{\bar{n}+1}-1}$  with  $c_1$  a positive constant. Finally, using the property

$$\lim_{n \rightarrow \infty} (u * v)_n = \lim_{n \rightarrow \infty} u_n \times \lim_{n \rightarrow \infty} v_n, \quad (43)$$

we dominate the error term (ii) by

$$\left| \sum_{n \geq 1}^{\bar{n}} (\alpha^k P(\mathbf{x}) * P(\alpha^k) - \pi(\alpha^k)) * t_n^f r^n \right| \leq \left( \sum_{n \geq 1} |[\alpha^k P(\mathbf{x}) * P(\alpha^k)]_n(\alpha^k) - \pi(\alpha^k)| r^n \right) \sup_{\mathbf{v}} \mathbb{E}_{\mathbf{v}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right]. \quad (44)$$

Additionally, we have

$$\begin{aligned} |[\alpha^k P(\mathbf{x}) * P(\alpha^k)]_n(\alpha^k) - \pi(\alpha^k)| &= |[\alpha^k P(\mathbf{x}) * (P(\alpha^k) - \pi(\alpha^k))]_n(\alpha^k) - \pi(\alpha^k) \sum_{i \geq n+1} \alpha^k P^i(\mathbf{x}, \alpha^k)| \\ &= |[\alpha^k P(\mathbf{x}) * (P(\alpha^k) - P(\mathbf{y}))]_n(\alpha^k) + [\alpha^k P(\mathbf{x}) * (P(\mathbf{y}) - \pi(\alpha^k))]_n(\alpha^k) \\ &\quad - \pi(\alpha^k) \sum_{i \geq n+1} \alpha^k P^i(\mathbf{x}, \alpha^k)| \\ &\leq [\alpha^k P(\mathbf{x}) * |P(\mathbf{y}) - P(\alpha^k)|]_n(\alpha^k) + [\alpha^k P(\mathbf{x}) * |P(\mathbf{y}) - \pi(\alpha^k)|]_n(\alpha^k) \\ &\quad + \pi(\alpha^k) \sum_{i \geq n+1} \alpha^k P^i(\mathbf{x}^u, \alpha^k), \end{aligned}$$

#### IV. From asymptotic properties of general point processes to the ranking of financial agents

for any  $n \in \mathbb{N}$ . Using Equation (43), we get

$$\begin{aligned}
\sum_{n \geq 1}^{\bar{n}} [|\alpha^k P(\mathbf{x}) * P(\alpha^k) - \pi(\alpha^k)|]_n(\alpha^k) r^n &\leq \sum_{n \geq 1}^{\bar{n}} [\alpha^k P(\mathbf{x}) * |P^Y(\alpha^k) - P(\alpha^k)|]_n(\alpha^k) r^n \\
&+ \sum_{n \geq 1}^{\bar{n}} [\alpha^k P(\mathbf{x}) * |P(\mathbf{y}) - \pi(\alpha^k)|]_n(\alpha^k) r^n + \pi(\alpha^k) \sum_{\substack{n \geq 1 \\ i \geq n+1}} \alpha^k P^i(\mathbf{x}, \alpha^k) r^n \\
&\leq \left( \sum_{n \geq 1} \alpha^k P^n(\mathbf{x}, \alpha^k) r^n \right) \left( \sum_{n \geq 1} \sup_{\mathbf{w} \in \alpha^k} |P^n(\mathbf{y}, \alpha^k) - P^n(\mathbf{w}, \alpha^k)| r^n \right) \\
&+ \left( \sum_{n \geq 1} \alpha^k P^n(\mathbf{x}, \alpha^k) r^n \right) \left( \sum_{n \geq 1} |P^n(\mathbf{y}, \alpha^k) - \pi(\alpha^k)| r^n \right) \\
&+ \pi(\alpha^k) \sum_{n \geq 1} \sum_{i \geq n+1} \alpha^k P^i(\mathbf{x}, \alpha^k) r^n = (1) + (2) + (3).
\end{aligned}$$

The term (2) is bounded by

$$(2) \leq \mathbb{E}_{\mathbf{x}}[r^{\tau_{\alpha^k}}] \sup_{\mathbf{y}} \left( \sum_{n \geq 1} |P^n(\mathbf{y}, \alpha) - \pi(\alpha)| r^n \right).$$

Since the Kendall theorem ensures that  $E_{\mathbf{x}}[r^{\tau_{\alpha^k}}] < \infty$  and  $\sum_{n \geq 1} |P^n(\mathbf{x}, \alpha^k) - \pi(\alpha^k)| r^n < \infty$  are equivalent, the quantity (1) is finite if and only if  $\sup_{\mathbf{v}} E_{\mathbf{v}}[r^{\tau_{\alpha^k}}] < \infty$ . The term (1) is majorated by

$$(1) \leq \mathbb{E}_{\mathbf{x}}[r^{\tau_{\alpha^k}}] \left( \sum_{n \geq 1} \sup_{\mathbf{w} \in \alpha^k} |P^n(\mathbf{w}, \alpha^k) - P^n(\mathbf{y}, \alpha^k)| r^n \right).$$

To ensure that the sequence  $v(\bar{n}) = \sum_{n \geq 1}^{\bar{n}} \sup_{\mathbf{w} \in \alpha^k} |P^n(\mathbf{w}, \alpha^k) - P^n(\mathbf{y}, \alpha^k)| r^n$  is bounded, the put a dependence  $k$  and  $\bar{n}$ . Let  $\epsilon^1(\bar{n}) > 0$ . By following the same arguments used in the proof of Inequality (42), there exists  $\bar{k}^1(\bar{n})$  such that for any  $k \geq \bar{k}^1(\bar{n})$ , we have

$$\sum_{n \geq 1} \sup_{\mathbf{w} \in \alpha^k} |P^n(\mathbf{y}, \alpha^k) - P^n(\mathbf{w}, \alpha^k)| r^n \leq \epsilon^1(\bar{n}) \frac{r^{\bar{n}+1} - 1}{r - 1}.$$

By taking  $\epsilon^1(\bar{n}) \leq c_1 \frac{r-1}{r^{\bar{n}+1}-1}$ , we get (1)  $\leq c_1 \sup_{\mathbf{x}} \mathbb{E}_{\mathbf{x}}[r^{\tau_{\alpha^k}}]$ . Furthermore, the term (3) can be dominated by (3)  $\leq \mathbb{E}_{\mathbf{x}}[r^{\tau_{\alpha^k}}]$ . Thus, we deduce that

$$(ii) \leq c_1 \mathbb{E}_{\mathbf{x}}[r^{\tau_{\alpha^k}}] \left( 1 + \sup_{\mathbf{v}} \sum_{n \geq 1} |P^n(\mathbf{v}, \alpha) - \pi(\alpha)| r^n \right) \sup_{\mathbf{v}} \mathbb{E}_{\mathbf{v}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right]. \quad (45)$$

By combining Inequalities (40), (41), (44) and (45), we have (39) when  $\sup_{\mathbf{x}} E_{\mathbf{x}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right]$  and  $\sup_{\mathbf{x}} E_{\mathbf{x}}[r^{\tau_{\alpha^k}}]$  are both finite. Since  $E_{\mathbf{x}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right] < \infty$  implies  $E_{\mathbf{x}}[r^{\tau_{\alpha^k}}] < \infty$ , we only need to prove

$$\mathbb{E}_{\mathbf{x}} \left[ \sum_{n=1}^{\tau_{\alpha^k}} f(U_n) r^n \right] < \infty.$$

This last inequality is satisfied thanks to Lemma IV.D.4. □

**Lemma IV.D.7.** *Under Assumptions 3.1, 3.3 and 3.4, the process  $\bar{U}$  is ergodic.*

*Proof of Lemma IV.D.7.* For simplicity, we write  $c_1$ ,  $c_2$  and  $c_3$  for positive constants and forget the dependence of  $\mathbb{E}_{\mathbf{x}}[X]$  on the initial state  $\mathbf{x}$  for any random variable  $X$ . Let  $N^\infty = (T_i^\infty, X_i^\infty)$  be the stationary process constructed in Lemma 3.1 and  $N = (T_i, X_i)$  be a point process whose intensity satisfies (2). We write  $U^\infty = (Q^{1^\infty}, Q^{2^\infty}, S^\infty)$  (resp.  $U = (Q^1, Q^2, S)$ ) for the order book state associated to  $N^\infty$  (resp.  $N$ ). We denote by  $\lambda^\infty$  (resp.  $\lambda$ ) the intensity of  $N^\infty$  (resp.  $N$ ). To prove the uniqueness, we need to show that  $N$  and  $N^\infty$  couple which is satisfied when

$$\int_0^\infty \sup_e \mathbb{E}[|\lambda_t(e) - \lambda_t^\infty(e)|] dt < \infty,$$

thanks to Lemma IV.D.3. We write  $f(t) = \sup_e \mathbb{E}[|\lambda_t(e) - \lambda_t^\infty(e)|]$  for any  $t \geq 0$ .

**Step (i):** For any  $\gamma = \frac{p}{q} > 1$  with  $p, q \in \mathbb{N}^*$  and  $\beta$  such that  $\frac{1}{\beta} + \frac{1}{\gamma} = 1$ . Let us first prove that

$$f(t) \leq u(t) + g_1(t)G\left(\int_0^t \bar{h}(t-s)f(s) ds\right), \quad (46)$$

with  $u(t) = c_3\mathbb{E}[\|U_t - U_t^\infty\|] + c_1\mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^{\frac{1}{\beta p}}[1 + 2B(t)]$ ,  $g_1(t) = c_2(1 + 2B(t))$ ,  $G(t) = t^{\frac{1}{\beta}}$ ,  $\bar{h}(t) = \sup_{e,u,x} \phi(e, u, t, x)$  and  $B(t) = \sup_{0 \leq k \leq n_\psi - 1} [B_k(t)]^{\frac{1}{\beta \gamma}}$  with  $B_k(t)$  defined in Equation (50). The quantities  $c_1$ ,  $c_2$  and  $c_3$  are positive constants. We have

$$\begin{aligned} f(t) &= \mathbb{E}[|\psi(e, U_t, t, r_t) - \psi(e, U_t^\infty, t, r_t^\infty)|] \\ &\leq \mathbb{E}[|\psi(e, U_t, t, r_t) - \psi(e, U_t^\infty, t, r_t)|] + \mathbb{E}[|\psi(e, U_t^\infty, t, r_t) - \psi(e, U_t^\infty, t, r_t^\infty)|] \\ &= (1) + (2), \end{aligned}$$

with  $r_t = \int_0^t \phi(e, U_t, t-s, X_s) dN_s$  and  $r_t^\infty = \int_0^t \phi(e, U_t^\infty, t-s, X_s^\infty) dN_s^\infty$ . Let us first handle the term (2). Using Assumption 3.4, we have

$$\begin{aligned} \mathbb{E}[|\psi(e, U_t^\infty, t, r_t) - \psi(e, U_t^\infty, t, r_t^\infty)|] &\leq \mathbb{E}[|\bar{\psi}(r_t) - \bar{\psi}(r_t^\infty)|] \\ &\leq K\mathbb{E}[|r_t - r_t^\infty| |1 + r_t^{n_1} + r_t^{\infty n_1}|] \\ &\leq K \underbrace{\mathbb{E}[|r_t - r_t^\infty|^\beta]^{\frac{1}{\beta}}}_{(i)} \underbrace{\mathbb{E}[|1 + r_t^{n_1} + r_t^{\infty n_1}|^\gamma]^{\frac{1}{\gamma}}}_{(ii)}. \end{aligned}$$

The term (i) can be dominated by

$$\begin{aligned}
 \mathbb{E}[|r_t - r_t^\infty|^\beta]^\frac{1}{\beta} &\leq \mathbb{E}\left[\left|\int_0^t \phi(e, U_t, t-s, X_s) dN_s - \phi(e, U_t^\infty, t-s, X_s) dN_s^\infty\right|^\beta\right]^\frac{1}{\beta} \\
 &\leq 2^\frac{\beta-1}{\beta} \mathbb{E}\left[\left|\int_0^t \phi(e, U_t, t-s, X_s) dN_s - \phi(e, U_t^\infty, t-s, X_s) dN_s\right|^\beta\right]^\frac{1}{\beta} \\
 &\quad + 2^\frac{\beta-1}{\beta} \mathbb{E}\left[\left(\int_0^t \bar{h}(t-s) |dN_s - dN_s^\infty|\right)^\beta\right]^\frac{1}{\beta} \\
 &\leq 2^\frac{\beta-1}{\beta} \mathbb{E}[\|U_t - U_t^\infty\|^\beta] \int_0^t \bar{h}(e, t-s, X_s) dN_s|^\beta]^\frac{1}{\beta} + 2^\frac{\beta-1}{\beta} \left[\int_0^t \bar{h}(t-s) f(s) ds\right]^\frac{1}{\beta} \\
 &\leq 2^\frac{\beta-1}{\beta} \mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^\frac{1}{\beta p} \mathbb{E}\left[\left|\int_0^t \bar{h}(e, t-s, X_s) dN_s\right|^{\beta q}\right]^\frac{1}{\beta q} + 2^\frac{\beta-1}{\beta} \left[\int_0^t \bar{h}(t-s) f(s) ds\right]^\frac{1}{\beta} \\
 &= c_1 \mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^\frac{1}{\beta p} + c_2 \left[\int_0^t \bar{h}(t-s) f(s) ds\right]^\frac{1}{\beta}, \tag{47}
 \end{aligned}$$

with  $\bar{h}(s) = \sup_{e,u,x} \phi(e, u, s, x)$ ,  $\tilde{h}(e, s, x) = \frac{2}{\min(\alpha_0, 1)} \sup_u \phi(e, u, s, x)$  and  $\min(\alpha_0, 1)$  represents the minimum distance between two elements in the countable space  $\mathbb{U}$ . The quantity  $\mathbb{E}\left[\left|\int_0^t \tilde{h}(e, t-s, X_s) dN_s\right|^{\beta q}\right]$  is bounded since

$$\begin{aligned}
 \mathbb{E}\left[\left|\int_0^t \tilde{h}(e, t-s, X_s) dN_s\right|^{\beta q}\right] &\leq \mathbb{E}\left[\left|\int_0^t \tilde{h}(e, t-s, X_s) dN_s\right|^q\right]^\frac{1}{\beta} \\
 &\leq \left\{ \sum_{\mathbf{k}_m \in \mathcal{D}(q)} \sum_{\tilde{x} \in E^m} \binom{q}{\mathbf{k}_m} \times \int_{(-\infty, t)^m} \mathbb{E}\left[\prod_{i=1}^m \tilde{h}(e, t-s_i, x_i) dN_{s_i}\right] \right\}^\frac{1}{\beta} \\
 &\leq \left\{ \sum_{\mathbf{k}_m \in \mathcal{D}(q)} \sum_{\tilde{x} \in E^m} \binom{q}{\mathbf{k}_m} \times \int_{(-\infty, t)^m} \prod_{i=1}^m \tilde{h}(e, t-s_i, x_i) \mathbb{E}[\lambda_{s_i}] ds_i \right\}^\frac{1}{q} < \infty.
 \end{aligned}$$

The term (ii) satisfies

$$\mathbb{E}\left[|1 + r_t^{n_1} + r_t^{\infty n_1}|^\gamma\right]^\frac{1}{\gamma} \leq 3^\frac{\gamma-1}{\gamma} (1 + \mathbb{E}[|r_t^{n_1}|^\gamma]^\frac{1}{\gamma} + \mathbb{E}[|r_t^{\infty n_1}|^\gamma]^\frac{1}{\gamma}), \tag{48}$$

with  $\gamma = \frac{p}{q}$  and  $p, q \in \mathbb{N}^*$ . We have

$$\begin{aligned}
 \mathbb{E}[|r_t^{n_1}|^\frac{p}{q}] &\leq \mathbb{E}[|r_t^{n_1}|^p]^\frac{1}{q} \\
 &= \mathbb{E}\left[\left(\int_0^t \phi(e, U_t, t-s, X_s) dN_s\right)^{n_1 p}\right]^\frac{1}{q} \\
 &= \left\{ \sum_{\mathbf{k}_m \in \mathcal{D}(\bar{p})} \sum_{\tilde{x} \in E^m} \binom{\bar{p}}{\mathbf{k}_m} \times \int_{(-\infty, t)^m} \mathbb{E}\left[\prod_{i=1}^m \bar{\phi}(t-s_i, x_i) dN_{s_i}\right] \right\}^\frac{1}{q}, \tag{49}
 \end{aligned}$$

with  $\bar{\phi}(t, x) = \sup_{e,u} \phi(e, u, t, x)$  and  $\bar{p} = n_1 p$ . Using (49) and the Brascamp-Lieb inequality, we have

$$\begin{aligned}
 \mathbb{E}[|r_t^{n_1}|^\frac{p}{q}] &\leq \left[ \sum_{\mathbf{k}_m \in \mathcal{D}(\bar{p})} \sum_{\tilde{x} \in E^m} \binom{\bar{p}}{\mathbf{k}_m} \times \int_{(-\infty, t)^m} \prod_{i=1}^m \bar{\phi}(t-s_i, x_i) \mathbb{E}[\lambda_{s_i}]^\frac{1}{m} ds_i \right]^\frac{1}{q} \\
 &= \left[ \sum_{\mathbf{k}_m \in \mathcal{D}(\bar{p})} \sum_{\tilde{x} \in E^m} \binom{\bar{p}}{\mathbf{k}_m} R_m(t) \right]^\frac{1}{q} = B_k(t)^\frac{1}{q}, \tag{50}
 \end{aligned}$$

with  $R_m(t) = \int_{(-\infty, t)^m} \prod_{i=1}^m \bar{\phi}(t - s_i, x_i) \mathbb{E}[\lambda_{s_i}]^{\frac{1}{m+m'}} ds_i$  and  $B_k(t) = \sum_{\mathbf{k}_m \in \mathcal{D}(\bar{p})} \sum_{\bar{x} \in E^m} \binom{\bar{p}}{\mathbf{k}_m} R_m(t)$ . Similarly, we also have

$$\mathbb{E}[|r_t^{\infty n_1}|^{\frac{p}{q}}] \leq B_k(t)^{\frac{1}{q}}. \quad (51)$$

Using Inequalities (48) and (50), we deduce that (ii) verifies

$$\mathbb{E}[|1 + r_t^{n_1} + r^{\infty n_1}|^{\frac{1}{\gamma}}] \leq 3^{\frac{\gamma-1}{\gamma}} (1 + 2 \sup_{0 \leq k \leq n_\psi - 1} [B_k(t)]^{\frac{1}{q\gamma}}). \quad (52)$$

By combining inequalities (47) and (52), we deduce that

$$(2) \leq 3^{\frac{\gamma-1}{\gamma}} \left[ c_1 \mathbb{E}[\|U_t - U_t^\infty\|^{\frac{1}{\beta p}}] + c_2 \left[ \int_0^t g(t-s)f(s) ds \right]^{\frac{1}{\beta}} \right] \left[ 1 + 2 \sup_{0 \leq k \leq n_\psi - 1} [B_k(t)]^{\frac{1}{q\gamma}} \right]. \quad (53)$$

Using Theorem 3.1, we have  $\sup_{e,t} \mathbb{E}[\sup_u \psi(e, u, t, r_t)]$  is finite. Thus, there exists  $K$  such that

$$(1) \leq c_3 \mathbb{E}[\|U_t - U_t^\infty\|]. \quad (54)$$

Thus using Equations (53) and (54), we prove (46).

**Step (ii):** By a density argument, there exist continuous sequences of functions  $(u^p)_{p \geq 1}$ ,  $(g_1^p)_{p \geq 1}$  and  $(\bar{h}^p)_{p \geq 1}$  such that  $u^p(t) \xrightarrow{p \rightarrow \infty} u(t)$  and  $u \leq u^p$ ,  $g_1^p(t) \xrightarrow{p \rightarrow \infty} g_1(t)$  and  $g_1 \leq g_1^p$  and  $\bar{h}^p \xrightarrow{p \rightarrow \infty} \bar{h}$  and  $\bar{h} \leq \bar{h}^p$ . Thus, we have

$$f(t) \leq u^p(t) + g_1^p(t) G \left( \int_0^t \bar{h}^p(s) f(s) ds \right).$$

Using a density argument again, we can find a sequence of functions  $(f^k)_{k \geq 1}$  converges uniformly towards  $f$ . By affording ourselves to use sub-sequences, we can always consider that

$$f^p(t) \leq \tilde{u}^p(t) + g_1^p(t) G \left( \int_0^t \bar{h}^p(s) f^p(s) ds \right),$$

with  $\tilde{u}^p(t) = u^p(t) + |f - f^p|_\infty$ . Using Theorem 3 in [21] and Inequality (46), we have

$$f^p(t) \leq v^p(t) F^p(t) \left\{ 1 + G \left[ H^{-1} \left( \int_0^t \bar{h}^p(s) g_1^p(s) ds \right) \right] \right\},$$

with  $H(s) = \int_0^s \frac{dt}{1+G(t)}$ ,  $v^p(t) = \max(G_1(\tilde{u}^p)(t), 1)$ ,  $F^p(t) = \max(G_1(g_1^p)(t), 1)$  and

$$G_1(w)(t) = w(t) \left( 1 + \int_0^t w(s) \bar{h}^p(s) e^{\int_s^t \bar{h}^p g_1^p du} ds \right).$$

By sending  $p$  to infinity, we deduce that

$$f(t) \leq v(t) F(t) \left\{ 1 + G \left[ H^{-1} \left( \int_0^t \bar{h}(s) g_1(s) ds \right) \right] \right\}, \quad (55)$$

with  $v(t) = \max(G_1(u)(t), 1)$  and  $F(t) = \max(G_1(g_1)(t), 1)$ .

**Step (iii):** Let us prove that  $\int_{\mathbb{R}_+} u(t) dt < \infty$ . Since  $B(t)$  is uniformly bounded, we only need to prove that

$$\begin{cases} \int_{\mathbb{R}_+} \mathbb{E}[\|U_t - U_t^\infty\|] dt < \infty, \\ \int_{\mathbb{R}_+} \mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^{\frac{1}{\beta p}} dt < \infty. \end{cases}$$

Since  $0 < \underline{\psi} = \inf_{u,t,r} \sup_e \psi(e, u, t, r) \leq \lambda_n$ , we have

$$\begin{aligned} \mathbb{E}\left[\int_{\mathbb{R}_+} \|U_t - U_t^\infty\| dt\right] &\leq \mathbb{E}\left[\sum_{n \geq 0} \|U_n - U_n^\infty\| \int_{T_n}^{T_{n+1}} dt\right] \\ &= \mathbb{E}\left[\sum_{n \geq 0} \|U_n - U_n^\infty\| \mathbb{E}[T_{n+1} - T_n | \mathcal{F}_n]\right] \\ &\leq \mathbb{E}\left[\sum_{n \geq 0} \|U_n - U_n^\infty\| \frac{1}{\underline{\psi}}\right]. \end{aligned}$$

Using Lemma IV.D.6, we have  $\mathbb{E}[\sum_{n \geq 0} \|U_n - U_n^\infty\|] < \infty$  which ensures that  $\mathbb{E}[\int_{\mathbb{R}_+} \|U_t - U_t^\infty\| dt] < \infty$ . By using a similar methodology and the fact that  $\sum_{n \geq 0} \mathbb{E}[\|U_n - U_n^\infty\|^{\beta p}] r^n < \infty$  with  $r > 1$ , see Lemma IV.D.6, we also have  $\int_{\mathbb{R}_+} \mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^{\frac{1}{\beta p}} dt < \infty$ .

**Step (iv):** Since  $g_1$  is bounded and  $\int_0^t \bar{h}(s) ds < \infty$ , the functions  $F(t)$  and  $\left\{1 + G\left[H^{-1}\left(\int_0^t \bar{h}(s) g_1(s) ds\right)\right]\right\}$  are bounded as well. Moreover,  $\int_{\mathbb{R}_+} u(t) dt < \infty$  thanks to the previous step. Thus, by applying Inequality (55), we have that  $\int_{\mathbb{R}_+} f(t) dt < \infty$  which completes the proof.  $\square$

### IV.D.3 Speed of convergence

**Lemma IV.D.8.** *We have the following error estimate:*

$$\|P_t(w, \cdot) - \bar{\pi}\|_{TV} \leq K_1 e^{-K_2 t}, \quad \forall w \in W,$$

with  $K_3 > 0$  and  $K_2 > 0$ .

*Proof of Lemma IV.D.8.* We forget the dependence of  $\mathbb{E}_x[X]$  on the initial state  $\mathbf{x}$  for any random variable  $X$ . We have

$$\begin{aligned} \|P_t(w, \cdot) - \bar{\pi}\|_{TV} &\leq \mathbb{P}[\sup_e |N_s - N_s^\infty| \neq 0, \quad \forall s \in (t, \infty)] \\ &= \left(1 - \mathbb{P}[\sup_e N_s = N_s^\infty, \quad \forall s \in (t, \infty)]\right) = (i). \end{aligned}$$

Using Lemma IV.D.2 and Jensen's Inequality, we have

$$(i) \leq 1 - e^{-\int_t^\infty f(s) ds},$$

with  $f(t) = \sup_e \mathbb{E}[|\lambda_t(e) - \lambda_t^\infty(e)|]$  for any  $t \geq 0$ . Using Inequality (55) and the boundedness of  $F$  and  $\left\{1 + G\left[H^{-1}\left(\int_0^t \bar{h}(s) g_1(s) ds\right)\right]\right\}$ , we have

$$(i) \leq c_1 \int_t^\infty u(t) dt, \tag{56}$$

with  $c_1$  a positive constant. Let us now prove that

$$u(t) \leq c_1 e^{-\alpha t}, \quad (57)$$

with  $\alpha$  a positive constant. We have

$$\mathbb{E}[\|U_t - U_t^\infty\|] = \mathbb{E}[\|U_{N(t)} - U_{N^\infty(t)}^\infty\|] \leq \mathbb{E}[\|U_{N(t)} - U_{N(t)}^\infty\|] + \mathbb{E}[\|U_{N(t)}^\infty - U_{N^\infty(t)}^\infty\|].$$

Using the fact that  $\sum_{n \geq 1} \mathbb{E}[\|U_n - U_n^\infty\|] r^n < \infty$ , there exists  $\alpha > 0$  such that  $\mathbb{E}[\|U_n - U_n^\infty\|] \leq A e^{-\alpha n}$ . Let us denote by  $U_t^{\infty, \delta}$  the  $\delta$ -translated process defined such that  $U_t^{\infty, \delta} = U_{t+\delta}^\infty$ . By applying Lemma IV.D.7 to the process  $U^{\infty, \delta}$ , we also have  $\sup_\delta (\sum_{n \geq 1} \mathbb{E}[\|U_n^{\infty, \delta} - U_n^\infty\|] r^n) < \infty$  which ensures that  $\mathbb{E}[\|U_n^{\infty, \delta} - U_n^\infty\|] \leq A e^{-\alpha n}$ . Using Lemma IV.D.9 below and the uniqueness of the stationary distribution, we have  $\frac{N(t)}{t} \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mathbb{E}_\pi[\Delta T_1]}$  and  $\frac{N^\infty(t)}{t} \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mathbb{E}_\pi[\Delta T_1]}$ , *a.s.* Thus, we deduce that

$$\mathbb{E}[\|U_t - U_t^\infty\|] \leq c_1 e^{-\alpha t}. \quad (58)$$

Using the same lines of argument, we also have

$$\mathbb{E}[\|U_t - U_t^\infty\|^{\beta p}]^{\frac{1}{\beta p}} \leq c_1 e^{-\alpha t}. \quad (59)$$

By combining Inequalities (58) and (59) and using the expression of  $u(t)$ , we recover Inequality (57) which ensures that

$$(i) \leq c_1 e^{-\alpha t}.$$

This completes the proof.  $\square$

**Lemma IV.D.9.** *For any initial state  $u \in \mathcal{U}$ , the process  $\Delta T_n$  satisfies*

$$\frac{\sum_{i=1}^n \Delta T_i}{n} \xrightarrow[n \rightarrow \infty]{} \mathbb{E}_\mu[\Delta T_1] \quad \textit{a.s.},$$

with  $\mu$  the unique stationary distribution of the point process  $N$ .

