HAL
open science

# Optimal High Frequency Trading in a Pro-Rata Microstructure with Predictive Information

Fabien Guilbaud, Huyên Pham

**To cite this version:**

Fabien Guilbaud, Huyên Pham. Optimal High Frequency Trading in a Pro-Rata Microstructure with Predictive Information. 2012. hal-00697125

**HAL Id: hal-00697125**

**https://hal.science/hal-00697125v1**

Preprint submitted on 14 May 2012

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

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

# Optimal high-frequency trading in a pro-rata microstructure with predictive information

Fabien GUILBAUD
Laboratoire de Probabilités et Modèles Aléatoires
CNRS, UMR 7599
Université Paris 7 Diderot
fabien.guilbaud@gmail.com

Huyên PHAM
Laboratoire de Probabilités et Modèles Aléatoires
CNRS, UMR 7599
Université Paris 7 Diderot
pham@math.univ-paris-diderot.fr
and CREST-ENSAE

May 14, 2012

**Abstract**

We propose a framework to study optimal trading policies in a one-tick pro-rata limit order book, as typically arises in short-term interest rate futures contracts. The high-frequency trader has the choice to trade via market orders or limit orders, which are represented respectively by impulse controls and regular controls. We model and discuss the consequences of the two main features of this particular microstructure: first, the limit orders sent by the high frequency trader are only partially executed, and therefore she has no control on the executed quantity. For this purpose, cumulative executed volumes are modelled by compound Poisson processes. Second, the high frequency trader faces the overtrading risk, which is the risk of brutal variations in her inventory. The consequences of this risk are investigated in the context of optimal liquidation.

The optimal trading problem is studied by stochastic control and dynamic programming methods, which lead to a characterization of the value function in terms of an integro quasi-variational inequality. We then provide the associated numerical resolution procedure, and convergence of this computational scheme is proved. Next, we examine several situations where we can on one hand simplify the numerical procedure by reducing the number of state variables, and on the other hand focus on specific cases of practical interest. We examine both a market making problem and a best execution problem in the case where the mid-price process is a martingale. We also detail a high frequency trading strategy in the case where a (predictive) directional information on the mid-price is available. Each of the resulting strategies are illustrated by numerical tests.

**Keywords:** Market making, limit order book, pro-rata microstructure, inventory risk, marked point process, stochastic control.

# 1 Introduction

In most of modern public security markets, the price formation process, or price discovery, results from competition between several market agents that take part in a public auction. In particular, day trading sessions, which are also called continuous trading phases, consist of continuous double auctions. In these situations, liquidity providers[1] continuously set bid and ask prices for the considered security, and the marketplace publicly displays a (possibly partial) information about these bid and ask prices, along with transactions prices. The action of continuously providing bid and ask quotes during day trading sessions is called *market making*, and this role was tradionnally performed by specialist firms. However, due to the recent increased availability of electronic trading technologies, as well as regulatory changes, a large range of investors are now able to implement such market making strategies. These strategies are part of the broader category of high frequency trading (HFT) strategies, which are characterized by the fact that they facilitate a larger number of orders being sent to the market per unit of time. HFT takes place in the continuous trading phase, and therefore in the double continuous auction context, but actual mechanisms that implement this general continuous double auction set-up directly influence the price formation process and, as a consequence, HFT strategies.

In this work, we shall focus on the case where the continuous double auction is implemented by a limit order book (LOB), operated under the pro-rata microstructure[2], see [10] and [1]. This microstructure can be encountered on some derivatives markets, and especially in short-term interest rate (STIR) futures markets, also known as financial futures, traded e.g. on LIFFE (London International Financial Futures and options Exchange) or on CME (Chicago Mercantile Exchange). We will describe this microstructure in depth in Section 2, but the general mechanism of this microstructure is as follows: an incoming market order is dispatched on all active limit orders at the best price, with each limit order contributing to execution in proportion to its volume. In particular, we will discuss the two main consequences of this microstructure on HFT strategies which are the oversizing of the best priced slices of the LOB and the overtrading risk.

Our main goal is to construct an HFT strategy, by means of optimal stochastic control, that targets the pro-rata microstructure. We allow both limit orders and market orders in this HFT strategy, modelled respectively as continuous and impulse controls, due to considerations about direct trading costs. From a modelling point of view, the key novelty is that we take into account partial execution for limit orders, which is crucial in the pro-rata case. For this purpose we introduce a Poissonian model for trades processes, that can be fitted to a large class of real-world execution processes, since we make few assumptions about the distributions of execution volumes. From a practical trading point of view, we allow the HFT to input predictive information about price evolution into the strategy, so that our algorithm can be seen as an information-driven HFT strategy (this situation is sometimes called HFT with superior information, see [4]). We derive the dynamic programming equation corresponding to this mixed impulse/regular control problem. Moreover, we are able to

[1] In this paper, we call liquidity provider any investor that currently trades with limit orders

[2] This microstructure differs from the price/time microstructure that can be encountered on most cash equity markets, see [9].

reduce the number of relevant state variables to one in two situations of practical interest: first, in the simple case where the mid-price is a martingale, and second, in the case where the mid-price is a Lévy process, in particular when the HFT has predictive information on price trend, in line with recent studies [5]. We provide a computational algorithm for the resolution of the dynamic programming equation, and prove the convergence of this scheme. We illustrate numerically the behavior of the strategy and perform a simulated data benchmarked backtest.

High-frequency trading has recently received sustained academic interest. The reference work for inventory-based high frequency trading is Avellaneda and Stoikov (2008) [2]. The authors present the HFT problem as an inventory management problem and define inventory risk as the risk of holding a non-zero position in a risky asset. They also provide a closed-form approximate solution in a stylized market model where the controls are continuous. Several works are available that describe optimal strategies for HFT on cash equities or foreign exchange, e.g. [11], [7], [9] or [12]. Guéant, Tapia and Lehalle ([7]) provide extensive analytical treatment of the Avellaneda and Stoikov model. Veraart ([12]) includes market orders (that are modelled as impulse controls) as well as limit order in the context of FX trading. Guilbaud and Pham ([9]) study market/limit orders HFT strategies on stocks with a focus on the price/time microstructure and the bid/ask spread modelling. More recently, Cartea, Jaimungal and Ricci ([4]) consider a HFT strategy that takes into account influence of trades on the LOB, and give the HFT superior information about the security price evolution. A growing literature is dedicated to modelling the dynamics of the limit order book itself, and its consequences for the price formation process. A popular approach is the Poisson Limit Order Book model as in Cont and de Larrard ([5]). These authors are able to retrieve a predictive information on price behavior (together with other LOB features) based on the current state of the order book. Finally, in empirical literature, much work is available for cash equities e.g. [8], but very few is dedicated to markets operating under pro-rata microstructure. We would like to mention the work by Field and Large ([6]), which provides a detailed empirical description of pro-rata microstructure.

This paper is organized as follows: in Section 2, we detail the market model and explain the high frequency trading strategy. In Section 3, we formulate the control problem, derive the corresponding dynamic programming equation (DPE) for the value function, and state some bounds and symmetry properties. We also simplify the DPE in two cases of practical interest, namely the case where the price is a martingale, and the case where the investor has predictive information on price trend available. In Section 4, we provide the numerical algorithm to solve the DPE, and we study the convergence of the numerical scheme towards the exact solution, by proving the monotonicity, stability and consistency for this scheme. We also provide numerical tests including computations of the optimal policies and performance analysis on a simulated data backtest. Finally, in Section 5, we show how to extend our model in the optimal liquidation case, i.e. when the investor's objective is to minimize the trading costs for unwinding her portfolio.

## 2 Market model

Let us fix a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. It is assumed that all random variables and stochastic processes are defined on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$.

**Prices in a one-tick microstructure.** We denote by $P$ the midprice, defined as a Markov process with generator $\mathcal{P}$ valued in $\mathbb{P}$. We shall assume that $P$ is a special semimartingale such that its predictable finite variation term $A$ satisfies the canonical structure: $dA_t \ll d < P >_t$, with a bounded density process:

$$\theta_t \;=\; \frac{dA_t}{d < P >_t}, \tag{2.1}$$

and the sharp bracket process $< P >$ is absolutely continuous with respect to the Lebesgue measure:

$$d < P >_t \;=\; \varrho(P_t)dt, \tag{2.2}$$

for some positive continuous function $\varrho$ on $P$. We denote by $\delta > 0$ the tick size, and we shall assume that the spread is constantly equal to $\delta$, i.e. the best ask (resp. bid) price is $P^a := P + \frac{\delta}{2}$ (resp. $P^b := P - \frac{\delta}{2}$). This assumption corresponds to the case of the so-called *one-tick microstructure* [6], which can be encountered e.g. on short term interest rates futures contracts.

**Trading strategies.** For most of investors, the brokerage costs are paid when a transaction occurs, but new limit order submission, update or cancel are free of charge. Therefore, the investor can submit or update her quotes at any time, with no costs associated to this operation: it is then natural to model the limit order strategy (*make* strategy) as a continuous time predictable control process. On the contrary, market orders lead to immediate execution, and are costly, so that continuous submission of market orders would lead to bankruptcy. Therefore, we choose to model the market order strategy (*take* strategy) as impulse controls. More precisely, we model trading strategies by a pair $\alpha = (\alpha^{make}, \alpha^{take})$ in the form:

$$\alpha^{make} \;=\; \big(L^a_t, L^b_t\big)_{t\geq 0}, \qquad \alpha^{take} \;=\; \big(\tau_n, \xi_n\big)_{n\in\mathbb{N}}.$$

The predictable processes $L^a$ and $L^b$, valued in $\{0, 1\}$ represent the possible *make regimes*: when $L^a_t = 1$ (resp. $L^b_t = 1$) this means that the investor has active limit orders at the best ask price (resp. best bid price) at time $t$, else, if $L^a_t = 0$ (resp. $L^b_t = 0$) this means that the investor has no active order at the best ask price (resp. best bid price) at time $t$. Practical implementation of such rule would be, for example, to send a limit order with a fixed quantity, when the corresponding control is 1, and cancel it when it turns to 0. On the other hand, $(\tau_n)_{n\in\mathbb{N}}$ is an increasing sequence of stopping times, representing the times when the investor chooses to trade at market, and $\xi_n$, $n \geq 0$ are $\mathcal{F}_{\tau_n}$-measurable random variables valued in $\mathbb{R}$, representing the quantity purchased if $\xi_n \geq 0$ or sold if $\xi_n < 0$.

**Execution processes in a pro-rata microstructure.** The pro-rata microstructure (see

[10] for extensive presentation and discussion) can be schematically described as follows[3]: when a market order comes in the pro-rata limit order book, its volume is dispatched among all active limit orders at best prices, proportionnally to each limit orders volumes, and therefore create several transactions (see Figure 1).

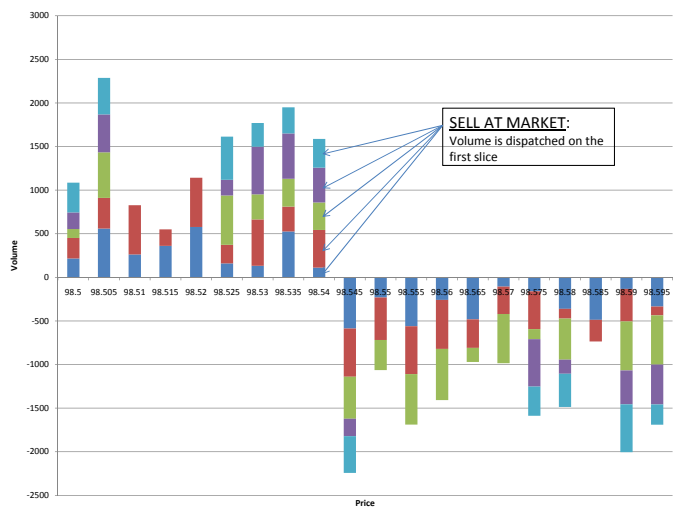



Figure 1: Schematic view of the pro-rata market microstructure.

