

Optimal market making under partial  
information  
and numerical methods for impulse control  
games  
with applications

by

Diego Zabaljauregui

Department of Statistics  
London School of Economics and Political Science

A thesis  
submitted for the degree of  
Doctor of Philosophy

London, United Kingdom, 2019

© Diego Zabaljauregui 2019

## **Declaration**

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party. I declare that my thesis consists of less than 100,000 words.

I confirm that Chapter 1 was jointly co-authored with my supervisor, Professor Luciano Campi, and I contributed 70% of this work.

I confirm that Chapter 2 was jointly co-authored with Professors René Aïd, Francisco Bernal, Mohamed Mnif and J.P. Zubelli, and I contributed 40% of this work.

## **Acknowledgments and dedication**

I would like to thank my supervisor, my department, my family and friends; with a special dedication to my wife and parents for their unconditional support.

## Abstract

The topics treated in this thesis are inherently two-fold. The first part considers the problem of a market maker who wants to optimally set bid/ask quotes over a finite time horizon, to maximize her expected utility. The intensities of the orders she receives depend not only on the spreads she quotes, but also on unobservable factors modelled by a hidden Markov chain. This stochastic control problem under partial information is solved by means of stochastic filtering, control and piecewise-deterministic Markov processes theory. The value function is characterized as the unique continuous viscosity solution of its dynamic programming equation. Afterwards, the analogous full information problem is solved and results are compared numerically through a concrete example. The optimal full information spreads are shown to be biased when the exact market regime is unknown, as the market maker needs to adjust for additional regime uncertainty in terms of P&L sensitivity and observable order flow volatility.

The second part deals with numerically solving nonzero-sum stochastic differential games with impulse controls. These offer a realistic and far-reaching modelling framework for applications within finance, energy markets and other areas, but the difficulty in solving such problems has hindered their proliferation. Semi-analytical approaches make strong assumptions pertaining very particular cases. To the author's best knowledge, there are no numerical methods available in the literature. A policy-iteration-type solver is proposed to solve an underlying system of quasi-variational inequalities, and it is validated numerically with reassuring results. In particular, it is observed that the algorithm does not enjoy global convergence and a heuristic methodology is proposed to construct initial guesses.

Eventually, the focus is put on games with a symmetric structure and a substantially improved version of the former algorithm is put forward. A rigorous convergence analysis is undertaken with natural assumptions on the players strategies, which admit graph-theoretic interpretations in the context of weakly chained diagonally dominant matrices. A provably convergent single-player impulse control solver, often outperforming classical policy iteration, is also provided. The main algorithm is used to compute with high precision equilibrium payoffs and Nash equilibria of otherwise too challenging problems, and even some for which results go beyond the scope of all the currently available theory.

# Table of Contents

<b>Introduction</b>	<b>1</b>
<b>1 Optimal market making with partial information</b>	<b>4</b>
Introduction . . . . .	4
1.1 Setting and main assumptions . . . . .	7
1.1.1 Preliminaries on the probability spaces . . . . .	7
1.1.2 Description of the model . . . . .	8
1.1.3 Jump measures . . . . .	11
1.1.4 Construction and characterization of the model . . . . .	12
1.2 Filtering problem . . . . .	14
1.3 Value function and HJB equation . . . . .	18
1.4 Full information . . . . .	25
1.4.1 Dimensionality reduction and the general system of ODEs . . . . .	26
1.4.2 Constrained inventory ODEs . . . . .	27
1.4.3 Unconstrained inventory ODEs . . . . .	29
1.4.4 General Verification Theorem . . . . .	30
1.4.5 Computing the optimal spreads: some particular cases . . . . .	33
1.5 Numerical analysis . . . . .	35
1.5.1 Comparing full and partial information optimal strategies . . . . .	37
1.5.2 Sample paths and a closer look at the filter . . . . .	41
<b>2 A policy iteration algorithm for nonzero-sum impulse games</b>	<b>44</b>
Introduction . . . . .	44
2.1 Analytical continuous-space problem . . . . .	46
2.1.1 General two-player nonzero-sum impulse games . . . . .	46
2.1.2 General system of quasi-variational inequalities . . . . .	47
2.1.3 Linear game and motivation . . . . .	49
2.2 Iterative algorithm for impulse games . . . . .	50
2.3 Numerical results . . . . .	54
2.3.1 Non-symmetric parabolic game . . . . .	54
2.3.2 Capped linear game . . . . .	58
2.3.3 Linear game with educated initial guess . . . . .	59
2.4 Concluding remarks . . . . .	61

<b>3</b>	<b>A fixed-point policy-iteration-type algorithm for symmetric impulse games</b>	<b>63</b>
	Introduction . . . . .	63
3.1	Analytical continuous-space problem:	
	the symmetric case . . . . .	64
3.2	Numerical discrete-space problem . . . . .	68
3.2.1	Discrete system of quasi-variational inequalities . . . . .	68
3.2.2	Iterative algorithm for symmetric games . . . . .	69
3.2.3	Iterative subroutine for impulse control . . . . .	72
3.2.4	Overall routine as a fixed-point policy-iteration-type method . . . . .	75
3.2.5	Convergence analysis . . . . .	77
3.2.6	Discretization schemes . . . . .	80
3.3	Numerical results . . . . .	82
3.3.1	Convergence to discrete solution on a fixed grid . . . . .	83
3.3.2	Convergence to analytical solution with refining grids . . . . .	87
3.3.3	Games without Nash equilibria . . . . .	87
3.3.4	Beyond the Verification Theorem . . . . .	88
3.4	Concluding remarks . . . . .	90
<b>A</b>	<b>Appendix: Matrix and graph-theoretic definitions and results</b>	<b>91</b>
	<b>References</b>	<b>94</b>

# List of Tables

1.1	Parameters for market making numerical tests . . . . .	37
2.1	Convergence tests: non-symmetric parabolic game . . . . .	55
2.2	Convergence tests: capped linear game . . . . .	59
2.3	Convergence tests: linear game (first) . . . . .	60
2.4	Convergence tests: linear game (second) . . . . .	61
3.1	Convergence of symmetric algorithm to analytical solution . . . . .	87

# List of Figures

1.1	Optimal spreads with full and partial information without terminal penalty .	38
1.2	Optimal spreads with full and partial information with terminal penalty . . .	40
1.3	Optimal spreads as functions of time and filter . . . . .	41
1.4	Sample paths of optimal spreads, inventory and filter . . . . .	42
2.1	Solution of non-symmetric parabolic game . . . . .	55
2.2	Sample paths of nonsymmetric parabolic game . . . . .	56
2.3	Monte Carlo simulations: equilibrium payoffs of non-symmetric parabolic game	57
2.4	Monte Carlo simulations: Nash equilibrium of non-symmetric parabolic game	58
2.5	Capped game as approximation of linear game . . . . .	59
2.6	Numerical solution of linear game . . . . .	60
2.7	Stagnation of accuracy of non-symmetric algorithm . . . . .	61
3.1	Nash equilibrium and payoff: linear game . . . . .	84
3.2	Nash equilibrium and payoff: quadratic running payoff . . . . .	84
3.3	Nash equilibrium and payoff: degree four running payoff . . . . .	84
3.4	Nash equilibrium and payoff: cubic running payoff . . . . .	85
3.5	Nash equilibrium and payoff: Ornstein–Uhlenbeck dynamics . . . . .	85
3.6	Nash equilibrium and payoff: Ornstein–Uhlenbeck dynamics with concave costs and gains . . . . .	86
3.7	Discontinuous impulses and irregular discontinuous payoffs (first) . . . . .	89
3.8	Discontinuous impulses and irregular discontinuous payoffs (second) . . . . .	89
3.9	Nash equilibrium and payoff: concave costs, quadratic running payoff and zero gain . . . . .	90



# Introduction

This thesis is divided in two main parts. The first one (Chapter 1) deals with the well-known and very relevant problem of *optimal market making* under a novel and realistic framework of *partial information*. The second one (Chapters 2 and 3), focuses on the development of numerical methods for a very general class of far-reaching models known as *nonzero-sum stochastic differential games with impulse controls*. From a mathematical viewpoint, *stochastic control* (in a broad sense) is the unifying underlying topic, spanning from classical ‘continuous’ controls, to impulse controls and impulse games; and covering the different levels of the problems: from the applications and motivation to the most technical aspects, from the theoretical analysis to the numerical methods and their implementation.

However, the problems, models and techniques used vary so widely, that a unique and comprehensive introduction is not only difficult, but likely counterproductive. For this reason, each chapter starts with a detailed introduction of its own. This section is therefore intended only to briefly motivate the problems to be studied, comment on the techniques used, and list the main resulting scientific contributions.

The optimal market making problem (Chapter 1) consists on determining how a liquidity provider for a given asset should (continuously) choose to set her bid/ask quotes if she wishes to maximize her expected utility. In doing so, she faces a complicated problem with both dynamic and static components, and several risks ranging from adverse price movements, exposure, execution costs and more. An increasingly popular approach, both academically and in practice, is the stochastic optimal control framework proposed by Avellaneda and Stoikov [AS08], where the liquidity taken from the market maker decays as a function of the spreads she quotes. Different variants of the first model have been proposed and thoroughly studied since, both for market making and optimal liquidation [AS08, BL14, CDJ17, CJ15, CJR14, FL12, FL13, Gué17, GL15, GLFT13].

Motivated by empirical evidence [CJ13], we consider a *hidden Markov-chain* model that unifies and generalizes features of all the previously mentioned ones, and admits the possibility that liquidity may also vary due to unobservable market conditions. We solve this combined control and partial information problem in full, using stochastic control, *stochastic filtering* and *piecewise-deterministic Markov processes* (PDMPs) theory [DF99]. The analogous and idealized problem in which the market maker can observe the market state is also solved, and results are compared numerically through a concrete example. Our main contributions are:

- A general and realistic Avellaneda–Stoikov-type model for optimal market making, incorporating for the first time the liquidity dependence on hidden market conditions and allowing for very general intensity (liquidity) shapes.

- A novel approach, within this framework, to characterize the value function of the market maker. The complexity of the dynamic programming equation in our model prevents us from using the most standard technique: proposing an ansatz and proving it valid by a verification theorem. Instead, we explicitly find the ansatz decomposition in terms of the value function of a new, diffusion-free, problem. This result is interesting in itself, as it opens the door to probabilistic and PDMPs numerical techniques when the dimension of the problem renders partial differential equations (PDEs) methods prohibitive.
- Ultimately, we characterize the ‘reduced’ value function as the unique continuous *viscosity solution* [FS06] of its formally derived equation.
- Having solved the idealized full information problem with a general verification theorem, we show that the corresponding optimal strategies become suboptimal when the market regime is unknown, as the market maker needs to manage a higher regime risk. We interpret the adjustment needed in terms of observable order flow volatility and sensitivity of the expected profit to observable regime changes. Ultimately we see and interpret why the bias of the full information optimal strategies becomes higher, the longer the waiting time in between orders.

Chapters 2 and 3 consider very general models known as nonzero-sum stochastic impulse control games [ABC<sup>+</sup>19, BCG19], in which players interact by shifting a certain continuous-time stochastic process at discrete (usually random) points in time. Such models have motivation and wide applicability within finance, energy markets, insurance and many other fields. Unfortunately, they are utterly challenging, what has rendered their study far too underdeveloped. In light of the evident need for numerical methods, and the nonexistence of any such one, Chapter 2:

- puts forward the very first iterative algorithm to solve nonzero-sum stochastic impulse games. The method is of *policy-iteration-type* [AF16, BMZ09] and solves an underlying system of differential *quasi-variational inequalities*, involving highly nonlinear, nonlocal and noncontractive operators, as well as non-smooth solutions.

The proposed algorithm is admittedly heuristic and no convergence analysis is carried out in this chapter. Instead, it is validated numerically over a series of experiments, considering different games over both fixed and refining grids, performing simulations and comparing results with the only (almost) fully analytically tractable game in the literature [ABC<sup>+</sup>19]. The results suggest that the algorithm can effectively be used to gain insight into general nonzero-sum stochastic impulse games.

In Chapter 3, we specialize our study by putting the focus on *symmetric* games. This subclass is broad enough to include plenty of interesting applications, such as competition between central banks [ABC<sup>+</sup>19, AF16, CZ99, JP93, MØ98], cash management problems [BCG19] and the generalization of many impulse control problems to the two-player instance.

For this type of games, an iterative algorithm which substantially improves the one in Chapter 2 is presented and rigorously studied, using techniques from *policy iteration for impulse control* [AF16], *fixed-point policy-iteration* [HFL12] and graph-theoretic notions relat-

ing to *weakly chained diagonally dominant (WCDD) matrices* and their recently introduced matrix sequence counterpart [Azi19a]. The main contributions of this chapter are:

- An iterative algorithm for the effective solution of symmetric nonzero stochastic impulse games, which improves the general one in terms of simplicity, efficiency, precision and stability.
- Providing the missing convergence analysis, which is performed under very natural assumptions on the players strategies and interpreted in terms of WCDD matrices and their extension in [Azi19a]. The latter is applied to impulse control for the first time, to the author's best knowledge.
- Proof of properties of contractiveness, boundedness of iterates and convergence to solutions, as well as sufficient conditions for convergence.
- As a by-product and to improve efficiency, an impulse control solver often outperforming classical policy iteration is also provided, and its convergence is proved.
- An extensive numerical validation is carried out, considering different performance metrics and addressing practical matters of implementation.
- Ultimately, equilibrium payoffs and Nash equilibria are computed with high precision, for games seemingly too challenging for the available analytical approaches, and even some for which results go beyond the scope of all the currently available theory (including discontinuous impulses and very irregular discontinuous payoffs). The latter also motivate further research into this field, with an emphasis on the need for a viscosity solutions framework.

# Chapter 1

## Optimal market making under partial information with general intensities

### Introduction

Given a financial market, a market maker (MM) can be understood as someone who provides liquidity for a certain asset. That is, she (almost) continuously posts bid/ask quotes for the asset, in the hope to profit from the bid-ask spread. In choosing how to do so, the MM faces a complicated problem on several levels; namely: the instantaneous margin/volume trade-off (the further away she quotes from the ‘fair price’, the less she gets executed, and vice-versa), adverse price movements and inventory risk (exposure), execution costs, and many others.

An increasingly popular mathematical approach to the MM problem, both in academia and in practice, is by means of stochastic optimal control. In particular, a line of research has focused on the modelling framework proposed by Avellaneda and Stoikov [AS08] (rooted in turn in [HS81]). Although widely motivated in the literature by order-driven markets such as equity markets, the shape of the limit order book is not explicitly taken into account. Thus, the framework is more easily understood and applied to over-the-counter (OTC) quote-driven markets such as the foreign exchange (FX) market, and we therefore choose to present it in this setting.

In this framework, the MM gives firm bid-ask quotes during a finite time interval by choosing bid/ask spreads with respect to a certain *reference price*.<sup>1</sup> Depending on the market, this price could be for example an aggregated mid-price or a dealer-to-dealer price, and it is frequently assumed to behave as an arithmetic Brownian motion. To explicitly model the margin/volume trade-off, the probability that the MM gets executed decays as a function of the corresponding spread. More precisely, it is assumed that the MM receives market orders according to counting processes of stochastic intensity. Most typically in the mathematical literature, intensities are assumed to decay exponentially on the spreads (mainly for tractability reasons). The goal of the MM is to find a strategy that allows her to maximize her expected terminal utility, which is taken as a constant absolute risk aversion (CARA) utility. The problem is then translated into a deterministic one: solving the

---

<sup>1</sup>Also referred to by some authors as *micro-price* [CJP15] or *efficient price* [DRR13] in the martingale case.

associated Hamilton–Jacobi–Bellman (HJB) equation, which is a partial-integro differential equation (PIDE) for the MM’s value function, and ultimately retrieving the optimal strategy in feedback form.

Within the described framework, a lot variants have been put forward and extensively studied. In [BL14, GLFT12], Bayraktar and Ludkovski, and Guéant, Lehalle and Fernandez-Tapia, apply a one-trading-side version to optimal liquidation in the risk-neutral and risk averse contexts, respectively. The introduction of a constraint on the inventory in [GLFT13], allowed the authors to rigorously solve the original problem of [AS08] with exponential intensities, by means of a verification theorem. This constraint has been widely used moving forward, with the exception of [FL12, FL13], where strategies are derived without a verification theorem.

Guéant and Lehalle continued to actively contribute to the area. In [GL15], they revisit the risk averse optimal liquidation problem for general intensities satisfying a certain ordinary differential inequality, and in [Gué17] the same is done for market making.<sup>2</sup> This assumption had been firstly introduced in [BL14, Sect.5.3] under risk-neutrality.

Other prolific contributors in the field are Cartea, Jaimungal and their coauthors [CJ15, CJP15, CJR14], who introduced a quadratic running penalty on the inventory to manage the ‘accumulated’ inventory risk for an otherwise risk-neutral MM. Constraints on the MM’s spreads have also been considered in some of them. It is worth noting that [Gué17] shows how the two seemingly different subclasses of models (risk averse and ‘risk-neutral’ with running penalty) can ultimately be characterized by a unique system of ordinary differential equations (ODEs) when considering two appropriate ansatz.

A relevant issue in practice that has not been included in the previous models is the following one. Empirical evidence (see, e.g., [CJ13]) suggests that liquidity taken by clients depends not only on the quoted spreads but also on other unobservable factors. Indeed, the confluence of factors such as market sentiment towards the asset and the competition with other market makers also affects the intensities at which the MM receives orders. This complicated effect can be modelled in a simplified fashion by making the intensities depend as well on a hidden finite-state Markov chain, effectively reflecting the regime or state of the market (different levels from very slow to very active). To the best of our knowledge, this has only been briefly done in the Avellaneda–Stoikov framework by proposing an approximation for the optimal strategy with exponential intensities [CJ13, Sect.5.1] or studying a simple two-states version with power-law intensities [BL14, Sect.5.4]. Both of these papers deal only with the ‘risk-neutral’ (possibly penalized) case and make the unrealistic assumption that the current market regime is known by the MM.<sup>3</sup>

When the state of the market is unknown, the problem becomes significantly more challenging for several reasons:

- It becomes a combined problem of stochastic control and filtering. The MM needs to dynamically make her best possible prediction of the market regime (or more precisely, its distribution), based on the information she has (i.e., the orders she has received so far and the evolution of the reference price), and adjust her spreads accordingly. This

---

<sup>2</sup>The latter paper also considers the multi-asset case.

<sup>3</sup> [BL09, Sect.3.3] (arXiv version) also studies optimal trade execution under partial information, albeit with uncontrolled intensity.

prediction is known as a filter.

- The associated HJB PIDE has higher dimension and more non-linearities.
- The standard approach used in all of the previously cited papers relies on reducing the HJB PIDE to a system of ODEs by means of an ansatz, proving that such a system has a classical (smooth) solution, and recovering the value function via a verification theorem. Under partial information however, the reduced HJB equation is still a complicated PIDE such that, in general, a classical solution may not exist (or it may be too difficult to prove otherwise). Hence, the ansatz argument breaks down.
- As a consequence, one needs to resort to the concept of viscosity solutions [FS06]. In addition, the numerical resolution unfailingly becomes a lot more involved than for simple systems of ODEs.
- On a technical level, the construction of the model is not straightforward. The MM needs to adjust her strategy based on her observable information, such as the arriving orders, but this flow of information is in turn affected by the MM's actions.

In this chapter (based on [CZ19]), we solve the problem of the MM under partial information. First, we start by unifying in one single formulation all the modelling features described so far. This allows us to simultaneously tackle all the models at once with a single approach, while generalising them at the same time. Indeed, our formulation allows for the interaction of any CARA utility (whether risk-neutral or risk averse) with running inventory penalty, terminal execution cost, inventory constraints and spread constraints. Further, motivated by practitioners' needs, we strongly generalize the intensity shapes to any continuous, decreasing to zero functions, adding modelling flexibility.<sup>4</sup> We even allow for the inventory to be unconstrained when no penalties are present. (This scenario is considered mainly for the sake of completeness and comparison, at almost no extra cost.)

Secondly, we let the intensities depend on a  $k$ -dimensional hidden Markov chain. Following [CEFS16a], we use a weak formulation<sup>5</sup> to construct a well-defined model (with exogenous information) and solve the filtering problem by means of the reference probability approach of stochastic filtering [Bré81, Chpt.VI]. The rigorous setting of the model and its full characterization are carried out in Section 1.1, while the filtering problem is solved in Section 1.2.

The optimization problem of the MM is then reformulated in terms of the usual state variables together with the  $k$ -dimensional observable distribution of the Markov chain (Section 1.3). At this point, the problem is too involved both analytically and numerically. We proceed by showing that, formally, there is an ansatz for the value function that reduces the dimensionality of the problem (i.e., the standard approach). However, as this methodology is not valid any more, we rigorously prove that the ansatz decomposition holds true (Theorem 1.3.1.1), and that the reduced value function can be characterized as the unique continuous viscosity solution of its formally derived equation (Theorem 1.3.1.3). The latter is done by

---

<sup>4</sup>This is done at the expense of renouncing to uniqueness in the optimal strategy. We also assume the decay to be 'fast enough' in certain cases.

<sup>5</sup>That is, with controlled probability measures.

harnessing results of piecewise-deterministic Markov process (PDMPs) [CEFS16a, DF99], as first defined in [Dav84].

Thirdly, we solve the idealized problem of a MM with full information, who can observe the Markov chain (Section 1.4). We show that a similar ansatz and the standard approach work out in this case. We prove that the MM's value function is a classical solution of its HJB equation via a general verification theorem (Theorem 1.4.4.1) and we recover the well known strategies for one regime as particular cases.

Finally, we compare the optimal strategies under full and partial information through a concrete example, by numerical analysis (Section 1.5). In particular, we show that the optimal full information spreads are biased when the exact regime is unknown, and using them becomes suboptimal. We interpret the adjustment needed in terms of observable order flow volatility and sensitivity of the expected profit to observable regime changes; and we show how this effect becomes higher, the longer the waiting time in between orders (leading to higher uncertainty for the MM).

## 1.1 Setting and main assumptions

### 1.1.1 Preliminaries on the probability spaces

We start by setting up our framework in an abstract fashion, deferring the question of existence of a model to Proposition 1.1.4.1, and a characterization of all such models to Proposition 1.1.4.2. As mentioned in the Introduction, we seek to construct a model under a weak formulation, so that the information flow remains exogenous (i.e., unaffected by the MM's actions).

Let  $T > 0$  be a given finite horizon and  $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T})$  a right-continuous filtered measurable space with  $\mathcal{F}_T = \mathcal{F}$ . Suppose it supports three adapted stochastic processes  $W, N^+, N^-$ . (More assumptions to be added in the sequel.) Let  $\mathbb{F}^{W, N^+, N^-} = (\sigma(W_u, N_u^+, N_u^- : 0 \leq u \leq t) \vee \mathcal{F}_0)_{0 \leq t \leq T}$  be the natural filtration of these processes enlarged ('completed') by  $\mathcal{F}_0$  and define the set of *admissible spreads* as

$$\mathcal{A} := \{\delta : [0, T] \times \Omega \rightarrow \overline{(\delta_*, \delta^*)} : \delta \text{ is } \mathbb{F}^{W, N^+, N^-}\text{-predictable and bounded}\}, \quad (1.1.1.1)$$

where  $-\infty \leq \delta_* < \delta^* \leq +\infty$  are fixed constants and  $\overline{(a, b)} = \mathbb{R} \cap [a, b]$  denotes the closure of the interval  $(a, b)$  in  $\mathbb{R}$ , for any  $-\infty \leq a < b \leq +\infty$ . Note that for  $\delta_* = -\infty$  and  $\delta^* = +\infty$ , the admissible spreads are not uniformly bounded. The self-imposed constraints  $\delta_*, \delta^*$  for the MM's spreads can be taken to be different for the bid and the ask if wanted, without any additional effort.

Consider on the former space a family  $(\mathbb{P}^\alpha)_{\alpha \in \mathcal{A}^2}$ ,  $\alpha = (\delta^-, \delta^+)$ , of equivalent probability measures such that the sigma algebra generated by their null sets is  $\mathcal{F}_0$ . Note that for each  $\alpha$ ,  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}^\alpha)$  is under the usual conditions and  $\mathcal{F}_0$  is the completed trivial sigma algebra. We refer to  $\mathbb{P}^\alpha$  as the *physical* or *historical* probability given the *admissible strategy* (or *control*)  $\alpha$ , and we write  $\mathbb{E}^\alpha$  for the expectation under  $\mathbb{P}^\alpha$ . We will drop the ' $\alpha$ ' from the notation when there is no room for ambiguity.

Henceforth, all the processes (resp. properties) considered are supposed to be defined (resp. hold) on the space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}^\alpha)$  for all  $\alpha \in \mathcal{A}^2$ , unless otherwise stated. For example, a

‘Brownian motion’ is a process on  $[0, T] \times \Omega$  that is an  $(\mathbb{F}, \mathbb{P}^\alpha)$ -Brownian motion for all  $\alpha$ . All subfiltrations are understood to have been augmented to satisfy the usual conditions (which in the case of the natural filtrations of Feller processes, it amounts simply to completing them). Càdlàg versions of the processes are used whenever available.

### 1.1.2 Description of the model

A market maker in a quote-driven market gives binding bid/ask quotes  $S^-, S^+$  resp. for a certain financial asset, during the time interval  $[0, T]$ . She receives unitary size market orders<sup>6</sup> to buy/sell according to counting processes  $N^-, N^+$  starting at 0, with a.s. no common jumps and stochastic intensities  $\lambda^-, \lambda^+$  resp. (see [Bré81, p.27, Def. D7]). Each intensity at time  $t$  depends on the corresponding spread  $\delta_t^\pm := \pm(S_t^\pm - S_t)$  between the MM’s quotes and a reference price  $S$ .  $S$  may be interpreted, e.g., as an aggregated mid-price or a dealer-to-dealer price, depending on the market. We assume  $S$  is a Brownian motion with drift, of the form

$$S_t = s_0 + \mu t + \sigma W_t, \quad (1.1.2.1)$$

where  $W$  is a Wiener process,  $s_0, \mu \in \mathbb{R}$  and  $\sigma \geq 0$ . Note that the MM can fully specify her quote  $S_t^\pm$  by choosing the spread  $\delta_t^\pm$ , since  $S_t^\pm = S_t \pm \delta_t^\pm$ . We assume  $\delta^-, \delta^+ \in \mathcal{A}$ . In particular, the MM chooses her spreads based on the observation of  $W$ ,  $N^-$  and  $N^+$ . We refer to  $\mathbb{F}^{N^-, N^+, W}$  as the *observable filtration*.

In addition, the intensities also depend on a hidden Markov chain  $Y$  with state space  $\{1, \dots, k\}$ , initial distribution  $\mu_0$  and deterministic generator matrix  $Q = (q_t^{ij})_{1 \leq i, j \leq k}$ . We assume  $Q$  is a continuous, stable and conservative  $Q(t)$ -matrix, i.e.:

**Assumptions 1.1.2.1 (Generator matrix).** *For all  $t \in [0, T]$  and  $1 \leq i \leq k$ :*

$$0 \leq q_t^{ij} < +\infty \text{ for all } 1 \leq j \leq k, j \neq i, \quad \sum_{j=1}^k q_t^{ij} = 0 \quad \text{and} \quad Q \in C[0, T].$$

The previous are standing assumptions in the literature of Markov chains and they guarantee in particular the existence of a unique (up to evanescence) càdlàg version of  $Y$ , and the existence of the predictable intensity kernel for its jump measure, as given in (1.1.3.1) below. In addition, we suppose that

$$Y \text{ has a.s. no common jumps with } N^- \text{ and } N^+. \quad (1.1.2.2)$$

In this model,  $Y_t$  represents the *market regime* at time  $t$  and results from the interaction of a range of different factors, such as market sentiment towards the asset and varying levels of competition with other liquidity providers. These effects are almost never explicitly modelled in the mathematical literature of market making. To this end, it is natural to assume that  $Y$  is not directly observable by the MM who can only see  $N^-, N^+$  and  $W$ . (As often assumed in filtering and control with hidden Markov chain models, the parameters  $k, Q$ , and  $\mu_0$  of  $Y$  are known to the MM, who needs to model/estimate them in practice.)

---

<sup>6</sup>Any constant size would work in the same way, but this convention simplifies the notation. See, e.g., [Gué17] for some formulas with arbitrary constant size under full information.



We define the MM's inventory process

$$N_t := n_0 + N_t^- - N_t^+ \quad (1.1.2.3)$$

and the cash account process

$$X_t^{\delta^-, \delta^+} := x_0 + \int_0^t (S_u + \delta_u^+) dN_u^+ - \int_0^t (S_u - \delta_u^-) dN_u^-, \quad (1.1.2.4)$$

for some fixed initial values  $n_0 \in \mathbb{Z}$  and  $x_0 \in \mathbb{R}$ . Note that

$$\mathbb{F}^{N^-, N^+} = \mathbb{F}^N \quad \text{and} \quad \mathbb{F}^{N^-, N^+, W} = \mathbb{F}^{N, W},$$

since  $N^-, N^+$  have a.s. no common jumps.

We shall make some very natural modelling assumptions on the intensities. These strongly generalize those in [BL14, Gué17, GL15] to a context with several market regimes, without the need of any conditions on the derivatives or even smoothness.

**Assumptions 1.1.2.2 (Orders intensities).** *There exist functions  $\Lambda_i^\pm : (\overline{\delta_*}, \delta^*) \rightarrow (0, +\infty)$  for  $i = 1, \dots, k$  and  $N^* \in \mathbb{N} \cup \{+\infty\}$  with  $-N^* \leq n_0 \leq N^*$ , such that:*

$$(i) \quad \lambda_t^\pm = \lambda_t^{\pm, \alpha} = \Lambda_{Y_{t-}}^\pm(\delta_t^\pm) \mathbb{1}_{\{\mp N_{t-} < N^*\}} \text{ if } \alpha = (\delta^-, \delta^+).$$

$$(ii) \quad \Lambda_i^\pm \text{ is continuous, decreasing (not necessarily strictly) and } \lim_{\delta \rightarrow +\infty} \delta^{\mathbb{1}_{\{\gamma=0\}}} \Lambda_i^\pm(\delta) = 0 \text{ if } \delta^* = +\infty, \text{ for all } 1 \leq i \leq k.$$

*Notation.* For the sake of readability, we will often write  $\mathbb{1}_t^\pm$  instead of  $\mathbb{1}_{\{\mp N_{t-} < N^*\}}$ .

$N^*$  is a self-imposed constraint on the MM's inventory (whichever its sign) as originally proposed in [GLFT13]. When  $N^* = \infty$  we are in the unconstrained context, while a finite  $N^*$  effectively means that the MM will not buy (resp. sell) whenever  $N_{t-} = N^*$  (resp.  $N_{t-} = -N^*$ ). In practice, the MM could achieve this either by abstaining from quoting or by quoting with an excessively large spread that prevents any transaction (i.e., a 'stub' or 'placeholder quote').<sup>7</sup>

*Remark 1.1.2.3.* The last assumption states in particular that the intensities should decrease to zero when the spreads grow arbitrarily large. Furthermore, when  $\gamma = 0$  we require that they decay faster than  $1/\delta$ . Loosely speaking, this states that for any  $\gamma \geq 0$  the 'expected' utility of the MM's instantaneous margin,  $F_i^\pm(\delta) := \Lambda_i(\delta)U_\gamma(\delta)$ , should vanish for 'stub quotes'. This is a reasonable assumption in practice, as the opposite could lead to unrealistic optimal strategies such as continuously quoting 'infinite spreads'.

<sup>7</sup>For simplicity and without loss of generality, only the first alternative is formalized in our model (as done in [Gué17, GL15, GLFT13]) and this is reflected in the admissible spreads being real-valued. We could also allow for different constraints depending on the sign of the inventory, but we refrain from this to simplify the notation.

*Remark.* Note that the above intensities are predictable, and since any admissible spreads  $\delta^-, \delta^+$  are bounded,  $\lambda^-, \lambda^+$  are in turn bounded. Accordingly,  $N^-, N^+, N$  and  $X$  are non-explosive (see, e.g., [Bré81, p.27 T8]). Furthermore, for any constant  $\lambda^* > 0$  such that  $\lambda^- + \lambda^+ \leq \lambda^*$ , it is easy to see that for all  $p \in \mathbb{N}$ ,  $t \in [0, T]$  and  $r \geq 0$ ,

$$\mathbb{E}^\alpha[|N_t|^p] \leq \mathbb{E}^\alpha[(n_0 + N_t^- + N_t^+)^p] \leq \mathbb{E}^\alpha[(n_0 + M)^p], \quad \mathcal{M}_{|N_t|}^\alpha(r) \leq \mathcal{M}_{n_0 + N_t^- + N_t^+}^\alpha(r) \leq \mathcal{M}_{n_0 + M}^\alpha(r), \quad (1.1.2.5)$$

where  $M \sim \text{Poisson}(\lambda^* T)$  and  $\mathcal{M}_R^\alpha$  is the moment generating function of a random variable  $R$ .<sup>8</sup>

The following are some examples which we shall come to back in Section 1.4.5, when revisiting the standard assumptions in the literature. We remark that exponential, power-law and logistic intensities are the explicit ones most commonly used in the mathematical literature, for tractability reasons.

*Examples 1.1.2.4.* 1.  $\Lambda_i^\pm(\delta) = a_i^\pm e^{-b_i^\pm \delta}$ , with  $a_i^\pm, b_i^\pm > 0$ ,  $-\infty \leq \delta_* < \delta^* \leq +\infty$ .

2.  $\Lambda_i^\pm(\delta) = \frac{a_i^\pm}{1 + c_i^\pm e^{b_i^\pm \delta}}$ , with  $a_i^\pm, b_i^\pm, c_i^\pm > 0$ ,  $-\infty \leq \delta_* < \delta^* \leq +\infty$ .

3.  $\Lambda_i^\pm(\delta) = a_i^\pm (\pi/2 + \arctan(-b_i^\pm \delta + c_i^\pm))$ , with  $a_i^\pm, b_i^\pm > 0$ ,  $c_i^\pm \in \mathbb{R}$ ,  $-\infty \leq \delta_* < \delta^* \leq +\infty$ .

4.  $\Lambda_i^\pm(\delta) = a_i^\pm (b_i^\pm \delta + c_i^\pm)^{-d_i^\pm}$ , with  $a_i^\pm, b_i^\pm, d_i^\pm > 0$ ,  $c_i^\pm \in \mathbb{R}$ ,  $\max_i \{-c_i^\pm / b_i^\pm\} < \delta_* < \delta^* \leq +\infty$ .

To fix ideas, consider Example 1 and let us observe that if  $Y$  represents decreasing levels of competition for the MM, the values of  $a_i^\pm$  (resp.  $b_i^\pm$ ) should increase (resp. decrease) as  $Y$  increases. The same is true if  $Y$  represents increasing levels of positive sentiment (or ‘bullishness’) towards the asset. In practice,  $Y$  will result from the combination of these effects and many more.

Suppose the MM can hedge any remaining inventory at time  $T$  at the reference price  $S_T$  minus a certain execution cost and that she has a terminal CARA utility function

$$U_\gamma(c) = \frac{1 - e^{-\gamma c}}{\gamma}, \quad \text{for } c \in \mathbb{R},$$

when  $\gamma > 0$  and  $U_0 = Id$ . The parameter  $\gamma$  is known as absolute risk aversion.

Neglecting discounting between 0 and  $T$ , we consider the optimization problem faced by the MM who tries to maximize the expected utility of her terminal *penalized profit and loss* (P&L),

$$\sup_{\delta^-, \delta^+ \in \mathcal{A}} \mathbb{E}^{\delta^-, \delta^+} \left[ U_\gamma \left( X_T^{\delta^-, \delta^+} + S_T N_T - \ell(N_T) - \frac{1}{2} \sigma^2 \zeta \int_0^T N_t^2 dt \right) \right], \quad (1.1.2.6)$$

where:

---

<sup>8</sup>By possibly enlarging the space, one can consider a counting process  $Z$  with no common jumps with  $N^-, N^+$  and stochastic intensity  $\lambda^* - \lambda^- - \lambda^+ \geq 0$ . Then the process  $Z + N^- + N^+$  is a Poisson with intensity  $\lambda^*$  that dominates  $N^+ + N^-$ . The claim follows immediately.

- (i)  $\zeta \geq 0$ . When  $\zeta > 0$  we are including in the model a *running penalty*, as firstly done in [CJ15], for the MM to further control her *accumulated inventory risk*. This is the same as subtracting the variance of the mark-to-market value of the inventory, weighted by  $\zeta$ .
- (ii)  $\ell : \mathbb{R} \rightarrow [0, +\infty)$  represents the final execution cost and is increasing on  $[0, +\infty)$ , decreasing on  $(-\infty, 0]$  and  $\ell(0) = 0$ . (And usually convex in practice.)
- (iii) If  $N^* = +\infty$ , we set  $\gamma = \zeta = 0 \equiv \ell$ .

The last restriction on the parameters states that the only case we will consider in which the inventory could be arbitrary large (whichever its sign) is that of a completely risk-neutral MM with negligible costs. The risk averse case is far more challenging and has not been treated in full mathematical detail even in the complete information case (see discussion in [GLFT13]). The model given by (iii) is of secondary interest in practice, yet, it will allow us to further understand the general problem and methodology and to have a more holistic view at almost no extra cost.

*Remark.* Previous works in optimal market making do not consider both penalty and CARA utility (with  $\gamma > 0$ ) in unison, and instead treat the two families of models separately [Gué17]. Besides the obvious interest in unifying the approach and generalizing existing models, allowing for  $\gamma > 0$  and  $\zeta > 0$  simultaneously adds flexibility to the risk managing capabilities of the MM. Indeed, in [Gué17] the author derives some HJB-type systems of ODEs for each problem with a unique risk aversion parameter  $\bar{\gamma}$ , and then relates them by the introduction of an auxiliary parameter  $0 \leq \bar{\xi} \leq \bar{\gamma}$ . The later is afterwards interpreted as measuring risk aversion to non-execution risk only and it is carried forward to the formulae of the ‘optimal strategies’. However, such strategies are not shown to result from any particular optimization problem for  $0 < \bar{\xi} < \bar{\gamma}$ . By looking at the equations of the full information version of our model (Section 1.4) with a single market regime, one can immediately see that the equations in [Gué17] are obtained via the reparametrization  $\gamma = \bar{\xi}$  and  $\zeta = \bar{\gamma} - \bar{\xi}$ . Thus, the formulation in (1.1.2.6) allows us, in particular, to establish the optimality of the strategies in [Gué17] for any value of the non-execution risk aversion and, in general, to differentiate between aversion to different types of risks. We can say that, in our model,  $\gamma$  represents aversion to all types of risks, while the penalty  $\zeta$  is used to further increase the aversion to price risk only.

### 1.1.3 Jump measures

It will be useful in the sequel to understand which are the jump measures and intensity kernels involved in our model. Let  $\mu^N(dt, dz)$  (or simply  $\mu^N$ ) be the jump (or counting) measure of  $N$  (see def. in [Bré81, p.234] or [JS02, p.69]). We use a similar notation for the jump measures of  $Y$  and the pair  $(Y, N)$ . These are all finite random measures since the processes are non-explosive for any  $\delta^+, \delta^- \in \mathcal{A}$  (see (1.1.2.5)), and they admit predictable intensity kernels (see [Bré81, p.235 D2]). If  $m_{z_0}(dz)$  is the Dirac measure at a point  $z_0 \in \mathbb{R}$ , then the  $(\mathbb{P}^\alpha, \mathbb{F})$ -predictable intensity kernels of  $\mu^N, \mu^Y$  and  $\mu^{(Y, N)}$  are respectively given by:

$$\begin{aligned}
 \eta_t^{\alpha, N}(dz) &= \eta^N(Y_{t-}, \delta_t^-, \delta_t^+, N_{t-}, dz) \\
 &= \lambda_t^- m_1(dz) + \lambda_t^+ m_{-1}(dz) \\
 &= \Lambda_{Y_{t-}}^-(\delta_t^-) \mathbb{1}_{\{N_{t-} < N^*\}} m_1(dz) + \Lambda_{Y_{t-}}^+(\delta_t^+) \mathbb{1}_{\{-N_{t-} < N^*\}} m_{-1}(dz), \\
 \eta_t^Y(dh) &= \eta^Y(Y_{t-}, dh) = \sum_{j \neq Y_{t-}} q_t^{Y_{t-}, j} m_{j-Y_{t-}}(dh)
 \end{aligned} \tag{1.1.3.1}$$

and

$$\eta_t^{\alpha, (Y, N)}(dh, dz) = \eta_t^{\alpha, N}(dz) \otimes m_0(dh) + \eta_t^Y(dh) \otimes m_0(dz). \tag{1.1.3.2}$$

We denote by  $\bar{\mu}_\alpha^N(dt, dz)$ ,  $\bar{\mu}_\alpha^Y(dt, dh)$  and  $\bar{\mu}_\alpha^{(Y, N)}(dt, dh, dz)$  the corresponding  $(\mathbb{F}, \mathbb{P}^\alpha)$ -compensated measures.

*Remark.* Note that equation (1.1.3.2) is a consequence of assuming that  $Y, N$  have a.s. no common jumps (which in turn is equivalent to (1.1.2.2)). Indeed, it is easy to see that  $Y, N$  have a.s. no common jumps if and only if

$$\mu^{(Y, N)}(dt, dh, dz) = \mu^N(dt, dz) \otimes m_0(dh) + \mu^Y(dt, dh) \otimes m_0(dz), \tag{1.1.3.3}$$

which implies (1.1.3.2).

### 1.1.4 Construction and characterization of the model

We now prove the existence of a model within our framework (Proposition 1.1.4.1). We construct it by means of a reference probability (an overview of which can be found, e.g., in [Bré81, Chpt.VI]) and, in particular, the Girsanov theorem for point processes.

For the following proposition, let us consider a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$  under the usual conditions, with  $\mathcal{F} = \mathcal{F}_T$  and  $\mathcal{F}_0$  the completed trivial sigma algebra. We refer to  $\mathbb{Q}$  as the *reference probability*. Suppose this space supports a two-dimensional counting process  $(N^-, N^+)$  (see def. in [Bré81, Chapt.II]), a Wiener process  $W$  and a Markov chain  $Y$  with finite state space  $\{1, \dots, k\}$ , initial distribution  $\mu_0$  and time-dependent generator matrix  $Q = (q_t^{ij})_{1 \leq i, j \leq k}$  satisfying Assumptions 1.1.2.1, such that  $N^-, N^+, Y$  verify (1.1.2.2). We set  $N = N^- - N^+$  as before and suppose  $N^\pm$  has intensity  $\mathbb{1}_t^\pm = \mathbb{1}_{\{\mp N_{t-} < N^*\}}$ . ( $N^* \in \mathbb{N} \cup \{+\infty\}$  is given.) Such a simpler model can be constructed for example as a product of canonical spaces, with the existence of the counting processes with the right intensities proved in [JP82, Thm.24 and Cor.31].<sup>9</sup> Let  $\Lambda_i^\pm : (\delta_*, \delta^*) \rightarrow \mathbb{R}$  for  $i = 1, \dots, k$  be functions under the Assumptions 1.1.2.2 ((ii)), for given  $-\infty \leq \delta_* < \delta^* \leq +\infty$  and  $N^* \in \mathbb{N} \cup \{+\infty\}$  with  $-N^* \leq n_0 \leq N^*$ , and define  $\mathcal{A}$  as in (1.1.1.1).

**Proposition 1.1.4.1.** *Let  $\alpha = (\delta^-, \delta^+) \in \mathcal{A}^2$  and define the process  $Z^\alpha$  as the stochastic exponential*

$$Z^\alpha := \mathcal{E} \left( \int_0^\cdot (\Lambda_{Y_{u-}}^-(\delta_u^-) - 1) (dN_u^- - \mathbb{1}_u^- du) + (\Lambda_{Y_{u-}}^+(\delta_u^+) - 1) (dN_u^+ - \mathbb{1}_u^+ du) \right).$$

*Then,*

---

<sup>9</sup>The finite-dimensional result has the same proof as in one dimension, starting from independent Poisson measures.

- (i)  $Z^\alpha$  is a strictly positive uniformly integrable (UI) martingale. In particular,  $\mathbb{E}^\mathbb{Q}[Z_T^\alpha] = 1$ .
- (ii)  $\mathbb{P}^\alpha$  defined by  $\frac{d\mathbb{P}^\alpha}{d\mathbb{Q}} = Z_T^\alpha$  is an equivalent probability measure such that, for  $(\mathbb{F}, \mathbb{P}^\alpha)$ ,  $N^\pm$  is a counting process with intensity  $\lambda_t^\pm = \Lambda_{Y_{t-}}^\pm(\delta_t^\pm) \mathbb{1}_t^\pm$ ,  $W$  is a Wiener process and  $Y$  is a Markov chain with state space  $\{1, \dots, k\}$ , initial distribution  $\mu_0$  and generator matrix  $Q$ . (That is, Assumptions 1.1.2.1, 1.1.2.2 and (1.1.2.2) are all verified.)

*Proof.*  $N^-, N^+$  are finite variation (FV) processes with no common jumps, hence  $[N^-, N^+] = 0$   $\mathbb{Q}$ -a.s. Therefore, the multiplicativity of the stochastic exponential [JS02, p.138] and the exponential formula for FV processes [Bré81, p.337 T4] yield the explicit expression  $Z^\alpha = Z^{\alpha,-} Z^{\alpha,+}$ , with

$$Z_t^{\alpha,\pm} = \exp \left( \int_0^t (1 - \Lambda_{Y_{u-}}^\pm(\delta_u^\pm)) \mathbb{1}_u^\pm du \right) \prod_{\substack{u \leq t: \\ \Delta N_u^\pm \neq 0}} \Lambda_{Y_{u-}}^\pm(\delta_u^\pm). \quad (1.1.4.1)$$

Then (i) follows from the strict positivity and boundedness of  $\Lambda_i^\pm(\delta^\pm)$  for  $1 \leq i \leq k$ , as in [Bré81, p.168 T4] with terminal time  $T$ . The only difference is that in our case the  $\mathbb{Q}$ -intensity of  $N^\pm$  is  $\mathbb{1}_{\{\mp N_{t-} < N^*\}} \leq 1$  instead of 1. The proof remains the same though, simply recalling that the  $\mathbb{Q}$ -moment generating function of  $N_T^\pm$  is dominated by the one of a standard Poisson random variable (see (1.1.2.5)). Uniform integrability is immediate.

(i) guarantees that  $\mathbb{P}^\alpha$ , as defined in (ii), is an equivalent probability measure. The shape of the  $\mathbb{P}^\alpha$ -intensities of  $N^-, N^+$  is due to [Bré81, p.166 T3].

The fact that  $W$  is still a Wiener process under  $\mathbb{P}^\alpha$  is a consequence of the Girsanov–Meyer Theorem [Pro04, p.132 Thm.35] and Levy’s characterization Theorem.