*Proof.* Since there exists  $\underline{\lambda} > 0$  such that  $\inf_{t, u, r} \sum_{e \in E} \lambda_t(e, u, r) > \underline{\lambda}$ , we have  $\mathbb{E}[\Delta T_n] \leq \frac{1}{\underline{\lambda}}$  for any  $n \geq 1$ . Thus,  $\Delta T_n$  admits a finite stationary distribution. Using the Theorem 17.1.2 in [123], we complete the proof.  $\square$

## IV.E Proof of Propositions 4.1 and 4.2

*Proof of Proposition 4.1.* The proof of Equation (9) is a direct application of Theorem 2 in [63]. Since  $(U_n)$  is  $f$ -geometrically ergodic, see Lemma IV.D.6,  $(Y_n)$  is  $g$ -geometrically ergodic and  $U_n$  and  $Y_n$  are independent, the process  $(U_n, Y_n)$  is  $\tilde{f}$ -geometrically ergodic with  $\tilde{f}(u, y) = f(u) + g(y)$ . Let  $g$  and  $h$  be two functions such that  $g^2, h^2 \leq \tilde{f}$ ,  $\mu$  the stationary distribution



#### IV. From asymptotic properties of general point processes to the ranking of financial agents

of  $(U, Y)$  and  $\bar{v} = v - \mathbb{E}_\mu[v]$  for any function  $v$ . By following the same lines of argument of Lemma 16.1.5 in [123], we have

$$|\mathbb{E}_\pi[\bar{h}(Z_n)\bar{g}(Z_{n+k})]| \leq R\mathbb{E}_\pi[\bar{f}(Z_0)]r^k,$$

with  $Z_n = (U_n, Y_n)$ ,  $r < 1$  and  $R$  a positive constant. The quantity  $\mathbb{E}_\pi[\bar{f}(Z_0)]$  is bounded by Lemma IV.C.2. Thus  $Z$  is a geometric mixing and Theorems 19.1 and 19.2 in [32] give the result.  $\square$

*Proof of Proposition 4.2.* Using Lemma IV.D.9 and Proposition 4.2, the proof of this result is analogous to the proof of Theorem 4.2 in [91].  $\square$

### IV.F Stationary distribution computation

*Proof of Proposition 5.1.* Let  $z \in Z$  and  $z' \in Z$  such that  $z \neq z'$ . Since  $\zeta$  is stationary under  $\mu$ , we have

$$\sum_{z' \in Z} \int_{A^{z'}} \mu(dw) P_t(w, A^z) = \int_{W_0} \mu(dw) P_t(w, A^z) = \mu(A^z), \quad \forall t \geq 0, \quad (60)$$

with  $P_t(w, \cdot)$  the probability distribution of  $\zeta_t^{0,w}$  starting from the initial condition  $w$  and  $A^z = \{(w_s)_{s \leq 0} \in W_0; \zeta_0^{0,w} = z\}$ . Since  $\int_{A^{z'}} \mu(dw) P_t(w, A^z) = \mathbb{P}_\mu[\zeta_t = z, \zeta_0 = z'] = \mathbb{P}_\mu[\zeta_0 = z'] \mathbb{P}_\mu[\zeta_t = z | \zeta_0 = z']$  and  $\mu(A^z) = \mathbb{P}_\mu[\zeta_0 = z]$ , the quantity  $\pi(z) = \mu(A^z)$  defined in Section 5.1 satisfies

$$\sum_{z' \in Z} \pi(z') \mathbb{P}_\mu[\zeta_t = z | \zeta_0 = z'] = \pi(z), \quad \forall t \geq 0,$$

which also leads to the following equation:

$$\sum_{z' \in Z} \pi(z') \tilde{Q}(z, z') = 0, \quad \sum_{z' \in Z} \pi(z') = 1,$$

with  $\tilde{Q}(z, z') = \lim_{\delta \rightarrow 0} \frac{\mathbb{P}_\mu[\zeta_\delta = z' | \zeta_0 = z]}{\delta}$ . The quantity  $\tilde{Q}(z, z')$  satisfies

$$\begin{aligned} \tilde{Q}(z, z') &= \lim_{\delta \rightarrow 0} \frac{\mathbb{P}_\mu[U_\delta = z' | U_0 = z]}{\delta} = \lim_{\delta \rightarrow 0} \frac{\mathbb{P}_\mu[\{T_1 \leq \delta, e_1 \in E(z, z')\} | \zeta_0 = z] + \epsilon}{\delta} \\ &= \lim_{\delta \rightarrow 0} \frac{\mathbb{E}_\mu[\mathbb{P}[\{T_1 \leq \delta, e_1 \in E(z, z')\} | \mathcal{F}_0] | \zeta_0 = z] + \epsilon}{\delta} \\ &= \mathbb{E}_\mu[\lim_{\delta \rightarrow 0} \frac{\mathbb{P}[\{T_1 \leq \delta, e_1 \in E(z, z')\} | \mathcal{F}_0]}{\delta} | \zeta_0 = z] + \lim_{\delta \rightarrow 0} \frac{\epsilon}{\delta} \\ &= \mathbb{E}_\mu[\sum_{e_1 \in E(z, z')} \lambda_0(e_1) | \zeta_0 = z] + \lim_{\delta \rightarrow 0} \frac{\epsilon}{\delta} \\ &= \sum_{e_1 \in E(z, z')} \mathbb{E}_\mu[\lambda_0(e_1) | \zeta_0 = z] + \lim_{\delta \rightarrow 0} \frac{\epsilon}{\delta}, \end{aligned}$$

where  $\epsilon$  is an error term associated to the cases when at least two events happen in the interval  $[0, \delta]$ . Since  $\sum_{e_1 \in E} \mathbb{E}_\mu[\lambda_0(e_1)]$  is finite, we have  $\epsilon \leq c_1 \delta^2$  with  $c_1$  a positive constant. We deduce that

$$\tilde{Q}(z, z') = \sum_{e_1 \in E(z, z')} \mathbb{E}_\mu[\lambda_0(e_1, u)] = Q(z, z'). \quad (61)$$

This completes the proof.  $\square$

## IV.G Proof of Proposition 5.2

*Proof of Proposition 5.2.* We write  $\lambda_s^{u, u'} = \sum_{e \in E(u, u')} \lambda_s(e)$  and  $E(u, u')$  the set of events that moves the order book from the state  $u$  to  $u'$ . We have

$$\frac{N_t^{u, u'}}{t} = \frac{\int_0^t \lambda_s \delta_{u, u'}^s ds}{t} + \left( \frac{N_t^{u, u'} - \int_0^t \lambda_s \delta_{u, u'}^s ds}{t} \right). \quad (62)$$

Since  $(\lambda_s)_{s \geq 0}$  is stationary under  $\bar{\pi}$  and  $\mathbb{E}_{\bar{\pi}}[\lambda_s] < \infty$ , the Theorem 2.1-chapter X in [63] ensures that

$$\begin{aligned} \frac{\int_0^t \lambda_s \delta_{u, u'}^s ds}{t} &\xrightarrow{t \rightarrow \infty} \mathbb{E}_{\bar{\pi}}[\lambda_0 \delta_{u, u'}^0] = \sum_{e \in E(u, u')} \mathbb{E}_{\bar{\pi}}[\lambda_0(e) \delta_{u, u'}^0] = \sum_{e \in E(u, u')} \mathbb{E}_{\bar{\pi}}[\lambda_0(e) \delta_{u, u'}^0 | U_0 = u] \mathbb{P}_{\bar{\pi}}[U_0 = u] \\ &= \mathbb{P}_{\bar{\pi}}[U_0 = u] \sum_{e \in E(u, u')} \mathbb{E}_{\bar{\pi}}[\lambda_0(e) | U_0 = u] \\ &= \mathbb{P}_{\bar{\pi}}[U_0 = u] Q(u, u'), \quad a.s. \end{aligned} \quad (63)$$

Moreover, since  $N_t^{u, u'} - \int_0^t \lambda_s \delta_{u, u'}^s ds$  is a martingale and  $\sup_{s \geq 0, u, u'} \mathbb{E}[\lambda_s^{u, u'}] < \infty$ , we have

$$\frac{N_t^{u, u'} - \int_0^t \lambda_s \delta_{u, u'}^s ds}{t} \xrightarrow{t \rightarrow \infty} 0, \quad a.s. \quad (64)$$

Hence, by combining (62), (63) and (64), we prove  $\frac{N_t^{u, u'}}{t} \xrightarrow{t \rightarrow \infty} \mathbb{P}_{\bar{\pi}}[U_0 = u] Q(u, u')$ , *a.s.* On the other hand, we have

$$\frac{t^u}{t} = \frac{\int_0^t \delta_u^s ds}{t}. \quad (65)$$

Since  $(U_s)_{s \geq 0}$  is stationary under  $\bar{\pi}$  and  $\mathbb{E}_{\bar{\pi}}[\delta_u^s] < \infty$ , the Theorem 2.1-chapter X in [63] ensures that

$$\frac{\int_0^t \delta_u^s ds}{t} \xrightarrow{t \rightarrow \infty} \mathbb{E}_{\bar{\pi}}[\delta_u^0] = \mathbb{P}_{\bar{\pi}}[U_0 = u], \quad a.s. \quad (66)$$

Thus, we deduce that

$$\frac{N_t^{u, u'}}{t^u} = \frac{\frac{N_t^{u, u'}}{t}}{\frac{t^u}{t}} \xrightarrow{t \rightarrow \infty} Q(u, u'), \quad a.s.,$$

which completes the proof.  $\square$

#### IV. From asymptotic properties of general point processes to the ranking of financial agents

*Proof of confidence interval computation.* By applying Theorem 4.1 to the sequence of  $\eta_s = \lambda_s \delta_{u,u'}^s$  and use basic inequalities to approximate  $t$  by its integer part  $\lfloor t \rfloor$ , we have

$$\sqrt{t} \left( \frac{N_t^{u,u'}}{t} - \mathbb{P}_{\bar{\pi}}[U_0 = u] Q(u, u') \right) \xrightarrow{\mathcal{L}} \sigma^1 W_1, \quad (67)$$

with  $\sigma_1^2 = \mathbb{E}_\mu[(\lambda_0 \delta_{u,u'}^0)^2] + 2 \sum_{k \geq 1} \mathbb{E}_\mu[\lambda_0 \delta_{u,u'}^0 \lambda_k \delta_{u,u'}^k]$  and  $W_t$  a standard brownian motion. Similarly, by using the same arguments, we also have

$$\sqrt{t} \left( \frac{t^u}{t} - \mathbb{P}_{\bar{\pi}}[U_0 = u] \right) \xrightarrow{\mathcal{L}} \sigma^2 W_1, \quad (68)$$

with  $\sigma_2^2 = \mathbb{E}_\mu[(\delta_u^0)^2] + 2 \sum_{k \geq 1} \mathbb{E}_\mu[\delta_u^0 \delta_u^k]$ . Using (67) and (68), we have with asymptotic probability 95% that

$$\begin{aligned} \mathbb{P}_{\bar{\pi}}[U_0 = u] Q(u, u') &\in \left[ \frac{N_t^{u,u'}}{t} + \frac{1.96\sigma_1}{\sqrt{t}}, \frac{N_t^{u,u'}}{t} - \frac{1.96\sigma_1}{\sqrt{t}} \right], \\ \mathbb{P}_{\bar{\pi}}[U_0 = u]^{-1} &\in \left[ \frac{t}{t^u} + \frac{1.96\sigma_2}{\sqrt{t}} \times \frac{t}{t^u}, \frac{t}{t^u} - \frac{1.96\sigma_2}{\sqrt{t}} \times \frac{t}{t^u} \right]. \end{aligned} \quad (69)$$

Equation (69) ensures that we have with probability 90%

$$Q(u, u') \in \left[ \left( \frac{N_t^{u,u'}}{t} + \frac{1.96\sigma_1}{\sqrt{t}} \right) \left( \frac{t}{t^u} + \frac{1.96\sigma_2}{\sqrt{t}} \times \frac{t}{t^u} \right), \left( \frac{N_t^{u,u'}}{t} - \frac{1.96\sigma_1}{\sqrt{t}} \right) \left( \frac{t}{t^u} - \frac{1.96\sigma_2}{\sqrt{t}} \times \frac{t}{t^u} \right) \right].$$

□

#### IV.H Proof of Remark 6.1

*Proof.* We assume that the insertion (resp. consumption) intensity  $\lambda^+$  (resp.  $\lambda^-$ ) is constant and focus on the best bid limit  $Q^1$ . The stationary distribution  $\pi^{old}$  of  $Q^1$  verifies

$$\pi^{old}(q) = \pi^{old}(0) (\rho^{old})^q, \quad \pi^{old}(0) = \left( 1 + \sum_{q=1}^{\infty} (\rho^{old})^q \right)^{-1}, \quad \rho^{old} = \frac{\lambda^+}{\lambda^-}, \quad (70)$$

with  $q \geq 1$  the size of  $Q^1$ . We add to the market a new agent whose insertion (resp. consumption) intensity  $\lambda^{+,a}$  (resp.  $\lambda^{-,a}$ ) is also constant. The stationary distribution  $\pi^{new}$  of  $Q^1$  in the new market satisfies

$$\pi^{new}(q) = \pi^{new}(0) (\rho^{new})^q, \quad \pi^{new}(0) = \left( 1 + \sum_{q=1}^{\infty} (\rho^{new})^q \right)^{-1}, \quad \rho^{new} = \frac{\lambda^+ + \lambda^{+,a}}{\lambda^- + \lambda^{-,a}}, \quad (71)$$

with  $q \geq 1$  the size of  $Q^1$ . Using Equations (70) and (71), we can write

$$\rho^{new} = \rho^{old} (1 + R(\lambda, \lambda^a)), \quad \pi^{new}(0) = \left( 1 + \sum_{q=1}^{\infty} (\rho^{old})^q (1 + R(\lambda, \lambda^a))^q \right)^{-1}, \quad (72)$$

with  $\lambda = (\lambda^+, \lambda^-)$ ,  $\lambda^a = (\lambda^{+,a}, \lambda^{-,a})$  and  $R(\lambda, \lambda^a) = (1 + \frac{\lambda^{+,a}}{\lambda^+}) / (1 + \frac{\lambda^{-,a}}{\lambda^-}) - 1$ . We want the new introduced agent to reduce the volatility of the old market which at the first order reads

$$\mathbb{E}_{\pi^{new}}[\eta_0^2] \leq \mathbb{E}_{\pi^{old}}[\eta_0^2]. \quad (73)$$

Using Equation (72), we can reformulate Inequality (73) in the following way:

$$\sum_q \frac{(\rho^{old})^q (1 + R(\lambda, \lambda^a))^q}{\left(1 + \sum_{j=1}^{\infty} (\rho^{old})^j (1 + R(\lambda, \lambda^a))^j\right)} \eta_0^2(q) \leq \sum_q \frac{(\rho^{old})^q}{\left(1 + \sum_{j=1}^{\infty} (\rho^{old})^j\right)} \eta_0^2(q), \quad (74)$$

for any function  $\eta_0$ . To satisfy Inequality (74) we need  $R(\lambda, \lambda^a) \geq 0$  which leads to

$$\frac{\lambda^{+,a}}{\lambda^{-,a}} \geq \frac{\lambda^+}{\lambda^-},$$

This condition is a well-known result which ensures that the new agent needs to have an insertion/consumption ratio greater than the one of the market.  $\square$

## IV.I Supplementary numerical results

The three next figures show the liquidity consumption and provision intensities at the first limit relative to the whole market according to the queue size, the corresponding stationary measure and the long term volatility, respectively for EssilorLuxottica, Michelin and Orange.

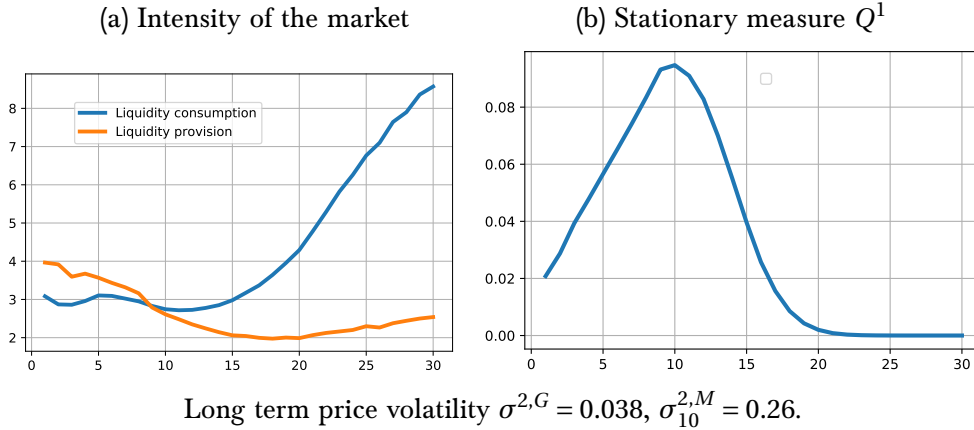


Figure IV.4 - (a) Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and (b) the corresponding stationary distribution of  $Q^1$  with respect to the queue size (in AES), proper to ExilorLuxottica.

IV. From asymptotic properties of general point processes to the ranking of financial agents

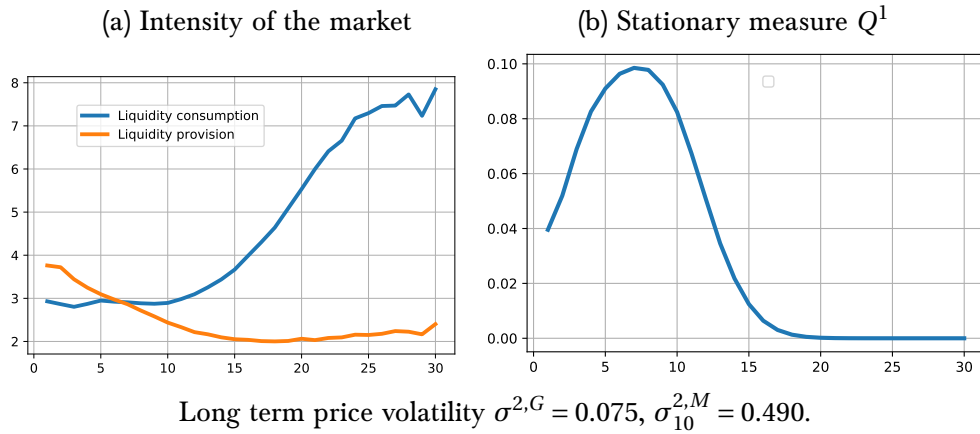


Figure IV.5 – (a) Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and (b) the corresponding stationary distribution of  $Q^1$  with respect to the queue size (in AES), proper to Michelin.

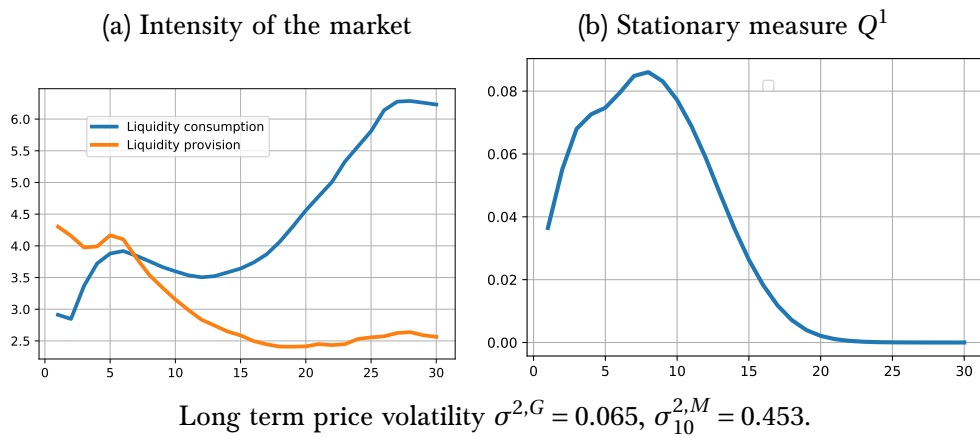


Figure IV.6 – (a) Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and (b) the corresponding stationary distribution of  $Q^1$  with respect to the queue size (in AES), proper to Orange.

For each of the market makers, we compute the liquidity consumption and provision intensities, and the corresponding stationary measure that we would obtain in a situation where the studied market maker withdraws from the market and the other market participants do not change their behaviour. We show respectively the results relative to EssilorLuxottica, Michelin and Orange.

Intensities and  $\sigma_{10}^{2,M}$  when one market maker leaves the market: stock EssilorLuxottica

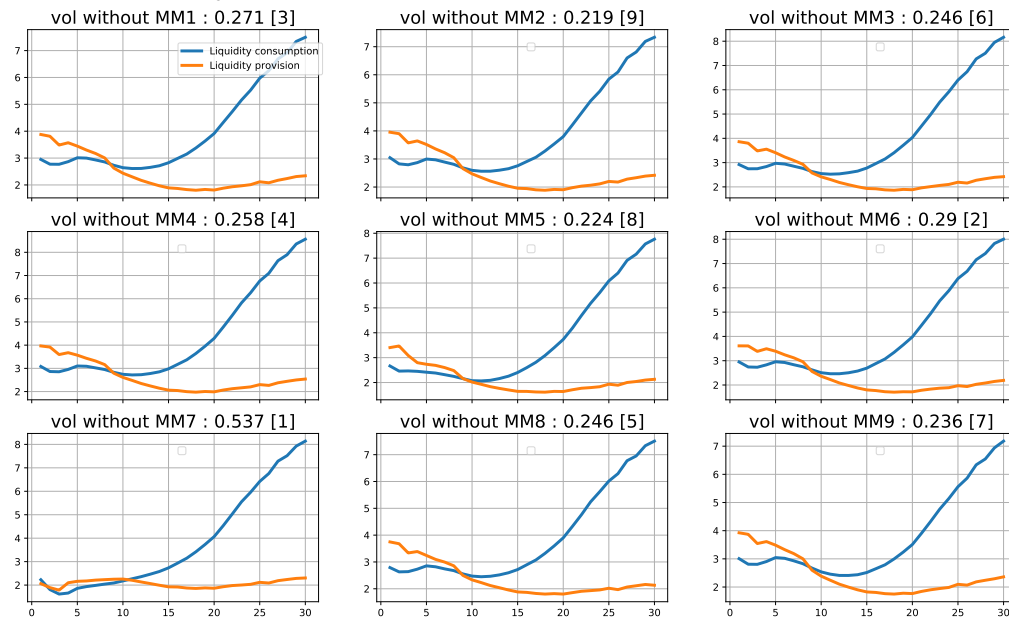


Figure IV.7 - Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and  $\sigma_{10}^{2,M}$  when one market maker is ejected from the market for the stock EssilorLuxottica.

IV. From asymptotic properties of general point processes to the ranking of financial agents

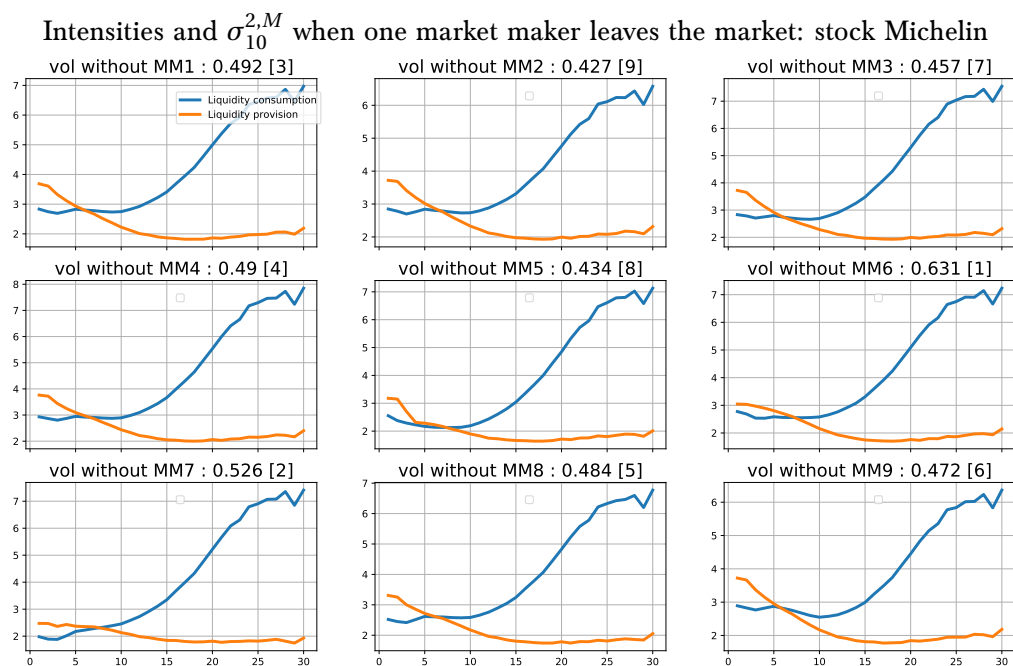


Figure IV.8 - Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and  $\sigma_{10}^{2,M}$  when one market maker is ejected from the market for the stock Michelin.

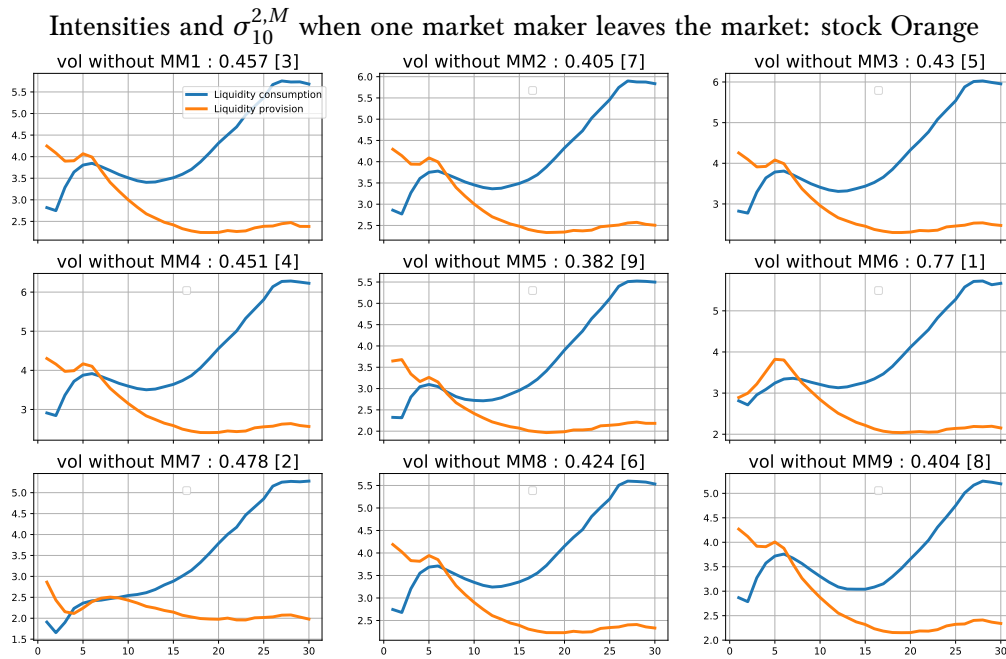


Figure IV.9 – Liquidity insertion and consumption intensities (in orders per second) with respect to the queue size (in AES) and  $\sigma_{10}^{2,M}$  when one market maker is ejected from the market for the stock Orange.





---

# Improving reinforcement learning algorithms: towards optimal learning rate policies

---

## Abstract

This paper investigates to what extent we can improve reinforcement learning algorithms. Our study is split in three parts. First, our analysis shows that the classical asymptotic convergence rate  $O(1/\sqrt{N})$  is pessimistic and can be replaced by  $O((\log(N)/N)^\beta)$  with  $\frac{1}{2} \leq \beta \leq 1$  and  $N$  the number of iterations. Second, we propose a dynamic optimal policy for the choice of the learning rate  $(\gamma_k)_{k \geq 0}$  used in stochastic algorithms. We decompose our policy into two interacting levels: the inner and the outer level. In the inner level, we present the PASS algorithm (for “PAst Sign Search”) which, based on a predefined sequence  $(\gamma_k^o)_{k \geq 0}$ , constructs a new sequence  $(\gamma_k^i)_{k \geq 0}$  whose error decreases faster. In the outer level, we propose an optimal methodology for the selection of the predefined sequence  $(\gamma_k^o)_{k \geq 0}$ . Third, we show empirically that our selection methodology of the learning rate outperforms significantly standard algorithms used in reinforcement learning (RL) in the three following applications: the estimation of a drift, the optimal placement of limit orders and the optimal execution of large number of shares.

## 1 Introduction

We consider a discrete state space  $\mathcal{Z} = \mathbb{N}$  or  $\mathcal{Z} = \{1, \dots, d\}$  with  $d \in \mathbb{N}^*$ . We are interested in finding  $q^* \in \mathcal{Q} = \mathbb{R}^{\mathcal{Z}}$  solution of

$$M(q, z) = \mathbb{E}[m(q, X(z), z)] = 0, \quad \forall z \in \mathcal{Z}, \quad (1)$$

with  $X(z) \in \mathcal{X}$  a random variable with an unknown distribution and  $m$  a function from  $\mathcal{Q} \times \mathcal{X} \times \mathcal{Z}$  to  $\mathcal{Q}$ . Although the distribution of  $X(z)$  is unspecified, we assume that we can observe some variables  $(Z_n)_{n \geq 1}$  valued in  $\mathcal{Z}$  and  $(X_n(Z_n))_{n \geq 1}$  drawn from the distribution of  $X(Z_n)$ . Reinforcement learning (RL) addresses this problem through the following iterative procedure:

$$q_{n+1}(Z_n) = q_n(Z_n) - \gamma_n(Z_n)m(q_n, X_{n+1}(Z_n), Z_n), \quad (2)$$

where  $q_0$  is a given initial condition and each  $\gamma_n$  is a component-wise non-negative vector valued in  $\mathbb{R}^{\mathcal{Z}}$ . The connection between RL, problem (1) and Algorithm (2) is detailed in Section 2. It is possible to recover the classical SARSA,  $Q$ -learning and double  $Q$ -learning algorithms used in RL by taking a specific expression for  $m$  and  $X_{n+1}$ . Note that Algorithm (2) is different from the standard Robbins-Monro (RM) algorithm used in stochastic approximation (SA)

$$q_{n+1} = q_n - \gamma_n \bar{m}(q_n, X_{n+1}), \quad (3)$$

with  $\bar{m}(q, x)(z) = m(q, x(z), z)$  and  $\gamma_n \geq 0$ , mainly because, as it is frequent in RL, we do not observe the entire variable  $(X_{n+1}(z))_{z \in \mathcal{Z}}$  but only its value according to the coordinate  $Z_n$ . Indeed, the way  $(Z_n)_{n \geq 1}$  visits the set  $\mathcal{Z}$  plays a key role in the convergence of Algorithm (2) which is not the case of Algorithm (3). RM algorithm was first introduced by Robbins and Monro in [138]. After that, it was studied by many authors who prove the convergence of  $q_n$  towards  $q^*$ , see [24, 28, 35, 104]. The asymptotic convergence rate has also been investigated in many papers, see [24, 103, 140]. They show that this speed is in general proportional to  $1/\sqrt{N}$  with  $N$  the number of iterations.

In this work, we give a special focus to RL problems. Nowadays RL cover a very wide collection of recipes to solve control problems in an exploration-exploitation context. This literature started in the seventies, see [151, 152], and became famous mainly with the seminal paper of Sutton, see [147]. It largely relied on the recent advances in the control theory developed in the late 1950s, see [22]. The key tool borrowed from this theory is the dynamic programming principle satisfied by the value function. This principle enables us to solve control problems numerically when the environment is known and the dimension is not too large. To tackle the curse of dimensionality, recent papers, see [143], use deep neural networks (DNN). For example, in [92], authors use DNN to derive optimal hedging strategies for finance derivatives and in [117] they use a similar method to solve a high dimensional optimal trading problem. To overcome the fact that environment is unspecified, it is common to use RM algorithm which estimates on-line quantities of interest. The combination of control theory and SA gave birth to numerous papers on RL.

Our contributions are as follows. We first conduct an error analysis to show that the classical asymptotic rate  $O(1/\sqrt{N})$  is pessimistic and can be enhanced in many situations. It is indeed possible to get a  $O((\log(N)/N)^\beta)$  asymptotic speed with  $1/2 \leq \beta \leq 1$  and  $N$  the number of iterations. Then, we present our main result. It consists in proposing a dynamic policy for the choice of the step size  $(\gamma_k)_{k \geq 0}$  used in (3). Our policy is decomposed into two interacting levels: the inner and the outer level. In the inner level, we propose the PASS algorithm, for ‘‘PAst Sign Search’’. This algorithm builds a new sequence  $(\gamma_k^i)_{k \geq 0}$ , using a predefined sequence  $(\gamma_k^o)_{k \geq 0}$  and the sign variations of  $m(q_n, X_{n+1}(Z_n), Z_n)$ . The error of  $(\gamma_k^i)_{k \geq 0}$  decreases faster than the one of  $(\gamma_k^o)_{k \geq 0}$ . In the outer level, we propose an optimal methodology for the construction of a piecewise constant predefined sequence  $(\gamma_k^o)_{k \geq 0}$ . These two levels are interacting in the sense that PASS influences the construction of  $(\gamma_k^o)_{k \geq 0}$ . Finally, we show that our selection methodology provides better convergence results than standard RL algorithms

in three numerical examples: the drift estimation, the optimal placement of limit orders and the optimal execution of a large number of shares.

The structure of this paper is as follows: Section 2 describes the relation between RL and SA. Section 3 reformulates the problem (1) in terms of a minimisation issue and defines with accuracy the different sources of error. This enables us to exploit the most recent convergence results for each source of error to show that the slow convergence speed  $O(1/\sqrt{N})$  can be replaced by  $O((\log(N)/N)^\beta)$  with  $1/2 \leq \beta \leq 1$  and  $N$  the number of iterations. Section 4 contains our main contribution. We start by defining the Algorithm PASS, comparing it with two other schemes and proving its convergence. Then, we describe our outer level policy and discuss its speed of convergence. Finally, we explain our selection methodology of the learning rate  $(\gamma_k^Q)_{k \geq 0}$  which combines the PASS Algorithm with the outer level to enhance the convergence of (3). The last Section 5 provides numerical examples taken from the optimal trading literature: optimal placement of a limit order and the optimization of the trading speed of a liquidation algorithm. Proofs and additional results are relegated to an appendix.

## 2 Reinforcement learning

We detail in this section the relation between SA and RL since we are interested in solving RL problems. RL aims at estimating the  $Q$ -function which quantifies the value for the player to choose the action  $a$  when the system is at  $z$ . Let  $t$  be the current time,  $U_t \in \mathcal{U}$  be a process defined on a filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  which represents the state of the system and  $A_t \in \mathcal{A}$  the agent action at time  $t$ . We assume that the process  $(U_t, A_t)$  is Markovian. The agent aims at maximizing

$$\mathbb{E}[\int_0^T \rho^s f(s, U_s, A_s) ds + \rho^T g(U_T)], \quad (4)$$

with  $g$  the terminal constraint,  $f$  the instantaneous reward,  $\rho$  a discount factor and  $T$  the final time. Let us fix a time step  $\Delta > 0$  and allow the agent to take actions only at times<sup>1</sup>  $k\Delta$  with  $k \in \mathbb{N}$ . The  $Q$ -function satisfies

$$Q(t, u, a) = \sup_A \mathbb{E}[\int_t^T \rho^{(s-t)} f(s, U_s, A_s) ds + \rho^{(T-t)} g(U_T) | U_t = u, A_t = a], \quad \forall (t, u, a) \in \mathbb{R}_+ \times \mathcal{U} \times \mathcal{A},$$

with  $A = \{A_t, t < T\}$  a possible control process for the agent. We view the agent control  $A$  as a feedback process (i.e. adapted to the filtration  $\mathcal{F}_t$ ). The  $Q$ -function satisfies the classical dynamic programming principle (DPP)

$$Q(t, u, a) = \mathbb{E}[R_{t+\Delta} + \rho^\Delta \sup_{a' \in \mathcal{A}} Q(t+\Delta, U_{t+\Delta}, a') | U_t = u, A_t = a], \quad (5)$$

<sup>1</sup>We recall the following classical result: when  $\Delta$  goes to zero, the value function and the optimal control of this problem converges towards the one where decisions are taken at any time.

with  $R_{t+\Delta} = \int_t^{t+\Delta} \rho^{(s-t)} f(s, U_s, A_s) ds$ . Equation (5) reads that the optimal expected gain when the agent starts at  $u$  and chooses action  $a$  at time  $t$  is the sum of the next expected reward  $R_{t+\Delta}$  plus the value of acting optimally starting from the new position  $U_{t+\Delta}$  at time  $t + \Delta$ . By reformulating (5), we obtain that  $Q$  solves equation

$$\mathbb{E}[m(q, X_{t+\Delta}^z, z)] = 0, \quad \forall z = (t, u, a) \in \mathcal{Z} = [0, T] \times \mathcal{U} \times \mathcal{A}, \quad (6)$$

where  $X_{t+\Delta}^z = (U_{t+\Delta}^z, R_{t+\Delta}^z) \in \mathcal{X} = \mathcal{U} \times \mathbb{R}$ ,  $U_s^z$  and  $R_s^z$  are respectively the conditional random variables  $U_s$  and  $R_s$  given the initial condition  $(U_t^z, A_t^z) = (u, a)$  with  $z = (t, u, a) \in \mathcal{Z}$  and  $m$  is defined as follows:

$$m(q, x, z^1) = H(q, x, z^1) - q(z^1), \quad H(q, x, z^1) = r + \rho^\Delta \sup_{a' \in \mathcal{A}} q(t^1 + \Delta, u, a'),$$

for any  $x = (u, r) \in \mathcal{X}$  and  $z = (t^1, u^1, a^1) \in \mathcal{Z}$ . Thus, one can use stochastic approximation tools to solve (6).