This pro-rata microstructure fundamentally differs from price-time microstructure [9] for two reasons: first, several limit orders at the best prices receive incoming market order flow, regardless of the time priority, and second, market makers tend to oversize their liquidity offering (that is, posting limit order with much higher volume than they actually want to trade) in order to increase their transaction volume. For example, on the three-months EURIBOR futures contracts, the liquidity available at the best prices is 200 times higher than the average transaction size. Therefore, to model a market making strategy, one must take into account the fact that limit orders are always oversized, so that the executed volume is a random variable on which the market maker has no control[4]. More precisely, let $N^a$ (resp. $N^b$) be a Poisson process of intensity $\lambda^a > 0$ (resp. $\lambda^b$), whose jump times represent the times when execution by a market order flow occurs at best ask (resp. best bid), and we assume that $N^a$ and $N^b$ are independent. Let $(\zeta_n^a)_{n\in\mathbb{N}^*}$ and $(\zeta_n^b)_{n\in\mathbb{N}^*}$ be two independent sequences of i.i.d integrable random variables valued in $(0,\infty)$, of distribution laws $\mu^a$ and $\mu^b$, which represent the transacted volume of the $n^{th}$ execution at best ask and best bid. We denote by $\nu^a(dt,dz)$ (resp. $\nu^b(dt,dz)$) the Poisson random measure associated to the marked point process $(N^a,(\zeta_n^a)_{n\in\mathbb{N}^*})$ (resp. $(N^b,(\zeta_n^b)_{n\in\mathbb{N}^*})$) of intensity measure $\lambda^a\mu^a(dz)dt$

[3]For a detailled description of actual trading rules, and a general overview of STIR futures trading, we refer to [1] and references therein.

[4]This differs from the price-time microstructure case, in which the market maker can control the upper bound of the execution volume by adjusting the limit order volume.

(resp. $\lambda^b \mu^b(dz)dt$), which is often identified with the compound Poisson processes

$$\vartheta_t^a = \sum_{n=1}^{N_t^a} \zeta_n^a = \int_0^t \int_0^\infty z\, \nu^a(dt,dz), \qquad \vartheta_t^b = \sum_{n=1}^{N_t^b} \zeta_n^b = \int_0^t \int_0^\infty z\, \nu^b(dt,dz). \quad (2.3)$$

representing the cumulative volume of transaction at ask, and bid. Notice that these processes model only the trades in which the investor has participated.

**Cash holdings and inventory.** The cash holdings process $X$ and the cumulated number of stocks $Y$ (also called inventory) hold by the investor evolve according to the following dynamics:

$$dX_t = L_t^a\big(P_{t^-} + \frac{\delta}{2}\big)d\vartheta_t^a - L_t^b\big(P_{t^-} - \frac{\delta}{2}\big)d\vartheta_t^b, \quad \tau_n \le t < \tau_{n+1} \quad (2.4)$$

$$dY_t = L_t^b d\vartheta_t^b - L_t^a d\vartheta_t^a, \qquad \tau_n \le t < \tau_{n+1} \quad (2.5)$$

$$X_{\tau_n} - X_{\tau_n -} = -\xi_n P_{\tau_n} - |\xi_n|\big(\frac{\delta}{2} + \varepsilon\big) - \varepsilon_0, \quad (2.6)$$

$$Y_{\tau_n} - Y_{\tau_n -} = \xi_n. \quad (2.7)$$

The equations (2.4)-(2.5) model the evolution of the cash holdings and inventory under a limit order (make) strategy, while equations (2.6)-(2.7) describe the jump on the cash holdings and inventory when posting a market order (take) strategy, subject to a per share fee $\varepsilon > 0$ and a fixed fee $\varepsilon_0 > 0$. In the sequel, we impose the natural admissibility condition that the size of the market order should not be larger than the current inventory, i.e. $|\xi_n| \le |Y_{\tau_n -}|$, $n \ge 0$, and we shall denote by $\mathcal{A}$ the set of all admissible make and take strategies $\alpha = (\alpha^{make}, \alpha^{take})$.

**Remark 2.1** Let us define the process $V_t = X_t + Y_t P_t$, which represents at time $t$ the marked-to-market value of the portfolio (or book value of the portfolio). From (2.4)-(2.5)-(2.6)-(2.7), we see that its dynamics is governed by:

$$dV_t = \frac{\delta}{2}(L_t^b d\vartheta_t^b + L_t^a d\vartheta_t^a) + Y_{t^-} dP_t, \quad (2.8)$$

$$V_{\tau_n} - V_{\tau_n -} = -|\xi_n|(\frac{\delta}{2} + \varepsilon) - \varepsilon_0. \quad (2.9)$$

In equation (2.9), we notice that a trade at market will always diminish the marked to market value of our portfolio, due to the fact that we have to "cross the spread", hence trade at a least favorable price. On the other hand, in equation (2.8), the term $\int \frac{\delta}{2}(L_t^b d\vartheta_t^b + L_t^a d\vartheta_t^a)$ is always positive, and represents the profit obtained from a limit order execution, while the term $\int Y_{t^-} dP_t$ represents the portfolio value when holding shares in the stock, hence inducing an inventory risk, which one wants to reduce its variance.

# 3 Market making optimization procedure

## 3.1 Control problem formulation

The market model in the previous section is fully determined by the state variables $(X, Y, P)$ controlled by the limit/market orders strategies $\alpha = (\alpha^{make}, \alpha^{take}) \in \mathcal{A}$. The market maker

wants to optimize her profit over a finite time horizon $T$ (typically short term), while keeping control of her inventory risk, and to get rid of any risky asset by time $T$. We choose a mean-variance optimization criterion, and the goal is to

$$\text{maximize} \quad \mathbf{E}\Big[X_T - \gamma \int_0^T Y_{t^-}^2 d < P >_t \Big] \text{ over all strategies } \alpha \in \mathcal{A}, \text{ s.t } Y_T = 0. \tag{3.1}$$

The integral $\int_0^T Y_{t^-}^2 d < P >_t$ is a quadratic penalization term for holding a non zero inventory in the stock, and $\gamma > 0$ is a risk aversion parameter chosen by the investor.

Let us now rewrite problem (3.1) in a more standard formulation. Notice indeed that one can remove mathematically the constraint $Y_T = 0$ on the inventory control, by introducing the liquidation function:

$$L(x, y, p) \quad = \quad x + yp - |y|\big(\frac{\delta}{2} + \varepsilon\big) - \varepsilon_0,$$

which represents the cash obtained after an immediate liquidation of the inventory via a market order. Thus, problem (3.1) is formulated equivalently as

$$\text{maximize} \quad \mathbf{E}\Big[L(X_T, Y_T, P_T) - \gamma \int_0^T Y_t^2 \varrho(P_t) dt\Big] \quad \text{over all strategies } \alpha \in \mathcal{A}, \tag{3.2}$$

where we used also (2.2). Let us then define the value function for the problem (3.2):

$$v(t, x, y, p) \quad = \quad \sup_{\alpha \in \mathcal{A}} \mathbf{E}_{t,x,y,p}\Big[L(X_T, Y_T, P_T) - \gamma \int_t^T Y_s^2 \varrho(P_s) ds\Big], \tag{3.3}$$

for $t \in [0, T]$, $(x, y, p) \in \mathbb{R}^2 \times \mathbb{P}$. Here, given $\alpha \in \mathcal{A}$, $\mathbf{E}_{t,x,y,p}$ denotes the expectation operator under which the process $(X, Y, P)$ solution to (2.4)-(2.5)-(2.6)-(2.7) with initial state $(X_{t^-}, Y_{t^-}, P_{t^-}) = (x, y, p)$, is taken. Problem (3.3) is a mixed impulse/regular control problem in Markov model with jumps that we shall study by dynamic programming methods.

First, we state some bounds on the value function.

**Proposition 3.1** *There exists some constant $K_P$ (depending only on the price process and $\gamma$) such that for all $(t, x, y, p) \in [0, T] \times \mathbb{R}^2 \times \mathbb{P}$,*

$$L(x, y, p) \ \leq \ v(t, x, y, p) \ \leq \ x + yp + \frac{\delta}{2}\big(\lambda^a \bar{\mu}^a + \lambda^b \bar{\mu}^b\big)(T - t) + K_P, \tag{3.4}$$

*where $\bar{\mu}^a = \int_0^\infty z\mu^a(dz)$, $\bar{\mu}^b = \int_0^\infty z\mu^b(dz)$ are the mean of the distribution laws $\mu^a$ and $\mu^b$.*

**Proof.** The lower bound in (3.4) is derived easily by considering the particular strategy, which consists of liquidating immediately all the current inventory via a market order, and then doing nothing else until the final horizon. Let us now focus on the upper bound. Observe that in the definition of the value function in (3.3), we can restrict obviously to controls $\alpha \in \mathcal{A}$ s.t.

$$\mathbb{E}\Big[\int_0^T Y_t^2 d < P >_t \Big] \quad < \quad \infty. \tag{3.5}$$

For such strategies, we have:

$$
\begin{aligned}
& \mathbf{E}_{t,x,y,p}\Big[L(X_T,Y_T,P_T)-\gamma\int_t^T Y_s^2 d<P>_s\Big] \\
\leq\ & \mathbf{E}_{t,x,y,p}\Big[V_T-\gamma\int_t^T Y_s^2 d<P>_s\Big] \\
\leq\ & x+yp+\mathbf{E}_{t,x,y,p}\Big[\frac{\delta}{2}\big(\vartheta^a_{T-t}+\vartheta^b_{T-t}\big)+\int_t^T Y_{s^-}dP_s-\gamma\int_t^T Y_s^2 d<P>_s\Big] \\
=\ & x+yp+\mathbf{E}_{t,x,y,p}\Big[\frac{\delta}{2}\big(\vartheta^a_{T-t}+\vartheta^b_{T-t}\big)+\int_t^T \big(Y_{s^-}\theta_s-\gamma Y_s^2\big)d<P>_s\Big].
\end{aligned}
$$

Here, the second inequality follows from the relation (2.8), together with the fact that $L^a, L^b \leq 1$, $\vartheta^a$, $\vartheta^b$ are increasing processes, and also that jumps of $V$ are negative by (2.9). The last equality holds true by (2.1) and the fact that $\int Y_- dM$ is a square-integrable martingale from (3.5), where $M$ is the martingale part of the semimartingale $P$. Since $\theta$ is bounded and $\gamma > 0$, this shows that for all strategies $\alpha$ satisfying (3.5), we have:

$$
\begin{aligned}
& \mathbf{E}_{t,x,y,p}\Big[L(X_T,Y_T,P_T)-\gamma\int_t^T Y_s^2 d<P>_s\Big] \\
\leq\ & x+yp+\frac{\delta}{2}\mathbb{E}\big[\vartheta^a_{T-t}+\vartheta^b_{T-t}\big]+K\mathbb{E}[<P>_T],
\end{aligned}
$$

for some positive constant $K$, which proves the required result by recalling the characteristics of the compound Poisson processes $\vartheta^a$ and $\vartheta^b$, and since $<P>_T$ is assumed to be square-integrable. □

**Remark 3.1** The terms of the upper bound in (3.4) has a financial interpretation. The term $x+yp$ represents the marked-to-market value of the portfolio evaluated at mid-price, whereas the term $K_P$ stands for a bound on profit for any directional frictionless strategy on the fictive asset that is priced $P$. The term $\frac{\delta}{2}\big(\lambda^a\bar{\mu}^a+\lambda^b\bar{\mu}^b\big)(T-t)$, always positive, represents the upper bound on profit due to market making, i.e. the profit of the strategy participating in every trade, but with no costs associated to managing its inventory.