As for  $Y$ , note first that its initial distribution does not change as  $Z_0^\alpha = 1$ , and the same goes for its infinitesimal generator operator. To see this, consider the  $\mathbb{Q}$ -generator operator  $\mathcal{A}_t^Y : \mathbb{R}^E \rightarrow \mathbb{R}^E$ ,  $\mathcal{A}_t^Y f(i) = \sum_{j=1}^k q_t^{ij} f(j)$ . We know the process  $M_t = f(Y_t) - f(Y_0) - \int_0^t \mathcal{A}_u^Y f(Y_u) du$  is a  $\mathbb{Q}$ -local martingale. Once again, by the Girsanov–Meyer Theorem.,  $M$  is also a  $\mathbb{P}^\alpha$ -local martingale ( $[M, Z] = 0$  since  $Y, N^-, N^+$  are FV processes satisfying (1.1.2.2)). Moreover, being bounded,  $M$  is a true  $\mathbb{P}^\alpha$ -martingale and  $Y$  solves, for  $(\mathbb{F}, \mathbb{P}^\alpha)$ , a well-posed martingale problem for  $(\mathcal{A}_t^Y, \mu_0)$ . This implies  $Y$  is a  $\mathbb{P}^\alpha$ -Markov chain with a uniquely determined law [EK09, p.184 Thm.4.2].<sup>10</sup> That  $Q$  is also the  $\mathbb{P}^\alpha$ -generator matrix follows from uniqueness.

(1.1.2.2) clearly remains unchanged under an equivalent change of probability measure.  $\square$

Reciprocally, suppose we start with a family of filtered probability spaces  $\{(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P}^\alpha)\}_{\alpha \in \mathcal{A}^2}$  as in Section 1.1.1, supporting a counting process  $(N^-, N^+)$  and a Markov chain  $Y$  satisfying condition (1.1.2.2), together with a Wiener process  $W$ . Suppose also that  $Y$  has finite state space  $\{1, \dots, k\}$ , initial distribution  $\mu_0$  and time-dependent generator matrix  $Q$ , and that our Assumptions 1.1.2.1 and 1.1.2.2 are in place. We put

<sup>10</sup>Although our martingale problem is non-homogeneous in time,  $(Q_t)$  is deterministic, so this does not represent a problem.

$N = N^- - N^+$  as always. We would like to characterize any such model in terms of a reference probability as in Proposition 1.1.4.1 by an inverse change of measure. However, we only claim the uniqueness of the reference probability on  $\mathbb{F}^N$ . We will come back to this result in the sequel.

**Proposition 1.1.4.2.** *Let  $\alpha = (\delta^-, \delta^+) \in \mathcal{A}^2$  and define*

$$\bar{Z}^\alpha := \mathcal{E} \left( \int_0^\cdot \left( 1/\Lambda_{Y_{u-}}^-(\delta_u^-) - 1 \right) (dN_u^- - \lambda_u^- du) + \left( 1/\Lambda_{Y_{u-}}^+(\delta_u^+) - 1 \right) (dN_u^+ - \lambda_u^+ du) \right).$$

*Then,*

- (i)  $\bar{Z}^\alpha$  is a strictly positive UI martingale. In particular,  $\mathbb{E}^\alpha[\bar{Z}_T] = 1$ .
- (ii)  $\mathbb{Q}^\alpha$  defined by  $\frac{d\mathbb{Q}^\alpha}{d\mathbb{P}^\alpha} = \bar{Z}_T^\alpha$  is a probability measure equivalent to  $\mathbb{P}^\alpha$  such that, for  $(\mathbb{F}, \mathbb{Q}^\alpha)$ ,  $N^\pm$  is a counting process with intensity  $\mathbb{1}_t^\pm$ ,  $W$  is a Wiener process and  $Y$  is a Markov chain with state space  $\{1, \dots, k\}$ , initial distribution  $\mu_0$  and generator matrix  $Q$ . Furthermore,  $Y, N, W$  are independent.
- (iii) If we define  $Z^\alpha$  as in Lemma 1.1.4.1 under  $\mathbb{Q}^\alpha$ , then  $Z^\alpha = 1/\bar{Z}^\alpha$  (i.e., the two changes of measure are inverse of each other).
- (iv) For all  $\tilde{\alpha} \in \mathcal{A}^2$ ,  $\mathbb{Q}^{\tilde{\alpha}} \equiv \mathbb{Q}^\alpha$  on  $\mathbb{F}_T^N$ .

*Proof.* (i) and (ii) are proved just as in Proposition 1.1.4.1 (the processes  $1/\Lambda_t^\pm(\delta^\pm)$  are strictly positive and bounded for  $1 \leq i \leq k$ , and the intensities  $\lambda^\pm$  are bounded). The  $\mathbb{Q}^\alpha$ -independence of  $Y, N, W$  is a consequence of [JYC09, p.543 Lem.9.5.4.1]. That is, the  $(\mathbb{Q}^\alpha, \mathbb{F})$ -martingales  $Y - \int_0^\cdot \sum_{j=1}^k q_u^{Y_{u-}, j} j du$ ,  $N - \int_0^\cdot (\mathbb{1}_u^- - \mathbb{1}_u^+) du$  and  $W$  have the predictable representation property with respect to  $\mathbb{F}^Y$ ,  $\mathbb{F}^N$  and  $\mathbb{F}^W$  resp. (see [Br 81, p.239 T8] for compensated counting measures) and they are  $(\mathbb{F}, \mathbb{Q}^\alpha)$ -orthogonal, which implies their independence.

The explicit expressions for the stochastic exponentials (see (1.1.4.1)) show straightforwardly that  $Z^\alpha \bar{Z}^\alpha = 1$ , proving (iii).

(iv) is due to [Br 81, p.64 T8]. □

## 1.2 Filtering problem

Since the MM cannot directly observe all the information in  $\mathbb{F}$  but only  $\mathbb{F}^{N, W}$  (in particular, she cannot observe  $Y$ ), in order to solve the optimization problem (1.1.2.6) under partial information we want to reduce it first to an equivalent one under full information. Throughout this section we work under  $\mathbb{P}^\alpha$  with  $\alpha = (\delta^-, \delta^+) \in \mathcal{A}^2$  fixed. We sometimes omit  $\alpha$  from the notation for simplicity.

Recall that for any c dl g bounded process  $M$  (not necessarily adapted) on a filtered probability space  $(\Sigma, \mathcal{H}, \mathbb{H}, \mathbb{P})$  satisfying the usual conditions, the *optional projection* of  $M$  on  $\mathbb{H}$  is the unique c dl g process  ${}^oM^{(\mathbb{P}, \mathbb{H})}$  such that  ${}^oM_t^{(\mathbb{P}, \mathbb{H})} = \mathbb{E}^\mathbb{P}[M_t | \mathcal{H}_t]$  a.s. for each  $t$ . Its existence is guaranteed by the Optional Projection Theorem (see, e.g., [JYC09, p.264])

or [Nik06, p.357-358]). (We omit  $\mathbb{P}$  and/or  $\mathbb{H}$  when clear from the context.) Let us consider the optional projections

$$\Pi^{\alpha,i} := {}^o(\mathbb{1}_{\{Y_t=i\}})^{(\mathbb{P}^\alpha, \mathbb{F}^{N,W})}, 1 \leq i \leq k, \quad \text{and} \quad \Pi^\alpha := (\Pi^{\alpha,1}, \dots, \Pi^{\alpha,k}).$$

In other words,  $\Pi^\alpha$  is the unique càdlàg version of the conditional distribution of  $Y$  given the observable information, i.e.,  $\Pi_t^{\alpha,i} = \mathbb{P}^\alpha(Y_t = i | \mathcal{F}_t^{N,W})$ . We now characterize the observable (that is, the  $\mathbb{F}^{N,W}$ -) predictable intensities of  $N^-, N^+$  and  $\mu^N$  in terms of  $\Pi^\alpha$ . Loosely speaking, the observable intensity of, say  $N^+$ , is obtained projecting:  $\mathbb{E}^\alpha[\lambda_t^+ | \mathcal{F}_t^{W,N}]$  [Bré81, p.32 Comment and Pseudo-Proof of T14]. The only technical difficulty is that of finding a predictable version of such process.  ${}^o(\lambda^+)$  has the desired projective property but it is not predictable in general.  ${}^o(\lambda^+)_{t-}$ , on the other hand, is predictable but does not normally enjoy the projective property. In fact, the process we are looking for is ‘in between’ these two.

**Proposition 1.2.1.1 (Observable intensities).** *The  $(\mathbb{P}^\alpha, \mathbb{F}^{N,W})$ -predictable intensities of  $N^\pm$  and  $\mu^N$ , resp., are*

$$\widehat{\lambda}_t^{\pm\alpha} := \mathbb{1}_t^\pm \sum_{i=1}^k \Pi_{u-}^{\alpha,i} \Lambda_i^\pm(\delta_u^\pm) \quad \text{and} \quad \widehat{\eta}_t^{\alpha,N}(dz) := \widehat{\lambda}_t^{-\alpha} m_1(dz) + \widehat{\lambda}_t^{+\alpha} m_{-1}(dz).$$

*Notation.* We set  $\widehat{\Lambda}^\pm(\pi, \delta) = \sum_{i=1}^k \pi^i \Lambda_i^\pm(\delta)$ ; thus,  $\widehat{\lambda}_t^{\pm\alpha} = \mathbb{1}_t^\pm \widehat{\Lambda}^\pm(\Pi_{u-}^\alpha, \delta_u^\pm)$ .

*Proof.* We prove it just for  $N^+$ , the others being analogous, and we omit  $\alpha$  for simplicity. It is clear that  $\widehat{\lambda}^+$  is predictable. We need to check that for any  $\mathbb{F}^{W,N}$ -predictable process  $\psi \geq 0$ ,

$$\mathbb{E} \left[ \int_0^T \psi_t \widehat{\lambda}_t^+ dt \right] = \mathbb{E} \left[ \int_0^T \psi_t dN_t^+ \right].$$

For any  $1 \leq i \leq k$ , each path of the càdlàg processes  $\Pi^i$  and  $\mathbb{1}_{\{Y_{t-}=i\}}$  has only countably many jumps. We can therefore interchange these processes and their left limits when integrating with respect to  $dt$ . By properties of the conditional expectation and Fubini’s Theorem,

$$\begin{aligned} \mathbb{E} \left[ \int_0^T \psi_t \widehat{\lambda}_t^+ dt \right] &= \mathbb{E} \left[ \int_0^T \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \Pi_{t-}^i \Lambda_i^+(\delta_t^+) dt \right] = \mathbb{E} \left[ \int_0^T \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \Pi_t^i \Lambda_i^+(\delta_t^+) dt \right] \\ &= \mathbb{E} \left[ \int_0^T \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \mathbb{E}[\mathbb{1}_{\{Y_t=i\}} | \mathcal{F}_t^{N,W}] \Lambda_i^+(\delta_t^+) dt \right] = \int_0^T \mathbb{E} \left[ \mathbb{E} \left[ \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \Lambda_i^+(\delta_t^+) | \mathcal{F}_t^{N,W} \right] \right] dt \\ &= \int_0^T \mathbb{E} \left[ \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \mathbb{1}_{\{Y_t=i\}} \Lambda_i^+(\delta_t^+) \right] dt = \mathbb{E} \left[ \int_0^T \psi_t \mathbb{1}_t^+ \sum_{i=1}^k \mathbb{1}_{\{Y_{t-}=i\}} \Lambda_i^+(\delta_t^+) dt \right] = \mathbb{E} \left[ \int_0^T \psi_t \lambda_t^+ dt \right] \\ &= \mathbb{E} \left[ \int_0^T \psi_t dN_t^+ \right]. \end{aligned}$$

□

We give now the *filtering (or Kushner-Stratonovich) equations* for the observable distribution of  $Y$ . These are coupled stochastic differential equations (SDEs) governing the dynamics of  $\Pi^\alpha$ .

*Notation.* We denote by  $\Delta \subset \mathbb{R}^k$  the  $(k-1)$ -simplex (i.e.,  $\Delta = \{\pi \in \mathbb{R}^k : 0 \leq \pi^i \leq 1 \text{ for all } i, \text{ and } \sum_i \pi^i = 1\}$  and by  $\Delta^\circ$  its interior relative to the hyperplane  $\{\pi \in \mathbb{R}^k : \sum_i \pi^i = 1\}$  (i.e.,  $\Delta^\circ = \{\pi \in \Delta : 0 < \pi^i < 1 \text{ for all } i\}$ ).

**Proposition 1.2.1.2 (Observable distribution of  $Y$ ).** *The process  $\Pi^\alpha = (\Pi^{\alpha,1}, \dots, \Pi^{\alpha,k})$  is the unique strong solution of the constrained system of SDEs*

$$\begin{aligned} d\Pi_t^i = & \sum_{j=1}^k q_t^{ji} \Pi_t^j dt + \Pi_{t-}^i \left( \frac{\Lambda_i^-(\delta_t^-)}{\sum_{j=1}^k \Pi_{t-}^j \Lambda_j^-(\delta_t^-)} - 1 \right) \left( dN_t^- - \mathbb{1}_t^- \sum_{j=1}^k \Pi_t^j \Lambda_j^-(\delta_t^-) dt \right) \\ & + \Pi_{t-}^i \left( \frac{\Lambda_i^+(\delta_t^+)}{\sum_{j=1}^k \Pi_{t-}^j \Lambda_j^+(\delta_t^+)} - 1 \right) \left( dN_t^+ - \mathbb{1}_t^+ \sum_{j=1}^k \Pi_t^j \Lambda_j^+(\delta_t^+) dt \right), \end{aligned} \quad (1.2.1.1)$$

such that  $\Pi_0 = \mu_0$  and  $\Pi_t \in \Delta$  for all  $t \in [0, T]$  a.s. Equivalently,

$$\begin{aligned} d\Pi_t^i = & \sum_{j=1}^k \left( q_t^{ji} \Pi_t^j + \Pi_t^i \Pi_t^j \left( \mathbb{1}_t^-(\Lambda_j^- - \Lambda_i^-)(\delta_t^-) + \mathbb{1}_t^+(\Lambda_j^+ - \Lambda_i^+)(\delta_t^+) \right) \right) dt \\ & + \Pi_{t-}^i \left( \frac{\Lambda_i^-(\delta_t^-)}{\sum_{j=1}^k \Pi_{t-}^j \Lambda_j^-(\delta_t^-)} - 1 \right) dN_t^- + \Pi_{t-}^i \left( \frac{\Lambda_i^+(\delta_t^+)}{\sum_{j=1}^k \Pi_{t-}^j \Lambda_j^+(\delta_t^+)} - 1 \right) dN_t^+, \end{aligned} \quad (1.2.1.2)$$

with  $\Pi_0 = \mu_0$  and  $\Pi_t \in \Delta$  for all  $t \in [0, T]$  a.s.

*Proof.* The equivalence between the constrained systems of SDEs (1.2.1.1) and (1.2.1.2) results from rearranging the terms and using  $\sum_{j=1}^k \Pi_t^j = 1$ .

Let us check that  $\Pi^\alpha$  solves (1.2.1.2). Clearly, the constraint and the initial condition are satisfied. The verification of the SDEs is due to [CEFS16a, Prop.3.3] (with more details in [CEFS16b, App.A, Lemma A.2 and Prop.3.3]), albeit some considerations need to be made.

On the one hand, the authors work with a pure jump model, with strategies adapted to the natural filtration of the driving jump process only (i.e., there is no diffusion) and constant generator matrix for the Markov chain. However, *mutatis mutandis* the former differences yield no major change in the proofs.

On the other hand, the main assumption of the authors, [CEFS16a, Asm.2.1], postulates the existence of some deterministic measure  $\tilde{\eta}^N(dz)$  on  $\mathbb{R}$  with compact support<sup>11</sup> such that for all  $i \in E$ ,  $\delta^+, \delta^- \in (\delta_*, \delta^*)$ ,  $n \in [-N^*, N^*] \cap \mathbb{Z}$ , the measure  $\eta^N(i, \delta^+, \delta^-, n, dz)$  is equivalent to  $\tilde{\eta}^N(dz)$  and the Radon-Nikodym derivative  $d\eta^N(i, \delta^+, \delta^-, n, \cdot)/d\tilde{\eta}^N$  is uniformly bounded and bounded away from zero  $\tilde{\eta}^N(dz)$ -a.s.

<sup>11</sup>In [CEFS16a] the support is assumed to be a subset of  $(-1, \infty)$ , but this is only for a ‘return (or yield) process’ as in their case.



Since the spread processes are fixed and bounded, we can assume without loss of generality that  $\delta_*$  is finite. Setting  $\tilde{\eta}^N(dz) := m_1(dz) + m_{-1}(dz)$  we see straightforwardly that the two measures are equivalent with derivative

$$\frac{d\eta^N(i, \delta^-, \delta^+, n, \cdot)}{d\tilde{\eta}^N}(z) = \Lambda_i^-(\delta^-) \mathbb{1}_{\{n < N^*\}} \mathbb{1}_{\{z=1\}} + \Lambda_i^+(\delta^+) \mathbb{1}_{\{-n < N^*\}} \mathbb{1}_{\{z=-1\}}, \quad (1.2.1.3)$$

uniformly bounded by  $\Lambda_i^-(\delta_*) + \Lambda_i^+(\delta_*)$ . However, our model allows for  $d\eta^N/d\tilde{\eta}^N(z) = 0$ , which is a consequence of having vanishing intensities  $\lambda^\pm$ . This poses no issue nonetheless, as  $d\eta^N/d\tilde{\eta}^N(z) > 0$  is only used in [CEFS16a, Prop.3.3] to guarantee  $\bar{Z}^\alpha > 0$  (see Proposition 1.1.4.2). This condition, also satisfied in our model,<sup>12</sup> allows to go from the physical probability  $\mathbb{P}^\alpha$  to a reference probability  $\mathbb{Q}^\alpha$  and backwards.

We turn now to the proof of uniqueness. We remark first that the jump height coefficients in (1.2.1.2) will typically not be Lipschitz (classical results for SDEs such as [Pro04, p.253 Thm.7] cannot be applied) and the paths of the spreads need not be continuous between the jump times of  $N$  (ruling out the most classical results of ODEs [CL55]). Nevertheless, we can still follow a pathwise ODEs approach. Let us fix a path and alternatively verify uniqueness inductively on the intervals  $[\tau_m, \tau_{m+1})$ , where  $\tau_0 := 0$  and  $0 < \tau_1 < \tau_2 < \dots < \tau_M = T$  are the jump times of  $N$  (including the terminal time  $T$  even if there is no jump at that point). Then  $A^{ji}$  is bounded for all  $1 \leq i, j \leq k$ . Now observe that any càdlàg process  $\tilde{\Pi}$ , solving the constrained system of SDEs (1.2.1.2), must solve pathwise for  $m = 0, 1, \dots, M-1$  the following system of ODEs in integral form, for  $t \in [\tau_m, \tau_{m+1})$ :

$$\tilde{\Pi}_t^i = R_m^i(\tilde{\Pi}_{\tau_m^-}) + \int_{\tau_m}^t \sum_{j=1}^k \left( q_u^{ji} \tilde{\Pi}_u^j + \tilde{\Pi}_u^i \tilde{\Pi}_u^j A_u^{ji} \right) du, \quad (1.2.1.4)$$

with  $R_m^i(\tilde{\Pi}_{\tau_m^-}) := \tilde{\Pi}_{\tau_m^-}^i \Lambda_i^\pm(\delta_{\tau_m}^\pm) / \sum_{j=1}^k \tilde{\Pi}_{\tau_m^-}^j \Lambda_j^\pm(\delta_{\tau_m}^\pm)$  if  $\Delta N_{\tau_m} = \mp 1$ ,  $m > 0$ , and  $R_0^i(\tilde{\Pi}_{\tau_0^-}) = \mu_0^i$ . Elementary algebra of bounded Lipschitz functions shows that  $f^i : [\tau_m, \tau_{m+1}) \times \Delta \rightarrow \mathbb{R}$ , defined by  $f^i(u, \pi) = \sum_{j=1}^k (q_u^{ji} \pi^j + \pi^i \pi^j A_u^{ji})$  is Lipschitz in  $\pi$  (uniformly in  $u$ ). Let  $K$  be the maximum Lipschitz constant of  $f^l$  for  $1 \leq l \leq k$  and suppose  $\Pi_{\tau_m^-}^\alpha = \tilde{\Pi}_{\tau_m^-}$  (clearly satisfied for  $m = 0$ ). Then (1.2.1.4) yields  $\|\Pi_t^\alpha - \tilde{\Pi}_t\| \leq K \int_{\tau_m}^t \|\Pi_u^\alpha - \tilde{\Pi}_u\| du$ , implying  $\Pi_t^\alpha = \tilde{\Pi}_t$  on  $[\tau_m, \tau_{m+1})$  by Grönwall's inequality. As a consequence, the equality on  $[0, T)$  follows by induction. It must clearly hold at time  $T$  as well, either by continuity or (if there is a jump) because  $\Pi_T^\alpha = R_M(\Pi_{\tau_M^-}^\alpha) = R_M(\tilde{\Pi}_{\tau_M^-}) = \tilde{\Pi}_T$ .  $\square$

*Remark 1.2.1.3.* Consider the identification  $\Delta \simeq [0, 1]^{k-1}$  (resp.  $\Delta^\circ \simeq (0, 1)^{k-1}$ ) obtained by the substitution  $\pi^k = 1 - \sum_{j < k} \pi^j$  (where the choice of the  $k$ -th coordinate over the rest is completely arbitrary). Then the constrained system of SDEs (1.2.1.1) (or equivalently, (1.2.1.2)) for  $(\Pi^{\alpha,1}, \dots, \Pi^{\alpha,k})$  becomes an ‘unconstrained’ system for  $(\Pi^{\alpha,1}, \dots, \Pi^{\alpha,k-1}) \in [0, 1]^{k-1}$ . Henceforth, we shall use this identification whenever convenient.

We finish this section with a short lemma. It states that the conditional distribution  $\Pi^\alpha$  can never reach the relative border of the simplex  $\Delta$ , provided it starts from the relative interior. This amounts to saying that all regimes have some positive probability at time zero.

<sup>12</sup>This was ultimately a consequence of the decomposition in Assumptions 1.1.2.2 ((i)), that allowed for the vanishing factors of the intensities to be secluded as the reference probability intensities.

**Lemma 1.2.1.4.** *If  $\mu_0 \in \Delta^\circ$ , then  $\Pi_t^\alpha \in \Delta^\circ$  for all  $0 \leq t \leq T$  a.s.*

*Proof.* We want to show that  $\Pi_t^{\alpha,i} > 0$  for all  $1 \leq i \leq k$ ,  $0 \leq t \leq T$  a.s. We proceed by induction on the jump times of  $N$ , for each path, as at the end of the proof of Proposition 1.2.1.2. Using the same notations, let  $1 \leq i \leq k$  and suppose  $\Pi_{\tau_m^-}^{\alpha,j} > 0$  for all  $1 \leq j \leq k$  (satisfied for  $m = 0$  by assumption). Then (1.2.1.4), Assumptions 1.1.2.1 and the fact that  $A^{ii} \equiv 0$  by definition, show that  $\Pi^{\alpha,i}$  is absolutely continuous on the interval  $[\tau_m, \tau_{m+1})$  and satisfies

$$\begin{aligned} (\Pi_t^{\alpha,i})' &= \sum_{j=1}^k \left( q_t^{ji} \Pi_t^{\alpha,j} + \Pi_t^{\alpha,i} \Pi_t^{\alpha,j} A_t^{ji} \right) = \Pi_t^{\alpha,i} \left( q_t^{ii} + \sum_{j \neq i} \Pi_t^{\alpha,j} A_t^{ji} \right) + \sum_{j \neq i} q_t^{ji} \Pi_t^{\alpha,j} \\ &\geq \Pi_t^{\alpha,i} \left( q_t^{ii} + \sum_{j \neq i} \Pi_t^{\alpha,j} A_t^{ji} \right), \end{aligned} \quad (1.2.1.5)$$

for  $dt$ -a.e.  $t \in [\tau_m, \tau_{m+1})$ , subject to  $\Pi_{\tau_m}^{\alpha,i} = R_m^i(\Pi_{\tau_m^-}^\alpha) > 0$ . Let us set  $s := \sup([\tau_m, \tau_{m+1}) \cap \{t \in [0, T] : \Pi_t^{\alpha,i} > 0\})$ . We need to prove that  $s = \tau_{m+1}$ . By the continuity of  $\Pi^{\alpha,i}$ , it must be  $s > \tau_m$ . Consequently, (1.2.1.5) and the absolute continuity of  $\log \Pi^{\alpha,i}$  on  $[\tau_m, t] \subset [\tau_m, s)$  yield

$$\Pi_t^{\alpha,i} \geq R_m^i(\Pi_{\tau_m^-}^\alpha) \exp \left( \int_{\tau_m}^t (q_u^{ii} + \sum_{j \neq i} \Pi_u^{\alpha,j} A_u^{ji}) du \right) \quad \text{for } dt\text{-a.e. } t \in [\tau_m, s).$$

If it were  $s < \tau_{m+1}$ , the continuity of  $\Pi^{\alpha,i}$  again and the former inequality would imply  $0 = \Pi_s^{\alpha,i} > 0$ , proving by contradiction that  $s = \tau_{m+1}$  and  $\Pi^{\alpha,i}$  is positive on the whole interval  $[\tau_m, \tau_{m+1})$ . Positivity on  $[0, T)$  now follows by induction, and it must clearly hold at time  $T$  as well, as either there is a jump or we can reason as we just did.  $\square$

In light of the previous lemma, we will assume from here onwards that

$$\mu_0 \in \Delta^\circ, \quad (1.2.1.6)$$

and therefore work with  $\Delta^\circ$  instead of  $\Delta$ .

### 1.3 Value function and HJB equation

In this section, we tackle the control problem of the MM with the filter as an additional state variable. We define the MM's value function and we aim to characterize it by means of an HJB equation.

Let  $t \in [0, T]$ . We consider our model 'starting at  $t$ ' instead of 0. Whenever a process is defined from time ' $t^-$  onwards' (i.e., from time  $t$  onwards and decreeing its left-limit value at  $t$ ) this implicitly means it is constant on  $[0, t)$ . In particular, we use this convention for all integrals (stochastic or not) of the form  $\int_t^\cdot$  and (with a slight abuse of notation) for the processes  $W - W_t$  and  $N^\pm - N_t^\pm$ . We work with  $t^-$  instead of  $t$  due to the jumps of the processes  $N^-, N^+$ . However, since  $N^-, N^+$  are quasi-left continuous,<sup>13</sup> for most intended

<sup>13</sup>For example, because they are increasing càdlàg processes admitting continuous compensators for (any one of) the physical probabilities [JS02, p.70 Prop.1.19 or p.77 Prop.2.9].

purposes one can drop the left limit with no harm. We define  $\mathbb{F}^{t,W,N} = (\sigma(W_r - W_t, N_r - N_t : t \leq r \leq u \vee t) \vee \mathcal{F}_0)_u$  and  $\mathbb{F}^{t,N}$  analogously.

Let  $s, x \in \mathbb{R}$ ,  $n \in \mathbb{Z} \cap [-N^*, N^*]$  and  $\pi \in \Delta^\circ \subset \mathbb{R}^k$ . The set  $\mathcal{A}_t$  of admissible spreads starting at  $t$  is the set of  $\delta \in \mathcal{A}$  which are independent of  $\mathcal{F}_{t-}^{N,W}$  (equivalently, the  $\delta \in \mathcal{A}$  which are  $\mathbb{F}^{t,W,N}$ -predictable). Consider for each  $\alpha = (\delta^-, \delta^+) \in \mathcal{A}_t^2$  the processes  $S^{t,s}, X^{\alpha,t,x,s}, N^{t,n}, \Pi^{\alpha,t,n,\pi}$  defined pathwise, outside some set  $A \in \mathcal{F}_0$ , by (1.1.2.1), (1.1.2.4), (1.1.2.3), (1.2.1.1) resp., replacing the initial conditions  $s_0, x_0, n_0, \mu_0$  at time 0 by  $s, x, n, \pi$  resp. at time  $t^-$ . We remark that  $\mathbb{F}^{t,W,N} = \mathbb{F}^{t,W,N^{t,n}}$  (since  $N^{t,n} = n + N^- - N^+ - (N_{t-}^- - N_{t-}^+)$ ) and all the processes defined in this section are adapted to this filtration. We further assume there exists a family of ‘physical’ probabilities  $(\mathbb{P}^{\alpha,t,n,\pi})_{\alpha \in \mathcal{A}_t^2}$  such that their null sets generate  $\mathcal{F}_0$ , and for  $(\mathbb{P}^{\alpha,t,n,\pi}, \mathbb{F}^{t,W,N})$  it holds that  $W - W_t$  is a Wiener process and  $N^\pm - N_{t-}^\pm$  has predictable intensity  $\widehat{\lambda}^{\pm, \alpha, t, n, \pi}$ , as defined in Proposition 1.2.1.1 in terms of  $N^{t,n}$  and  $\Pi^{\alpha,t,n,\pi}$ .

We define the penalized P&L from  $t$  to  $T$  (see (1.1.2.6) for parameter restrictions) as

$$P_{t,T}^{\alpha,s,x,n} := X_T^{\alpha,t,x,s} + S_T^{t,s} N_T^{t,n} - \ell(N_T^{t,n}) - \frac{1}{2} \sigma^2 \zeta \int_t^T (N_u^{t,n})^2 du, \quad (1.3.1.1)$$

and the *value function* of problem (1.1.2.6) as

$$V(t, s, x, n, \pi) := \sup_{\alpha \in \mathcal{A}_t^2} \mathbb{E}^{\alpha,t,n,\pi} \left[ U_\gamma \left( P_{t,T}^{\alpha,s,x,n} \right) \right]. \quad (1.3.1.2)$$

Our goal is to compute optimal or ‘close to optimal’ strategies.<sup>14</sup> The Dynamic Programming Principle and Ito’s Lemma allow us to formally derive (see, e.g., [Bou07]) the *Hamilton–Jacobi–Bellman* (or *dynamic programming*) partial-integro differential equation associated to  $V$ :

$$\begin{aligned} 0 = & v_t + \mu v_s + \frac{1}{2} \sigma^2 v_{ss} + \sum_{i,j=1}^k q_t^{ji} \pi^j v_{\pi^i} + \frac{1}{2} \sigma^2 \zeta n^2 (\gamma v - 1) \\ & + \mathbb{1}_{\{n < N^*\}} \sup_{\delta^- \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^- - \Lambda_i^-) (\delta^-) \pi^j \pi^i v_{\pi^i} + d_{\delta^-}^-(v) \sum_{i=1}^k \pi^i \Lambda_i^-(\delta^-) \right\} \\ & + \mathbb{1}_{\{-n < N^*\}} \sup_{\delta^+ \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^+ - \Lambda_i^+) (\delta^+) \pi^j \pi^i v_{\pi^i} + d_{\delta^+}^+(v) \sum_{i=1}^k \pi^i \Lambda_i^+(\delta^+) \right\}, \end{aligned} \quad (1.3.1.3)$$

with terminal condition  $v(T, s, x, n, \pi) = U_\gamma(x + sn - \ell(n))$ , where:

$$\begin{aligned} & d_\delta^\pm(v)(t, s, x, n, \pi) \\ & = v \left( t, s, x \pm (s \pm \delta), n \mp 1, \frac{1}{\sum_{j=1}^k \pi^j \Lambda_j^\pm(\delta)} (\pi^1 \Lambda_1^\pm(\delta), \dots, \pi^k \Lambda_k^\pm(\delta)) \right) - v(t, s, x, n, \pi), \end{aligned}$$

and we convene the following:

---

<sup>14</sup>By ‘close to optimal’ we mean that for each  $\varepsilon > 0$  there exists a strategy such that the supremum in (1.3.1.2) is attained up to  $\varepsilon$ .

*Notation.* (i) The derivatives with respect to  $\pi = (\pi^1, \dots, \pi^k)$  should be understood via the identification of Remark 1.2.1.3. (ii) Although it is not meaningful to evaluate  $v$  on the inventories  $\pm N^* \pm 1$ , this only happens in equation (1.3.1.3) when the corresponding term vanishes. This slight abuse of notation can be found throughout previous works and we will be using it as well.

Equation (1.3.1.3) can also be seen as a coupled system of PIDEs indexed in  $n \in \mathbb{Z} \cap [-N^*, N^*]$ . (We will talk about system of equations or simply ‘equation’ indistinctly). Nonlinearity aside, (1.3.1.3) is rather complex, in particular due to being of second order, high-dimensional and with derivatives in almost all of these dimensions. Tackling it directly (either analytically or numerically) is utterly challenging. Consequently, it has become common practice for optimal market making and optimal liquidation models *à la* Avellaneda–Stoikov [AS08] to propose an ansatz for the solution [AS08, BL14, CDJ17, CJ15, CJR14, FL12, FL13, Gué17, GL15, GLFT13]. This approach however, relies heavily on the existence of a classical solution for the resulting simplified equation, so that the ansatz is ultimately proved valid by a suitable verification theorem. (See Section 1.4 for more details.) When the simplified equation does not admit (or cannot be guaranteed to admit) a classical solution, and a viscosity approach needs to be used instead, the previous argument breaks down.

If we attempted to solve our problem by the standard approach, a plausible ansatz for the value function could be

$$V(t, s, x, n, \pi) = U_\gamma(x + sn + \Theta(t, n, \pi)). \quad (1.3.1.4)$$

Formal substitution yields the following equation for  $\Theta$ :

$$\begin{aligned} 0 = & \theta_t + \mu n - \frac{1}{2} \sigma^2 n^2 (\gamma + \zeta) + \sum_{i,j=1}^k q_t^{ji} \pi^j \theta_{\pi^i} \\ & + \mathbb{1}_{\{n < N^*\}} \sup_{\delta^- \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^- - \Lambda_i^-)(\delta^-) \pi^j \pi^i \theta_{\pi^i} + U_\gamma(\delta^- + d_{\delta^-}(\theta)) \sum_{i=1}^k \pi^i \Lambda_i^-(\delta^-) \right\} \\ & + \mathbb{1}_{\{-n < N^*\}} \sup_{\delta^+ \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^+ - \Lambda_i^+)(\delta^+) \pi^j \pi^i \theta_{\pi^i} + U_\gamma(\delta^+ + d_{\delta^+}(\theta)) \sum_{i=1}^k \pi^i \Lambda_i^+(\delta^+) \right\}, \end{aligned} \quad (1.3.1.5)$$

with terminal condition  $\theta(T, n, \pi) = -\ell(n)$ , where:

$$d_\delta^\pm(\theta)(t, n, \pi) = \theta\left(t, n \mp 1, \frac{1}{\sum_{j=1}^k \pi^j \Lambda_j^\pm(\delta)} (\pi^1 \Lambda_1^\pm(\delta), \dots, \pi^k \Lambda_k^\pm(\delta))\right) - \theta(t, n, \pi).$$

The new system of PIDEs is of first order and no longer depends on the variables  $s$  and  $x$  (there is no diffusion anymore). This is a considerable simplification; one that will permit effective numerical solution in Section 1.5. But it is not good enough for us to assert existence of a classical solution. Notwithstanding, we are able to rigorously prove the decomposition (1.3.1.4) and explicitly find  $\Theta$  as a new ‘value function’ (Theorem 1.3.1.1). When the control space is compact, this ultimately allows us to characterize  $\Theta$  as the unique solution of the

terminal condition PIDE (1.3.1.5) in the viscosity sense (Theorem 1.3.1.3), further simplified in the unconstrained inventory case. These two theorems constitute the main theoretical results of this chapter. They allow us to safely postulate reasonable candidates for optimal (or  $\epsilon$ -optimal) strategies for the MM, i.e., those given by spreads that (at least approximately) realize the suprema in (1.3.1.5).

**Theorem 1.3.1.1.** *There exists a unique function  $\Theta : [0, T] \times (\mathbb{Z} \cap [-N^*, N^*]) \times \Delta^\circ \rightarrow \mathbb{R}$  such that the decomposition (1.3.1.4) holds true. Furthermore, there exists a family of equivalent probability measures  $\tilde{\mathbb{P}}^{\alpha, t, n, \pi} \sim \mathbb{P}^{\alpha, t, n, \pi}$ ,  $\alpha = (\delta^-, \delta^+) \in \mathcal{A}_t^2$ , such that*

(i)  $\Pi^{\alpha, t, n, \pi}$  is the unique strong solution of (1.2.1.1) with initial condition  $(t^-, \pi)$  under  $\tilde{\mathbb{P}}^{\alpha, t, n, \pi}$ .

(ii)  $\mu^{N^{t, n}}$  has  $(\tilde{\mathbb{P}}^{\alpha, t, n, \pi}, \mathbb{F}^{t, W, N})$ -predictable intensity kernel

$$\tilde{\eta}_u^N(dz) := e^{-\gamma \delta_u^-} \widehat{\lambda}_u^{\alpha, t, n, \pi} m_1(dz) + e^{-\gamma \delta_u^+} \widehat{\lambda}_u^{\alpha, t, n, \pi} m_{-1}(dz).$$

(iii)  $\Theta = U_\gamma^{-1} \circ \Psi = -\frac{1}{\gamma} \log(1 - \gamma \Psi)$  with

$$\Psi(t, n, \pi) := \sup_{\alpha \in \mathcal{A}_t^2} \tilde{\mathbb{E}}^{\alpha, t, n, \pi} \left[ U_\gamma \left( \tilde{P}_{t, T}^{\alpha, n, \pi} \right) \right], \quad (1.3.1.6)$$

where  $\tilde{P}_{t, T}^{\alpha, n, \pi} := \int_t^T \left\{ U_\gamma(\delta_u^-) \widehat{\lambda}_u^{\alpha, t, n, \pi} + U_\gamma(\delta_u^+) \widehat{\lambda}_u^{\alpha, t, n, \pi} + \mu N_u^{t, n} - \frac{1}{2} \sigma^2 (\gamma + \zeta) (N_u^{t, n})^2 \right\} du - \ell(N_T^{t, n})$  and  $\tilde{\mathcal{A}}_t := \{\delta \in \mathcal{A} : \delta \text{ is } \mathbb{F}^{t, N^{t, n}}\text{-predictable}\}$ .

*Proof.* For shortness, we make an abuse of notation and omit  $(t, n, \pi)$  from the probability measures and expectations. Let us start by proving (1.3.1.4) and finding  $(\tilde{\mathbb{P}}^\alpha)_{\alpha \in \mathcal{A}_t^2}$  with the desired properties. Using integration by parts we can re-write the penalized P&L (1.3.1.1) as

$$\begin{aligned} P_{t, T}^{\alpha, s, x, n} &= x + sn + \int_t^T \left\{ \mu N_u^{t, n} - \frac{1}{2} \sigma^2 (\gamma + \zeta) (N_u^{t, n})^2 \right\} du - \ell(N_T^{t, n}) \\ &\quad + \sigma \int_t^T N_u^{t, n} dW_u + \frac{1}{2} \sigma^2 \gamma \int_t^T (N_u^{t, n})^2 du + \int_t^T \delta_u^- dN_u^- + \delta_u^+ dN_u^+ \\ &=: x + sn + \bar{P}_{t, T}^{\alpha, n}, \end{aligned} \quad (1.3.1.7)$$

Consider first the case  $\gamma = 0$ . The integrals with respect to  $W, N^-, N^+$  all have bounded integrands (and predictable for  $N^-, N^+$ ), except in the case of unconstrained inventory:  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$  (see (1.1.2.6)). Regardless, we still have  $\mathbb{E}^\alpha \left[ \int_t^T (N_u^{t, n})^2 du \right] = \int_t^T \mathbb{E}^\alpha [(N_u^{t, n})^2] du < +\infty$  by (1.1.2.5). Choosing  $\tilde{\mathbb{P}}^\alpha := \mathbb{P}^\alpha$  the conclusion follows by taking expectation and by Propositions 1.2.1.1 and 1.2.1.2.

Consider now  $\gamma > 0$ . Hence, we are in the case  $|N| \leq N^* < +\infty$ . We define

$$A^\alpha := \mathcal{E} \left( -\gamma \sigma \int_t^\cdot N_u^{t, n} dW_u \right) = \exp \left( -\gamma \sigma \int_t^\cdot N_u^{t, n} dW_u - \frac{1}{2} \sigma^2 \gamma^2 \int_t^\cdot (N_u^{t, n})^2 du \right).$$

By Novikov's condition,  $A^\alpha$  is a strictly positive UI martingale with  $\mathbb{E}^\alpha[A_T] = 1$ , and therefore defines an equivalent probability measure  $\mathbb{A}^\alpha \sim \mathbb{P}^\alpha$  via  $\frac{d\mathbb{A}^\alpha}{d\mathbb{P}^\alpha} = A_T^\alpha$ . Note that the Girsanov–Meyer Theorem. [Pro04, p.132 Thm.35] ensures the  $\mathbb{F}^{t,W,N}$ -intensities of  $N^\pm - N_t^\pm$  remain the same when changing to  $\mathbb{A}^\alpha$ . Let us set

$$\begin{aligned} B^\alpha &:= \mathcal{E} \left( -\gamma \int_t^\cdot U_\gamma(\delta_u^-) d\overline{N}_u^{-\alpha} + U_\gamma(\delta_u^+) d\overline{N}_u^{+\alpha} \right) \\ &= \mathcal{E} \left( \int_t^\cdot (e^{-\gamma\delta_u^-} - 1) d\overline{N}_u^{-\alpha} + (e^{-\gamma\delta_u^+} - 1) d\overline{N}_u^{+\alpha} \right), \end{aligned}$$

where  $\overline{N}_u^{\pm\alpha}$  denote the corresponding  $(\mathbb{P}^\alpha, \mathbb{F}^{t,W,N})$ -compensated (or equivalently,  $(\mathbb{A}^\alpha, \mathbb{F}^{t,W,N})$ -compensated) processes. By the same arguments of Propositions 1.1.4.1 and 1.1.4.2,  $B^\alpha$  is a strictly positive UI martingale with  $\mathbb{E}^{\mathbb{A}^\alpha}[B_T^\alpha] = 1$  and defines an equivalent probability measure  $\tilde{\mathbb{P}}^\alpha \sim \mathbb{A}^\alpha \sim \mathbb{P}^\alpha$  via  $\frac{d\tilde{\mathbb{P}}^\alpha}{d\mathbb{A}^\alpha} = B_T^\alpha$ , such that (ii) holds true. Note that (i) is also trivially verified due to the equivalence of the probability measures.

Suppose for the time being that  $\Psi$  is defined as in (1.3.1.6) but taking supremum over the whole set of admissible controls  $\mathcal{A}_t^2 \supseteq \tilde{\mathcal{A}}_t^2$  instead. We will see afterwards that this makes no difference. To see (1.3.1.4), observe that the identity  $U_\gamma(a+b) = U_\gamma(b)e^{-\gamma a} + U_\gamma(a)$  and (1.3.1.7) yield  $U_\gamma(P_{t,T}^{\alpha,s,x,n}) = U_\gamma(\overline{P}_{t,T}^{\alpha,n})e^{-\gamma(x+sn)} + U_\gamma(x+sn)$ , giving

$$V(t, s, x, n, \pi) = \sup_{\alpha \in \mathcal{A}_t^2} \mathbb{E}^\alpha \left[ U_\gamma(\overline{P}_{t,T}^{\alpha,n,\pi}) \right] e^{-\gamma(x+sn)} + U_\gamma(x+sn).$$

On the other hand, by the same identity,

$$U_\gamma(x+sn+\Theta) = (U_\gamma \circ \Theta)e^{-\gamma(x+sn)} + U_\gamma(x+sn) = \Psi e^{-\gamma(x+sn)} + U_\gamma(x+sn).$$

As a consequence, (1.3.1.4) is equivalent to the equality  $\Psi(t, n, \pi) = \sup_{\alpha \in \mathcal{A}_t^2} \mathbb{E}^\alpha \left[ U_\gamma(\overline{P}_{t,T}^{\alpha,n}) \right]$ . We check instead the stronger statement

$$\mathbb{E}^\alpha \left[ \exp \left( -\gamma \overline{P}_{t,T}^{\alpha,n} \right) \right] = \tilde{\mathbb{E}}^\alpha \left[ \exp \left( -\gamma \tilde{P}_{t,T}^{\alpha,n} \right) \right], \text{ for all } \alpha \in \mathcal{A}_t^2. \quad (1.3.1.8)$$

Using the explicit exponential formula (see equation (1.1.4.1)) and by straightforward computations:

$$\begin{aligned} \exp \left( -\gamma \overline{P}_{t,T}^{\alpha,n} \right) &= \exp \left( -\gamma \left( \int_t^T \left\{ \mu N_u^{t,n} - \frac{1}{2} \sigma^2 (\gamma + \zeta) (N_u^{t,n})^2 \right\} du - \ell(N_T^{t,n}) \right) \right) \\ &\quad \times A_T^\alpha \exp \left( -\gamma \int_t^T \delta_u^- dN_u^- + \delta_u^+ dN_u^+ \right) \\ &= \exp \left( -\gamma \tilde{P}_{t,T}^{\alpha,n} \right) A_T^\alpha \exp \left( \gamma \int_t^T \left\{ U_\gamma(\delta_u^-) \widehat{\lambda}_u^{-\alpha,t,n,\pi} + U_\gamma(\delta_u^+) \widehat{\lambda}_u^{+\alpha,t,n,\pi} \right\} du \right) \\ &\quad \times \prod_{\substack{t \leq u \leq T: \\ \Delta N_u^- \neq 0}} \exp(-\gamma \delta_u^-) \prod_{\substack{t \leq u \leq T: \\ \Delta N_u^+ \neq 0}} \exp(-\gamma \delta_u^+) \\ &= \exp \left( -\gamma \tilde{P}_{t,T}^{\alpha,n} \right) A_T^\alpha B_T^\alpha, \end{aligned}$$

which yields (1.3.1.8) after taking  $\mathbb{P}^\alpha$ -expectation.