**Actions of the agent.** Note that Equation (6) shows that one can study  $Q$  only on the time grid<sup>2</sup>  $D_T = \{n\Delta, n \leq T/\Delta\}$ . Thus, we define  $A_k$  and  $U_k$  such that  $A_k = A_{k\Delta}$  and  $U_k = U_{k\Delta}$  for any  $k \in \mathbb{N}$ . The key variable to study is not the agent decision  $A_k$  but  $Z_k = (k, U_k, A_k)$ . Thus, the rest of the paper formulates the results in terms of  $Z_k$  only.

Moreover, it is important in practice to visit the space  $D_T \times \mathcal{U} \times \mathcal{A}$  sufficiently enough. Thus, to learn  $Q$ , it is common to not choose the maximising action<sup>3</sup>, but to set the conditional distribution of the random variable  $A_k$  such that

$$\mathbb{P}[A_k = a | \mathcal{F}_k] = \frac{e^{\bar{\epsilon}_k(Z_k^a)}}{\sum_{a'} e^{\bar{\epsilon}_k(Z_k^{a'})}}, \quad \forall a \in \mathcal{A}, \quad (7)$$

with  $Z_k^a = (k, U_k, a)$ ,  $\bar{\epsilon}_k(Z_k^a) = \beta(Z_k^a) \sum_{a'} \epsilon_k(U_{k+1}^{Z_k^a}, a')$ ,  $\beta: \mathcal{Z} \rightarrow \mathbb{R}_+$  a discount function and

$$\epsilon_k(z) = \begin{cases} |m(q_{r_k(z)}, X_{r_k(z)+1}(z), z)|, & \text{when the state } z \text{ is already visited,} \\ b, & \text{otherwise,} \end{cases}$$

where  $b > 0$  encourages the exploration,  $q_k$  satisfies (2) and  $r_k(z)$  is the last observation time of the state  $z$ . Beside, to give more importance to the maximizing action, one may consider the following policy:

$$\mathbb{P}[A_k = a | \mathcal{F}_k] = \frac{e^{\beta_k(U_k) q_k(Z_k^a)}}{\sum_{a'} e^{\beta_k(U_k) q_k(Z_k^{a'})}}, \quad \forall a \in \mathcal{A}, \quad (8)$$

with  $\beta_k: \mathcal{U} \rightarrow \mathbb{R}_+$  a discount function. Any mixture of these two procedures can also be considered.

---

<sup>2</sup>Here, we take  $T = n^* \Delta$  with  $n^* \in \mathbb{N}^*$ . Such approximation is not restrictive.

<sup>3</sup>The maximising action  $a^*$  for a state  $u$  is defined such that  $a^* = \operatorname{argmax}_{a \in \mathcal{A}} q(u, a)$ .

### 3 Improvement of the asymptotic convergence rate

In [25, Part 2, Section 4], [103, Section 10] and [104, Section 7], the authors show a central limit theorem for the procedure (3) which ensures a convergence rate of  $O(1/\sqrt{N})$  where  $N$  is the number of iterations. In this section, we extend such convergence rate to Algorithm (2) and aim at understanding how one can improve it. For this, we decompose our total error into two components: estimation error and optimization error.

#### 3.1 Error decomposition

In this section, the space  $\mathcal{Z} = \{1, \dots, d\}$  is finite with  $d \in \mathbb{N}^*$ . In such case, we view  $q$  and  $M(q)$  as vectors of  $\mathbb{R}^{\mathcal{Z}}$  (i.e.  $M(q)[z] = M(q, z)$  for any  $z \in \mathcal{Z}$ ). Moreover, the process  $(Z_n)_{n \geq 1}$  is Markovian. We consider the following assumptions.

**Assumption 3.1** (Existence of a solution). *There exists a solution  $q^*$  of Equation (1).*

**Assumption 3.2** (Pseudo convexity). *The function  $M$  satisfies*

$$\langle M(q) - M(q'), q - q' \rangle \geq 0, \quad \forall q \in \mathbb{R}^{\mathcal{Z}}, \forall q' \in \mathbb{R}^{\mathcal{Z}}.$$

The above assumption is natural since the gradient of any real valued convex function  $f$  satisfies Assumption 3.2.

**Assumption 3.3** (Pseudo strong convexity). *There exists  $L > 0$  such that*

$$\langle M(q) - M(q'), q - q' \rangle \geq L \|q - q'\|^2, \quad \forall q \in \mathbb{R}^{\mathcal{Z}}, \forall q' \in \mathbb{R}^{\mathcal{Z}}.$$

Assumption 3.3 replaces the strong convexity condition. Under Assumption 3.1, the function  $q^*$  is solution of the minimization problem

$$\min_{q \in \mathcal{Q}} g(q). \tag{9}$$

with  $g(q) = \|M(q)\|^2$ .

**Remark 3.1.** *Note that, in the special case where  $M$  is the gradient of a given function  $f$  (i.e.  $\nabla f = M$ ), the quantity  $q^*$  minimises a convex and differentiable cost  $\tilde{g} = \sum_{z \in \mathcal{Z}} \mathbb{E}[L(q, X(z), z)]$  with*

$$L(q, x, z) = \int_0^1 (q(z) - q^*(z)) \times m(q^* + r(q - q^*), x, z) dr, \quad \forall q \in \mathcal{Q}, \forall z \in \mathcal{Z}, \forall x \in \mathcal{X}.$$

and  $\nabla \tilde{g} = M$ . Thus, we can replace  $g$  by  $\tilde{g}$  and all the results of this section hold. In the rest of this section, we use  $g$  instead of  $\tilde{g}$ .

In our context we do not have a direct access to the distribution of  $X(z)$  because we are in the case of “on-line” RL. Nevertheless, we assume that at time  $n$  we keep memory of a training sample of  $n(z)$  independent variables  $(X_i^z)_{i=1\dots n(z)}$  drawn from the distribution  $X(z)$  where  $n(z)$  is the number of times the Markov chain  $Z_n$  visited  $z$ . We define  $q^n$  as a solution of

$$\min_{q \in \mathcal{Q}} g_n(q), \quad (10)$$

with  $g_n(q) = \|M^n(q)\|^2$  and  $M^n(q) = \mathbb{E}^n[m(q, X(z), z)] = (\sum_{j=1}^{n(z)} m(q, X_j(z), z))/n(z)$  the expected value under the empirical measure  $\mu = (\sum_{j=1}^{n(z)} \delta_{X_j(z)})/n(z)$ . We finally define  $q_k^n$  as an approximate solution of the problem (10) returned by an optimization algorithm after  $k$  iterations. Thus, we can bound the error  $g(q_k^n)$  by

$$0 \leq \mathbb{E} \left[ \left( g(q_k^n) - g(q^*) \right) (z) \right] \leq \underbrace{\mathbb{E} \left[ \left( g(q^n) - g(q^*) \right) (z) \right]}_{\text{estimation error}} + \underbrace{\mathbb{E} \left[ \left| g(q^n) - g(q_k^n) \right| (z) \right]}_{\text{optimization error}},$$

since  $q^*$  minimizes  $g$ .

## 3.2 Convergence rate of the estimation error

### 3.2.1 Slow convergence rate

We have the following result.

**Proposition 3.1.** *We assume that the Markov chain  $Z_n$  is irreducible. There exists  $c_1 > 0$  such that*

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)] \leq c_1 \frac{1}{\sqrt{n}}.$$

The proof of this result is given in Appendix V.D. This result allows us to derive the following bound for the estimation error

$$\begin{aligned} \mathbb{E}[(g(q^n) - g(q^*))(z)] &= \mathbb{E}[(g(q^n) - g_n(q^n))(z)] + \underbrace{\mathbb{E}[(g_n(q^n) - g_n(q^*))(z)]}_{\leq 0} + \mathbb{E}[(g(q^*) - g_n(q^*))(z)] \\ &\leq 2\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)] \leq 2c_1 \frac{1}{\sqrt{n}}. \end{aligned} \quad (11)$$

This bound is known to be pessimistic.

### 3.2.2 Fast convergence rate

We obtain the following fast statistical convergence rate.

**Proposition 3.2.** *Assume that the Markov chain  $Z_n$  is irreducible and*

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)|\bar{n}(z)] \leq c' \left( \frac{\log(\bar{n}(z))}{\bar{n}(z)} \right)^\beta, \quad (12)$$

with  $\frac{1}{2} \leq \beta \leq 1$ ,  $c' > 0$  and  $\bar{n}(z) = n(z) \wedge 1$ . Then, there exists  $c_2 > 0$  such that

$$\mathbb{E}[(g(q^n) - g(q^*))(z)] \leq c_2 \left( \frac{\log(n)}{n} \right)^\beta.$$

The proof of this proposition is given in Appendix V.E. The condition (12) is established when

- The loss function  $g$  satisfies regularity conditions, of which the most important are: Lipschitz continuity and convexity, see [19]. Moreover under the strong convexity assumption, the constant  $\beta$  is equal to 1.
- The data distribution satisfies some noise conditions, see for instance [19, 150] in the pattern recognition problem.
- The function  $m$  has a bounded moment  $\alpha$  (i.e.  $\mathbb{E}[m^\alpha(q)] < \infty$ ) with  $\alpha > 1$ , see [58].

### 3.3 Convergence rate of the optimization error

We turn now to the optimization error. This means that expected value in (9) is replaced by the empirical mean which is known. In such case, one can use many algorithms to find  $q_n$ . We present in the table below the most important properties of some gradient methods. Note that the results of Table V.1 remain valid when strong convexity is replaced by Assumption 3.3 and in general when convexity is replaced by Assumption 3.2.<sup>4</sup>

Algorithm	Cost of one iteration	Iterations to achieve an $\epsilon$ precision		Time to reach an $\epsilon$ precision	
		Convex	Strongly convex	Convex	Strongly convex
GD	$O(d^2)$	$O\left(\frac{1}{\epsilon}\right)$	$O\left(\log\left(\frac{1}{\epsilon}\right)\right)$	$O\left(\frac{d^2}{\epsilon}\right)$	$O\left(d^2 \log\left(\frac{1}{\epsilon}\right)\right)$
SGD	$O(d)$	$O\left(\frac{1}{\epsilon^2}\right)$	$O\left(\frac{1}{\epsilon}\right)$	$O\left(\frac{d}{\epsilon^2}\right)$	$O\left(\frac{d}{\epsilon}\right)$
Proximal	$O(d)$	$O\left(\frac{1}{\epsilon}\right)$	$O\left(\log\left(\frac{1}{\epsilon}\right)\right)$	$O\left(\frac{d}{\epsilon}\right)$	$O\left(d \log\left(\frac{1}{\epsilon}\right)\right)$
Acc. prox.	$O(d)$	$O\left(\frac{1}{\sqrt{\epsilon}}\right)$	$O\left(\log\left(\frac{1}{\epsilon}\right)\right)$	$O\left(\frac{d}{\sqrt{\epsilon}}\right)$	$O\left(d \log\left(\frac{1}{\epsilon}\right)\right)$
SAGA	$O(d)$	$O\left(\frac{1}{\epsilon}\right)$	$O\left(\log\left(\frac{1}{\epsilon}\right)\right)$	$O\left(\frac{d}{\epsilon}\right)$	$O\left(d \log\left(\frac{1}{\epsilon}\right)\right)$
SVRG	$O(d)$		$O\left(\log\left(\frac{1}{\epsilon}\right)\right)$		$O\left(d \log\left(\frac{1}{\epsilon}\right)\right)$

Table V.1 – Asymptotic properties of some gradient methods. Note that  $d$  is the dimension of the state space  $\mathcal{X}$  and  $\epsilon$  is a desired level of accuracy. Here  $\epsilon$  corresponds to  $1/n$ . GD stands for Gradient Descent, SDG for Stochastic Gradient Descent, *Proximal* for Stochastic proximal gradient descent [53, 141], *Acc. prox.* for accelerated proximal stochastic gradient descent [130, 141], SAGA for Stochastic accelerated gradient approximation [61], and SVRG for stochastic variance reduced gradient [97].

<sup>4</sup>Assumption 3.3 is the analogous of the strong convexity condition and convexity can be replaced by Assumption 3.2 in many cases.



### 3.4 Conclusion

Following the formalism of [36], we have decomposed our initial error into

- **Estimation error:** its convergence is  $O(1/\sqrt{N})$  in pessimistic cases. In the other situations, the convergence is faster (i.e.  $O((\log(N)/N)^\beta)$ ) with  $1/2 \leq \beta \leq 1$ .
- **Optimization error:** the convergence is exponential under suitable conditions. In unfavourable cases, the convergence rate is  $O(1/N)$ .

The comparison of these error sources shows that the estimation error is the dominant component. Thus, we can overcome the  $O(1/\sqrt{N})$  asymptotic speed, in some situations, by improving the estimation error.

## 4 Optimal policy for the learning rate $\gamma$ when is countable

In this section, we take  $\mathcal{Z} = \mathbb{N}$  and consider the following type of algorithms:

$$q_{n+1}(Z_n) = q_n(Z_n) - \gamma_n(Z_n)m(q_n, X_{n+1}(Z_n), Z_n), \quad \forall n \in \mathbb{N}.$$

One can recover the classical SARSA, Q-learning and double Q-learning algorithms used in RL by considering a specific expression for  $m$  and  $X_{n+1}$ . In such algorithms the choice of  $\gamma_n$  is a crucial point. One can find in the literature general conditions that guarantee the convergence such that

$$\sum_{k \geq 0} \gamma_k(z) = \infty, \quad a.s., \quad \sum_{k \geq 0} \gamma_k^2(z) < \infty, \quad a.s., \quad \forall z \in \mathcal{Z}. \quad (13)$$

However, since the set of processes  $(\gamma_n)_{n \geq 0}$  satisfying these conditions is large in general and may even be empty when  $(Z_n)_{n \geq 0}$  is not recurrent. Thus, to be more specific, many authors suggest to take  $\gamma_n$  proportional to  $1/n^\alpha$  for stochastic approximation algorithms. The exponent  $\alpha$  may vary from 0 to 1 depending on the algorithm used, see to cite [68, 128]. Nonetheless, such a choice may be sub optimal. For example, Figure V.1.a shows that the blue curve is a way higher than the orange one. Here, the blue curve represents the variation of the logarithm of the  $L^2$ -error when  $\gamma_n = \eta/n$ , whereas the orange curve stands for a constant learning rate (i.e.  $\gamma_n = \gamma$ ). We choose the constant  $\eta$  that ensures the fastest convergence for the blue curve.

In this paper, we propose to use a stochastic learning rate  $(\gamma_k)_{k \geq 0}$ ; our learning policy is decomposed into two interacting levels: the inner and the outer level. In the inner level, we use the PASS algorithm, for ‘‘PAsT Sign Search’’. This algorithm builds a new sequence  $(\gamma_k^i)_{k \geq 0}$ , based on a predefined sequence  $(\gamma_k^o)_{k \geq 0}$  and the sign variations of  $m(q_n, X_{n+1}(Z_n), Z_n)$ , whose error decreases faster than the predefined one. In the outer level, we propose an optimal methodology for the selection of a piecewise constant predefined sequence  $(\gamma_k^o)_{k \geq 0}$ . These two levels are interacting in the sense that the PASS algorithm influences the construction of the sequence  $(\gamma_k^o)_{k \geq 0}$ .

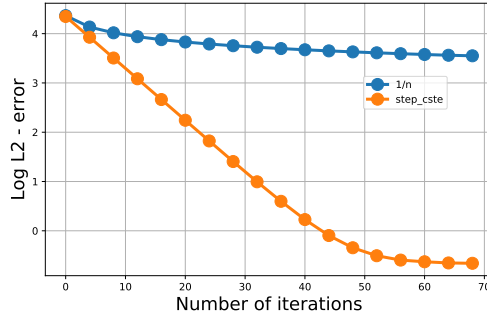


Figure V.1 –  $L^2$ -error for the estimation of the drift when  $\gamma_k$  is constant in orange and when  $\gamma_k \propto \frac{1}{k}$  in blue.

## 4.1 The inner level

### 4.1.1 The algorithms

In this part, we introduce three algorithms. We start with our benchmark which is the standard algorithm used in RL. Then, we present a second algorithm inspired from SAGA [61], which is a method used to accelerate the convergence of the stochastic gradient descent. SAGA reduces the optimization error exponentially fast. Finally, we describe the PASS algorithm that modifies the learning rate  $(\gamma_k)_{k \in \mathbb{N}}$  based on the sign variations of  $m(q_n, X_{n+1}(Z_n), Z_n)$ . The main idea is to increase  $(\gamma_k)_{k \in \mathbb{N}}$  as long as the sign of  $m(q_n, X_{n+1}(Z_n), Z_n)$  remains unchanged. Then, we reinitialize or lower  $\gamma_k$  using a predefined sequence  $(\gamma_k^o)_{k \in \mathbb{N}}$  when the sign of  $m(q_n, X_{n+1}(Z_n), Z_n)$  switches. This algorithm can be seen as an adaptation of the line search strategy, which determines the maximum distance to move along a given search direction, to stochastic iterative methods. Actually, the line search method requires a complete knowledge of the cost function because it demands to evaluate several times the difference  $g(q_k + \gamma M(q^k)) - g(q_k)$  for different values of  $\gamma$  with  $g$  and  $M$  defined in Section 3.1. However, stochastic iterative models have neither access to  $g$  nor  $M$ . They can only compute  $m(q_n, X_{n+1}(Z_n), Z_n)$  when the state  $z = Z_k$  is visited. Moreover, to get a new observation they need to wait<sup>5</sup> for the next visit of the state  $z = Z_k$ . Nevertheless, they have instantaneous access to the previously observed values. Thus, the main idea here is to use these past observations although it adds a small memory cost. Some theoretical properties of these algorithms are investigated in Section 4.1.3.

**Algorithm 2** (RL). *We start with an arbitrary  $q^0 \in \mathcal{Q}$  and define by induction  $q^k$*

$$q^{k+1}(Z_k) = q^k(Z_k) - \gamma_k(Z_k)m(q^k, X_{k+1}(Z_k), Z_k).$$

<sup>5</sup>This waiting time may be very long depending on the dimension of the state space  $\mathcal{Z}$  and the properties of the process  $(Z_k)_{k \geq 0}$ .

**Algorithm 3** (SAGA). *We start with an arbitrary  $q^0 \in \mathcal{Q}$ ,  $M^0 = 0^6$ ,  $n^0 = 0$  and define by induction  $q^k$ ,  $M^k$  and  $n^k$*

$$\begin{aligned} q^{k+1}(Z_k) &= q^k(Z_k) - \gamma_k(Z_k) \left[ m(q^k, X_{k+1}(Z_k), Z_k) - M^k[Z_k, i] + \frac{(\sum_{j=1}^M n^k[Z_k, j] M^k[Z_k, j])}{(\sum_{j=1}^M n^k[Z_k, j])} \right], \\ M^{k+1}[Z_k, i] &= m(q^k, X_{k+1}(Z_k), Z_k), \\ n^{k+1}[k, i] &= 1, \end{aligned}$$

with  $i$  picked from the distribution  $p = (\sum_{i=1}^M n^k[k, i] \delta_i) / (\sum_{j=1}^M n^k[k, j])$ . The quantity  $n^k[k, j]$  records the  $j$ -th time when the process  $Z$  visits the state  $Z_k$  before the  $k$ -th iteration.

For the next algorithm, we give ourselves a predefined sequence  $(\gamma_k)_{k \geq 0}$ , an increasing function  $h: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and a non-increasing function  $l: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ . The function  $h$  is used to increase the learning rate and hence to accelerate the descent, while the function  $l$  is used to go back to a slower pace. A typical way to do it is to use a predefined learning rate.

**Algorithm 4** (PASS). *We start with an arbitrary  $q^0$  and define by induction  $q^k$  and  $\hat{\gamma}_k$*

- If  $m(q_n, X_{n+1}(Z_n), Z_n) \times m(q_{r_1^n}, X_{r_1^n+1}(Z_{r_1^n}), Z_{r_1^n}) \geq 0$ , then do

$$\begin{aligned} q_{n+1}(Z_n) &= q_n(Z_n) - h(\hat{\gamma}_n(Z_n), \gamma_n(Z_n)) m(q_n, X_{n+1}(Z_n), Z_n), \\ \hat{\gamma}_{n+1}(Z_n) &= h(\hat{\gamma}_n(Z_n), \gamma_n(Z_n)), \end{aligned}$$

with  $r_1^n$  is the index of the last observation when the process  $X$  visits the state  $X_n$ .

- Else, do

$$\begin{aligned} q_{n+1}(X_n) &= q_n(X_n) - l(\hat{\gamma}_n(X_n), \gamma_n(X_n)) m(q_n, X_n, X_{n+1}), \\ \hat{\gamma}_{n+1}(X_n) &= l(\hat{\gamma}_n(X_n), \gamma_n(X_n)). \end{aligned}$$

#### 4.1.2 Assumptions

In this section, we present the assumptions needed to prove our main result about the convergence of Algorithms RL, SAGA and PASS. We consider the following assumptions:

**Assumption 4.1** (Existence of a solution). *There exists a solution  $q^*$  of (1).*

We write  $m^*$  for the vector  $m^* = m(q^*) \in \mathcal{X} \times \mathcal{Z}$  with  $q^*$  solution of (1).

**Assumption 4.2** (Pseudo strong convexity 2). *There exists a constant  $L > 0$  such that*

$$(\mathbb{E}_k[m(q_k, X_{k+1}(Z_k), Z_k) - m^*(X_{k+1}(Z_k), Z_k)])(q_k(Z_k) - q^*(Z_k)) \geq L(q_k(Z_k) - q^*(Z_k))^2, \quad (14)$$

with  $q^*$  of (1) and  $\mathbb{E}_k[X] = \mathbb{E}[X | \mathcal{F}_k]$  for any random variable  $X$ .

---

<sup>6</sup>Here  $M^0$  is the zero function in the sense that  $M^0[z, i] = 0$  for any  $z \in \mathcal{Z}$  and  $i \in \{1, \dots, M\}$ .

Recall that Assumption 4.2 is natural in the deterministic framework. For instance, if we take a strongly convex function  $f$  and call  $m$  its gradient (i.e.  $m = \nabla f$ ). Then,  $m$  satisfies Assumption 4.2. Additionally, the pseudo-gradient property (PG) considered in [28, Section 4.2] is close to Assumption 3.2. However, Assumption 3.2 is slightly less restrictive than PG since it involves only the norm of the component  $(q_k - q^*)(Z_k)$  instead of the norm of the vector  $(q_k - q^*)$ . To get tighter approximations, we will also need the quantity  $L_k$  defined as follows:

$$L_k = \begin{cases} \frac{\mathbb{E}_k[m(q_k, X_{k+1}(Z_k), Z_k) - m^*(X_{k+1}(Z_k), Z_k)]}{q^k(Z_k) - q^*(Z_k)}, & \text{If } q^k(Z_k) - q^*(Z_k) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note that  $L_k \geq 0$  under Assumption 4.2. It is also the biggest constant that satisfies (14) for a fixed  $k$ . In particular, this means that  $L_k \geq L$ .

**Assumption 4.3** (Lipschitz continuity of  $m$ ). *There exists a positive constant  $B > 0$  such that for any random variables  $X$  and  $X'$  valued in  $\mathcal{X}$  we have*

$$\mathbb{E}_k[(m(q_k, X, Z_k) - m^*(X', Z_k))^2] \leq B\{1 + (q_k(Z_k) - q^*(Z_k))^2 + \mathbb{E}_k[(X - X')^2]\}, \quad (15)$$

with  $\mathbb{E}_k[X] = \mathbb{E}[X|\mathcal{F}_k]$  for any random variable  $X$ .

Assumption 4.3 guarantees that  $m$  is Lipschitz. Authors in [28, Section 4.2] use a similar condition. To get better bounds, we introduce  $B_k$

$$B_k = \frac{\mathbb{E}_k[(m(q_k, X, Z_k) - m^*(X', Z_k))^2]}{1 + (q_k(Z_k) - q^*(Z_k))^2 + \mathbb{E}_k[(X - X')^2]}.$$

We have  $B_k \leq B$  since  $B_k$  is the smallest constant satisfying (15) for a fixed  $k$ . We finally need an assumption on the learning  $(\gamma_k)_{k \geq 0}$  that describes indirectly how the process  $Z$  communicates with its different states.

**Assumption 4.4** (Learning rate explosion). *For any  $z \in \mathcal{Z}$ , there exists a non-negative deterministic sequence  $(\gamma_k^d(z))_{k \in \mathbb{N}}$  such that*

$$\gamma_k(z) \geq \gamma_k^d(z) \quad \text{and} \quad \sum_{k \geq 1} \gamma_k^d(z) = \infty, \quad \forall z \in \mathcal{Z}.$$

When the process  $Z$  is Markovian and  $\gamma_k(z)$  is bounded, Assumption 4.4 ensures that  $Z$  is recurrent. The boundedness assumption of  $\gamma_k(z)$  is not restrictive and can always be fulfilled. Indeed, since  $\gamma_k(z)$  is bounded, there exists a positive constant  $A$  such that  $\gamma_k(z) \leq A$ , a.s for all  $k \geq 1$ . Thus, we get  $\sum_{k \geq 1} \mathbb{E}[\gamma_k(z) \mathbf{1}_{Z_k=z}] \leq A \sum_{k \geq 1} \mathbb{P}[Z_k = z]$ . Since the left hand side of the previous inequality diverges under Assumption 4.4, we have  $\sum_{k \geq 1} \mathbb{P}[Z_k = z] = \infty$  which proves that  $Z$  is recurrent.

### 4.1.3 Main results

In this section, we compare the algorithms RL, SAGA and PASS and prove the convergence of PASS. Let  $c$  be a positive constant and  $k \in \mathbb{N}$ . We define the error function for the different algorithms as follows:

$$e^k(z) = \begin{cases} (q_k(z) - q^*(z))^2, & \text{for algorithms RL and PASS,} \\ \sum_{j=1}^M \bar{n}^k[z, j] (M^k[z, j] - m^*(z))^2 + c(q_k(z) - q^*(z))^2, & \text{for algorithm SAGA,} \end{cases}$$

for all  $z \in \mathcal{Z}$  with  $\bar{n}^k[z, j] = (n^k[z, j]) / (\sum_{i=1}^M n^k[z, i])$  for any  $j \in \{1, \dots, M\}$  and the total error  $E^k$  such that  $E^k = \|e^k\|$ . We also use the following notations:

$$p(x) = 2Lx - Bx^2, \quad p_k(x) = 2L_k x - B_k x^2, \quad \bar{\gamma}_k = \operatorname{argsup}_{l \in \mathbb{R}} p_k(l) = \frac{L_k}{B_k}, \quad \forall x \in \mathbb{R}.$$

**Proposition 4.1.** *Let  $z \in \mathcal{Z}$ . Under Assumptions 4.1, 4.2 and 4.3, and when there exists  $r_1 \geq 1$  such that*

$$\gamma_2 \leq l(\gamma_1, \gamma_2), \quad \text{and} \quad h(\gamma_1, \gamma_2) \leq r_1 \gamma_2, \quad \forall (\gamma_1, \gamma_2) \in \mathbb{R}^2,$$

*we have*

$$\mathbf{1}_A \mathbb{E}_k[e^{k+1}(z)] \leq \mathbf{1}_A [\alpha_k e^k(z) + M_k], \quad (16)$$

*with  $A = \{Z_k = z\}$ . The values of the constant  $\alpha_k$  and  $M_k$  vary from an algorithm to another as follows:*

$$\alpha_k(z_1) = \begin{cases} [1 - p(\gamma_k(z))], & \text{for algorithm RL,} \\ \max\left(1 - p(\gamma_k(z)) + \frac{B}{Mc}, 1 - \left(\frac{1}{M} - 6\gamma_k^2(z)c\right)\right), & \text{for algorithm SAGA,} \\ \{1 - p_k(\underline{\gamma}_k^*) + d^1 \gamma_k^2(z) \mathbf{1}_{c_k \geq 1}\}, & \text{for algorithm PASS,} \end{cases} \quad (17)$$

*with  $c_k = \frac{\mathbb{E}_k[\hat{\gamma}_k(z)m(q^k, X_{k+1}(z), z)]}{\bar{\gamma}_k(z)\mathbb{E}_k[m(q^k, X_{k+1}(z), z)]}$ ,  $\underline{\gamma}_k^* = c_k(z)\bar{\gamma}_k \vee \bar{\gamma}_k$ ,  $d^1 = (r_1 - 1)^2 B_k$  and*

$$M_k = \begin{cases} B\gamma_k^2(z)(2 + v_k), & \text{for algorithm RL,} \\ 3B\gamma_k^2(z)(2 + v_k), & \text{for algorithm SAGA,} \\ B(c_k \bar{\gamma}_k)^2(2 + v_k), & \text{for algorithm PASS,} \end{cases} \quad (18)$$

*where  $v_k$  is the variance of  $Z_k$ .*

Equation (16) reveals that the performances of algorithms RL, SAGA and PASS depend on the interaction between two competing terms:

- On the one hand the slope  $\alpha_k$  controls the decrease of the error from one step to the other,
- On the other hand the quantity  $M_k$  gathers two sources of imprecision: the estimation error and the optimization error. Both sources of imprecision have a term in the variance  $v_n$  (because the distribution of  $Z$  is unknown) and a positive constant (coming from the noise generated by the noisy nature of observations).

There is a competition between these two terms: to decrease  $M_k$  we need to send  $\gamma_k$  towards zero while the reduction of  $\alpha_k$  requires a relatively small but still non-zero value of  $\gamma_k$ . Thus,  $\gamma_k$  should satisfy a trade-off in order to ensure the convergence of the algorithms. The RM conditions (13) are a way to address this trade-off. Now, in order to analyse the properties of each algorithm, we compare for a given  $\gamma_k$  their respective values for  $\alpha_k$  and  $M_k$  in Table V.2. For sake of clarity, we choose to present the variable  $(1 - \alpha_k)$  instead of  $\alpha_k$  in this table; note that a large value of  $1 - \alpha_k$  means that  $\alpha_k$  is small and thus induces a fast convergence.

Algorithms	$(1 - \alpha_k)$		$M_k$	
	Value	Comparison with Algo 1	Value	Comparison with Algo 1
RL (Algo 1)	$2\gamma_k(z_1)L - B\gamma_k^2(z_1)$	—	$B\gamma_k^2(2 + \nu_k)$	—
SAGA (Algo 2)	$(2\gamma_k(z_1)L - B\gamma_k^2(z_1) - \frac{B}{\bar{M}c}) \vee (\frac{1}{\bar{M}} - 6\gamma_k^2(z_1)c)$	smaller	$3B\gamma_k^2(2 + \nu_k)$	larger
PASS (Algo 4)	$2\gamma_k^*L_k - B_k(\gamma_k^*)^2 + d^1\gamma_k^2\mathbf{1}_{\alpha_k \geq 1}$	larger	$B(c_k\tilde{\gamma}_k)^2(2 + \nu_k)$	larger

Table V.2 – Comparison of the algorithms RL, SAGA and PASS.

The result below holds only for Algorithms 2 and 4.