## 3.2 Dynamic programming equation

For any $(\ell^a,\ell^b)\in\{0,1\}^2$, we introduce the non-local operator associated with the limit order control:

$$
\mathcal{L}^{\ell^a,\ell^b} \ = \ \mathcal{P}+\ell^a\Gamma^a+\ell^b\Gamma^b, \tag{3.6}
$$

where

$$
\begin{aligned}
\Gamma^a\phi(t,x,y,p) &= \lambda^a\int_0^\infty \big[\phi\big(t,x+z(p+\frac{\delta}{2}),y-z,p\big)-\phi(t,x,y,p)\big]\mu^a(dz) \\
\Gamma^b\phi(t,x,y,p) &= \lambda^b\int_0^\infty \big[\phi\big(t,x-z(p-\frac{\delta}{2}),y+z,p\big)-\phi(t,x,y,p)\big]\mu^b(dz),
\end{aligned}
$$

for $(t,x,y,p)$ $[0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{P}$. Let us also consider the impulse operator associated with admissible market order controls, and defined by:

$$\mathcal{M}\phi(t,x,y,p) \;=\; \sup_{e\in[-|y|,|y|]} \phi\big(t,x-ep-|e|(\frac{\delta}{2}+\varepsilon)-\varepsilon_0,y+e,p\big).$$

The dynamic programming equation (DPE) associated to the control problem (3.3) is a quasi-variational inequality (QVI) in the form:

$$\min\Big[-\frac{\partial v}{\partial t}-\sup_{(\ell^a,\ell^b)\in\{0,1\}^2}\mathcal{L}^{\ell^a,\ell^b}v \;+\; \gamma g\;,\; v-\mathcal{M}v\Big] \;=\; 0,\;\text{ on }\;[0,T)\times\mathbb{R}^2\times\mathbb{P}, \quad (3.7)$$

together with the terminal condition:

$$v(T,.) \;=\; L, \qquad \text{on }\;\mathbb{R}^2\times\mathbb{P}, \quad (3.8)$$

where we denoted by $g$ the function: $g(y,p)=y^2\varrho(p)$. This DPE may be written explicitly as:

$$\begin{aligned}\min\Big[-\frac{\partial v}{\partial t}-\mathcal{P}v-\lambda^a\Big(\int_0^\infty\big[v\big(t,x+z(p+\frac{\delta}{2}),y-z,p\big)-v(t,x,y,p)\big]\mu^a(dz)\Big)_+ \quad (3.9)\\ -\lambda^b\Big(\int_0^\infty\big[v\big(t,x-z(p-\frac{\delta}{2}),y+z,p\big)-v(t,x,y,p)\big]\mu^b(dz)\Big)_+ +\gamma y^2\varrho(p)\;;\\ v(t,x,y,p)-\sup_{e\in[-|y|,|y|]}v\big(t,x-ep-|e|(\frac{\delta}{2}+\varepsilon)-\varepsilon_0,y+e,p\big)\Big] \;=\; 0,\end{aligned}$$

for $(t,x,y,p)\in[0,T)\times\mathbb{R}^2\times\mathbb{P}$, together with the terminal condition:

$$v(T,x,y,p) \;=\; x+yp-|y|\big(\frac{\delta}{2}+\varepsilon\big)-\varepsilon_0,\quad \forall(x,y,p)\in\mathbb{R}^2\times\mathbb{P}. \quad (3.10)$$

By standard methods of dynamic programming, one can show that the value function in (3.3) is the unique viscosity solution under growth conditions determined by (3.4) to the DPE (5.2)-(3.10) of dimension 3 (in addition to the time variable).

### 3.3 Dimension reduction in the Lévy case

We now consider a special case on the mid-price process where the market making control problem can be reduced to the resolution of a one-dimensional variational inequality involving only the inventory state variable. We shall suppose actually that $P$ is a Lévy process so that

$$\mathcal{P}I_{\mathbb{P}} \;=\; c_{_P}, \quad \text{and} \quad \varrho \;\text{ is a constant}, \quad (3.11)$$

where $I_{\mathbb{P}}$ is the identity function on $\mathbb{P}$, i.e. $I_{\mathbb{P}}(p)=p$, and $\varrho>0$, $c_{_P}$ are real constants depending on the characteristics triplet of $P$. Two practical examples are:

• **Martingale case**: The mid-price process $P$ is a martingale, so that $\mathcal{P}I_{\mathbb{P}}=0$. This martingale assumption in a high-frequency context reflects the idea that the market maker has no information on the future direction of the stock price.

• **Trend information**: To remove the martingale assumption, one can introduce some knowledge about the price trend. A typical simple example is when $P$ follows an arithmetic Brownian motion (Bachelier model). A more relevant example is described by a pure jump process $P$ valued in the discrete grid $\delta\mathbb{Z}$ with tick $\delta > 0$, and such that

$$\begin{aligned}
\mathbf{P}\big(P_{t+h} - P_t = \delta \,|\mathcal{F}_t\big) &= \pi^+ h + o(h) \\
\mathbf{P}\big(P_{t+h} - P_t = -\delta \,|\mathcal{F}_t\big) &= \pi^- h + o(h) \\
\mathbf{P}\big(|P_{t+h} - P_t| > \delta \,|\mathcal{F}_t\big) &= o(h),
\end{aligned}$$

where $\pi^+, \pi^- > 0$, and $o(h)$ is the usual notation meaning that $\lim_{h\to 0} o(h)/h = 0$. Relation (3.11) then holds with $c_P = \varpi\delta$, where $\varpi = \pi^+ - \pi^-$ represents a constant information about price direction, and $\varrho = (\pi^+ + \pi^-)\delta^2$. In a high-frequency context, this model is of practical interest as it provides a way to include a (predictive) information about price direction. For example, work have been done in [5] to infer the future prices movements (at the scale of a few seconds) from the current state of the limit order book in a Poisson framework. In this work, as well as in our real data tests, the main quantities of interest are the volume offered at the best prices in the limit order book, also known as the imbalance.

In this Lévy context, the value function $v$ is decomposed into the form:

$$v(t,x,y,p) \;=\; L(x,y,p) + w(t,y), \tag{3.12}$$

where $w$ is solution to the integral variational inequality:

$$\min\Big[-\frac{\partial w}{\partial t} - yc_P + \gamma\varrho y^2 - \mathcal{I}^a w - \mathcal{I}^b w \;,\; w - \tilde{\mathcal{M}} w\Big] \;=\; 0, \quad \text{on } [0,T)\times\mathbb{R}, \tag{3.13}$$

together with the terminal condition:

$$w(T,y) \;=\; 0, \quad \forall y \in \mathbb{R}, \tag{3.14}$$

where $\mathcal{I}^a$ and $\mathcal{I}^b$ are the nonlocal integral operators:

$$\begin{aligned}
\mathcal{I}^a w(t,y) &= \lambda^a\Big(\int_0^\infty \Big[w(t,y-z) - w(t,y) + z\frac{\delta}{2} + (\frac{\delta}{2}+\varepsilon)(|y| - |y-z|)\Big]\mu^a(dz)\Big)_+ \\
\mathcal{I}^b w(t,y) &= \lambda^b\Big(\int_0^\infty \Big[w(t,y+z) - w(t,y) + z\frac{\delta}{2} + (\frac{\delta}{2}+\varepsilon)(|y| - |y+z|)\Big]\mu^b(dz)\Big)_+,
\end{aligned}$$

and $\tilde{\mathcal{M}}$ is the nonlocal operator:

$$\tilde{\mathcal{M}} w(t,y) \;=\; \sup_{e\in[-|y|,|y|]} \Big[w(t,y+e) - (\frac{\delta}{2}+\varepsilon)(|y+e| + |e| - |y|) - \varepsilon_0\Big].$$

The interpretation of the decomposition (3.12) is the following. The term $L(x,y,p)$ represents the book value that the investor would obtain by liquidating immediately with a market order, and $w$ is an additional correction term taking into account the illiquidity effects induced by the bid-ask spread and the fee, as well as the execution risk when submitting limit orders. Moreover, in the Lévy case, this correction function $w$ depends only on time and inventory. From (3.4), we have the following bounds on the function $w$:

$$0 \;\le\; w(t,y) \;\le (\frac{\delta}{2}+\varepsilon)|y| + \frac{\delta}{2}\big(\lambda^a\bar{\mu}^a + \lambda^b\bar{\mu}^b\big)(T-t) + K_P, \quad \forall (t,y)\in[0,T]\times\mathbb{R}.$$

Actually, we have a sharper upper bound in the Lévy context.

**Proposition 3.2** *Under* (3.11)*, we have:*

$$0 \;\le\; w(t,y) \;\;\le\;\; (T-t)\Big[\frac{c_P^2}{4\gamma\rho} + \lambda^a(\delta+\epsilon)\bar\mu^a + \lambda^b(\delta+\epsilon)\bar\mu^b\Big],$$

*for all* $(t,x,y,p) \in [0,T]\times\mathbb{R}^2\times\mathbb{P}$.