It remains to see that

$$\Psi(t, n, \pi) := \sup_{\alpha \in \mathcal{A}_t^2} \tilde{\mathbb{E}}^\alpha \left[ U_\gamma \left( \tilde{P}_{t,T}^{\alpha, n, \pi} \right) \right] = \sup_{\alpha \in \tilde{\mathcal{A}}_t^2} \tilde{\mathbb{E}}^\alpha \left[ U_\gamma \left( \tilde{P}_{t,T}^{\alpha, n, \pi} \right) \right] =: \tilde{\Psi}(t, n, \pi).$$

Clearly  $\Psi \geq \tilde{\Psi}$ . Let us check  $\Psi \leq \tilde{\Psi}$ . As done in Proposition 1.1.4.2 we can define a family of ‘reference’ equivalent probability measures  $\tilde{\mathbb{Q}}^\alpha \sim \tilde{\mathbb{P}}^\alpha$ ,  $\alpha \in \mathcal{A}_t^2$ , such that for  $(\tilde{\mathbb{Q}}^\alpha, \mathbb{F}^{t, W, N})$  it holds:  $W - W_t$  is a Wiener process independent of the counting process  $(N^- - N_{t-}^-, N^+ - N_{t-}^+)$  and  $N^\pm - N_{t-}^\pm$  has predictable intensity  $\mathbb{1}_u^{\pm, t, n} := \mathbb{1}_{\{\mp N_{u-}^{t, n} < N^*\}}$  (in particular, its law does not depend on  $\alpha$ ). Furthermore, the inverse change of measure is given by  $d\tilde{\mathbb{P}}^\alpha/d\tilde{\mathbb{Q}}^\alpha = \tilde{Z}_T^\alpha$  with

$$\begin{aligned} \tilde{Z}^\alpha := \\ \mathcal{E} \left( \int_t^\cdot \left( \widehat{\Lambda}^-(\Pi_{u-}^{\alpha, t, n, \pi}, \delta_u^-) - 1 \right) (dN_u^- - \mathbb{1}_u^{-, t, n} du) + \left( \widehat{\Lambda}^+(\Pi_{u-}^{\alpha, t, n, \pi}, \delta_u^+) - 1 \right) (dN_u^+ - \mathbb{1}_u^{+, t, n} du) \right). \end{aligned}$$

Let us fix  $\alpha \in \mathcal{A}_t^2$ . Denote by  $\mathcal{D} = \mathcal{D}([t, T], \mathbb{R})$  the Skorokhod space of càdlàg functions with its usual sigma algebra and by  $\mathbb{P}^W, \mathbb{P}^N$  the laws (or pushforward measures) induced on  $\mathcal{D}$  by  $W - W_t, N^{t, n}$  resp. when starting from  $\mathbb{Q}^\alpha$ . These laws do not depend on  $\alpha$  and characterize the joint law of  $(W - W_t, N^{t, n})$  on  $\mathcal{D}^2$  as  $\mathbb{P}^W \otimes \mathbb{P}^N$ , due to the independence of the two processes. Since  $\alpha$  is  $\mathbb{F}^{t, W, N}$ -predictable, by a monotone class argument one can show there exists a jointly measurable process  $f : [t, T] \times \mathcal{D}^2 \rightarrow \mathbb{R}$  such that  $\alpha_u = f_u(W - W_t, N^{t, n})$  and for  $\mathbb{P}^W$ -almost every  $w \in \mathcal{D}$ , the process  $\tilde{\alpha}_u := f_u(w, N^{t, n})$  is in  $\tilde{\mathcal{A}}_t^2$ . Note also that we can write  $\tilde{Z}_T^\alpha U_\gamma \left( \tilde{P}_{t,T}^{\alpha, n, \pi} \right) = g(\alpha, N^{t, n}) = g(f(W - W_t, N^{t, n}), N^{t, n})$  for some function  $g$ . By Fubini’s theorem,

$$\begin{aligned} \tilde{\mathbb{E}}^\alpha \left[ U_\gamma \left( \tilde{P}_{t,T}^{\alpha, n, \pi} \right) \right] &= \tilde{\mathbb{E}}^{\mathbb{Q}^\alpha} \left[ \tilde{Z}_T^\alpha U_\gamma \left( \tilde{P}_{t,T}^{\alpha, n, \pi} \right) \right] = \int \mathbb{E}^{\mathbb{P}^N} [g(f(w, \cdot), \cdot)] d\mathbb{P}^W(w) \\ &\leq \int \tilde{\Psi} d\mathbb{P}^W(w) = \tilde{\Psi}. \end{aligned}$$

Since  $\alpha \in \mathcal{A}_t^2$  was arbitrary, we conclude that  $\Psi = \tilde{\Psi}$ .  $\square$

Just as it occurs under full information (see Section 1.4), for a fully risk-neutral MM with negligible costs (i.e.  $N^* = +\infty$ ,  $\gamma = \zeta = 0 \equiv \ell$ ),  $\Theta$  can be further decomposed. Note, from their definition, that in this case  $\widehat{\lambda}^{\pm, \alpha, t, n, \pi}$  and  $\Pi^{\alpha, t, n, \pi}$  do not depend on  $n$ . As it was proved in Theorem 1.3.1.1, when  $\gamma = 0$  the family  $(\tilde{\mathbb{P}}^{\alpha, t, n})$  can be taken as the original physical probabilities, and these do not depend on  $n$  either. The following corollary is now immediate.

**Corollary 1.3.1.2.** *If  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$  then*

$$V(t, s, x, n, \pi) = x + sn + \mu n(T - t) + \Phi(t, \pi),$$

$$\text{with } \Phi(t, \pi) = \sup_{\alpha \in \tilde{\mathcal{A}}_t^2} \mathbb{E}^{t, \pi} \left[ \int_t^T \left\{ \delta_u^- \widehat{\lambda}_u^{-, \alpha, t, \pi} + \delta_u^+ \widehat{\lambda}_u^{+, \alpha, t, \pi} + \mu (\widehat{\lambda}_u^{-, \alpha, t, \pi} - \widehat{\lambda}_u^{+, \alpha, t, \pi}) \right\} du \right].$$

In the context of the previous corollary, formal substitution in (1.3.1.3) or (1.3.1.5) yields the following PIDE for  $\Phi$ :

$$\begin{aligned}
 0 = & \phi_t + \sum_{i,j=1}^k q_t^{ji} \pi^j \phi_{\pi^i} \\
 & + \sup_{\delta^- \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^- - \Lambda_i^-)(\delta^-) \pi^j \pi^i \phi_{\pi^i} + (\delta^- + \mu(T-t) + d_{\delta^-}^-(\phi)) \sum_{i=1}^k \pi^i \Lambda_i^-(\delta^-) \right\} \\
 & + \sup_{\delta^+ \in (\delta_*, \delta^*)} \left\{ \sum_{i,j=1}^k (\Lambda_j^+ - \Lambda_i^+)(\delta^+) \pi^j \pi^i \phi_{\pi^i} + (\delta^+ - \mu(T-t) + d_{\delta^+}^+(\phi)) \sum_{i=1}^k \pi^i \Lambda_i^+(\delta^+) \right\},
 \end{aligned} \tag{1.3.1.9}$$

with terminal condition  $\phi(T, \pi) = 0$ , where:

$$d_{\delta}^{\pm}(\phi)(t, \pi) = \phi\left(t, \frac{1}{\sum_{j=1}^k \pi^j \Lambda_j^{\pm}(\delta)} (\pi^1 \Lambda_1^{\pm}(\delta), \dots, \pi^k \Lambda_k^{\pm}(\delta))\right) - \phi(t, \pi).$$

We now want to prove that  $\Theta$  (resp.  $\Phi$ ) is the unique continuous viscosity solution of the terminal condition PIDE (1.3.1.5) (resp. (1.3.1.9)). (See, e.g., [Son88, Def.2.1] for the relevant definition, or more in general [DF99, Def.7.3], recalling that in our case we have no boundary conditions other than that at terminal time.) A complication inevitably arises, as classical viscosity techniques [Bou07, FS06, ØS09] cannot be applied directly to weak formulation models such as ours. However, the decomposition of Theorem 1.3.1.1 (resp. Corollary 1.3.1.2) not only reduces the dimensionality of the problem, but also states that the MM may neglect the diffusion component of the state process altogether, focusing solely on the time-space state variable  $(u, N^{t,n}, \Pi^{\alpha,t,n,\pi})$  (resp.  $(u, \Pi^{\alpha,t,n,\pi})$ ). This is a PDMP as introduced in [Dav84] (detailed treatments also found in [BR11, Dav93]). Using results from PDMPs theory, our continuous-time problem is identified with a control problem for a discrete-time Markov decision model, as in [BR09, BR10, BR11, CEFS16a], and linked again with viscosity solutions of HJB PIDEs as in [CEFS16a, DF99].

An inevitable drawback is that the PDMPs approach relies on the use of the so-called *randomized* (or *relaxed*) controls and requires the control space to be compact. Hence, for the following theorem we will assume  $-\infty < \delta_* < \delta^* < +\infty$ . Assuming a uniform lower constraint  $\delta_* > -\infty$  is hardly a problem. On the contrary,  $\delta_* = 0$  (or even some small positive number) is the most meaningful in practice, as negative spreads imply the MM is willing to offer her clients better prices than the reference price  $S$ . ( $\delta_* = -\infty$  is motivated in the literature by mathematical convenience rather than modelling accuracy.) A uniform upper bound  $\delta^* < +\infty$ , on the other hand, is harder to assess a priori. Fortunately, in most situations encountered in practice, the unconstrained optimization will yield bounded optimal spreads nonetheless, and the MM can dispense with  $\delta_*, \delta^*$  if she wishes to do so (see Section 1.5 for an example).

**Theorem 1.3.1.3.** *Assume  $-\infty < \delta_* < \delta^* < \infty$ . For  $N^* < \infty$  (resp.  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ ), let  $\Theta$  (resp.  $\Phi$ ) be as in Theorem 1.3.1.1 (resp. Corollary 1.3.1.2). Then  $\Theta$  (resp.  $\Phi$ ) is the unique continuous viscosity solution of the terminal condition PIDE (1.3.1.5) (resp. (1.3.1.9)).*



*Proof.* Case  $N^* < +\infty$  and  $\gamma > 0$ : Let us write  $\Psi = \frac{1+\Upsilon}{\gamma}$ , with

$$\Upsilon(t, n, \pi) = \sup_{\alpha \in \tilde{\mathcal{A}}_t^2} \tilde{\mathbb{E}}^{\alpha, t, n, \pi} \left[ \exp \left( -\gamma (\tilde{P}_{t,T}^{\alpha, n, \pi} + \ell(N_T^{t, n})) \right) \left( -\exp(\gamma \ell(N_T^{t, n})) \right) \right].$$

Then  $\Upsilon$  can be regarded as the value function of an optimization problem in the standard Bolza-Lagrange formulation, i.e.,

$$\Upsilon(a) = \sup_{\alpha \in \tilde{\mathcal{A}}_t^2} \tilde{\mathbb{E}}^{\alpha, a} \left[ \int_t^T D_{t,u}^{\alpha, a} f(u, A_u^{\alpha, a}, \alpha_u) du + D_{t,T}^{\alpha, a} g(A_T^{\alpha, a}) \right],$$

where the state variable  $A_u^{\alpha, a} = (u, N_u^{t, n}, \Pi_u^{\alpha, t, n, \pi})$  is a PDMP with bounded state space  $[0, T] \times (\mathbb{Z} \cap [-N^*, N^*]) \times \Delta^\circ$  and initial condition  $a = (t, n, \pi)$ ,  $D_{t,u}^{\alpha, a} = \exp \left( -\int_t^u \rho(v, A_v^{\alpha, a}, \alpha_v) dv \right)$  is the discount factor and  $f, g, \rho$  are bounded functions. (These functions are bounded thanks to the control space being bounded.) The continuity of  $\Upsilon$  can be proved now in the same way as in [CEFS16a, Thm.4.10] albeit in a more straightforward manner. This is due to the boundedness of  $f, g, \rho$ , the fact that  $\Pi^{\alpha, t, n, \pi}$  never visits the relative border of the simplex  $\Delta$ , and that there is no exit time of the state space other than the terminal time. Assumptions [CEFS16a, Asm.4.7] are clearly verified in our model, and [CEFS16a, Asm.2.1] has already been accounted for at the beginning of the proof of Proposition 1.2.1.2.<sup>15</sup> We remark that in our case the bounding function (see [CEFS16a, Lem.4.6]) can be taken simply as  $b(t, n, \pi) = \exp(\eta(T - t))$ , for an  $\eta > 0$  large enough to prove contractiveness.

Having proved the continuity, the same proof of [CEFS16a, Thm.5.3] (or [DF99, Thm.7.5]) shows that  $\Upsilon$  is the unique continuous viscosity solution of its standard HJB equation. We remark once again that our case is simpler, in that  $\Upsilon$  is bounded and we do not have any boundary conditions other than the terminal time condition. In particular, there is no need for additional assumptions on the growth of  $\Upsilon$ .

Finally, the result for  $\Theta$  is obtained via the two increasing diffeomorphic transformations  $\Psi = \frac{1+\Upsilon}{\gamma}$  and  $\Theta = U_\gamma^{-1} \circ \Psi$ .

Case  $N^* < +\infty$  and  $\gamma = 0$ : The only difference with the previous case is that the Bolza-Lagrange representation of the problem is obtained directly, since

$$\Theta(t, n, \pi) = \sup_{\alpha \in \tilde{\mathcal{A}}_t^2} \tilde{\mathbb{E}}^{\alpha, t, n, \pi} \left[ \int_t^T \left\{ \delta_u^- \widehat{\lambda}_u^{\alpha, t, n, \pi} + \delta_u^+ \widehat{\lambda}_u^{\alpha, t, n, \pi} + \mu N_u^{t, n} - \frac{1}{2} \sigma^2 \zeta(N_u^{t, n})^2 \right\} du - \ell(N_T^{t, n}) \right],$$

with no need for any transformation.

Case  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ : The same as the latter case but working instead with the state variable  $(u, \Pi^{\alpha, t, n, \pi})$  and the value function  $\Phi$ .  $\square$

## 1.4 Full information

In this section we consider the idealized case of a MM with full information. We assume the MM has inside information in such a way that she can observe the full filtration  $\mathbb{F}$ , and in

<sup>15</sup>An additional detail now is that [CEFS16a, Asm.2.1] is also used in [CEFS16a, Lem.4.1], which states that the drift coefficient in equation (1.2.1.2) is Lipschitz in the state variable, uniform in time and control. This is routinely verified in our case, under our new assumption:  $-\infty < \delta_* < \delta^* < \infty$ .

particular she can observe  $Y$ . For example, if  $Y$  represents different levels of competition amongst liquidity providers, this would practically mean the MM has information regarding her competitors' quotes. We will see that in this case the value function turns out to be a regular, classical solution, of its HJB equation. Afterwards, we will compare the results with those obtained in the more realistic setting of partial information.

### 1.4.1 Dimensionality reduction and the general system of ODEs

We consider problem (1.1.2.6) but under full information, with the set of admissible spreads:

$$\mathcal{U} := \{\delta : [0, T] \rightarrow \overline{(\delta_*, \delta^*)} : \delta \text{ is } \mathbb{F}\text{-predictable and bounded from below}\}.$$

Note that in this section, and when  $\delta^* = +\infty$ , we do not assume upper-boundedness of the spreads. Let  $t \in [0, T]$ ,  $s, x \in \mathbb{R}$  and  $(n, i) \in \mathcal{I} := (\mathbb{Z} \cap [-N^*, N^*]) \times \{1, \dots, k\}$ . Consider the processes  $S^{t,s}, X^{\alpha,t,x,s}, N^{t,n}, P^{\alpha,s,x,n}$  as defined in Section 1.3 and  $Y^{t,i}$  a Markov chain with deterministic generator matrix  $Q$ , state space  $\{1, \dots, k\}$  and such that  $Y_{t-}^{t,i} = i$ . We assume the physical probabilities  $\mathbb{P}^{\alpha,t,n,i}$  are defined for every  $\alpha \in \mathcal{U}^2$  and that Assumptions 1.1.2.1, 1.1.2.2 and (1.1.2.2) (starting at time  $t^-$ ) are still in place. The value function in this case is

$$V(t, s, x, n, i) := \sup_{\alpha \in \mathcal{U}_t^2} \mathbb{E}^{t,n,i} \left[ U_\gamma(P_{t,T}^{\alpha,s,x,n}) \right], \quad (1.4.1.1)$$

By means of the Dynamic Programming Principle and Ito's Lemma one can formally derive the following HJB PIDE for  $V$ :

$$\begin{aligned} 0 = & v_t(t, s, x, n, i) + \mu v_s(t, s, x, n, i) + \frac{1}{2} \sigma^2 v_{ss}(t, s, x, n, i) \\ & + \frac{1}{2} \sigma^2 \zeta n^2 (\gamma v(t, s, x, n, i) - 1) + \sum_{j=1}^k q_t^{ij} v(t, s, x, n, j) \\ & + \mathbb{1}_{\{n < N^*\}} \sup_{\delta^- \in \overline{(\delta_*, \delta^*)}} \Lambda_i^-(\delta^-) \left( v(t, s, x - (s - \delta^-), n + 1, i) - v(t, s, x, n, i) \right) \\ & + \mathbb{1}_{\{-n < N^*\}} \sup_{\delta^+ \in \overline{(\delta_*, \delta^*)}} \Lambda_i^+(\delta^+) \left( v(t, s, x + (s + \delta^+), n - 1, i) - v(t, s, x, n, i) \right), \end{aligned} \quad (1.4.1.2)$$

with terminal condition  $v(T, s, x, n, i) = U_\gamma(x + sn - \ell(n))$ .

In this new context, instead of formally proving a decomposition of  $V$  as in Theorem 1.3.1.1, it is more straightforward to propose an ansatz and ultimately prove it valid with a verification theorem. (This is the standard approach used in the Avellaneda–Stoikov framework.) Let us consider an ansatz for the value function analogous to those used for the one regime case:

$$V(t, s, x, n, i) = U_\gamma(x + sn + \Theta(t, n, i)),$$

for some function  $\Theta : [0, T] \times \mathcal{I} \rightarrow \mathbb{R}$ ,  $C^1$  in time. Substituting in (1.4.1.2) and using Assumptions 1.1.2.1, we see that  $\Theta$  must satisfy a system of ODEs indexed in  $(n, i) \in \mathcal{I}$ :

$$\begin{aligned}
 0 = & \theta_t(t, n, i) + \mu n - \frac{1}{2} \sigma^2 n^2 (\zeta + \gamma) + \sum_{j \neq i} q_t^{ij} U_\gamma(\theta(t, n, j) - \theta(t, n, i)) \\
 & + \mathbb{1}_{\{n < N^*\}} H_i^-(\theta(t, n+1, i) - \theta(t, n, i)) + \mathbb{1}_{\{-n < N^*\}} H_i^+(\theta(t, n-1, i) - \theta(t, n, i)),
 \end{aligned} \tag{1.4.1.3}$$

with terminal condition  $\theta(T, n, i) = -\ell(n)$ , where:

1.  $H_i^\pm(d) := \sup_{\delta \in (\delta_*, \delta^*)} h_i^\pm(\delta, d)$ , for  $d \in \mathbb{R}$ .
2.  $h_i^\pm(\delta, d) := \Lambda_i^\pm(\delta) U_\gamma(\delta + d)$ , for  $d \in \mathbb{R}$ ,  $\delta \in (\delta_*, \delta^*)$ .

The original problem is simplified in this way, both by the dimension of the state variable and by the complexity of the equations, provided we can show that problem (1.4.1.3) admits a solution. With this aim in mind, let us prove first the following property of the Hamiltonian functions  $H_1^\pm, \dots, H_k^\pm$ .

**Lemma 1.4.1.1.** *For all  $1 \leq i \leq k$ , it holds:*

- (i) *For each compact  $K \subset \mathbb{R}$ , there exists  $[a, b] \subseteq (\delta_*, \delta^*)$  such that  $H_i^\pm(d) = \max_{\delta \in [a, b]} h_i^\pm(\delta, d)$ , for all  $d \in K$ .*
- (ii)  *$H_i^\pm$  is locally Lipschitz.*

*Proof.* Fix  $1 \leq i \leq k$  and let  $K \subset \mathbb{R}$  be a compact set. We verify first that (i) is a consequence of Assumptions 1.1.2.2. Let  $C > 0$  such that  $|d| \leq C$  for all  $d \in K$ . If  $\delta_* = -\infty$ , then we can take any  $a < \min\{-C, \delta^*\}$  since  $h_i^\pm(\delta, d) < h_i^\pm(a, d)$  for all  $\delta < a$  and  $d \in K$ . On the other hand, if  $\delta^* = +\infty$ , take some  $c > \max\{\delta_*, 2C\}$ . It holds that  $h_i^\pm(c, d) \geq \Lambda_i^\pm(c) U_\gamma(C) =: \varepsilon > 0$  and we can choose  $b > c$  such that  $h_i^\pm(\delta, d) \leq h_i^\pm(\delta, C) < \varepsilon$  for all  $\delta \geq b$  and  $d \in K$ . Replacing supremum by maximum is now immediate due to the continuity of  $h_i^\pm(\cdot, d)$  on  $[a, b]$  for all  $d$ .

(ii) is routinely verified using that the family  $\{h_i^\pm(\delta, \cdot)\}_{\delta \in [a, b]}$  is equi-Lipschitz on  $K$ .  $\square$

We want to prove now that the Cauchy problem (1.4.1.3) admits a unique global classical solution  $\theta$  which is  $C^1$  in time. To this purpose, we will treat the cases of the finite system ( $N^* < \infty$ ) and infinite system ( $N^* = \infty$ ) of equations separately.

## 1.4.2 Constrained inventory ODEs

For  $N^* < \infty$  we are dealing with a finite system of ODEs. We know that under certain regularity conditions the Cauchy problem (1.4.1.3) is guaranteed to have a classical solution on some neighbourhood  $(\tau, T] \subset [0, T]$  of  $T$ . Nonetheless, it is not always the case that such a local solution can be extended to a global one on  $[0, T]$ . Following [Gué17, GL15], we start by proving a comparison principle for (1.4.1.3) that will allow us, in particular, to show the existence of a global solution. The argument used is standard for comparison principles of HJB-type equations.

**Proposition 1.4.2.1 (Comparison Principle).** *Let  $I \subseteq [0, T]$  be an interval containing  $T$  and let  $\bar{\theta}, \underline{\theta} : I \times \mathcal{I} \rightarrow \mathbb{R}$  be classical ( $C^1$  with respect to time) super- and sub-solutions resp. of (1.4.1.3). That is,*

$$0 \leq -\bar{\theta}_t(t, n, i) - \mu n + \frac{1}{2}\sigma^2 n^2(\zeta + \gamma) - \sum_{j \neq i} q_t^{ij} U_\gamma(\bar{\theta}(t, n, j) - \bar{\theta}(t, n, i)) \\ - \mathbb{1}_{\{n < N^*\}} H_i^-(\bar{\theta}(t, n+1, i) - \bar{\theta}(t, n, i)) - \mathbb{1}_{\{-n < N^*\}} H_i^+(\bar{\theta}(t, n-1, i) - \bar{\theta}(t, n, i)), \quad (1.4.2.1)$$

$$0 \geq -\underline{\theta}_t(t, n, i) - \mu n + \frac{1}{2}\sigma^2 n^2(\zeta + \gamma) - \sum_{j \neq i} q_t^{ij} U_\gamma(\underline{\theta}(t, n, j) - \underline{\theta}(t, n, i)) \\ - \mathbb{1}_{\{n < N^*\}} H_i^-(\underline{\theta}(t, n+1, i) - \underline{\theta}(t, n, i)) - \mathbb{1}_{\{-n < N^*\}} H_i^+(\underline{\theta}(t, n-1, i) - \underline{\theta}(t, n, i)) \quad (1.4.2.2)$$

and

$$\underline{\theta}(T, \cdot, \cdot) \leq -\ell \leq \bar{\theta}(T, \cdot, \cdot). \quad (1.4.2.3)$$

Then

$$\underline{\theta} \leq \bar{\theta}.$$

*Proof.* Suppose first  $I = [\tau, T]$  for some  $0 \leq \tau < T$  and let  $\varepsilon > 0$ . Since  $N^* < \infty$ , there exists  $(t_\varepsilon, n_\varepsilon, i_\varepsilon) \in [\tau, T] \times \mathcal{I}$  such that

$$\bar{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon) - \underline{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon) + \varepsilon(T - t_\varepsilon) = \min_{(t, n, i) \in [\tau, T] \times \mathcal{I}} \bar{\theta}(t, n, i) - \underline{\theta}(t, n, i) + \varepsilon(T - t). \quad (1.4.2.4)$$

If  $t_\varepsilon < T$ , then we must have

$$\bar{\theta}_t(t_\varepsilon, n_\varepsilon, i_\varepsilon) - \underline{\theta}_t(t_\varepsilon, n_\varepsilon, i_\varepsilon) \geq \varepsilon.$$

Let us see that the left-hand side is non-positive. By (1.4.2.1) and (1.4.2.2),

$$\begin{aligned} & \bar{\theta}_t(t_\varepsilon, n_\varepsilon, i_\varepsilon) - \underline{\theta}_t(t_\varepsilon, n_\varepsilon, i_\varepsilon) \\ & \leq \sum_{j \neq i_\varepsilon} q_t^{ij} \left( U_\gamma(\underline{\theta}(t_\varepsilon, n_\varepsilon, j) - \underline{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) - U_\gamma(\bar{\theta}(t_\varepsilon, n_\varepsilon, j) - \bar{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) \right) \\ & + \mathbb{1}_{\{n < N^*\}} \left( H_i^-(\underline{\theta}(t_\varepsilon, n_\varepsilon + 1, i_\varepsilon) - \underline{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) - H_i^-(\bar{\theta}(t_\varepsilon, n_\varepsilon + 1, i_\varepsilon) - \bar{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) \right) \\ & + \mathbb{1}_{\{-n < N^*\}} \left( H_i^+(\underline{\theta}(t_\varepsilon, n_\varepsilon - 1, i_\varepsilon) - \underline{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) - H_i^+(\bar{\theta}(t_\varepsilon, n_\varepsilon - 1, i_\varepsilon) - \bar{\theta}(t_\varepsilon, n_\varepsilon, i_\varepsilon)) \right). \end{aligned}$$

$H_i^\pm$  increasing (resp.  $U_\gamma$  increasing) and (1.4.2.4) imply that the last two terms (resp. the first one) are non-positive. We must have then that  $t_\varepsilon = T$ , and due to (1.4.2.4) and (1.4.2.3) for all  $(t, n, i) \in [0, T] \times \mathcal{I}$ :

$$\begin{aligned} \bar{\theta}(t, n, i) - \underline{\theta}(t, n, i) + \varepsilon(T - t) & \geq \bar{\theta}(T, n_\varepsilon, i_\varepsilon) - \underline{\theta}(T, n_\varepsilon, i_\varepsilon) + \varepsilon(T - T) \geq 0 \\ \bar{\theta}(t, n, i) & \geq \underline{\theta}(t, n, i) - \varepsilon(T - t). \end{aligned}$$

Since  $\varepsilon > 0$  was arbitrary, we obtain the desired result.

The case  $I = (\tau, T]$  is now a consequence of comparing  $\underline{\theta}$  and  $\bar{\theta}$  on intervals of the form  $[t_n, T] \subseteq (\tau, T]$  with  $t_n \searrow \tau$ .  $\square$

We can now prove the existence and uniqueness of a classical global solution of the Cauchy problem (1.4.1.3).

**Theorem 1.4.2.2.** *There exists a unique  $\Theta : [0, T] \times \mathcal{I} \rightarrow \mathbb{R}$ ,  $C^1$  in time, which (classically) solves the Cauchy problem (1.4.1.3).*

*Proof.* For each  $1 \leq i \leq k$ , Lemma 1.4.1.1 tells us that  $H_i^\pm : \mathbb{R} \rightarrow \mathbb{R}$  is locally Lipschitz. We also assumed  $Q : [0, T] \rightarrow \mathbb{R}^{k \times k}$  continuous (Assumptions 1.1.2.1). Hence, by the Cauchy-Lipschitz Theorem, the terminal condition system of ODEs (1.4.1.3) admits a unique  $C^1$  local solution  $(\Theta(\cdot, n, i))_{(n, i) \in \mathcal{I}}$ , defined on some maximal interval  $I \subset [0, T]$  containing  $T$ .

Suppose that  $I \subsetneq [0, T]$ . Then  $I = (\tau, T]$  for some  $0 \leq \tau < T$  and  $\|\Theta(t)\| \rightarrow \infty$  as  $t \searrow \tau$ . We claim that  $\Theta$  is actually bounded, resulting in a contradiction. Indeed, take

$$\bar{K} = \max_{(i, n) \in \mathcal{I}} \left\{ \left| \mu n - \frac{1}{2} \sigma^2 n^2 (\zeta + \gamma) + \mathbb{1}_{\{n < N^*\}} H_i^-(0) + \mathbb{1}_{\{-n < N^*\}} H_i^+(0) \right| \right\}$$

and

$$\underline{K} = \ell(N^*),$$

and define  $\bar{\theta} : I \times \mathcal{I} \rightarrow \mathbb{R}$  by  $\bar{\theta}(t, n, i) = \bar{K}(T - t)$  and  $\underline{\theta} := -\bar{\theta} - \underline{K}$ . Then  $\bar{\theta}$  (resp.  $\underline{\theta}$ ) is a super- (resp. sub-) solution of the Cauchy problem (1.4.1.3). By the Comparison Principle (1.4.2.1),

$$-\bar{K}T - \underline{K} \leq \underline{\theta} \leq \Theta \leq \bar{\theta} \leq \bar{K}T,$$

proving that  $\Theta$  is bounded.  $\square$

### 1.4.3 Unconstrained inventory ODEs

We consider now  $N^* = +\infty$ , for which (1.4.1.3) becomes an infinite system of ODEs. Recall that in this case we assumed  $\gamma = \zeta = 0 \equiv \ell$ , i.e., the MM is fully risk-neutral and has negligible costs. This allows us to further reduce the dimensionality of the state variable by the additional ansatz

$$\Theta(t, n, i) = \mu n(T - t) + \Phi_i(t), \quad (1.4.3.1)$$

for some  $\Phi = (\Phi_i)_{i=1}^k \in C^1([0, T], \mathbb{R}^k)$ . Substituting in (1.4.1.3), we get that  $\Phi$  must solve the finite linear system of ODEs

$$\begin{cases} \phi'(t) = -Q(t)\phi(t) + b(t) \\ \phi(T) = 0, \end{cases} \quad (1.4.3.2)$$

with  $b_i(t) = -H_i^-(\mu(T - t)) - H_i^+(-\mu(T - t))$ ,  $i = 1, \dots, k$ . By continuity of  $Q$  and  $b$  (see Lemma 1.4.1.1), the previous system is known to have a unique global solution  $\Phi \in C^1([0, T], \mathbb{R}^k)$  which can be computed by the variation of parameters method. Straightforward verification now gives the following:

**Proposition 1.4.3.1.** *Let  $\Phi \in C^1([0, T], \mathbb{R}^k)$  be the unique solution of (1.4.3.2). Then  $\Theta : [0, T] \times \mathcal{I} \rightarrow \mathbb{R}$  such that  $\Theta(t, n, i) = \mu n(T - t) + \Phi_i(t)$ , is the unique,  $C^1$  in time, (classical) solution of the Cauchy problem (1.4.1.3) with  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ .*

### 1.4.4 General Verification Theorem

We give now the complete solution for the general model under full information. For the next theorem we note that given the function  $\Theta$  defined in Theorem 1.4.2.2 for  $N^* < +\infty$  (resp. Proposition 1.4.3.1 for  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ ) the difference or ‘jump’ terms of equation (1.4.1.3) are bounded by continuity. That is,  $\Theta(t, n \mp 1, i) - \Theta(t, n, i)$  is bounded on  $[0, T] \times \mathcal{I}$  for  $N^* < +\infty$  (resp.  $\Theta(t, n \mp 1, i) - \Theta(t, n, i) = \mp \mu(T - t)$  is bounded on  $[0, T]$ ).

**Theorem 1.4.4.1 (Verification Theorem).** *Let  $\Theta$  be as in Theorem 1.4.2.2 for  $N^* < +\infty$  and as in Proposition 1.4.3.1 for  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ . Then the value function in (1.4.1.1) is*

$$V(t, s, x, n, i) = U_\gamma(x + sn + \Theta(t, n, i)).$$

Furthermore, given  $C > 0$  such that  $|\Theta(t, n \mp 1, i) - \Theta(t, n, i)| \leq C$  for all  $(t, n, i) \in [0, T] \times \mathcal{I}$  with  $-N^* \leq n \mp 1 \leq N^*$ , there exist Borel measurable functions  $\bar{\delta}_1^\pm, \dots, \bar{\delta}_k^\pm : [-C, C] \rightarrow (\bar{\delta}_*, \bar{\delta}^*)$  such that  $\bar{\delta}_i^\pm(d) \in \arg \max_{\delta \in (\bar{\delta}_*, \bar{\delta}^*)} h_i^\pm(\delta, d)$  for all  $d \in [-C, C]$ ,  $1 \leq i \leq k$ ; and for any such functions, the strategy  $(\bar{\delta}_u^+, \bar{\delta}_u^-)$ , with

$$\bar{\delta}_u^\pm := \bar{\delta}_{Y_{u-}^{t,i}}^\pm \left( \Theta(u, N_{u-}^{t,n} \mp 1, Y_{u-}^{t,i}) - \Theta(u, N_{u-}^{t,n}, Y_{u-}^{t,i}) \right),$$

is optimal.

*Notation.* Note that we make a slight abuse of notation, writing  $\bar{\delta}_i^\pm$  for the Borel functions in the theorem and  $(\bar{\delta}_u^\pm)$  for the spread processes.

*Proof.* Let  $1 \leq i \leq k$ . We check first that we can choose a maximizer of  $h_i^\pm(\cdot, d)$  in a measurable way with respect to  $d$ . Lemma 1.4.1.1 tells us that there exists  $[a_i, b_i] \subseteq (\bar{\delta}_*, \bar{\delta}^*)$  such that  $h_i^\pm(\delta, d)$  attains its maximum in  $[a_i, b_i]$  for all  $d \in [-C, C]$ . Due to Assumptions 1.1.2.2 ((ii)),  $h_i^\pm$  is continuous and, in particular, a Carathéodory function [AB06, Def.4.50]. Thus, the Measurable Maximum Theorem [AB06, Thm.18.19] guarantees the existence of a Borel selector of the  $\arg \max$ ,  $\bar{\delta}_i^\pm : [-C, C] \rightarrow [a_i, b_i] \subseteq (\bar{\delta}_*, \bar{\delta}^*)$ . (Note that the weak measurability assumption is trivially verified in this case.)

From here onwards, let  $\bar{\delta}_1^\pm, \dots, \bar{\delta}_k^\pm : [-C, C] \rightarrow (\bar{\delta}_*, \bar{\delta}^*)$  be some Borel selectors as above. By Lemma 1.4.1.1 again, these functions must be bounded from below, and the spread processes  $(\bar{\delta}_u^\pm)$  defined as in the theorem are clearly admissible.

We define now  $\tilde{V}(t, s, x, n, i) := U_\gamma(x + ns + \Theta(t, n, i))$  and we want to show that  $\tilde{V} = V$ . Let us fix the initial time and values,  $t \in [0, T]$ ,  $s, x \in \mathbb{R}$  and  $(n, i) \in \mathcal{I}$ , and consider an arbitrary strategy  $\alpha = (\delta_u^-, \delta_u^+) \in \mathcal{U}_t^2$ . For shortness, we omit these initial conditions and strategy from the notation of the processes and the expectation. We denote by  $\mathcal{S} := \{j - l : 1 \leq l, j \leq k, l \neq j\}$ , the set of jump heights of  $Y$ , and

$$R_{t,v} := -\frac{1}{2}\sigma^2\zeta \int_t^v N_u^2 du, \quad Z_{t,v} := \exp(-\gamma R_{t,v}).$$

By the identity  $U_\gamma(a + b) = U_\gamma(a)e^{-\gamma b} + U_\gamma(b)$ , we can rewrite the utility of the MM’s penalized P&L as

$$\begin{aligned} U_\gamma(P_{t,T}) &= U_\gamma(X_T + S_T N_T - \ell(N_T)) Z_{t,T} + U_\gamma(R_{t,T}) \\ &= \tilde{V}(T, S_T, X_T, N_T, Y_T) Z_{t,T} + U_\gamma(R_{t,T}). \end{aligned} \tag{1.4.4.1}$$

Using integration by parts and Ito's Lemma (recalling (1.1.2.2)), we re-express the last two terms as

$$U_\gamma(R_{t,T}) = -\frac{1}{2}\sigma^2\zeta \int_t^T Z_{t,u}N_u^2 du$$

and

$$\begin{aligned} & \tilde{V}(T, S_T, X_T, N_T, Y_T)Z_{t,T} \\ &= \tilde{V}(t, s, x, n, i) + \int_t^T \tilde{V}(u, S_u, X_u, N_u, Y_u)dZ_{t,u} + \int_t^T Z_{t,u}d\tilde{V}(u, S_u, X_u, N_u, Y_u) \\ &= \tilde{V}(t, s, x, n, i) + \frac{1}{2}\sigma^2\zeta\gamma \int_t^T V(u, S_u, X_u, N_u, Y_u)Z_{t,u}N_u^2 du + \int_t^T Z_{t,u}\tilde{V}_t(u, S_u, X_{u-}, N_{u-}, Y_{u-})du \\ &+ \mu \int_t^T Z_{t,u}\tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-})du + \sigma \int_t^T Z_{t,u}\tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-})dW_u \\ &+ \frac{1}{2}\sigma^2 \int_t^T Z_{t,u}\tilde{V}_{ss}(u, S_u, X_{u-}, N_{u-}, Y_{u-})du \\ &+ \int_t^T Z_{t,u}(\tilde{V}(u, S_u, X_{u-} - (S_u - \delta_u^-), N_{u-} + 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))dN_u^- \\ &+ \int_t^T Z_{t,u}(\tilde{V}(u, S_u, X_{u-} + S_u + \delta_u^+, N_{u-} - 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))dN_u^+ \\ &+ \int_t^T Z_{t,u} \int_S (\tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-} + h) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))\mu^Y(dt, dh), \end{aligned}$$

where  $\mu^Y$  is the jump measure of  $Y$ . Substituting in (1.4.4.1),

$$\begin{aligned} U_\gamma(P_{t,T}) &= \tilde{V}(t, s, x, n, i) + \sigma \int_t^T Z_{t,u}\tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-})dW_u \\ &+ \int_t^T Z_{t,u} \left\{ \tilde{V}_t(u, S_u, X_{u-}, N_{u-}, Y_{u-})du + \mu\tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-})du \right. \\ &+ \frac{1}{2}\sigma^2\tilde{V}_{ss}(u, S_u, X_{u-}, N_{u-}, Y_{u-})du + \frac{1}{2}\sigma^2\zeta N_u^2(\gamma V(u, S_u, X_u, N_u, Y_u) - 1)du \\ &+ \int_S (\tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-} + h) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))\mu^Y(dt, dh) \\ &+ (\tilde{V}(u, S_u, X_{u-} - (S_u - \delta_u^-), N_{u-} + 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))dN_u^- \\ &\left. + (\tilde{V}(u, S_u, X_{u-} + S_u + \delta_u^+, N_{u-} - 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-}))dN_u^+ \right\}. \end{aligned} \tag{1.4.4.2}$$

Next, we want to verify that the process  $\left( \int_t^v Z_{t,u}\tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-})dW_u \right)_{t \leq v \leq T}$  is a zero mean martingale. Since  $Z_{t,u}$  is bounded (either  $N^* < \infty$  or  $\gamma = 0$ ) it suffices to check

$$\int_t^T \mathbb{E} \left[ \tilde{V}_s(u, S_u, X_u, N_u, Y_u)^2 \right] du < +\infty. \tag{1.4.4.3}$$

If  $\gamma = 0$ ,

$$\tilde{V}_s(u, S_u, X_u, N_u, Y_u)^2 = N_u^2$$

and (1.4.4.3) is a consequence of (1.1.2.5).

If  $\gamma \neq 0$  (hence  $N^* < \infty$ ), let  $C$  be a lower bound for  $(\delta_u^-)$  and  $(\delta_u^+)$ . Integration by parts and Hölder inequality yield

$$\begin{aligned} \mathbb{E} \left[ \tilde{V}_s(u, S_u, X_u, N_u, Y_u)^2 \right] &= \gamma^2 \mathbb{E} \left[ \exp \left( -2\gamma (X_u + S_u N_u + \theta(u, N_u, Y_u)) \right) N_u^2 \right] \\ &= \gamma^2 \exp(-2\gamma(x + sn)) \\ &\times \mathbb{E} \left[ \exp \left( -2\gamma \int_t^u \delta_w^- dN_w^- + \delta_w^+ dN_w^+ + \mu N_w dw + \sigma N_w dW_w + \theta(w, N_w, Y_w) \right) N_u^2 \right] \\ &\leq \gamma^2 N^{*2} \exp \left( 2\gamma(x + sn + \mu N^*(T - t) + \|\theta\|_{L^\infty([t, T] \times \mathcal{I})} + 2\sigma^2 \gamma N^{*2}(T - t)) \right) \\ &\times \mathbb{E} \left[ \exp \left( -4\gamma C(N_T^- + N_T^+ - N_t^- - N_t^+) \right) \right]^{\frac{1}{2}} \mathbb{E} \left[ \exp \left( -4\gamma \sigma \int_t^T N_u dW_u - 8\gamma^2 \sigma^2 \int_t^T N_u^2 du \right) \right]^{\frac{1}{2}} \end{aligned}$$

and (1.4.4.3) is again consequence of (1.1.2.5) and Novikov's condition (trivially satisfied).

In the same way, recalling that  $(Z_{t,u})$  is bounded,  $(\delta_u^\pm)$  is bounded from below and that  $Q$  is bounded by continuity, one can check that

$$\int_t^T \mathbb{E} \left[ Z_{t,u} |\tilde{V}(u, S_u, X_{u-} \pm (S_u \pm \delta_u^\pm), N_{u-} \mp 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-})| \lambda_u^\pm \right] du < +\infty$$

and

$$\int_t^T \mathbb{E} \left[ Z_{t,u} \sum_{j \neq Y_{u-}} q_t^{Y_{u-}-j} |\tilde{V}(u, S_u, X_{u-}, N_{u-}, j) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-})| \right] du < +\infty.$$

Taking expectation in (1.4.4.2), the Brownian term vanishes and integration with respect to  $dN^-$ ,  $dN^+$  and  $d\mu^Y$  is replaced by integration with respect to their dual predictable projections (see, e.g., [Bré81, p.27 T8 and p.235 C4]). That is,

$$\begin{aligned} \mathbb{E} \left[ U_\gamma(P_{t,T}) \right] &= \tilde{V}(t, s, x, n, i) + \int_t^T \mathbb{E} \left[ Z_{t,u} \left\{ \tilde{V}_t(u, S_u, X_{u-}, N_{u-}, Y_{u-}) \right. \right. \\ &\quad + \mu \tilde{V}_s(u, S_u, X_{u-}, N_{u-}, Y_{u-}) + \frac{1}{2} \sigma^2 \tilde{V}_{ss}(u, S_u, X_{u-}, N_{u-}, Y_{u-}) \\ &\quad + \frac{1}{2} \sigma^2 \zeta N_u^2 (\gamma \tilde{V}(u, S_u, X_u, N_u, Y_u) - 1) + \sum_{j=1}^k q_t^{Y_{u-}-j} \tilde{V}(u, S_u, X_{u-}, N_{u-}, j) \\ &\quad + \mathbb{1}_{u-}^- \Lambda_{Y_{u-}}^- (\delta_u^-) (\tilde{V}(u, S_u, X_{u-} - (S_u - \delta_u^-), N_{u-} + 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-})) \\ &\quad \left. \left. + \mathbb{1}_{u-}^+ \Lambda_{Y_{u-}}^+ (\delta_u^+) (\tilde{V}(u, S_u, X_{u-} + S_u + \delta_u^+, N_{u-} - 1, Y_{u-}) - \tilde{V}(u, S_u, X_{u-}, N_{u-}, Y_{u-})) \right\} \right] du \\ &\leq \tilde{V}(t, s, x, n, i), \end{aligned}$$



as  $\tilde{V}$  solves (1.4.1.2), with the equality attained for  $(\delta_u^-) = (\overline{\delta_u^-})$  and  $(\delta_u^+) = (\overline{\delta_u^+})$  by definition. We conclude that  $V = \tilde{V}$  and that the pair  $(\overline{\delta_u^-}, \overline{\delta_u^+})$  is optimal.  $\square$

*Remark.* For  $N^* < \infty$ , since the MM will not buy (resp. sell) whenever  $(N_{u-})$  hits  $N^*$  (resp.  $-N^*$ ), the value of  $(\overline{\delta_u^-})$  (resp.  $(\overline{\delta_u^+})$ ) at these stopping times is essentially irrelevant. From a strict mathematical perspective, the only constraint is that whichever the value we choose, the process needs to remain admissible.

### 1.4.5 Computing the optimal spreads: some particular cases

We have shown in Theorem 1.4.4.1 that the optimal spreads for the full information problem (1.4.1.1) can be computed in feedback form in terms of  $(t, n, i) = (t, N_{t-}, Y_{t-})$  at each time  $t \in [0, T]$ . Practically, this means finding  $\Theta$  (typically, numerically) that solves the terminal condition system of ODEs (1.4.1.3), and finding the spreads by maximization:

$$\overline{\delta^\pm}(t, n, i) \in \arg \max_{\delta \in (\overline{\delta_*}, \delta^*)} \Lambda_i^\pm(\delta) U_\gamma(\delta + \Theta(t, n \mp 1, i) - \Theta(t, n, i)). \quad (1.4.5.1)$$

(See the proof of Lemma 1.4.1.1 to see how to reduce the maximization to a compact domain in the cases of  $\delta_* = -\infty$  or  $\delta^* = +\infty$ .) The functions in (1.4.5.1) may admit multiple maximizers in general. In [BL14, Gué17, GL15] stronger assumptions are imposed on the orders intensities which guarantee, in particular, the uniqueness of the maximizers.

We now extend these assumptions to our context and give the corresponding characterization of the spreads. Henceforth, Assumption 1.1.2.2 ((ii)) is replaced by the following:

**Assumptions 1.4.5.1.**  $\Lambda_i^\pm \in C^2(\overline{(\delta_*, \delta^*)})$ ,  $\Lambda_i^{\pm'} < 0$ ,  $\Lambda_i^\pm \Lambda_i^{\pm''} < c(\Lambda_i^{\pm'})^2$  for some  $0 < c < 2$  and  $\lim_{\delta \rightarrow +\infty} \Lambda_i^\pm(\delta) = 0$  if  $\delta^* = +\infty$ , for all  $1 \leq i \leq k$ .<sup>16</sup>

*Remark.* The intensity functions of Examples 1.1.2.4 1, 2 and 3 all verify these stronger assumptions, while Example 4 only does so for  $d_i^\pm > 1$ . As a matter of fact, by solving the differential inequality in Assumptions 1.4.5.1, one sees that any  $\Lambda_i^\pm$  within this new framework has a strict upper bound of the form of Example 4. In particular, for  $\delta^* = +\infty$ ,  $\lim_{\delta \rightarrow +\infty} \delta \Lambda_i^\pm(\delta) = 0$  is necessarily satisfied.

Under Assumptions 1.4.5.1, the maximization of (1.4.5.1) (for fixed  $(t, n, i)$ ) can be replaced essentially by solving a contractive fixed point equation and flooring and capping the results at  $\delta_*$  and  $\delta^*$  respectively. To make this precise, even when the intensity functions are not defined beyond  $(\delta_*, \delta^*)$ , let us set

$$d_{i*}^\pm := \begin{cases} -U_\gamma^{-1}\left(\frac{\Lambda_i^\pm(\delta_*)}{\Lambda_i^{\pm'}(\delta_*)}\right) - \delta_* & \text{if } \delta_* > -\infty \\ +\infty & \text{if } \delta_* = -\infty, \end{cases} \quad d_i^{\pm*} := \begin{cases} -U_\gamma^{-1}\left(\frac{\Lambda_i^\pm(\delta^*)}{\Lambda_i^{\pm'}(\delta^*)}\right) - \delta^* & \text{if } \delta^* < +\infty \\ -\infty & \text{if } \delta^* = +\infty. \end{cases}$$

The following result shows how, up to constraints, the optimal spreads are given by a first term that maximizes the ‘expected’ utility of the MM’s instantaneous margin (see Remark

<sup>16</sup>Lateral derivatives are considered on the domain border. The strict inequality  $c < 2$  is needed when  $\gamma = 0$  and spreads are unconstrained.

1.1.2.3), plus an additional risk adjustment taking into account the inventory held and the prospect of the market shifting. Note how the first term depends on  $\Lambda_i^{\pm'}/\Lambda_i^{\pm}$ , the percentage sensitivity of the liquidity to spread changes.