**Theorem 4.1.** *Let the Assumptions 4.1, 4.2, 4.3 and 4.4 be in force. Algorithms 2 and 4 verify*

- *When  $\sum_{k \geq 0} \mathbb{E}[\gamma_k^2](z) < \infty$  for all  $z \in \mathbb{Z}$ , we have*

$$\mathbb{E}_0[E^n] \xrightarrow{n \rightarrow \infty} 0. \quad (19)$$

- *When  $\gamma_k(z) = \gamma \mathbf{1}_{\{Z_k=z\}}$  for all  $(k, z) \in \mathbb{N} \times \mathcal{Z}$  with  $\gamma > 0$  a positive constant, for any positive number  $r$  there exist two constants  $\beta \in [0, 1)$  and  $\bar{M} > 0$  which satisfy*

$$\mathbb{E}_0[E^n] \leq c(1+r)(1 - \tilde{\alpha}^n) + \tilde{\alpha}^n \bar{M}E^0, \quad \forall \tilde{\alpha} < \max(\beta, \alpha), \forall n \in \mathbb{N}, \quad (20)$$

$$\text{with } c = \frac{\gamma^2 M}{1 - \alpha}, \quad \alpha = 1 - \gamma \text{ and } M = \sup_n M_n.$$

## 4.2 The upper level

In practice, to apply PASS we need an appropriate predefined sequence  $(\gamma_k)_{k \in \mathbb{N}}$ . It is possible to take  $\gamma_k$  proportional to  $1/k^\alpha$  with  $\alpha \in (0, 1]$  as proposed in [128, 138]. However, in this section, we present a piecewise constant policy for the selection of the learning rate  $(\gamma_k)_{k \in \mathbb{N}}$ . To do so, we construct a sequence of integers  $(k_i)_{i \in \mathbb{N}}$  such that  $\gamma_k$  is constant within each interval  $I_i = [k_i, k_{i+1}[$  with  $i \in \mathbb{N}$ . Additionally, we choose the “best” constant  $\gamma_{k_i}$  such that the error  $\mathbb{E}[E^n]$  decreases the fastest possible by at least a factor  $\alpha$  during the time period  $I_i$  in the sense that  $\mathbb{E}[E^{k_{i+1}}] \leq \alpha \mathbb{E}[E^{k_i}]$ . Here the factor  $\alpha$  is a fixed constant.

First, note that the function  $f(x) = 2Lx - B^2$  with  $L$  defined in Assumption 4.2 and  $B$  introduced in Assumption 4.3 reaches its maximum value at the point  $\tilde{\gamma}_1 = \frac{L}{B} \geq 0$ . Since we need  $f(\gamma_0)$  to be the largest possible for a fast convergence, we take  $\gamma_0$  as close as possible to

$\bar{\gamma}_1$ . Let  $r > 0$  be a precision factor and  $\bar{\gamma}_2 = \frac{2E^0L}{\bar{m}(1+r)+BE^0}$  where  $E^0$  is the initial total error,  $\bar{m} = \sup_k M_k/\gamma_k^2$  is uniformly bounded of  $\bar{\gamma}_2$  and the ratio  $M_k/\gamma_k^2$  is defined in Proposition 4.1. To guarantee a decrease of the error by a factor  $\alpha$  during the first iterations, we also need  $\gamma_0 \leq \bar{\gamma}_2$ . Thus, we define recursively the sequence  $(k_i)_{i \in \mathbb{N}}$  and  $(\gamma_{k_i})_{i \in \mathbb{N}}$  as follows:

$$\begin{cases} k_{i+1} = k_i + \lceil \frac{\log(\frac{\alpha r}{2(1+r)\bar{M}})}{\log(\alpha_{k_i})} \rceil, & \gamma_{k_{i+1}} = \alpha \gamma_{k_i}, & \forall i \geq 0, \\ k_0 = 0 & \gamma_{k_0} = \bar{\gamma}_1 \vee \bar{\gamma}_2, \end{cases} \quad (21)$$

with  $\alpha_{k_i} = 1 - \gamma_{k_i}$  and  $\bar{M}$  defined in Theorem 4.1. The proposition below provides a non-asymptotic approximation of the number of steps  $k_i$  needed to reduce the error by the factor  $\alpha^i$  with  $i \in \mathbb{N}$ .

**Proposition 4.2.** *Under Assumptions 4.1, 4.2 and 4.3, we have*

$$k_i \leq L_1 i + L_2(i - i^*)_+ + L_3(\alpha^{-(i-i^*)_+} - 1).$$

with  $i^* = \sup\{i \geq 0, \alpha_{k_i} \leq \beta\}$ ,  $\beta$  defined in Theorem 4.1,  $L_1 = \frac{-L}{\log(\beta)}$ ,  $L_2 = \frac{L\bar{b}}{2qL(1-\bar{b}\bar{\gamma}^1)}$ ,  $L_3 = \frac{L}{2qL\gamma_0(1-\alpha^{-1})}$ ,  $L = |\log(\frac{\alpha r}{2(1+r)\bar{M}})|$ ,  $\bar{b} = \frac{B}{2L}$  and  $q = \frac{1}{2-\beta}$ .

The proof of the above result is given in Appendix V.H. Proposition 4.2 shows that the convergence speed due to (21) is exponentially fast as long as  $i \leq i^*$  (i.e.  $k_i \leq L_1 i$ ). In such case, the leading term is  $L_1 i$ . This means that we need around  $L_1 i$  steps to reduce the error by a factor  $\alpha^{i/2}$ .<sup>7</sup> However, when  $i > i^*$ , the dominating term becomes  $L_3(\alpha^{-(i-i^*)_+} - 1)$ . This means that after a number  $k$  of iterations proportional to  $k \propto \alpha^{-i}$  we reduce the error by  $1/\sqrt{k} = \alpha^{i/2}$ . Then, we recover the slow convergence rate of  $O(1/\sqrt{k})$ .

Since the constants  $L$  and  $B$  are unknown in practice, a first solution consists in starting with arbitrary values for  $B$  and  $L$  and generating a sequence of learning rates. If the error  $m(q_n, X_{n+1}(Z_n), Z_n)$  increases, one can take a larger value for  $B$  and a smaller one for  $L$  otherwise he keeps  $B$  and  $L$  unchanged. A second solution consists in directly averaging the error  $m(q_n, X_{n+1}(Z_n), Z_n)$  over the last  $p$  visit times with  $p \in \mathbb{N}$  fixed by the controller. If this average error does not decrease, Equation (21) reduces the value of the step size by a factor  $\alpha$ .

### 4.3 Extension

The results of this section still hold when the descent sequence (2) is replaced by

$$q_{n+1}(z) = q_n(z) - \gamma_n(z)m(q_n, X_{n+1}(z), z), \quad \forall z \in \mathcal{Z}, \forall n \in \mathbb{N}. \quad (22)$$

When  $\gamma_n(z) = 0$  if  $z \neq Z_n$ , we recover (2). Thus, Equation (22) is slightly more general. Moreover, Algorithm (22) appears in many contexts: stochastic iterative algorithms, gradient methods, fixed point iterative techniques, etc. The results of this section can be directly transposed to (22). We present in Appendix V.A an adaptation of the Algorithm PASS for (22).

<sup>7</sup>Recall that  $\alpha^i$  is the square of the error.

## 5 Some examples

### 5.1 Methodology

In this section, we compare four algorithms. The two first ones are two different versions of RL. In the first version, the learning rate  $\gamma_k$  is taken such that  $\gamma_k = \frac{\eta}{k}$  with  $\eta > 0$  selected to provide the best convergence results. In the second version, the step size follows the upper level policy, described in Section 4.2. The third algorithm is SAGA where the step size is derived from the upper level policy. Finally, we use the PASS algorithm (for “PAsT Sign Search”) presented and studied theoretically in the previous sections. We consider three numerical examples to compare the convergence speed of these algorithms: drift estimation, optimal placement of limit orders and the optimal liquidation of shares.

### 5.2 Drift estimation

**Formulation of the problem.** We observe a process  $(S_n)_{n \geq 0}$  which satisfies

$$S_{n+1} = S_n + f_{n+1} + W_n, \quad (23)$$

with  $W_n$  a centred noise with finite variance. We want to estimate the quantities  $f_i$  with  $i \in \{1, \dots, n_{max}\}$ . Using (23) and  $\mathbb{E}[W_i] = 0$ , we get

$$\mathbb{E}[S_{i+1} - S_i - f_{i+1}] = 0, \quad \forall i \in \{0, \dots, n_{max} - 1\}.$$

Thus, we can estimate  $f_i$  using stochastic iterative algorithms. The pseudo-code of our implementation of PASS for this problem can be found in the Appendix V.B under the name Implementation 1.

**Numerical results.** Figure VI.3 shows the variation of the  $L^2$ -error when the number of iterations increases. We can see that the algorithm PASS outperforms standard stochastic approximation algorithms. Moreover, other algorithms behave as expected: the standard RL decreases very slowly (but we know it will drive the asymptotic error to zero), the constant learning rate and SAGA provides better results than RL, while PASS seems to have captured the best of the two worlds for this application: very fast acceleration at the beginning and the asymptotic error goes to zero.



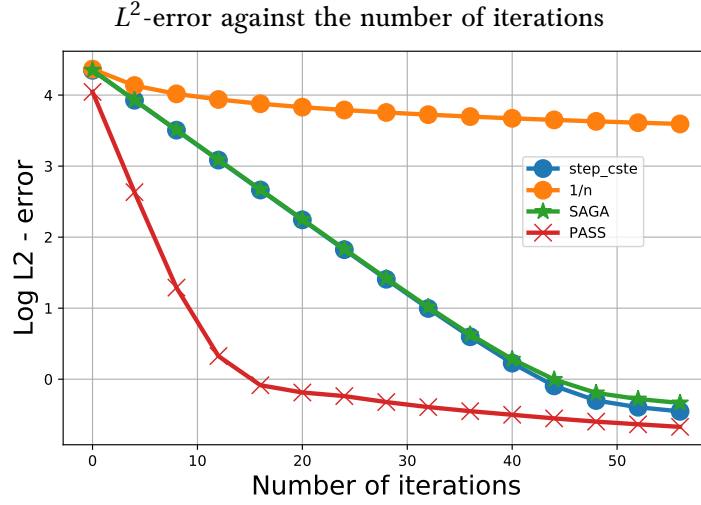


Figure V.2 – The  $L^2$ -error between  $f^k$  and  $f$  for different numerical methods averaged over 1000 simulated paths.

### 5.3 Optimal placement of a limit order

**Formalisation of the problem.** We consider an agent who aims at buying a unit quantity using limit orders and market orders (see Chapter I for detailed explanations). In such case, the agents wonder how to find the right balance between fast execution and avoiding trading costs associated to the bid-ask spread. The agent state at time  $t$  is modelled by  $X_t = (Q^{Before}, Q^{After}, P)$  with  $Q^{Before}$  the number of shares placed before the agent’s order,  $Q^{After}$  the queue size after the agent’s order and  $P_t$  the mid price, see Figure V.3. The agents wants to minimise the quantity

$$\mathbb{E}[F(X_{\tau \vee T^{\text{exec}}}) + \int_0^{\tau \vee T^{\text{exec}}} c ds],$$

where

- $T^{\text{exec}} = \inf\{t \geq 0, P_t = 0\}$  the first time when the limit order gets a transaction.
- $\tau$  the first time when a market order is sent.
- $X = (Q^{Before}, Q^{After}, P)$  the state of the order book.
- $F(u)$  is the price of the transaction (i.e.  $F(u) = p + \psi$  when the agents crosses the spread and  $F(u) = p$  otherwise).

We show in Section 2 that the  $Q$ -function is solution of (6). Thus, we can use Algorithms RL, SAGA and PASS to estimate it. The pseudo-code of our implementation of PASS is available as Implementation 2 in Appendix V.B.

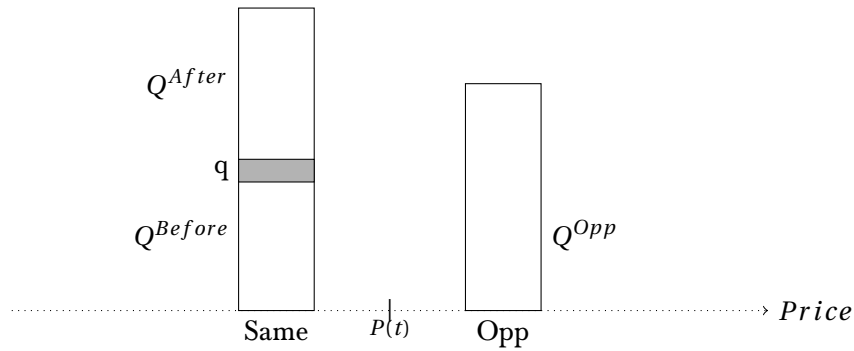


Figure V.3 – The state space of our limit order control problem.

**Numerical results.** Figure V.4 shows three control maps: the x-axis reads the quantity on “same side” (i.e.  $Q^{same} = Q^{Before} + Q^{After}$ ) and the y-axis reads the position of the limit order in the queue, i.e.  $Q^{Before}$ . The color and numbers gives the control associated to a pair  $(Q^{same}, Q^{Before})$ : 1 (blue) means “stay in the book”, while 0 (red) means “cross the spread” to obtain a transaction. The panel (at the left) gives the reference optimal controls obtained with a finite difference scheme, the middle panel the optimal controls obtained for a RL algorithm where the step-size  $(\gamma_k)_{k \geq 0}$  is derived from the upper level policy, and the right panel the optimal control obtained with our optimal policy (i.e. upper level and inner level combined). It shows that after few iterations our optimal policy already found the optimal controls. Figure V.5 compares the log of the  $L^2$  error, averaged over 100 trajectories, between the different algorithms. We see clearly that our methodology improves basic stochastic approximation algorithms. Again, the other algorithms behave as expected: SAGA is better than a constant learning rate that is better than the standard RL (at the beginning, since we know that asymptotically RL will drive the error to zeros whereas a constant learning rate does not).

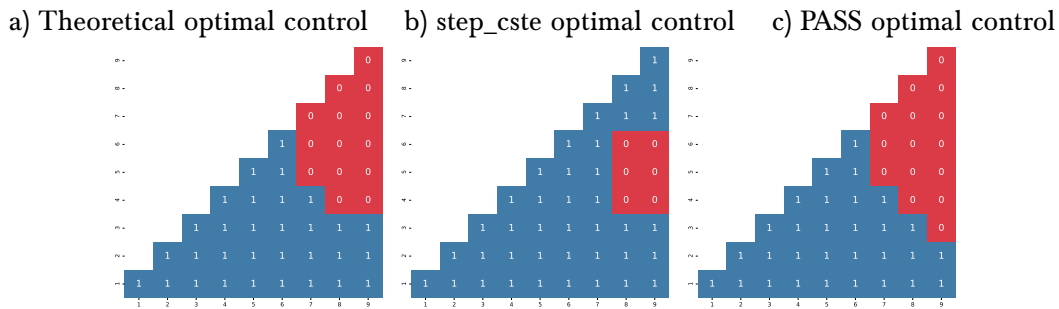


Figure V.4 – Comparison optimal control after 300 iteration for different methods: left is the optimal control, middle is RL with a step size derived from the upper level and right is our optimal policy for the step size (i.e. upper level and inner level combined).

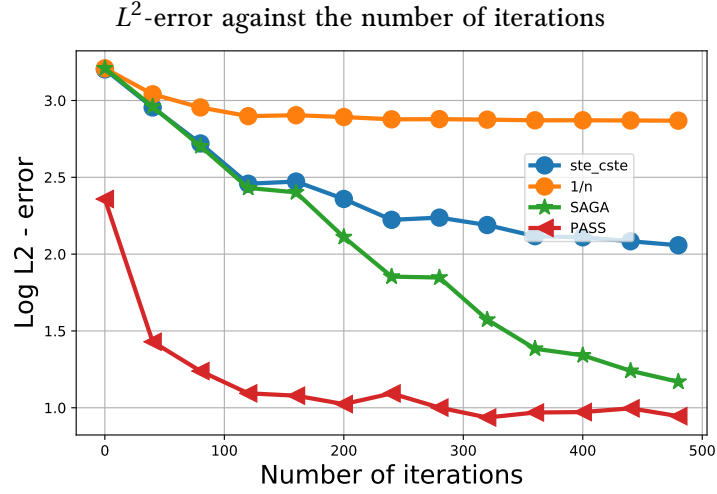


Figure V.5 – The log  $L^2$ -error against the number of iterations averaged over 1000 simulated paths.

#### 5.4 Optimal execution

**Formalisation of the problem.** An investor wants to buy a given quantity  $q_0$  of a tradable instrument (see [52] and [145] for details about this application). The price  $S_t$  of this instrument satisfies the following dynamics:

$$dS_t = \alpha dt + \sigma dB_t,$$

where  $\alpha \in \mathbb{R}$  is the drift and  $\sigma$  is the price volatility. The state of the investor is described by two variables its inventory  $Q_t$  and its wealth  $X_t$  at time  $t$ . The evolution of these two variables reads

$$\begin{cases} dQ_t = v_t dt, & Q_0 = q_0, \\ dW_t = -v_t(S_t + \kappa v_t) dt, & W_0 = 0, \end{cases} \quad (24)$$

with  $v_t$  the trading speed of the agent and  $\kappa > 0$ . The term  $\kappa v_t$  corresponds to the temporary price impact. The investor wants to maximize the following quantity

$$W_T + Q_T(S_T - AQ_T) - \phi \int_t^T Q_s^2 ds,$$

it represents its final wealth  $X_T$  at time  $T$ , plus the value of liquidating its inventory minus a running quadratic cost. The value function  $V$  is defined such that

$$V(t, w, q, s) = \sup_v \mathbb{E}[W_T + Q_T(S_T - AQ_T) - \phi \int_t^T Q_s^2 ds | W_t = w, Q_t = q, S_t = s].$$

We remark that  $v(t, w, q, s) = V(t, w, q, s) - w - qs$  verifies

$$v(t, w, q, s) = \sup_v \mathbb{E} \left[ \underbrace{(W_T - W_t) + (Q_T S_T - Q_t S_t)}_{=M_T^t} - A Q_T^2 - \phi \int_t^T Q_s^2 ds \mid W_t = w, Q_t = q, S_t = s \right]. \quad (25)$$

Using (24), we can see that the variable  $M_T^t$  is independent of the initial values  $W_t$ ,  $S_t$  and  $Q_t$ . This means that  $v$  is a function of only two variables the time  $t$  and the inventory  $q$ . The dynamic programming principle ensures that  $v$  satisfies

$$v(t, q) = \sup_v \mathbb{E} \left[ M_{t+\Delta}^t - \phi \int_t^{t+\Delta} Q_s^2 ds + v(t+\Delta, Q_{t+\Delta}) \mid Q_t = q \right]. \quad (26)$$

We fix a maximum inventory  $\bar{q}$ . Let  $k = (k_T, k_q) \in (\mathbb{N}^*)^2$ ,  $\Delta = T/k_T$ ,  $D_T = \{t_i^{k_T}; i \leq k_T\}$  and  $D_q = \{q_i^{k_q}; i \leq k_q\}$  with  $t_i^{k_T} = i\Delta$  and  $q_i^{k_q} = -\bar{q} + 2i\bar{q}/k_q$ . To estimate  $v$  we use the numerical scheme  $(v_n^k)_{n \geq 1, k \in (\mathbb{N}^*)^2}$  defined such that

$$v_{n+1}^k(Z_n) = v_n^k(Z_n) + \gamma_n(Z_n) \left[ \sup_{v \in A(Z_n)} \{M_{n+1}^v - \phi \Delta Q_n^2 + v_n^k(Z_{n+1}^v) - v_n^k(Z_n)\} \right],$$

with  $Z_n = (n\Delta, Q_{n\Delta})$  and  $A(Z_n) \in D_q$  is the set of admissible actions.<sup>8</sup> When the final time  $T$  is reached (i.e.  $n = k_T$ ), we pick a new initial inventory from the set  $D_q$  and start again its liquidation. At a first sight, it is not clear that  $v_n^k$  approximates  $v$ . However, we show in Appendix V.I that  $v_n^k$  converges pointwise to  $v$  on  $D_T \times D_q$  when  $n \rightarrow \infty$  and  $k \rightarrow \infty$ . see Appendix V.B for a detailed implementation of the algorithm with the corresponding pseudo-code (as Implementation 3).

**Numerical results.** Figure V.6 shows the value function  $v$  for different values of the elapsed time  $t$  and the remaining inventory  $Q_t$ . The panel (at the left) gives the reference value function. It is computed by following the same approach of [46]. The middle panel the value function obtained after 120 000 iterations for RL algorithm where the step-size  $(\gamma_k)_{k \geq 0}$  is derived from the upper level of our optimal policy, and the right panel the value function obtained with our optimal policy (i.e. upper level and inner level combined). It shows that our optimal strategy leads to better performance results. We also plot, in Figure V.7, a simulated path for the variations of the log  $L^2$  error for different algorithms. Here again, we notice that our methodology improves the basic RL algorithm and that the ordering of other approaches is similar to the one of the “drift estimation” approximation (i.e. SAGA and the constant learning rate are very similar).

<sup>8</sup>We do not allow controls that leads to states where the inventory exceeds  $\bar{q}$ .

a) Theoretical value function    b) step\_ cste value function    c) PASS value function

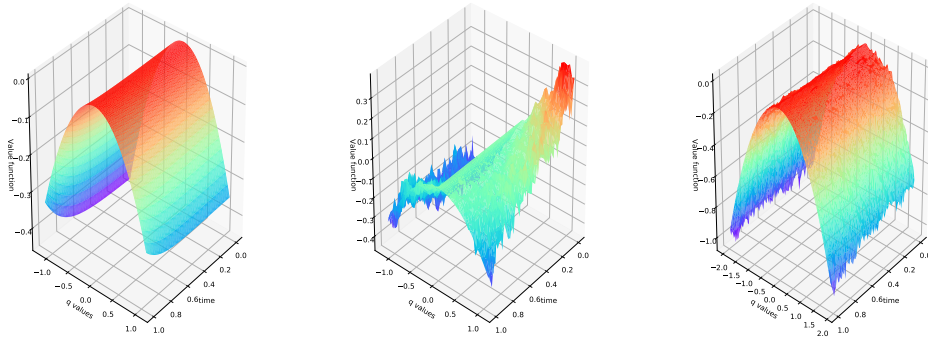


Figure V.6 – Comparison value function between methods.

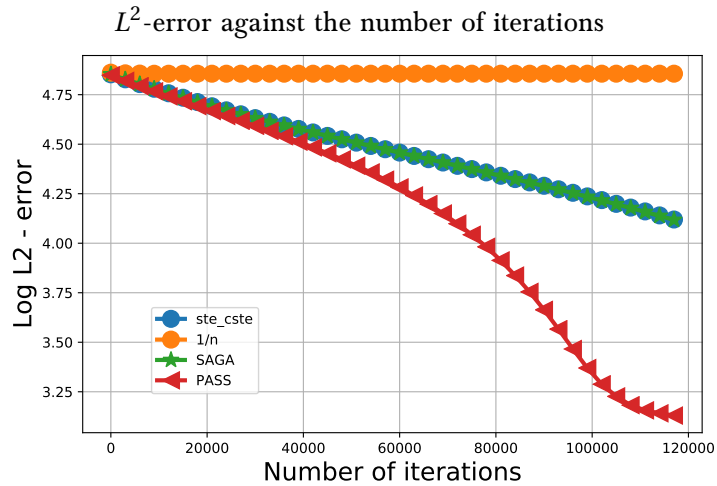


Figure V.7 – The log  $L^2$ -error against the number of iterations averaged over 1000 simulated paths.

### V.A Extension of the PASS Algorithm

First, we can adapt Algorithm PASS to (22) by considering component-wise version which consists in applying Algorithm 4 introduced in Section 4.1.1 to all the visited coordinates of  $q_n$ . Next, we also propose a more direct approach with the algorithm below.

**Algorithm 5** (PASS for (22)). *We start with an arbitrary  $q_0$  and define by induction  $q_k$  and  $\hat{\gamma}_k$*

- If  $\langle m(q_n, X_{n+1}), m(q_{n-1}, X_n) \rangle \geq 0$ , then do

$$\begin{aligned} q_{n+1} &= q_n - h(\hat{\gamma}_n, \gamma_n)m(q_n, X_{n+1}), \\ \hat{\gamma}_{n+1} &= h(\hat{\gamma}_n, \gamma_n). \end{aligned}$$

- Else, do

$$\begin{aligned} q_{n+1} &= q_n - l(\hat{\gamma}_n, \gamma_n)m(q_n, X_{n+1}), \\ \hat{\gamma}_{n+1} &= l(\hat{\gamma}_n, \gamma_n). \end{aligned}$$

## V.B Implementations

We give here the pseudo code used for each one of the three numerical examples considered in Section 5.

**Drift estimation.** We consider the following expression for the functions  $h$  and  $l$ :

$$h(\gamma, \gamma_{base}) = \min(\gamma + \gamma_{base}, 3\gamma_{base}), \quad l(\gamma, \gamma_{base}) = \max(\gamma - \gamma_{base}, \gamma_{base}), \quad \forall (\gamma, \gamma_{base}) \in \mathbb{R}_+.$$

We use Implementation 1 for the numerical experiments.

**Optimal placement of limit orders.** We consider the following expression for the functions  $h$  and  $l$ :

$$h(\gamma, \gamma_{base}) = \max(\min(\gamma + 2/3\gamma_{base}, 3\gamma_{base}), \gamma_{base}), \quad l(\gamma, \gamma_{base}) = \max(\gamma - 2/3\gamma_{base}, \gamma_{base}), \quad \forall (\gamma, \gamma_{base}) \in \mathbb{R}_+. \quad (27)$$

We use Algorithm 2 for the numerical experiments. Note that we do not need to send a market order to know our expected future gain.

**Optimal execution of a large number of shares.** To solve this problem we use the same functions  $h$  and  $l$  considered in the previous problem, see (27). Then, we apply Algorithm 3. In this problem, it is crucial to select actions according to the policy (7) in order to encourage exploration. The coefficient  $\bar{\beta}$  used by the agent to select its actions is taken constant equal to  $\bar{\beta} = 5$ . We consider the same policy for all the tested algorithms.

## V.C Additional results: optimal execution

Authors in [46] show that  $v$  satisfies

$$v(t, q) = h_0(t) + h_1(t)q - h_2(t)/2q^2, \quad \forall t \in [0, T], \forall q \in \mathbb{R}, \quad (28)$$

with  $h_0$ ,  $h_1$  and  $h_2$  are three real valued functions. Here, we want to estimate  $h_0$ ,  $h_1$  and  $h_2$  using the ersatz (28) and Equation (26). We use Equations (26), (49) and (28) and we neglect the terms of second order (i.e.  $o(\Delta_t)$ ) to get

$$v(t, q) = \sup_v \mathbb{E}[v\Delta_t(h_1 - qh_2) - v^2\Delta_t\kappa + q\Delta S - \phi q^2 + v(t + \Delta_t, q)]. \quad (29)$$

---

**Algorithm 1** PAsT Sign Search (PASS) for (RL) drift estimation problem

---

```

1: Algorithm parameters: step size  $(\gamma^o)_{n \geq 0} \in (0, 1]$ , number of episodes  $n$ 
   initial guess  $q_0$ , past error value  $E_{past}$ 
   Initialise  $\hat{\gamma}_0 = \gamma_0^o$ 
2: for episode in  $1 : n$  do
3:   for  $t \in \{0, \dots, n_{\max} - 1\}$  do
4:     Observe  $\Delta X_{next} = S_{t+1} - S_t$ 
5:     if the first visit time to  $t$  then
6:        $q_0(t) \leftarrow q_0(t) - \hat{\gamma}_0(t)m(q_0, \Delta X_{next}, t)$ 
7:     else if  $m(q_0, t, \Delta X_{next}) \times E_{past}(t) \geq 0$  then
8:        $\hat{\gamma}_0(t) \leftarrow h(\hat{\gamma}_0(t), \gamma^o(t))$ 
9:        $q_0(t) \leftarrow q_0(t) - \hat{\gamma}_0(t)m(q_0, \Delta X_{next}, t)$ 
10:    else if  $m(q_0, t, \Delta X_{next}) \times E_{past}(t) < 0$  then
11:       $\hat{\gamma}_0(t) \leftarrow l(\hat{\gamma}_0(t), \gamma^o(t))$ 
12:       $q_0(t) \leftarrow q_0(t) - \hat{\gamma}_0(t)m(q_0, \Delta X_{next}, t)$ 
13:    end if
14:     $E_{past}(t) \leftarrow m(q_0, t, \Delta X_{next})$ 
15:  end for
16:  Save the norm  $\|E\|$  of the vector  $E_{past}(t)$ .
17:  if the average value of  $\|E\|$  over the last  $w = 5$  episodes is not reduced by  $p = 1\%$ 
   then
18:     $\gamma^o(t) \leftarrow \max(\gamma^o(t)/2, 0.01)$  (this is done each  $w$  episodes)
19:  end if
20: end for

```

---

By maximising the above expression over  $v$ , we get the optimal control  $v^* = \frac{h_1 - qh_2}{2\kappa}$ . Finally, Using (28), we can split Equation (29) in three parts

$$\begin{cases} h_2(t) = h_2(t + \Delta_t) - 2\phi + \frac{(h_2(t + \Delta_t))^2}{2\kappa}, & h_2(T) = 2A, \\ h_1(t) = h_1(t + \Delta_t) + \mathbb{E}[\Delta S] - \frac{h_2(t + \Delta_t)h_1(t + \Delta_t)}{2\kappa}, & h_1(T) = 0, \\ h_0(t) = h_0(t + \Delta_t) + \frac{(h_1(t + \Delta_t))^2}{4\kappa}, & h_0(T) = 0. \end{cases} \quad (30)$$

Thus, we can use Algorithm RL to estimate  $h_0$ ,  $h_1$  and  $h_2$ . The results are given in Figures V.8 and V.9.

---

**Algorithm 2** PAsT Sign Search (PASS) for (RL) optimal placement problem

---

```

1: Algorithm parameters: step size  $(\gamma^o)_{n \geq 0} \in (0, 1]$ , number of episodes  $n$ 
   initial guess  $q_0$ , past error value  $E_{past}$ 
   Initialise  $\hat{\gamma}_0 = \gamma_0^o$ 
2: for episode in  $n$  do
3:   Select initial state  $X_0$ 
4:   for each step within episode do
5:     Take the action stay in the order book
6:     Observe the new order book state  $X_{next}$ 
7:     for  $a \in \{0, 1\}$  do
8:       if the first visit time to  $X_{next}$  then
9:          $q_0(X_0, a) \leftarrow q_0(X_0, a) - \hat{\gamma}_0(X_0, a)m^a(q_0, X_0, X_{next})$ 
10:      else if  $m^a(q_0, X_0, X_{next}) \times E_{past}(X_0, a) \geq 0$  then
11:         $\hat{\gamma}_0(X_0, a) \leftarrow h(\hat{\gamma}_0(X_0, a), \gamma^o(X_0, a))$ 
12:         $q_0(X_0, a) \leftarrow q_0(X_0, a) - \hat{\gamma}_0(X_0, a)m^a(q_0, X_0, X_{next})$ 
13:      else if  $m^a(q_0, X_0, X_{next}) \times E_{past}(X_0, a) < 0$  then
14:         $\hat{\gamma}_0(X_0, a) \leftarrow l(\hat{\gamma}_0(X_0, a), \gamma^o(X_0, a))$ 
15:         $q_0(X_0, a) \leftarrow q_0(X_0, a) - \hat{\gamma}_0(X_0, a)m^a(q_0, X_0, X_{next})$ 
16:      end if
17:       $E_{past}(X_0, a) \leftarrow m^a(q_0, X_0, X_{next})$ 
18:    end for
19:     $X_0 \leftarrow X_{next}$ 
20:  end for
21:  Save the norm  $\|E\|$  of the vector  $E_{past}(t)$ .
22:  if the average value of  $\|E\|$  over the last  $w = 40$  episodes is not reduced by  $p = 5\%$ 
then
23:     $\gamma^o(t) \leftarrow \max(\gamma^o(t)/2, 0.01)$  (this is done each  $w$  episodes)
24:  end if
25: end for

```

---

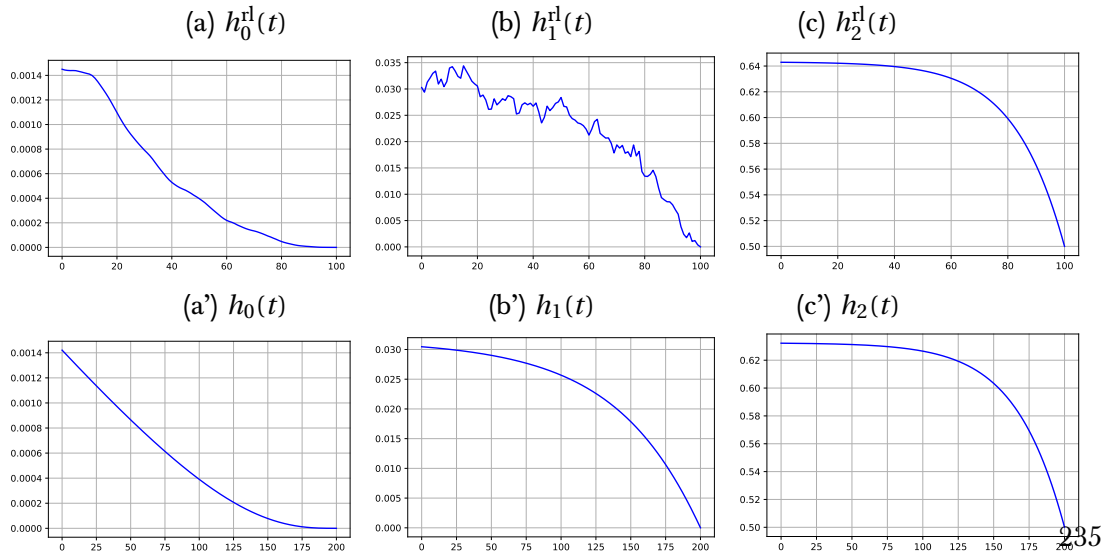


Figure V.8 – (a) The quantity  $h_0^{\text{rl}}(t)$ , (b)  $h_1^{\text{rl}}(t)$  and (c)  $h_2^{\text{rl}}(t)$  computed using the (RL) procedure. (a') The quantity  $h_0(t)$ , (b')  $h_1(t)$  and (c')  $h_2(t)$  computed using an explicit numerical scheme. The (RL) approximations are close to the one of the numerical scheme.