**Proof.** For any $(x,y,p)\in\mathbb{R}^2\times\mathbb{P}$, we notice that

$$\begin{aligned} & L(x,y,p) - \sup_{e\in[-|y|,|y|]} L(x-ep-|e|(\frac{\delta}{2}+\varepsilon)-\varepsilon_0, y+e, p) \\ = \;\; & \varepsilon_0 + (\frac{\delta}{2}+\varepsilon)\Big[-|y| + \inf_{e\in[-|y|,|y|]} |e| + |y+e|\Big] \;=\; \varepsilon_0 > 0. \end{aligned} \tag{3.15}$$

We also observe that for all $z\ge 0$:

$$\begin{aligned} L(x+z(p+\frac{\delta}{2}), y-z, p) - L(x,y,p) \;\; &= \;\; z\frac{\delta}{2} + (\frac{\delta}{2}+\varepsilon)\big(|y|-|y-z|\big) \\ &\le \;\; (\delta+\varepsilon)z, \end{aligned} \tag{3.16}$$

and similarly:

$$L(x-z(p-\frac{\delta}{2}), y+z, p) - L(x,y,p) \;\;\le\;\; (\delta+\varepsilon)z. \tag{3.17}$$

Let us then consider the function $\phi(t,x,y,p) = L(x,y,p) + (T-t)u$, for some real constant $u$ to be determined later. Then, $\phi(T,.) = L$, and by (3.16)-(3.17), we easily check that:

$$\begin{aligned} & -\frac{\partial\phi}{\partial t} - \sup_{(\ell^a,\ell^b)\in\{0,1\}^2} \mathcal{L}^{\ell^a,\ell^b}\phi \;+\; \gamma g \\ \ge \;\; & u - \lambda^a(\delta+\varepsilon)\bar\mu^a - \lambda^b(\delta+\varepsilon)\bar\mu^b - yc_P + \gamma y^2\rho. \end{aligned}$$

The r.h.s. of this last inequality is a second order polynomial in $y$ and therefore it is always nonnegative iff:

$$c_P^2 - 4\gamma\rho(u - \lambda^a(\delta+\varepsilon)\bar\mu^a - \lambda^b(\delta+\varepsilon)\bar\mu^b) \;\;\le\;\; 0,$$

which is satisfied once the constant $u$ is large enough, namely:

$$u \;\;\ge\;\; \hat u \;:=\; \frac{c_P^2}{4\gamma\rho} + \lambda^a(\delta+\epsilon)\bar\mu^a + \lambda^b(\delta+\epsilon)\bar\mu^b.$$

For such choice of $u=\hat u$, and denoting by $\hat\phi$ the associated function: $\hat\phi(t,x,y,p) = L(x,y,p) + (T-t)\hat u$ we have

$$-\frac{\partial\hat\phi}{\partial t} - \sup_{(\ell^a,\ell^b)\in\{0,1\}^2} \mathcal{L}^{\ell^a,\ell^b}\hat\phi + \gamma g \;\;\ge\;\; 0,$$

which shows, together with (3.15), that $\hat\phi$ is a supersolution of (3.7)-(3.8). From comparison principle for this variational inequality, we deduce that

$$v \;\;\le\;\; \hat\phi \;\text{ on }\; [0,T]\times\mathbb{R}^2\times\mathbb{P},$$

which shows the required upper bound for $w = v - L$. □

Finally, from (3.13)-(3.14), and in the case where $\lambda^a = \lambda^b$, $\mu^a = \mu^b$, and by stressing the dependence of $w$ in $c_P$, we see that $w$ satisfies the symmetry relation:

$$w(t,y,c_P) \;\;=\;\; w(t,-y,-c_P), \quad \forall (t,y)\in[0,T]\times\mathbb{R}. \tag{3.18}$$

# 4 Numerical resolution

In this section, we focus on the numerical resolution of the integral variational inequality (3.13)-(3.14), which characterizes the reduced value function of the market-making problem in the Lévy case.

## 4.1 Numerical scheme

We provide a computational scheme for the integral variational inequality (3.13). We first consider a time discretization of the interval $[0,T]$ with time step $h = T/N$ and a regular time grid $\mathbb{T}_N = \{t_k = kh\,,\, k = 0,\ldots,N\}$. Next, we discretize and localize the inventory state space on a finite regular grid: for any $M > 0$ and $N_Y \in \mathbb{N}$, and denoting by $\Delta_Y = \dfrac{M}{N_Y}$, we set:

$$\mathbb{Y}_M \;=\; \big\{y_i = i\Delta_Y,\; i = -N_Y,\ldots,N_Y\big\}.$$

We denote by $\mathrm{Proj}_M(y) := -M \vee (y \wedge M)$, and consider the discrete approximating distribution of $\mu^a$ and $\mu^b$, defined by:

$$\hat{\mu}^a \;=\; \sum_{i\in\mathbb{Z}^+} \mu^a([i\Delta_Y;(i+1)\Delta_Y))\delta_{i\Delta_Y}\;\;, \qquad \hat{\mu}^b \;=\; \sum_{i\in\mathbb{Z}^+} \mu^b([i\Delta_Y;(i+1)\Delta_Y))\delta_{i\Delta_Y},$$

with $\delta_x$ the Dirac measure at $x$. We then introduce the operator associated to the explicit time-space discretization of the integral variational inequality (3.13): for any real-valued function $\varphi$ on $[0,T]\times\mathbb{R}$, $t\in[0,T]$, and $y\in\mathbb{R}$, we define:

$$\mathcal{S}^{h,\Delta_Y,M}(t,y,\varphi) \;=\; \max\Big[\mathcal{T}^{h,\Delta_Y,M}(t,y,\varphi)\;;\;\tilde{\mathcal{M}}^{h,\Delta_Y,M}(t,y,\varphi)\Big],$$

where

$$\begin{aligned}
\mathcal{T}^{h,\Delta_Y,M}(t,y,\varphi) \;=\;& \varphi(t,y) - h\gamma\varrho y^2 + hyc_P \\
&+ \lambda^a h\Big(\int_0^\infty \big[\varphi(t,\mathrm{Proj}_M(y-z)) - \varphi(t,y)\big]\hat{\mu}^a(dz) \\
&\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\big]\mu^a(dz)\Big)_+ \\
&+ \lambda^b h\Big(\int_0^\infty \big[\varphi(t,\mathrm{Proj}_M(y+z)) - \varphi(t,y)\big]\hat{\mu}^b(dz) \\
&\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y+z|)\big]\mu^b(dz)\Big)_+,
\end{aligned}$$

and

$$\begin{aligned}
&\tilde{\mathcal{M}}^{h,\Delta_Y,M}(t,y,\varphi) \\
&= \sup_{e\in\mathbb{Y}_M\cap[-|y|,|y|]} \big[\varphi(t,\mathrm{Proj}_M(y+e)) - (\frac{\delta}{2}+\varepsilon)(|y+e|+|e|-|y|) - \varepsilon_0\big]. \qquad (4.1)
\end{aligned}$$

By recalling that $x_+ = \max_{\ell\in\{0,1\}} \ell x$, we see that the operator $\mathcal{T}^{h,\Delta_Y,M}$ may be written also as:

$$\begin{aligned}
\mathcal{T}^{h,\Delta_Y,M}(t,y,\varphi) &= -h\gamma\varrho y^2 + hyc_P + \max_{\ell^a,\ell^b\in\{0,1\}} \Big[\varphi(t,y)(1-\lambda^a h\ell^a - \lambda^b h\ell^b) \qquad (4.2)\\
&\quad + \lambda^a h\ell^a \Big( \int_0^\infty \varphi(t,\mathrm{Proj}_M(y-z))\hat{\mu}^a(dz)\\
&\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\big]\mu^a(dz)\Big)\\
&\quad + \lambda^b h\ell^b \Big( \int_0^\infty \varphi(t,\mathrm{Proj}_M(y+z))\hat{\mu}^b(dz)\\
&\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y+z|)\big]\mu^b(dz)\Big)\Big].
\end{aligned}$$

Note that the integral terms involving $\varphi$ in $\mathcal{S}^{h,\Delta_Y,M}(t,y,\varphi)$ are in fact finite sums, and therefore are readily computable. We also assume, for simplicity sake, that the terms $\int_0^\infty \frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\mu^a(dz)$ and $\int_0^\infty \frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y+z|)\mu^b(dz)$ are exactly computable.

We then approximate the solution $w$ to (3.13)-(3.14) by the function $w^{h,\Delta_Y,M}$ on $\mathbb{T}_N \times \mathbb{Y}_M$ solution to the computational scheme:

$$\begin{aligned}
w^{h,\Delta_Y,M}(t_N,.) &= 0 \qquad (4.3)\\
w^{h,\Delta_Y,M}(t_k,y) &= \mathcal{S}^{h,\Delta_Y,M}(t_{k+1},y,w^{h,\Delta_Y,M})\ ,\ k=0,\ldots,N-1\,,\ y\in\mathbb{Y}_M. \qquad (4.4)
\end{aligned}$$

## 4.2 Convergence of the numerical scheme

In this section, we study the convergence of the numerical scheme (4.3)-(4.4) by showing the monotonicity, stability and consistency properties of this scheme. We denote by $C_b^1([0,T]\times\mathbb{R})$ the set of bounded continuously differentiable functions on $[0,T]\times\mathbb{R}$, with bounded derivatives.

**Proposition 4.1** *(Monotonicity)*
*For any $h>0$ s.t. $h<\dfrac{1}{\lambda^a+\lambda^b}$ the operator $\mathcal{S}^{h,\Delta_Y,M}$ is non-decreasing in $\varphi$, i.e. for any $(t,y)\in[0,T]\times\mathbb{R}$ and any $\varphi,\psi\in C_b^1([0,T]\times\mathbb{R})$ , s.t. $\varphi\le\psi$ :*

$$\mathcal{S}^{h,\Delta_Y,M}(t,y,\varphi) \ \le\ \mathcal{S}^{h,\Delta_Y,M}(t,y,\psi)$$

**Proof.** From the expression (4.2), it is clear that $\mathcal{T}^{h,\Delta_Y,M}(t,y,\varphi)$, and then also $\mathcal{S}^{h,\Delta_Y,M}(t,y,\varphi)$ is monotone in $\varphi$ once $1-\lambda^a h-\lambda^b h>0$. □

**Proposition 4.2** *(Stability)*
*For any $h,\Delta_Y,M>0$ there exists a unique solution $w^{h,\Delta_Y,M}$ to (4.3)-(4.4), and the sequence $(w^{h,\Delta_Y,M})$ is uniformly bounded: for any $(t,y)\in\mathbb{T}_N\times\mathbb{Y}_M$,*

$$-\varepsilon_0 \ \le\ w^{h,\Delta_Y,M}(t,y) \ \le\ (T-t)\Big[\frac{c_P^2}{4\gamma\rho}+\lambda^a(\delta+\epsilon)\bar{\mu}^a+\lambda^b(\delta+\epsilon)\bar{\mu}^b\Big].$$

**Proof.** Existence and uniqueness of $w^{h,\Delta_Y,M}$ follows from the explicit backward scheme (4.3)-(4.4). Let us now prove the uniform bounds. We consider the function

$$\Psi^\star(t) \quad = \quad (T-t)\left[\frac{c_P^2}{4\gamma\rho} + \lambda^a(\delta+\epsilon)\bar{\mu}^a + \lambda^b(\delta+\epsilon)\bar{\mu}^b\right]$$

and notice that $\Psi^\star(t) \geq \mathcal{S}^{h,\Delta_Y,M}(t+h,y,\Psi^\star)$ by the same arguments as in Proposition 3.2. Moreover, we have, by definition, $w^{h,\Delta_Y,M}(T,y) = \Psi^\star(T) = 0$, and therefore, a direct recurrence from (4.3)-(4.4) shows that $w^{h,\Delta_Y,M}(t,y) \leq \Psi^\star(t)$ for all $(t,y) \in \mathbb{T}_n \times \mathbb{Y}_M$, which is the required upper bound for $w^{h,\Delta_Y,M}$.