**Proposition 1.4.5.2.** *Under Assumptions 1.4.5.1, the functions  $\overline{\delta}_1^{\pm}, \dots, \overline{\delta}_k^{\pm}$  of Theorem 1.4.4.1 are uniquely characterized by*

$$\overline{\delta}_i^{\pm}(d) = \delta^* \text{ if } d < d_i^{\pm*}, \quad \overline{\delta}_i^{\pm}(d) = \delta_* \text{ if } d > d_{i*}^{\pm} \quad (1.4.5.2)$$

and if  $d \in (\overline{d_i^{\pm*}}, d_{i*}^{\pm})$ , then  $\overline{\delta}_i^{\pm}(d)$  is the unique solution of the fixed point equation

$$\overline{\delta}_i^{\pm}(d) = -U_{\gamma}^{-1} \left( \frac{\Lambda_i^{\pm}(\overline{\delta}_i^{\pm}(d))}{\Lambda_i^{\pm'}(\overline{\delta}_i^{\pm}(d))} \right) - d. \quad (1.4.5.3)$$

Additionally, the Hamiltonians of equation (1.4.1.3) can be expressed as

$$H_i^{\pm}(d) = \begin{cases} h_i^{\pm}(\delta^*, d) & \text{if } d \leq d_i^{\pm*} \\ -\frac{\Lambda_i^{\pm 2}(\widehat{\delta}^{\pm}(d))}{\Lambda_i^{\pm'}(\widehat{\delta}^{\pm}(d)) - \gamma \Lambda_i^{\pm}(\widehat{\delta}^{\pm}(d))} & \text{if } d_i^{\pm*} < d < d_{i*}^{\pm} \\ h_i^{\pm}(\delta_*, d) & \text{if } d \geq d_{i*}^{\pm}. \end{cases}$$

*Proof.* For each  $1 \leq i \leq k$ , define

$$f_i^{\pm}(\delta, d) = \delta + U_{\gamma}^{-1} \left( \frac{\Lambda_i^{\pm}(\delta)}{\Lambda_i^{\pm'}(\delta)} \right) + d, \quad \text{with } \delta \in (\delta_*, \delta^*), \quad d \in \mathbb{R}.$$

By Assumptions 1.4.5.1 and straightforward computations (see computations, e.g., in [Gué17, Lemma 3.1]), one verifies that  $\text{sgn}(\frac{\partial h_i^{\pm}}{\partial \delta}) = -\text{sgn}(f_i^{\pm})$  and  $f_i^{\pm}$  is strictly increasing on each variable.<sup>17</sup> It follows that for any  $d < d_i^{\pm*}$  and  $\delta < \delta^*$  (resp.  $d > d_{i*}^{\pm}$  and  $\delta > \delta_*$ ),  $f_i^{\pm}(\delta, d) < f_i^{\pm}(\delta^*, d_i^{\pm*}) = 0$  (resp.  $f_i^{\pm}(\delta, d) > f_i^{\pm}(\delta_*, d_{i*}^{\pm}) = 0$ ), which proves (1.4.5.2).

In the same way, if  $d \in (d_i^{\pm*}, d_{i*}^{\pm})$  and  $-\infty < \delta_* < \delta^* < +\infty$ , then  $f_i^{\pm}(\delta_*, d) < 0$  and  $f_i^{\pm}(\delta^*, d) > 0$ . Hence, the continuity of  $f_i^{\pm}(\cdot, d)$  implies (1.4.5.3) (the uniqueness of the solution being due to the strict monotonicity of  $f_i^{\pm}(\cdot, d)$ ). The cases of unconstrained spreads are proved as in [Gué17].

Lastly, the cases  $d = d_i^{\pm*} > -\infty$  and  $d = d_{i*}^{\pm} < +\infty$  follow in the same manner; and the new expressions for the Hamiltonians are immediate, putting  $H_i^{\pm}(d) = h_i^{\pm}(\overline{\delta}_i^{\pm}(d), d)$ .  $\square$

As done in [Gué17, Lemma 3.1], (1.4.5.3) can be replaced by the explicit formula

$$\overline{\delta}_i^{\pm}(d) = (\Lambda_i^{\pm})^{-1} \left( \gamma H_i^{\pm}(d) - H_i^{\pm'}(d) \right),$$

but the computation of  $(\Lambda_i^{\pm})^{-1}$ ,  $H_i^{\pm}$  and  $H_i^{\pm'}$  still has to be carried out numerically, in general. We state now a few simplifications that arise in some particular cases, i.e., when the intensities are exponential or when  $N^* = \infty$ ,  $\gamma = \zeta = 0 \equiv \ell$ . These are all derived by straightforward substitution. We refer to [Gué17, Sect.4] for some asymptotic approximations when  $t \ll T$  in the one regime case, with unconstrained spreads and  $N^* < \infty$ .

<sup>17</sup>sgn denotes the sign function, with  $\text{sgn}(0) := 0$ .

**Corollary 1.4.5.3.** *If  $N^* = +\infty$  and  $\gamma = \zeta = 0 \equiv \ell$ , then*

$$\bar{\delta}^\pm(t, i) = \delta^* \text{ if } \mp \mu(T - t) < d_i^{\pm*}, \quad \bar{\delta}_i^\pm(d) = \delta_* \text{ if } \mp \mu(T - t) > d_{i*}^\pm, \quad (1.4.5.4)$$

*and if  $\mp \mu(T - t) \in \overline{(d_i^{\pm*}, d_{i*}^\pm)}$ , then  $\bar{\delta}^\pm(t, i)$  is the unique solution of the fixed point equation*

$$\bar{\delta}^\pm(t, i) = -\frac{\Lambda_i^\pm(\bar{\delta}^\pm(t, i))}{\Lambda_i^{\pm'}(\bar{\delta}^\pm(t, i))} \pm \mu(T - t). \quad (1.4.5.5)$$

*If additionally  $\Lambda_i^\pm(\delta) = a_i^\pm e^{-b_i^\pm \delta}$ , with  $a_i^\pm, b_i^\pm > 0$  for each  $1 \leq i \leq k$ , then*

$$\bar{\delta}^\pm(t, i) = \delta_* \vee \left( \frac{1}{b_i^\pm} \pm \mu(T - t) \right) \wedge \delta^*.$$

In other words, if  $N^* = \infty$  and  $\gamma = \zeta = 0 \equiv \ell$ , then solving (1.4.1.3) is no longer necessary and one only needs to solve the fixed point equations (1.4.5.5). Further, the MM's spreads do not depend on  $n$ . This is to be expected, since in this case she is neutral to all types of inventory risks and the terminal execution cost is neglected. However, she does need to re-adjust for changes in the regime as these impact on her probability of getting orders. If additionally, the orders intensities are exponential, then the spreads are computed straightforwardly avoiding the need of any numerical scheme. Note also how this simple model makes evident a second component of the optimal spreads through equation (1.4.5.5): a drift adjustment by which the MM takes into account the overall tendency of the asset's price.

Now we look at the results obtained for exponential intensities in general, and in particular for  $N^* < +\infty$ . Equation (1.4.5.3) becomes an explicit formula in this case, as the percentage liquidity sensitivities,  $\Lambda_i^{\pm'}/\Lambda_i^\pm$ , are constant.

**Corollary 1.4.5.4.** *If  $\Lambda_i^\pm(\delta) = a_i^\pm e^{-b_i^\pm \delta}$ , with  $a_i^\pm, b_i^\pm > 0$  for each  $1 \leq i \leq k$ , then*

$$\bar{\delta}^\pm(t, n, i) = \delta_* \vee \left( -U_\gamma^{-1}(-1/b_i^\pm) - (\Theta(t, n \mp 1, i) - \Theta(t, n, i)) \right) \wedge \delta^*.$$

This means that for exponential intensities in general, it is no longer necessary to solve any fixed point equations but only the system of ODEs (1.4.1.3). The latter still needs to be solved numerically in general. (Alternatively, see [Gué17] for some asymptotic approximations.)

*Remark.* Apart from the case  $N^* = +\infty$ ,  $\gamma = \zeta = 0 \equiv \ell$ , a noteworthy scenario in which (1.4.1.3) can be further simplified to a linear system of ODEs (with constant coefficients) is when  $k = 1$ ,  $N^* < +\infty$ ,  $\delta_* = -\infty$ ,  $\delta^* = +\infty$  and  $\Lambda^\pm(\delta) = ae^{-b\delta}$ , with  $a, b > 0$ . The reduction is achieved via the transformation  $\Theta(t, n) = \frac{1}{b} \log \Psi(t, n)$ . See [GLFT13] for more details.

## 1.5 Numerical analysis

In this section we present our numerical results, focusing on the difference in optimal behaviours between the partial and full information frameworks, and the intuition behind the

filter. To exemplify our findings in a concrete manner, while keeping the presentation as simple as possible, we will assume throughout this section that:

- The MM has risk aversion parameter  $\gamma = 0$ .
- There are only two possible states: state 1 represents a ‘bad’ regime with low liquidity taken by clients, and state 2 a ‘good’ regime with high liquidity.
- The transition rate matrix  $Q$  is constant.
- The intensities are symmetric, proportional and exponential, i.e.,  $\Lambda_1^\pm(\delta) = \Lambda_1(\delta) = ae^{-b\delta}$  and  $\Lambda_2^\pm(\delta) = \Lambda_2(\delta) = m\Lambda_1(\delta)$ , with  $a, b > 0$ ,  $m > 1$ .

The latter assumption allows us to perform the optimizations in equation (1.3.1.5) analytically. Practically, proportional intensities mean that while there is more active trading on the good regime, the way in which the clients react to movements in the spreads remains unaffected. As in [CJ15], we will allow for the three type of penalties to manage inventory risk: constraints on the maximum long and short positions, accumulated inventory penalty and a quadratic (possibly negligible) terminal penalty (or cost) for the MM. That is,

- $N^* < +\infty$ ,  $\zeta \geq 0$  and  $\ell(n) = cn^2$ , for some  $c \geq 0$ .

Let us write  $\pi := \pi_1$  for the conditional probability of being in the bad regime given the observable information. Note that  $\pi$  is a scalar in this section, since  $\pi_2 = 1 - \pi_1$  neglects the need of the additional variable. The PIDE (1.3.1.5) at the point  $(t, n, \pi)$  reads:

$$0 = \theta_t + \mu n - \frac{1}{2}\sigma^2 n^2 \zeta + 2\hat{q}(\pi)\theta_\pi + \frac{a}{b}\hat{m}(\pi) \left( \mathbb{1}_{\{n < N^*\}} e^{-b\widehat{\delta}^-} + \mathbb{1}_{\{n > -N^*\}} e^{-b\widehat{\delta}^+} \right), \quad (1.5.0.1)$$

with terminal condition  $\theta(T, n, \pi) = -cn^2$  and partial information optimal spreads given by

$$\widehat{\delta}^\pm(t, n, \pi) = \frac{1}{b} - 2\frac{w(\pi)}{\hat{m}(\pi)}\theta_\pi(t, n, \pi) - (\theta(t, n \mp 1, \pi/\hat{m}(\pi)) - \theta(t, n, \pi)), \text{ where:} \quad (1.5.0.2)$$

- (i)  $\hat{q}(\pi) = q^{11}\pi + q^{21}(1 - \pi)$  is the observable transition rate to the bad regime.
- (ii)  $\hat{m}(\pi) = \pi + (1 - \pi)m$  is the observable intensity increase from the bad regime (as a ratio).
- (iii)  $w(\pi) = (m - 1)\pi(1 - \pi)$  is the observable variance, of the square root, of the percentage intensity increase from the bad regime; i.e., a measure of observable order flow volatility.

The previous equations are valid for  $\delta_* \ll 0 \ll \delta^*$ , or more precisely, for  $\delta_*, \delta^*$  such that  $\delta_* \leq \widehat{\delta}^\pm \leq \delta^*$  holds true over the whole domain. Otherwise, one needs to floor and cap the optimal spreads and change the Hamiltonians accordingly, as done in Proposition 1.4.5.2.

A finite differences scheme of simple implementation to solve (1.5.0.1) consists of reversing time and using an explicit upwind Euler scheme over a uniform grid for  $(t, \pi)$  and  $n = -N^*, -N^* + 1, \dots, N^*$ . The terms of the form  $\theta(t, n \mp 1, \pi/\hat{m}(\pi))$ , where  $\pi/\hat{m}(\pi)$  will typically fall outside of the grid, can be approximated by linear interpolation as in [DF14]. The limiting equations can be used for  $\pi \rightarrow 0^+$  and  $\pi \rightarrow 1^-$  provided that  $-q^{11} > 0$ ,

which we assume henceforth. Such a scheme can be shown to be consistent, stable and monotone under an appropriate CFL condition. In light of equation (1.3.1.5) satisfying a comparison principle (see [CEFS16a, Thm.5.3] and why it holds for our model in the proof of Theorem 1.3.1.3), we know the scheme converges to the unique continuous viscosity solution [BS91], and we can recover the expected penalized P&L of the MM (Theorems 1.3.1.1 and 1.3.1.3). The optimal strategy to be followed by the MM is then given in feedback form by  $(\widehat{\delta}^-(t, N_{t-}, \Pi_{t-}), \widehat{\delta}^+(t, N_{t-}, \Pi_{t-}))$ .<sup>18</sup>

We will focus our attention on the optimal ask spread, with analogous observations holding for the bid spread. The parameter values in Table 1.1 were used for all the experiments presented in this section. For those present in the classical one regime models, we have chosen values used in previous works (in the  $\gamma = 0$  case [CJ15]) to make the comparison clearer. In particular, the value of  $c$  will be taken as either 0 or 0.01, to be further specified in each experiment. The time horizon will always be the one displayed on the corresponding axis. Note that we work in a symmetric market, which justifies analyzing one side only.

Parameter	$\mu$	$\sigma$	$\zeta$	$a$	$b$	$N^*$	m	$-q^{11} = q^{21}$	$-\delta_* = \delta^*$
Value	0	0.1	0.1	2	25	3	5	5	10

Table 1.1: Standing parameter values for numerical tests presented.

### 1.5.1 Comparing full and partial information optimal strategies

Under the standing assumptions of this section, the full information equation (1.4.1.3), for a function  $\tilde{\theta}$ , becomes:

$$0 = \tilde{\theta}_t + \mu n - \frac{1}{2}\sigma^2 n^2 \zeta + \tilde{q}_i \left( \tilde{\theta}(t, n, 1) - \tilde{\theta}(t, n, 2) \right) + \frac{a}{b} \tilde{m}_i \left( \mathbb{1}_{\{n < N^*\}} e^{-b\tilde{\delta}^-} + \mathbb{1}_{\{n > -N^*\}} e^{-b\tilde{\delta}^+} \right), \quad (1.5.1.1)$$

with terminal condition  $\tilde{\theta}(T, n, i) = -cn^2$  and full information optimal spreads given by

$$\tilde{\delta}^\pm(t, n, i) = \frac{1}{b} - (\tilde{\theta}(t, n \mp 1, i) - \tilde{\theta}(t, n, i)), \text{ where:} \quad (1.5.1.2)$$

- (i)  $\tilde{q}_i = q^{11} \mathbb{1}_{\{1\}}(i) + q^{21} \mathbb{1}_{\{2\}}(i)$  is the effective transition rate to the bad regime.
- (ii)  $\tilde{m}_i = \mathbb{1}_{\{1\}}(i) + \mathbb{1}_{\{2\}}(i)m$  is the effective intensity increase from the bad regime (as a ratio).

The previous equations are valid for  $\delta_* \ll 0 \ll \delta^*$ , as in the partial information case.

Although similar, equation (1.5.1.1) for the bad regime  $i = 1$  (resp. good regime  $i = 2$ ) is not the limiting equation of (1.5.0.1) for  $\pi \rightarrow 1^-$  (resp.  $\pi \rightarrow 0^+$ ). Indeed, a MM with full information can expect to make a larger profit. Thus, in general,  $\tilde{\theta} > \theta$  even in these

<sup>18</sup>As previously mentioned, these are technically only candidates for optimal (or in general,  $\varepsilon$ -optimal) strategies. We do not rigorously prove their optimality character here, but merely note that well known results of discrete time dynamic programming, together with the convergence of the discrete solutions of (1.5.0.1) to the analytical one, suggest that they can be safely used as such.

extreme cases.<sup>19</sup> However, the corresponding optimal strategies do (at least approximately) agree, as Figures 1.1 and 1.2 show.

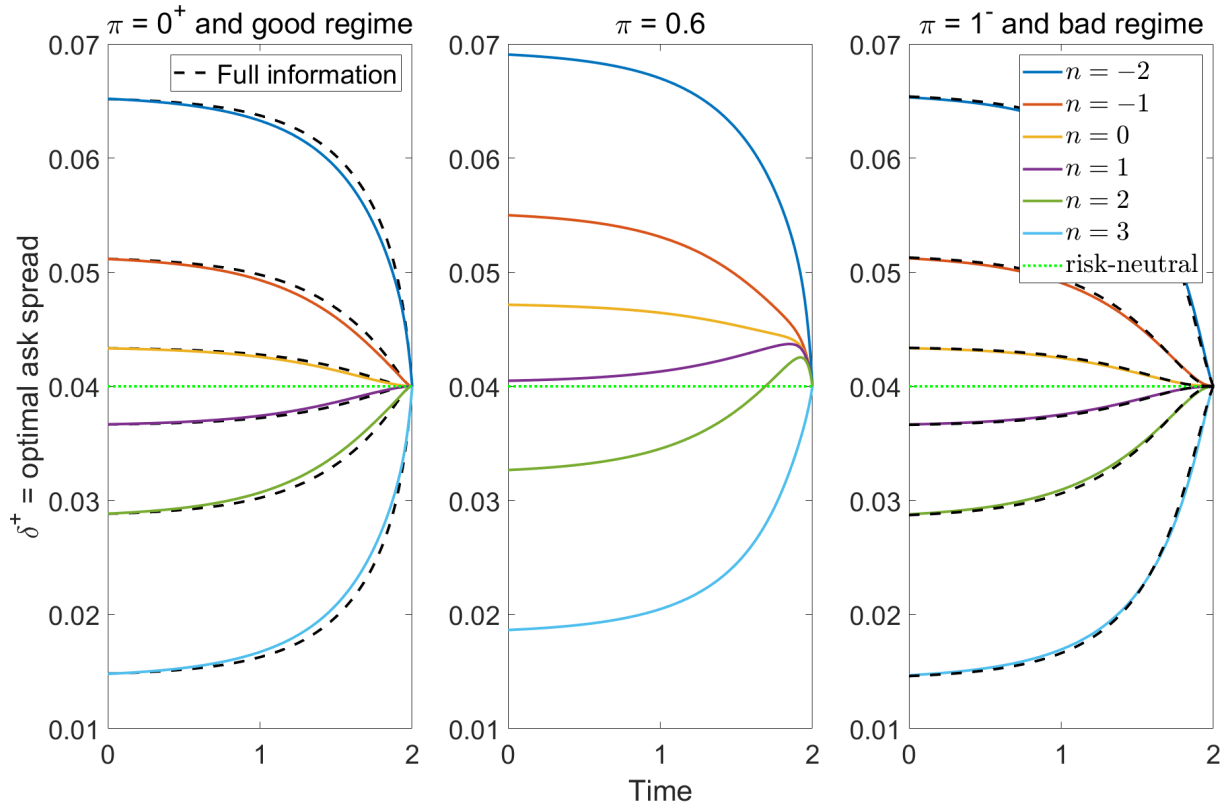


Figure 1.1: Optimal ask spread as a function of time, under full and partial information (dashed and solid lines resp.), and for different inventory and filter values. Inventory levels increase from top to bottom in all cases. Terminal execution cost is neglected ( $c = 0$ ).

Figure 1.1 shows the optimal ask spread under partial information as a function of time, for different inventory levels (solid lines). It also displays how it changes for different filter values ( $\pi = 0, 0.6$  and  $1$  from left to right) and how it compares to the optimal ask spread under full information (dashed lines) in the good regime (left) and bad regime (right). The inventory levels increase from top to bottom in all cases, and the terminal execution cost is neglected ( $c = 0$ ). We have chosen to display  $\pi = 0.6$  due to the effect of the filter being more pronounced around this value; it results in an optimal spread which is, for  $t \ll T$ , between 6% and 30% higher (depending on the inventory level) than the corresponding ones under full information (a sizeable difference in practice). We will come back to this in Section 1.5.2. All other values show similar intermediate behaviours. Recall that when the inventory reaches the minimum  $n = -N^* = -3$ , the MM will not sell any more (either abstaining from quoting or giving a ‘stub’ quote) until her inventory increases again, which is why there is no spread plotted for this position. We see that some of the features already present in one regime models are preserved for two regimes, both under full and partial information; namely:

<sup>19</sup>Our numerical findings were indeed consistent with this intuitive statement.

- For  $t \ll T$  the spread does not depend on time (approximately) and its asymptotic value becomes higher as the inventory decreases to the lower constraint  $n = -N^*$ . Indeed, for flat or short positions, any additional ask order will increase inventory risk, moving it closer to the minimum allowed and raising the accumulated inventory penalty. The MM manages this risk by increasing the spread, thus demanding a higher instantaneous profit and reducing the probability of getting executed. Similarly for long positions, the higher the exposure the lower the spread, as the MM seeks to unwind her inventory.
- In the case of negligible terminal cost, the spread converges to a terminal value independent of the inventory. This is the optimal spread of a fully risk-neutral MM with negligible costs (Corollary 1.4.5.3), who only maximizes ‘expected instantaneous margin’. (This value is given by  $\widehat{\delta}^+(T, \cdot, \cdot) = 1/b$ , the reciprocal percentage sensitivity of the liquidity to spread changes; cf. Proposition 1.4.5.2). The reason being that as the time horizon approaches, it becomes less and less likely for the MM to get executed again. The accumulated penalty cannot change too much at this point, and the risk of reaching the inventory constraints diminishes. The MM therefore takes a bit more risk and makes a last attempt at increasing her expected P&L, either by increasing execution probability (for null or short positions) or increasing her instantaneous profit if executed (for long positions).

Nevertheless, the following differences can be observed:

- Under full information (and for symmetric markets), the spreads are always symmetric in the inventory with respect the risk-neutral level, i.e.,  $1/b - \widetilde{\delta}^+(\cdot, n, \cdot) = \widetilde{\delta}^+(\cdot, 1-n, \cdot) - 1/b$ , for  $0 < n \leq N^*$ . However, under partial information this symmetry is broken, as the MM increases her spreads. This holds in fact for any  $0 < \pi < 1$  and it is suggested by the comparison of equations (1.5.0.2) and (1.5.1.2), since  $\theta_\pi < 0$  (an increasing probability of being in the bad regime lowers the expected P&L; see argument of [BL09, Lemma 3.3]). As a result, the full information spreads are downward biased when the exact regime is unknown. Loosely speaking, this bias is approximately proportional to the product of  $\theta_\pi$  (sensitivity of the expected P&L to observable regime changes) and  $w(\pi)$  (observable order flow volatility), and inversely proportional to the observable intensity increase  $\hat{m}(\pi)$ . Intuitively, a partially informed MM faces not only the risk of the market regime shifting but also uncertainty on the current state, and must increase her spreads accordingly. This is done by considering the cost of any observable changes on the P&L and the fluctuations in the order flow, and discounting by liquidity increases.
- The good and bad regimes differ in the rate of order flow at the reference price (i.e., the amplitude parameters  $a$  and  $am$  of the orders intensities). Consequently, one could erroneously think that the partial information strategy should essentially be a one regime strategy for some intermediate parameter, such as  $a\hat{m}(\pi)$ . However, in the full information framework, an increase in the liquidity taken by clients results in the asymptotic spreads (i.e., for  $t \ll T$ ) moving closer to the risk-neutral level (see [CJ15, GLFT13]). On the other hand, this is not true any more in the partial

information case, due to the regime risk adjustment previously mentioned that shifts the spreads.

- Lastly, we remark that there is a distinguished change in behaviour close to expiry, with the spreads overshooting above the risk-neutral level for some inventories, before approaching it again. Intuitively, the overshooting is due to the added regime risk adjustment, and the convergence is due to the vanishing of additional inventory risks. (Recall that the MM's main concern becomes her terminal hedging cost, should there be one, and this does not depend on the market regime.) A somewhat similar effect can be found under full information in asymmetric markets [CJ15].

Figure 1.2 serves the same comparison as Figure 1.1, but including in this case a terminal execution cost. The observations remain mostly the same, except for a change in the terminal behaviour of the spreads, just as in the classical one regime case. Here, the spreads diverge from the risk-neutral no-costs value instead of approaching it, as the MM makes a last instant attempt to cover her hedging cost.

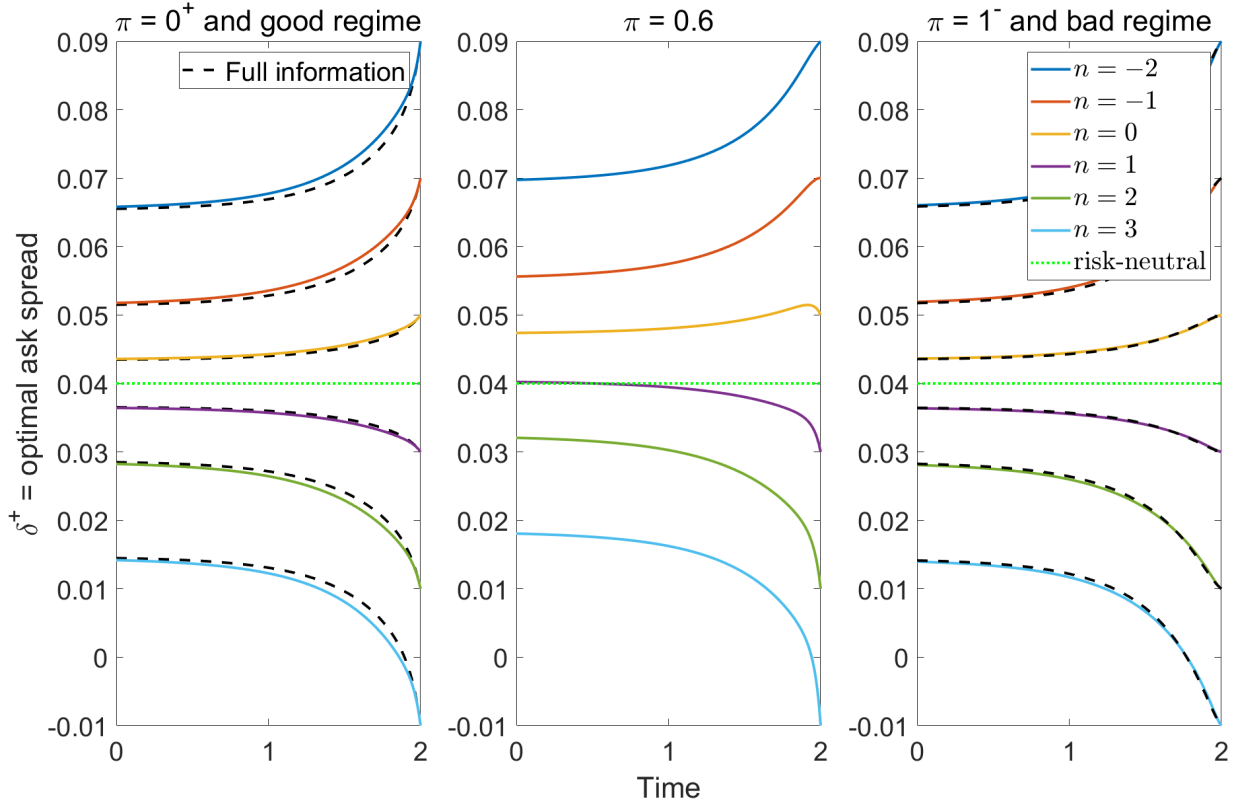


Figure 1.2: Optimal ask spread as a function of time, under full and partial information (dashed and solid lines resp.), and for different inventory and filter values. Inventory levels increase from top to bottom in all cases. Terminal execution cost:  $\ell(n) = -cn^2$ , with  $c = 0.01$ .

Figure 1.3 shows the partial information optimal ask spread as a function of time and filter, for three different inventory positions: extreme short (left), flat (center) and extreme



long (right), and no terminal penalty. We can see that the spread is concave on the filter, and the maximum concavity is reached for a null inventory and  $\pi \approx 0.6$ , gradually decreasing as the position becomes short or long. We had observed already in equation (1.5.0.2) that the partially informed MM increases her spreads to manage higher regime risk, and that this correction is approximately proportional to the expected P&L sensitivity,  $\theta_\pi$ , times the observed order flow volatility  $w(\pi) = (m-1)\pi(1-\pi)$ . Heuristically, it seems reasonable that the change in spread should resemble the concavity of  $w(\pi)$  for a risk averse MM, and that the cost of regime uncertainty is the highest for a flat position (since deviations from it increase in turn price exposure). The question of why the maximum variation is reached around  $\pi \approx 0.6$  is once again deferred to Section 1.5.2.

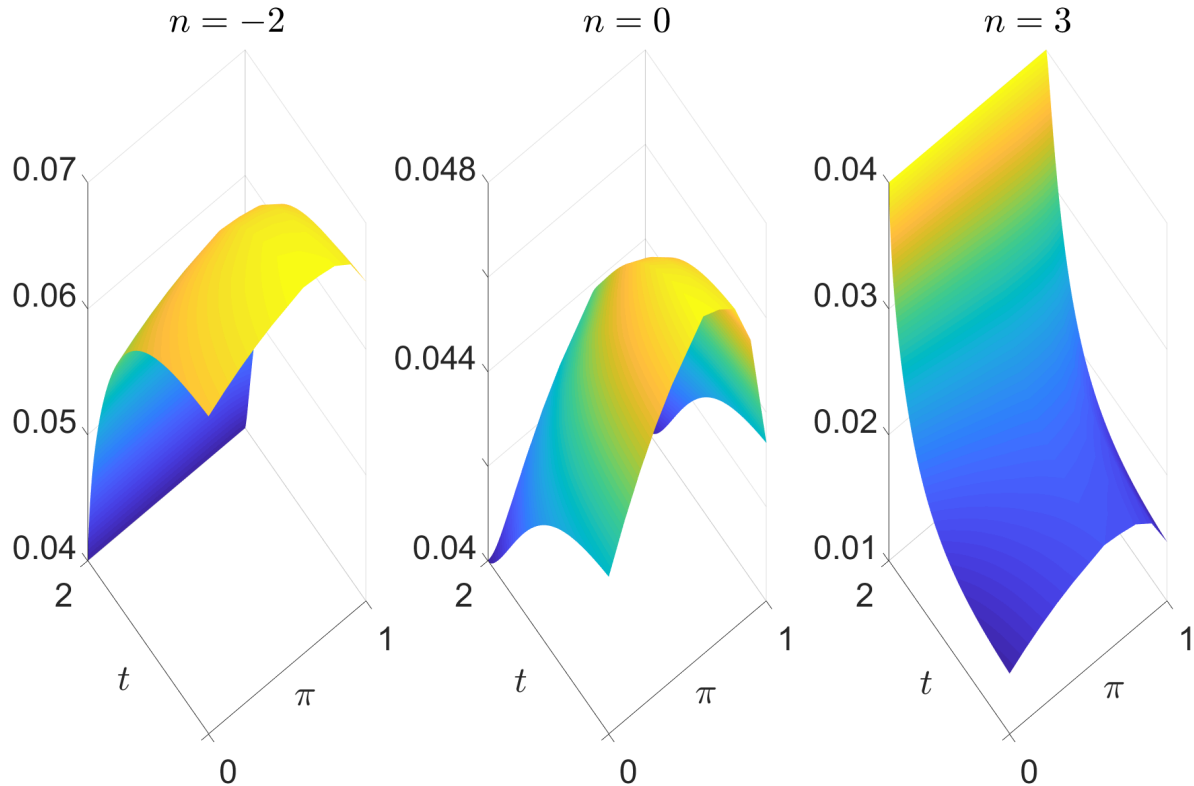


Figure 1.3: Optimal partial information ask spread as a function of time and filter (observable probability of bad regime), for different inventory levels. Terminal execution cost is neglected ( $c = 0$ ).

### 1.5.2 Sample paths and a closer look at the filter

There are different ways to simulate point processes with stochastic intensity. A classical approach of simple implementation is the so-called *thinning method* [Oga81]. This method is particularly well suited to our framework, as it can be combined with the filtering theory developed in Section 1.2 by making use of the observable intensity of  $N$  (see Theorem 1.3.1.1 and [GKM11, Alg.3.2]). The interested reader is referred to [LTT18] for optimizations in terms of the thinning bound.

Figure 1.4 shows four sample paths resulting from an optimal behaviour of the MM with incomplete information. They were obtained by jointly simulating the inventory  $N$  (middle) starting from  $n_0 = 0$ , the filter  $\Pi = \Pi^{\alpha,1}$  (bottom) starting from  $\pi_0 = 0.5$ , and the optimal strategy  $\alpha = (\widehat{\delta}^-(t, N_{t-}, \Pi_{t-}), \widehat{\delta}^+(t, N_{t-}, \Pi_{t-}))$  (top), for  $c = 0$ .

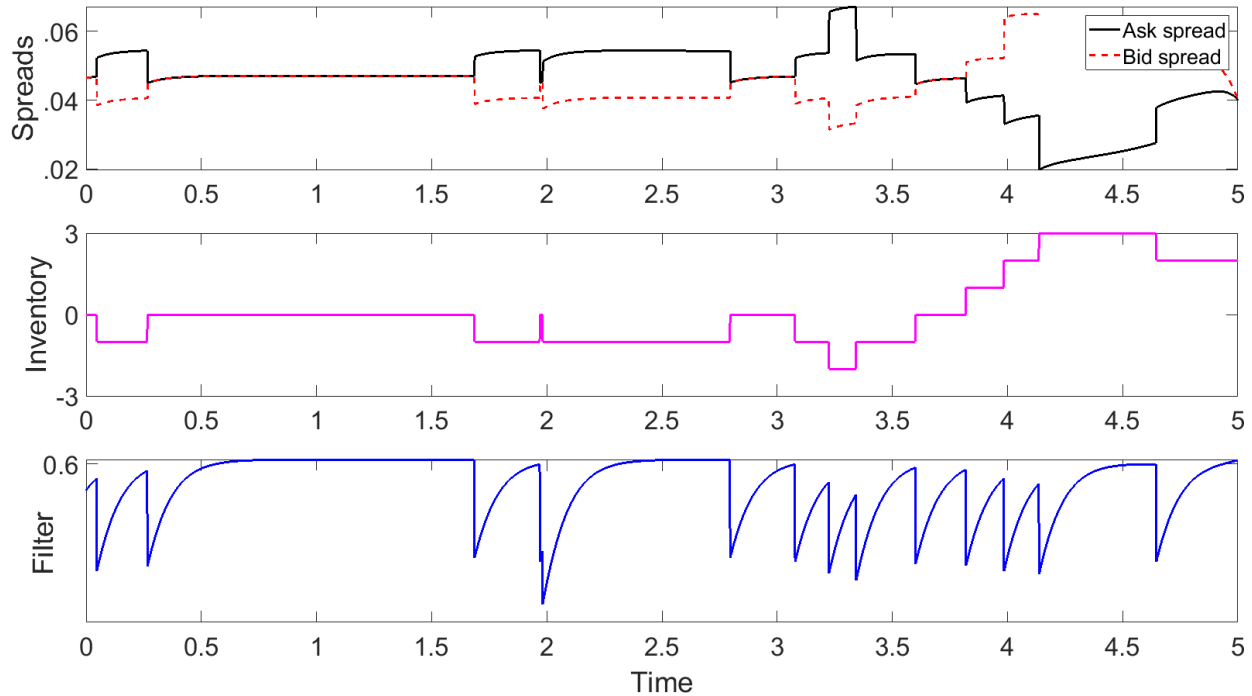


Figure 1.4: Sample paths of optimal partial information bid/ask spreads (first), inventory (second) and filter (third), with  $c = 0$ ,  $n_0 = 0$  and  $\pi_0 = 0.5$ .

We want to analyze the behaviour of  $\Pi$ . Let us recall that  $\Pi_t$  represents the probability of being in the bad regime (a slow market with low liquidity) given the information observed by the MM at time  $t$ . The MM makes her assessment based on the orders (both buy and sell) that she received so far, which is the same as looking at the evolution of her inventory. If we take a close look at what happens between any two consecutive orders, we see that there is a pattern that repeats itself. As time passes without the MM receiving any order, she deems more likely that the market is in the bad regime, and so  $\Pi_t$  increases. If enough time goes by in this way, then  $\Pi_t$  gets asymptotically closer to some preponderant ‘equilibrium’ value which is around 0.6. (Recall that this is the same value we have encountered before as having the greatest effect on the spreads.) But as soon as an order arrives, the MM revisits her probabilities, as being in the good regime of high liquidity seems a lot more likely. This is why we see that  $\Pi_t$  jumps downwards with each trade (i.e., at the jump times of the inventory).

So what makes  $\pi \approx 0.6$  so special? If we look at the SDE (1.2.1.2) governing the filter dynamics, it now reads:

$$d\Pi_t = \left( \hat{q}(\Pi_t) + w(\Pi_t)a(\mathbb{1}_t^- e^{-\widehat{b\delta_t^-}} + \mathbb{1}_t^+ e^{-\widehat{b\delta_t^+}}) \right) dt + \Pi_{t-} \left( \frac{1}{\hat{m}(\Pi_{t-})} - 1 \right) (dN_t^- + dN_t^+). \quad (1.5.2.1)$$

In between jumps, each path of  $\Pi_t$  evolves according to an ODE. If enough time passes by without another order arriving, we can consider the asymptotic behaviour for  $t \rightarrow T$ . For  $c = 0$ , this leads to

$$\frac{d\Pi_t}{dt} \approx \hat{q}(\Pi_t) + w(\Pi_t)\frac{a}{e}\beta, \quad \text{with } \beta = 1 \text{ or } \beta = 2,$$

depending on the current inventory level. The right-hand side of this ODE is a concave parabola in  $\Pi_t$  with a negative root and another one,  $\pi^*$ , between 0 and 1. For the parameter values in Table 1.1, it is either  $\pi^* \approx 0.572$  for  $\beta = 1$ , or  $\pi^* \approx 0.636$  for  $\beta = 2$ . Therefore,  $\pi^*$  is an attractor (i.e., a stable equilibrium) for the filter, which explains the tendency of  $\Pi_t$  towards 0.6 and the central role of this value. Indeed, as the filter approaches one of its attractors, the MM exhausts all the information she has and can make no further assessment on the state of the market until she is hit by another order. In a sense, this is when she faces the most uncertainty, and she increases her spreads accordingly to manage regime risk.

# Chapter 2

## A policy iteration algorithm for nonzero-sum stochastic impulse games

### Introduction

Stochastic differential games model the interaction between players whose objective functions depend on the evolution of a certain continuous-time stochastic process. The subclass of impulse games focuses on the case where the players only act at discrete (usually random) points in time by shifting the process. In doing so, each of them incurs into costs and possibly generates ‘gains’ (not necessarily nonnegative) for the others at the same time. They constitute a generalization of the well-known (single-player) optimal impulse control problems [ØS09, Chpt.7-10], which have found a wide range of applications in finance, energy markets and insurance [Bré81, CCTZ06, EH88, Kor99], among plenty of other fields.

From a deterministic numerical viewpoint, an impulse control problem entails the resolution of a differential quasi-variational inequality (QVI) to compute the value function and, when possible, retrieve an optimal strategy. Policy-iteration-type algorithms [AF16, CØS02, CMS07] undoubtedly occupy an ubiquitous place in this respect, and more so in the infinite horizon case.

The presence of a second player makes matters much more challenging, as one needs to find two optimal (or *equilibrium*) payoffs dependent on one another, and the optimal strategies take the form of *Nash equilibria* (NEs), i.e., strategies such that no player can benefit from unilaterally changing her own. And while impulse controls give a more realistic setting than ‘continuous’ controls in applications as the aforementioned ones, they normally lead to less tractable and very technical models.

It is not surprising then, that the literature in impulse games is limited and mainly focused on the zero-sum case [Azi19b, Cos13, EAM18]: when the cost for one player equals the gain for the other, and the same is true for the objective functions. The more general and flexible nonzero-sum instance, which dispenses with the former constraints, has only recently began to receive more attention. The authors of [ABC<sup>+</sup>19] consider for the first time a general two-player game where both participants act through impulse controls,<sup>1</sup> and

---

<sup>1</sup> [CWW13, WW17] also consider nonzero-sum impulse games but assuming the intervention times of the players are known from the outset.

characterize certain type of equilibrium payoffs and NEs via a system of QVIs by means of a Verification Theorem. Using this result, they provide the first example of an (almost) fully analytically solvable game, motivated by central banks competition over the exchange rate. The result is generalized to arbitrary  $N$  players in [BCG19], which also gives a semi-analytical solution (i.e., depending on several parameters found numerically) to a concrete cash management problem.<sup>2</sup> A different, more probabilistic, approach is taken in [FK19] to find a semi-analytical solution of a strategic pollution control problem and to prove another Verification Theorem.

The previous examples, and the lack of others,<sup>3</sup> give testimony of how difficult it is to explicitly solve nonzero-sum impulse games. The analytical approaches require an educated guess to start with and (with the exception of the henceforth referred to as *the linear game* in [ABC<sup>+</sup>19]) several parameters need to be solved for in general from highly-nonlinear systems of equations coupled with order conditions. All of this can be very difficult, if not prohibitive, when the structure of the game is not simple enough. Further, all of them (as well as the majority of the concrete examples in the impulse control literature) assume linear costs. In general, for nonlinear costs, the state to which each player wants to shift the process when intervening is not unique, but depends on the starting point. This effectively means that infinite parameters may need to be solved for, drastically discouraging this methodology.

While the need for numerical schemes able to handle nonzero-sum impulse games is obvious, unlike in the single-player case, this is an utterly underdeveloped line of research. Focusing on the purely deterministic approach, solving the system of QVIs derived in [ABC<sup>+</sup>19] involves handling coupled free boundary problems, further complicated by the presence of nonlinear, nonlocal and noncontractive operators. Additionally, solutions will typically be irregular even in the simplest cases such as the linear game. Moreover, the absence of a viscosity solutions framework such as that of impulse control [Sey09] means that it is not possible to know whether the system of QVIs should have a solution or not (not to mention some form of uniqueness) unless one can explicitly solve it. This is further exacerbated by the fact that even defining such system requires a priori assumptions on the solutions (the *unique impulse property*). This is also the case in [FK19].

In this chapter (based on [ABM<sup>+</sup>19, Zab19]), we make the first attempt (to the best of our knowledge) at numerically solving nonzero-sum impulse games. In a nutshell, the iterative algorithm we put forward treats the game at each iteration as a combination of a fixed point problem and a slowly relaxing single-player impulse control problem. This allows us to take advantage of the machinery for the latter. Because no convergence analysis is provided in this chapter, the proposed algorithm is admittedly heuristic. Instead, we report on a range of numerical experiments which show convergence on fixed grids as well as convergence of the error with respect to the discretization. In fact, errors are quite satisfactory and, in the examples we have tackled, the relative errors easily drop well below 0.1%. The algorithm can thus be used to assist further development of the field, as well as to gain insight into applications modelled by nonzero-sum impulse games. In Chapter 3, which focuses on *symmetric* games, a modified algorithm will be presented together with a

---

<sup>2</sup> [BCG19] also studies the mean field limit game.

<sup>3</sup> [ABC<sup>+</sup>19] also gives semi-analytical solutions to modifications of the linear game when changing the payoffs in a non-symmetric way. To the best of the author's knowledge, these are the collection of examples available at the time of writing.

rigorous convergence analysis.

Let us outline the organization of the rest of the chapter. We start in Section 2.1 by properly setting up nonzero-sum stochastic impulse games and recalling the main result of [ABC<sup>+</sup>19], namely, the Verification Theorem with the corresponding system of QVIs. For the sake of completeness, the linear game and the underlying problem of competing central banks are also presented. Section 2.2 motivates, lists, and discusses our algorithm. Section 2.3 presents numerical evidence supporting it, in the light of which we draw some conclusions in Section 2.4.

## 2.1 Analytical continuous-space problem

In this section we start by reviewing a general formulation of two-player nonzero-sum stochastic differential games with impulse controls, as considered in [ABC<sup>+</sup>19], together with the main theoretical result of the authors: a characterization of certain NEs via a deterministic system of QVIs. The indexes of the players are denoted  $i, j \in \{1, 2\}$  with  $i \neq j$ . Since no other type of games is considered in this thesis, we will often speak simply of ‘games’ for brevity.

Throughout the thesis, we restrict our attention to the one-dimensional infinite-horizon case. Some of the most technical details, concerning the well-posedness of the model, are left out for the sake of brevity and can be found in [ABC<sup>+</sup>19, Sect.2].

### 2.1.1 General two-player nonzero-sum impulse games

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space under the usual conditions supporting a standard one-dimensional Wiener process  $W$ . We consider two players that observe the evolution of a *state variable*  $X$ , modifying it when convenient through *controls* of the form  $u_i = \{(\tau_i^k, \delta_i^k)\}_{k=1}^\infty$  for  $i = 1, 2$ . The stopping times  $(\tau_i^k)$  are their *intervention times* and the  $\mathcal{F}_{\tau_i^k}$ -measurable random variables  $(\delta_i^k)$  are their *intervention impulses*. Given controls  $(u_1, u_2)$  and a starting point  $X_{0-} = x \in \mathbb{R}$ , we assume  $X = X^{x; u_1, u_2}$  has dynamics

$$X_t = x + \int_0^t \mu(X_s) ds + \int_0^t \sigma(X_s) dW_s + \sum_{k: \tau_1^k \leq t} \delta_1^k + \sum_{k: \tau_2^k \leq t} \delta_2^k, \quad (2.1.1.1)$$

for some given drift and volatility functions  $\mu, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ , locally Lipschitz with linear growth.<sup>4</sup>

Equation (2.1.1.1) states that  $X$  evolves as an Itô diffusion in between the intervention times, and that each intervention consists in shifting  $X$  by applying an impulse. It is assumed that the players choose their controls by means of threshold-type *strategies* of the form  $\varphi_i = (\mathcal{I}_i, \delta_i)$ , where  $\mathcal{I}_i \subseteq \mathbb{R}$  is a closed set called *intervention (or action) region* and  $\delta_i : \mathbb{R} \rightarrow \mathbb{R}$  is an *impulse function* assumed to be continuous. The complement  $\mathcal{C}_i = \mathcal{I}_i^c$  is called *continuation (or waiting) region*.<sup>5</sup> That is, player  $i$  intervenes if and only if the state variable reaches her intervention region, by applying an impulse  $\delta_i(X_{t-})$  (or equivalently, shifting  $X_{t-}$  to

<sup>4</sup>See [ABC<sup>+</sup>19, Def.2.2] for a precise recursive definition in terms of the strategies.

<sup>5</sup>In [ABC<sup>+</sup>19], strategies are described in terms of continuation regions instead.

$X_{t-} + \delta_i(X_{t-})$ ). Further, we impose a priori constraints on the impulses: for each  $x \in \mathbb{R}$  there exists a set  $\emptyset \neq \mathcal{Z}_i(x) \subseteq \mathbb{R}$  such that  $\delta_i(x) \in \mathcal{Z}_i(x)$  if  $x \in \mathcal{I}_i$ .<sup>6</sup> We also assume the game has no end and player 1 has the priority should they both want to intervene at the same time.