**Algorithm 3** PAsT Sign Search (PASS) for (RL) optimal execution problem

- 1: Algorithm parameters: step size  $(\gamma^o)_{n \geq 0} \in (0, 1]$ , number of episodes  $n$   
 initial guess  $q_0$ , past error value  $E_{past}$   
 Initialise  $\hat{\gamma}_0 = \gamma_0^o$
- 2: **for** episode in  $n$  **do**
- 3:     Select the initial inventory  $Q_0$
- 4:     **for**  $t \in \{0, \dots, n_T - 1\}$  **do**
- 5:         Observe the new price state  $S_{next}$  and set  $X_0 = (t, Q_0)$
- 6:         Observe the new price state  $S_{next}$
- 7:         **if** the first visit time to  $X_0$  **then**
- 8:              $q_0(X_0) \leftarrow q_0(X_0) - \hat{\gamma}_0(X_0)m(q_0, S_{next}, X_0)$
- 9:         **else if**  $m(q_0, S_{next}, X_0) \times E_{past}(X_0) \geq 0$  **then**
- 10:              $\hat{\gamma}_0(X_0) \leftarrow h(\hat{\gamma}_0(X_0), \gamma^o(X_0))$
- 11:              $q_0(X_0) \leftarrow q_0(X_0) - \hat{\gamma}_0(X_0)m(q_0, S_{next}, X_0)$
- 12:         **else if**  $m(q_0, S_{next}, X_0) \times E_{past}(X_0) < 0$  **then**
- 13:              $\hat{\gamma}_0(X_0) \leftarrow l(\hat{\gamma}_0(X_0), \gamma^o(X_0))$
- 14:              $q_0(X_0) \leftarrow q_0(X_0) - \hat{\gamma}_0(X_0)m(q_0, S_{next}, X_0)$
- 15:         **end if**
- 16:          $E_{past}(X_0) \leftarrow m(q_0, S_{next}, X_0)$
- 17:         Select an action  $A$  and observe  $Q_{next}$
- 18:          $Q_0 \leftarrow Q_{next}$
- 19:     **end for**
- 20:     Save the norm  $\|E\|$  of the vector  $E_{past}(t)$ .
- 21:     **if** the average value of  $\|E\|$  over the last  $w = 300$  episodes is not reduced by  $p = 1\%$   
    **then**
- 22:          $\gamma^o(t) \leftarrow \max(\gamma^o(t) - 0.01, 0.01)$  (this is done each  $w$  episodes)
- 23:     **end if**
- 24: **end for**

(a)  $v^{rl}(t, q)$

(b)  $v^{theo}(t, q)$

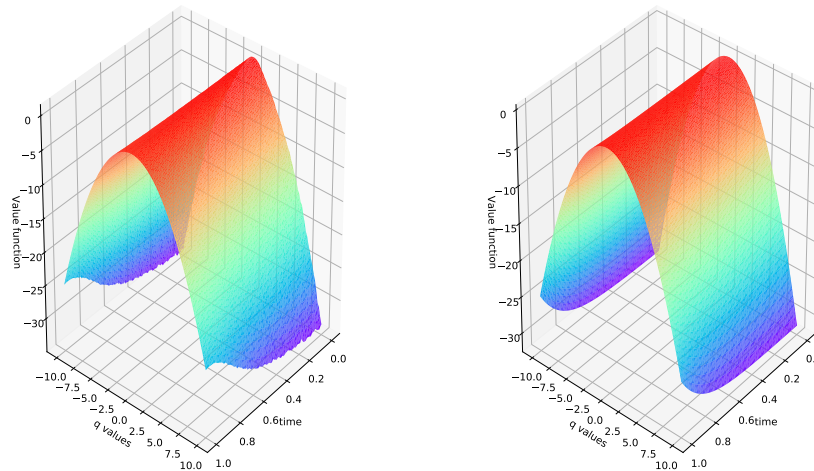


Figure V.9 – (a) The quantity  $v^{rl}(0, q)$  is estimated using the (RL) approach and (b) the quantity  $v^{rl}(t, q)$  is computed by following [46]. Both approximations  $v^{rl}(0, q)$  and  $v^{theo}(0, q)$  are close.

## V.D Proof of Proposition 3.1

*Proof of Proposition 3.1.* Let  $z \in \mathcal{Z}$ . Standard uniform convergence results ensure that

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)|n(z)] \leq c \frac{1}{\sqrt{n(z) \wedge 1}}, \quad a.s.$$

with  $c > 0$  a positive constant. Since the Markov chain  $(Z_n)_{n \geq 1}$  is irreducible and the set  $\mathcal{Z}$  is finite, the sequence  $(Z_n)_{n \geq 1}$  is positive recurrent and we have

$$\frac{n(z)}{n} = \frac{\sum_{k=1}^n \mathbf{1}_{Z_k=z}}{n} \xrightarrow{n \rightarrow \infty} \mathbb{P}_\mu[Z_n = z] > 0 = p(z),$$

with  $\mu$  the unique invariant distribution of  $(Z_n)_{n \geq 1}$ . Thus, we have

$$u_n(z) = \mathbb{E} \left[ \sqrt{\frac{n}{n(z) \wedge 1}} \right] \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{p(z)}} > 0.$$

This shows that  $u_n(z)$  is bounded by  $u_\infty(z)$  and ensures that

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)] \leq c_1(z) \frac{1}{\sqrt{n}},$$

with  $c_1(z) = c u_\infty(z)$ . The constant  $c_1$  can be taken independent of  $z$  since  $\mathcal{Z}$  is finite.  $\square$

## V.E Proof of Proposition 3.2

*Proof of Proposition 3.2.* Let  $z \in \mathcal{Z}$ . We follow the same approach used in the proof of Proposition 3.1 to get

$$v_n(z) = \mathbb{E} \left[ \left( \frac{\log(\bar{n}(z))/\log(n)}{\bar{n}(z)/n} \right)^\beta \right] \leq \mathbb{E} \left[ \left( \frac{1}{\bar{n}(z)/n} \right)^\beta \right] \xrightarrow{n \rightarrow \infty} \frac{1}{p(z)^\beta} > 0.$$

This shows that  $v_n(z)$  is bounded by  $v_\infty(z)$  and ensures that

$$\mathbb{E}[\sup_q |g(q) - g_n(q)|(z)] \leq c_2(z) \frac{1}{\sqrt{n}},$$

with  $c_2(z) = c' v_\infty(z)$ . The constant  $c_2$  can be taken independent of  $z$  since  $\mathcal{Z}$  is finite. Using the same manipulations of (11), we complete the proof.  $\square$

## V.F Proof of Proposition 4.1

*Proof of Proposition 4.1.* Let  $k \geq 0$ ,  $A$  the set  $A = \{Z_k = z\}$  and  $m(q_k) \in \mathbb{R}^{\mathcal{Z}}$  such that  $m(q_k)(z') = m(q_k, X_{k+1}(z'), z')$  for any  $z' \in \mathcal{Z}$ . We split the proof in three steps. In each one of these steps, we prove (16) for a given algorithm.

**Step (i):** In this step, we prove (16) for Algorithm RL. We have

$$\begin{aligned} \mathbf{1}_A \mathbb{E}_k[(q_{k+1} - q^*)^2(z)] &= \mathbf{1}_A \mathbb{E}_k[(q_k - \gamma_k m(q_k) - q^*)^2(z)] \\ &= \mathbf{1}_A \left\{ (q_k - q^*)^2(z) \underbrace{- 2\gamma_k \mathbb{E}_k[m(q_k)](q_k - q^*)}_{=(i)} + \gamma_k^2 \underbrace{\mathbb{E}_k[m(q_k)^2(z)]}_{=(ii)} \right\}. \end{aligned}$$

Using Assumption 4.2 and  $\mathbb{E}_k[m^*] = 0$ , we get  $(i) \leq -2L\gamma_k |q_k - q^*|^2(z)$ . Since  $\mathbb{E}_k[m^*] = 0$ , Assumption 4.3 gives

$$\begin{aligned} (ii) &= \mathbb{E}_k[(m(q_k) - \mathbb{E}_k[m(q_k)])^2(z)] + (\mathbb{E}_k[m(q_k) - m^*])^2(z) \\ &\leq B(1 + v_k) + B(1 + (q_k - q^*)^2(z)). \end{aligned}$$

Thus, we deduce that

$$\mathbf{1}_A \mathbb{E}_k[(q_{k+1} - q^*)^2(z)] \leq \mathbf{1}_A \left\{ \underbrace{(1 - 2\gamma_k L + B\gamma_k^2)}_{=-p(\gamma_k)} (q_k - q^*)^2(z) + \gamma_k^2(z) \underbrace{B(2 + v_k)}_{=M_k} \right\},$$

which shows (16) for Algorithm RL.

**Step (ii):** Here we show (16) for Algorithm SAGA. Let  $\bar{M}^k(z) = \sum_{j=1}^M \bar{n}^k[z, j] M^k[z, j]$  and  $\bar{n}^k[z, j] = n^k[z, j] / \sum_{j'=1}^M n^k[z, j']$ . Using  $\mathbb{E}_k[m^*] = 0$  and  $\mathbb{E}_k[M^k[z, i]] = \bar{M}^k(z)$ , we have

$$\begin{aligned} \mathbf{1}_A \mathbb{E}_k[(q_{k+1} - q^*)^2(z)] &= \mathbf{1}_A \left\{ (q_k - q^*)^2(z) + 2(\mathbb{E}_k[q_{k+1}] - q_k)(q_k - q^*)(z) + \mathbb{E}_k[(q_{k+1} - q_k)^2(z)] \right\} \\ &= \mathbf{1}_A \left\{ (q_k - q^*)^2(z) - 2(\gamma_k \mathbb{E}_k[m(q_k) - m^*])(q_k - q^*)(z) \right. \\ &\quad \left. + \gamma_k^2 \mathbb{E}_k[(m(q_k)(z) - M^k[z, i] + \bar{M}^k(z))^2] \right\} \\ &\stackrel{\text{Assumption 4.2}}{\leq} \mathbf{1}_A \left\{ (1 - 2L\gamma_k)(q_k - q^*)^2(z) + \gamma_k^2 \underbrace{\mathbb{E}_k[(\bar{n}^k(q^k)(z) - M^k[z, i] + \bar{M}^k(z))^2]}_{=(1)} \right\}. \end{aligned} \tag{31}$$

We first dominate the term (1). Since  $\mathbb{E}_k[m^*](z) = 0$  and

$$\mathbb{E}_k \left[ (M^k[z, i] - m^*)^2(z) \right] = \sum_j \bar{n}^k[z, j] \mathbb{E}_k \left[ (M^k[z, j] - m^*(z))^2 \right],$$

we have

$$\begin{aligned}
 (1) &= \mathbb{E}_k \left| \left( m(q_k) - \mathbb{E}_k[m^*](z) - (M^k[z, i] - m^*(z)) + \left( \sum_j \bar{n}^k[z, j](M^k[z, j] - m^*(z)) \right) \right) \right|^2 \\
 &\leq 3 \left[ \mathbb{E}_k [m(q_k) - m^*]^2(z) + \mathbb{E}_k [(M^k[z, i] - m^*(z))^2] + \mathbb{E}_k \left[ \left( \sum_j \bar{n}^k[z, j](M^k[z, j] - m^*(z)) \right)^2 \right] \right] \\
 &\stackrel{\text{Jensen's inequality}}{\leq} 3 \left[ \mathbb{E}_k [(m(q_k) - m^*)^2(z)] + \mathbb{E}_k [(M^k[z, i] - m^*)^2(z)] + \sum_j \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] \right] \\
 &= 3 \left[ \mathbb{E}_k [(m(q_k) - \mathbb{E}_k[m(q_k)])^2(z)] + (\mathbb{E}_k[m(q_k) - m^*](z))^2 + 2 \sum_j \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] \right] \\
 &\stackrel{\text{Assumption 4.3}}{\leq} 3 \left[ B(2 + v_k + (q_k - q^*)^2(z)) + 2 \sum_j \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] \right] \\
 &= 3B(2 + v_k) + 3B(q_k - q^*)^2(z) + 6 \sum_j \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2]. \tag{32}
 \end{aligned}$$

By combining (31) and (32), we get

$$\begin{aligned}
 \mathbf{1}_A \mathbb{E}_k [|q_{k+1} - q^*|^2(z)] &\leq \mathbf{1}_A \left\{ (1 - 2\gamma_k(z)L + 3B\gamma_k^2(z))(q_k - q^*)^2(z) \right. \\
 &\quad \left. + 6 \sum_j \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] + 3B\gamma_k^2(v_k + 2)(z) \right\}. \tag{33}
 \end{aligned}$$

Moreover, we have

$$\begin{aligned}
 \mathbf{1}_A \mathbb{E}_k \left[ \sum_{j=1}^M \bar{n}^k[z, j] (M^{k+1}[z, j] - m^*(z))^2 \right] &= \mathbf{1}_A \left\{ \frac{1}{M} \mathbb{E}_k [(m(q_k) - m^*)^2(z)] \right. \\
 &\quad \left. + \left(1 - \frac{1}{M}\right) \sum_{j=1}^M \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] \right\} \\
 &\leq \mathbf{1}_A \left\{ \frac{B}{M} (1 + (q_k - q^*)^2(z)) \right. \\
 &\quad \left. + \left(1 - \frac{1}{M}\right) \sum_{j=1}^M \bar{n}^k[z, j] \mathbb{E}_k [(M^k[z, j] - m^*(z))^2] \right\}. \tag{34}
 \end{aligned}$$

Thus using (33) and (34), we conclude

$$\begin{aligned} \mathbf{1}_A \mathbb{E}_k[e^{k+1}(z)] &\leq \mathbf{1}_A \left\{ \underbrace{\left(1 - 2\gamma_k L + 3B\gamma_k^2 + \frac{B}{Mc}\right)}_{=\alpha_1} c(q_k - q^*)^2(z) \right. \\ &\quad \left. + \underbrace{\left(1 - \frac{1}{M} + 6\gamma_k^2(z)c\right)}_{\alpha_2} \sum_j \bar{n}^k[z, j] \mathbb{E}_k \left[ \left(M^k[z, j] - m^*(z)\right)^2 \right] \right\} + 3B\gamma_k^2(v_k + 2)(z) \\ &\leq \mathbf{1}_A \alpha e^k(z) + 3B\gamma_k^2(v_k + 2)(z), \end{aligned}$$

with  $\alpha = \max(\alpha_1, \alpha_2) \in [0, 1)$ .

**Step (iii):** In this final step, we show (16) for Algorithm PASS. We have

$$\begin{aligned} \mathbf{1}_A \mathbb{E}_k[(q_{k+1} - q^*)^2(z)] &= \mathbf{1}_A \mathbb{E}_k[(q_k - \hat{\gamma}_k m(q_k) - q^*)^2(z)] \\ &= \mathbf{1}_A \left\{ (q_k - q^*)^2(z) \underbrace{- 2\mathbb{E}_k[\hat{\gamma}_k m(q_k)](q_k - q^*)}_{=(i)}(z) + \underbrace{\mathbb{E}_k[\hat{\gamma}_k^2(m(q_k))^2(z)]}_{=(ii)} \right\}. \end{aligned}$$

For the term (i), using Assumption 4.2 and  $\mathbb{E}_k[m^*] = 0$ , we have (i) =  $-2c_k \bar{\gamma}_k (q_k - q^*)^2(z)$  with  $c_k = \frac{\mathbb{E}_k[\hat{\gamma}_k m(q_k)(z)]}{\bar{\gamma}_k \mathbb{E}_k[m(q_k)(z)]}$ . Using Assumption 4.3 and  $\mathbb{E}_k[m^*] = 0$ , we get

$$\begin{aligned} (ii) &= c_k^2 \bar{\gamma}_k^2 \mathbb{E}_k[m(q_k)^2(z)] = c_k^2 \bar{\gamma}_k^2 \left( \mathbb{E}_k[(m(q_k) - \mathbb{E}_k[m(q_k)])^2(z)] + (\mathbb{E}_k[m(q_k) - m^*])^2(z) \right) \\ &\leq c_k^2 \bar{\gamma}_k^2 \left( B_k(1 + v_k) + B_k(1 + (q_k - q^*)^2(z)) \right). \end{aligned}$$

Thus, we deduce that

$$\mathbf{1}_A \mathbb{E}_k[(q_{k+1} - q^*)^2(z)] \leq \mathbf{1}_A \left\{ \underbrace{(1 - 2c_k \bar{\gamma}_k L_k + B_k(c_k \bar{\gamma}_k)^2)}_{=-p_k(c_k \bar{\gamma}_k)} (q_k - q^*)^2(z) + c_k^2 \bar{\gamma}_k^2(z) \underbrace{B_k(2 + v_k)}_{=M_k} \right\}. \quad (35)$$

We write  $\underline{\gamma}_k$  for the quantity  $\underline{\gamma}_k = c_k \bar{\gamma}_k \vee \bar{\gamma}_k$ . Since  $\hat{\gamma}_k \in [\gamma_k, r_1 \gamma_k]$ , we have  $c_k \bar{\gamma}_k \in [\gamma_k, r_1 \gamma_k]$ . When  $c_k \bar{\gamma}_k \in ]\gamma_k, \bar{\gamma}_k]$ , we have  $p_k(\underline{\gamma}_k) = p_k(c_k \bar{\gamma}_k) > p_k(\gamma_k) \geq 0$ . When  $c_k \bar{\gamma}_k \in ]\bar{\gamma}_k, r_1 \gamma_k]$  (i.e.  $c_k \geq 1$ ), we use that

$$|p_k(c_k \bar{\gamma}_k) - p_k(\underline{\gamma}_k)| = |p_k(c_k \bar{\gamma}_k) - p_k(\bar{\gamma}_k)| \leq B_k(c_k \bar{\gamma}_k - \bar{\gamma}_k)^2 \leq B_k(r_1 \gamma_k - \gamma_k)^2 = \gamma_k^2 d^1,$$

with  $d^1 = (r_1 - 1)^2 B_k$ . Thus, using (35), we conclude

$$\mathbf{1}_A \mathbb{E}_k[(q^{k+1} - q^*)^2(z)] \leq \mathbf{1}_A \left\{ (1 - p_k(\underline{\gamma}_k) + \gamma_k^2 d^1 \mathbf{1}_{c_k \geq 1}) (q^k - q^*)^2(z) + c_k^2 \bar{\gamma}_k^2(z) \underbrace{B_k(2 + v_k)}_{=M_k} \right\}.$$

This completes the proof.  $\square$

## V.G Proof of Theorem 4.1

### V.G.1 Preparation for the proof of Theorem 4.1

We introduce the following notations. Let  $j \in \mathbb{N}$  and  $(\mu_n)_{n \geq 1}$  a real sequence, we write  $(\mu_n^j)_{n \geq 1}$  for the delayed sequence  $\mu_n^j = \mu_{j+n}$  for any  $n \geq 1$ . Additionally, we define recursively the sequence  $(a_n^\mu)_{n \geq 1}$  as follows:

$$a_1^\mu = 1, \quad \text{and} \quad a_{n+1}^\mu = \mu_n \sum_{l=1}^n a_{n+1-l} a_l^\mu, \quad \forall n \geq 1. \quad (36)$$

**Lemma V.G.1.** *By convention, an empty sum is equal to zero. Let  $(v_n)_{n \geq 1}$  be the sequence defined as follows:*

$$v_n = \epsilon_n + \mu_n \left( \sum_{j=1}^{n-1} a_{n-j} v_j \right), \quad \forall n \geq 1,$$

where  $(\epsilon_n)_{n \geq 0}$  and  $(\mu_n)_{n \geq 0}$  are two real sequences and  $\sum_{j \geq 0} a_j = 1$ . Then, we have

$$v_n = \sum_{j=1}^n a_{n+1-j}^\mu \epsilon_j, \quad \forall n \geq 1. \quad (37)$$

*Proof of Lemma V.G.1.* Let us prove the result by induction on  $n \geq 1$ . By definition, Equation (37) is satisfied for  $n = 1$ . By applying the induction hypothesis (37) to all  $j \leq n$ , we get

$$\begin{aligned} v_{n+1} &= 1 \times \epsilon_{n+1} + \mu_{n+1} \left( \sum_{j=1}^n a_{n+1-j} v_j \right) = a_1^{\mu^{n+1}} \epsilon_{n+1} + \mu_{n+1} \left( \sum_{j=1}^n a_{n+1-j} \sum_{l=1}^j a_{j+1-l}^\mu \epsilon_l \right) \\ &= a_1^{\mu^{n+1}} \epsilon_{n+1} + \left[ \sum_{l=1}^n \mu_{n+1} \left( \sum_{j=l}^n a_{n+1-j} a_{j+1-l}^\mu \right) \epsilon_l \right] \\ &= a_1^{\mu^{n+1}} \epsilon_{n+1} + \left[ \sum_{l=1}^n \mu_{n+1} \left( \sum_{j=1}^{n-l+1} a_{n-l+2-j} a_j^\mu \right) \epsilon_l \right] \\ &= a_1^{\mu^{n+1}} \epsilon_{n+1} + \sum_{l=1}^n a_{n+2-l}^{\mu^l} \epsilon_l = \sum_{j=1}^{n+1} a_{n+2-j}^{\mu^j} \epsilon_j. \end{aligned}$$

□

**Lemma V.G.2.** *Let  $n \in \mathbb{N}$ ,  $(\mu_n)_{n \geq 0}$  be a positive sequence,  $(a_n^\mu)_{n \geq 0}$  be the sequence defined in (36) and  $r_n = 1 - \mu_n$ .*

- When  $\sum_{n \geq 0} r_n = \infty$ , we have

$$a_n^\mu \xrightarrow{n \rightarrow \infty} 0. \quad (38)$$

- When  $\sum_{n \geq 0} r_n = \infty$  and there exists  $\beta \in [0, 1)$  such that  $\sum_{n \geq 0} \beta^n a_n < \infty$ , we have

$$\beta^n a_n \xrightarrow{n \rightarrow \infty} 0. \quad (39)$$

*Proof of Lemma V.G.2.* We need to introduce the following notations. Let  $n \in \mathbb{N}^*$  and  $(u_n)_{n \geq 1}$  be a real sequence. We define respectively the translation operator  $T_1$  and the aggregation operator  $T_2$  as follows  $T_1(u)_n = u_{n-1} \mathbf{1}_{n \geq 2}$  and  $T_2(u)_n = (\sum_{l \geq 1} a_l u_l) \mathbf{1}_{n=1}$ . We denote by  $(\mathbf{w}^k)_{k \geq 1}$  the sequence

$$\mathbf{w}_n^1 = \epsilon_1 \mathbf{1}_{n=1}, \quad \mathbf{w}^{k+1} = (T_1 + \mu_{k+1} T_2)(\mathbf{w}^k), \quad \forall k \geq 1, \forall n \geq 1. \quad (40)$$

Our construction ensures that  $v_{k+1} = \mathbf{w}_1^{k+1} = \mu_{k+1} T_2(\mathbf{w}^k)_1$ . In the sequel, we first prove (38) then show (39).

**Step (i):** Here, we demonstrate (38). We first handle the case where  $a_1 = 1$ . In such case, we have

$$v_n = \mu_n v_{n-1} = \left( \prod_{i=2}^n \mu_i \right) v_1 = \left( \prod_{i=2}^n (1 - r_i) \right) v_1 \leq e^{-\sum_{i=2}^n r_i} v_1 \xrightarrow{n \rightarrow \infty} 0,$$

since  $\sum_{l \geq 2} r_l = \infty$ . Now we place ourselves in the case where  $a_1 < 1$ . Using (40), we have  $\mathbf{w}^n = [\prod_{l=1}^{n-1} (T_1 + \mu_{l+1} T_2)] \mathbf{w}^1$  which gives

$$v_n = \mu_n T_2 \left[ \prod_{l=1}^{n-1} (T_1 + \mu_{l+1} T_2) \right] (\mathbf{w}^1)_1 = \mu_n T_2 \left[ \sum_{\mathbf{k}=(k_1, \dots, k_m) \in \mathcal{P}_{n-1}} \sum_{\mathbf{i}=(i_1, \dots, i_m) \in \{1, 2\}^m} \bar{\mu}_i^{\mathbf{k}} T_{i_m}^{k_m} \dots T_{i_1}^{k_1} \right] (\mathbf{w}^1)_1,$$

with  $\mathcal{P}_{n-1} = \{\mathbf{k} = (k_1, \dots, k_m) \in (\mathbb{N}^*)^m; s.t. k_1 + \dots + k_m = n-1\}$  the set containing all the partitions of  $(n-1)$  and  $\bar{\mu}_i^{\mathbf{k}} = \prod_{l \in \mathcal{A}_i^{\mathbf{k}}} \mu_l$ . Note that each integer  $l \in \{1, \dots, n-1\}$  belongs to a group  $p_l$  (i.e.  $p_l = \sum_{r=1}^{l} k_r \leq l < \sum_{r=1}^{p_l+1} k_r$ ) and the set  $\mathcal{A}_i^{\mathbf{k}} = \{l \in \{1, \dots, n-1\}; s.t. i_{p_l} = 2\}$  refers to the groups where we select the operator  $T_2$  instead of  $T_1$ .

Let  $\epsilon > 0$  and  $m \in \mathbb{N}^*$ . Using that  $\|\mathbf{w}^1\|_\infty = \sup_{l \geq 1} |\mathbf{w}_l^1|$  is finite, we have  $T_2 T_1^m (\mathbf{w}^1)_1 = \sum_{l \geq m+1} a_l \mathbf{w}_l^1 \leq (\sum_{l \geq m+1} a_l) \|\mathbf{w}^1\|_\infty$ . Since  $\sum_{l \geq 1} a_l < \infty$ , there exists  $n_0 \in \mathbb{N}$  such that when  $m \geq n_0$ , we have  $T_2 T_1^m (\mathbf{w}^1)_1 \leq \frac{\epsilon}{2}$ . We write  $\bar{a}$  for the quantity  $\bar{a} = \max_{l \leq n_0} a_l$ . Since  $a_1 < 1$  and  $\sum_{l \geq 1} a_l = 1$ , we have  $\bar{a} < 1$ . Thus there exists  $n_1 \in \mathbb{N}$  such that for all  $m \geq n_1$ , we have  $\bar{a}^{n_1} \leq \frac{\epsilon}{2}$ .

We denote by  $k^{T_1} = \sum_{r \geq 1, s.t. i_r = 1} k_r$  the integer that counts the number of times the operator  $T_1$  is chosen. When  $k^{T_1} \leq n_0 n_1$ , we have  $\#\{l, l \notin \mathcal{A}_i^{\mathbf{k}}\} \leq n_0 n_1$ , which gives

$$\bar{\mu}_i^{\mathbf{k}} = \prod_{l \in \mathcal{A}_i^{\mathbf{k}}} \mu_l = \prod_{l \in \mathcal{A}_i^{\mathbf{k}}} (1 - r_l) \leq e^{-\left(\sum_{l \in \mathcal{A}_i^{\mathbf{k}}} r_l\right)} = e^{-\left(\sum_{l=0}^n r_l\right) + \left(\sum_{l \notin \mathcal{A}_i^{\mathbf{k}}} r_l\right)} \leq e^{-\left(\sum_{l=0}^n r_l\right) + n_0 n_1}, \quad (41)$$

since  $r_l \leq 1$ . Thus, there exists  $n_2 \in \mathbb{N}$  such that for all  $m \geq n_2$ , we have  $\bar{\mu}_i^{\mathbf{k}} \leq \frac{\epsilon}{2}$ .

We take now  $n \geq n_3 = (n_0 n_1) \wedge n_2$ . When  $k^{T_1} > n_0 n_1$ , we differentiate two cases. First, if there exists  $r$  such that  $k_r \geq n_0$  and  $i_r = 1$ . In such case, since  $T_2(u) \leq T_\infty(u)$  with  $T_\infty(u)_n = \|u\|_\infty \mathbf{1}_{n=1}$ , we have

$$T_2(T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 = T_2(T_{i_m}^{k_m} \dots T_2^{k_{r+1}} T_1^{n_0} T_1^{k_r - n_0} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 \leq \frac{\epsilon}{2} T_2(T_{i_m}^{k_m} \dots T_\infty^{k_{r+1}} T_1^{k_r} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1. \quad (42)$$

Second, when we cannot find such  $r$ , then necessarily  $k^{T_2} \geq \frac{n}{n_0} \geq n_1$  with  $k^{T_2} = \sum_{r \geq 1, s.t. i_r = 2} k_r$  represents the number of times the operator  $T_2$  is selected. Since  $\mathbf{w}_n^1 = \epsilon \mathbf{1}_{n=1}$  and  $k_l \leq k_0$  whenever  $i_l = 2$ , the coefficients  $a_l$  with  $l > k_0$  never appear. In such situation, we get  $T_2(u) \leq \bar{a} T_\infty(u)$  using basic inequalities. The inequalities  $T_2(u) \leq \bar{a} T_\infty(u)$  and  $\bar{a}^{k^{T_2}} \leq \bar{a}^{n_1} \leq \frac{\epsilon}{2}$  ensure

$$T_2(T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 \leq \bar{a}^{k^{T_2}} T_2(T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 \leq \frac{\epsilon}{2} T_2(T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1, \quad (43)$$

with  $\bar{i}_m = i_m \mathbf{1}_{i_m=1} + \infty \mathbf{1}_{i_m=2}$ . When  $k^{T_1} \leq n_0 n_1$ , we use (41) and  $n \geq n_2$  to get

$$\bar{\mu}_i^k T_2(T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 \leq \frac{\epsilon}{2} (T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1. \quad (44)$$

By combining the inequalities (42), (43), (44) and  $T_2(u) \leq T_\infty(u)$ , we conclude

$$\begin{aligned} v_n &\leq \frac{\epsilon}{2} \left[ \sum_{k^{T_1} \leq n_0 n_1} (T_{i_m}^{k_m} \dots T_{i_1}^{k_1})(\mathbf{w}^1)_1 \right] + \frac{\epsilon}{2} \left[ \sum_{k^{T_1} > n_0 n_1} (T_{\bar{i}_m}^{k_m} \dots T_{\bar{i}_1}^{k_1})(\mathbf{w}^1)_1 \right] \\ &\leq \frac{\epsilon}{2} \left[ \prod_{l=1}^{n-1} (T_1 + \mu_{l+1} T_2) \right] (\mathbf{w}^1)_1 + \frac{\epsilon}{2} \left[ \prod_{l=1}^{n-1} (T_1 + \mu_{l+1} T_\infty) \right] (\mathbf{w}^1)_1 \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq \epsilon. \end{aligned}$$

which proves (38).