On the other hand, we notice that $\mathcal{S}^{h,\Delta_Y,M}(t,0,\varphi) \geq \varphi(t,0)$ for any function $\varphi$ on $[0,T]\times\mathbb{R}$, and $t \in [0,T]$, by considering the"diffusive" part of the numerical scheme with the particular controls $\ell^a = \ell^b = 0$. Therefore, since $w^{h,\Delta_Y,M}(T,0) = 0$, we obtain by induction on (4.3)-(4.4) that $w^{h,\Delta_Y,M}(t,0) \geq 0$ for any $t \in \mathbb{T}_N$. Finally, considering the obstacle part of the numerical scheme, with the particular control $e = -y$, shows that $w^{h,\Delta_Y,M}(t,y) \geq w^{h,\Delta_Y,M}(t,0) - \varepsilon_0 \geq -\varepsilon_0$ for any $(t,y) \in \mathbb{T}_N \times \mathbb{Y}_M$, which proves the required lower bound for $w^{h,\Delta_Y,M}$. □

**Proposition 4.3** *(Consistency)*
*For all $(t,y) \in [0,T)\times\mathbb{R}$ and $\varphi \in C_b^1([0,T]\times\mathbb{R})$, we have*

$$\begin{aligned} &\lim_{\substack{(h,\Delta_Y,M)\to(0,0,\infty)\\(t',y')\to(t,y)}} \frac{1}{h}\Big[\varphi(t',y') - \mathcal{T}^{h,\Delta_Y,M}(t'+h,y',\varphi)\Big] \qquad (4.5)\\ = \quad & -\frac{\partial\varphi}{\partial t}(t,y) - yc_P + \gamma\varrho y^2 - \mathcal{I}^a\varphi(t,y) - \mathcal{I}^b\varphi(t,y)\end{aligned}$$

*and*

$$\lim_{\substack{(h,\Delta_Y,M)\to(0,0,\infty)\\(t',y')\to(t,y)}} \tilde{\mathcal{M}}^{h,\Delta_Y,M}(t'+h,y',\varphi) \quad = \quad \tilde{\mathcal{M}}\varphi(t,y) \qquad (4.6)$$

**Proof.** The consistency relation (4.6) follows from the continuity of the function $(t,y,e) \to \varphi(t,y+e) - (\frac{\delta}{2}+\varepsilon)(|y+e|+|e|-|y|) - \varepsilon_0$. On the other hand, we have for all $(t',y') \in [0,T)\times\mathbb{R}$,

$$\begin{aligned}\frac{1}{h}\Big[\varphi(t',y') - \mathcal{T}^{h,\Delta_Y,M}(t'+h,y',\varphi)\Big] \quad &= \quad \frac{1}{h}\big[\varphi(t',y') - \varphi(t'+h,y')\big] - y'c_P + \gamma\rho y'^2 \qquad (4.7)\\ &\quad - \mathcal{I}_a^{h,\Delta_Y,M}(t'+h,y',\varphi) - \mathcal{I}_b^{h,\Delta_Y,M}(t'+h,y',\varphi),\end{aligned}$$

where

$$\begin{aligned}\mathcal{I}_a^{h,\Delta_Y,M}(t,y,\varphi) \quad &= \quad \lambda^a\Big(\int_0^\infty \big[\varphi(t,\mathrm{Proj}_M(y-z)) - \varphi(t,y)\big]\hat{\mu}^a(dz)\\ &\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\big]\mu^a(dz)\Big)_+\\ \mathcal{I}_a^{h,\Delta_Y,M}(t,y,\varphi) \quad &= \quad \lambda^b\Big(\int_0^\infty \big[\varphi(t,\mathrm{Proj}_M(y+z)) - \varphi(t,y)\big]\hat{\mu}^b(dz)\\ &\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y+z|)\big]\mu^b(dz)\Big)_+.\end{aligned}$$

The three first terms of (4.7) converge trivially to $-\dfrac{\partial\varphi}{\partial t}(t,y) - yc_P + \gamma\varrho y^2$ as $h$ goes to zero and $(t',y')$ goes to $(t,y)$. Hence, in order to get the consistency relation, it remains to prove the convergence of $\mathcal{I}_a^{h,\Delta_Y,M}(t'+h,y',\varphi)$ to $\mathcal{I}^a\varphi(t,y)$ as $(h,\Delta_Y,M)$ goes to $(0,0,\infty)$, and $(t',y')$ goes to $(t,y)$ (an identical argument holds for $\mathcal{I}_b^{h,\Delta_Y,M}(t'+h,y',\varphi)$). By writing that $|x_+ - x'_+| \leq |x - x'|$, we have

$$\begin{aligned}
& \Big|\mathcal{I}_a^{h,\Delta_Y,M}(t'+h,y',\varphi) - \mathcal{I}^a\varphi(t,y)\Big| \\
\leq\ & \lambda^a\big|\varphi(t'+h,y') - \varphi(t,y)\big| \\
& + \lambda^a\Big|\int_0^\infty \varphi(t'+h,\mathrm{Proj}_M(y'-z))\hat{\mu}^a(dz) \ - \ \int_0^\infty \varphi(t,y-z)\mu^a(dz)\Big| \\
\leq\ & \lambda^a\big|\varphi(t'+h,y') - \varphi(t,y)\big| \\
& + \lambda^a\Big|\int_0^{M+y'} \varphi(t'+h,y'-z)\hat{\mu}^a(dz) \ - \ \int_0^{M+y'} \varphi(t,y-z)\mu^a(dz)\Big| \\
& + \lambda^a\Big|\int_{M+y'}^\infty \varphi(t'+h,-M)\hat{\mu}^a(dz) \ - \ \int_{M+y'}^\infty \varphi(t,y-z)\mu^a(dz)\Big| \\
\leq\ & \lambda^a\big|\varphi(t'+h,y') - \varphi(t,y)\big| \\
& + \lambda^a\int_0^\infty \big|\varphi(t'+h,y'-\kappa(z)) \ - \ \varphi(t,y-z)\big|\mu^a(dz) \\
& + 2\lambda^a\|\varphi\|_\infty\mu^a\big([M+y',\infty)\big),
\end{aligned}$$

where we denote by $\kappa(z) = \lfloor\dfrac{z}{\Delta_Y}\rfloor\Delta_Y$. Here $\lfloor z\rfloor$ denotes the largest integer smaller than $z$. Since the smooth function $\varphi$ has bounded derivatives, say bounded by $\|\varphi^{(1)}\|_\infty$, it follows that

$$\begin{aligned}
\Big|\mathcal{I}_a^{h,\Delta_Y,M}(t'+h,y',\varphi) - \mathcal{I}^a\varphi(t,y)\Big| \ \leq\ & \lambda^a\|\varphi^{(1)}\|_\infty\big(h + 2|y'-y| + \Delta_Y\big) \\
& + 2\lambda^a\|\varphi\|_\infty\mu^a\big([M+y',\infty)\big),
\end{aligned}$$

which proves that

$$\lim_{\substack{(h,\Delta_Y,M)\to(0,0,\infty)\\(t',y')\to(t,y)}} \mathcal{I}_a^{h,\Delta_Y,M}(t'+h,y',\varphi) \ = \ \mathcal{I}^a\varphi(t,y),$$

hence completing the consistency relation (4.5). □

**Theorem 4.1** *(Convergence)*
*The solution $w^{h,\Delta_Y,M}$ to the numerical scheme ((4.3)-(4.4)) converges locally uniformly to $w$ on $[0,T)\times\mathbb{R}$, as $(h,\Delta_Y,M)$ goes to $(0,0,\infty)$.*

**Proof.** Given the above monotonicity, stability and consistency properties, the convergence of the sequence $(w^{h,\Delta_Y,M})$ towards $w$, which is the unique bounded viscosity solution to (3.13)-(3.14), follows from [3]. We report the arguments for sake of completeness. From

the stability property, the semi-relaxed limits:

$$\begin{aligned} w_*(t,y) &= \liminf_{\substack{(h,\Delta_Y,M)\to(0,0,\infty)\\(t',y')\to(t,y)}} w^{h,\Delta_Y,M}(t',y'), \\ w^*(t,y) &= \limsup_{\substack{(h,\Delta_Y,M)\to(0,0,\infty)\\(t',y')\to(t,y)}} w^{h,\Delta_Y,M}(t',y'), \end{aligned}$$

are finite lower-semicontinuous and upper-semicontinuous functions on $[0,T]\times\mathbb{R}$, and inherit the boundedness of $(w^{h,\Delta_Y,M})$. We claim that $w_*$ are $w^*$ are respectively viscosity super and subsolution of (3.13)-(3.14). Assuming for the moment that this claim is true, we obtain by the strong comparison principle for (3.13)-(3.14) that $w^* \leq w_*$. Since the converse inequality is obvious by the very definition of $w_*$ and $w^*$, this shows that $w_* = w^* = w$ is the unique bounded continuous viscosity solution to (3.13)-(3.14), hence completing the proof of convergence.

In the sequel, we prove the viscosity supersolution property of $w_*$ (a symmetric argument for the viscosity subsolution property of $w^*$ holds true). Let $(\bar{t},\bar{y}) \in [0,T)\times\mathbb{R}$ and $\varphi$ a test function in $C_b^1([0,T]\times\mathbb{R})$ s.t. $(\bar{t},\bar{y})$ is a strict global minimimum point of $w_*-\varphi$. Then, one can find a sequence $(t'_n,y'_n)$ in $[0,T)\times\mathbb{R}$, and a sequence $(h_n,\Delta_Y^n,M_n)$ such that:

$$\begin{aligned} (t'_n,y'_n) \to (\bar{t},\bar{y}), \quad (h_n,\Delta_Y^n,M_n) \to (0,0,\infty), \quad w^{h_n,\Delta_Y^n,M_n} \to w_*(\bar{t},\bar{y}), \\ (t'_n,y'_n) \text{ is a global minimum point of } w^{h_n,\Delta_Y^n,M_n}-\varphi. \end{aligned}$$

Denoting by $\zeta_n = (w^{h_n,\Delta_Y^n,M_n}-\varphi)(t'_n,y'_n)$, we have $w^{h_n,\Delta_Y^n,M_n} \geq \varphi+\zeta_n$. From the definition of the numerical scheme $\mathcal{S}^{h_n,\Delta_Y^n,M_n}$, and its monotonicity, we then get:

$$\begin{aligned} \zeta_n+\varphi(t'_n,y'_n) &= w^{h_n,\Delta_Y^n,M_n}(t'_n,y'_n) \\ &= \mathcal{S}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,w^{h_n,\Delta_Y^n,M_n}) \\ &\geq \mathcal{S}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi+\zeta_n) \;=\; \mathcal{S}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi)+\zeta_n \\ &= \max\Big[\mathcal{T}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi)\;,\;\tilde{\mathcal{M}}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi)\Big]+\zeta_n, \end{aligned}$$

which implies

$$\min\Big[\frac{\varphi(t'_n,y'_n)-\mathcal{T}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi)}{h_n}\;,\;\varphi(t'_n,y'_n)-\tilde{\mathcal{M}}^{h_n,\Delta_Y^n,M_n}(t'_n+h_n,y'_n,\varphi)\Big] \;\geq\; 0.$$

By the consistency properties (4.5)-(4.6), and by sending $n$ to infinity in the above inequality, we obtain the required viscosity supersolution property:

$$\min\Big[-\frac{\partial\varphi}{\partial t}(\bar{t},\bar{y})-\bar{y}c_P+\gamma\varrho\bar{y}^2-\mathcal{I}^a\varphi(\bar{t},\bar{y})-\mathcal{I}^b\varphi(\bar{t},\bar{y})\;,\;\varphi(\bar{t},\bar{y})-\tilde{\mathcal{M}}\varphi(\bar{t},\bar{y})\Big] \;\geq\; 0.$$

□

### 4.3 Numerical tests

In this section, we provide numerical results for the (reduced-form) value function and optimal policies in the martingale case and the trend information case, and a backtest on

| Parameter | Value |
|---|---|
| $\delta$ | 12.5 EUR/contract |
| $\varepsilon$ | 1.05 EUR/contract |
| $\varepsilon_0$ | 0 |
| $\lambda$ | $0.05s^{-1}$ |
| $\mu$ | $\exp(1/\bar{\mu})$ |
| $\bar{\mu}$ | 20 contracts |
| $\gamma$ | $2.5.10^{-5}$ |
| $T$ | 100 s |

(a) Market and risk parameters

| Parameter | Value |
|---|---|
| $N_Y$ | 100 |
| $N_T$ | 500 |
| $N_\varpi$ | 20 |

(b) Discretization parameters

Figure 2: Parameters for numerical results.

simulated data for the trend information case. Parameters for these numerical tests are shown in Figure 2.