Given a starting point and a pair strategies, the (expected) *payoff* of player  $i$  is given by

$$J_i(x; \varphi_1, \varphi_2) := \mathbb{E} \left[ \int_0^\infty e^{-\rho_i s} f_i(X_s) ds - \sum_{k=1}^\infty e^{-\rho_i \tau_i^k} c_i(X_{(\tau_i^k)^-}, \delta_i^k) + \sum_{k=1}^\infty e^{-\rho_i \tau_j^k} g_i(X_{(\tau_j^k)^-}, \delta_j^k) \right],$$

with  $X = X^{x; u_1, u_2} = X^{x; \varphi_1, \varphi_2}$ . For player  $i$ ,  $\rho_i > 0$  represents her (subjective) *discount rate*,  $f_i : \mathbb{R} \rightarrow \mathbb{R}$  her *running payoff*,  $c_i : \mathbb{R}^2 \rightarrow (0, +\infty)$  her *cost of intervention* and  $g_i : \mathbb{R}^2 \rightarrow \mathbb{R}$  her *gain due to her opponent's intervention* (not necessarily non-negative). The functions  $f_i, c_i, g_i$  are assumed to be continuous.

Throughout this thesis, only admissible strategies are considered. Briefly,  $(\varphi_1, \varphi_2)$  is *admissible* if it gives well-defined payoffs for all  $x \in \mathbb{R}$ ,  $\|X\|_\infty$  has finite moments and, although each player can intervene immediately after the other, infinite simultaneous interventions are precluded.<sup>7</sup> As an example, if the running payoffs have polynomial growth, the ‘never intervene strategies’  $\varphi_1 = \varphi_2 = (\emptyset, \emptyset \hookrightarrow \mathbb{R})$  are admissible and the game can be played.

Given a game, we want to know whether it admits some Nash equilibrium and how to compute it. Recall that a pair of strategies  $(\varphi_1^*, \varphi_2^*)$  is a *Nash equilibrium* (NE) if for every admissible  $(\varphi_1, \varphi_2)$ ,

$$J_1(x; \varphi_1^*, \varphi_2^*) \geq J_1(x; \varphi_1, \varphi_2^*) \quad \text{and} \quad J_2(x; \varphi_1^*, \varphi_2^*) \geq J_2(x; \varphi_1^*, \varphi_2),$$

i.e., no player can gain from a unilateral change of strategy. If one such NE exists, we refer to  $(V_1, V_2)$ , with  $V_i(x) = J_i(x; \varphi_1^*, \varphi_2^*)$ , as a pair of *equilibrium payoffs*.

### 2.1.2 General system of quasi-variational inequalities

To present the system of QVIs derived in [ABC<sup>+</sup>19], we need to define first the *intervention operators*. For any  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  and  $x \in \mathbb{R}$ , the *loss operator* of player  $i$  is defined as

$$\mathcal{M}_i V_i(x) := \sup_{\delta \in \mathcal{Z}_i(x)} \{V_i(x + \delta) - c_i(x, \delta)\}.^8 \quad (2.1.2.1)$$

When applied to an equilibrium payoff, the loss operator  $\mathcal{M}_i$  gives a recomputed present value for player  $i$  due to the cost of her own intervention. Given the optimality of the NEs, one would intuitively expect that  $\mathcal{M}_i V_i \leq V_i$  for equilibrium payoffs and that the equality is attained only when it is optimal for player  $i$  to intervene. Under this logic:

<sup>6</sup>In [ABC<sup>+</sup>19],  $\mathcal{Z}_i(x)$  is the same for every  $x \in \mathbb{R}$ . The generalization in this chapter is a standard one in impulse control and will prove useful in the sequel. The results in [ABC<sup>+</sup>19] still hold with the same proofs.

<sup>7</sup>More precisely, these would be  $\mathbb{R}$ -admissible strategies. See [ABC<sup>+</sup>19, Def.2.5] for more details.

<sup>8</sup>Although we could have  $\mathcal{M}_i V_i(x) = +\infty$ , this will be excluded when enforcing the system of QVIs (2.1.2.3).

**Definition 2.1.2.1.** We say that the pair  $(V_1, V_2)$  has the *unique impulse property (UIP)* if for each  $i = 1, 2$  and  $x \in \{\mathcal{M}_i V_i = V_i\}$ , there exists a unique impulse, denoted  $\delta_i^*(x) = \delta_i^*(V_i)(x) \in \mathcal{Z}_i(x)$ , that realizes the supremum in (2.1.2.1).<sup>9</sup>

If  $(V_1, V_2)$  enjoys the UIP, we define the *gain operator* of player  $i$  as

$$\mathcal{H}_i V_i(x) := V_i(x + \delta_i^*(x)) + g_i(x, \delta_i^*(x)), \quad \text{for } x \in \{\mathcal{M}_j V_j = V_j\} \quad (2.1.2.2)$$

When applied to equilibrium payoffs, the gain operator  $\mathcal{H}_i$  gives a recomputed present value for player  $i$  due to her opponent's intervention.

Finally, let us denote by  $\mathcal{A}$  the infinitesimal generator of  $X$  when uncontrolled, i.e.,

$$\mathcal{A}V(x) := \frac{1}{2}\sigma^2(x)V''(x) + \mu(x)V'(x),$$

for any  $V : \mathbb{R} \rightarrow \mathbb{R}$  which is  $C^2$  at some open neighborhood of a given  $x \in \mathbb{R}$ . We assume this regularity holds whenever we compute  $\mathcal{A}V(x)$  for some  $V$  and  $x$ . The following Verification Theorem, due to [ABC<sup>+</sup>19, Thm.3.3], states that if a regular enough solution  $(V_1, V_2)$  to a certain system of QVIs exists, then it must be a pair of equilibrium payoffs, and a corresponding NE can be retrieved. We state here a simplified version that applies to the one-dimensional infinite-horizon games at hand.<sup>10</sup>

**Theorem 2.1.2.2 (General system of QVIs).** *Given a game as in Section 2.1.1, let  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  be pair of functions with the UIP, such that for any  $i, j \in \{1, 2\}$ ,  $i \neq j$ :*

$$\begin{cases} \mathcal{M}_j V_j - V_j \leq 0 & \text{on } \mathbb{R} \\ \mathcal{H}_i V_i - V_i = 0 & \text{on } \{\mathcal{M}_j V_j - V_j = 0\} =: \mathcal{I}_j^* \\ \max \{\mathcal{A}V_i - \rho_i V_i + f_i, \mathcal{M}_i V_i - V_i\} = 0 & \text{on } \{\mathcal{M}_j V_j - V_j < 0\} =: \mathcal{C}_j^* \end{cases} \quad (2.1.2.3)$$

and  $V_i \in C^2(\mathcal{C}_j^* \setminus \partial \mathcal{C}_i^*) \cap C^1(\mathcal{C}_j^*) \cap C(\mathbb{R})$  has polynomial growth and bounded second derivative on some reduced neighbourhood of  $\partial \mathcal{C}_i^*$ . Suppose further  $((\mathcal{I}_i^*, \delta_i^*))_{i=1,2}$  are admissible strategies.<sup>11</sup>

Then,  $(V_1, V_2)$  are equilibrium payoffs attained at a NE  $((\mathcal{I}_i^*, \delta_i^*))_{i=1,2}$ .

The first equation of system (2.1.2.3) states that at an equilibrium, a player cannot increase her own payoff by a unilateral intervention. One therefore expects that the equality  $\mathcal{M}_j V_j = V_j$  will only hold when player  $j$  intervenes, or in other words, when the value she gains can exactly compensate the cost of her intervention. Consequently, the second equation says that a gain results from the opponent's intervention. Finally, the last one,

<sup>9</sup>We do not require the UIP to hold outside of  $\{\mathcal{M}_i V_i = V_i\}$ , as this is not the case for equilibrium payoffs in many examples, such as the linear game with constant costs/gains. Proofs in [ABC<sup>+</sup>19] carry through unaltered.

<sup>10</sup>Unlike in [ABC<sup>+</sup>19], there is no terminal condition in the system of QVIs and the assumption that  $\partial \mathcal{C}_i^*$  be a Lipschitz surface is trivially satisfied for an open  $\mathcal{C}_i^* \subseteq \mathbb{R}$ , as it is a countable union of disjoint open intervals.

<sup>11</sup>For consistency with the strategies' definition, one should assume that  $\delta^*$  has been continuously extended to  $\mathbb{R}$ . The conclusion is unaffected by the choice of the extension.



means that when the opponent does not intervene, each player faces a single-player impulse control problem.

We conclude this section with some final observations that we will come back to in Chapter 3:

*Remark 2.1.2.3.* An immediate consequence of assuming strictly positive costs is that intervening at any state with a null impulse reduces the payoff of the acting player and is therefore suboptimal. This is also displayed in system (2.1.2.3): if at some state  $x$ ,  $\mathcal{M}_j V_j(x)$  was realized for  $\delta = 0$ , then  $\mathcal{M}_j V_j(x) = V_j(x) + c_j(x, 0) < V_j(x)$ . At the same time, allowing for vanishing costs often leads to degenerate games in the current framework [ABC<sup>+</sup>19, Sect.4.4]. Hence, assuming  $c_j > 0$  is quite reasonable.

*Remark 2.1.2.4.* Consider the case of nonnegative impulses and cost functions being strictly concave in the impulse, as in [Cos13]. (Concave costs are also assumed in [EAM18].) That is,  $c_i(x, \delta + \bar{\delta}) < c_i(x, \delta) + c_i(x + \delta, \bar{\delta})$  for all  $x \in \mathbb{R}$ ,  $\delta, \bar{\delta} \geq 0$ . This models the situation in which simultaneous interventions are more expensive than a single one to the same effect. In such cases, it is easy to see that in the context of Theorem 2.1.2.2, player  $i$  will only shift the state variable towards her continuation region.<sup>12</sup>

### 2.1.3 Linear game and motivation

For completeness, we finish this Section by explicitly recalling the linear game of [ABC<sup>+</sup>19], the only almost fully analytically solvable game in the literature at the time of writing, and the application that motivates it. Recall that the choice of a linear structure (as opposed to piecewise linear-quadratic as in the original impulse control formulations [AF16, CZ99, JP93, MØ98]) is made for tractability reasons.

**Central banks competing over the foreign exchange rate (FEX):** Consider two central banks that compete to influence the FEX between their respective currencies, both seeking to devalue (resp. appreciate) their own one. In this application, either central bank intervenes when it deems its currency too strong (resp. too weak). Each bank's strategy is made up of the ensuing devaluation (resp. appreciation) amount and the threshold FEX triggering it. Both of these are to be optimally determined in advance, by solving the game. Let the process  $X$  model the FEX, evolving as a Brownian motion when there are no interventions. In a simplified model, we assume the running payoffs of the banks are  $X - s_1$  and  $s_2 - X$  respectively ( $s_2 > s_1$ ) and that the costs/gains upon interventions have a fixed component and another one proportional to the size of the change induced. We further assume that both banks discount their winnings/losses at the same rate  $\rho > 0$  and they have the same cost/gain parameters. More specifically,

$$X_t = x + \sigma W_t + \sum_{k: \tau_1^k \leq t} \delta_1^k + \sum_{k: \tau_2^k \leq t} \delta_2^k, \quad \text{and}$$

<sup>12</sup>Let  $x \in \mathcal{I}_i^*$  and suppose  $y_i^* := x + \delta_i^*(x) \in \mathcal{I}_i^*$ . Set  $y_i^{**} := y_i^* + \delta_i^*(y_i^*)$ . Then, by the UIP, the definitions of  $\delta_i^*(x)$  and  $\mathcal{I}_i^*$ , and the concavity of the cost:  $V_i(y_i^{**}) - c_i(x, \delta_i^*(x) + \delta_i^*(y_i^*)) \leq V_i(y_i^*) - c_i(x, \delta_i^*(x)) = V_i(y_i^{**}) - c_i(y_i^*, \delta_i^*(y_i^*)) - c_i(x, \delta_i^*(x)) < V_i(y_i^{**}) - c_i(x, \delta_i^*(x) + \delta_i^*(y_i^*))$ , which is a contradiction.

$$J_i(x; \varphi_1, \varphi_2) := \mathbb{E}_x \left[ \int_0^\infty e^{-\rho s} (-1)^{i-1} (X_s - s_i) ds - \sum_{k=1}^\infty e^{-\rho \tau_i^k} (c + \lambda |\delta_i^k|) + \sum_{k=1}^\infty e^{-\rho \tau_j^k} (\tilde{c} + \tilde{\lambda} |\delta_j^k|) \right],$$

where  $0 \leq \tilde{c} \leq c$ ,  $0 \leq \tilde{\lambda} \leq \lambda$  and  $(c, \lambda) \neq (\tilde{c}, \tilde{\lambda})$ . These parametric restrictions ensure the existence of a NE (see [ABC<sup>+</sup>19] and Section 3.3.3 in Chapter 3).<sup>13</sup>

An exact solution of this game can be found by an application of Theorem 2.1.2.2. To this purpose, the system of QVIs (2.1.2.3) is heuristically solved, yielding candidates for equilibrium payoffs and a NE. This is done, first, by making some educated guesses. Second, by solving the ordinary differential equations (ODEs) in (2.1.2.3) where appropriate. Finally, by imposing the regularity requirements of Theorem 2.1.2.2 through pasting conditions. Upon verification of the remaining hypotheses, the following turns out to be a solution of the game:

$$V_2(x) = \begin{cases} \varphi^{A_1, A_2}(x_1^*) + \tilde{c} + \tilde{\lambda}(x_1^* - x) & \text{if } x \in (-\infty, \bar{x}_1] \\ \varphi^{A_1, A_2}(x) & \text{if } x \in (\bar{x}_1, \bar{x}_2) \\ \varphi^{A_1, A_2}(x_2^*) - c - \lambda(x - x_2^*) & \text{if } x \in [\bar{x}_2, +\infty), \end{cases} \quad V_1(x) = V_2(2\tilde{s} - x),$$

where:

$$\begin{aligned} \varphi^{A_1, A_2}(x) &= A_1 e^{\theta x} + A_2 e^{-\theta x} + \frac{1}{\rho}(s_2 - x), \\ \tilde{s} &:= \frac{s_1 + s_2}{2}, \quad \theta := \sqrt{2\rho/\sigma^2}, \quad \eta := (1 - \lambda\rho)/\rho, \\ \bar{x}_i &:= \tilde{s} + \frac{(-1)^i}{\theta} \log \left( \sqrt{\frac{\eta + \xi}{\eta - \xi}} (\sqrt{\Gamma + 1} + \sqrt{\Gamma}) \right), \\ x_i^* &:= \tilde{s} + \frac{(-1)^i}{\theta} \log \left( \sqrt{\frac{\eta - \xi}{\eta + \xi}} (\sqrt{\Gamma + 1} + \sqrt{\Gamma}) \right), \\ A_i &:= \exp \left( (-1)^i \theta \tilde{s} \right) \frac{\sqrt{\eta^2 - \xi^2}}{2\theta} \left( (-1)^{i+1} \sqrt{\Gamma + 1} - \sqrt{\Gamma} \right), \\ \Gamma &:= \frac{\theta(c - \tilde{c})}{4\xi} + \frac{\theta c(\lambda - \tilde{\lambda})}{4\eta\xi} + \frac{\lambda - \tilde{\lambda}}{2\eta}, \end{aligned}$$

and  $\xi \in [0, \eta]$  is the unique zero of  $F(y) := 2y - \eta \log \left( \frac{\eta+y}{\eta-y} \right) + \theta c$ .

## 2.2 Iterative algorithm for impulse games

Compared with the single-value-function problems of impulse control, general two-player games are distinctly more challenging. Recall that a deterministic approach based on Theorem 2.1.2.2 entails that:

---

<sup>13</sup> $1 - \rho\lambda > 0$  is also imposed in [ABC<sup>+</sup>19], but this is only necessary to preclude the case in which the players never intervene.

- two equilibrium payoffs, governed by a system of QVIs, must be solved for,
- the dependence between a pair of solutions  $(V_1, V_2)$  is highly nonlinear due to the presence of the gain operators  $\mathcal{H}_i(V_j)V_i$ ,
- the gain operators are only well-defined for pairs with the UIP,
- solutions will typically be less regular. For example, if  $(V_1, V_2)$  is the solution in Section 2.1.3 of the linear game, then  $V_i$  is singular at  $\bar{x}_j$  (i.e., each equilibrium payoff is non differentiable at the border of the opponent's intervention region), in spite of the game having linear running payoffs, costs and gains. Note that Theorem 2.1.2.2 contemplates this lack of regularity within the smoothness assumptions (compare to the classical Verification Theorems for single-player problems [ØS09] where greater regularity is assumed).

Algorithm 2.2.1.1 below is, as far as we know, the first ever numerical attempt at two-player nonzero-sum stochastic impulse games. It is admittedly heuristic and supported only by the numerical evidence reported in Section 2.3. A rigorous analysis concerning an improved algorithm, specific to symmetric games, will be carried out in Chapter 3.

The remainder of this Section is organized as follows. We start by giving a quick overview of the discretization scheme we shall be using for system of QVIs (2.1.2.3). (A more detailed description can be found in Chapter 3.) We then explain the idea and motivate the underlying heuristics of our iterative method. Finally, we list the new algorithm.

*Notation.*  $\mathbb{R}^G$  denotes the set of functions  $v : G \rightarrow \mathbb{R}$ . We shall identify grid points with indexes, functions in  $\mathbb{R}^G$  with vectors and linear operators with matrices. The (partial) order considered in  $\mathbb{R}^G$  and  $\mathbb{R}^{G \times G}$  is the usual pointwise order for functions (elementwise for vectors and matrices), and the same is true for the supremum, maximum and arg-maximum induced by it. All set complements are taken with respect to the grid.

**Discretization:** From now on we localize the domain and work on a discrete finite grid  $G \subseteq \mathbb{R}$ . In all of the numerical experiments described in the sequel we have taken  $G$  as an equispaced grid of  $M$  steps between certain  $x_{\min} < 0 < x_{\max}$  with  $|x_{\min}|, |x_{\max}|$  and  $M$  large enough (more about this in Chapter 3).

We approximate the impulse constraint sets  $\mathcal{Z}(x)$  with  $x \in \mathbb{R}$  by finite sets  $\emptyset \neq Z(x)$  with  $x \in G$  and we set  $Z := \prod_{x \in G} Z(x)$ . For  $i, j \in \{1, 2\}$ ,  $i \neq j$ ,  $v_1, v_2 \in \mathbb{R}^G$  and  $x \in G$ , we define discretized versions of the loss and gain operators as

$$M_i v_i(x) := \max_{\delta \in Z} \{v_i[x + \delta(x)] - c_i(x, \delta(x))\}, \quad H_i(v_j)v_i(x) := v_i[x + \delta_j^*(x)] + g_i(x, \delta_j^*(x))$$

$$\text{with } \delta_j^*(x) = \delta_j^*(v_j)(x) := \min \left( \arg \max_{\delta \in Z} \{v_j[x + \delta(x)] - c_j(x, \delta(x))\} \right),$$

where  $v[y]$  denotes linear interpolation of  $v$  on  $y$  using the closest nodes on the grid, and  $v[y] = v(x_{\min})$  (resp.  $v[y] = v(x_{\max})$ ) if  $y < x_{\min}$  (resp.  $y > x_{\max}$ ); i.e., ‘no extrapolation’. For the experiments presented in Section 2.3 we assumed  $\mathcal{Z}(x) = \mathbb{R}$  for all  $x \in \mathbb{R}$  and we set  $Z(x) = G - x = \{y - x : y \in G\}$  for simplicity.

Next, we choose a finite difference scheme for the ODEs

$$0 = \mathcal{A}V_i - \rho_i V_i + f_i = \frac{1}{2}\sigma^2 V_i'' + \mu V_i' - \rho_i V_i + f_i$$

which is consistent, monotone and stable, adding artificial Dirichlet or Neumann-type boundary conditions (BCs). In all of our experiments, we have chosen an upwind (or positive coefficients) finite difference scheme, where

$$V_i'(x) \approx \frac{V_i(x + \text{sgn}(\mu(x))h) - V_i(x)}{\text{sgn}(\mu(x))h}, \quad V_i''(x) \approx \frac{V_i(x+h) - 2V_i(x) + V_i(x-h)}{h^2}$$

and  $V_i(x_{\min} - h)$ ,  $\tilde{V}_i(x_{\max} + h)$  were solved for using artificial Neumann conditions. (See Chapter 3 for more details on this choice as well as that of the grid extension.) The described procedure leads to a discretization of the ODEs as

$$L_i v_i + f_i = 0,$$

with  $-L_i$  strictly diagonally dominant (SDD)  $L_0$ -matrices (see definitions in Appendix A). Note that the values of  $f_i$  at  $x_{\min}, x_{\max}$  need to be modified to account for the boundary conditions.

**Heuristics:** Having discretized the system of QVIs (2.1.2.3), we start from an initial guess  $(v_1^0, v_2^0)$  to approximate its solution. We seek an iterative procedure to consecutively compute  $(v_1^{k+1}, v_2^{k+1})$ , given  $(v_1^k, v_2^k)$  at the  $k$ -th iteration.

Let  $i, j \in \{1, 2\}$ ,  $i \neq j$ . A natural idea is, first, to partition the grid into the ‘approximate intervention region’ of player  $j$ ,  $\{M_j v_j^k - v_j^k \geq 0\}$  (equivalent to  $\{M_j v_j^k - v_j^k = 0\}$  for a solution), and its complement, the ‘approximate continuation region’; and then to compute  $v_i^{k+1}$  either by calculating a gain as  $H_i(v_j^k)v_i^k$  in the former region, or by solving an impulse control problem in the latter. Unfortunately, our numerical experiments with these hard threshold definitions lead to unstable behaviours more often than not, and a successive relaxation procedure proved to be more useful. Consequently, the approximate intervention region of player  $j$  is set as  $\{M_j v_j^k - v_j^k \geq -r^k\}$  instead, where  $(r^k)$  are small positive numbers decreasing to zero.

It remains to schedule the relaxation procedure and to specify a stopping criteria for the algorithm. Regarding the former, having computed  $(v_1^{k+1}, v_2^{k+1})$ ,  $r^k$  is linearly relaxed. (This relaxation procedure is chosen for simplicity.). Then the largest pointwise residual to the system of QVIs (incurred by either payoff) is calculated across the grid, taking into account a numerical tolerance  $tol > 0$ . We denote this residual  $R^{k+1}$  and we consider the algorithm has ‘converged’ when the residual drops below a certain threshold. We use the largest residual (instead of the distance between consecutive approximations) because in the absence of a convergence analysis it is a practical metric reflecting whether a solution of the discrete system of QVIs has been found (as opposed to the algorithm stagnating or converging to something other than a solution). Additionally, the residuals give valuable information at each grid node. It is not without its caveats nonetheless, as it will be seen in the experiments that floating point arithmetic can often lead to a ‘spiking’ of the residuals at the nodes where the discrete solutions reproduce singularities of the analytical ones. Thus,

the largest residual on its own can at times be artificially big and misleading, and further inspection over all grid nodes is recommended in such cases (see Section 2.3 and more details in Chapter 3). We should also mention that in Chapter 3, aided by a convergence analysis, we will take the opposite view and the largest residual will become a tool for secondary assessment.

As a final remark, we note that numerical experiments show Algorithm 2.2.1.1 does not enjoy global convergence (i.e. it is not guaranteed to converge from arbitrary initial guesses). Providing a good enough pair  $(v_1^0, v_2^0)$  is thus a practical issue; in Section 2.3, a natural way of constructing educated guesses is explained.

In the following,  $a^+ := \max\{a, 0\}$  denotes the positive part of a given  $a \in \mathbb{R}$ ,  $\mathbb{1}_S$  denotes the indicator function of  $S \subseteq \mathbb{R}$  and  $\|\cdot\|_\infty$  is computed over  $G$ . We take for granted that we have a convergent solver for single-player impulse control problems. In Chapter 3 this point will be addressed in detail, showing how one can use either classical policy iteration or fixed-point policy iteration (the latter was used in the experiments presented in Section 2.3). Additionally, we note that in practice a maximum number of iterations should be used as an additional stopping criteria. We neglect this to simplify the presentation.

---

**Algorithm 2.2.1.1** Iterative algorithm for general impulse games

---

Set  $tol > 0$  and relaxation parameters  $0 < \alpha < 1$ ,  $r^0 > 0$ .

- 1: Choose initial guess:  $(v_1^0, v_2^0) \in \mathbb{R}^G \times \mathbb{R}^G$
  - 2: Set  $R^0 = +\infty$
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:   **for**  $i=1, 2$  **do**
  - 5:      $j = 3 - i$
  - 6:      $I_j^k = \{M_j v_j^k - v_j^k \geq -r^k\}$
  - 7:      $v_i^{k+1} = H_i(v_j^k) v_i^k$  on  $I_j^k$
  - 8:     Solve impulse control:  $\max \{L_i v_i^{k+1} + f_i, M_i v_i^{k+1} - v_i^{k+1}\} = 0$  on  $(I_j^k)^c$
  - 9:   **end for**
  - 10:    $r^{k+1} = \alpha r^k$
  - 11:    $I_j^{k+1, tol} = \{M_j v_j^{k+1} - v_j^{k+1} \geq -tol\}$
  - 12:    $R^{k+1} = \left\| \max_{i,j \in \{1,2\}, i \neq j} \left\{ (M_i v_i^{k+1} - v_i^{k+1})^+, |H_i(v_j^{k+1}) v_i^{k+1} - v_i^{k+1}| \mathbb{1}_{I_j^{k+1, tol}} + \right. \right.$   
 $\left. \left| \max \{L_i v_i^{k+1} + f_i, M_i v_i^{k+1} - v_i^{k+1}\} \right| \mathbb{1}_{(I_j^{k+1, tol})^c} \right\} \right\|_\infty$
  - 13:   **if**  $R^{k+1} < tol$  **then**
  - 14:     break from the iteration
  - 15:   **end if**
  - 16: **end for**
- 

Line 8 of Algorithm 2.2.1.1 deserves special attention. Although at this step we want to solve an impulse control problem restricted to the subgrid  $D = (I_j^k)^c$  (for fixed  $k, j$ ), we still need the information in  $I_j^k$  in two ways; namely:

- To compute the non-local operator  $M_i$ .
- To restrict to  $D$  the system  $L_i v_i^{k+1} + f_i = 0$ .

It is also for this reason that the gain  $v_i^{k+1} = H_i(v_j^k)v_i^k(x)$  is computed in advance in line 7, effectively making the QVI in line 8 a constrained QVI over the whole grid. Thus, one needs to algebraically modify it into a classical QVI on  $D$  before an impulse control solver can be applied. See more details in Chapter 3.

## 2.3 Numerical results

In this section we assess the performance of Algorithm 2.2.1.1. The system of QVIs (2.1.2.3) is discretized as explained in Section 2.2 on an equispaced grid of  $M + 1$  nodes, and the computational domain is always the plotted one. We assume  $\mathcal{Z}(x) = \mathbb{R}$  for all  $x \in \mathbb{R}$  and set  $Z(x) = G - x = \{y - x : y \in G\}$ . Throughout this section,  $tol = 10^{-8}$ ,  $\alpha = 0.8$  and  $r_0 = 1$ . The largest pointwise residual at convergence is denoted by  $R^\infty$ .

As mentioned before, the only almost fully analytically solvable game currently available is the linear game of Section 2.1.3. Therefore, we shall eventually focus on that problem for validation purposes. However, we introduce two other games first for which we do not have an analytical solution. For these games, we solve the corresponding QVIs systems on a fixed grid and find approximate NEs (as much as this can be claimed from a numerical viewpoint). Additionally, they illustrate how an initial guess for the linear game can be constructed and provide further numerical evidence supporting Algorithm 2.2.1.1.

### 2.3.1 Non-symmetric parabolic game

This is a version of the linear game where the running payoffs are replaced by (non-symmetric) concave parabolae with roots  $r_i^L$  and  $r_i^R$ :

$$f_i = -(x - r_i^L)(x - r_i^R), \text{ with } r_i^L < r_i^R. \quad (2.3.1.1)$$

Let us motivate this game. We seek to use as initial guess the pair of value functions of the single-player versions of the game. That is, the impulse control problems in which one of the players never intervenes. Removing the action of one of the players, however, may not always lead to a well-posed problem. For example, for running payoffs without maxima one can easily end up with ‘infinite-valued value functions’ or ‘infinite-valued optimal impulses’. This is indeed the case for the linear game. In order to skirt that difficulty it is convenient to consider modifications with running payoffs that attain a maximum value, like those in (2.3.1.1).

Figure 2.1 shows a numerical solution,  $(v_1, v_2)$ , to a parabolic game (pair of solid curves) along with the initial guess (pair of dashed curves). The latter are, in turn, the value functions of the corresponding single-player games. It is intuitively clear that  $(v_1, v_2)$  approximate, over the grid, functions which indeed satisfy the assumptions of the Verification Theorem 2.1.2.2. The NE exhibited in this Theorem,  $(\varphi_1^*, \varphi_2^*)$ , can be retrieved from the pair of equilibrium payoffs, giving  $\varphi_1^* = ([1.068, +\infty), -1.848 - x)$  and  $\varphi_2^* = ((-\infty, -3.048], -0.120 - x)$ .

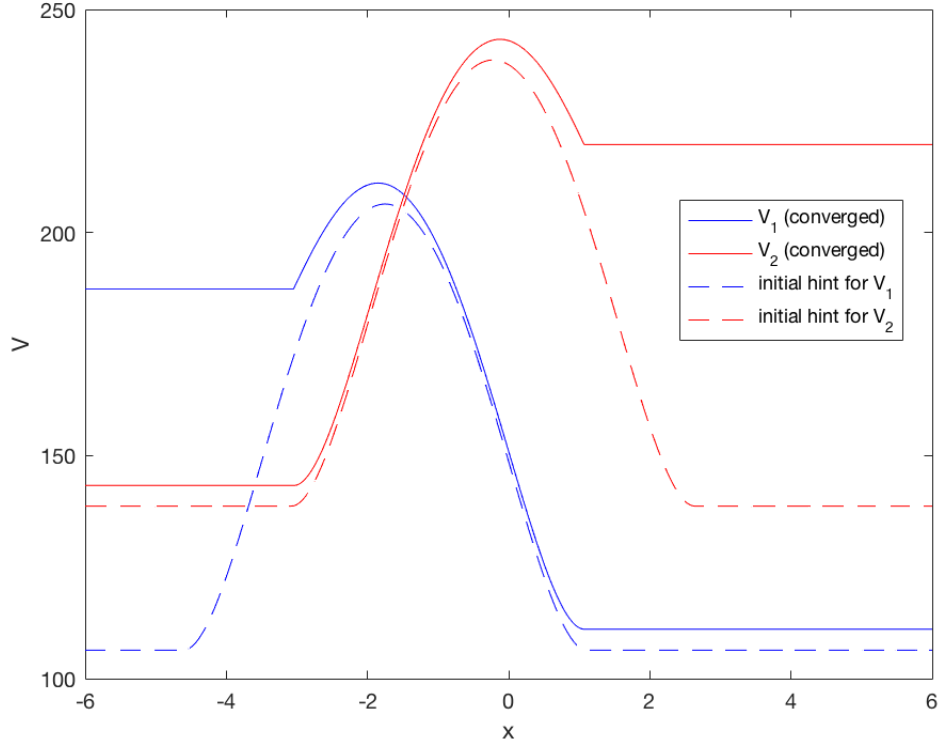


Figure 2.1: Equilibrium payoffs of non-symmetric parabolic game:  $f_1 = -(x + 4.5)(x - 1)$  and  $f_2 = -(x + \pi)(x - 2.7)$ ,  $\rho = .03$ ,  $\sigma = .25$ ,  $c = 100$ ,  $\tilde{c} = 30$ ,  $\lambda = \tilde{\lambda} = 0$ . Overlaid, initial guesses (solutions of the respective single-player games). Here,  $M = 1000$ ; check Table 2.1 for other values.

As it turns out, for this parabolic game Algorithm 2.2.1.1 also converges from the ‘zero guess’ ( $v_1^0 = v_2^0 = 0$ ). On Table 2.1, the value of  $R^\infty$  over a wide range of finite difference grids (with 301, 601,  $\dots$ , 3001 nodes) is compiled. In all cases, either initial guess leads to convergence (up to tolerance  $tol$ ). Nonetheless, the algorithm takes in general fewer iterations when starting from the value functions of the single-player parabolic games. (We stress that those on Table 2.1 are the outer iterations of Algorithm 2.2.1.1. Within each one there is an inner loop of an impulse control solver as in Chapter 3.)

M	$R^\infty$	its. <sup>a</sup>	its. <sup>b</sup>	Table 2.1: Largest residual to QVIs at convergence ( $R^\infty$ ) vs. number of grid nodes ( $M + 1$ ) for parabolic game in Figure 2.1. Iterations to convergence within $tol = 10^{-8}$ starting from: zero guess (its. <sup>a</sup> ) and value functions of single-player problems (its. <sup>b</sup> ).
300	$1.4 \times 10^{-12}$	54	53	
600	$4.0 \times 10^{-9}$	74	77	
1200	$3.3 \times 10^{-9}$	144	77	
1800	$9.7 \times 10^{-9}$	95	77	
2400	$7.3 \times 10^{-9}$	123	77	
3000	$5.9 \times 10^{-9}$	215	103	

We conclude this example by illustrating the interplay between the equilibrium payoffs found with Algorithm 2.2.1.1, the NE derived from them and the evolution of the optimally controlled state variable. Once the optimal strategies  $(\varphi_1^*, \varphi_2^*)$  are available, they can be

executed on specific realizations of the game. Sticking to the parameters and numerical solution in Figure 2.1, Figure 2.2 depicts one exemplary path of the state process in the time interval  $[0, 1000]$ , starting from  $x = 0$  and subjected to the pair of optimal strategies. For numerical purposes, let us consider the finite horizon payoffs:

$$\hat{J}_i(x; \varphi_1, \varphi_2, T) := \mathbb{E} \left[ \int_0^T e^{-\rho_i s} f_i(X_s) ds - \sum_{\substack{k: \\ \tau_i^k \leq T}} e^{-\rho_i \tau_i^k} c_i(X_{(\tau_i^k)^-}, \delta_i^k) + \sum_{\substack{k: \\ \tau_j^k \leq T}} e^{-\rho_i \tau_j^k} g_i(X_{(\tau_j^k)^-}, \delta_j^k) \right]. \quad (2.3.1.2)$$

For ‘good enough’ games and strategies, one intuitively expects that  $\hat{J}_i(x; \varphi_1, \varphi_2, T) \rightarrow J_i(x; \varphi_1, \varphi_2)$  as  $T \rightarrow \infty$ . In fact, after  $T \gtrsim 300$ , the integrals in (2.3.1.2) for the parabolic game and NE of Figure 2.1 have essentially attained their asymptotic value. Thus, we simply take  $J_i(x; \varphi_1^*, \varphi_2^*) \approx \hat{J}_i(x; \varphi_1^*, \varphi_2^*, T = 1000)$ . With this clarification, Figure 2.3 shows in particular  $v_1(0) = J^1(0; \varphi_1^*, \varphi_2^*)$ ,  $v_2(0) = J^2(0; \varphi_1^*, \varphi_2^*)$ ,  $v_1(-1) = J^1(-1; \varphi_1^*, \varphi_2^*)$ , and  $v_2(-1) = J^2(-1; \varphi_1^*, \varphi_2^*)$ , obtained by Monte Carlo simulations.<sup>14</sup> They compare fairly well with the values of  $v_1(0)$ ,  $v_2(0)$ ,  $v_1(-1)$  and  $v_2(-1)$  in Figure 2.1 obtained by our algorithm. (Even better agreement could be obtained by increasing  $M$  in that figure and reducing the discretization bias and statistical error of the Monte Carlo simulation, but this is good enough to make our point.)

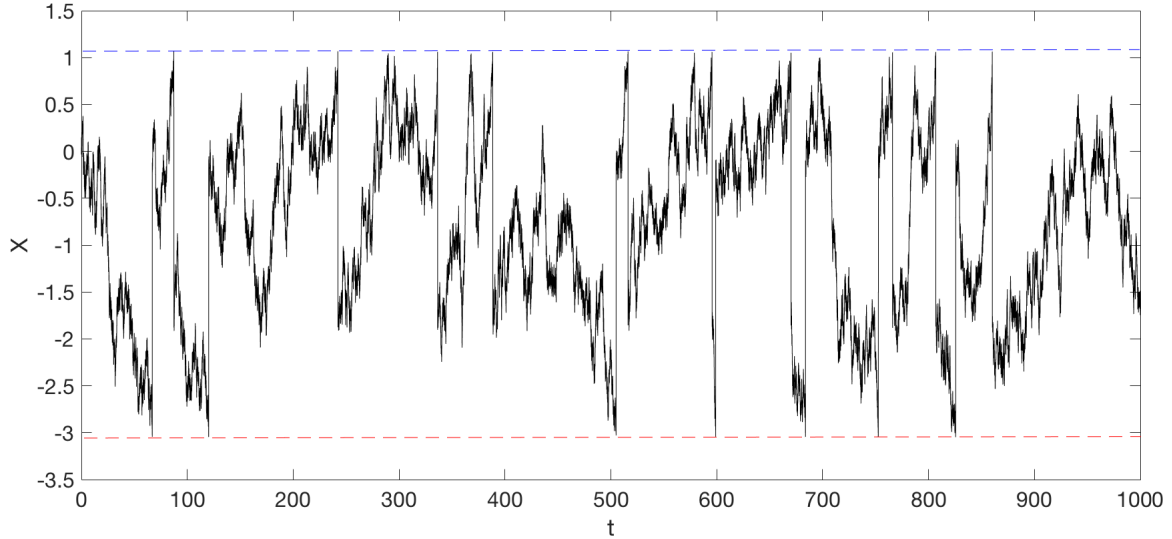


Figure 2.2: Sample path from  $x = 0$  with parameters and solution from Figure 2.1. Blue and red dashed lines are intervention thresholds for players 1 and 2, respectively.

<sup>14</sup>The expected values in (2.3.1.2) are approximated by the mean over  $N = 200$  realizations integrated with the Euler-Maruyama method with time step  $\Delta t = 0.001$ .



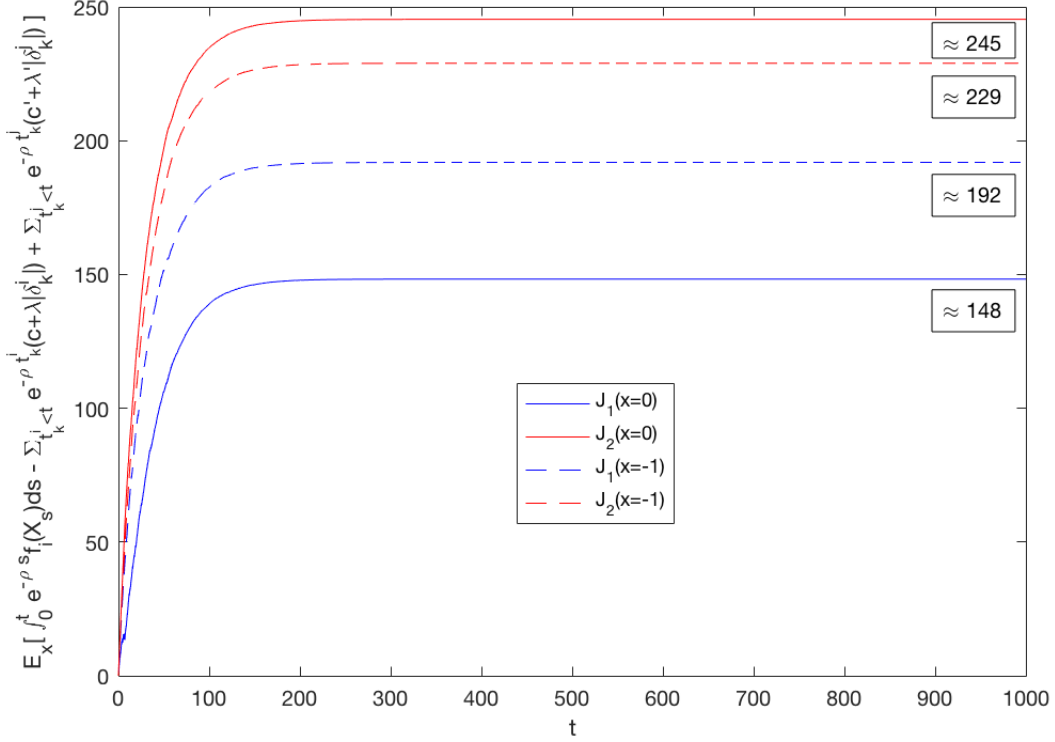


Figure 2.3: Approximations of the equilibrium payoffs at  $x = 0$  (solid curves) and  $x = -1$  (dashed curves), obtained by Monte Carlo simulation (see text for details). Parameters and optimal strategies from Figure 2.1. Compare with  $v_1(0)$ ,  $v_2(0)$ ,  $v_1(-1)$  and  $v_2(-1)$  there.

The NE itself can be visually explored in the following way. For a given starting point  $x$ , we keep the optimal strategy for one of the two players, and slightly alter the strategy of the other one. For concreteness, let us assume that player 1 changes her strategy to  $\varphi_1 = (1 \pm 0.25\mathcal{U})\varphi_1^*$  while player 2 maintains  $\varphi_2 = \varphi_2^*$ , where  $\mathcal{U}$  is the uniform distribution, ‘ $\pm$ ’ means ‘with equal chance’.<sup>15</sup> Then, we proceed to calculate  $J^1(x; \varphi_1, \varphi_2^*)$  by Monte Carlo simulations as before. By definition of the NE,  $J^1(x; \varphi_1, \varphi_2^*)$  cannot be larger than  $v_1(x)$ . Within numerical tolerance, this is indeed observed in Figure 2.4, where the blue empty circles ( $J^1(x; \varphi_1, \varphi_2^*)$ ) do not lie over the blue curve ( $v_1(x)$ ). Full red circles represent  $J^2(x; \varphi_1, \varphi_2^*)$ : note that the player who sticks to her optimal strategy may indeed improve over  $v_2(x)$ , should her opponent depart from a NE. (When player 2 is the one who changes, the red empty circles, red curve and full blue circles apply instead.) We stress, however, that Monte Carlo simulations such as these cannot *prove* that a pair of strategies form a NE. (At most, they can disprove it.) This is why Algorithm 2.2.1.1 becomes a valuable tool in solving the underlying system of QVIs and applying the Verification Theorem 2.1.2.2.

<sup>15</sup>Changes are applied to each of the two parameters defining  $\varphi_1^*$ .

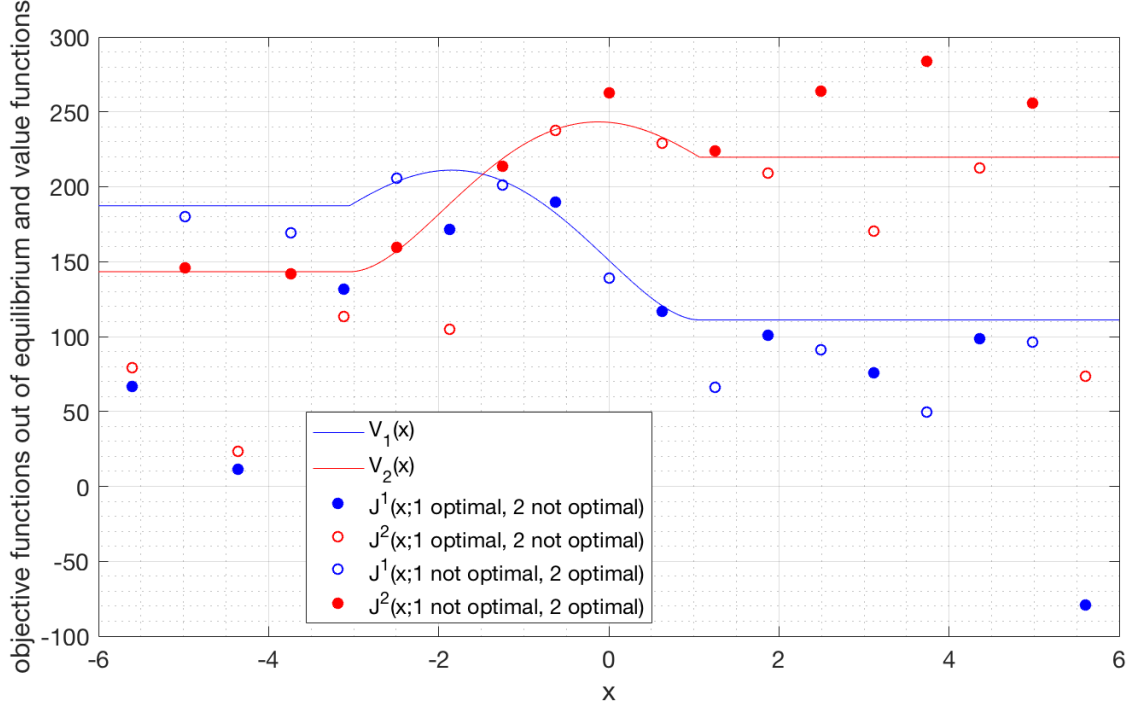


Figure 2.4: Empty blue (resp. red) circles mean that the player 1 (resp. 2) has departed from her optimal strategy while her opponent has not. By virtue of NE, she cannot be (within numerical tolerance) better off than  $v_1$  (resp.  $v_2$ ). Full circles: payoff of the player who does not change her optimal strategy. Parameters and optimal strategies from Figure 2.1. (Note that results are subject to numerical error.)

### 2.3.2 Capped linear game

In this game, we cap the running payoffs of the linear game at  $K > 0$ :

$$f_i = (-1)^{i-1}(x - s_i) \wedge K.^{16} \quad (2.3.2.1)$$

We shall always take  $K = 5$ . Once again, the corresponding single-player games are well-posed and their solutions can be used as initial guess for the capped linear game. As in the previous example, the capped game also seems to converge from the zero guess. Some convergence results are compiled in Table 2.2. Note that convergence falters with  $M = 600$  and  $M = 3300$ ; we will come back to this later.

<sup>16</sup> $a \wedge b$  denotes the minimum of  $a$  and  $b$  for any  $a, b \in \mathbb{R}$ .