**Step (ii):** Let us prove (39). Note that  $(h_n)_{n \geq 0} = (\beta^n a_n^\mu)_{n \geq 0}$  verifies

$$h_0 = 0, \quad h_{n+1} = \mu_{n+1} \underbrace{\sum_{l=0}^n (\beta^{n+1-l} a_{n+1-l}^\mu)}_{=g_{n+1-l}} h_l, \quad \forall n \geq 0,$$

with  $(g_n)_{n \geq 0}$  the sequence defined such that  $g_n = \beta^n a_n^\mu$ . This means that  $h_n = g_n^\mu$ . Thus a straightforward application of (38) to the sequence  $(g_n^\mu)_{n \geq 0}$  gives (39). This completes the proof.  $\square$

**Lemma V.G.3.** Let  $\mathcal{Z}$  be a finite space and  $(z, z_1) \in \mathcal{Z}^2$ . Under Assumption 4.4, there exist two constants  $\beta \in [0, 1)$  and  $d > 0$  such that

$$a_n(z_1) = \mathbb{P}_z[\tau_{\mathcal{Z}_1 \geq n}] \leq \beta^n d, \quad \forall n \in \mathbb{N}. \quad (45)$$

*Proof of Lemma V.G.3.* Let  $(z, z_1) \in \mathcal{Z}^2$  and  $A_{z_1} = \{z' \in \mathcal{Z}; \mathbb{P}[Z_1 = z_1 | Z_0 = z'] = 1\}$  the states that leads to  $z_1$  with probability 1. To avoid the states  $z' \in \mathcal{Z}$  where the quantity  $\mathbb{P}[Z_1 \neq z_1 | Z_0 = z'] = 0$ , we first show (45) for the Markov chain  $\tilde{Z}_n = Z_n \mathbf{1}_{Z_n \notin A_{z_1}} + z_1 \mathbf{1}_{Z_n \in A_{z_1}}$ . Then, we prove (45) in the general case.



**Step (i):** We prove (45) by induction on  $n \in \mathbb{N}^*$  for the Markov chain  $\tilde{Z}$ . We denote by  $\tau_{\tilde{Z}_1}^1$  the first visit time of the Markov chain  $\tilde{Z}$  to the state  $z_1$ . For  $n = 1$ , we have

$$\mathbb{P}_z[\tau_{\tilde{Z}_1}^1 = 1] = \mathbb{P}_z[\tilde{Z}_1 = z_1] \leq \max_{z' \in \tilde{\mathcal{Z}}} [\max(\mathbb{P}_{z'}[\tilde{Z}_1 = z_1], \mathbb{P}_{z'}[\tilde{Z}_1 \neq z_1])] = \beta.$$

Note that  $\beta \in [0, 1)$  since  $\mathbb{P}_{z'}[\tilde{Z}_1 = z_1] < 1$ . Let us take now  $n > 1$ . We have

$$\mathbb{P}_z[\tau_{\tilde{Z}_1}^1 \leq n] = \mathbb{P}_z[\tau_{\tilde{Z}_1}^1 = 1] + \sum_{z' \neq z_1} P(z, z') \mathbb{P}_z[\tau_{\tilde{Z}_1}^1 \leq n-1],$$

with  $P(z, z') = \mathbb{P}[\tilde{Z}_1 = z' | \tilde{Z}_0 = z]$ . The induction assumption gives  $\mathbb{P}_z[\tau_{\tilde{Z}_1}^1 \leq n-1] \geq 1 - \beta^{n-1}$  which ensures

$$\mathbb{P}_z[\tau_{\tilde{Z}_1}^1 \leq n] \geq (1 - P_{z, z_1^c}) + (1 - \beta^{n-1}) P_{z, z_1^c} \geq 1 - \beta^n,$$

with  $P_{z, z_1^c} = \mathbb{P}[\tilde{Z}_1 \neq z_1 | \tilde{Z}_0 = z]$ . Thus, we get  $\mathbb{P}_z[\tau_{\tilde{Z}_1}^1 > n] \leq \beta^n$ .

**Step (ii):** Let us now prove (45) for  $Z$ . Since  $Z$  reaches  $z_1$  at most one iteration after  $\tilde{Z}$ , we have

$$\mathbb{P}_z[\tau_{\mathcal{Z}_1}^1 \leq n] \leq \mathbb{P}_z[\tau_{\tilde{Z}_1}^1 \leq n-1] \leq \beta^{n-1} = \underbrace{\frac{1}{\beta}}_{=d} \beta^n, \quad \forall n \geq 1.$$

Recall that  $\tau_{\mathcal{Z}_1}^1$  is the first visit time of the Markov chain  $Z$  to the state  $z_1$ . This completes the proof.  $\square$

## V.G.2 Propagation of the error

As a first step, we consider the case where the process  $(Z_k)_{k \geq 1}$  is Markovian. For any  $z_1 \in \mathcal{Z}$ , we write  $\tau_{z_1}^1 = \inf\{l > 0, Z_l = z_1\}$  for the first visit time of  $(Z_k)_{k \geq 1}$  to the state  $z_1$ . Moreover, we define the sequence  $(a_k)_{k \geq 0}$  such that  $a_k(z_1) = \mathbb{P}[\tau_{z_1}^1 \geq k | \mathcal{F}_0]$ . We have the following result.

**Proposition V.G.1.** *Let  $(Z_k)_{k \geq 1}$  be a Markov process,  $z_1 \in \mathcal{Z}$  and  $n \in \mathbb{N}^*$ . Under Assumptions 4.1, 4.2, 4.3 and 4.4, we have*

$$\mathbb{E}_0[e^n(z_1)] \leq \epsilon_n + \sum_{j=1}^{n-1} a_{n-j} \bar{\alpha}_j \mathbb{E}_0[e^j(z_1)],$$

with  $\epsilon_n = e^1(z_1) a_n(z_1) + \sum_{j=1}^{n-1} a_{n-j} M_j$  and  $\bar{\alpha}_j = \frac{\mathbb{E}_0[\alpha_j e^j(z_1)]}{\mathbb{E}_0[e^j(z_1)]} \mathbf{1}_{\mathbb{E}_0[e^j(z_1)] \neq 0}$ . The variables  $\alpha_j$  and  $M_j$  are given by (17) and (18).

*Proof of Proposition V.G.1.* Let  $(z, z_1) \in \mathcal{Z}^2$  and  $n \in \mathbb{N}^*$ . Using the last-exit decomposition, see Section 8.2.1 in [123], we have

$$\begin{aligned}
 \mathbb{E}_z[e^n(z_1)] &= \mathbb{E}_z[e^n(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[e^n(z_1)\mathbf{1}_{\{\tau_{z_1}^1 \geq n-j, Z_j=z_1\}}] \\
 &= \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[e^{j+1}(z_1)\mathbf{1}_{\{\tau_{z_1}^1 \geq n-j, Z_j=z_1\}}] \\
 &\stackrel{\text{Proposition 4.1}}{\leq} \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)\mathbf{1}_{\{\tau_{z_1}^1 \geq n-j, Z_j=z_1\}}] \\
 &= \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)\mathbf{1}_{\{Z_j=z_1\}}\mathbb{E}[\mathbf{1}_{\tau_{z_1}^1 \geq n-j} | \mathcal{F}_j]] \\
 &\leq e^1(z_1)\mathbb{E}_z[\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{P}_{z_1}[\tau_{z_1}^1 \geq n-j] \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)] \\
 &= \epsilon_n + \sum_{j=1}^{n-1} a_{n-j} \bar{\alpha}_j \mathbb{E}_z[e^j(z_1)],
 \end{aligned}$$

with  $a_j(z_1) = \mathbb{P}_{z_1}[\tau_{z_1}^1 \geq j]$ ,  $\tau_{z_1}^1 = \inf\{l > 0, Z_l = z_1\}$  and  $\alpha_n$  and  $M_n$  defined in Proposition 4.1. In the second equality, we use that  $e^n(z_1)$  does not change as long as the state  $z_1$  is not reached. This completes the proof.  $\square$

We turn now to the general case. However, we use the same kind of arguments needed in the Markov setting. Let  $n \in \mathbb{N}^*$  and  $(z_1, \dots, z_n) \in \mathcal{Z}^n$ . For any  $i \in \{1, \dots, n\}$ , we denote by  $A^i$  the set  $A^i = \{Z_1 = z_1, \dots, Z_i = z_i\}$ . Moreover, we write  $\tau_{z_1}^n = \inf\{l > 0, Z_{l+n} = z_1\}$  and  $a_k^n = \mathbb{P}[\tau_{z_1}^{n-k} \geq k | \mathcal{F}_n]$ . We have the following result.

**Proposition V.G.2.** *Under Assumptions 4.1, 4.2, 4.3 and 4.4, we have*

$$\mathbb{E}_0[e^n(z_1)\mathbf{1}_{A^n}] \leq \epsilon_n + \sum_{j=1}^{n-1} a_{n-j}^n \bar{\alpha}_j \mathbb{E}_0[e^j(z_1)\mathbf{1}_{A^n}],$$

with  $\epsilon_n = e^1(z_1)a_n^0(z_1) + \sum_{j=1}^{n-1} a_{n-j}^n M_j$  and  $\bar{\alpha}_j = \frac{\mathbb{E}_0[\alpha_j e^j(z_1)\mathbf{1}_{A^n}]}{\mathbb{E}_0[e^j(z_1)\mathbf{1}_{A^n}]}$   $\mathbf{1}_{\mathbb{E}_0[e^j(z_1)\mathbf{1}_{A^n}] \neq 0}$ . The variables  $\alpha_j$  and  $M_j$  are given by (17) and (18).

*Proof of Proposition V.G.2.* We have

$$\begin{aligned}
 \mathbb{E}_z[e^n(z_1)\mathbf{1}_{A^n}] &= \mathbb{E}_z[e^n(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n, A^n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[e^n(z_1)\mathbf{1}_{\{\tau_{z_j}^j \geq n-j, Z_j=z_1, A^n\}}] \\
 &= \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n, A^n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[e^{j+1}(z_1)\mathbf{1}_{\{\tau_{z_j}^j \geq n-j, Z_j=z_1, A^n\}}] \\
 &\stackrel{\text{Proposition 4.1}}{\leq} \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n, A^n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)\mathbf{1}_{\{\tau_{z_j}^j \geq n-j, Z_j=z_1, A^n\}}] \\
 &\leq \mathbb{E}_z[e^1(z_1)\mathbf{1}_{\tau_{z_1}^1 \geq n, A^n}] + \sum_{j=1}^{n-1} \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)\mathbf{1}_{\{Z_j=z_1, A^n\}} \mathbb{E}[\mathbf{1}_{\{\tau_{z_j}^j \geq n-j\}} | \mathcal{F}_n]]] \\
 &\leq e^1(z_1) \mathbb{E}_z[\mathbf{1}_{\tau_{z_1}^1 \geq n}] + \sum_{j=1}^{n-1} \mathbb{P}[\tau_{z_j}^j \geq n-j | \mathcal{F}_n] \mathbb{E}_z[(\alpha_j e^j(z_1) + M_j)\mathbf{1}_{A^n}] \\
 &\leq \epsilon_n + \sum_{j=1}^{n-1} a_{n-j}^n \bar{\alpha}_j \mathbb{E}_z[e^j(z_1)\mathbf{1}_{A^n}],
 \end{aligned}$$

In the second equality, we use that  $e^n(z_1)$  does not change as long as the state  $z_1$  is not reached. This completes the proof.  $\square$

### V.G.3 Proof of Theorem 4.1

*Proof of Theorem 4.1.* We split the proof in two steps. We first show (19) and (20) when  $\mathcal{Z}$  is finite then when  $\mathcal{Z}$  is countable.

**Step (i):** In this part, we prove (19) and (20) when the space  $\mathcal{Z}$  is finite. For clarity, we first proof the results when  $(Z_k)_{k \geq 0}$  is Markovian then we explain how to show it for the general case.

**Sub-step (i-1):** Here we demonstrate (19) when the space  $\mathcal{Z}$  is finite. Let  $z_1 \in \mathcal{Z}$  and  $n \in \mathbb{N}^*$ . Using Proposition V.G.1, the sequence  $v_n(z_1) = \bar{\alpha}_n(z_1) \mathbb{E}[e^n(z_1)]$  verifies

$$v_n(z_1) \leq \bar{\epsilon}_n(z_1) + \bar{\alpha}_n(z_1) \sum_{j=1}^{n-1} a_{n-j}(z_1) v_j(z_1),$$

with  $\bar{\epsilon}_n(z_1) = \bar{\alpha}_n(z_1) \epsilon_n(z_1)$ . For clarity and since there is no ambiguity, we forget the dependencies to  $z_1$  in the rest of the proof. Thus, using Lemma V.G.1, we get

$$v_n = \sum_{j=0}^n a_{n-j}^{\mu_j} \bar{\epsilon}_j,$$

with  $(a_n^\mu)_{n \geq 0}$  defined in (36). We can assimilate the sequence  $\bar{\epsilon}$  to the measure  $\mu = \sum_{k \geq 0} \bar{\epsilon}_k \delta_k$  with  $\delta_k$  the Dirac measure at  $k$ . Since  $\epsilon_k = e^1(z_1) a_k + \sum_{k=1}^{n-1} a_{n-k} \mathbb{E}[M_k]$ ,  $\sum_k a_k < \infty$ ,  $\sum_{k \geq 0} \mathbb{E}[\gamma_k^2] < \infty$  and  $\mathbb{E}[M_k] = O(\mathbb{E}[\gamma_k^2])$  for all the algorithms, we get  $\sum_{k \geq 0} \epsilon_k < \infty$  using the properties of

the Cauchy product between two sequences. This property ensures that the measure  $\mu$  has a finite mass. Additionally, using Lemma V.G.2, we have  $a_n^{\mu^j} \xrightarrow{n \rightarrow \infty} 0$  for any  $j \geq 0$ . Thus, the dominated convergence theorem ensures that  $v_n \xrightarrow{n \rightarrow \infty} 0$ .

When  $(Z_k)_{k \geq 0}$  is not Markovian, one can show using the same methodology above and Proposition V.G.2 that for any sequence  $\mathbf{z} = (z_k)_{k \in \mathbb{N}^*} \in \mathcal{Z}$  the quantity  $\tilde{u}_n = \mathbb{E}[e^n(z_1)\mathbf{1}_{A_{\mathbf{z}}^n}]$  converges towards 0 with  $A_{\mathbf{z}}^n = \{Z_1 = z_1, \dots, Z_n = z_n\}$ . Since  $e^n \mathbf{1}_{A_{\mathbf{z}}^n} \leq e^0 \mathbf{1}_{A_{\mathbf{z}}^n}$ , we use the dominated convergence theorem to get

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^n] = \sum_{\mathbf{z} \in S} \lim_{n \rightarrow \infty} \mathbb{E}[e^n \mathbf{1}_{A_{\mathbf{z}}^n}] = 0,$$

where  $S$  is the set of sequences valued in  $\mathcal{Z}$  and  $A_{\mathbf{z}}^n = \{Z_1 = z_1, \dots, Z_n = z_n\}$  for any  $\mathbf{z} \in \mathcal{S}$ .

**Sub-step (i-2):** In this step, we show (20) when the space  $\mathcal{Z}$  is finite. Since  $\gamma_k(z) = \gamma \mathbf{1}_{\{Z_k=z\}}$  for all  $(k, z) \in \mathbb{N} \times \mathcal{Z}$ , the quantity  $\alpha_k$  is constant (i.e.  $\alpha_k = \alpha$ ). Let  $M = \sup_n M_n$ . Using a direct induction and Proposition V.G.1, we get  $u_n = \mathbb{E}[e^n(z_1)] \leq c$  with  $c$  solution of the equation  $c = \gamma^2 M + \alpha c$  (i.e.  $c = \frac{\gamma^2 M}{1-\alpha}$ ). We define  $(m_n)_{n \geq 0}$  as follows:

$$m_n = \underbrace{\gamma^2 M + (e^1(z_1)) a_n(z_1)}_{=b_n} + \overbrace{\alpha \sum_{j=1}^n a_{n+1-j} v_j}^{=\tilde{a}_n}, \quad \forall n \geq 0.$$

By direct induction, we have  $u_n \leq m_n$  for all  $n \geq 0$ . Lemma V.G.1 gives

$$m_n = \sum_{l=0}^n a_{n-l}^{\mu^l} b_l = \gamma^2 M \left( \sum_{l=0}^n a_{n-l}^{\mu^l} \right) + \left( \sum_{l=0}^n a_{n-l}^{\mu^l} \tilde{a}_l \right), \quad \forall n \geq 0,$$

Using a direct induction and (36), we check that  $a_{n-l}^{\mu^l} = \alpha^{n-l} a_{n-l}^{*n-l}$ . The sequence  $a^{*l}$  is defined recursively such that  $a_n^{*1} = a_n$  and  $a_n^{*l+1} = (a^{*l} * a)_n = \sum_{j=1}^{n-1} a_{n+1-j}^{*l} a_j$  for any  $n \in \mathbb{N}$ .

Using this expression of  $a_{n-l}^{\mu^l}$ , we get

$$\gamma^2 M \left( \sum_{l=0}^n a_{n-l}^{\mu^l} \right) \leq \gamma^2 M \left( \sum_{l=0}^n \alpha^l \right) = \gamma^2 M \frac{1-\alpha^n}{1-\alpha} = (1-\alpha^n)c.$$

Thanks to Lemma V.G.3, we know that  $a_n(z_1) \leq \beta^n d$  with  $\beta \in [0, 1)$  and  $d > 0$  which gives  $\sum_{l \geq 0} r^n \tilde{a}_n(z_1) < \infty$  for any  $r < \frac{1}{\beta}$ . Let  $r < \min(1/\beta, 1/\alpha)$ . Using the properties of the Cauchy product between sequences, we have

$$\sum_{n \geq 1} r^n \left( \sum_{l=0}^n a_{n-l}^{\mu^l} \tilde{a}_l \right) \leq \sum_{n \geq 1} \left( \sum_{l=0}^n r^{n-l} \alpha^{n-l} r^l \tilde{a}_l \right) \leq \left( \sum_{n \geq 1} r^n \alpha^n \right) \left( \sum_{n \geq 1} r^n \tilde{a}_n \right) < \infty. \quad (46)$$

Inequality (46) ensures the existence of  $\bar{M} > 0$  such that  $(\sum_{l=0}^n a_{n-l}^{\mu^l} \tilde{a}_l) \leq \bar{\alpha}^n \bar{M}$  for any  $\tilde{a} < \max(\beta, \alpha)$ . This shows that  $m_n \leq (1-\bar{\alpha}^n)c + \bar{\alpha}^n \bar{M} e^1(z_1)$  which completes the proof when  $\mathcal{Z}$  is finite.

To handle the case where  $(Z_k)_{k \geq 0}$  is not Markovian, we first show using Proposition V.G.2 and the methodology above that for any sequence  $\mathbf{z} = (z_k)_{k \in \mathbb{N}^*} \in \mathcal{Z}$  we have  $\mathbb{E}[e^n \mathbf{1}_{A^n}] \leq (1 - \tilde{\alpha}^n)c + \tilde{\alpha}^n \bar{M} e^1(z_1) \mathbb{E}[\mathbf{1}_{A^n}]$  with  $A_{\mathbf{z}}^n = \{Z_1 = \mathbf{z}_1, \dots, Z_n = \mathbf{z}_n\}$ . Then, we use the same approach of Section Sub-step (i-1) to deduce that  $u_n \leq (1 - \tilde{\alpha}^n)c + \tilde{\alpha}^n \bar{M} e^1(z_1)$ .

**Step (ii):** Now we assume that  $\mathcal{Z}$  is countable. We prove the result when  $(Z_k)_{k \geq 0}$  is Markovian then we explain how to extend it in the general case.

**Sub-step(ii-1):** Let  $\epsilon > 0$ . Since  $E^1 = \sum_{k \geq 1} e^1(z_k) < \infty$ , there exists  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$ , we have  $\sum_{k \geq k_0} e^1(z_k) < \frac{\epsilon}{2}$ . We write  $A_{k_0}$  for the set  $A_{k_0} = \{z_k, k \leq k_0\}$ . Since  $A_{k_0}$  is finite, we use Sub-step (i-1) to show the existence of  $k_1 \in \mathbb{N}$  such that for all  $k \geq k_1$ , we have  $\sum_{k \leq k_0} e^{k_1}(z_k) < \frac{\epsilon}{2}$ . Let us now take  $k \geq k_1$ . Using  $(e^l(z))_{l \geq 1}$  is non-increasing for any  $z \in \mathcal{Z}$ , we get  $E^k = \sum_{k \geq k_0} e^1(z_k) + \sum_{k < k_0} e^1(z_k) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ .

**Sub-step(ii-2):** Let  $M = \sup_n M_n$  and  $\epsilon = (1 - \alpha)\gamma^2 Mr > 0$ . Since  $E^1 = \sum_{k \geq 1} e^1(z_k) < \infty$ , there exists  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$ , we have  $\sum_{k \geq k_0} e^1(z_k) < \epsilon$ . We write  $A_{k_0}$  for the set  $A_{k_0} = \{z_k, k \leq k_0\}$ . Since  $A_{k_0}$  is finite, we use Sub-step (i-2) to show that for all  $n \in \mathbb{N}$ , we have  $e^n(z_k) < (1 - \tilde{\alpha}^n)c + \tilde{\alpha}^n \bar{M} e^1(z_1)$  with  $z_k \in A_{k_0}$ . Using  $\epsilon = (1 - \alpha)\gamma^2 Mr \leq (1 - \alpha)\frac{\gamma^2 M}{1 - \alpha} r = (1 - \alpha)cr$  and  $(e^l(z))_{l \geq 1}$  is non-increasing for any  $z \in \mathcal{Z}$ , we deduce that  $\|e^n\|_\infty = \max(\sup_{k > k_0} e^1(z_k), \sup_{k \leq k_0} e^1(z_k)) \leq \max((1 - \tilde{\alpha}^n)c + \tilde{\alpha}^n \bar{M} e^1(z_1), \epsilon) \leq (1 - \tilde{\alpha}^n)c(1 + r) + \tilde{\alpha}^n \bar{M} e^1(z_1)$ .  $\square$

## V.H Proof of Proposition 4.2

**Lemma V.H.1.** *Under Assumptions 4.1, 4.2 and 4.3, we have*

$$\mathbb{E}[E^{k_{i+1}}] \leq \alpha^i \frac{\bar{M}\gamma_0(1+r)}{u - B\gamma_0}, \quad \forall i \in \mathbb{N}. \quad (47)$$

*Proof of Lemma V.H.1.* Let us show the result by induction on  $i \in \mathbb{N}$ . Since the proof for the initialisation step and the induction step are similar, we only show the induction step. We assume that (47) holds for a given  $i - 1$ . We consider the delayed error sequence  $u_n^{k_i} = \mathbb{E}[e^{n+k_i} | \mathcal{F}_{k_i}]$ . By following the same methodology used in the proof of Proposition V.G.1, we get

$$u_n^{k_i} \leq \epsilon_n^{k_i} + \alpha_{k_i} \sum_{j=1}^{k_i-1} a_{k_i+k-j} u_j^{k_i}, \quad \forall k \in \mathbb{N}^*.$$

This allows us to use Theorem 4.1 in order to exhibit a dominating sequence  $\bar{M}$  which verifies

$$\mathbb{E}_{z_i}[u_n^{k_i}] \leq (1 - \alpha_{k_i}^n) \left(1 + \frac{r}{2}\right) c_i + \alpha_{k_i}^n \bar{M} u_1^{k_i}, \quad \forall n \in \mathbb{N},$$

with  $c_i = \frac{\gamma_{k_i} \bar{M}}{(1-\alpha_{k_i})} \leq \alpha^{i+1} \frac{\bar{M} \gamma_0}{u - B \gamma_0}$  and  $u_1^{k_i} \leq \alpha^i \frac{\bar{M}(1+r) \gamma_0}{u - B \gamma_0}$  under the induction hypothesis. Thus,

by taking  $k_{i+1} = k_i + \lceil \frac{\log(\frac{\alpha r}{2(1+r)\bar{M}})}{\log(\alpha_{k_i})} \rceil$  and  $n_{i+1} = k_{i+1} - k_i$ , we check that  $\alpha_{k_i}^{n_{i+1}} \bar{M} \leq \frac{\alpha r}{2(1+r)}$

which gives

$$\mathbb{E}_{z_i}[e^{k_{i+1}}] = \mathbb{E}_{z_i}[u_{n_{i+1}}^{k_i}] \leq (1 + \frac{r}{2})c_i + \frac{r}{2}u_1^{k_i} \leq \alpha^{i+1} \frac{\bar{M}(1+r)\gamma_0}{u - B\gamma_0}.$$

This completes the proof.  $\square$

*Proof of Proposition 4.2.* Let  $L = |\log(\frac{\alpha r}{2(1+r)\bar{M}})|$  and  $i^* = \sup\{i \geq 0, \alpha_{k_i} \leq \beta\}$ . When  $\alpha_{k_i} \geq \beta$  (i.e.  $f(\gamma_{k_i}) \leq (1-\beta)$ ), we use the inequality  $\log(1-x) \leq -qx$ , for all  $x \in [0, 1-\beta]$  with  $q = \frac{1}{2-\beta}$ , to get

$$\begin{aligned} k_i &\leq L \sum_{\gamma_{k_j} \geq \underline{\gamma}} \frac{-1}{\log(\beta)} + L \sum_{\gamma_{k_j} < \underline{\gamma}} \frac{-1}{\log(1-f(\gamma_{k_i}))} \\ &\leq L \sum_{i \leq i^*} \frac{-1}{\log(\beta)} + L \sum_{i > i^*} \frac{1}{qf(\gamma_{k_i})} = L \sum_{i \leq i^*} \frac{-1}{\log(\beta)} + \frac{L}{qu} \sum_{i > i^*} \frac{1}{\gamma_{k_i}(1-\bar{b}\gamma_{k_i})}, \end{aligned}$$

with  $\bar{b} = \frac{b}{u}$ . Then, we use that  $\frac{1}{x(1-\bar{b}x)} = \frac{1}{x} + \frac{\bar{b}}{1-\bar{b}x}$  to deduce

$$k_i \leq Li \frac{-1}{\log(\beta)} + (i - i^*)_+ \frac{L\bar{b}}{qu(1-\bar{b}\bar{\gamma}^1)} + \frac{L}{qu} \sum_{i > i^*} \frac{1}{\gamma_{k_i}}.$$

Since  $\gamma_{k_i} = \alpha^i \gamma_0$ , we conclude that

$$k_i \leq L_1 i + L_2 (i - i^*)_+ + L_3 (\alpha^{-(i-i^*+1)_+} - 1).$$

with  $L_1 = \frac{-L}{\log(\beta)}$ ,  $L_2 = \frac{L\bar{b}}{qu(1-\bar{b}\bar{\gamma}^1)}$  and  $L_3 = \frac{L}{qu\gamma_0(1-\alpha^{-1})}$ . This completes the proof.  $\square$

## VI Proof of the convergence of $(v_n^k)_{n \geq 1, k \in (\mathbb{N}^*)^2}$

In this section we prove the following result.

**Proposition VI.1.** *The sequence  $(v_n^k)_{n \geq 1, k \geq 1}$  converges pointwise towards  $v$  on  $D_T \times D_q$  when  $n \rightarrow \infty$  and  $k \rightarrow \infty$ .*

*Proof of Proposition VI.1.* We prove this result in three steps. First, we show that  $v$  can be approximated by a numerical scheme  $\bar{v}^k$ . Then, we replace  $\bar{v}^k$  by another scheme  $v^k$  that also converges towards  $v$ . Finally, we show that  $v_n^k$  tends to  $v^k$  when  $n \rightarrow \infty$ .

**Step (i):** We start with our initial control problem where the agents may choose its trading speed at any time. It was studied by many authors, see for example [46], who show that the optimal trading speed verifies

$$v(t, q) = \frac{h_1(t) - qh_2(t)}{2\kappa}, \quad (48)$$

with  $h_1 : [0, T] \rightarrow \mathbb{R}$  and  $h_2 : [0, T] \rightarrow \mathbb{R}_+$  a positive function. The above equation shows that when  $q$  is large enough  $v$  has the opposite sign of  $q$ . Let us take such a large enough  $\tilde{q}$  and place ourselves in  $S_q = [-\tilde{q}, \tilde{q}]$  with  $\tilde{q} = 2\tilde{q}$ . In such case, Equation (48) ensures that  $v$  is bounded by  $\bar{v}$  on  $S_q$ . This show that we can replace the initial set of controls  $\mathbb{R}$  by the subset  $D \subset [-\bar{v}, \bar{v}]$  where the sign of  $v$  positive is (resp. negative) when the inventory is  $-\tilde{q}$  (resp.  $\tilde{q}$ ). Now, we approximate this problem in a classical way using the numerical scheme  $\bar{v}^k$  defined such that

$$\bar{v}^k(n_t, n_q) = \sup_{v \in D_a} \mathbb{E}[M_{n_t+1}^{n_t} - \phi Q_{n_t \Delta_t}^2 \Delta + \bar{v}^k(n_t + 1, n_{q+1}^v) | Q_{n_t \Delta_t} = n_q \Delta_q], \quad \forall (n_t, n_q) \in D_T \times D_q,$$

with  $M_{n_t+1}^{n_t} = M_{\Delta_t(n_t+1)}^{n_t}$ ,  $n_{q+1}^v$  the index such that  $Q_{(n_t+1)\Delta_t}^v = n_{q+1}^v \Delta_q$  and  $D_a = D \cap \{i \Delta_q, i \in \mathbb{Z}\}$ . Note that the set  $D_a$  is contained in our admissible set of controls  $A(n_t \Delta_t, n_q \Delta_q)$  when  $\Delta$  is small enough. This shows that we can substitute  $A(n_t \Delta_t, n_q \Delta_q)$  to  $D_a$  without any restriction. The convergence of  $(\bar{v}^k)_{k \geq \infty}$  towards  $v$  on the set  $D_T \times D_q$  when  $k \rightarrow \infty$  is standard.

**Step (ii):** We denote by  $v^k$  the numerical scheme

$$v^k(n_t, n_q) = \mathbb{E}[\sup_{v \in D_a} \{M_{n_t+1}^{n_t} - \phi Q_{n_t \Delta_t}^2 \Delta + v^k(n_t + 1, n_{q+1}^v)\} | Q_{n_t \Delta_t} = n_q \Delta_q], \quad \forall (n_t, n_q) \in D_T \times D_q.$$

Let us show that  $\bar{v}^k$  and  $v^k$  have the same limit. For this, we use a backward recurrence. For the moment, we assume that  $|\sup_v \mathbb{E}[M_{n_t+1}^{n_t}] - \mathbb{E}[\sup_v M_{n_t+1}^{n_t}]| \leq K \Delta_t^2$  and we will prove it at the end of Step (ii). We want to show that  $|\bar{v}^k(n_t, n_q) - v^k(n_t, n_q)| \leq K(T-t)\Delta_t$  for all  $(n_t, n_q) \in D_T \times D_q$ . At the terminal time  $\bar{v}^k$  and  $v^k$  coincide. We move now to the induction part. We have

$$\begin{aligned} |\bar{v}^k(n_t, n_q) - v^k(n_t, n_q)| &= |\sup_v \mathbb{E}[M_{n_t+1}^{n_t} - \phi q^2 \Delta_t + \bar{v}^k(n_t + 1, n_{q_t+1})] \\ &\quad - \mathbb{E}[\sup_v M_{n_t+1}^{n_t} - \phi q^2 \Delta_t + v^k(n_t + 1, n_{q_t+1})]| \\ &\leq |\sup_v \mathbb{E}[M_{n_t+1}^{n_t}] - \mathbb{E}[\sup_v M_{n_t+1}^{n_t}]| + |\mathbb{E}[\sup_v \{\bar{v}^k(n_t + 1, n_{q_t+1}) - v^k(n_t + 1, n_{q_t+1})\}]| \\ &\leq \mathbb{E}[\sup_v \{K \Delta_t^2 + K(T-t-\Delta_t)\Delta_t\}] \\ &= K(T-t)\Delta_t. \end{aligned}$$

In the third inequality, we use the induction assumption to complete the proof. Now, let us show that  $|\sup_v \mathbb{E}[M_{n_t+1}^{n_t}] - \mathbb{E}[\sup_v M_{n_t+1}^{n_t}]| \leq K\Delta_t^2$ . We write  $t = n_t\Delta_t$ ,  $\Delta Q = Q_{t+\Delta_t} - Q_t = v\Delta_t$ ,  $\Delta S = S_{t+\Delta_t} - S_t$  and  $\Delta \bar{S} = \int_t^{t+\Delta_t} (S_s - S_t) ds$ . Thus, we have

$$\begin{aligned} M_{n_t+1}^{n_t} &= (W_{t+\Delta} - W_t) + (Q_{t+\Delta}S_{t+\Delta} - Q_tS_t) \\ &= -v\Delta\bar{S} - \Delta QS_t - \kappa v^2\Delta_t + Q_t\Delta S + \Delta QS_t + \Delta Q\Delta S \\ &= -v\Delta\bar{S} - \kappa v^2\Delta_t + Q_t\Delta S + v\Delta_t\Delta S. \end{aligned} \tag{49}$$

The above equation shows  $\sup_v M_{n_t+1}^{n_t} = \frac{(\Delta_t\Delta S - \Delta\bar{S})^2}{4\kappa\Delta_t} + Q_t\Delta S$ . Using  $\mathbb{E}[\Delta S] = \alpha\Delta_t$  and

$$\mathbb{E}[\Delta\bar{S}] = \int_t^{t+\Delta} \alpha(s-t) ds = \alpha\Delta_t^2/2,$$

we get

$$\sup_v \mathbb{E}[M_{n_t+1}^{n_t}] = \frac{\alpha^2\Delta_t^3}{16\kappa} + \alpha Q_t\Delta_t \quad \text{and} \quad \mathbb{E}[\sup_v M_{n_t+1}^{n_t}] = \frac{\alpha^2\Delta_t^3}{16\kappa} + \frac{\sigma^2\Delta_t^2}{12\kappa} + \alpha Q_t\Delta_t.$$

Thus we deduce that  $|\sup_v \mathbb{E}[M_{n_t+1}^{n_t}] - \mathbb{E}[\sup_v M_{n_t+1}^{n_t}]| \leq K\Delta_t^2$  with  $K = \frac{\sigma^2}{12\kappa}$ .