This set of parameters are chosen to be consistent with calibration data on the front maturity for 3-months EURIBOR future, see for example [6]. Within this section, and for this set of parameters, we will denote by $w^h$ the value function and by $\alpha^\star$ the make/take strategy associated with the backward numerical scheme (4.3)-(4.4). Given a generic controlled process $Z$ and a control $\alpha \in \mathcal{A}$, we will denote $Z^\alpha$ the process controlled by $\alpha$. Unless specified otherwise, such processes will be supposed to start at zero: typically, we assume that the investor starts from zero cash and zero inventory at date $t = 0$ in the following numerical tests. Finally, we will write indifferently $w^h(t, y, c_P)$ or $w^h(t, y) := w^h(t, y, 0)$ to either stress or omit the dependence in $c_P$.

• **The martingale case**: in the martingale case, we performed the algorithm (4.3)-(4.4) with parameters shown above.

Figure 3 displayed the reduced-form value function $w^h$ on $[0, T] \times [-N_Y; N_Y]$. This result illustrates the linear bound (3.2) as noticed in proposition 3.2, and also the symmetry of $w^h$ as pointed out in (3.18). We also observe the monotonicity over $\mathbb{R}^+$ and $\mathbb{R}^-$ of the value function $w^h(t, .)$.

In Figure 4, we display the optimal make and take policies. The optimal take policy (on the left side) is represented as the volume to buy or sell with a market order, as a function of the time and inventory $(t, y) \in [0, T] \times [-N_Y; N_Y]$. We notice that a market order only occurs when the inventory becomes to large, and therefore, the take policy can be interpreted as a "stop-loss" constraint, i.e. an emergency rebalancing of the portfolio when the inventory risk is too large.

The optimal make policy is represented as the regime of limit orders posting as a function of the time and inventory $(t, y) \in [0, T] \times [-N_Y; N_Y]$. For sake of simplicity, we represented the sum of $\ell^a$ and $\ell^b$ on the map. The meaning of this surface is as follows: 0 means that there is no active limit orders on either sides, 2 means that there is active limit orders on both bid and ask sides, and 1 means that there is only one active limit order either on the bid or the ask side, depending on the sign of $y$ (if $y < 0$ only the bid side is active, and if $y > 0$ only the ask side is active). We notice that when close to maturity $T$, the

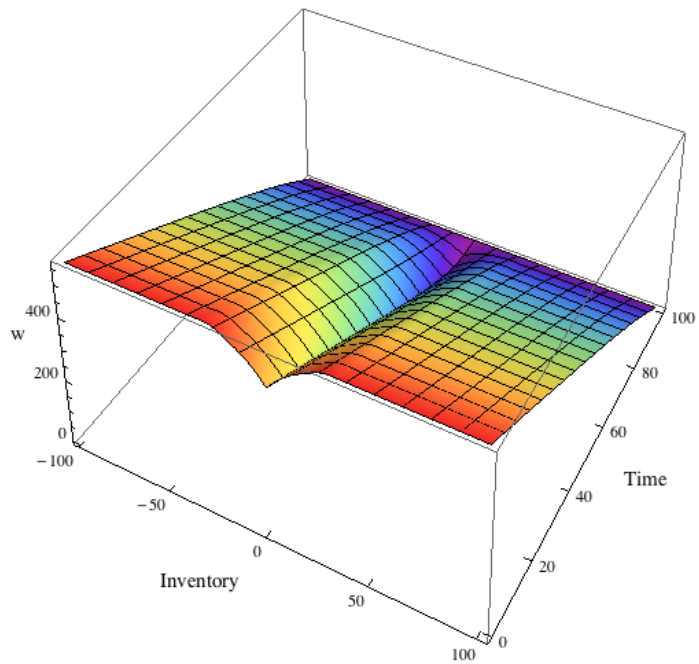
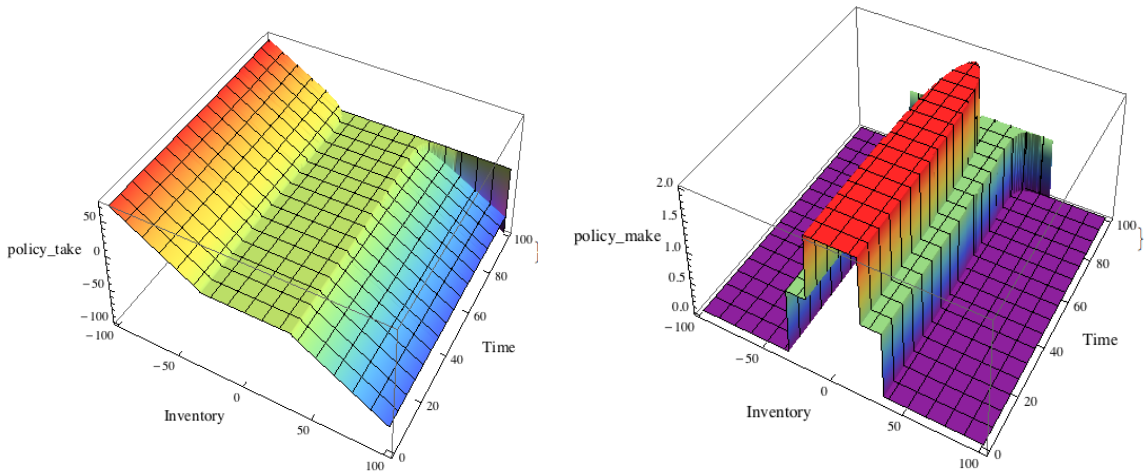



Figure 3: Reduced form value function $w^h$.




(a) Optimal take policy.

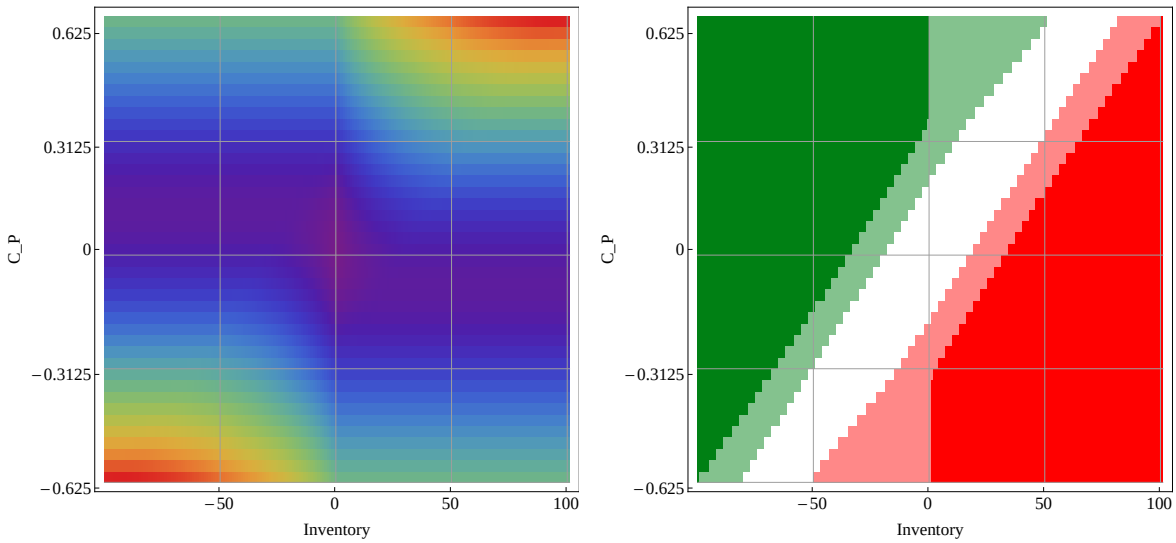
(b) Optimal make policy

Figure 4: Numerical results for the martingale case: representation of optimal make and take policies $\alpha^\star$.

optimal strategy tends to be more agressive, in the sense that it will seek to get rid of any positive or negative inventory, to match the terminal liquidation constraint. Moreover, we notice that close to date 0, the dependence in $t$ seems to be negligible, which indicates that a"stationary regime" may be attained for large $T$.

• **The trend information case:** in this case, we provide a backtest of the optimal strategy

on simulated data in addition to the plots of the value function $w^h$ and optimal policy $\alpha^\star$.




(a) Value function $w^h$ at date $t = 0$.

(b) Optimal policy $\alpha^\star$ at date $t = 0$.

Figure 5: Value function and optimal policy for the trend information case.

Figure 5 displays the value function and optimal policy at date $t = 0$, in the plane $(y, c_P)$. The value function has central symmetry properties as expected in (3.18), and should be read as follows: dark green zones represent situation where a market order to buy must be sent, light green means that a limit order is active only at bid, white means that limit orders are active on both sides, light red means that a limit order is active only at ask, and dark red means that a market order to sell must be sent. The value function also increases with $|c_P|$. This effect can be interpreted as the gain in performance due to the superior information on price trend $c_P$. The interpretation of this extra performance due to $c_P$ is that the optimal policy avoids part of the adverse selection risk by using this predictive information about price movements. Let us provide a qualitative example: assume that after the high frequency trader acquired a positive inventory, the adverse selection effect implies that price should go down; therefore, using the fact that in this case we should have $c_P < 0$, the optimal strategy will be either to cancel the bid limit order (light red zone) and keep ask limit order active, or depending on the value of $|c_P|$, send a market order to get rid of our positive inventory (dark red zone).

We performed a benchmarked backtest on simulated data and a performance analysis in this case. The benchmark strategy is made of constant controls (a.k.a *symmetric* or *constant* strategy):

$$\begin{aligned} \alpha^{benchmark} &:= (\alpha^{make\,,benchmark}, 0) \\ \alpha^{make\,,benchmark} &:= (1, 1) \end{aligned}$$

In order to make our simulated data backtest closer to the reality, we chosed to slightly deviate from the original price model, and use a varying price trend. We simulate a price

process model given by

$$\hat{P}_t = \hat{P}_0 + \delta(N_t^+ - N_t^-),$$

where $N^+$ and $N^-$ are the Euler scheme simulation of Cox processes of respective intensities $\pi^+$ and $\pi^-$ defined as follows

$$\begin{aligned} \pi^+ + \pi^- &\equiv K = \varrho/\delta^2 \\ d\pi_t^+ - d\pi_t^- := d\varpi_t &= -\theta\varpi_t dt + \sigma dB_t \end{aligned}$$

where $K > 0$, $\theta > 0$ and $\sigma > 0$ are positive constants, and $B$ is an independent Brownian motion. Note that we choosed the sum $\pi^+ + \pi^-$ to be the constant $K$, for simplicity sake: it means that, disregarding the direction of price variation, the mean number of price change per second is assumed to be constant $\mathbf{P}\left(|P_{t+h} - P_t| = \delta\right) = Kh + o(h)$, which provides an easy way to calibrate the parameter $K$ while reducing the dimension of the simulation. The interpretation of this simulation model is as follows: we add an exogenous risk factor $B$, which drives the price trend information $\varpi$ as an Ornstein-Uhlenbeck process. Notice that this supplementary risk factor $B$ is not taken into account in our optimization procedure and thus has a penalizing impact on the strategy's performance: therefore it does not spoil the backtest. This model choice for the process $(\varpi_t)$ is an convenient way to simulate the real-world situation, where the high-frequency trader continuously updates her predictive information about short-term price movements, based e.g. on the current state of the limit order book. Therefore, qualitatively speaking, our optimization procedure is consistent with this simulation model if we choose $\theta$ and $\sigma$ s.t. the variation of the (reduced-form) value function $w$ due to predictive information is very small compared to the variation of the value function due to other market events (e.g. an execution event). This assumption is consistent with HFT practice since the HF trader is able to adapt very quickly to a modification of this predictive information. Backtest parameters involved in this simulation are shown in Figure 6.

| Parameter | Value |
|---|---|
| $K$ | 1.0 |
| $\theta$ | 2 |
| $\sigma$ | 0.01 |
| $N_{MC}$ | 10000 |

Figure 6: Backtest parameters

Let us denote by $\hat{\vartheta}^a$ and $\hat{\vartheta}^b$ the Euler scheme simulation of the compound poisson processes $\vartheta^a$ and $\vartheta^b$, with dynamics (2.3). Therefore, for $\alpha \in \{\alpha^\star, \alpha^{benchmark}\}$, we were able to compute the Euler scheme simulation $\hat{X}^\alpha$ (resp. $\hat{Y}^\alpha$) of $X^\alpha$ (resp. $Y^\alpha$), starting at 0 at $t = 0$, by replacing $\vartheta^a$ (resp. $\vartheta^b$) by $\hat{\vartheta}^a$ (resp. $\hat{\vartheta}^b$) in equation (2.4) (resp. (2.5)).