M	$R^\infty$	its.	Table 2.2: Iterations needed for convergence and largest terminal residual ( $R^\infty$ ) in capped linear game using zero initial guess (a hyphen stands for lack of convergence within $tol$ ). Same parameters as in Figure 2.5.
600	—	$\infty$	
900	$6.5 \times 10^{-10}$	100	
1200	$8.8 \times 10^{-9}$	124	
1500	$7.4 \times 10^{-9}$	78	
2700	$4.1 \times 10^{-9}$	96	
3000	$6.7 \times 10^{-9}$	125	
3300	—	$\infty$	

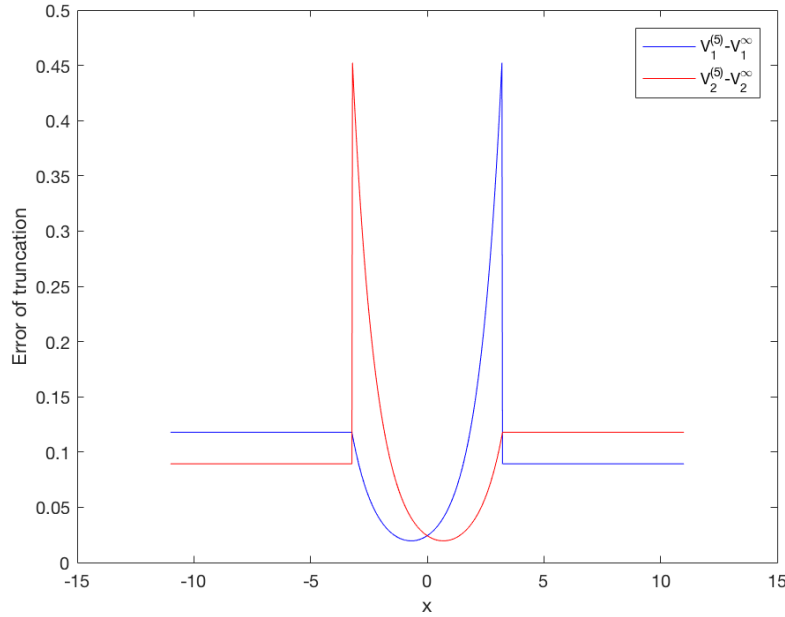


Figure 2.5: Difference (in absolute value, less than 1%) between (numerical, with  $M = 1200$ ) equilibrium payoffs of the capped linear game and those (exact) of the linear game from Section 2.1.3, justifying the former payoffs being used as initial guess for the latter game. Parameters:  $\sigma = .25$ ,  $\rho = .03$ ,  $c = 100$ ,  $\tilde{c} = 30$ ,  $\lambda = 0.5$ ,  $\tilde{\lambda} = 0.3$ ,  $s_1 = -\pi/3$ ,  $s_2 = \pi/3$ ,  $K = 5$ .

The equilibrium payoffs found for the capped linear game are a good approximation to those of the linear game shown in Section 2.1.3 (see Figure 2.5). This seems to make sense: due to the action of the opponent and for  $\sigma \ll K$ , the discarded portion of the payoff is not very relevant in practice.

### 2.3.3 Linear game with educated initial guess

Finally, we tackle the linear game, for which an exact solution is available (see Section 2.1.3). Contrary to the previous examples, Algorithm 2.2.1.1 does not seem to enjoy unconditional convergence here. In fact, when the zero guess was used, it failed to converge more often than not (not reported). In order to construct an adequate initial guess, we first solve for some

equilibrium payoffs of the capped linear game. Using them as the initial guess, convergence was achieved in every experiment.

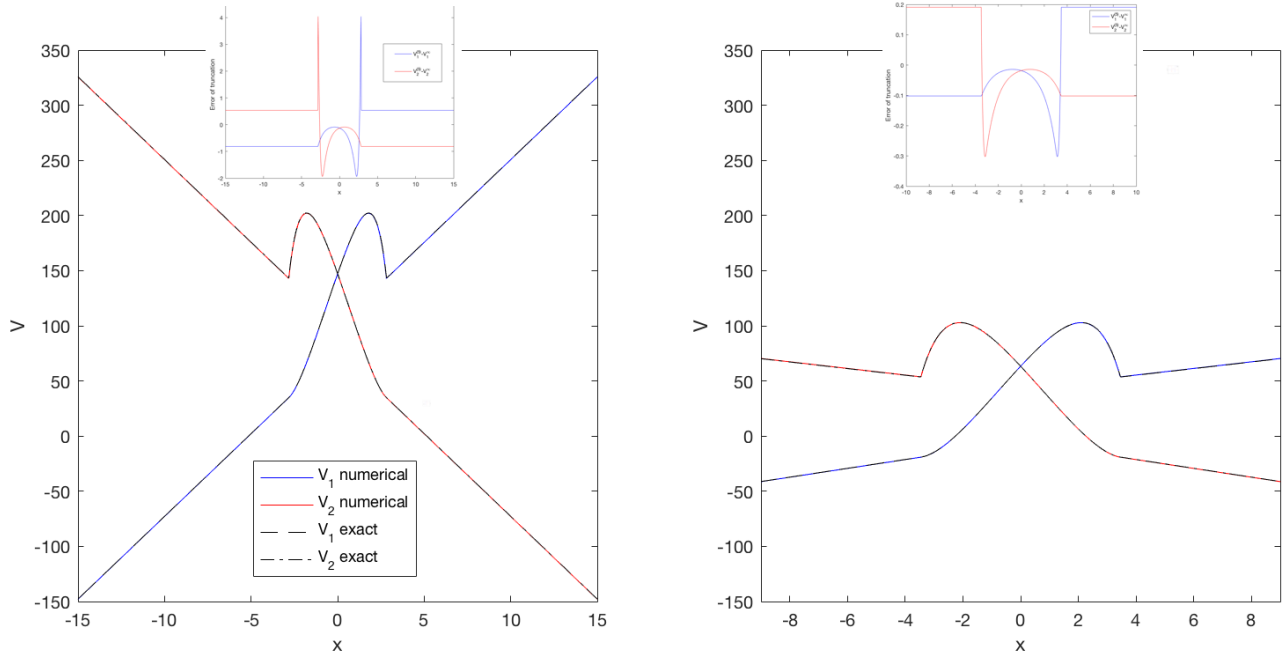


Figure 2.6: Equilibrium payoffs of two instances of linear game. Initial guess: solutions of capped linear games ( $K = 5$ ). Overlaid, error of the initial guess. Parameters:  $\rho = .02$ ,  $\sigma = .15$ ,  $s_1 = -3$ ,  $s_2 = 3$ ,  $c = 100$ ,  $\tilde{c} = 0$ ,  $\lambda = \tilde{\lambda} = 15$ ,  $M = 1000$  (left);  $\rho = .03$ ,  $\sigma = .25$ ,  $s_1 = -2$ ,  $s_2 = 2$ ,  $c = 100$ ,  $\tilde{c} = 30$ ,  $\lambda = 4$ ,  $\tilde{\lambda} = 3$ ,  $M = 1000$  (right).

The results of two experiments are plotted on Figure 2.6. The numerical approximations can hardly be distinguished from the exact solutions with the naked eye. As it was done with the parabolic game, once again we can retrieve an approximate NE from the numerical solution. For the left figure, this gives  $\varphi_1^* = ((-\infty, -2.82], 1.53 - x)$  and  $\varphi_2^* = ([2.82, +\infty), -1.53 - x)$ . When compared with the exact equilibrium, the errors on the corresponding abscissae are smaller than the grid step size.

M	$R^\infty$	error	its.	Table 2.3: Convergence of Algorithm 2.2.1.1 for linear game. Same parameters as in Figure 2.6 (left).
500	$3.8 \times 10^{-10}$	0.687	126	
1000	$1.2 \times 10^{-9}$	0.805	154	
1500	$7.2 \times 10^{-9}$	0.512	157	
2000	$3.5 \times 10^{-9}$	0.365	172	
2500	—	—	$\infty$	

On Tables 2.3 and 2.4, the convergence of the numerical approximation provided by Algorithm 2.2.1.1 to the true solution is demonstrated. However,  $R^\infty$  fails to drop below  $tol$  for some discretizations as the algorithm stagnates, with the highest residuals found on and around the nodes where the payoffs display a singularity. Figure 2.7 illustrates this situation.

M	$R^\infty$	error	its.	Table 2.4: Convergence of Algorithm 2.2.1.1 for linear game (same parameters as in Figure 2.6 (right)).
600	—	—	$\infty$	
800	$9.5 \times 10^{-10}$	0.023	183	
1000	$3.7 \times 10^{-9}$	0.330	177	
1400	$8.7 \times 10^{-9}$	0.196	159	
1800	$6.8 \times 10^{-9}$	0.121	226	
2200	$5.6 \times 10^{-9}$	0.073	224	
2600	—	—	$\infty$	

It can be seen that errors continue to be acceptable (far less than 1% relative error in the worst case), but further refinements of the grid did not yield any major improvements. The algorithm for symmetric games, presented and studied in Chapter 3, is far better-behaved in this respect as well. It will be seen that for games in which one intuitively expects a solution to exist, cases of stagnation were always resolved by refining the grid.

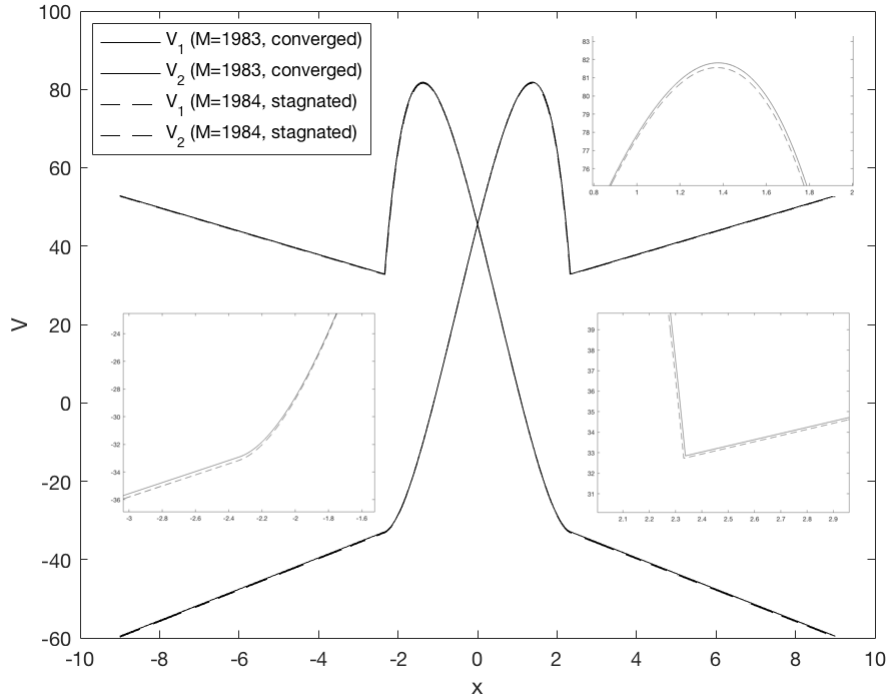


Figure 2.7: Stagnation of accuracy: the pair with  $M = 1983$  is fully converged (within numerical tolerance); the pair with  $M = 1984$  (overlaid) is not. The insets zoom in on both pairs of functions. Parameters:  $\rho = .02$ ,  $\sigma = .15$ ,  $s_1 = -1$ ,  $s_2 = 1$ ,  $c = 100$ ,  $\tilde{c} = 30$ ,  $\lambda = 4$ ,  $\tilde{\lambda} = 3$ .

## 2.4 Concluding remarks

In this chapter, we have designed and tested a novel policy iteration algorithm (the first one as far as we know) to numerically solve nonzero-sum stochastic impulse control games.

The approach consists in solving a system of QVIs which characterizes certain equilibrium payoffs and NEs, exploiting the Verification Theorem of [ABC<sup>+</sup>19].

Our algorithm iteratively computes the approximate solution by partitioning, for each of the players, the discretized localized spatial domain into approximate continuation and intervention regions of the opponent. These are defined through a relaxation parameter that evolves along the iterations. In the continuation region, we solve a constrained single-player impulse control problem, whereas in the complement, a gain is computed. A way of producing an educated initial guess to start the iterations, which relies on solving standard impulse control problems, has been presented along with the new algorithm.

The algorithm was validated numerically with satisfactory results, by considering different games and both fixed and refining grids, but it was noted that cases of stagnation could not be resolved by simple grid refinements. In particular, we have tested the convergence to an analytical solution of the linear game, the only one almost fully tractable at the time of writing. Results show that the algorithm offers a means of gaining quantitative insight into applications modelled by general nonzero-sum stochastic impulse games. In the case of symmetric games however, we shall see in the next chapter a modified algorithm which substantially improves the current one, and we will provide the missing convergence analysis.

# Chapter 3

## A fixed-point policy-iteration-type algorithm for symmetric nonzero-sum stochastic impulse games

### Introduction

To the author's best knowledge, the only numerical method available in the literature of nonzero-sum stochastic impulse games is our algorithm in [ABM<sup>+</sup>19] (Algorithm 2.2.1.1 in Chapter 2), which tackles the system of QVIs (2.1.2.3) by sequentially solving single-player impulse control problems combined with a relaxation scheme. Unfortunately, the choice of the latter scheme is not obvious in general and it was verified that the convergence of the algorithm was reliant on a good initial guess. It was also observed that stagnation could put a cap on the accuracy of the results, without any simple solution to it. Lastly, while a numerical validation was performed, no rigorous convergence analysis was provided at the time.

Restricting attention to the one-dimensional infinite horizon two-player case, this chapter (based on [Zab19]) puts the focus on certain nonzero-sum impulse games which display a symmetric structure between the players. This class is broad enough to include many interesting applications; no less than the competing central banks problem (whether in its linear form [ABC<sup>+</sup>19] or others considered in the single bank formulation [AF16, CZ99, JP93, MØ98]), the cash management problem [BCG19] (reducing its dimension by a simple change of variables) and the generalization of many impulse control problems to the two-player case.

For this class of games, an iterative algorithm is presented which substantially improves Algorithm 2.2.1.1 by harnessing the symmetry of the problem, removing the strong dependence on the initial guess and dispensing with the relaxation scheme altogether. The result is a simpler and more intuitive, precise and efficient routine, for which a convergence analysis is provided. It is shown that the overall routine admits a representation that strongly resembles, both algorithmically and in its properties, that of the combined fixed-point policy-iteration methods [HFL12, Cli07], albeit with nonexpansive operators. Still, a certain contraction property can still be established.

To perform the analysis, assumptions are imposed on the discretization scheme used on

the system of QVIs and the discrete admissible strategies. These naturally generalize those of the impulse control case [AF16] and admit graph-theoretic interpretations in terms of weakly chained diagonally dominant (WCDD) matrices and their matrix sequences counterpart [Azi19a]. We establish a clear parallel between these discrete type assumptions, the behaviour of the players and the mimicking of the Verification Theorem.

Section 3.1 deals with the analytical problem. We give a precise definition of the class of symmetric nonzero-sum impulse games and establish some preliminary results.

Section 3.2 moves on the analogue discrete problem. Section 3.2.1 specifies a general and abstract discrete version of the system of QVIs, such that any discretization scheme compliant with the assumptions to be imposed will be subject to the same results. Section 3.2.2 presents the iterative algorithm subject of this chapter, and shows how the impulse control problems that need to be sequentially solved have a unique solution and can be handled by policy iteration. However, the latter has costly efficiency drawbacks, which is why Section 3.2.3 provides a general solver for impulse control problems which outperforms classical policy iteration in many practical situations. It consists of an instance of fixed-point policy-iteration that is noncompliant with the standard assumptions [HFL12] and, as far as the author knows, it was not used in the context of impulse control before, other than heuristically in Chapter 2. We prove its convergence under our framework.

Section 3.2.4 characterizes the overall iterative algorithm as a fixed-point policy-iteration-type method, allowing for reformulations of the original problem and results pertaining the solutions. (The necessary matrix and graph-theoretic definitions and results are collected in Appendix A for the reader's convenience.) Section 3.2.5 carries on with the convergence analysis, and shows to which extent different sets of reasonable assumptions are enough to guarantee convergence to solutions, convergence of strategies and boundedness of iterates. A result of theoretical interest, giving sufficient conditions for convergence, is proved. Discretization schemes within the standing framework are provided in Section 3.2.6.

Section 3.3 presents all the numerical results. In Section 3.3.1, a variety of symmetric nonzero-sum impulse games, many seemingly too complicated to be handled analytically, are explicitly solved for equilibrium payoffs and NEs strategies with great precision. This is done on a fixed grid, while considering different performance metrics and addressing practical matters of implementation. In the absence of a viscosity solutions framework to establish convergence to analytical solutions as the grid is refined, Section 3.3.2 performs a numerical validation using the only examples of symmetric solvable games in the literature. Section 3.3.3 addresses the case of games without NEs. Section 3.3.4 tackles games such that the results go beyond the scope of all the currently available theory, displaying discontinuous impulses and very irregular payoffs. The latter give insight and motivate further research into this field. Lastly, Section 3.4 concludes.

### 3.1 Analytical continuous-space problem: the symmetric case

In Section 2.1 we established a general framework for two-player nonzero-sum stochastic impulse control games. We now want to focus our study on games which present a certain



type of symmetric structure between the players, generalising the linear game in Section 2.1.3 and the cash management game [BCG19].<sup>1</sup> We shall make a slight abuse of terminology with respect to the more common use of the term ‘symmetric’ in games theory, but this will be consistent all throughout.

*Notation.* The type of games presented in Section 2.1.1 are fully defined by setting the drift, volatility, impulse constraints, discount rates, running payoffs, costs and gains. In other words, any such game can be represented by a tuple  $\mathcal{G} = (\mu, \sigma, \mathcal{Z}_i, \rho_i, f_i, c_i, g_i)_{i=1,2}$ .

**Definitions 3.1.1.1.** We say that a game  $\mathcal{G} = (\mu, \sigma, \mathcal{Z}_i, \rho_i, f_i, c_i, g_i)_{i=1,2}$  is *symmetric (with respect to zero)* if

- (S1)  $\mu$  is odd and  $\sigma$  is even (i.e.,  $\mu(x) = -\mu(-x)$  and  $\sigma(x) = \sigma(-x)$  for all  $x \in \mathbb{R}$ ).
- (S2)  $-\mathcal{Z}_2(-x) = \mathcal{Z}_1(x) \subseteq [0, +\infty)$  for all  $x \in \mathbb{R}$  and  $\mathcal{Z}_1(x) = \{0\} = \mathcal{Z}_2(-x)$  for all  $x \geq 0$ .
- (S3)  $\rho_1 = \rho_2$ ,  $f_1(x) = f_2(-x)$ ,  $c_1(x, \delta) = c_2(-x, -\delta)$  and  $g_1(x, -\delta) = g_2(-x, \delta)$ , for all  $\delta \in \mathcal{Z}_1(x)$ ,  $x \in \mathbb{R}$ .

We say that the game is *symmetric with respect to  $s$*  (for some  $s \in \mathbb{R}$ ), if the  *$s$ -shifted* game  $(\mu(x+s), \sigma(x+s), \mathcal{Z}_i(x+s), \rho_i, f_i(x+s), c_i(x+s, \delta), g_i(x+s, \delta))_{i=1,2}$  is symmetric. We refer to  $x = s$  as a *symmetry line* of the game.

Condition (S1) is necessary for the state variable to have symmetric dynamics. In particular, together with (S3), it guarantees symmetry between solutions of the Hamilton–Jacobi–Bellman (HJB) equations of the players when there are no interventions, i.e.,

$$V = V^*(x) \text{ solves } \mathcal{A}V - \rho_1 V + f_1 = 0 \quad \text{if and only if} \quad V = V^*(-x) \text{ solves } \mathcal{A}V - \rho_2 V + f_2 = 0.$$

*Examples 3.1.1.2.* The most common examples of Itô diffusions satisfying this assumption are the scaled Brownian motion (symmetric with respect to zero) and the Ornstein–Uhlenbeck (OU) process (symmetric with respect to its long term mean).

Condition (S3) is self-explanatory, while (S2) is only partly so. Indeed, although symmetric constraints on the impulses  $\mathcal{Z}_1(x) = -\mathcal{Z}_2(-x)$  should clearly be a requirement, the rest of (ii) is in fact motivated by the numerical method to be presented and the type of problems it can handle. On the one hand, the third equation of the QVIs system (2.1.2.3) implies that a stochastic impulse control problem for player  $i$  needs to be solved on  $\mathcal{C}_j^*$ . The unidirectional impulses assumption is a common one for the convergence of policy iteration algorithms in impulse control.<sup>2</sup> However, it is often too restrictive for many interesting applications, such as when the controller would benefit the most from keeping the state variable within some bounded interval instead of simply keeping it ‘high’ or ‘low’ (see, e.g., [Bas19] and [AF16, Sect.6.1]). Interestingly enough, assuming unidirectional impulses turns out to be less restrictive when there is a second player present, with an opposed objective. Indeed, it is often the case that each player needs not to intervene in one of the two directions, and

<sup>1</sup>The latter can be reduced to one dimension with the change of variables  $x = x_1 - x_2$ . Additionally, we will restrict attention to unidirectional impulses, as these yield the ‘most relevant’ NE found in [BCG19].

<sup>2</sup>See this assumption in [CMS07, Sect.4] or [ØS09, Sect.10.4.2], its graph-theoretic counterpart in [AF16, Asm.(H2) and Thm.4.3], and a counterexample of convergence in its absence in [AF16, Ex.4.9].

can instead rely on her opponent doing so, while capitalising a gain rather than paying a cost. See examples in Section 3.3.1 with quadratic and degree four running payoffs.

On the other hand,  $\mathcal{Z}_1(x) = \{0\} = \mathcal{Z}_2(-x)$  for all  $x \geq 0$  means that we can assume without loss of generality that the admissible intervention regions do not cross over the symmetry line; i.e.,  $\mathcal{I}_1 \subseteq (-\infty, 0)$  and  $\mathcal{I}_2 \subseteq (0, +\infty)$  for every pair of strategies. (See Remark 2.1.2.3.) This guarantees in particular that the players never want to intervene at the same time and the priority rule can be disregarded.

There are different reasons why the last mentioned condition is less restrictive than it first appears to be. It is not uncommon to assume connectedness of either intervention or continuation regions (or other conditions implying them) both in impulse control [Ega08] and nonzero-sum games [DAFM18, Sect.1.2.1]. The same can be said for assumptions that prevent the players from intervening in unison [DAFM18, Sect.1.2.1], [Cos13, Rmk.6.5].<sup>3</sup> In the context of symmetric games and payoffs (see Lemma 3.1.1.7) such assumptions would necessarily imply the intervention regions need to be on opposed sides of the symmetry line. Additionally, without any further requirements, strategies such that  $\mathcal{I}_1 \supseteq (-\infty, 0]$  and  $\mathcal{I}_2 \supseteq [0, +\infty)$  would be inadmissible in the present framework, as per yielding infinite simultaneous impulses.

**Definitions 3.1.1.3.** Given a symmetric game, we say that  $((\mathcal{I}_i, \delta_i))_{i=1,2}$  are *symmetric strategies (with respect to zero)* if  $\mathcal{I}_1 = -\mathcal{I}_2$  and  $\delta_1(x) = -\delta_2(-x)$ . Given a symmetric game with respect to some  $s \in \mathbb{R}$ , we say that  $((\mathcal{I}_i, \delta_i))_{i=1,2}$  are *symmetric strategies with respect to  $s$*  if  $((\mathcal{I}_i - s, \delta_i(x + s)))_{i=1,2}$  are symmetric, and we refer to  $x = s$  as a *symmetry line* of the strategies.

**Definition 3.1.1.4.** We say that  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  are *symmetric functions (with respect to zero)* if  $V_1(x) = V_2(-x)$ . We say that they are *symmetric functions with respect to  $s$*  (for some  $s \in \mathbb{R}$ ) if  $V_1(x + s), V_2(x + s)$  are symmetric, and we refer to  $x = s$  as a *symmetry line* for  $V_1, V_2$ .

*Remark 3.1.1.5.* Definition 3.1.1.3 singles out strategies that share the same symmetry line with the game. For the linear game, for example, the authors find infinitely many NEs [ABC<sup>+</sup>19, Prop.4.7], each presenting symmetry with respect to some point  $s$ , but only one for  $s = 0$  (hence, symmetric in the sense of Definition 3.1.1.3). At the same time, the latter is the only one for which the corresponding equilibrium payoffs  $V_1, V_2$  have a symmetry line as per Definition 3.1.1.4. The same is true for the cash management game [BCG19].

*Remark 3.1.1.6.* Throughout the chapter we will work only with games symmetric with respect to zero, to simplify the notation. Working with any other symmetry line amounts simply to shifting the game and results back and forth.

**Lemma 3.1.1.7.** *For any symmetric game, strategies  $(\varphi_1, \varphi_2)$  and functions  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$ :*

(i) *If  $V_1, V_2$  are symmetric, then  $\mathcal{M}_1 V_1, \mathcal{M}_2 V_2$  are symmetric.*

---

<sup>3</sup>In [Cos13, Rmk.6.5], assumptions are given to guarantee the zero-sum analogue to  $\mathcal{M}_i V_i \leq \mathcal{H}_i V_i$ , with a strict inequality if such assumptions are slightly strengthened. These inequalities, in the context of Theorem 2.1.2.2, imply that the equilibrium intervention regions cannot intersect.

- (ii) If  $V_1, V_2$  are symmetric and have the UIP, then  $\delta_1^*(x) = -\delta_2^*(-x)$  and  $\mathcal{H}_1 V_1, \mathcal{H}_2 V_2$  are symmetric.
- (iii) If  $(\varphi_1, \varphi_2)$  are symmetric, then  $J_1(\cdot; \varphi_1, \varphi_2), J_2(\cdot; \varphi_1, \varphi_2)$  are symmetric.
- (iv) If  $V_1, V_2$  are as in Theorem 2.1.2.2 and  $(\varphi_1^*, \varphi_2^*)$  is the corresponding NE of the theorem, then  $(\varphi_1^*, \varphi_2^*)$  are symmetric if and only if  $V_1, V_2$  are symmetric.

*Proof.* (i) and (ii) are straightforward from the definitions.

To see (iii), one can check with the recursive definition of the state variable [ABC<sup>+</sup>19, Def.2.2] that  $X^{-x; \varphi_1, \varphi_2}$  has the same law as  $-X^{x; \varphi_1, \varphi_2}$  (recall that the continuation regions are simply disjoint unions of open intervals). Noting also that intervention times and impulses are nothing but jump times and sizes of  $X$ , one concludes that  $J_1(x; \varphi_1, \varphi_2) = J_2(-x; \varphi_1, \varphi_2)$ , as intended.

Finally, (iv) is a consequence of (i), (ii) and (iii).  $\square$

*Convention 3.1.1.8.* In light of Lemma 3.1.1.7, for any symmetric game we will often lose the player index from the notations and refer always to quantities corresponding to player 1,<sup>4</sup> henceforth addressed simply as ‘the player’. Player 2 shall be referred to as ‘the opponent’. Statements like ‘ $V$  has the UIP’ or ‘ $V$  is a symmetric equilibrium payoff’ are understood to refer to  $(V(x), V(-x))$ . Likewise, ‘ $(\mathcal{I}, \delta)$  is admissible’ or ‘ $(\mathcal{I}, \delta)$  is a NE’ refer to the pair  $(\mathcal{I}, \delta(x)), (-\mathcal{I}, -\delta(-x))$ .

Due to their general lack of uniqueness, it is customary in game theory to restrict attention to specific type of NEs, depending on the problem at hand (see for instance [HS81] for a treatment within the classical theory). Motivated by Lemma 3.1.1.7 (iii) and (iv), and by Remark 3.1.1.5, one can arguably state that symmetric NEs are the most meaningful for symmetric games. Furthermore, Lemma 3.1.1.7 implies that for symmetric games, one can considerably reduce the complexity of the full system of QVIs (2.1.2.3) provided the conjectured NE (or equivalently, the pair of payoffs) is symmetric. Using Convention 3.1.1.8, Theorem 2.1.2.2 and Lemma 3.1.1.7 give:

**Corollary 3.1.1.9 (Symmetric system of QVIs).** *Given a symmetric game as in Definition 3.1.1.1, let  $V : \mathbb{R} \rightarrow \mathbb{R}$  be a function with the UIP, such that:*

$$\begin{cases} \mathcal{H}V - V = 0 & \text{on } -\{\mathcal{M}V - V = 0\} =: -\mathcal{I}^* \\ \max\{\mathcal{A}V - \rho V + f, \mathcal{M}V - V\} = 0 & \text{on } -\{\mathcal{M}V - V < 0\} =: -\mathcal{C}^* \end{cases} \quad (3.1.1.1)$$

and  $V \in C^2(-\mathcal{C}^* \setminus \partial\mathcal{C}^*) \cap C^1(-\mathcal{C}^*) \cap C(\mathbb{R})$  has polynomial growth and bounded second derivative on some reduced neighbourhood of  $\partial\mathcal{C}^*$ . Suppose further that  $(\mathcal{I}^*, \delta^*)$  is an admissible strategy.

*Then,  $V$  is a symmetric equilibrium payoff attained at a symmetric NE  $(\mathcal{I}^*, \delta^*)$ .*

---

<sup>4</sup>Note that  $g$  will denote  $g(x, \delta) := g_1(x, -\delta)$ , as  $\delta \geq 0$  for player 1, yet  $g_1$  depends on the (negative) impulse of player 2.

Note that system (3.1.1.1) also omits the equation  $\mathcal{M}V - V \leq 0$  as per being redundant. Indeed, by Definition 3.1.1.1 and Remark 2.1.2.3, at a NE the player does not intervene above 0, nor the opponent below it. Thus,  $\mathcal{M}V - V \leq \max\{\mathcal{A}V - \rho V + f, \mathcal{M}V - V\} = 0$  on  $-\mathcal{C}^* \supset (-\infty, 0]$  and  $\mathcal{M}V - V < 0$  on  $[0, +\infty)$ .

System (3.1.1.1) simplifies a numerical problem which is very challenging even in cases of linear structure (see Chapter 2). In light of the previous, we will focus our attention on symmetric NEs only and numerically solving the reduced system of QVIs (3.1.1.1).

## 3.2 Numerical discrete-space problem

In this section we consider a discrete version of the symmetric system of QVIs (3.1.1.1) over a fixed grid, and propose and study an iterative method to solve it. As it is often done in numerical analysis for stochastic control, for the sake of generality we proceed first in an abstract fashion without making reference to any particular discretization scheme. Instead, we give some general assumptions any such scheme should satisfy for the results presented to hold. Explicit discretization schemes within our framework are presented in Section 3.2.6 and used in Section 3.3.

### 3.2.1 Discrete system of quasi-variational inequalities

From now on we work on a discrete symmetric grid

$$G : x_{-N} = -x_N < \cdots < x_{-1} = -x_1 < x_0 = 0 < x_1 < \cdots < x_N.$$

Recall that  $\mathbb{R}^G$  denotes the set of functions  $v : G \rightarrow \mathbb{R}$  and  $S : \mathbb{R}^G \rightarrow \mathbb{R}^G$  denotes the symmetry operator,  $Sv(x) = v(-x)$ . In general, by an ‘operator’ we simply mean some  $F : \mathbb{R}^G \rightarrow \mathbb{R}^G$ , not necessarily linear nor affine unless explicitly stated. We shall identify grid points with indexes, functions in  $\mathbb{R}^G$  with vectors and linear operators with matrices; e.g.,  $S = (S_{ij})$  with  $S_{ij} = 1$  if  $x_i = -x_j$  and 0 otherwise. The (partial) order considered in  $\mathbb{R}^G$  and  $\mathbb{R}^{G \times G}$  is the usual pointwise order for functions (elementwise for vectors and matrices), and the same is true for the supremum, maximum and arg-maximum induced by it.

We want to solve the following discrete nonlinear system of QVIs for  $v \in \mathbb{R}^G$ :

$$\begin{cases} Hv - v = 0 & \text{on } -\{Mv - v = 0\} =: -I^* \\ \max\{Lv + f, Mv - v\} = 0 & \text{on } -\{Mv - v < 0\} =: -C^*, \end{cases} \quad (3.2.1.1)$$

where  $f \in \mathbb{R}^G$  and  $L : \mathbb{R}^G \rightarrow \mathbb{R}^G$  is a linear operator. The nonlinear operators  $M, H : \mathbb{R}^G \rightarrow \mathbb{R}^G$  are as follows: let  $\emptyset \neq Z(x) \subseteq \mathbb{R}$  be a finite set for each  $x \in G$ , with  $Z(x) = \{0\}$  if  $x \geq 0$ . Set  $Z := \prod_{x \in G} Z(x)$  and for each  $\delta \in Z$  let  $B(\delta) : \mathbb{R}^G \rightarrow \mathbb{R}^G$  be a linear operator,  $c(\delta) \in (0, +\infty)^G$  and  $g(\delta) \in \mathbb{R}^G$ , the three of them being *row-decoupled* (i.e., row  $x$  of  $B(\delta), c(\delta), g(\delta)$  depends only on  $\delta(x) \in Z(x)$ ). Then

$$Mv := \max_{\delta \in Z} \{B(\delta)v - c(\delta)\}, \quad Hv = H(\delta^*)v := SB(\delta^*)Sv + g(S\delta^*) \quad (3.2.1.2)$$

$$\text{and } \delta^* = \delta^*(v) := \max_{\delta \in Z} \left( \arg \max \{B(\delta)v - c(\delta)\} \right). \quad (3.2.1.3)$$

Some remarks are in order. Firstly, in the same fashion as the continuous-space case, the sets  $I^*, C^*$  form a partition of the grid and represent the (discrete) intervention and continuation regions of the player, while  $-I^*, -C^*$  are such regions for the opponent.

Secondly, the general representation of  $M$  follows [CMS07, AF16]. For the standard choices of  $B(\delta)$ , our definition of  $H$  is the only one for which a discrete version of Lemma 3.1.1.7 holds true (see Section 3.2.6). However, since  $B$  and  $g$  are row-decoupled,  $SB(\delta^*)S$  and  $g(S\delta^*)$  cannot be, as each row  $x$  depends on  $\delta^*(-x)$ . For this reason and the lack of maximization over  $-I^*$ , there is no obvious way to reduce problem (3.2.1.1) to a classical Bellman problem:

$$\sup_{\varphi} \{ -A(\varphi)v + b(\varphi) \} = 0, \quad (3.2.1.4)$$

like in the impulse control case [AF16], to apply Howard's policy iteration [BMZ09, Ho1]. Furthermore, unlike in the control case, even with unidirectional impulses and good properties for  $L$  and  $B(\delta)$ , system (3.2.1.1) may have no solution as in the analytical case [ABC<sup>+</sup>19].

Thirdly, we have defined  $\delta^*$  in (3.2.1.3) by choosing one particular maximizing impulse for each  $x \in G$ . The main motivation behind fixing one is to have a well defined discrete system of QVIs for every  $v \in \mathbb{R}^G$ . (This is not the case for the analytical problem (3.1.1.1) where the gain operator  $\mathcal{H}$  is not well defined unless  $V$  has the UIP.) Being able to plug in any  $v$  in (3.2.1.1) and obtain a residual will be useful in practice, when assessing the convergence of the algorithm (see Section 3.3). Whether a numerical solution verifies, at least approximately, a discrete UIP (and the remaining technical conditions of the Verification Theorem) becomes something to be checked separately.

*Remark 3.2.1.1.* Choosing the maximum arg-maximum in (3.2.1.3) is partly motivated by ensuring a discrete solution will inherit the property of Remark 2.1.2.4. (The proof remains the same, for the discretizations of Section 3.2.6.) We will also motivate it in terms of the proposed numerical algorithm in Remark 3.2.6.4. Note that in [ABM<sup>+</sup>19] (Chapter 2) the minimum arg-maximum was used instead for both players. Nevertheless, the replication of property (ii), Lemma 3.1.1.7, dictates that it is only possible to be consistent with [ABM<sup>+</sup>19] for one of the two players (in this case, the opponent).

### 3.2.2 Iterative algorithm for symmetric games

This section introduces the iterative algorithm developed to solve system (3.2.1.1), which builds on Algorithm 2.2.1.1 by harnessing the symmetry of the problem and dispenses with the need for a relaxation scheme altogether. It is presented with a pseudocode that highlights the mimicking of system (3.2.1.1) and the intuition behind the algorithm; namely:

- The player starts with some suboptimal strategy  $\varphi^0 = (I^0, \delta^0)$  and payoff  $v^0$ , to which the opponent responds symmetrically, resulting in a gain for the player (first equation of (3.2.1.1); lines 1, 2 and 4 of Algorithm 3.2.2.1).
- The player improves her strategy by choosing the optimal response, i.e., by solving a single-player impulse control problem through a policy-iteration-type algorithm (second equation of (3.2.1.1); line 5 of Algorithm 3.2.2.1).

- This procedure is iterated until reaching a stopping criteria (lines 6-8 of Algorithm 3.2.2.1).

*Notation 3.2.2.1.* In the following:  $G_{<0}$  and  $G_{\leq 0}$  represent the sets of grid points which are negative and nonpositive respectively, and  $\Phi$  the set of (*discrete*) *strategies*

$$\Phi := \{\varphi = (I, \delta) : I \subseteq G_{<0} \text{ and } \delta \in Z\}. \quad (3.2.2.1)$$

Set complements are taken with respect to the whole grid,  $Id : \mathbb{R}^G \rightarrow \mathbb{R}^G$  is the identity operator; and given a linear operator  $O : \mathbb{R}^G \rightarrow \mathbb{R}^G \simeq \mathbb{R}^{G \times G}$ ,  $v \in \mathbb{R}^G$  and subsets  $I, J \subseteq G$ ,  $v_I \in \mathbb{R}^I$  denotes the restriction of  $v$  to  $I$  and  $O_{IJ} \in \mathbb{R}^{I \times J}$  the submatrix/operator with rows in  $I$  and columns in  $J$ .

---

**Algorithm 3.2.2.1** Iterative algorithm for symmetric games

---

Set  $tol, scale > 0$ .

- 1: Choose initial guess:  $v^0 \in \mathbb{R}^G$
  - 2: Set  $I^0 = \{Lv^0 + f \leq Mv^0 - v^0\} \cap G_{<0}$  and  $\delta^0 = \delta^*(v^0)$
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:  $v^{k+1/2} = \begin{cases} H(\delta^k)v^k & \text{on } -I^k \\ v^k & \text{on } (-I^k)^c \end{cases}$
  - 5:  $(v^{k+1}, I^{k+1}, \delta^{k+1}) = \text{SOLVEIMPULSECONTROL}(v^{k+1/2}, (-I^k)^c)$
  - 6: **if**  $\|(v^{k+1} - v^k) / \max\{|v^{k+1}|, scale\}\|_\infty < tol$  **then**
  - 7:     break from the iteration
  - 8: **end if**
  - 9: **end for**
- 

The *scale* parameter in line 5 of Algorithm 3.2.2.1, used throughout the literature by Forsyth, Labahn and coauthors [AF16, FL07, HFL12, HFL13], prevents the enforcement of unrealistic levels of accuracy for points  $x$  where  $v^{k+1}(x) \approx 0$ . Additionally, note that having chosen the initial guess for the payoff  $v^0$ , the initial guess for the strategy is taken as the one induced by  $v^0$ . (The alternative expression for the intervention region gives the same as  $\{Mv^0 - v^0 = 0\}$  for a solution of (3.2.1.1).)

Line 5 of Algorithm 3.2.2.1 assumes we have a subroutine  $\text{SOLVEIMPULSECONTROL}(w, D)$  that solves the constrained QVI problem:

$$\max\{Lv + f, Mv - v\} = 0 \text{ on } D, \quad \text{subject to } v = w \text{ on } D^c \quad (3.2.2.2)$$

for fixed  $G_{\leq 0} \subseteq D \subseteq G$  (approximate continuation region of the opponent) and  $w \in \mathbb{R}^G$  (approximate payoff due to the opponent's intervention). Although we only need to solve for  $\tilde{v} = v_D$ , the value of  $v_{D^c} = w_{D^c}$  impacts the solution both when restricting the equations and when applying the nonlocal operator  $M$ . Hence, the approximate payoff  $v^{k+1/2}$  fed to the subroutine serves to pass on the gain that resulted from the opponent's intervention and as an initial guess if desired (more on this on Remark 3.2.3.3).

The remaining of this section consists in establishing an equivalence between problem (3.2.2.2) and a classical (unconstrained) QVI problem of impulse control. This allows us to

prove the existence and uniqueness of its solution. In particular, we will see that `SOLVEIMPULSECONTROL` can be defined, if wanted, by policy iteration. However, we will present in the next section an alternative method that performs better in many practical situations and, in particular, in the examples treated in Section 3.3. Let us suppose from here onwards that the following assumptions hold true (see Appendix A for the relevant Definitions A.1.1.1):

- (A0) For each strategy  $\varphi = (I, \delta) \in \Phi$  and  $x \in I$ , there exists a walk in  $\text{graph} B(\delta)$  from row  $x$  to some row  $y \in I^c$ .
- (A1)  $-L$  is a strictly diagonally dominant (SDD)  $L_0$ -matrix and, for each  $\delta \in Z$ ,  $Id - B(\delta)$  is a weakly diagonally dominant (WDD)  $L_0$ -matrix.

*Remark 3.2.2.2 (Interpretation).* Assumptions (A0),(A1) are (H2),(H3) in [AF16]. For an impulse operator (say, ' $B(\delta)v(x) = v(x + \delta)$ '), (A0) asserts that the player always wants to shift states in her intervention region to her continuation region through finitely many impulses. (This does not take into account the opponent's response.) On the other hand, if problem (3.2.2.2) was rewritten as a fixed point problem, (A1) would essentially mean that the uncontrolled operator is contractive while the controlled ones are nonexpansive (see [CMS07] and [AF16, Sect.4]).

**Theorem 3.2.2.3.** *Assume (A0),(A1). Then, for every  $G_{\leq 0} \subseteq D \subseteq G$  and  $w \in \mathbb{R}^G$ , there exists a unique  $v^* \in \mathbb{R}^G$  that solves the constrained QVI problem (3.2.2.2). Further,  $v_D^*$  is the unique solution of*

$$\max \left\{ \tilde{L}\tilde{v} + \tilde{f}, \tilde{M}\tilde{v} - \tilde{v} \right\} = 0, \quad (3.2.2.3)$$

where  $\tilde{L} := L_{DD}$ ,  $\tilde{f} := f_D + L_{DD^c}w_{D^c}$ ,  $\tilde{Z} := \prod_{x \in D} Z(x)$ ,  $\tilde{c}(\tilde{\delta}) := c(\tilde{\delta}) - B(\tilde{\delta})_{DD^c}w_{D^c}$  and  $\tilde{B}(\tilde{\delta}) := B(\tilde{\delta})_{DD}$  for  $\tilde{\delta} \in \tilde{Z}$ ; and  $\tilde{M}\tilde{v} := \max_{\tilde{\delta} \in \tilde{Z}} \{ \tilde{B}(\tilde{\delta})\tilde{v} - \tilde{c}(\tilde{\delta}) \}$  for  $\tilde{v} \in \mathbb{R}^D$ .

Additionally, for any initial guess, the sequence  $(\tilde{v}^k) \subseteq \mathbb{R}^D$  defined by policy iteration [AF16, Thm.4.3] applied to problem (3.2.2.3), converges exactly to  $v_D^*$  in at most  $|\tilde{\Phi}|$  iterations, with  $\tilde{\Phi} := \{ \tilde{\varphi} = (I, \tilde{\delta}) : I \subseteq G_{<0} \text{ and } \tilde{\delta} \in \tilde{Z} \}$  the set of restricted admissible strategies.<sup>5</sup>

*Proof.* The equivalence between problems (3.2.2.2) and (3.2.2.3) is due to simple algebraic manipulation and  $B(\delta), c(\delta)$  being row-decoupled for every  $\delta \in Z$ . ( $B(\tilde{\delta}), c(\tilde{\delta})$  are thus defined in the obvious way for each  $\tilde{\delta} \in \tilde{Z}$ .)

The rest of the proof is mostly as in [AF16, Thm.4.3]. Let  $\tilde{I}d = Id_{DD}$ . Each intervention region  $I$ , can be identified with its indicator  $\tilde{\psi} = \mathbb{1}_I \in \{0, 1\}^D$  since  $D \supseteq I$ , and each  $\tilde{\psi}$  can be identified in turn with a diagonal matrix having  $\tilde{\psi}$  as main diagonal:  $\tilde{\Psi} = \text{diag}(\tilde{\psi}) \in \mathbb{R}^{D \times D}$ . Then problem (3.2.2.3) takes the form of the classical Bellman problem

$$\max_{\tilde{\varphi} \in \tilde{\Phi}} \left\{ -A(\tilde{\varphi})v + b(\tilde{\varphi}) \right\} = 0, \quad (3.2.2.4)$$

if we take

$$A(\tilde{\varphi}) = -(\tilde{I}d - \tilde{\Psi})\tilde{L} + \tilde{\Psi}(\tilde{I}d - \tilde{B}(\tilde{\delta})) \quad \text{and} \quad b(\tilde{\varphi}) = (\tilde{I}d - \tilde{\Psi})\tilde{f} - \tilde{\Psi}\tilde{c}.$$

---

<sup>5</sup> $|A|$  denotes the cardinal of set  $A$ .

Note that  $\tilde{\Phi}$  can be identified with the Cartesian product

$$\tilde{\Phi} = \left( \prod_{x \in G_{<0}} \{0, 1\} \times Z(x) \right) \times \left( \prod_{x \in D \setminus G_{<0}} \{0\} \times Z(x) \right)$$

and  $A(\tilde{\varphi}), b(\tilde{\varphi})$  are row-decoupled for every  $\tilde{\varphi} \in \tilde{\Phi}$ . Since  $\tilde{\Phi}$  is finite, all we need to show is that the matrices  $A(\tilde{\varphi})$  are monotone (Definitions A.1.1.1 and [BMZ09, Thm.2.1]). Let us check the stronger property (Thm. A.1.1.4 and Prop. A.1.1.3) of being weakly chained diagonally dominant (WCDD)  $L_0$ -matrices (see Definitions A.1.1.1).

If (A0) and (A1) also held true for the restricted matrices and strategies, the conclusion would follow. While (A1) is clearly inherited, (A0) may fail to do so, but only in non-problematic cases. To see this, let  $\tilde{\varphi} = (I, \tilde{\delta}) \in \tilde{\Phi}$ ,  $x \in I \subseteq D$  and let  $\delta \in Z$  be some extension of  $\tilde{\delta}$ . Note that row  $x$  of  $A(\tilde{\varphi})$  is WDD. We want to show that there is a walk in  $\text{graph}A(\tilde{\varphi})$  from  $x$  to an SDD row.

By (A0) there must exist some walk  $x \rightarrow y_1 \rightarrow \dots \rightarrow y_n \in I^c$  in  $\text{graph}B(\delta)$ . If this is in fact a walk from  $x$  to  $I^c \cap D$  in  $\text{graph}\tilde{B}(\tilde{\delta})$ , then it verifies the desired property (just as in [AF16, Thm.4.3]). If not, then there must be a first  $1 \leq m \leq n$  such that the subwalk  $x \rightarrow y_1 \rightarrow \dots \rightarrow y_m$  is in  $\text{graph}\tilde{B}(\tilde{\delta})$  but  $y_{m+1} \notin D$ . Since  $y_m \rightarrow y_{m+1}$  is an edge in  $\text{graph}B(\delta)$ , we have  $B(\delta)_{y_m, y_{m+1}} \neq 0$  and the WDD row (by (A1))  $y_m$  of  $\tilde{I}d - \tilde{B}(\tilde{\delta})$  is in fact SDD. Meaning that the subwalk  $x \rightarrow y_1 \rightarrow \dots \rightarrow y_m$  verifies the desired property instead.  $\square$

*Remark 3.2.2.4. (Practical considerations)* 1. While convergence is guaranteed to be exact, floating point arithmetic can bring about stagnation [HFL13]. A stopping criteria like that of Algorithm 3.2.2.1 should be used in those cases, with a tolerance  $\ll \text{tol}$ . 2. The solution of system (3.2.2.3) does not change if one introduces a scaling factor  $\lambda > 0$ :  $\max \{ \tilde{L}\tilde{v} + \tilde{f}, \lambda(\tilde{M}\tilde{v} - \tilde{v}) \} = 0$  [AF16, Lem.4.1]. This problem-specific parameter is typically added in the implementation to enhance performance [HFL12, AF16]. It can intuitively be thought as a units adjustment.