**Step (iii):** Theorem 4.1, proves that  $v_n^k$  converges towards  $v^k$ . Thus by composition we have  $v_n^k$  converges pointwise towards  $v$  when  $n \rightarrow \infty$  and  $k \rightarrow \infty$  which completes the proof.  $\square$





---

# Price impact study of the electricity intra-day market

---

## Abstract

We propose the first analysis of the price behaviour in the intra-day electricity market. This market is unique since it has very specific operating rules. We first conduct a global descriptive study of the price variations and the trading activity in such market. Then, we investigate the price impact generated by a power plant breakdown, study the price structure and finally propose a model to reproduce these properties.

## 1 Introduction

The growth of electronic and fragmented exchanges generated deep changes in the financial markets during the last 20 years. In particular, it facilitated the access to high-quality data that describes the actions and interactions of market participants. This data has been thoroughly analysed in many equity markets, see [12, 14, 43, 73, 118, 127, 149]. However, none of these studies concern the intra-day electricity market since agents had access to such data only “few” years ago and the current operating rules in such market reduce the usefulness of this data. This market is maintained in a rudimentary competitive state by strong operating constraints that inhibit the development of substantial market making or arbitrage. Indeed, an agent needs, for example, to buy and sell back the asset during a short time window to generate profit since the lifetime of the asset is lower than one day (i.e. the longest trading period opens at 3 p.m and ends the day after at 11.30 p.m). Additionally, the trading activity in such markets is low compared to the one of electronic venues (i.e. small daily traded volume and low trading frequency) since they attract mainly market participants with economically driven decisions, see the website of Epex Spot.<sup>1</sup> This paper is, to the best of our knowledge, the first one to conduct a detailed, in-depth academic study of the price structure in the intraday electricity market which is a unique example of exchanges that are close to the old non-electronic markets and where empirical studies are possible since transactions are recorded.

---

<sup>1</sup><http://www.epexspot.com/>, visited on May 3<sup>rd</sup>, 2019

The price impact is an important feature of financial markets since it describes how the price absorbs a new information. It is defined as the average price move conditioned on a starting event. It is natural to think that some particular starting events will impact the price in a specific way. For example, one may expect that initiating buy orders tend to increase the price while initiating sell orders tend to drive it down. The empirical characterization of the price impact and its theoretical understanding are a major issue from practitioners point of view, see [26, 38, 102, 142]. On the one hand, a robust prediction of the market impact allows an accurate estimation of the transaction costs used by financial agents to build their trading strategies, see [12, 14, 102, 127]. On the other hand, the market impact shape may penalize big investment funds since it represents trading costs and therefore it contributes to selecting the size distribution of agents, see [26, 142]. Finally, a precise prediction of the price impact reflects the lack or excess of demand and thus it indicates to market participants the potential existence of profit opportunities, see [38].

The vast majority of conducted studies addressed the problem of the market impact for the initiating event: liquidation of metaorders, see [12, 14, 27, 43, 73, 118, 149]. The initiating event considered in this paper is the announcement of a power plant breakdown in the electricity intra-day market run by Epex Spot in Europe. This incident is similar to a metaorder execution since the owner of the plant needs to buy large quantities during the whole breakdown period. Indeed, after a power plant failure, the electricity distributor suffers from a lack of energy provision. Since he must provide his clients with their exact consumption at every minute and the electricity is not storable at large scales, the distributor sends immediately buy orders to compensate this lack. The literature has already paid attention to electricity intra-day markets, mainly with the scope of a renewable power plant owner seeking to reduce his imbalance, see for instance Garnier and Madlener [69], Henriot [89], and Aïd et al. [4]. The latter focuses on optimal trading in such market and uses the linear permanent/instantaneous market impact model of Almgren and Chriss [10] to incorporate the transaction costs that a producer may face while buying energy after a power plant outage.

The impact on the price is not the only important aspect of financial markets. In this paper, we also investigate the three following properties: the trade signs autocorrelation, the response function, see [39], and the price diffusivity. For the first two aspects (i.e. the trade signs and the response function), their characterization allows a better understanding and modelling of the price impact, see [39], which is in itself a crucial issue as discussed before. Additionally, these two first aspects can also be used as signals that trigger trading decisions. Here, we describe the price diffusivity through the Hurst exponent. This quantity determines the balance between two competitive ingredients of the price variations: the mean reversion (i.e. sub-diffusivity) and the price capacity to follow a trend (i.e. super-diffusivity). The Hurst exponent is also useful in the price modelling since it fixes the regularity of the fractional Brownian motion (fBm) that generates the dynamics of the price volatility, see [96]. Finally, the Hurst exponent value may indicate the existence or not of a long memory for the price process, see [70]. Based on our empirical study, we propose a model, inspired from the propagator approach in [39], to whom we add a drift component and the impact of exogenous information. This model is consistent with the four investigated properties and allows the estimation and prediction of the quantities of interest.

After a first descriptive analysis that describes the trading activity, the price distribution and the price seasonality in the electricity intra-day market, we show that the price impact generated by a power plant breakdown is linear. Then, we exhibit a short term correlation between the trade signs (i.e. the next trade sign depends mainly on the current one) and a power law decrease of the response function. After that, we study the diffusivity of the price process and show that the value of its Hurst exponent is  $H = 0.38$ . Finally, we propose a model for the observed price which reproduces the four investigated market properties.

The paper is organised as follows. First, we describe the intra-day electricity market and conduct a general descriptive analysis in Section 2. Then, we estimate the price impact generated by a power plant breakdown and investigate its variations for different price regimes in Section 3. In Section 4, we compute the autocorrelation function between the trade signs, the response function and the Hurst exponent of the price process and its volatility. Finally, we propose a price model consistent with the numerical results in Section 5 and draw our conclusions in Section 6. Proofs and additional results are relegated to an appendix.

## 2 Descriptive analysis of the intra-day electricity market

### 2.1 Description of the intra-day electricity market

Because of the non-storability of electricity, consumption must be equal to production when aggregated over an interconnected physical network. Thus, within each country, the Transmission System Operator (TSO) is the one in charge of implementing the equilibrium. To do so, the TSO pays agents to be responsible for the equilibrium over a part of the network, since the latter has millions of consumption units. These agents must ensure equilibrium otherwise they face penalties given by the TSO.

In continental Europe, all such registered agents can access a continuous-time market called the intra-day market, where they can buy or sell forward contracts (that is, a commitment to deliver energy during a specified time frame) for the next hours. Here we give a description of the electricity intra-day market runned by Epex Spot in France. A comparison with, say, Nordic markets can be found in the modeling article by Benth and Koekebakke [23]. Every day at 3 p.m, Epex Spot opens trading sessions<sup>2</sup> where agents trades forward contracts. The specificity of these forward contracts is that they should be delivered on the following day. Each trading session is associated to a unique time period and buying a forward contract in one given session ensures the supply of 1 MWh of energy during the related time period. Each of those sessions remains open until 30 minutes before the beginning of the delivery. When a session for a given delivery period is open, trading can occur continuously and is run using an order book mechanism.

While a session for a given delivery period is open, trading can occur continuously and is run using order books. Thus, an agent that is responsible for the equilibrium between production and consumption faces a choice when there is an outage in one of the plants in his portfolio:

---

<sup>2</sup>There are trading sessions for the 24 hours and 48 half hours in both France and Germany. For Germany, there are also 96 sessions for the quarters of an hour

he can switch another plant on if it is available or he can buy power in the intra-day markets. The choice is dictated by the costs of each solution and by the delay required to start the plants. Because of transparency rules, producers must inform all market participants of the outages in their production as soon as it happens. For that reason, market participant, who knows the production means of other members, may infer, with a reasonably high rate of success, whether a competitor facing an outage will buy energy in the intra-day market or not.

We end this presentation by describing the agents in the French and German intra-day electricity markets. According to the website of Epex Spot<sup>1</sup>, in 2019, there are 98 markets participants in France and 193 in Germany. Those participants are energy utilities, oil companies, aggregators of renewable production, banks and financial institutions. Market makers are thus present in the intraday markets.

### 2.2 Description of the database

We use two databases. The first one comes from the electricity intra-day market runned by Epex Spot in France. This database records, during two years from the 1st January 2017 to the 31th January 2019, the trades executed in the intra-day electricity market (i.e. the transaction price, the size of the transaction and the region of the transaction) event by event with a minute accuracy for each forward contract. The market is composed of 24 forward contracts and each one of them has a unique delivery hour (i.e. maturity) between 1 a.m and 12 p.m. Every day, trading sessions open at 3 p.m. In these sessions, agents can buy or sell forward contracts for maturities that vary from 1 a.m to 12 p.m of the day after. The second database reports all the power plants breakdowns that happened in the same period in the France market. This database enables us to select the observations of the intra-day market that occur during the breakdown periods. The first database accounts for 1 146 777 events and the second 12 109 breakdowns.

### 2.3 Features of the intra-day electricity market

**Breakdowns distribution.** There are two types of breakdowns: planned breakdowns and unexpected ones. For the planned ones, agents have enough time to take in advance measures that reduces the impact of the lack of energy provision generated by the breakdown. For the unexpected ones, agents react right after the power plant failure to compensate the under-supply of energy. Our database contains 25 853 planned breakdowns and 7831 unexpected ones. On average breakdowns last few hours and they generate an average loss of energy provision of 202 MW. Table VI.1 summarizes basic statistics about the power plant breakdowns.

## 2. Descriptive analysis of the intra-day electricity market

Type	Number of breakdowns	Mean energy loss (MW)	Duration (Hours)
Planned	2716	260	14
Unexpected	9393	177	10.5
All	12 109	202	11

Table VI.1 – Most important statistics of the power plant breakdowns.

Next, we focus on unexpected breakdowns and analyse for each forward contract the average number of transactions and the average exchanged volume during the breakdowns. The results are presented in Figure VI.1. We consider the maturities 18h, ... , 22h since they have the highest average daily volume, see Figure VI.1.

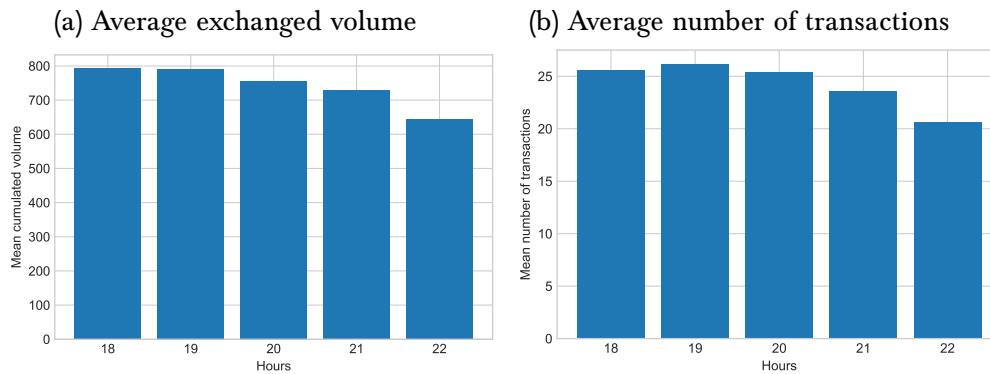


Figure VI.1 – (a) Average exchanged volume during breakdowns for each maturity, (b) Average number of transactions during breakdown for each maturity.

**Intra-day volumes.** Figure VI.2.a describes the daily average exchanged volume for each forward contract. The daily exchanged volume is at its lowest level for early maturities when the need of energy supply is low, it attains its highest level during afternoon maturities and then stabilizes for late maturities. It is important to note that for each contract the market activity accelerates close to the delivery hour (i.e. maturity), see Figure VI.2.b for the forward contract 20h.

**Price seasonality.** For each forward contract, Figure VI.3 shows some seasonality features of the weighted price (i.e. price normalized by the transaction volume). On average, the weighted price is low for early maturities, it begins to increase starting from the forward contract 9 a.m, reaches its highest level for afternoon maturities and stabilizes again for late delivery hours. Figure VI.3.b shows that the aggregated weighted price decreases during the weekend since the economic activity is reduced. Finally, Figure VI.3.b reveals that the weighted price during is low during the summer since the clients consumption is weak and becomes high in the winter when the clients need more energy supply.

## VI. Price impact study of the electricity intra-day market

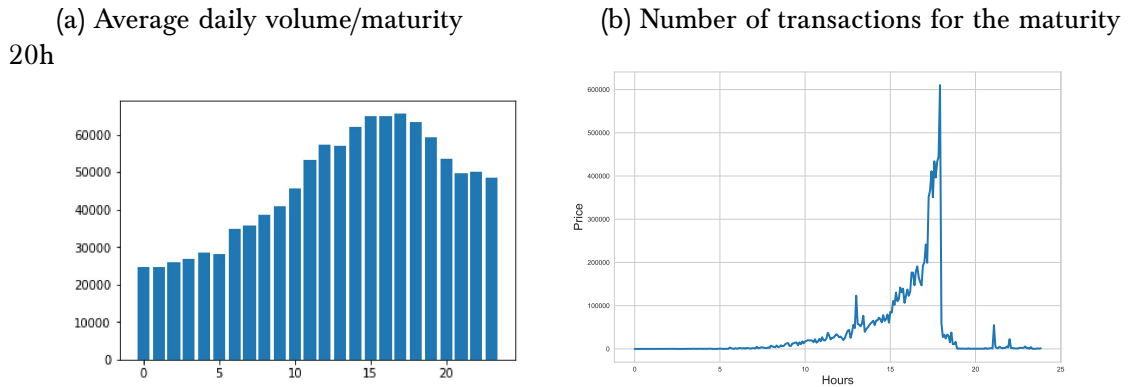


Figure VI.2 – (a) Average volume for breakdowns, (b) Average number of breakdowns for each maturity.



Figure VI.3 – (a) Average weighted price by maturity, (b) Average weighted price by days of the week, (c) Average weighted price by month.

### 3 Estimation methodology and definitions

Although the price impact curve during the execution may vary from a market to another one, its shape provides insight into the distribution of the agents composing the market. Here, we focus on the three following properties of this curve:

- **Linear:** a linear price impact means that the interaction between agents actions is weak since the impact mainly comes from exogenous information like market announcements. Such shape is expected in elementary market where the arbitrage and market making activity is not well developed or when the impact of all the agents is aggregated, see [135].
- **Convex:** a convex price impact means that agents detect the liquidation of a metaorder and react to it. In such case, the interaction between market participants decisions is strong and agents are ready to provide liquidity when other participants require it. This is expected in liquid equity markets, see for example [12, 14, 43, 118, 127].
- **Concave:** a concave price impact also means that agents identify the execution of a metaorder and reply to it. Thus, the interaction between agents decisions is strong however market participants avoid providing liquidity now. Such behaviour is expected when there is a financial crisis.

### 3.1 Estimation methodology

First, we select only unexpected ones since agents cannot take measures in advance to reduce the impact the power plant failure. Then, we require a minimum level of 150 MW energy loss for each breakdown otherwise its effect is not significant. After this first filter, we compute, for each one of the remaining breakdowns, the quantity

$$\Delta P_t = \epsilon_t \frac{P_t - P_{t_0}}{P_{t_0}}, \quad \forall t \in [t_0, t_1],$$

where  $P_t$  is the transaction price,  $\epsilon_t$  is the sign of the transaction (i.e.  $\epsilon_t = 1$  for buy market orders and  $-1$  otherwise) and  $t_0$  and  $t_1$  are defined as follows:

- $t_0$  is the first transaction time after the beginning of the breakdown.
- $t_1$  is the minimum time between the ending time of the breakdown, the end of the day and the closure time of the forward contract.

Note that the transaction price is not defined everywhere. In such case,  $P_t$  is the price of the last transaction before  $t$ . Finally, we define the price impact  $I_t$  as the empirical average of  $\Delta P_t$  over all the breakdowns and the rescaled time  $s$  such that  $s(t) = \frac{t-t_0}{t_1-t_0}$  for any  $t \in [t_0, t_1]$ .

## 3.2 Price impact profiles

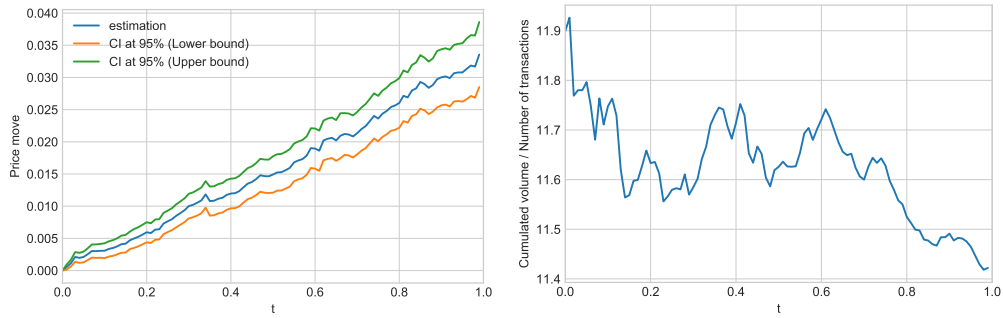
### 3.2.1 Aggregated price impact profiles

Figure VI.4.a shows that the aggregated price impact curve (i.e. for all the forward contracts) is linear. We choose to plot the aggregated impact since we have more observations however we check that our conclusion remain valid for the most important maturities (i.e. with high exchanged volume). As discussed before a linear price impact means that the interaction



between agents decisions is weak and their actions are mainly triggered by exogenous information like market announcements. Note that Figures VI.4.b and VI.4.c consolidates the shape of Figure VI.4.a. Indeed, the size of the transactions increases during the execution but the trading speed decreases (i.e. less transactions). This creates a balance that keep the slope of the price impact curve constant during the execution.<sup>3</sup>

(a) Price impact profile for aggregated maturities (b) Trading speed profile for aggregated maturities



(c) Transactions size profile for aggregated maturities

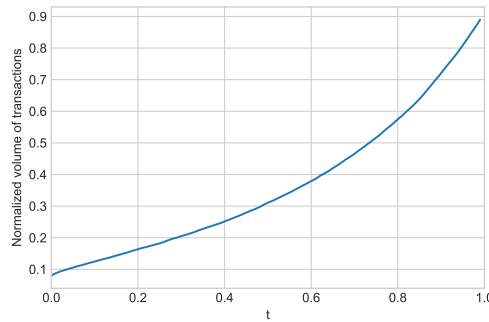


Figure VI.4 - (a) Price impact profile , (b) trading speed profile and (c) transactions size profile for aggregated maturities .

### 3.2.2 Price impact profiles for different regimes

In this section, we investigate the connection between the price impact curve and the amplitude of the energy lost during a breakdown (i.e. size of the breakdown). We choose this variable since it is the analogous of the size of the meta-order when dealing with execution issues. Many empirical studies have already explored the relation between the price impact at the end of the execution and the size of the meta-order, see []. They show that the price impact  $I$  generated by a meta-order of size  $Q$  follows the “square root law” (i.e.  $I \approx \left(\frac{Q}{V_D}\right)^\delta$  with  $\delta > 0$  and the daily volume).

<sup>3</sup>For example, in a situation, where both the transaction size and the trading speed decreases during the execution, the slope of the price impact curve will mechanically decrease.

We start by decomposing our initial data in three buckets with the same number of observations using quantiles. Figure VI.5 plots the price impact associated to each one of these buckets. Figure VI.5 shows that the global shape of the impact curve is unchanged when the energy loss vary since we observe a linear impact for all the buckets. However, the magnitude of impact increases with the size of the breakdown which is an expected result. Indeed, to compensate the energy loss of a huge breakdown, market participants need to buy large quantities from the intra-day market which increases the price.

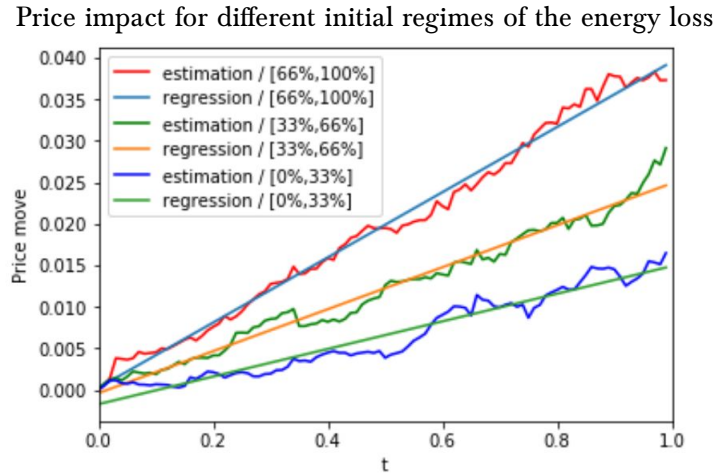


Figure VI.5 – Price impact for different initial price regimes.

## 4 Price study

### 4.1 Trade signs autocorrelation

Let  $l \in \mathbb{N}$ . In this section, we study the autocorrelation function  $C(l)$  between the trade signs  $\epsilon_l$  defined such that

$$C(l) = \mathbb{E}[\epsilon_l \epsilon_0] - \mathbb{E}[\epsilon_l] \mathbb{E}[\epsilon_0],$$

with  $\epsilon_l = 1$  for buy market orders and  $-1$  otherwise. Figure VI.6 shows the existence of a negative correlation between the first trade and its successor. Figure VI.6 also suggests the absence of memory since the long-term correlation between the trade signs is weak.

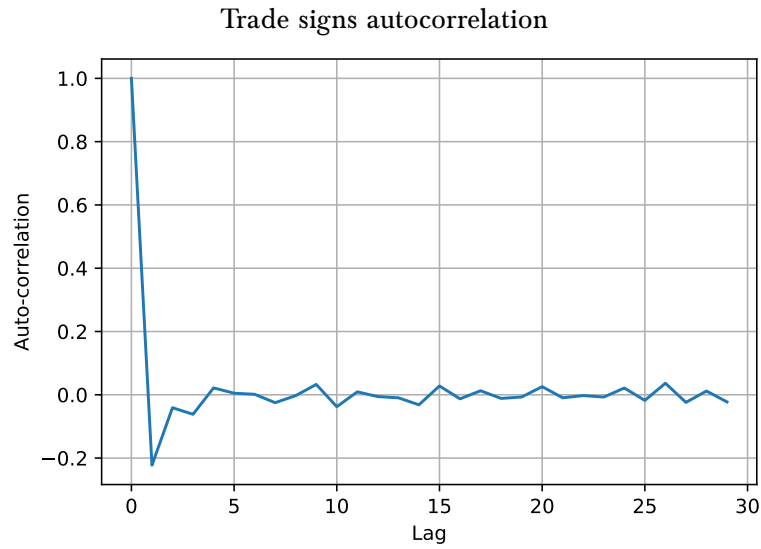


Figure VI.6 - The autocorrelation plot between trade signs.

#### 4.2 Response function autocorrelation

Let  $l \in \mathbb{N}$ . In this section, we study the response function  $R(l)$  between the price moves and the trade signs  $\epsilon_l$  defined such that

$$R(l) = \mathbb{E}[(\bar{p}_l - \bar{p}_0)\bar{\epsilon}_0],$$

with  $\bar{z} = z - \mathbb{E}[z]$  for any process  $z$ . Figure VI.7 reveals a power law decrease for the response function. Figure VI.7 also shows that the response function is higher than a bottom level close to 0.001.

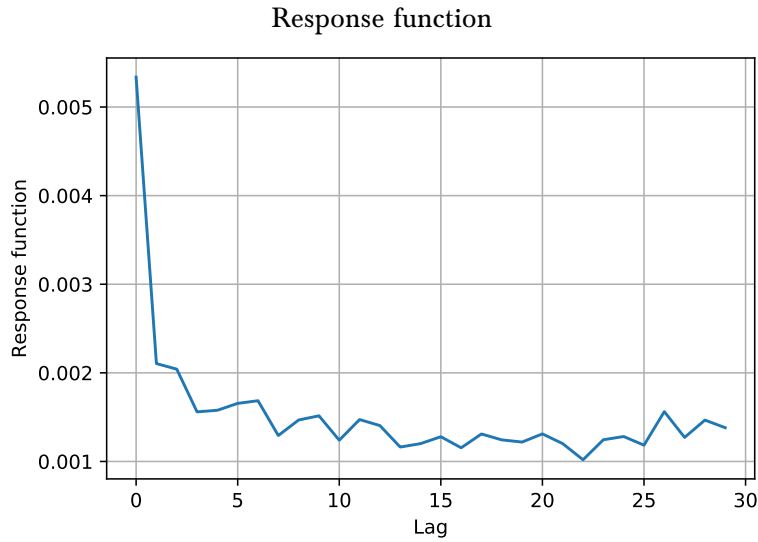


Figure VI.7 – Response function associated to the maturity 22h.

### 4.3 Price diffusion

In this section, we study the regularity of the price process and the volatility through their Hurst exponent  $H$ . Let  $Z_t$  be a process with the following dynamics:

$$dZ_t = \sigma dW_t^H, \quad \forall t \geq 0,$$

where  $\sigma$  is a constant volatility and  $W_t^H$  is a fractional Brownian motion (fBm) whose Hurst exponent is  $H$ . Then, there exists a constant  $C \geq 0$  such that

$$\mathbb{E}[|Z_{t+\Delta} - Z_t|^q] = C\Delta^{Hq}, \quad \forall t \geq 0, \forall \Delta \geq 0, \forall q \geq 0.$$

This shows that

$$y_\Delta^q = A'_q \Delta + C', \quad (1)$$

with  $y_\Delta^q = \log(\mathbb{E}[|p_{t+\Delta} - p_t|^q])$ ,  $A'_q = Hq$  and  $C' = \log(C)$ . Figures VI.8.a plots  $y_\Delta^q$  against  $\Delta$  when the underlying process  $Z$  is the price. We see that the linear relation 1 is verified. Figures VI.8.b plots  $A'_q$  for different values of  $q$  when the underlying process  $Z$  is the price and estimates the slope of this curve which represents the Hurst exponent. For both processes, the estimated Hurst exponent is close to  $0.48 \approx 0.5$  which shows that  $W^H$  is a Brownian motion.

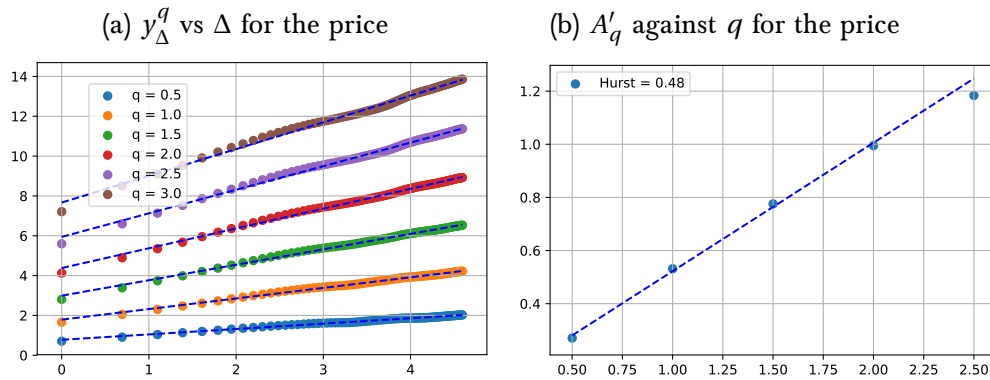


Figure VI.8 – (a) Plot of  $y_\Delta^q$  against  $\Delta$  during the years 2017-2018, (b) Plot of  $A'_q$  against  $q$  and estimation of the Hurst exponent for the price process associated to the maturity 22h.

## 5 Price modelling

In the literature, many authors have addressed the problem of modeling the market impact. First introduced models use a linear price impact, see [10, 13, 52, 80, 105], since it is the simplest theoretical setting where round-trip execution strategies (zero terminal inventory) with negative execution costs (profit making) are excluded. Additionally, when the impact is linear, it is possible to solve the optimal liquidation problem of a metaorder and to derive closed-form formulas for the solutions, see [10, 13, 80, 52]. After that, econophysicists start to tackle this problem and found that the market impact is non-linear using an empirical approach. In many of their articles they show that the price impact curve is concave and it can be well approximated by a square root function, see [12, 14, 43, 73, 118]. This new feature led to the development of several models to reproduce a non-linear impact. These approaches can be decomposed in three following categories:

- **Hawkes/propagator models:** the underlying idea behind Hawkes and propagator models is similar. Indeed, in the Hawkes framework, the arrival rate of each event depends fully on all the past market events, see [3, 14, 95]. While in the propagator model, it is the next price move which fully depends on all the past market events, see [39]. A link between these two types of models has been established in [95]. Additionally, we can solve, under both setting, optimal execution problems and derive optimal liquidation strategies, see [7, 6, 8].
- **Equilibrium models:** the market dynamics arise from interactions between rational agents acting optimally: each agent chooses the decision that maximizes its own utility. They were first introduced by economists with representative agents models, see for example the seminal work of Kyle [105] and the pioneer paper of Glosten and Milgrom [72] but also [66, 134, 139]. Since that, these models were extended to more general frameworks where a martingale condition is in force, see [64, 95, 100]. Such condition ensures no-arbitrage opportunities and a zero overall profit, see [64, 95, 100].

- **Statistical models:** the order flows of events are modeled by an appropriate random process in order to reproduce salient features of real markets, see [2, 3, 20, 55, 57, 106, 107, 144]. Such setting enables efficient computation of quantities of interest such as: trading costs, market impact or execution probabilities. Additionally, authors added, within such framework, rational agents that maximizes their own utility in order to solve liquidation problems, see Chapter II, or investigated the impact of transaction costs on Walrasian auctions<sup>4</sup>, see [62].

These three methodologies propose a large panel of possibilities and it is always possible to design, for each approach, a model that reproduces a relevant property of the price impact curve.

## 5.1 Description of the model

We can build a price model consistent with our observations. This model is inspired from the propagator model introduced in [39]. We show how to carry computations and calibrate this model in our specific setting. Such model is useful for prediction and estimation of the quantities of interest. We consider the following dynamics for the price at time  $n$ :

$$p_n = p_0 + \mu_n n + \sum_{k \leq n-1} G(n-k) \epsilon_k, \quad (2)$$

with  $\mu$  a real constant representing the price trend,  $G: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  the impact function that represents the influence of a single transaction on the price,  $\epsilon_n$  is the sign of the  $n$ -th trade.

## 5.2 Properties of the model

**Price impact.** Empirical evidences for the electricity market, see Figure VI.6, suggest that

$$C(1) \leq 0, \quad C(l) = 0, \quad \forall l \geq 2. \quad (3)$$

In such case, the price impact at time  $n$  verifies

$$I_n = \mathbb{E}[(p_n - p_0)\epsilon_n] = \mu_n \bar{e}n + G(1)C(1), \quad \forall n > 1,$$

with  $\bar{e} = \mathbb{E}[e_n]$ . Thus, we recover a linear price impact consistent with numerical observations, see Figures VI.4.a and VI.5. More generally, we always have

$$I_n = \mu_n \bar{e}n + \sum_{k=0}^{n-1} G(k)C(k), \quad \forall n \geq 0.$$

When the trades autocorrelation function and the kernel function satisfy a power law decrease which means

$$C(n) = \frac{K_0}{(h+n)^\gamma}, \quad G(n) = \frac{K_1}{(h+n)^\beta}, \quad (4)$$

<sup>4</sup>In Walrasian auctions the asset price is the point where the supply and demand curves intersect.