We performed $N_{MC}$ simulation of the above processes. For each simulation $\omega \in [1...N_{MC}]$ and for $\alpha \in \{\alpha^\star, \alpha^{benchmark}\}$, we stored the following quantities: the terminal wealth after terminal liquidation $\hat{V}_T^\alpha(\omega) := L(\hat{X}^\alpha(\omega), \hat{Y}^\alpha(\omega), \hat{P}(\omega))$, called "performance"

in what follows ; the total executed volume $\hat{Q}^{total,\alpha}(\omega) := \sum_{[0,T]} |\hat{Y}_t^\alpha(\omega) - \hat{Y}_{t-}^\alpha(\omega)|$ ; and the volume executed at market $\hat{Q}^{market,\alpha}(\omega) := \sum_{[0,T]} |\xi_n(\omega)^\alpha|$. Finally, we denote by $m(.)$ the empirical mean, by $\Sigma(.)$ the empirical standard deviation, by skew(.) the empirical skewness, and by kurt(.) the empirical kurtosis, taken over $\omega \in [1...N_{MC}]$.

| Quantity | Definition | $\alpha^\star$ | $\alpha^{benchmark}$ |
|---|---|---|---|
| Info ratio over $T$ | $m(\hat{V}_{\dot{T}})/\sigma(\hat{V}_{\dot{T}})$ | 0.238 | 0.104 |
| Profit per trade | $m(\hat{V}_{\dot{T}})/m(\hat{Q}^{total,\cdot})$ | 1.37 | 3.86 |
| Risk per trade | $\sigma(\hat{V}_{\dot{T}})/m(\hat{Q}^{total,\cdot})$ | 5.73 | 37.21 |
| Mean performance | $m(\hat{V}_{\dot{T}})$ | 376.08 | 773.15 |
| Standard deviation of perf | $\sigma(\hat{V}_{\dot{T}})$ | 1574.97 | 7462.96 |
| Skew of perf | skew$(\hat{V}_{\dot{T}})$ | 0.027 | -0.0468 |
| Kurtosis of perf | kurt$(\hat{V}_{\dot{T}})$ | 3.21 | 7.48 |
| Mean total executed volume | $m(\hat{Q}^{total,\cdot})$ | 274.77 | 200.27 |
| Mean at market volume | $m(\hat{Q}^{market,\cdot})$ | 101.75 | 0. |
| Ratio market over total exec | $m(\hat{Q}^{market,\cdot})/m(\hat{Q}^{total,\cdot})$ | 0.37 | 0. |

Figure 7: Synthetis table for backtest. Categories are, from top to bottom: relative performance metrics, period-adjusted performance metrics, absolute performance metrics and absolute activity metrics.

Figure 7 displayed a synthesis of descriptive statistics for this backtest. We first notice that the information ratio over $T$ of $\alpha^\star$ is more than twice that of $\alpha^{benchmark}$. Second, the per trade metrics can be compared to the half-spread $\frac{\delta}{2} = 6.25$ EUR/contract, and we see that although the mean profit per trade is smaller for the optimal strategy, the risk associated to each trade is dramatically reduced compared to the benchmark. This

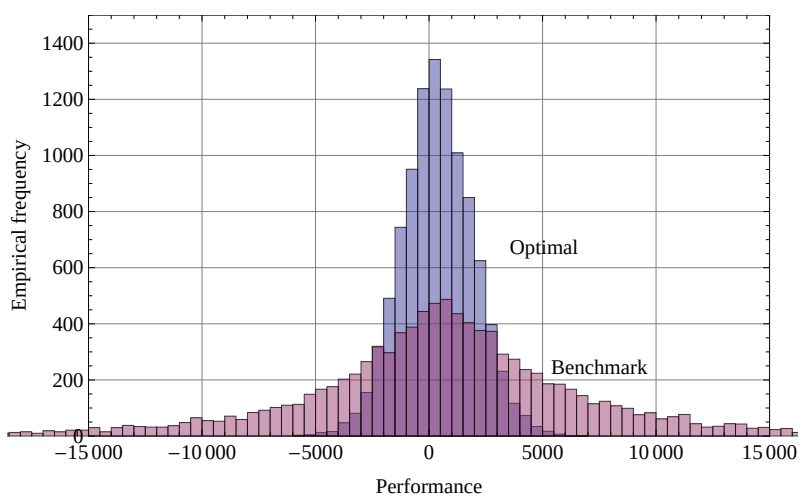



Figure 8: Empirical distribution of performance $\hat{V}_{\dot{T}}$. The graph shows the number of occurences for each bin on $N_{MC} = 10000$ simulations.

is confirmed by the empirical distribution of performance, also shown in Figure 8, where

the dark blue represents the performance distribution of the optimal strategy and the light purple represents the performance distribution of the benchmark strategy. We see that not only benchmark has higher standard deviation, but also higher excess kurtosis and heavy tails: this is due to the fact that inventory can be very large for the benchmark strategy, and therefore it bears a non-negligible market risk (or inventory risk). Finally, we see that about 37% of the trades are done with market orders, which indicates that this feature of the strategy is relevant when exposed to adverse selection risk (the risk that the mid-price moves unfavourably after a limit order execution).

| $\gamma$ | $\sigma(\hat{V}_{\dot{T}})$ | $m(\hat{V}_{\dot{T}})$ |
|---|---|---|
| 6.67.10-04 | 13.36 | 0.09 |
| 4.44.10-04 | 351.16 | 20.98 |
| 2.96.10-04 | 495.43 | 30.97 |
| 1.98.10-04 | 649.11 | 39.28 |
| 1.32.10-04 | 849.05 | 106.14 |
| 8.78.10-05 | 1048.73 | 177.27 |
| 5.85.10-05 | 1264.10 | 253.50 |
| 3.90.10-05 | 1428.10 | 309.95 |
| 2.60.10-05 | 1546.30 | 351.86 |
| 1.73.10-05 | 1635.61 | 368.15 |
| 1.16.10-05 | 1639.65 | 332.51 |

Figure 9: Varying risk aversion parameter $\gamma$: data.

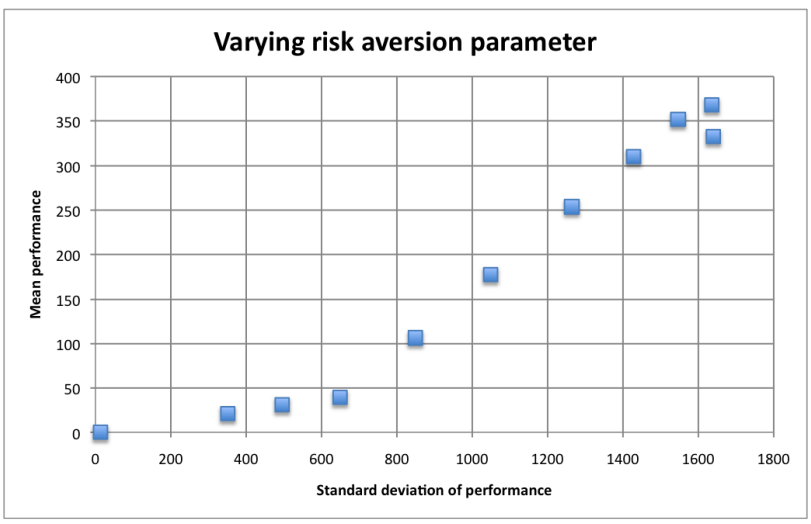



Figure 10: Varying risk aversion parameter $\gamma$: plot.

Our last numerical test is devoted to displaying the influence of the risk aversion parameter $\gamma$. All other parameters remaining the same, we tested several values of $\gamma$ (as indicated in Figure 9), and characterized the performance of the corresponding strategy by the pair $(\sigma(\hat{V}_{\dot{T}}), m(\hat{V}_{\dot{T}}))$, which gives the *efficient frontier* plot displayed in Figure 10. As

expected, a reduction of $\gamma$ increases the standard deviation of the strategy: this is due to the fact that a small $\gamma$ allows for large open position i.e. large inventory, and therefore the market risk is greater. For small $\gamma$, performance is also better since the investor can sustain large inventories, and therefore is less impatient to get rid of it: in particular, the proportion of volume executed at market is increasing in $\gamma$. We see that in real trading conditions, $\gamma$ must be chosen as small as possible, i.e. the value allowing the greatest performance, but maintaining the market risk sustainable.

# 5 Best execution problem and overtrading risk

In this section, we apply our market model framework to a best execution problem. The trading objective of the investor is to liquidate $Y_0 > 0$ assets over the finite time interval $[0, T]$. She is not allowed to purchase stock during the liquidation period, and may only buy back the asset in case of short position. In this context, the investor posts continuously a limit sell order (with a volume much larger that the required quantity $Y_0$) at the best ask price, and also runs market (sell) orders strategy until she reaches either a negative inventory or the terminal date. By doing so, she hopes to trade as much as possible at the ask price, and therefore avoiding to *cross the spread*.