### 3.2.3 Iterative subroutine for impulse control

Due to Theorem 3.2.2.3, a sensible choice for SOLVEIMPULSECONTROL is the classical policy iteration algorithm [AF16, Thm.4.3] applied to (3.2.2.3) (i.e., [BMZ09, Ho-1] applied to (3.2.2.4)), adding an appropriately chosen scaling factor  $\lambda$  to improve efficiency (Remark 3.2.2.4 2). It does, however, bear some drawbacks in practice. At each iteration, one needs to solve the system  $-A(\tilde{\varphi}^k)v^{k+1} + b(\tilde{\varphi}^k) = 0$  for some  $\tilde{\varphi}^k \in \tilde{\Phi}$ . While the matrix  $\tilde{L}$  typically has a good sparsity pattern in applications (often tridiagonal), the presence of  $\tilde{B}(\tilde{\delta}^k)$  prevents  $A(\tilde{\varphi}^k)$  from inheriting the same structure and makes the resolution of the previous system a lot more costly, even when using a good ordering technique. An exact resolution often becomes prohibitive and an iterative method, such as GMRES or BiCGSTAB with preconditioning, is relied upon.

To circumvent the previous problem, we propose to choose SOLVEIMPULSECONTROL as an instance of a very general class of algorithms known as *fixed-point policy iteration* [HFL12, Cli07]. As far as the author knows, this application to impulse control was never done



in the past other than heuristically in Chapter 2. Instead of solving  $-A(\tilde{\varphi}^k)v^{k+1} + b(\tilde{\varphi}^k) = 0$  at the  $k$ -th iteration, we will solve

$$\underbrace{(\tilde{I}d - \tilde{\Psi}^k)\tilde{L} - \tilde{\Psi}^k}_{-\tilde{\mathbb{A}}(\tilde{\varphi}^k)} v^{k+1} + \underbrace{\tilde{\Psi}^k \tilde{B}(\tilde{\delta}^k)}_{\tilde{\mathbb{B}}(\tilde{\varphi}^k)} v^k + \underbrace{b(\tilde{\varphi}^k)}_{\tilde{\mathbb{C}}(\tilde{\varphi}^k)} = 0, \quad (3.2.3.1)$$

(scaled by  $\lambda$ ) where the previous iterate value  $v^k$  is given and  $\tilde{\Psi}^k$  is the diagonal matrix with  $\psi^k$  as diagonal. In other words, we split the original policy matrix  $A(\tilde{\varphi}) = \tilde{\mathbb{A}}(\tilde{\varphi}) - \tilde{\mathbb{B}}(\tilde{\varphi})$  and we apply a one-step fixed-point approximation,

$$\tilde{\mathbb{A}}(\tilde{\varphi}^k)\tilde{v}^{k+1} = \tilde{\mathbb{B}}(\tilde{\varphi}^k)\tilde{v}^k + \tilde{\mathbb{C}}(\tilde{\varphi}^k), \quad (3.2.3.2)$$

at each iteration of Howard's algorithm. The resulting method can be expressed as follows (*tol* and *scale* are as in Algorithm 3.2.2.1):

---

**Subroutine 3.2.3.1**  $(v, I, \delta) = \text{SOLVEIMPULSECONTROL}(w, D)$

---

**Inputs**  $w \in \mathbb{R}^G$  and solvency region  $G_{\leq 0} \subseteq D \subseteq G$

**Outputs**  $v \in \mathbb{R}^G$ ,  $I \subseteq G_{<0}$ ,  $\delta \in Z$

---

Set scaling factor  $\lambda > 0$  and  $0 < \widetilde{tol} \ll tol$ .

// Restrict constrained problem

- 1:  $\tilde{L} := L_{DD}$ ,  $\tilde{f} := f_D + L_{DD^c}w_{D^c}$
- 2:  $\tilde{Z} := \prod_{x \in D} Z(x)$ ,  $\tilde{c}(\tilde{\delta}) := c(\tilde{\delta}) - B(\tilde{\delta})_{DD^c}w_{D^c}$ ,  $\tilde{B}(\tilde{\delta}) := B(\tilde{\delta})_{DD}$  for  $\tilde{\delta} \in \tilde{Z}$
- 3:  $\tilde{M}\tilde{v} := \max_{\tilde{\delta} \in \tilde{Z}} \{\tilde{B}(\tilde{\delta})\tilde{v} - \tilde{c}(\tilde{\delta})\}$  for  $\tilde{v} \in \mathbb{R}^D$ ,  $\tilde{I}d := Id_{DD}$

// Solve restricted problem

- 4: Choose initial guess:  $\tilde{v}^0 \in \mathbb{R}^D$ ,  $I^0 \subseteq G_{<0}$
  - 5: **for**  $k = 0, 1, 2, \dots$  **do**
  - 6:  $\tilde{L}_{ij}^k = \begin{cases} \tilde{L}_{ij} & \text{if } x_i \in D \setminus I^k \\ -\tilde{I}d_{ij} & \text{if } x_i \in I^k \end{cases} \quad \tilde{f}^k = \begin{cases} \tilde{f} & \text{on } D \setminus I^k \\ \tilde{M}\tilde{v}^k & \text{on } I^k \end{cases}$
  - 7:  $\tilde{v}^{k+1}$  solution of  $\tilde{L}^k\tilde{v} + \tilde{f}^k = 0$
  - 8:  $I^{k+1} = \{\tilde{L}\tilde{v}^{k+1} + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^{k+1} - \tilde{v}^{k+1})\}$
  - 9: **if**  $\|(\tilde{v}^{k+1} - \tilde{v}^k) / \max\{\tilde{v}^{k+1}, scale\}\|_\infty < \widetilde{tol}$  **then**
  - 10:  $v = \begin{cases} \tilde{v}^{k+1} & \text{on } D \\ w_{D^c} & \text{on } D^c \end{cases}$ ,  $I = I^{k+1}$ ,  $\delta = \delta^*(v)$  and break from the iteration
  - 11: **end if**
  - 12: **end for**
- 

Lines 1-3 of Subroutine 3.2.3.1 deal with restricting the constrained problem, while the rest give a routine that can be applied to any QVI of the form (3.2.2.3). Starting from some suboptimal  $\tilde{v}^0$  and  $I^0$ , one computes a new payoff  $\tilde{v}^1$  by solving the coupled equations  $\tilde{M}\tilde{v}^0 - \tilde{v}^1 = 0$  on  $I^0$  and  $\tilde{L}\tilde{v}^1 + \tilde{f} = 0$  outside  $I^0$ . A new intervention region  $I^1 = \{\tilde{L}\tilde{v}^1 + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^1 - \tilde{v}^1)\}$  is defined and the procedure is iterated.

Algorithmically, the difference with classical policy iteration is that  $\tilde{v}^{k+1}$  is computed in Line 7 with a fixed obstacle  $\tilde{M}v^k$ , changing a quasivariational inequality for a variational one. The resulting method is intuitive and simple to implement, and the linear system (3.2.3.1) (Line 7) inherits the sparsity pattern of  $\tilde{L}$ . This will normally result in more precision and lower space and time complexity.<sup>6</sup> For example, for an SDD tridiagonal  $\tilde{L}$ , the system can be solved (exactly in exact arithmetic, and stably in floating point one) in  $O(n)$  operations, with  $n = |D|$  [How60, Sect.9.5]. The matrix-vector multiply  $\tilde{B}(\tilde{\delta}^k)\tilde{v}^k$  can take at most  $O(n^2)$  operations, but will reduce to  $O(n)$  for standard discretizations of impulse operators.

It is also worth mentioning that Subroutine 3.2.3.1 differs from the so-called iterated optimal stopping [CØS02, ØS09] in that the latter solves  $\max \{ \tilde{L}\tilde{v}^{k+1} + \tilde{f}, \tilde{M}\tilde{v}^k - \tilde{v}^{k+1} \} = 0$  exactly at the  $k$ -th iteration (by running a full subroutine of Howard's algorithm with fixed obstacle), while the former only performs one approximation step.

To establish the convergence of Subroutine 3.2.3.1 in the present framework, we add the following assumption:

(A2)  $B(\delta)$  has nonnegative diagonal elements for all  $\delta \in Z$ .

*Remark 3.2.3.1.* (A2) and the requirement of (A1) that  $Id - B(\delta)$  be a WDD  $L_0$ -matrix are equivalent to  $B(\delta)$  being substochastic (see Appendix A). This is standard for impulse operators (see Section 3.2.6) and other applications of fixed-point policy iteration [HFL12, Sect.4-5].

**Theorem 3.2.3.2.** *Assume (A0)–(A2) and set  $I^0 = \emptyset$ . Then, for every  $G_{\leq 0} \subseteq D \subseteq G$  and  $w \in \mathbb{R}^G$ , the sequence  $(\tilde{v}^k)$  defined by SOLVEIMPULSECONTROL( $w, D$ ) is monotone increasing for  $k \geq 1$  and converges to the unique solution of (3.2.2.3).*

*Proof.* We can assume without loss of generality that  $\lambda = 1$ . Subroutine 3.2.3.1 takes the form of a fixed-point policy iteration algorithm as per (3.2.3.2). Assumptions (A0),(A1) ensure (3.2.2.3) has a unique solution (Theorem 3.2.3.1) and that this scheme satisfies [HFL12, Cond.3.1 (i),(ii)]. That is,  $\tilde{A}(\tilde{\varphi})$  and  $\tilde{A}(\tilde{\varphi}) - \tilde{B}(\tilde{\varphi})$  are nonsingular  $M$ -matrices (see proof of Theorem 3.2.3.1 and Appendix A) and all coefficients are bounded since  $\tilde{\Phi}$  is finite. In [HFL12, Thm.3.4] convergence is proved under one additional assumption of  $\|\cdot\|_\infty$ -contractiveness [HFL12, Cond.3.1 (iii)], which is not verified in our case. However, the same computations show that the scheme satisfies

$$\tilde{A}(\tilde{\varphi}^k)(\tilde{v}^{k+1} - \tilde{v}^k) \geq \tilde{B}(\tilde{\varphi}^{k-1})(\tilde{v}^k - \tilde{v}^{k-1}) \text{ for all } k \geq 1. \quad (3.2.3.3)$$

Since  $I^0 = \emptyset$ , and due to (A1) and (A2),  $\tilde{B}(\tilde{\varphi}^0) = 0$  and  $\tilde{B}(\tilde{\varphi}^k) \geq 0$  for all  $k$ . Thus,  $(\tilde{v}^k)_{k \geq 1}$  is increasing by monotonicity of  $\tilde{A}(\tilde{\varphi}^k)$ . Furthermore, it must be bounded, since for all  $k \geq 1$ :

$$\tilde{A}(\tilde{\varphi}^k)\tilde{v}^{k+1} = \tilde{B}(\tilde{\varphi}^k)\tilde{v}^k + \tilde{C}(\tilde{\varphi}^k) \leq \tilde{B}(\tilde{\varphi}^k)\tilde{v}^{k+1} + \tilde{C}(\tilde{\varphi}^k),$$

which gives  $\tilde{v}^{k+1} \leq (\tilde{A}(\tilde{\varphi}^k) - \tilde{B}(\tilde{\varphi}^k))^{-1}\tilde{C}(\tilde{\varphi}^k) \leq \max_{\tilde{\varphi} \in \tilde{\Phi}} (\tilde{A}(\tilde{\varphi}) - \tilde{B}(\tilde{\varphi}))^{-1}\tilde{C}(\tilde{\varphi})$ . Hence,  $(\tilde{v}^k)$  converges.  $\square$

---

<sup>6</sup>This refers to the overall time complexity of Subroutine 3.2.3.1, including the linear solves, and not just the number of outer loops.

*Remark 3.2.3.3.* In light of Theorem 3.2.3.2, moving forward we will set  $I^0 = \emptyset$  in Subroutine 3.2.3.1. It is natural however to choose  $\tilde{v}^0 = w_D$  and  $I^0 = \{\tilde{L}\tilde{v}^0 + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^0 - \tilde{v}^0)\}$ . The experiments performed with the latter choice displayed (non-monotone) convergence and usually a faster one; but this is not proved here. Additionally, exact convergence was often observed.

### 3.2.4 Overall routine as a fixed-point policy-iteration-type method

The system of QVIs (3.2.1.1) cannot be reduced in any apparent way to a Bellman formulation (3.2.1.4) (see comments preceding equation). Notwithstanding, we shall see that Algorithm 3.2.2.1 does take a very similar form to a fixed-point policy iteration algorithm as in (3.2.3.2) for some appropriate  $\mathbb{A}, \mathbb{B}, \mathbb{C}$ . Further, assumptions resembling those of the classical case [HFL12] will be either satisfied or imposed to study its convergence. This is independent of whether SOLVEIMPULSECONTROL is chosen as in Subroutine 3.2.3.1 or Howard's algorithm (Theorem 3.2.2.3), although we shall assume that the outputs  $I^k, \delta^k$  in the latter case are defined in the same way as in the former. The matrix and graph-theoretic definitions and properties used throughout this section can be found in Appendix A.

*Notation 3.2.4.1.* We identify each intervention region  $I \subseteq G_{<0}$  with its indicator function  $\psi = \mathbb{1}_I \in \{0, 1\}^G$  and each  $\psi$  with a diagonal matrix having  $\psi$  as main diagonal:  $\Psi = \text{diag}(\psi) \in \mathbb{R}^{G \times G}$ . The sequences  $(v^k)$  and  $(\varphi^k)$ , with  $\varphi^k = (\psi^k, \delta^k)$ , are the ones generated by Algorithm 3.2.2.1. We consider  $v^* \in \mathbb{R}^G$  fixed and  $\varphi^* = (\psi^*, \delta^*(v^*))$  the induced strategy with  $\psi^* := \{Lv^* + f \leq Mv^* - v^*\} \cap G_{<0}$ .

**Proposition 3.2.4.2.** *Assume (A0)–(A2). Then,*

$$\mathbb{A}(\varphi^k, \varphi^{k+1})v^{k+1} = \mathbb{B}(\varphi^k)v^k + \mathbb{C}(\varphi^k, \varphi^{k+1}), \text{ where:} \quad (3.2.4.1)$$

- (i)  $\psi^k = \mathbb{1}_{\{Lv^k + f \leq Mv^k - v^k\} \cap G_{<0}}$  and  $\delta^k \in \arg \max_{\delta \in Z} \{B(\delta)v^k - c(\delta)\}$ .
- (ii)  $\mathbb{A}(\varphi, \bar{\varphi}) := Id - (Id - \bar{\Psi} - S\Psi S)(Id + L) - \bar{\Psi}B(\bar{\delta})$  is a WCDD  $L_0$ -matrix, and thus a nonsingular  $M$ -matrix.
- (iii)  $\mathbb{B}(\varphi) := S\Psi B(\delta)S = \text{diag}(S\psi)SB(\delta)S$  is substochastic.
- (iv)  $\mathbb{C}(\varphi, \bar{\varphi}) := (Id - \bar{\Psi} - S\Psi S)f - \bar{\Psi}c(\bar{\delta}) + S\Psi Sg(S\delta)$ .

*Proof.* Using that  $(v^{k+1}, I^{k+1}, \delta^{k+1}) = \text{SOLVEIMPULSECONTROL}(v^{k+1/2}, (-I^k)^c)$  solves the constrained QVI problem (3.2.2.2) for  $D = (-I^k)^c$  and  $w = v^{k+1/2}$  (Theorem 3.2.2.3 or 3.2.3.2), the recurrence relation (3.2.4.1) results from simple algebraic manipulation.

Given  $\varphi, \bar{\varphi} \in \Phi$ , (A0) and (A1) ensure  $\mathbb{A}(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix, while (A1) and (A2) imply  $\mathbb{B}(\varphi)$  is substochastic.  $\square$

The following corollary is immediate by induction. It gives a representation of the sequence of payoffs in terms of the improving strategies throughout the algorithm.

**Corollary 3.2.4.3.** *Assume (A0)–(A2). Then,*

$$v^{k+1} = \left( \prod_{j=k}^0 \mathbb{A}^{-1} \mathbb{B}(\varphi^j, \varphi^{j+1}) \right) v^0 + \sum_{n=0}^k \left( \prod_{j=k}^{n+1} \mathbb{A}^{-1} \mathbb{B}(\varphi^j, \varphi^{j+1}) \right) \mathbb{A}^{-1} \mathbb{C}(\varphi^n, \varphi^{n+1}).^7 \quad (3.2.4.2)$$

We now establish some properties of the strategy-dependent matrix coefficients that will be useful in the sequel. Given a WDD (resp. substochastic) matrix  $A \in \mathbb{R}^{G \times G}$ , we define its set of ‘non-trouble states’ (or rows) as

$$J[A] := \{x \in G : \text{row } x \text{ of } A \text{ is SDD}\} \quad (\text{resp. } \hat{J}[A] := \{x \in G : \text{row } x \text{ of } A \text{ sums less than one}\}),$$

and its *index of connectivity*  $\text{con}A$  (resp. *index of contraction*  $\widehat{\text{con}}A$ ) by computing for each state the least length that needs to be walked on  $\text{graph}A$  to reach a non-trouble one, and then taking the maximum over all states (more details in Appendix A). This recently introduced concept gives an equivalent characterization of the WCDD property for a WDD matrix as  $\text{con}A < +\infty$ , and can be efficiently checked for sparse matrices in  $O(|G|)$  operations [Azi19a]. On the other hand, if  $A$  is substochastic then  $\widehat{\text{con}}A < \infty$  if and only if its spectral radius verifies  $\rho(A) < 1$  (Theorem A.1.1.6). The proof of the following lemma can be found in Appendix A.

**Lemma 3.2.4.4.** *Assume (A0)–(A2). Then for all  $\varphi, \bar{\varphi} \in \Phi$ ,  $\mathbb{A}^{-1} \mathbb{B}(\varphi, \bar{\varphi})$  is substochastic,  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WDD  $L_0$ -matrix and  $\widehat{\text{con}}[\mathbb{A}^{-1} \mathbb{B}(\varphi, \bar{\varphi})] \leq \text{con}[(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})]$ .*

As previously mentioned, system (3.2.1.1) may have no solution. The matrix coefficients introduced in this section allow us to algebraically characterize the existence of such solutions through strategy-dependent linear systems of equations. For each strategy  $\varphi \in \Phi$ , let  $\mathbb{O}(\varphi) : \mathbb{R}^G \rightarrow \mathbb{R}^G$  be the operator that applies  $Id + L$ ,  $M$  and  $H$  on the continuation, intervention and opponent’s intervention regions, respectively. That is,  $\mathbb{O}(\varphi) = Id - (\mathbb{A} - \mathbb{B})(\varphi^*, \varphi^*) = (Id - \Psi - S\Psi S)(Id + L) + \Psi B(\delta) + S\Psi B(\delta)S$ . Then the following equivalences are immediate.

**Proposition 3.2.4.5.** *Assume (A0)–(A2). Then the following statements are equivalent:*

- (i)  $v^*$  solves the system of QVIs (3.2.1.1).
- (ii)  $\mathbb{A}(\varphi^*, \varphi^*)v^* = \mathbb{B}(\varphi^*)v^* + \mathbb{C}(\varphi^*, \varphi^*)$ .
- (iii)  $v^* = \mathbb{O}(\varphi^*)v^* + \mathbb{C}(\varphi^*, \varphi^*)$ .

As mentioned in Remark 3.2.2.2, Assumption (A0) constrains the type of strategies the player can use, but without taking into account the opponent’s response. This is enough for the single-player constrained problems to have a solution and, therefore, for Algorithm 3.2.2.1 to be well defined. But we cannot expect this restriction to be sufficient in the study of the two-player game and the convergence of the overall routine.

In order to improve the result of Proposition 3.2.4.5 let us consider the following stronger version of (A0) reflecting the interaction between the player and the opponent.

<sup>7</sup>For any index  $i \leq k$ ,  $\prod_{j=k}^i A^j := A^k A^{k-1} \dots A^i$ .

(A0') For each pair of strategies  $\varphi, \bar{\varphi} \in \Phi$ , and for each  $x \in \bar{I} \cup (-I)$ , there exists a walk in  $\text{graph}(\bar{\Psi}B(\bar{\delta}) + S\Psi B(\delta)S)$  from row  $x$  to some row  $y \in \bar{C} \cap C$ , where  $\bar{C} = \bar{I}^c$ ,  $C = I^c$ .

*Remark 3.2.4.6. (Interpretation)* If  $\bar{\varphi}, -\varphi$  are the strategies used by the player and the opponent respectively,<sup>8</sup> then (A0') asserts that states in their intervention regions will eventually be shifted to the common continuation region. This precludes infinite simultaneous interventions and emulates the admissibility condition of the continuous-state case. Fixing  $I = \emptyset$  we recover (A0). Additionally, (A0') together with (A1) imply that  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix, hence an  $M$ -matrix. This is another one of the assumptions of the classical fixed-point policy iteration [HFL12].

Under this new assumption, the  $\varphi^* = \varphi^*(v^*)$ -dependent systems of Proposition 3.2.4.5 will admit a unique solution. Then solving the original problem (3.2.1.1) amounts to finding  $v^* \in \mathbb{R}^G$  that solves its induced linear system of equations.

**Proposition 3.2.4.7.** *Assume (A0'), (A1), (A2). In the context of Proposition 3.2.4.5, the following statements are also equivalent:*

$$(iv) \ v^* = (\mathbb{A} - \mathbb{B})^{-1}\mathbb{C}(\varphi^*, \varphi^*).$$

$$(v) \ v^* = (Id - \mathbb{A}^{-1}\mathbb{B})^{-1}\mathbb{A}^{-1}\mathbb{C}(\varphi^*, \varphi^*) = \sum_{n \geq 0} (\mathbb{A}^{-1}\mathbb{B})^n \mathbb{A}^{-1}\mathbb{C}(\varphi^*, \varphi^*). \quad (\text{cf. equation (3.2.4.2).})$$

*Proof.* Both expressions result from rewriting and solving the systems of Proposition 3.2.4.5. Assumptions (A0'), (A1) guarantee that  $(\mathbb{A} - \mathbb{B})(\varphi^*, \varphi^*)$  is WCDD and, hence, nonsingular. Then (v) is due to Lemma 3.2.4.4, Theorem A.1.1.6 and the matrix power series expansion  $(1 - X)^{-1} = \sum_{n \geq 0} X^n$ , when  $\rho(X) < 1$ .  $\square$

### 3.2.5 Convergence analysis

We now study the convergence properties of Algorithm 3.2.2.1. Henceforth, the UIP refers to the obvious discrete analogue of Definition 2.1.2.1, where we replace the domain  $\mathbb{R}$ , the impulse constraints  $\mathcal{Z}$  and the operator  $\mathcal{M}$  by their discretizations  $G$ ,  $Z$  and  $M$  respectively.

The obvious first question to address is whether when Algorithm 3.2.2.1 converges, it does so to a solution of the system of QVIs (3.2.1.1). Unlike in the classical Bellman problem (3.2.1.4), problem (3.2.1.1) is intrinsically dependent on the particular strategy chosen by the player (see Propositions 3.2.4.5 and 3.2.4.7). Accordingly, we start with a lemma addressing what can be said about the convergence of the strategies  $(\varphi^k)$  when the payoffs  $(v^k)$  converge.

*Notation 3.2.5.1.*  $\partial I^* := \{Lv^* + f = Mv^* - v^*\} \cap G_{<0}$  denotes the ‘border’ of the intervention region  $\{Lv^* + f \leq Mv^* - v^*\} \cap G_{<0}$  defined by  $v^*$ .

**Lemma 3.2.5.2.** *Assume (A0)-(A2) and suppose  $v^k \rightarrow v^*$ . Then:*

$$(i) \ \psi^k \rightarrow \psi^* \text{ in } (\partial I^*)^c \text{ and } Mv^k \rightarrow Mv^*.$$

---

<sup>8</sup>The slight abuse of notation  $-\varphi$  stands for the strategy symmetric to  $\varphi$ , i.e.,  $-\varphi = (-I, -\delta(-x))$ .

(ii) If  $\bar{\psi}, \bar{\delta}$  are any two limit points of  $(\psi^k), (\delta^k)$  resp.,<sup>9</sup> then

$$\bar{\delta} \in \arg \max_{\delta \in Z} \{B(\delta)v^* - c(\delta)\}, \quad \bar{\psi} = 0 \text{ on } G_{>=0} \quad \text{and} \quad \bar{\psi} \in \arg \max_{i \in \{0,1\}} \{O_i v^*\} \text{ on } G_{<0},$$

with  $O_0 v = Lv + f$  and  $O_1 v = Mv - v$ .

(iii) If  $v^*$  has the UIP, then  $\delta^k \rightarrow \delta^*(v^*)$  and  $Hv^k \rightarrow Hv^*$ .

*Proof.* That  $Mv^k \rightarrow Mv^*$  is clear by continuity of the operators  $B(\delta)$  and finiteness of  $Z$ .

Let  $x \in (\partial I^*)^c$  and suppose  $Lv^*(x) + f(x) < Mv^*(x) - v^*(x)$  (the other case being analogue). By continuity of  $L$  and  $M$  there must exist some  $k_0$  such that  $Lv^k(x) + f(x) < Mv^k(x) - v^k(x)$  for all  $k \geq k_0$ , which implies  $\psi^k(x) = 1 = \psi^*(x)$  for  $k \geq k_0$ .

The statement about  $\bar{\psi}, \bar{\delta}$  is proved as before by considering appropriate subsequences. Consequently, if  $v^*$  has the UIP, then necessarily  $\delta^k \rightarrow \delta^*(v^*)$  and  $Hv^k \rightarrow Hv^*$ .  $\square$

As a corollary we can establish that, should the sequence  $(v^k)$  converge, its limit must solve problem (3.2.1.1). If convergence is not exact however (i.e., in finite iterations), then we will ask that  $v^*$  verifies some of the properties of the Verification Theorem in Corollary 3.1.1.9. Namely, the UIP and a discrete analogue of the continuity in the border of the opponent's intervention region. We emphasize that our main motivation in solving system (3.2.1.1) relies in Corollary 3.1.1.9 and its framework. Additionally, in most practical situations and for fine-enough grids, one can intuitively expect the discretization of an equilibrium payoff as in Corollary 3.1.1.9 to inherit the UIP. Lastly, we note that the exact equality  $Lv^* + f = Mv^* - v^*$  will typically not be verified for any point in the grid in practice, giving  $\partial I^* = \emptyset$ .

**Corollary 3.2.5.3.** *Assume (A0)–(A2) and suppose  $v^k \rightarrow v^*$ . Then:*

(i) *If the convergence is exact, then  $v^*$  solves the system of QVIs (3.2.1.1).*

(ii) *If  $v^*$  has the UIP and  $Lv^* + f = Hv^* - v^*$  on  $-\partial I^*$ , then  $v^*$  solves (3.2.1.1).*

*Proof.* (i) is immediate from the definition of Algorithm 3.2.2.1.

In the general case, since  $\{0,1\}^G$  is finite, there is a subsequence of  $(\psi^k, \psi^{k+1})$  that converges to some pair  $(\psi, \bar{\psi})$ . Passing to such subsequence, by Lemma 3.2.5.2, the UIP of  $v^*$  and equation (3.2.4.1), we get that  $v^*$  solves the system  $\mathbb{A}(\varphi, \bar{\varphi})v^* = \mathbb{B}(\varphi)v^* + \mathbb{C}(\varphi, \bar{\varphi})$  for  $\varphi = (\psi, \delta^*(v^*))$ ,  $\bar{\varphi} = (\bar{\psi}, \delta^*(v^*))$  and  $\psi, \bar{\psi}$  coincide with  $\psi^*$  except possibly on  $\partial I^*$ . Thus, it only remains to show that  $v^*$  also solves the equations of the system (3.2.1.1) for any  $x \in \partial I^* \cup (-\partial I^*)$ .

For  $x \in \partial I^*$ , the previous is true by definition. Suppose now  $x \in -\partial I^* \subseteq \bar{I}^c$ . We have  $\psi^*(-x) = 1$ . If  $\psi(-x) = 1$ , there is nothing to prove. If  $\psi(-x) = 0$ , then  $x \in \bar{I}^c \cap (-I)^c$  and  $0 = Lv^*(x) + f(x) = Hv^*(x) - v^*(x)$ , where the last equality holds true by assumption.  $\square$

Lemma 3.2.5.2 shows to what extent the convergence of the payoffs imply the convergence of the strategies. The following theorem, of theoretical interest, establishes a reciprocal under the stronger assumption (A0'). In general, since the set of strategies  $\Phi$  is finite, the

<sup>9</sup>By 'limit point' we mean the limit of a convergent subsequence.

sequence of strategy-dependent coefficients of the fixed-point equations (3.2.4.1) will always be bounded and with finitely many limit points. However, if the approximating strategies are such that the former coefficients converge, then Algorithm 3.2.2.1 is guaranteed to converge. Further, instead of looking at the convergence of  $(\mathbb{A}, \mathbb{B}, \mathbb{C})(\varphi^k, \varphi^{k+1})$ , we can instead consider the weaker condition of  $(\mathbb{A}^{-1}\mathbb{B}, \mathbb{A}^{-1}\mathbb{C})(\varphi^k, \varphi^{k+1})$  converging.

**Theorem 3.2.5.4.** *Assume (A0'), (A1), (A2). If  $(\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1}))$  and  $(\mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1}))$  converge, then  $(v^k)$  converges.*

*Proof.* Set  $b = \lim_k \mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1})$ . Since  $\Phi$  is finite, there must exist  $k_0 \in \mathbb{N}$  and  $\varphi, \bar{\varphi} \in \Phi$  such that  $\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1}) = \mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})$  and  $\mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1}) = b$  for all  $k \geq k_0$ . Moreover, under our assumptions,  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix. Then Lemma 3.2.4.4 and Theorem A.1.1.6 imply that  $\mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})$  is contractive for some matrix norm. Lastly, note that the sequence of payoffs  $(v^k)_{k \geq k_0}$  now satisfies the classical (constant-coefficients) contractive fixed-point recurrence  $v^{k+1} = \mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})v^k + b$ , which converges to the unique fixed-point of the equation.  $\square$

The classical fixed-point policy-iteration framework [HFL12, Cli07] assumes uniform contractiveness in  $\|\cdot\|_\infty$  of the sequence of operators. This is a natural norm to consider in a context where matrices have properties defined row by row, such as diagonal dominance.<sup>10</sup> However, the authors mention convergence in experiments where only  $\|\cdot\|_\infty$ -non-expansiveness held true. The latter is the typical case in our context, for the matrices  $\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1})$ , which is why Theorem 3.2.5.4 relies on the fact that a spectral radius strictly smaller than one guarantees contractiveness in some matrix norm.

It is natural to ask whether there is some contractiveness condition that may account for the observations in [HFL12, Cli07] and that can be generalized to our context to further the study of Algorithm 3.2.2.1. Imposing a uniform bound on the spectral radii would not only be hard to check, but also difficult to manipulate, as the spectral radius is not sub-multiplicative.<sup>11</sup> Instead, we can consider the sequential indexes of contraction and connectivity, which naturally generalize those of the previous section by means of walks in the graph of a sequence of matrices (see Appendix A for more details). As before, they can be identified with one another (see Lemma A.1.1.5) and, given substochastic matrices, the sequential index of contraction tells us how many we need to multiply before the result becomes  $\|\cdot\|_\infty$ -contractive (Theorem A.1.1.7). Thus, let us consider a uniform bound on the following sequential indexes of connectivity:

(A0'') There exists  $m \in \mathbb{N}_0$  such that for any sequence of strategies  $(\bar{\varphi}^k) \subseteq \Phi$ ,

$$\text{con} \left[ \left( (\mathbb{A} - \mathbb{B})(\bar{\varphi}^k, \bar{\varphi}^{k+1}) \right)_k \right] \leq m.$$

*Remark 3.2.5.5.* Given  $\varphi, \bar{\varphi} \in \Phi$ , by considering the sequence  $\varphi, \bar{\varphi}, \varphi, \bar{\varphi}, \dots$ , we see that (A0'') implies (A0'). In fact, (A0'') can be interpreted as precluding infinite simultaneous impulses even when the players can adapt their strategies (cf. Remark 3.2.4.6) and imposing

<sup>10</sup>Recall that this norm can be computed as the maximum absolute value row sum.

<sup>11</sup> $\rho(AB) \leq \rho(A)\rho(B)$  does not hold in general when the matrices  $A$  and  $B$  do not commute.

that the number of shifts needed for any state to reach the common continuation region is bounded.

Under this stronger assumption, we have:

**Proposition 3.2.5.6.** *Assume (A0''),(A1),(A2). Then  $(v^k)$  is bounded.*

*Proof.* In a similar way to Lemma 3.2.4.4, one can check that under (A0''),(A1),(A2) we have the following uniform bound for the sequential indexes of contraction:

$$\widehat{\text{con}} \left[ \left( \mathbb{A}^{-1} \mathbb{B} \right) (\varphi^k, \varphi^{k+1}) \right]_k \leq m,$$

for any sequence of strategies  $(\varphi^k) \subseteq \Phi$ . In other words, multiplying any  $m + 1$  of the previous substochastic matrices results in a  $\|\cdot\|_\infty$ -contractive one. Furthermore, since  $\Phi^{m+1}$  is finite, there must be a uniform contraction constant  $C_1 < 1$ . Let  $C_2 > 0$  be a uniform bound for  $\mathbb{A}^{-1} \mathbb{C}$ . By the representation in Corollary 3.2.4.2,

$$\|v^{k+1}\|_\infty \leq \|v^0\|_\infty + C_2 \sum_{n=0}^k C_1^{\left[\frac{k-n}{m+1}\right]} \leq \|v^0\|_\infty + (m+1)C_2 \sum_{n=0}^\infty C_1^n < +\infty.^{12}$$

□

Given  $n_0 \in \mathbb{N}$  and  $k > (m+1)n_0$ , the same argument of the previous proof shows that one can decompose  $(v^k)$  as  $v^{k+1} = u^k + F(\varphi^{k-(m+1)n_0}, \dots, \varphi^k) + w^k$ , for a fixed function  $F$ ,  $\|u^k\|_\infty \leq C_1^{\left[\frac{k}{m+1}\right]} \|v^0\| \rightarrow 0$  and  $\|w^k\|_\infty \leq (m+1)C_2 \sum_{n=n_0}^\infty C_1^n$ . The latter is small if  $n_0$  is large. Hence, one could heuristically expect that the trailing strategies are often the ones dominating the convergence of the algorithm. In fact, in all the experiments carried out with a discretization satisfying (A0''),(A1),(A2), a dichotomous behaviour was observed: the algorithm either converged or at some point reached a cycle between a few payoffs. In the latter case, and restricting attention to instances in which one heuristically expects a solution to exist (more details in Section 3.3), it was possible to reduce the residual to the QVIs and the distance between the iterates by refining the grid.

The previous motivates the study of Algorithm 3.2.2.1 when the grid is sequentially refined, instead of fixed. Such an analysis however, would likely entail the need of a viscosity solutions framework as in [ABL18, BS91], which does not currently exist in the literature of nonzero-sum stochastic impulse games. Consequently, this analysis and the stronger convergence results that may come out of it are inevitably outside the scope of this thesis.

### 3.2.6 Discretization schemes

Let us conclude this section by showing how one can discretize the symmetric system of QVIs (3.1.1.1) to obtain (3.2.1.1) in a way that satisfies the assumptions present throughout the chapter. Recall that we work on a given symmetric grid  $G : x_{-N} = -x_N < \dots < x_{-1} = -x_1 < x_0 = 0 < x_1 < \dots < x_N$ .

---

<sup>12</sup>For any  $x \in \mathbb{R}$ ,  $[x]$  denotes its integer part.



Firstly, we want a discretization  $L$  of the operator  $\mathcal{A} - \rho Id$  such that  $-L$  is an SDD  $L_0$ -matrix as per (A1). A standard way to do this is to approximate the first (resp. second) order derivatives with forward and backward (resp. central) differences in such a way that we approximate the ODE  $\frac{1}{2}\sigma^2 V'' + \mu V' - \rho V + f = 0$  with an *upwind* (or *positive-coefficients*) scheme. More precisely, for each  $x = x_i \in G$  we approximate the first derivative with a forward (resp. backward) difference if its coefficient in the previous equation is nonnegative (resp. negative) in  $x_i$ ,

$$V'(x_i) \approx \frac{V(x_{i+1}) - V(x_i)}{x_{i+1} - x_i} \quad \text{if } \mu(x_i) \geq 0 \quad \text{and} \quad V'(x_i) \approx \frac{V(x_i) - V(x_{i+1})}{x_i - x_{i+1}} \quad \text{if } \mu(x_i) < 0$$

and the second derivative by

$$V''(x_i) \approx \frac{V(x_{i+1}) - V(x_i)}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})} - \frac{V(x_i) - V(x_{i-1})}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})}.$$

In the case of an equispaced grid with step size  $h$ , this reduces to

$$V'(x) \approx \frac{V(x + \text{sgn}(\mu(x))h) - V(x)}{\text{sgn}(\mu(x))h} \quad \text{and} \quad V''(x) \approx \frac{V(x+h) - 2V(x) + V(x-h)}{h^2}.^{13}$$

For the previous stencils to be defined in the extreme points of the grid, we consider two additional points  $x_{-N-1}, x_{N+1}$  and replace  $V(x_{-N-1}), V(x_{N+1})$  in the previous formulas by some values resulting from artificial boundary conditions. A common choice is to impose Neumann conditions to solve for  $V(x_{-N-1}), V(x_{N+1})$  using the first order differences from before. For example, in the equispaced grid case, given  $\text{LBC}, \text{RBC} \in \mathbb{R}$  we solve for  $V(x_{-N} - h)$  (resp.  $V(x_N + h)$ ) from the Neumann condition

$$\text{LBC} = V' \left( x_{-N} - h \mathbb{1}_{\{\mu(x_{-N}) \geq 0\}} \right) \quad (\text{resp. } \text{RBC} = V' \left( x_N + h \mathbb{1}_{\{\mu(x_N) < 0\}} \right)),$$

yielding  $V(x_{-N} - h) \approx V(x_{-N}) - \text{LBC}h$  (resp.  $V(x_N + h) \approx V(x_N) + \text{RBC}h$ ). The choice of  $\text{LBC}, \text{RBC}$  is problem-specific and intrinsically linked to that of  $x_N$ , although it does not affect the properties of the discrete operators. See more details in Section 3.3.

The described procedure leads to a discretization of the ODE as  $Lv + f = 0$ , with  $L$  satisfying the properties we wanted. (The *strict* diagonal dominance is a consequence of  $\rho > 0$ .) Note that the values of  $f$  at  $x_{-N}, x_N$  need to be modified to account for the boundary conditions.

*Remark 3.2.6.1.* One could increase the overall order of approximation by using central differences as much as possible for the first order derivatives, provided the scheme remains upwind (see [FL07, WF08] for more details). This is not done here in order to simplify the presentation.

We now approximate the impulse constraint sets  $\mathcal{Z}(x)$  ( $x \in \mathbb{R}$ ) by finite sets  $\emptyset \neq Z(x) \subseteq [0, +\infty)$  ( $x \in G$ ), such that  $Z(x) = \{0\}$  if  $x \geq 0$ , and define the impulse operators

$$B(\delta)v(x) = v\llbracket x + \delta(x) \rrbracket, \text{ for } v \in \mathbb{R}^G, \delta \in Z, x \in G,$$

<sup>13</sup>sgn denotes the sign function,  $\text{sgn}(x) = 1$  if  $x \geq 0$  and  $-1$  otherwise.

where  $v\llbracket y \rrbracket$  denotes linear interpolation of  $v$  on  $y$  using the closest nodes on the grid, and  $v\llbracket y \rrbracket = v(x_{\pm N})$  if  $\pm y > \pm x_{\pm N}$  (i.e., ‘no extrapolation’). This univocally defines the discrete loss and gain operators  $M$  and  $H$  as per (3.2.1.2), as well as the optimal impulse  $\delta^*$  according to (3.2.1.3). The set of discrete strategies  $\Phi$  is defined as in (3.2.2.1).

This general discretization scheme satisfies assumptions (A0)–(A2) and one can impose some regularity conditions on the sets  $\mathcal{Z}(x)$  and  $Z(x)$  such that the solutions of the discrete QVI problems (3.2.2.2) converge locally uniformly to the unique viscosity solution of the analytical impulse control problem, as the grid is refined.<sup>14</sup> See [Azi17, ABL18] for more details.

*Example 3.2.6.2.* In the case where  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$ , a natural and most simple choice for  $Z(x)$  is  $Z(x_i) = \{0, x_{i+1} - x_i, \dots, x_N - x_i\}$  for  $i < 0$ . In this case,  $B(\delta)v(x) = v(x + \delta(x))$  and  $Hv(x) = v(x - \delta^*(-x)) + g(x, -\delta^*(-x))$ . This choice, however, does not satisfy (A0’).

In order to preclude infinite simultaneous interventions it is enough to constrain the size of the impulses so that the symmetric point of the grid cannot be reached. That is,  $Z(x) \subseteq [0, -2x]$  for any  $x \in G_{<0}$ . In this case, the scheme satisfies the stronger conditions (A0’), (A1), (A2) (and in particular, (A0’)). Note that we can take  $m = N$  in (A0’), as each positive impulse will lead to a state which is at least one node closer to  $x_0 = 0$ , where no player intervenes. Practically, it makes sense to make this choice when one suspects (or wants to check whether) there is a symmetric NE with no ‘far-reaching impulses’, in the previous sense.

*Example 3.2.6.3.* If  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$ , the analogous of Example 3.2.6.2 is now  $Z(x_i) = \{0, x_{i+1} - x_i, \dots, x_{-i-1} - x_i\}$  for  $i < 0$ .

*Remark 3.2.6.4.* Consider Example 3.2.6.3 in the context of Remark 3.2.1.1. As in Proposition 3.2.5.6 and due to Theorem A.1.1.7, the less impulses needed between the two players to reach the common continuation region, the faster that the composition of the fixed-point operators of Algorithm 3.2.2.1 becomes contractive. Hence, one could intuitively expect that when close enough to the solution, the choice of the maximum arg-maximum in (3.2.1.3) improves the performance of Algorithm 3.2.2.1. This is another motivation for such choice.

### 3.3 Numerical results

This section presents numerical results obtained on a series of experiments. See the introduction of Chapter 2, Section 2.1.3 and Section 3.1 for the motivation and applications behind some of them. We do not assume additional constraints on the impulses in the analytical problem. All the results presented were obtained on equispaced grids with step size  $h > 0$  (to be specified) and with a discretization scheme as in Section 3.2.6 and Example 3.2.6.3. The extreme points of the grid are displayed on each graph.

For the games with linear costs and gains of the form  $c(x, \delta) = c^0 + c^1\delta$  and  $g(x, \delta) = g^0 + g^1\delta$ , with  $c^0, c^1, g^0, g^1$  constant, the artificial boundary conditions were taken as LBC

---

<sup>14</sup>Additional technical conditions include costs bounded away from zero and a comparison principle for the analytical QVI.

$= c^1$  and  $\text{RBC} = g^1$  for a sufficiently extensive grid. They result from the observation that on a hypothetical symmetric NE of the form  $\varphi^* = ((-\infty, \bar{x}], \delta^*(x) = y^* - x)$ , with  $\bar{x} < 0$ ,  $\bar{x} < y^* \in \mathbb{R}$ , the equilibrium payoff verifies  $V(x) = V(y^*) - c^0 - c^1(y^* - x)$  for  $x < \bar{x}$  and  $V(x) = V(-y^*) + g^0 + g^1(x + y^*)$  for  $x > -\bar{x}$ . For other examples, LBC, RBC and the grid extension were chosen by heuristic guesses and/or trial and error. However, in all the examples presented the error propagation from poorly chosen LBC, RBC was minimal.

The initial guess was set as  $v_0 = 0$  and its induced strategy in all cases. `SOLVEIMPULSECONTROL` was chosen as Subroutine 3.2.3.1 with  $\text{tol} = 10^{-15}$  and  $\lambda = 1$ .<sup>15</sup> Its convergence was exact however, in all the examples, and faster (in terms of time elapsed and number of operations) when it was compared with Howard's policy iteration (not reported). Instead of fixing a terminal tolerance  $\text{tol}$  beforehand, we display the highest accuracy that was attained in each case and the number of iterations needed for it.

Section 3.3.1 considers a fixed grid and games where the results point to the existence of a symmetric NE as per Corollary 3.1.1.9. Not having an analytical solution to compare with, results are assessed by means of the percentage difference between the iterates

$$\text{Diff} := \left\| (v^{k+1} - v^k) / \max\{|v^{k+1}|, \text{scale}\} \right\|_{\infty},$$

with  $\text{scale} = 1$  as in [AF16], and the maximum pointwise residual to the system of QVIs (3.2.1.1), defined for  $v \in \mathbb{R}^G$  by setting  $I = \{Lv + f \leq Mv - v\} \cap G_{<0}$ ,  $C = I^c$  and

$$\text{maxResQVIs}(v) := \left\| \max\{Lv + f, Mv - v\} \mathbb{1}_{-C} + (Hv - v) \mathbb{1}_{-I} \right\|_{\infty}.$$

Section 3.3.2 considers the only symmetric games in the literature with (semi-) analytical solution: the central bank linear game of Section 2.1.3 and the cash management game [BCG19], and computes the errors made by discrete approximations, showing in particular the effect of refining the grid. Not considered here is the strategic pollution control game [FK19], due to its inherent non-symmetric nature. Section 3.3.3 comments on results obtained for games without NEs. Finally, Section 3.3.4 shows results that go beyond the scope of the currently available theory for impulse games.

### 3.3.1 Convergence to discrete solution on a fixed grid

Throughout this section the grid step size is fixed as  $h = 0.01$ , unless otherwise stated (although results were corroborated by further refinements). Each figure specifies the structure,  $\mathcal{G} = (\mu, \sigma, \rho, f, c, g)$ , of the symmetric game solved and shows the numerical solutions at the terminal iteration for the equilibrium payoff,  $v^k$ , and NE. Graphs plot payoff versus state of the process. The intervention region is displayed in red over the graph of the payoff for presentation purposes.

As a general rule, we focus on games with higher costs  $c$  than gains  $g$ , as the opposite typically leads to players attempting to apply infinite simultaneous impulses [ABC<sup>+</sup>19] (i.e., inducing a gain from the opponent's intervention is 'cheap') leading to degenerate games. The following games resulted in exact convergence in finite iterations, which guarantees a

<sup>15</sup>For very fine grids, one should increase the value of  $\widetilde{\text{tol}}$  to avoid stagnation as per Remark 3.2.2.4, and the same being true for Subroutine 3.2.3.1.

solution of (3.2.1.1) was reached (Corollary 3.2.5.3), although very small acceptable errors were reached much sooner.