## VI. Price impact study of the electricity intra-day market

---

with  $(K_0, K_1) \in \mathbb{R}_+^2$ ,  $h \in \mathbb{R}_+$  and  $(\gamma, \beta) \in \mathbb{R}_+^2$ , we get

$$I_n = \mu_n \bar{\epsilon}_n + \sum_{k=0}^{n-1} \frac{K_0 K_1}{(h+n)^\gamma (h+n)^\beta} \sim \mu_n \bar{\epsilon}_n + \frac{K_0 K_1}{1-\gamma-\beta} (h+n)^{1-\gamma-\beta}.$$

We deduce that the shape of the price impact can vary from a linear to a power law function depending on the interaction between three factors: the price trend associated to  $\mu_n$ , the trade signs autocorrelation function  $C(n)$  and the impact function  $G$ .

**Response function.** We first assume that  $C$  satisfies (3) which is consistent with empirical evidences. Since  $C(1) \leq 0$  and  $G$  is non-negative, we find

$$R_l = \mathbb{E}[(\bar{p}_{n+l} - \bar{p}_n) \epsilon_n] = G(l) + C(1)[G(l+1) + G(l-1) - G(1)] = G(l) + C(1)[G(l+1) + G(l-1)] + a_1,$$

with  $\bar{p}_n = p_n - \mathbb{E}[p_n]$  and  $a_1 = C(1)G(1)$ . Note that the constant  $a_1$  represents the basis response level. Figure VI.4 shows that  $a_1$  is close to 0.01. Additionally, when the impact function  $G$  has a power law decrease which means that there exists  $K_1 \geq 0$  and  $\beta \geq 0$  such that

$$G(n) \underset{n \rightarrow \infty}{\sim} \frac{K_1}{n^\beta}, \quad (5)$$

we get  $R_l \underset{l \rightarrow \infty}{\sim} \frac{(1+2C(1))K_1}{l^\beta} + a_1$ . This ensures the power law decrease of  $R_l - a_1$  which is consistent with the empirical observations, see Figure VI.4. In the general case, both the impact function and the response function provide information about each other. Indeed, using Equation (2), we have

$$\begin{aligned} R_l &= \sum_{k=0}^{l-1} G(l-k)C(k) + \sum_{k \geq l} [G(l+k) - G(k)]C(k) \quad (a) \\ &= \sum_{k=1}^l G(k)[C(l-k) - C(k)] + \sum_{k \geq l+1} G(k)[C(k-l) - C(k)] \quad (b). \end{aligned} \quad (6)$$

Using Equation (6).a we can recover the response function using the impact function  $G$ . In particular, when both  $G$  and  $C$  verifies (4), we get  $R_l \underset{l \rightarrow \infty}{\sim} \frac{K_R}{l^{1-\gamma-\beta}}$  with  $K_R$  a non-negative constant. Conversely, when the response function is given, Equation (6).b shows that the impact function is solution of a linear system.

**Price and volatility diffusion.** We assume that  $C$  verifies (3) and  $G$  satisfies (5). These choices are consistent with the numerical observations. In such case, we use Equation (3) to get

$$\begin{cases} \bar{p}_n &= \sum_{k=0}^{n-1} G(n-k) \bar{\epsilon}_k, \\ \bar{p}_{n+l} - \bar{p}_n &= \sum_{k=0}^{l-1} G(n+l-k) \bar{\epsilon}_k + \sum_{k \geq l+1} [G(n+l-k) - G(n-k)] \bar{\epsilon}_k, \end{cases}$$

with  $\bar{p}_k = p_k - \mathbb{E}[p_k]$ ,  $\bar{\epsilon}_k = \epsilon_k - \mathbb{E}[\epsilon_k]$ . This gives

$$\Delta_l = \mathbb{E}[(\bar{p}_{n+l} - \bar{p}_n)^2] = \sum_{k < l} [(\bar{G}_k^l)^2 + 2C(1)\bar{G}_k^l \bar{G}_{k-1}^l],$$

with  $\bar{G}_k^l = G(l-k) - G(-k)\mathbf{1}_{k<0}$ . Since  $G$  satisfies (4), we deduce that  $\Delta_l \underset{l \rightarrow \infty}{\sim} \frac{K_\Delta}{l^h}$  with  $h = 1 - 2\beta$  and  $K_\Delta$  a non-negative constant. The empirical study of Section 4.3 gives an estimation of the exponent  $h$ . Thus, we can use this estimation to select  $\beta$ . In the more general case, we have

$$\Delta_l = \sum_{k \leq l-1} (\bar{G}_k^l)^2 + 2 \sum_{k < k' \leq l-1} C(k' - k) \bar{G}_k^l \bar{G}_{k'}^l.$$

When  $G$  and  $C$  satisfy (4), we deduce that  $\Delta_l \underset{l \rightarrow \infty}{\sim} \frac{K_\Delta}{l^h}$  with  $h = 1 - 2\beta - \gamma$  and  $K_\Delta$  a given non-negative constant. Using an empirical estimation of the exponents  $h$  and  $\gamma$ , see Section 4, we can approximate  $\beta$ .

## 6 Conclusion

We propose the first analysis of the price behaviour in the intra-day electricity market. This market is unique since it has very specific operating rules.

We find that the price impact is linear.

There is a very short term correlation between trade signs: Absence of memory. The response function has a power law. The price satisfies the EMH hypothesis since the Hurst exponent is close to 0.5.

We propose that reproduces these different properties which is inspired from . This model for prevision and estimation of the quantities of interests.





---

## Bibliography

---

- [1] Frédéric Abergel, Côme Huré, and Huyên Pham. Algorithmic trading in a microstructural limit order book model. *arXiv preprint arXiv:1705.01446*, 2017.
- [2] Frédéric Abergel and Aymen Jedidi. A mathematical approach to order book modeling. *International Journal of Theoretical and Applied Finance*, 16(05):1350025, 2013.
- [3] Frédéric Abergel and Aymen Jedidi. Long-time behavior of a Hawkes process-based limit order book. *SIAM Journal on Financial Mathematics*, 6(1):1026–1043, 2015.
- [4] René Aïd, Pierre Gruet, and Huyên Pham. An optimal trading problem in intraday electricity markets. *Mathematics and Financial Economics*, 10(1):49–85, 2016.
- [5] Yacine Ait-Sahalia and Mehmet Sağlam. High frequency market making: Optimal quoting. *Available at SSRN 2331613*, 2017.
- [6] Aurélien Alfonsi and Pierre Blanc. Dynamic optimal execution in a mixed-market-impact hawkes price model. *Finance and Stochastics*, 20(1):183–218, 2016.
- [7] Aurélien Alfonsi and Pierre Blanc. Extension and calibration of a hawkes-based optimal execution model. *Market Microstructure and Liquidity*, 2(02):1650005, 2016.
- [8] Aurélien Alfonsi and Alexander Schied. Capacitary measures for completely monotone kernels via singular control. *SIAM Journal on Control and Optimization*, 51(2):1758–1780, 2013.
- [9] Robert Almgren and Neil Chriss. Value under liquidation. *Risk*, 12(12):61–63, 1999.
- [10] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, pages 5–39, 2000.
- [11] Robert Almgren and Julian Lorenz. Bayesian adaptive trading with a daily cycle. *The Journal of Trading*, 1(4):38–46, January 2006.
- [12] Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li. Direct estimation of equity market impact. *Risk*, 18(7):58–62, 2005.
- [13] Marco Avellaneda and Sasha Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.

- [14] Emmanuel Bacry, Adrian Iuga, Matthieu Lasnier, and Charles-Albert Lehalle. Market impacts and the life cycle of investors orders. *Market Microstructure and Liquidity*, 1(2), December 2015.
- [15] Nicolas Baradel, Bruno Bouchard, and Ngoc Minh Dang. Optimal trading with online parameter revisions. *Market Microstructure and Liquidity*, page 1750003, 2016.
- [16] Nicolas Baradel, Bruno Bouchard, and Ngoc Minh Dang. Optimal control under uncertainty and bayesian parameters adjustments. *SIAM Journal on Control and Optimization*, 56(2):1038–1057, 2018.
- [17] Guy Barles and Panagiotis E Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.
- [18] Matthew Baron, Jonathan Brogaard, Björn Hagströmer, and Andrei Kirilenko. Risk and return in high-frequency trading. *Journal of Financial and Quantitative Analysis*, 54:993–1024, 2019.
- [19] Peter L Bartlett, Michael I Jordan, and Jon D McAuliffe. Convexity, classification, and risk bounds. *Journal of the American Statistical Association*, 101(473):138–156, 2006.
- [20] Christian Bayer, Ulrich Horst, and Jinniao Qiu. A functional limit theorem for limit order books with state dependent price dynamics. *The Annals of Applied Probability*, 27(5):2753–2806, 2017.
- [21] Paul R Beesack. On some Gronwall-type integral inequalities in  $n$  independent variables. *Journal of Mathematical Analysis and Applications*, 100(2):393–408, 1984.
- [22] Richard Bellman and Stuart Dreyfus. Functional approximations and dynamic programming. *Mathematical Tables and Other Aids to Computation*, pages 247–251, 1959.
- [23] Fred Espen Benth and Steen Koekebakker. Stochastic modeling of financial electricity contracts. *Energy Economics*, 30(3):1116–1157, 2008.
- [24] Albert Benveniste, Michel Métivier, and Pierre Priouret. Algorithmes adaptatifs et approximations stochastiques: théorie et applications à l'identification, au traitement du signal et à la reconnaissance des formes. 1987.
- [25] Albert Benveniste, Michel Métivier, and Pierre Priouret. *Adaptive algorithms and stochastic approximations*, volume 22. Springer Science & Business Media, 2012.
- [26] Jonathan B Berk and Richard C Green. Mutual fund flows and performance in rational markets. *Journal of political economy*, 112(6):1269–1295, 2004.
- [27] Nataliya Bershova and Dmitry Rakhlin. The non-linear market impact of large trades: Evidence from buy-side order flow. *Quantitative Finance*, 13(11):1759–1778, 2013.
- [28] Dimitri P Bertsekas and John N Tsitsiklis. *Neuro-dynamic programming*, volume 5. Athena Scientific Belmont, MA, 1996.

- 
- [29] Dimitris Bertsimas and Andrew Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.
- [30] Paul Besson, Stéphanie Pelin, and Matthieu Lasnier. To cross or not to cross the spread: that is the question! Technical report, Kepler Cheuvreux, Paris, June 2016.
- [31] Bruno Biais, Fany Declerck, and Sophie Moinas. Who supplies liquidity, how and when? Technical report, BIS Working Paper, May 2016.
- [32] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [33] Imran H Biswas, Espen R Jakobsen, and Kenneth H Karlsen. Viscosity solutions for a system of integro-pdes and connections to optimal switching and control of jump-diffusion processes. *Applied mathematics and optimization*, 62(1):47–80, 2010.
- [34] Pierre Blanc, Jonathan Donier, and Jean-Philippe Bouchaud. Quadratic Hawkes processes for financial prices. *Quantitative Finance*, 17(2):171–188, 2017.
- [35] Julius R Blum et al. Approximation methods which converge with probability one. *The Annals of Mathematical Statistics*, 25(2):382–386, 1954.
- [36] Léon Bottou and Olivier Bousquet. The tradeoffs of large scale learning. In *Advances in neural information processing systems*, pages 161–168, 2008.
- [37] Bruno Bouchard and Nizar Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.
- [38] Jean-Philippe Bouchaud, Doyne J Farmer, and Fabrizio Lillo. How markets slowly digest changes in supply and demand. In *Handbook of financial markets: dynamics and evolution*, pages 57–160. Elsevier, 2009.
- [39] 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, 2004.
- [40] Olivier Bousquet. Concentration inequalities and empirical processes theory applied to the analysis of learning algorithms. 2002.
- [41] Pierre Brémaud and Laurent Massoulié. Stability of nonlinear Hawkes processes. *The Annals of Probability*, pages 1563–1588, 1996.
- [42] Jonathan Brogaard, Matthew Baron, and Andrei Kirilenko. The Trading Profits of High Frequency Traders. Technical report, August 2012.
- [43] Xavier Brokmann, Emmanuel Serie, Julien Kockelkoren, and J-P Bouchaud. Slow decay of impact in equity markets. *Market Microstructure and Liquidity*, 1(02):1550007, 2015.

- [44] Eric Budish, Peter Cramton, and John Shim. The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response. *Quarterly Journal of Economics*, 130(4):1547–1621, November 2015.
- [45] Brian Bulthuis, Julio Concha, Tim Leung, and Brian Ward. Optimal execution of limit and market orders with trade director, speed limiter, and fill uncertainty. *International Journal of Financial Engineering*, 4(02n03):1750020, 2016.
- [46] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.
- [47] Álvaro Cartea, Ryan Donnelly, and Sebastian Jaimungal. Algorithmic trading with model uncertainty. *SIAM Journal on Financial Mathematics*, 8(1):635–671, 2017.
- [48] Álvaro Cartea, Ryan Donnelly, and Sebastian Jaimungal. Enhancing trading strategies with order book signals. *Applied Mathematical Finance*, 25(1):1–35, 2018.
- [49] Álvaro Cartea, Luhui Gan, and Sebastian Jaimungal. Trading co-integrated assets with price impact. *Mathematical Finance*, 29(2):542–567, 2019.
- [50] Álvaro Cartea and Sebastian Jaimungal. Optimal execution with limit and market orders. *Quantitative Finance*, 15(8):1279–1291, 2015.
- [51] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 10(3):339–364, 2016.
- [52] Álvaro Cartea, Sebastian Jaimungal, and José Penalva. *Algorithmic and High-Frequency Trading (Mathematics, Finance and Risk)*. Cambridge University Press, October 2015.
- [53] Patrick L Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- [54] Rama Cont and Adrien De Larrard. Order book dynamics in liquid markets: limit theorems and diffusion approximations. *Available at SSRN 1757861*, 2012.
- [55] Rama Cont and Adrien De Larrard. Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1):1–25, 2013.
- [56] Rama Cont and Arseniy Kukanov. Optimal order placement in limit order markets. *Quantitative Finance*, 17(1):21–39, 2017.
- [57] Rama Cont, Sasha Stoikov, and Rishi Talreja. A stochastic model for order book dynamics. *Operations research*, 58(3):549–563, 2010.
- [58] Corinna Cortes, Spencer Greenberg, and Mehryar Mohri. Relative deviation learning bounds and generalization with unbounded loss functions. *Annals of Mathematics and Artificial Intelligence*, 85(1):45–70, 2019.

- 
- [59] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media, 2007.
- [60] Khalil Dayri and Mathieu Rosenbaum. Large Tick Assets: Implicit Spread and Optimal Tick Size. *Market Microstructure and Liquidity*, 01(01):1550003, 2015.
- [61] Aaron Defazio, Francis Bach, and Simon Lacoste-Julien. Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. In *Advances in neural information processing systems*, pages 1646–1654, 2014.
- [62] Jonathan Donier and Jean-Philippe Bouchaud. From walras’ auctioneer to continuous time double auctions: A general dynamic theory of supply and demand. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(12):123406, 2016.
- [63] Joseph L Doob. *Stochastic processes*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1990.
- [64] J Doyne Farmer, Austin Gerig, Fabrizio Lillo, and Henri Waelbroeck. How efficiency shapes market impact. *Quantitative Finance*, 13(11):1743–1758, 2013.
- [65] Pietro Fodra and Huy en Pham. Semi-markov model for market microstructure. *Applied Mathematical Finance*, 22(3):261–295, 2015.
- [66] Thierry Foucault. Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, 2(2):99–134, 1999.
- [67] Daniel Fricke and Austin Gerig. Too fast or too slow? determining the optimal speed of financial markets. *Quantitative Finance*, 18(4):519–532, 2018.
- [68] S ebastien Gadat and Fabien Panloup. Optimal non-asymptotic bound of the ruppert-polyak averaging without strong convexity. *arXiv preprint arXiv:1709.03342*, 2017.
- [69] Ernesto Garnier and Reinhard Madlener. Balancing forecast errors in continuous-trade intraday markets. *Energy Systems*, 6(3):361–388, 2015.
- [70] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.
- [71] Jim Gatheral, Alexander Schied, and Alla Slynko. Exponential resilience and decay of market impact. In *Econophysics of order-driven markets*, pages 225–236. Springer, 2011.
- [72] Lawrence R Glosten and Paul R Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.
- [73] Carla Gomes and Henri Waelbroeck. Is market impact a measure of the information value of trades? market response to liquidity vs. informed metaorders. *Quantitative Finance*, 15(5):773–793, 2015.

- [74] Richard C Grinold and Ronald N Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk; 2nd ed.* McGraw-Hill, New York, NY, 2000.
- [75] Olivier Guéant. Optimal execution and block trade pricing: a general framework. *Applied Mathematical Finance*, 22(4):336–365, 2015.
- [76] Olivier Guéant. *The Financial Mathematics of Market Liquidity: From optimal execution to market making*, volume 33. CRC Press, 2016.
- [77] Olivier Guéant. Optimal market making. *Applied Mathematical Finance*, 24(2):112–154, 2017.
- [78] Olivier Gueant and Charles-Albert Lehalle. General intensity shapes in optimal liquidation. *Mathematical Finance*, 25(3):457–495, 2015.
- [79] Olivier Guéant, Charles-Albert Lehalle, and Joaquin Fernandez-Tapia. Optimal portfolio liquidation with limit orders. *SIAM Journal on Financial Mathematics*, 3(1):740–764, 2012.
- [80] 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, 2013.
- [81] Fabien Guilbaud and Huyen Pham. Optimal high frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94, 2013.
- [82] Xin Guo, Adrien De Larrard, and Zhao Ruan. Optimal placement in a limit order book: an analytical approach. *Mathematics and Financial Economics*, 11(2):189–213, 2017.
- [83] Xin Guo, Zhao Ruan, and Lingjiong Zhu. Dynamics of order positions and related queues in a limit order book. *arXiv preprint arXiv:1505.04810*, 2015.
- [84] Xin Guo, Howard Shek, Tze Leung Lai, and Samuel Po-Shing Wong. *Quantitative trading: algorithms, analytics, data, models, optimization*. Chapman and Hall/CRC, 2017.
- [85] Xin Guo and Mihail Zervos. Optimal execution with multiplicative price impact. *SIAM Journal on Financial Mathematics*, 6(1):281–306, 2015.
- [86] Martin Hairer. Ergodic properties of a class of non-markovian processes trends in stochastic analysis (London Mathematical Society Lecture note series vol 353), 2009.
- [87] Larry Harris. Maker-taker pricing effects on market quotations. *Unpublished working paper. University of Southern California, San Diego, CA*, 2013.
- [88] Terrence Hendershott, Charles M Jones, and Albert J Menkveld. Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1):1–33, 2011.
- [89] Arthur Henriot. Market design with centralized wind power management : handling low-predictability in intraday markets. *The Energy Journal*, 35(1):99–117, 2014.

- 
- [90] Weibing Huang, Charles-Albert Lehalle, and Mathieu Rosenbaum. Simulating and analyzing order book data: The queue-reactive model. *Journal of the American Statistical Association*, 10(509), December 2015.
- [91] Weibing Huang and Mathieu Rosenbaum. Ergodicity and diffusivity of Markovian order book models: a general framework. *SIAM Journal on Financial Mathematics*, 8(1):874–900, 2017.
- [92] James M Hutchinson, Andrew W Lo, and Tomaso Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3):851–889, 1994. bibtex\*[publisher=Wiley Online Library].
- [93] Jean Jacod. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(3):235–253, 1975.
- [94] Antoine Jacquier and Hao Liu. Optimal liquidation in a level-i limit order book for large-tick stocks. *SIAM Journal on Financial Mathematics*, 9(3):875–906, 2018.
- [95] Thibault Jaisson. Market impact as anticipation of the order flow imbalance. *Quantitative Finance*, 15(7):1123–1135, 2015.
- [96] Thibault Jaisson, Mathieu Rosenbaum, et al. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed hawkes processes. *The Annals of Applied Probability*, 26(5):2860–2882, 2016.
- [97] Rie Johnson and Tong Zhang. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in neural information processing systems*, pages 315–323, 2013.
- [98] Charles M Jones. What do we know about high-frequency trading? *Columbia Business School Research Paper*, (13-11), 2013.
- [99] Boyan Jovanovic and Albert J Menkveld. Middlemen in limit order markets. *Available at SSRN 1624329*, 2016.
- [100] Paul Jusselin and Mathieu Rosenbaum. No-arbitrage implies power-law market impact and rough volatility. *arXiv preprint arXiv:1805.07134*, 2018.
- [101] Andrei Kirilenko, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun. The flash crash: high frequency trading in an electronic market. *The Journal of Finance*, 72(3):967–998, 2017.
- [102] Robert Kissell, Morton Glantz, and Roberto Malamut. *Optimal trading strategies: quantitative approaches for managing market impact and trading risk*. PublicAffairs, 2003.
- [103] Harold Kushner and G George Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35. Springer Science & Business Media, 2003.



- [104] Harold Joseph Kushner and Dean S Clark. *Stochastic approximation methods for constrained and unconstrained systems*. Springer, 1978.
- [105] Albert P. Kyle. Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335, 1985.
- [106] Aimé Lachapelle, Jean-Michel Lasry, Charles-Albert Lehalle, and Pierre-Louis Lions. Efficiency of the Price Formation Process in Presence of High Frequency Participants: a Mean Field Game analysis. *Mathematics and Financial Economics*, 10(3):223–262, June 2016.
- [107] Peter Lakner, Josh Reed, and Sasha Stoikov. High frequency asymptotics for the limit order book. *Market Microstructure and Liquidity*, 2(01):1650004, 2016.
- [108] Sophie Laruelle, Charles-Albert Lehalle, and Gilles Pagès. Optimal posting price of limit orders: learning by trading. *Mathematics and Financial Economics*, 7(3):359–403, 2013.
- [109] Günter Last. On dependent marking and thinning of point processes. *Stochastic processes and their applications*, 45(1):73–94, 1993.
- [110] Charles-Albert Lehalle. Towards dynamic measures of liquidity. *AMF Scientific Advisory Board Review*, 1(1):55–62, September 2014.
- [111] Charles-Albert Lehalle and Eyal Neuman. Incorporating signals into optimal trading. *Finance and Stochastics*, 23(2):275–311, 2019.
- [112] Peter A W Lewis and Gerald S Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Naval research logistics quarterly*, 26(3):403–413, 1979.
- [113] Alexander Lipton, Umberto Pesavento, and Michael G Sotiropoulos. Trade arrival dynamics and quote imbalance in a limit order book. *arXiv preprint arXiv:1312.0514*, 2013.
- [114] Andrew W. Lo, Craig MacKinlay, and June Zhan. Econometric models of limit-order executions. *Journal of Financial Economics*, 65(1):31–71, July 2002.
- [115] Ananth Madhavan. Market microstructure: A survey. *Journal of Financial Markets*, 3(3):205–258, 2000.
- [116] Ananth Madhavan. Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal*, 68(4):20–35, 2012.
- [117] Iuliia Manziuk and Olivier Guéant. Deep reinforcement learning for market making in corporate bonds: beating the curse of dimensionality, June 2019.
- [118] Iacopo Mastromatteo, Bence Toth, and Jean-Philippe Bouchaud. Agent-based models for latent liquidity and concave price impact. *Physical Review E*, 89(4):042805, 2014.

- 
- [119] Albert J. Menkveld. High Frequency Trading and The New-Market Makers. *Journal of Financial Markets*, 16(4):712–740, November 2013.
- [120] Albert J Menkveld. High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740, 2013.
- [121] Sean P Meyn and Richard L Tweedie. Stability of Markovian processes i: Criteria for discrete-time chains. *Advances in Applied Probability*, 24(3):542–574, 1992.
- [122] Sean P Meyn and Richard L Tweedie. Stability of Markovian processes iii: Foster-lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993.
- [123] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.
- [124] Ciamac C. Moallemi. The Value of Queue Position in a Limit Order Book. In *Market Microstructure: Confronting Many Viewpoints*, Paris, December 2014. Louis Bachelier Institute.
- [125] Ciamac C Moallemi and Kai Yuan. A model for queue position valuation in a limit order book. *Columbia Business School Research Paper*, (17-70), 2016.
- [126] Maxime Morariu-Patrichi and Mikko S Pakkanen. State-dependent hawkes processes and their application to limit order book modelling. *arXiv preprint arXiv:1809.08060*, 2018.
- [127] Esteban Moro, Javier Vicente, Luis G Moyano, Austin Gerig, Doyne J Farmer, Gabriella Vaglica, Fabrizio Lillo, and Rosario N Mantegna. Market impact and trading profile of hidden orders in stock markets. *Physical Review E*, 80(6):066102, 2009.
- [128] Eric Moulines and Francis R Bach. Non-asymptotic analysis of stochastic approximation algorithms for machine learning. In *Advances in Neural Information Processing Systems*, pages 451–459, 2011.
- [129] Ioane Muni Toke. Stationary distribution of the volume at the best quote in a poisson order book model. *International Journal of Theoretical and Applied Finance*, 20(06):1750039, 2017.
- [130] Atsushi Nitanda. Stochastic proximal gradient descent with acceleration techniques. In *Advances in Neural Information Processing Systems*, pages 1574–1582, 2014.
- [131] James R. Norris. *Markov Chains (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, July 1998.
- [132] Yosihiko Ogata. On Lewis’ simulation method for point processes. *IEEE Transactions on Information Theory*, 27(1):23–31, 1981.

- [133] Bernt Karsten Øksendal and Agnes Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, 2005.
- [134] Christine A. Parlour. Price Dynamics in Limit Order Markets. *The Review of Financial Studies*, 11(4):789–816, 1998.
- [135] Felix Patzelt and Jean-Philippe Bouchaud. Universal scaling and nonlinearity of aggregate price impact in financial markets. *Physical Review E*, 97(1):012304, 2018.
- [136] Marcello Rambaldi, Emmanuel Bacry, and Jean-François Muzy. Disentangling and quantifying market participant volatility contributions. *Quantitative Finance*, pages 1–13, 2019.
- [137] Ryan Riordan and Andreas Storkenmaier. Latency, liquidity and price discovery. *Journal of Financial Markets*, 15(4):416–437, 2012.
- [138] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.
- [139] Ioanid Roşu. A dynamic model of the limit order book. *The Review of Financial Studies*, 22(11):4601–4641, 2009.
- [140] Jerome Sacks. Asymptotic distribution of stochastic approximation procedures. *The Annals of Mathematical Statistics*, 29(2):373–405, 1958.
- [141] Mark Schmidt, Nicolas L Roux, and Francis R Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. In *Advances in neural information processing systems*, pages 1458–1466, 2011.
- [142] Yonathan Schwarzkopf and J Doyne Farmer. Time evolution of the mutual fund size distribution. *arXiv preprint arXiv:0807.3800*, 2008.
- [143] Justin Sirignano and Konstantinos Spiliopoulos. Dgm: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375:1339–1364, 2018.
- [144] Eric Smith, Doyne J Farmer, Laszlo Gillemot, and Supriya Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative finance*, 3(6):481–514, 2003.
- [145] Laruelle Sophie and Lehalle Charles-albert. *Market Microstructure in Practice*. World Scientific, 2018.
- [146] Sasha Stoikov. Time is Money: Estimating the Cost of Latency in Trading. In Frédéric Abergel, Jean-Philippe Bouchaud, Thierry Foucault, Charles-Albert Lehalle, and Mathieu Rosenbaum, editors, *Market Microstructure: Confronting Many Viewpoints*, Paris, December 2014. Louis Bachelier Institute.
- [147] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.

- [148] Bence Toth, Zoltan Eisler, Fabrizio Lillo, Julien Kockelkoren, J-P Bouchaud, and J Doyne Farmer. How does the market react to your order flow? *Quantitative Finance*, 12(7):1015–1024, 2012.
- [149] Bence Tóth, Yves Lempriere, Cyril Deremble, Joachim De Lataillade, Julien Kockelkoren, and Jean-Philippe Bouchaud. Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2):021006, 2011.
- [150] Alexander B Tsybakov et al. Optimal aggregation of classifiers in statistical learning. *The Annals of Statistics*, 32(1):135–166, 2004.
- [151] Christopher John Cornish Hellaby Watkins. Learning from delayed rewards. 1989.
- [152] Paul J Werbos. Building and understanding adaptive systems: A statistical/numerical approach to factory automation and brain research. *IEEE Transactions on Systems, Man, and Cybernetics*, 17(1):7–20, 1987.
- [153] Peng Wu, Marcello Rambaldi, Jean-François Muzy, and Emmanuel Bacry. Queue-reactive hawkes models for the order flow. *arXiv preprint arXiv:1901.08938*, 2019.

**Titre:** Contrôle optimal et apprentissage statistique pour la modélisation du carnet d'ordres

**Mots clés:** Carnet d'ordres, trading optimal, modèle de file d'attente, processus de Hawkes, propriétés ergodiques, apprentissage par renforcement.

**Résumé:** L'objectif principal de cette thèse est d'étudier certains problèmes de contrôle rencontrés par les agents financiers et de comprendre les interactions entre les différentes quantités d'intérêt tels que les flux d'ordres, la liquidité, les prix et la volatilité à différentes échelles de temps (c'est-à-dire partant de la microseconde jusqu'à plusieurs années). On cherche à la fois à proposer des modèles utiles et à construire des estimateurs pour leurs paramètres. Ces modèles permettent de reproduire les principales caractéristiques des marchés et de concevoir des outils utiles pour les agents et les régulateurs.

Dans les premiers chapitres de cette thèse, on a étudié des problèmes de contrôle rencontrés par différents types d'agents: brokers institutionnels, market makers et les traders haute fréquence. La résolution de ces problèmes a produit des stratégies d'achat et de vente en accord avec les comportements de ces agents. On utilise ensuite ces stratégies pour simuler un marché où les brokers institutionnels, les market makers et les traders haute fréquence interagissent de manière optimale les uns avec les autres. Le modèle introduit dans les chapitres précédents est réaliste et utile, mais il présente certaines restrictions. Par exemple, les décisions prises par les agents dépendent

uniquement de l'état du carnet d'ordres alors que les stratégies de certains acteurs peuvent être plus complexes. Pour dépasser ces difficultés, on présente un modèle plus général qui permet de calculer en fonction de flux individuels plusieurs indicateurs pertinents. Il est notamment possible de classer les market makers en fonction de leur contribution à la volatilité

Pour résoudre les problèmes de contrôle soulevés dans la première partie de la thèse, nous avons développé des schémas numériques. Une telle approche est possible lorsque la dynamique du modèle est connue. Lorsque l'environnement est inconnu, on utilise généralement des méthodes d'approximation stochastique. Dans le cinquième chapitre, nous proposons un algorithme permettant d'accélérer la convergence de telles méthodes.

Les approches considérées dans les chapitres précédents sont adaptées pour des marchés liquides utilisant le mécanisme du carnet d'ordres. Cependant, cette méthodologie n'est plus nécessairement pertinente pour des marchés régis par des règles de fonctionnement spécifiques. Pour répondre à cette problématique, nous proposons, dans un premier temps, d'étudier le comportement des prix sur le marché très particulier de l'électricité.

**Title:** Optimal control and statistical learning for order book modelling

**Keywords:** Limit order book, optimal trading, queuing model, Hawkes processes, ergodic properties, reinforcement learning.

**Abstract:** The goal of this thesis is to study some control problems faced by market participants and to understand the interactions between financial processes such as order flows, liquidity, prices and volatility at different time scales (i.e. starting from the microsecond accuracy to several years). We aim at simultaneously proposing useful models and building estimation procedures for them. These models have to reproduce the most important features of the markets and enable us to design helpful tools for financial agents and regulators.

In the first three chapters of this thesis, we tackle control problems faced by different types of agents namely institutional brokers, market makers and high frequency traders. The resolution of these problems enables us to propose trading strategies consistent with typical market participants behaviours. Then, we use these strategies to simulate a market where institutional brokers, market makers and high frequency traders interact optimally with each other's.

The market model introduced in the previous chapters is realistic and useful however it has some limitations. For example, agents decisions depend on the market history

only through the last order book state while strategies of some market participants may be more intricate. To overcome such restriction, we propose a general order book model which allows us to compute several relevant microstructural indicators in terms of the individual flows. It is notably possible to rank market makers according to their own contribution to volatility.

In the first part of the thesis, we develop numerical schemes to solve agents control problems. This is possible when the dynamics of the model is known. To tackle control problems in an unknown environment, it is common to use stochastic approximation algorithms. In the fifth chapter, we propose a method that accelerates the convergence of such algorithms.

The approaches built in the previous chapters are appropriate for liquid markets that use an order book mechanism. However our methodologies may not be suitable for exchanges with very specific operating rules. To investigate this issue, as a first step, we study the price behaviour of the very particular intra-day electricity market.