Mathematically, this means that the investor uses a subset $\mathcal{A}_\ell$ of strategies $\alpha = (\alpha^{make} = (L^a, L^b), \alpha^{take})$ in $\mathcal{A}$ such that:

$$\begin{aligned}
(L^a_t, L^b_t) &= \begin{cases} (1,0) & \text{for } t < \tau, \\ (0,0) & \text{for } t \geq \tau \end{cases} \\
\alpha^{take} &= (\tau_n, \zeta_n)_n \cup (\tau, -Y_\tau), \quad \text{with } \tau_n < \tau,\ \zeta_n < 0,
\end{aligned}$$

where $\tau = \inf\{t \geq 0 : Y_t \leq 0\} \wedge T$. The value function associated to this liquidation problem is then defined by

$$v_\ell(t,x,y,p) = \sup_{\alpha \in \mathcal{A}_\ell} \mathbf{E}_{t,x,y,p}\Big[L(X_T, Y_T, P_T) - \gamma \int_t^T Y_s^2 \varrho(P_s) ds\Big], \tag{5.1}$$

for $(t,x,y,p) \in [0,T] \times \mathbb{R}^2 \times \mathbb{P}$. With the notation in (3.6), the operator corresponding to the limit order in $\mathcal{A}_\ell$ is given by $\mathcal{L}^{1,0} = \mathcal{P} + \Gamma^a$, while the impulse operator associated to the market order in $\mathcal{A}_\ell$ is defined by:

$$\mathcal{M}_\ell\, \varphi(t,x,y,p) = \sup_{e \in [-|y|,0]} \varphi\big(t, x - ep - |e|(\frac{\delta}{2} + \varepsilon) - \varepsilon_0, y + e, p\big)$$

The dynamic programming equation associated to (5.1) takes the form:

$$\min\big[ -\frac{\partial v_\ell}{\partial t} - \mathcal{P} v_\ell - \Gamma^a v_\ell \ + \ \gamma g \ , \ v_\ell - \mathcal{M}_\ell v_\ell \big] = 0, \text{ on } [0,T) \times \mathbb{R} \times (0,\infty) \times \mathbb{P},$$

together with the terminal and boundary conditions:

$$v_\ell = L, \quad \text{on } \big(\{T\} \times \mathbb{R} \times \mathbb{R} \times \mathbb{P}\big) \cup \big([0,T) \times \mathbb{R} \times \mathbb{R}_- \times \mathbb{P}\big).$$

The above boundary condition for nonpositive inventory is related to the *overtrading risk*, which is the risk that the investor sold too much assets via the (oversized) limit order at the best ask price. This risk occurs typically in execution problems on pro-rata limit order book, see [6].

Again, in the Lévy case (3.11), the value function $v_\ell$ is reduced into:

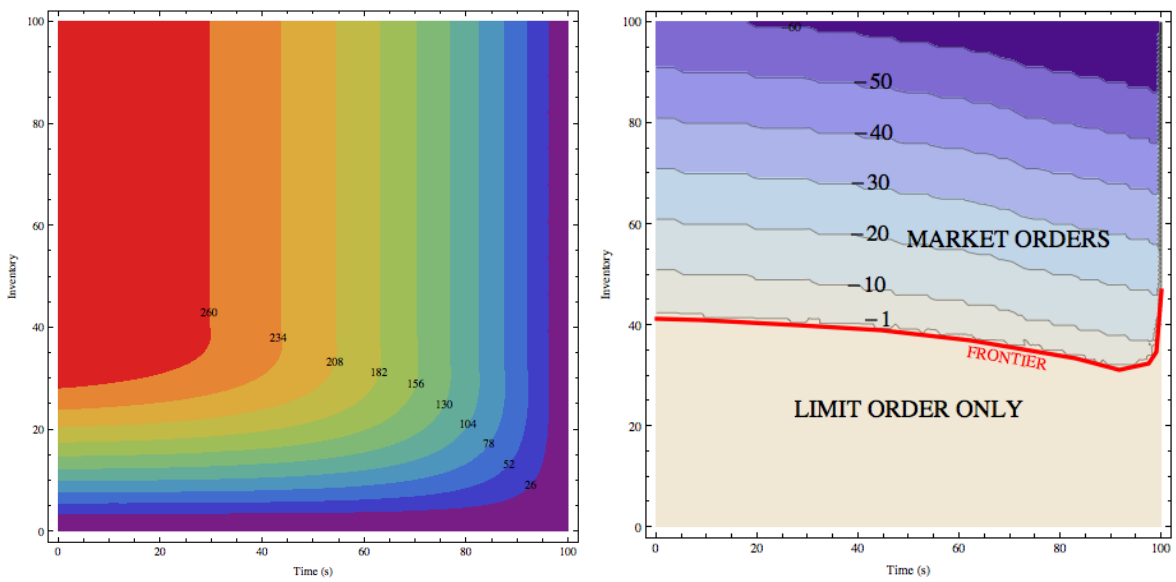
$$v_\ell(t,x,y,p) \quad = \quad L(x,y,p) + w_\ell(t,y),$$

where $w_\ell$ is solution to the integro-variational inequality:

$$\begin{aligned} &\min\Big\{ -\frac{\partial w_\ell}{\partial t} - yc_P + \gamma\varrho y^2 \\ &-\lambda^a \int_0^\infty \Big[ w_\ell(t,y-z) - w_\ell(t,y) + z\frac{\delta}{2} + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\Big]\mu^a(dz)\ ; \\ &w_\ell(t,y) - \sup_{e\in[-y,0]} \Big[ w_\ell(t,y+e) - (\frac{\delta}{2}+\varepsilon)(|y+e|+|e|-|y|) - \varepsilon_0\Big]\Big\} \quad = \quad 0, \end{aligned}$$

for $(t,y) \in [0,T) \times (0,\infty)$, together with the terminal and boundary conditions:

$$w_\ell(t,y) \quad = \quad 0, \quad \forall (t,y) \in \big(\{T\}\times\mathbb{R}\big) \cup \big([0,T)\times\mathbb{R}_-\big).$$




(a) Value Function $w_\ell$ (b) Optimal policy (take)

Figure 11: Numerical results for the simple liquidation problem (for $c_P = 0$). On the left side, level lines are indicated for the value function $w_\ell$. On the right side, numbers indicated on the figure represent the quantity to sell in the optimal market order control.

The associated numerical scheme reads now as follows:

$$\begin{aligned} w_\ell^h(t_N,y) &= 0, \quad y\in\mathbb{R}, \\ w_\ell^h(t_k,y) &= 0, \quad k=0,\ldots,N-1,\ y\le 0, \\ w_\ell^h(t_k,y) &= \max\Big[\mathcal{T}_\ell^{h,\Delta_Y,M}(t,y,\varphi)\ ;\ \mathcal{M}_\ell^{h,\Delta_Y,M}(t,y,\varphi)\Big],\ k=0,\ldots,N-1\,,\ y\in\mathbb{Y}_M^+, \end{aligned}$$

where $\mathbb{Y}_M^+ = \mathbb{Y}_M \cap \mathbb{R}_+$,

$$\begin{aligned}\mathcal{T}_\ell^{h,\Delta_Y,M}(t,y,\varphi) &= \varphi(t,y) - h\gamma\varrho y^2 + hyc_P \\ &\quad + \lambda^a h\Big(\int_0^\infty \big[\varphi(t,\mathrm{Proj}_M(y-z)) - \varphi(t,y)\big]\hat{\mu}^a(dz) \\ &\qquad + \int_0^\infty \big[\frac{\delta}{2}z + (\frac{\delta}{2}+\varepsilon)(|y|-|y-z|)\big]\mu^a(dz)\Big)\end{aligned}$$

and

$$\begin{aligned}&\mathcal{M}_\ell^{h,\Delta_Y,M}(t,y,\varphi) \\ =\ &\sup_{e\in\mathbb{Y}_M\cap[-y,0]} \big[\varphi(t,\mathrm{Proj}_M(y+e)) - (\frac{\delta}{2}+\varepsilon)(|y+e|+|e|-|y|) - \varepsilon_0\big].\end{aligned}$$

In this case, the optimal policy shown in Figure 11 is simple to describe. The state space is delimited in two zones: when the inventory is small, the HFT must wait for her limit sell order to be executed; and when the inventory is large, the HFT must send a market sell order to avoid the market risk related to holding a large position.

The frontier between the two zones (indicated in bold red in Figure 11) can be interpreted as an *optimal trading curve*, a concept that is extensively documented (see e.g. [7]) in the optimal execution literature. The optimal trading curve is the inventory that the investor should hold, seen as a function of time, in order to minimize overall trading costs. Therefore, in the typical setting, the execution strategy consists in trading via market orders to get as close as possible to the optimal trading curve. Similarly, in our case, we can see on Figure 11 that the optimal strategy will behave similarly for large inventories (i.e. when above the trading curve): indeed, we observe that the quantities to sell are such that the market orders strategy would keep the inventory close to the optimal trading curve, if no limit orders were allowed. Now, in our case, we observe two specific features of the optimal strategy: 1) the optimal trading curve does not reach 0 at maturity, and therefore the HFT has to get rid of her inventory at market at final date to match the constraint $Y_T = 0$. This is due to the fact that a supplemental gain is always achievable when the limit order is executed. Therefore, this features leads to an execution strategy where the final trade is bigger than intermediary trades; 2) below the optimal trading curve, i.e. in the region where the HFT trades via limit orders only, the sell limit order is always active, and can lead to an execution. Therefore, the inventory is always below the optimal trading curve, and the distance between the current inventory and the optimal trading curve equals the volume executed via limit orders. This differs from classic pattern-based best execution strategies, for example the U-shaped execution strategy that consists in trading a large quantity at the beginning and at the end of the liquidation, and trade regularly small quantities in between. Indeed, the optimal strategy does not provide a fixed pattern for every execution, but provide the optimal action to take given the observation of the inventory that is still to be sold and the market characteristics as e.g. the mean traded volume at ask per second $\lambda^a\bar{\mu}^a$, or trades volume distributions at ask $\mu^a$.

Finally, let us notice that this strategy can be interpreted as a convenient way to avoid the cost of crossing the spread during the liquidation of a portfolio, but we did not take into

account the impact of the market order on the transaction price. In the case of a pro-rata microstructure, available volumes offered at best prices are usually about 200 times larger than the mean volume of market orders (see [6]), and therefore it is consistent to consider that there is no impact on the price for our market orders. Yet, the model can easily be modified by adding an impact component in the obstacle operator $\mathcal{M}_\ell$ to take care of this effect. We also did not model the possibility that the intensities $\lambda^a$ and $\lambda^b$ of execution processes may vary, and postpone this investigation for future research.

## References


[1] Aikin, S. (2006): "Trading STIR Futures : An Introduction to Short-Term Interest Rate Futures", Harriman House Publishing, ISBN: 978-1897597811

[2] Avellaneda M. and S. Stoikov (2008): "High frequency trading in a limit order book", *Quantitative Finance*, 8(3), 217-224.

[3] Barles G. and P. Souganidis (1991): "Convergence of approximation schemes for fully nonlinear second order equations", *Asymptotic Analysis*, 4, 271-283.

[4] Cartea A, Jaimungal S. and J. Ricci (2011): "Buy low sell high: a high frequency trading perspective", preprint

[5] Cont R. and A. de Larrard (2010): "Price Dynamics in a Markovian Limit Order Market", preprint

[6] Field J. and J. Large (2008): "Pro-Rata Matching and One-Tick Futures Markets", preprint

[7] Guéant O., Fernandez Tapia J. and C.-A. Lehalle (2011): "Dealing with inventory risk", preprint.

[8] Gould M.D., Porter M.A, Williams S., McDonald M., Fenn D.J. and S.D. Howison (2010): "The limit order book: a survey", preprint.

[9] Guilbaud F. and H. Pham (2011): "Optimal high frequency trading with limit and market orders", preprint.

[10] Karel Janecek, Martin Kabrhel (2007): "Matching Algorithms of International Exchanges", working paper.

[11] Kühn C. and M. Stroh (2010): "Optimal portfolios of a small investor in a limit order market: a shadow price approach", *Mathematics and Financial Economics*, 3(2), 45-72.

[12] Veraart L.A.M. (2011): "Optimal Investment in the Foreign Exchange Market with Proportional Transaction Costs", *Quantitative Finance*, 11(4): 631-640. 2011.