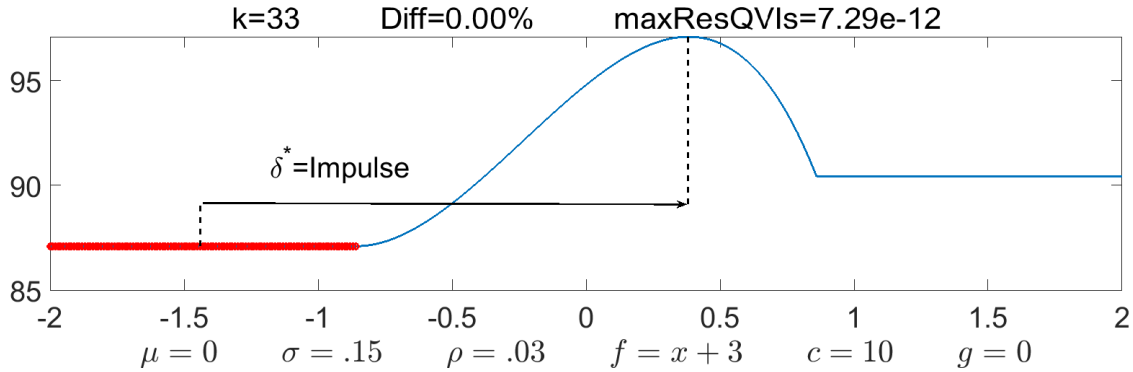


Figure 3.1

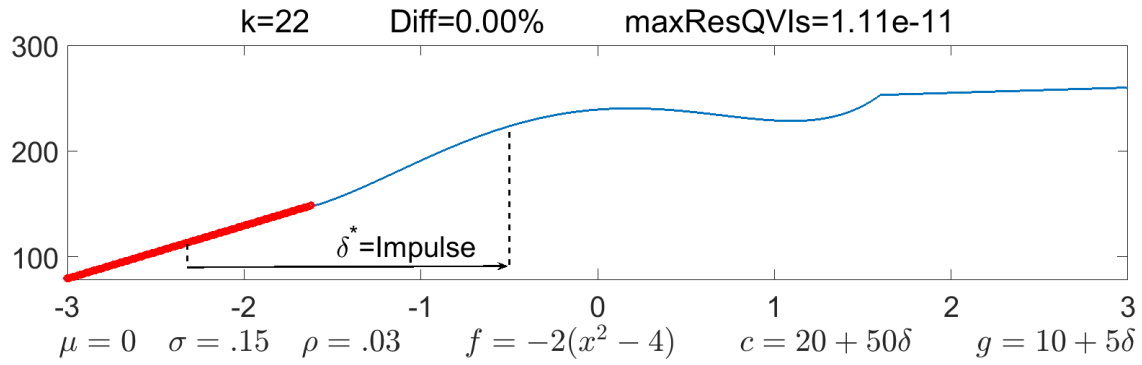


Figure 3.2

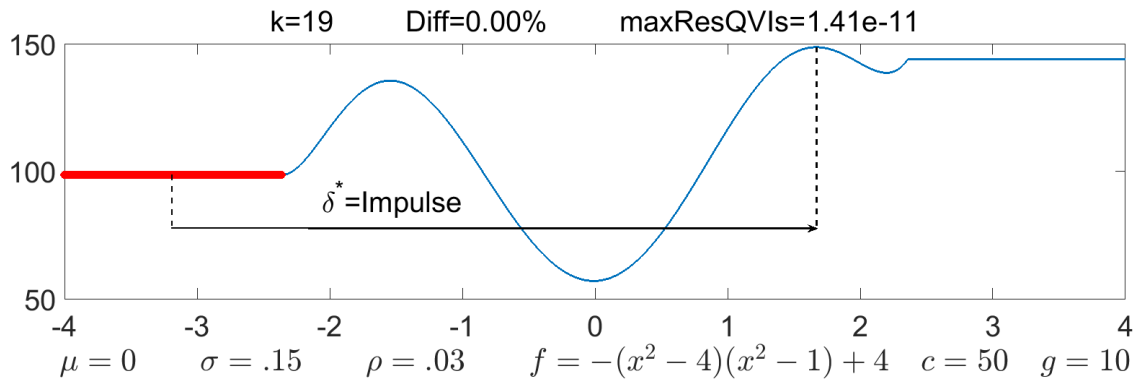


Figure 3.3

The following is an example in which the accuracy stagnates. At that point, the iterates start going back and forth between a few values. Although we cannot guarantee that we are close to a solution of (3.2.1.1), the results seem quite convincing, with both Diff and

maxResQVIs reasonably low. In fact, simply halving the step size to  $h = 0.005$  produces a substantial improvement of Diff=9.16e-11% and maxResQVIs=9.19e-11 in  $k = 33$  iterations.

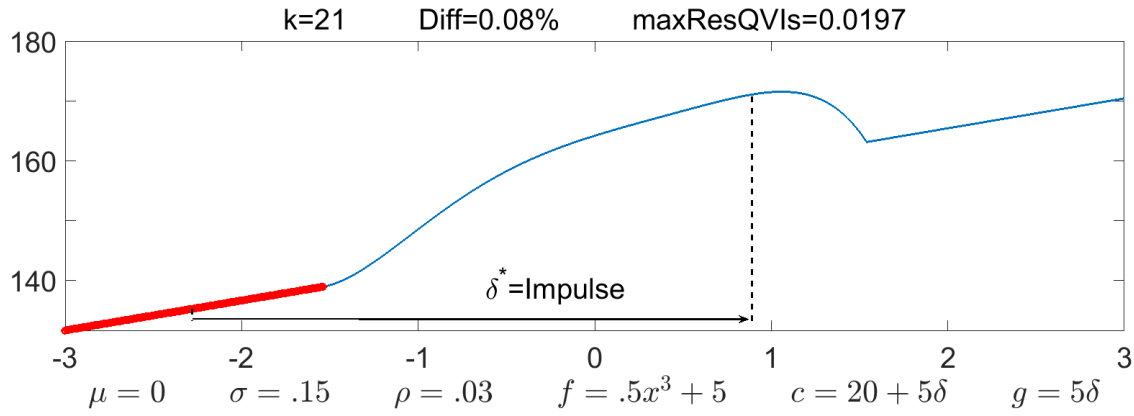


Figure 3.4

The previous games have a state variable evolving as a scaled Brownian motion. We now move on to a mean reverting OU process with zero long term mean. (Recall that any other value can be handled simply by shifting the game.) In general, the experiments with these dynamics converged exactly in a very small number of iterations.

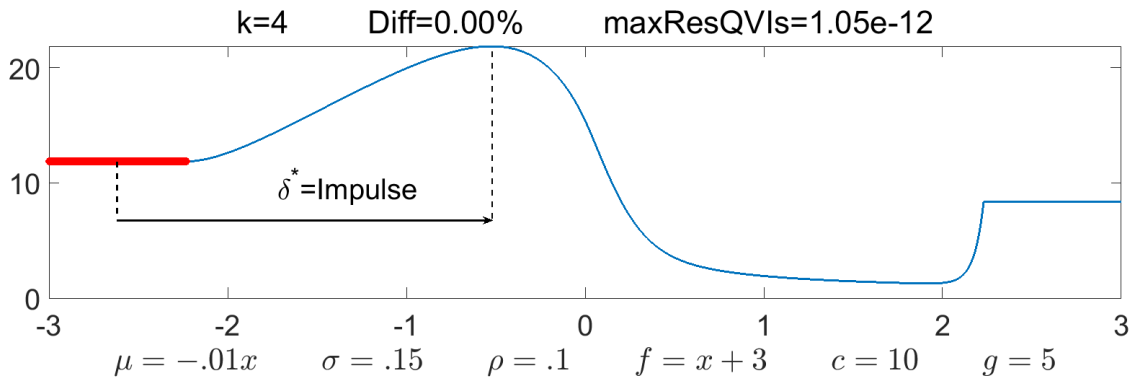


Figure 3.5

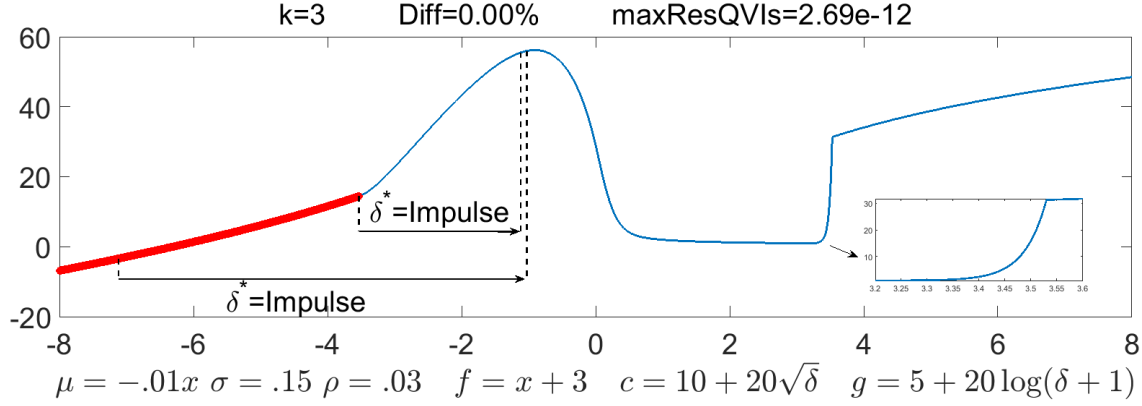


Figure 3.6

Note how all the games treated in this section exhibit a typical feature known to hold for simpler symmetric games (see Chapter 2 and [ABC<sup>+</sup>19, BCG19]): the equilibrium payoff of the player is only  $C^1$  at the border of the intervention region  $\partial\mathcal{I}^* = \{AV - \rho V + f = \mathcal{M}V - V\} \cap \mathbb{R}_{<0}$ , and only continuous at the border of the opponent's intervention region  $-\partial\mathcal{I}^*$ . In floating point arithmetic, the former makes the discrete approximation of  $\partial\mathcal{I}^*$  particularly elusive, while the latter can lead to high errors when close to  $-\partial\mathcal{I}^*$ . As a consequence, Subroutine 3.2.3.1 (or any equivalent) will often misplace a few grid nodes between intervention and continuation regions, which will in turn make the residual resQVIs ‘spike’ on the opponent's side. Thus, a large value of maxResQVIs can at times be misleading, and further inspection of the pointwise residuals is advisable.

As a matter of fact, halving the step size to  $h = 0.005$  in the last example results in exact convergence with terminal maxResQVIs = 2510, but the residuals on all grid nodes other than the ‘border’ of the opponent's intervention region and a contiguous one are less than  $1.43\text{e-}11$ . This is an extreme example propitiated by the almost vertical shape of the solution close to such border. Thus, while it is useful in practice to consider a stopping criteria for Algorithm 3.2.2.1 based on maxResQVIs, this phenomenon needs to be minded.

The last example also shows how impulses at a NE can lead to different endpoints depending on the state of the process. This is often the case when costs are nonlinear. In fact:

**Lemma 3.3.1.1.** *Let  $(\mathcal{I}^*, \delta^*)$  and  $V$  be a symmetric NE and equilibrium payoff as in Corollary 3.1.1.9. Assume that  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$  and  $c = c(x, \delta) \in C^2(\mathbb{R} \times (0, +\infty))$ , and consider the re-parametrization  $\bar{c} = \bar{c}(x, y) := c(x, y - x)$ . Suppose that  $y^* := x + \delta^*(x)$  is constant for all  $x \in (\mathcal{I}^*)^\circ$ .<sup>16</sup> Then  $\bar{c}_{xy}(x, y^*) \equiv 0$  on  $(\mathcal{I}^*)^\circ$ .*

The result is immediate since  $\mathcal{M}V(x) = \sup_{y>x} \{V(y) - \bar{c}(x, y)\} = V(y^*) - \bar{c}(x, y^*)$  on  $(\mathcal{I}^*)^\circ$  and  $V \in C^1(-\mathcal{C}^*)$ . ( $y^* \notin -\mathcal{I}^*$  or there would be infinite simultaneous impulses.) While the sufficient condition  $\bar{c}_{xy}(x, y^*) \equiv 0$  for some  $y^*$  is verified for linear costs, it is not in general and certainly not for  $c(x, \delta) = 10 + 20\sqrt{\delta}$  as in the last example.

<sup>16</sup> $A^\circ$  denotes the interior of the set  $A \subseteq \mathbb{R}$ .

### 3.3.2 Convergence to analytical solution with refining grids

A convergence analysis from discrete to analytical solution with refining grids is outside the scope of this thesis, and seems to be far too challenging when a viscosity solutions framework is yet to be developed. Instead, we present here a numerical validation using the solutions of the linear and cash management games. We focus first and foremost in the former, as it has an almost fully analytical solution, with only one parameter to be found numerically as opposed to four for the latter. The structure of the game is defined with parameter values used in [ABC<sup>+</sup>19] (also in Chapter 2). To minimize rounding errors from floating point arithmetic, we proceed as in [AF16] considering grids made up entirely of machine numbers. The results are displayed in Table 3.1.

For each step size, Algorithm 3.2.2.1 either converged or was terminated upon stagnation. Regardless, we can see the errors made when approximating the analytical solution are quite satisfactory in all cases, and overall decrease as  $h \rightarrow 0$ . Moreover, we see once again how the ‘spiking’ of the residual can be misleading (the highest value was always at the ‘border’ of the opponent’s intervention region).

h	%error	Its.	maxResQVIs	Table 3.1: Convergence to analytical solution when refining equispaced symmetric grid with step size $h$ and endpoint $x_N = 4$ . Game: $\mu = 0$ , $\sigma = .15$ , $\rho = .02$ , $f = x + 3$ , $c = 100 + 15\delta$ , $g = 15\delta$ . %error := $\ (v - V)/V\ _\infty$ , with $V$ exact solution and $v$ discrete approximation after Its iterations. $\ \cdot\ _\infty$ is computed over the grid.
1	6.67%	17	8.88e-16	
1/2	8.33%	13	5.33e-15	
1/4	0.23%	4	13.2	
1/8	0.21%	8	15.1	
1/16	0.16%	8	30	
1/32	0.07%	21	21.2	
1/64	0.0043%	37	0.343	

An exact symmetric NE for this game (up to five significant figures) is given by the intervention region  $\mathcal{I}^* = (-\infty, -2.8238]$  and impulse function  $\delta^*(x) = 1.5243 - x$ , while the approximation given by Algorithm 3.2.2.1 with  $h = 1/64$  is  $((-\infty, -2.8125], 1.5313 - x)$ , with absolute errors on the parameters of no more than the step size.

The cash management game [BCG19] with unidirectional impulses can be embedded in our framework by reducing its dimension with the change of variables  $x = x_1 - x_2$ , changing minimization by maximization and relabelling the players. With the parameter values of [BCG19, Fig.1b], it translates into the game:  $\mu = 0$ ,  $\sigma = 1$ ,  $\rho = .5$ ,  $f = -|x|$ ,  $c = 3 + \delta$ ,  $g = -1$ . The authors found numerically a symmetric NE approximately equal to  $((-\infty, -5.658], 0.686 - x)$ , while Algorithm 3.2.2.1 with  $x_N = 8$  and  $h = 1/64$  gives  $((-\infty, -5.6563], 0.6875 - x)$ . The absolute difference on the parameters is once again below the grid step size.

### 3.3.3 Games without Nash equilibria

It is natural to ask how Algorithm 3.2.2.1 behaves on games without symmetric NEs, and whether anything can be inferred from the results. For the linear game, two cases without NEs (symmetric or not) are addressed in [ABC<sup>+</sup>19]: ‘no fixed cost’ and ‘gain greater than

cost'. Both of them yield degenerate 'equilibria' where the players perform infinite simultaneous interventions. When tested for several parameters, Algorithm 3.2.2.1 converged in finitely many iterations (although rather slowly) and yielded the exact same type of 'equilibria'.<sup>17</sup> For a fine enough grid, the previous can be identified heuristically by some node in the intervention region that would be shifted to its symmetric one (*infinite alternated interventions*), or to its immediate successor over and over again, until reaching the continuation region (*infinite one-sided interventions*).<sup>18</sup>

In the first case, we recovered the limit 'equilibrium payoffs' of [ABC<sup>+</sup>19, Props.4.10,4.11]. Intuitively, the players in this game take advantage of free interventions, whether by no cost or perfect compensation, in order to shift the process as desired. Note that when  $c = g$ , the impulses that maximize the net payoff are not unique.

In the second case, grid refinements showed the discrete payoffs to diverge towards infinity at every point. This is again consistent with the theory: each player forces the other one to act, producing a positive instantaneous profit. Iterating this procedure infinitely often leads to infinite payoffs for every state.

Tested games in which Algorithm 3.2.2.1 failed to converge (and not due to stagnation nor a poor choice of the grid extension) were characterized by iterates reaching a cycle, typically with high values of Diff and maxResQVIs regardless of the grid step size. In many cases, the cycles would visit at least one payoff inducing infinite simultaneous impulses. While this might, potentially and heuristically, be indicative of the game admitting no symmetric NE, there is not much more than can be said at this stage.

### 3.3.4 Beyond the Verification Theorem

We now present two games in which the solution of the discrete QVIs system (3.2.1.1) found with Algorithm 3.2.2.1 (by exact convergence) does not comply with the continuity and smoothness assumptions of Corollary 3.1.1.9. However, the results are sensible enough to heuristically argue they may correspond to NEs beyond the scope of the Verification Theorem. In both cases  $h = 0.01$ . Finer grids yielded the same qualitative results.

The first game considers costs convex in the impulse. When far enough from her continuation region, it is cheaper for the player to apply several (finitely many) simultaneous impulses instead of one, to reach the state she wishes to (cf. Remark 2.1.2.4). In this game, the optimal impulse  $\delta^*$  becomes discontinuous, and its discontinuity points are those in the intervention region where the equilibrium payoff is non-differentiable. These, in turn, translate into discontinuities on the opponent's intervention region (one smoothness degree less, as with the border of the regions).

---

<sup>17</sup>There were also cases of stagnation, improved by refining the grid as in Section 3.3.1.

<sup>18</sup>More precisely, due to our choice of  $Z(x)$  a node can be shifted at most to that immediately preceding its symmetric one.



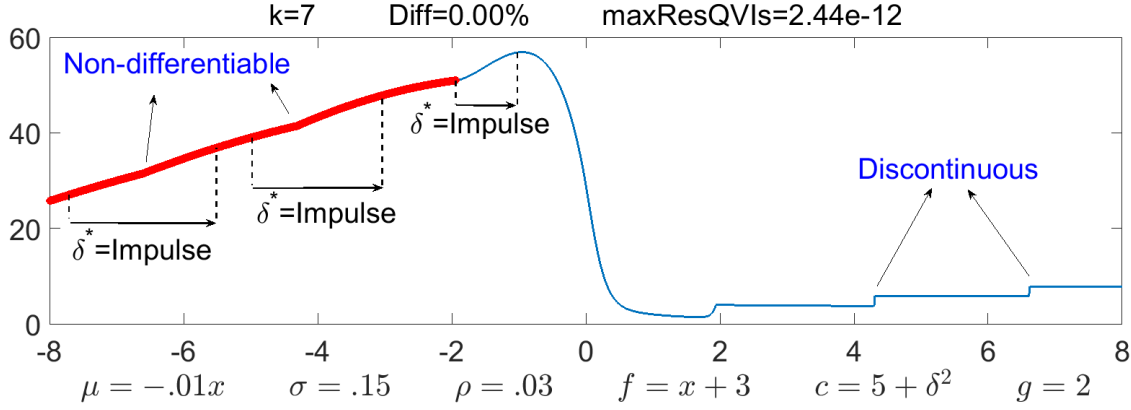


Figure 3.7

The second game considers linear gains, quadratic running payoffs and costs concave on the impulse. The latter makes the player shift the process towards her continuation region (cf. Remark 2.1.2.4). However, when far enough from the border, instead of shifting the process directly to her ‘preferred area’, the player chooses to pay a bit more to force her opponent’s intervention, inducing a gain and letting the latter pay for the final move. Once again, this causes  $\delta^*$  to be discontinuous and leads to a non-differentiable (resp. discontinuous) point for the equilibrium payoff in the intervention (resp. opponent’s intervention) region.

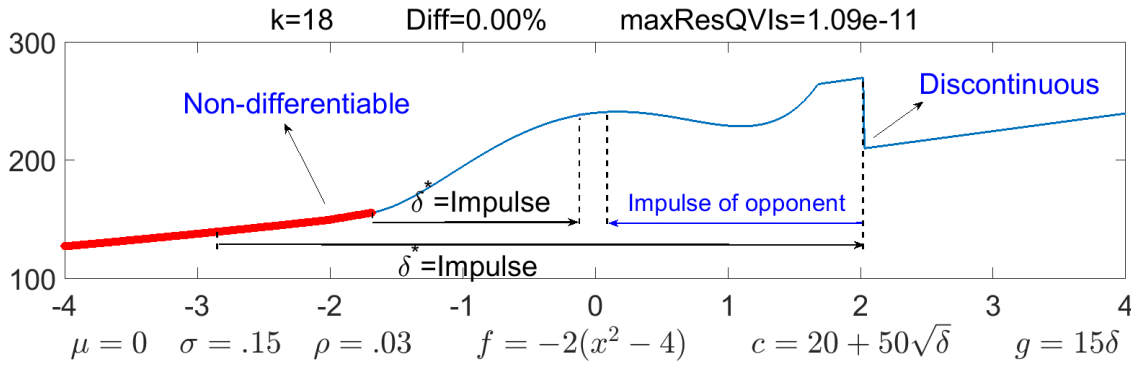


Figure 3.8

Under the previous reasoning, one could intuitively guess that setting  $g = 0$  in this game should remove the main incentive the player has to force her opponent to act. This is in fact the case, as shown below. As a result,  $\delta^*$  becomes continuous and the equilibrium payoff falls back into the domain of the Verification Theorem.

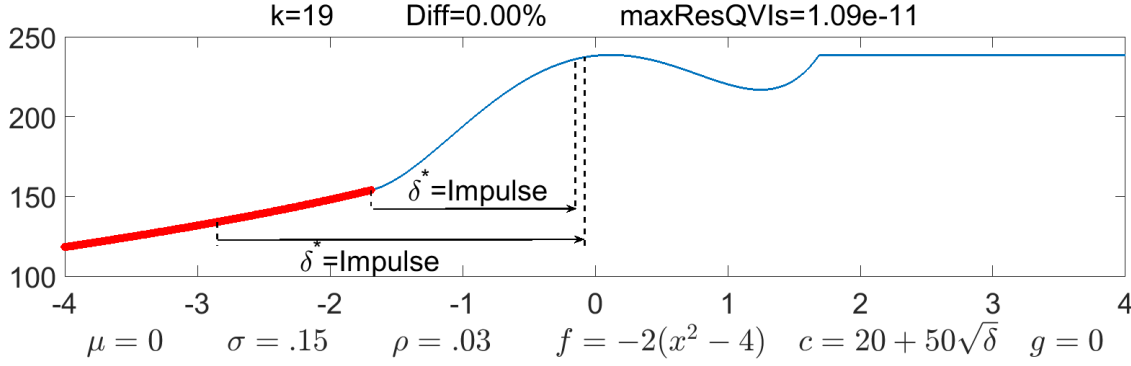


Figure 3.9

We remark that, should the previous solutions correspond indeed to NEs, then the alternative semi-analytical approach of [FK19] could not have produced them either, as that method can only yield continuous equilibrium payoffs.

### 3.4 Concluding remarks

This chapter presents a fixed-point policy-iteration-type algorithm to solve systems of quasi-variational inequalities resulting from a Verification Theorem for symmetric nonzero-sum stochastic impulse games. In this context, the method substantially improves the only algorithm available in the literature (Chapter 2) while providing the convergence analysis that we were missing. Graph-theoretic assumptions relating to weakly chained diagonally dominant matrices, which naturally parallel the admissibility of the players strategies, allow to prove properties of contractiveness, boundedness of iterates and convergence to solutions. A result of theoretical interest giving sufficient conditions for convergence is also proved. Additionally, a provably convergent impulse control solver is provided for added efficiency over classical policy iteration.

Equilibrium payoffs and Nash equilibria of games too challenging for the available analytical approaches are computed with high precision on a discrete setting. Numerical validation with analytical solutions is performed when possible, with reassuring results, but it is noted that grid refinements may be needed at times to overcome stagnation. Thus, formalising the approximating properties of the discrete solutions as well as deriving stronger convergence results for the algorithm may need a viscosity solutions framework currently missing in the theory. This is further substantiated by the irregularity of the solutions, particularly those found which escape the available theoretical results. This motivates further research while providing a tool that can effectively be used to gain insight into these very challenging problems.

# Appendix A

## Matrix and graph-theoretic definitions and results

For the reader's convenience, this appendix summarizes some important algebraic and graph-theoretic definitions and results used primarily throughout Chapter 3. More details can be found in the references given below. Henceforth,  $A \in \mathbb{R}^{N \times N}$  is a real matrix,  $Id \in \mathbb{R}^{N \times N}$  is the identity,  $\rho(\cdot)$  denotes the spectral radius and  $\mathbb{R}^N, \mathbb{R}^{N \times N}$  are equipped with the elementwise order. We talk about rows and 'states' interchangeably.

**Definitions A.1.1.1.** (D1)  $A$  is a *Z-matrix* if it has nonpositive off-diagonal elements.

(D2)  $A$  is an *L-matrix* (resp. *L<sub>0</sub>-matrix*) if it is a Z-matrix with positive (resp. nonnegative) diagonal elements.

(D3)  $A$  is an *M-matrix* if  $A = sId - B$  for some matrix  $B \geq 0$  and scalar  $s \geq \rho(B)$ .

(D4)  $A$  is *monotone* if it is nonsingular and  $A^{-1} \geq 0$ . Equivalently,  $A$  is monotone if  $Ax \geq 0$  implies  $x \geq 0$  for any  $x \in \mathbb{R}^N$ .

(D5) The  $i$ -th row of  $A$  is *weakly diagonally dominant (WDD)* (resp. *strictly diagonally dominant or SDD*) if  $|A_{ii}| \geq \sum_{j \neq i} |A_{ij}|$  (resp.  $>$ ).

(D6)  $A$  is *WDD* (resp. *SDD*) if every row of  $A$  is WDD (resp. SDD).

(D7) The *directed graph of A* is the pair  $\text{graph}A := (V, E)$ , where  $V := \{1, \dots, N\}$  is the set of *vertexes* and  $E \subseteq V \times V$  is the set of *edges*, such that  $(i, j) \in E$  iff  $A_{ij} \neq 0$ .

(D8) A *walk*  $p$  in  $\text{graph}A = (V, E)$  from vertex  $i$  to vertex  $j$  is a nonempty finite sequence  $(i, i_1), (i_1, i_2), \dots, (i_{k-1}, j) \subseteq E$ , which we denote by  $i \rightarrow i_1 \rightarrow \dots \rightarrow i_{k-1} \rightarrow j$ .  $|p| := k$  is called the *length* of the walk  $p$ .

(D9)  $A$  is *weakly chained diagonally dominant (WCDD)* if it is WDD and for each WDD row of  $A$  there is a walk in  $\text{graph}A$  to an SDD row (identifying vertexes and rows).

(D10)  $A$  is (right) *substochastic or sub-Markov* (resp. *stochastic or Markov*) if  $A \geq 0$  and each row sums at most one (resp. exactly one). Equivalently,  $A$  is substochastic if  $A \geq 0$  and  $\|A\|_\infty \leq 1$ . (Recall that  $\|A\|_\infty$  is the maximum row-sum of absolute values.)

(D11) If  $A$  is a WDD (resp. substochastic) matrix, its set of ‘non-trouble states’ (or rows) is  $J[A] := \{i : \text{the } i\text{-th row of } A \text{ is SDD}\}$  (resp.  $\hat{J}[A] := \{i : \sum_j A_{ij} < 1\}$ ). For each  $i$ , we write  $P_i[A] := \{\text{walks in graph } A \text{ from } i \text{ to some } j \in J[A]\}$  (resp. analogously for  $\hat{P}_i[A]$ ). The *index of connectivity* (resp. *index of contraction*) of  $A$  [Azi19a] is

$$\text{con}A := \left( \sup_{i \notin J[A]} \left\{ \inf_{p \in P_i[A]} |p| \right\} \right)^+ \quad (\text{resp. analogously for } \widehat{\text{con}}A).^1$$

It is clear that  $\text{SDD} \implies \text{WCDD} \implies \text{WDD}$ , and by definition,  $\text{L-matrix} \implies \text{L}_0\text{-matrix} \implies \text{Z-matrix}$ . Also by definition, if  $A$  is WDD then:  $A$  is WCDD  $\iff \text{con}A < +\infty$ .

**Proposition A.1.1.2.** (e.g., [AF16, Lem.3.2]) *Any WCDD matrix is nonsingular.*

**Proposition A.1.1.3.** (e.g., [Azi19a, Prop.2.15 and 2.17])

*Nonsingular M-matrix  $\iff$  monotone L-matrix  $\iff$  monotone Z-matrix.*

**Theorem A.1.1.4.** (e.g., [Azi19a, Thm.2.24]) *WCDD  $\text{L}_0$ -matrix  $\iff$  WDD nonsingular M-matrix.*<sup>2</sup>

**Proposition A.1.1.5.** (see proof of [Azi19a, Lem.2.22])  *$A$  is substochastic if and only if  $\text{Id} - A$  is a WDD  $\text{L}_0$ -matrix and  $A$  has non-negative diagonal elements. In such case,  $\hat{J}[A] = J[\text{Id} - A]$ , they have the same directed graphs (except possibly for self-loops  $i \rightarrow i$ ) and  $\widehat{\text{con}}A = \text{con}[\text{Id} - A]$ .*

For the following theorem, recall the characterization of the spectral radius  $\rho(A) = \inf\{\|A\| : \|\cdot\| \text{ is a matrix norm}\}$  and Gelfand’s formula  $\rho(A) = \lim_{n \rightarrow +\infty} \|A^n\|^{1/n}$ , for any matrix norm  $\|\cdot\|$ . Note also that if  $A$  is substochastic, then  $A^n$  is also substochastic for any  $n \in \mathbb{N}_0$ ,  $\|A^n\|_\infty \leq 1$  and  $\rho(A) \leq 1$ .

**Theorem A.1.1.6.** ([Azi19a, Thm.2.5 and Cor.2.6]) *Suppose  $A$  is substochastic. Then*

$$\widehat{\text{con}}A = \inf\{n \in \mathbb{N}_0 : \|A^{n+1}\|_\infty < 1\}.$$

*In particular,  $\widehat{\text{con}}A < +\infty$  if and only if  $\rho(A) < 1$ .*

The indexes of contraction and connectivity can be generalized in a natural way to sequences  $(A_k) \subseteq \mathbb{R}^{N \times N}$  by considering walks  $i_1 \rightarrow i_2 \rightarrow \dots$  such that  $i_k \rightarrow i_{k+1}$  is an edge in  $\text{graph}A_k$  (see [Azi19a, App.B] for more details). Theorem A.1.1.6 extends in the following way:

**Theorem A.1.1.7.** ([Azi19a, Thm.B.2]) *Suppose  $(A_k)$  are substochastic matrices and consider the sequence of products  $(B_k)$ , where  $B_k := A_1 \dots A_k$ . Then,*

$$\widehat{\text{con}}[(A_k)_k] = \inf\{k \in \mathbb{N}_0 : \|B_{k+1}\|_\infty < 1\}.$$

<sup>1</sup> $(\cdot)^+$  denotes positive part,  $\inf \emptyset = +\infty$  and  $\sup \emptyset = -\infty$ . The index is the least length that needs to be walked on  $\text{graph}A$  to reach the non-trouble states when starting from an arbitrary trouble one.

<sup>2</sup> [Azi19a, Thm.2.24] is formulated in terms of L-matrices instead. However, it is trivial to see that:  $\text{WCDD } \text{L}_0\text{-matrix} \iff \text{WCDD L-matrix}$ .

**Proof of Lemma 3.2.4.4.** For briefness, we omit the dependence on  $\varphi, \bar{\varphi}$  from the notation.

$\mathbb{A}^{-1}\mathbb{B} \geq 0$  since  $\mathbb{B} \geq 0$  and  $\mathbb{A}$  is monotone. To see that its rows sum up to one, let  $\mathbf{1} \in \mathbb{R}^G$  be the vector of ones. It is easy to check that under (A1)–(A2),  $\mathbb{A}\mathbb{A}^{-1}\mathbb{B}\mathbf{1} = \mathbb{B}\mathbf{1} \leq \mathbb{A}\mathbf{1}$ , which implies  $\mathbb{A}^{-1}\mathbb{B}\mathbf{1} \leq \mathbf{1}$ . The WDD  $L_0$ -property of  $\mathbb{A} - \mathbb{B}$  is due to (A1).

Note that by Proposition A.1.1.5, we can consider  $\mathbb{A}^{-1}\mathbb{B}$  and  $Id - \mathbb{A}^{-1}\mathbb{B}$  interchangeably. Let us show that  $J[(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})] \subseteq \hat{J}[\mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})]$ . By monotonicity and  $L_0$ -property,  $\mathbb{A}^{-1}$  must have strictly positive diagonal elements. Indeed, if there was some index  $i$  such that  $\mathbb{A}_{ii}^{-1} = 0$ , then  $1 = [\mathbb{A}\mathbb{A}^{-1}]_{ii} = \sum_{j \neq i} \mathbb{A}_{ij}\mathbb{A}_{ji}^{-1} \leq 0$ . Consider now some  $x_i \in \hat{J}[\mathbb{A}^{-1}\mathbb{B}]^c$ , i.e.,  $\sum_j [\mathbb{A}^{-1}\mathbb{B}]_{ij} = 1$ . We want to see that  $\sum_j [\mathbb{A} - \mathbb{B}]_{ij} = 0$ . We have  $0 = \sum_j [Id - \mathbb{A}^{-1}\mathbb{B}]_{ij} = \sum_j [\mathbb{A}^{-1}(\mathbb{A} - \mathbb{B})]_{ij} = \sum_k \mathbb{A}_{ik}^{-1} \sum_j [\mathbb{A} - \mathbb{B}]_{kj}$ . Since  $\mathbb{A}_{ik}^{-1}$  and  $\sum_j [\mathbb{A} - \mathbb{B}]_{kj}$  are non-negative for all  $k$ , one of the two must be zero for each  $k$ . But  $\mathbb{A}_{ii}^{-1} > 0$ , giving what we wanted.

The final conclusion follows from the previous properties and the fact that  $\mathbb{A}(Id - \mathbb{A}^{-1}\mathbb{B}) = \mathbb{A} - \mathbb{B}$ .  $\square$

# References

- [AB06] C.D. Aliprantis and K.C. Border, *Infinite dimensional analysis: a hitchhiker's guide*, 3rd ed., Springer, 2006.
- [ABC<sup>+</sup>19] R. Aïd, M. Basei, G. Callegaro, L. Campi, and T. Vargiolu, *Nonzero-sum stochastic differential games with impulse controls: a verification theorem with applications*, Mathematics of Operations Research (2019).
- [ABL18] P. Azimzadeh, E. Bayraktar, and G. Labahn, *Convergence of implicit schemes for Hamilton–Jacobi–Bellman quasi-variational inequalities*, SIAM Journal on Control and Optimization **56** (2018), no. 6, 3994–4016.
- [ABM<sup>+</sup>19] R. Aïd, F. Bernal, M. Mnif, D. Zabaljauregui, and J.P. Zubelli, *A policy iteration algorithm for nonzero-sum stochastic impulse games*, ESAIM: Proceedings and Surveys **65** (2019), 27–45.
- [AF16] P. Azimzadeh and P.A. Forsyth, *Weakly chained matrices, policy iteration, and impulse control*, SIAM Journal on Numerical Analysis **54** (2016), no. 3, 1341–1364.
- [AS08] M. Avellaneda and S. Stoikov, *High-frequency trading in a limit order book*, Quantitative Finance **8** (2008), no. 3, 217–224.
- [Azi17] P. Azimzadeh, *Impulse control in finance: numerical methods and viscosity solutions*, Ph.D. thesis, University of Waterloo, 2017.
- [Azi19a] ———, *A fast and stable test to check if a weakly diagonally dominant matrix is a nonsingular  $m$ -matrix*, Mathematics of Computation **88** (2019), no. 316, 783–800.
- [Azi19b] ———, *A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions*, Applied Mathematics & Optimization **79** (2019), no. 2, 483–514.
- [Bas19] M. Basei, *Optimal price management in retail energy markets: an impulse control problem with asymptotic estimates*, Mathematical Methods of Operations Research **89** (2019), no. 3, 355–383.
- [BCG19] M. Basei, H. Cao, and X. Guo, *Nonzero-sum stochastic games with impulse controls*, arXiv preprint arXiv:1901.08085 (2019).

- [BL09] E. Bayraktar and M. Ludkovski, *Optimal trade execution in illiquid markets*, arXiv preprint arXiv:0902.2516v1 (2009).
- [BL14] ———, *Liquidation in limit order books with controlled intensity*, Mathematical Finance **24** (2014), no. 4, 627–650.
- [BMZ09] O. Bokanowski, S. Maroso, and H. Zidani, *Some convergence results for Howard’s algorithm*, SIAM Journal on Numerical Analysis **47** (2009), no. 4, 3001–3026.
- [Bou07] B. Bouchard, *Introduction to stochastic control of mixed diffusion processes, viscosity solutions and applications in finance and insurance*, Lecture Notes Preprint (2007).
- [BR09] N. Bäuerle and U. Rieder, *MDP algorithms for portfolio optimization problems in pure jump markets*, Finance and Stochastics **13** (2009), no. 4, 591–611.
- [BR10] ———, *Optimal control of piecewise deterministic Markov processes with finite time horizon*, Modern trends in controlled stochastic processes: theory and applications **123** (2010), 143.
- [BR11] ———, *Markov decision processes with applications to finance*, Springer Science & Business Media, 2011.
- [Bré81] P. Brémaud, *Point processes and queues: martingale dynamics*, vol. 50, Springer, 1981.
- [BS91] G. Barles and P.E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic analysis **4** (1991), no. 3, 271–283.
- [CCTZ06] A. Cadenillas, T. Choulli, M. Taksar, and L. Zhang, *Classical and impulse stochastic control for the optimization of the dividend and risk policies of an insurance firm*, Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics **16** (2006), no. 1, 181–202.
- [CDJ17] Á. Cartea, R. Donnelly, and S. Jaimungal, *Algorithmic trading with model uncertainty*, SIAM Journal on Financial Mathematics **8** (2017), no. 1, 635–671.
- [CEFS16a] K. Colaneri, Z. Eksi, R. Frey, and M. Szölgyenyi, *Optimal liquidation under partial information with price impact*, arXiv preprint arXiv:1606.05079v5 (2016).
- [CEFS16b] ———, *Shall i sell or shall i wait? Optimal liquidation under partial information with price impact*, arXiv preprint arXiv:1606.05079v3 (2016).
- [CJ13] Á. Cartea and S. Jaimungal, *Modelling asset prices for algorithmic and high-frequency trading*, Applied Mathematical Finance **20** (2013), no. 6, 512–547.
- [CJ15] ———, *Risk metrics and fine tuning of high-frequency trading strategies*, Mathematical Finance **25** (2015), no. 3, 576–611.

- [CJP15] Á. Cartea, S. Jaimungal, and J. Penalva, *Algorithmic and high-frequency trading*, Cambridge University Press, 2015.
- [CJR14] Á. Cartea, S. Jaimungal, and J. Ricci, *Buy low, sell high: A high frequency trading perspective*, SIAM Journal on Financial Mathematics **5** (2014), no. 1, 415–444.
- [CL55] E.A. Coddington and N. Levinson, *Theory of ordinary differential equations*, Tata McGraw-Hill Education, 1955.
- [Cli07] S. Clift, *Linear and non-linear monotone methods for valuing financial options under two-factor, jump-diffusion models*, Ph.D. thesis, University of Waterloo, 2007.
- [CMS07] J.-P. Chancelier, M. Messaoud, and A. Sulem, *A policy iteration algorithm for fixed point problems with nonexpansive operators*, Mathematical Methods of Operations Research **65** (2007), no. 2, 239–259.
- [CØS02] J.-P. Chancelier, B. Øksendal, and A. Sulem, *Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control*, Tr. Mat. Inst. Steklova **237** (2002), no. 0, 149–172.
- [Cos13] A. Cosso, *Stochastic differential games involving impulse controls and double-obstacle quasi-variational inequalities*, SIAM Journal on Control and Optimization **51** (2013), no. 3, 2102–2131.
- [CWW13] D. Chang, H. Wang, and Z. Wu, *Maximum principle for non-zero sum differential games of bsdes involving impulse controls*, Proceedings of the 32nd Chinese Control Conference, IEEE, 2013, pp. 1564–1569.
- [CZ99] A Cadenillas and F Zapatero, *Optimal central bank intervention in the foreign exchange market*, Journal of Economic theory **87** (1999), no. 1, 218–242.
- [CZ19] L. Campi and D. Zabaljauregui, *Optimal market making under partial information with general intensities*, arXiv preprint arXiv:1902.01157 (2019).
- [DAFM18] T. De Angelis, G. Ferrari, and J. Moriarty, *Nash equilibria of threshold type for two-player nonzero-sum games of stopping*, The Annals of Applied Probability **28** (2018), no. 1, 112–147.
- [Dav84] M.H.A. Davis, *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, Journal of the Royal Statistical Society. Series B (Methodological) (1984), 353–388.
- [Dav93] ———, *Markov models and optimization*, Chapman and Hall/CRC, 1993.
- [DF99] M.H.A. Davis and M. Farid, *Piecewise-deterministic processes and viscosity solutions*, Stochastic analysis, control, optimization and applications, Springer, 1999, pp. 249–268.



- [DF14] D.-M. Dang and P.A. Forsyth, *Continuous time mean-variance optimal portfolio allocation under jump diffusion: An numerical impulse control approach*, Numerical Methods for Partial Differential Equations **30** (2014), no. 2, 664–698.
- [DRR13] S. Delattre, C.Y. Robert, and M. Rosenbaum, *Estimating the efficient price from the order flow: a brownian cox process approach*, Stochastic Processes and their Applications **123** (2013), no. 7, 2603–2619.
- [EAM18] B. El Asri and S. Mazid, *Zero-sum stochastic differential game in finite horizon involving impulse controls*, Applied Mathematics & Optimization (2018), 1–33.
- [Ega08] M. Egami, *A direct solution method for stochastic impulse control problems of one-dimensional diffusions*, SIAM Journal on Control and Optimization **47** (2008), no. 3, 1191–1218.
- [EH88] J.F. Eastham and K.J. Hastings, *Optimal impulse control of portfolios*, Mathematics of Operations Research **13** (1988), no. 4, 588–605.
- [EK09] S.N. Ethier and T.G. Kurtz, *Markov processes: characterization and convergence*, vol. 282, John Wiley & Sons, 2009.
- [FK19] G. Ferrari and T. Koch, *On a strategic model of pollution control*, Annals of Operations Research **275** (2019), no. 2, 297–319.
- [FL07] P.A. Forsyth and G. Labahn, *Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance*, Journal of Computational Finance **11** (2007), no. 2, 1.
- [FL12] P. Fodra and M. Labadie, *High-frequency market-making with inventory constraints and directional bets*, arXiv preprint arXiv:1206.4810 (2012).
- [FL13] ———, *High-frequency market-making for multi-dimensional Markov processes*, arXiv preprint arXiv:1303.7177 (2013).
- [FS06] W.H. Fleming and H.M. Soner, *Controlled Markov processes and viscosity solutions*, 2nd ed., vol. 25, Springer Science & Business Media, 2006.
- [GKM11] K. Giesecke, H. Kakavand, and M. Mousavi, *Exact simulation of point processes with stochastic intensities*, Operations Research **59** (2011), no. 5, 1233–1245.
- [GL15] O. Guéant and C.-A. Lehalle, *General intensity shapes in optimal liquidation*, Mathematical Finance **25** (2015), no. 3, 457–495.
- [GLFT12] O. Guéant, C.-A. Lehalle, and J. Fernandez-Tapia, *Optimal portfolio liquidation with limit orders*, SIAM Journal on Financial Mathematics **3** (2012), no. 1, 740–764.
- [GLFT13] ———, *Dealing with the inventory risk: a solution to the market making problem*, Mathematics and Financial Economics **7** (2013), no. 4, 477–507.

- [Gué17] O. Guéant, *Optimal market making*, Applied Mathematical Finance **24** (2017), no. 2, 112–154.
- [HFL12] Y. Huang, P.A. Forsyth, and G. Labahn, *Combined fixed point and policy iteration for Hamilton–Jacobi–Bellman equations in finance*, SIAM Journal on Numerical Analysis **50** (2012), no. 4, 1861–1882.
- [HFL13] ———, *Inexact arithmetic considerations for direct control and penalty methods: American options under jump diffusion*, Applied Numerical Mathematics **72** (2013), 33–51.
- [How60] R.A. Howard, *Dynamic programming and markov processes*, The MIT Press, Cambridge, MA (1960).
- [HS81] T. Ho and H.R. Stoll, *Optimal dealer pricing under transactions and return uncertainty*, Journal of Financial Economics **9** (1981), no. 1, 47–73.
- [JP82] J. Jacod and P. Protter, *Quelques remarques sur un nouveau type d’équations différentielles stochastiques*, Séminaire de Probabilités XVI 1980/81, Springer, 1982, pp. 447–458.
- [JP93] M. Jeanblanc-Picqué, *Impulse control method and exchange rate*, Mathematical Finance **3** (1993), no. 2, 161–177.
- [JS02] J. Jacod and A. Shiryaev, *Limit theorems for stochastic processes*, 2nd ed., vol. 288, Springer Science & Business Media, 2002.
- [JYC09] M. Jeanblanc, M. Yor, and M. Chesney, *Mathematical methods for financial markets*, Springer Science & Business Media, 2009.
- [Kor99] R. Korn, *Some applications of impulse control in mathematical finance*, Mathematical Methods of Operations Research **50** (1999), no. 3, 493–518.
- [LTT18] V. Lemaire, M. Thieullen, and N. Thomas, *Exact simulation of the jump times of a class of piecewise deterministic Markov processes*, Journal of Scientific Computing **75** (2018), no. 3, 1776–1807.
- [MØ98] G. Mundaca and B. Øksendal, *Optimal stochastic intervention control with application to the exchange rate*, Journal of Mathematical Economics **2** (1998), no. 29, 225–243.
- [Nik06] A. Nikeghbali, *An essay on the general theory of stochastic processes*, Probability Surveys **3** (2006), 345–412.
- [Oga81] Y. Ogata, *On Lewis’ simulation method for point processes*, IEEE Transactions on Information Theory **27** (1981), no. 1, 23–31.
- [ØS09] B. Øksendal and A. Sulem, *Applied stochastic control of jump processes*, 3rd ed., Springer, 2009.

- [Pro04] P.E. Protter, *Stochastic integration and differential equations*, 2nd ed., Springer, 2004.
- [Sey09] R.C. Seydel, *Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions*, *Stochastic Processes and their Applications* **119** (2009), no. 10, 3719–3748.
- [Son88] H.M. Soner, *Optimal control of jump-Markov processes and viscosity solutions*, *Stochastic differential systems, stochastic control theory and applications*, Springer, 1988, pp. 501–511.
- [WF08] J. Wang and P.A. Forsyth, *Maximal use of central differencing for Hamilton–Jacobi–Bellman PDEs in finance*, *SIAM Journal on Numerical Analysis* **46** (2008), no. 3, 1580–1601.
- [WW17] S. Wang and Z. Wu, *Stochastic maximum principle for optimal control problems of forward-backward delay systems involving impulse controls*, *Journal of Systems Science and Complexity* **30** (2017), no. 2, 280–306.
- [Zab19] D. Zabaljauregui, *A fixed-point policy-iteration-type algorithm for symmetric nonzero-sum stochastic impulse games*, arXiv preprint arXiv:1909.03574 (2019).