HAL
archives-ouvertes.fr

Optimal posting price of limit orders: learning by trading

Sophie Laruelle, Charles-Albert Lehalle, Gilles Pagès

▶ To cite this version:

Optimal posting price of limit orders: learning by trading

Sophie Laruelle [*]   Charles-Albert Lehalle [†]   Gilles Pagès [‡]

Abstract

Considering that a trader or a trading algorithm interacting with markets during continuous auctions can be modeled by an iterating procedure adjusting the price at which he posts orders at a given rhythm, this paper proposes a procedure minimizing his costs. We prove the *a.s.* convergence of the algorithm under assumptions on the cost function and give some practical criteria on model parameters to ensure that the conditions to use the algorithm are fulfilled (using notably the co-monotony principle). We illustrate our results with numerical experiments on both simulated data and using a financial market dataset.

Keywords *Stochastic approximation, order book, limit order, market impact, statistical learning, high-frequency optimal liquidation, compound Poisson process, co-monotony principle.*

2010 AMS classification: 62L20, secondary: 62P05, 60G55, 65C05.

1 Introduction

In recent years, with the growth of electronic trading, most of the transactions in the markets occur in *Limit Order Books.* During the matching of electronic orders, traders send orders of two kinds to the market: passive (*i.e.* limit or patient orders) which will not give birth to a trade but will stay in the order book (sell orders at a higher price than the *higher bid price* or buy orders at a lower price than the *lower ask price* are passive orders) and aggressive orders (*i.e.* market or impatient orders) which will generate a trade (sell orders at a lower price than the higher passive buy price or buy orders at a higher price than the lowest passive price). When a trader has to buy or sell a large number of shares, he cannot just send his large order at once (because it would consume all of the available liquidity in the order book, impacting the price at his disadvantage); he has first to schedule his trading rate to strike balance between the market risk and the market impact cost of being too aggressive (too many orders exhaust the order book and makes the price move). Several theoretical frameworks have been proposed for optimal scheduling of large orders (see [3], [7], [24], [2]). Once this optimal trading rate is known, the trader has to send smaller orders in the (electronic) limit order book by alternating limit (*i.e.* patient or passive) orders and market (*i.e.* urgent or aggressive) orders. The optimal mix of limit and market orders for a trader has not been investigated in the quantitative literature even if it has been studied from a global economic efficiency viewpoint (see for instance [9]). It has not either

---

[*]Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, UPMC, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: `sophie.laruelle@upmc.fr`

[†]Head of Quantitative Research, Crédit Agricole Cheuvreux, CALYON group ; 9 quai Paul Doumer, 92920 Paris La Défense. E-mail: `clehalle@cheuvreux.com`

[‡]Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, UPMC, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: `gilles.pages@upmc.fr`

1

been investigated from the viewpoint of one trader trying to optimize his own interactions with other market participants. One of the difficulties from the trader prospective is that transactions obtained by inserting a passive order in a limit order book is a functional of its distance to the mid-price, giving birth to a large number of possible tactics in terms of reassessment of the price of such orders.

In this paper, we study the optimal distance to submit passive orders in a limit order book, without needing a model of the limit order book dynamics (see *e.g.* [1] or [11] for such models of limit order books).

Optimal submission strategies have been studied in the microstructure literature using utility framework and optimal control (see [4], [10], [5], [12] and [13]). The authors of such papers consider an agent who plays the role of a market maker, *i.e.* he provides liquidity on the exchange by quoting bid and ask prices at which he is willing to buy and sell a specific quantity of assets. Strategies for bid and ask orders are derived by maximizing his utility function.

Our approach is different: we consider an agent who wants to buy (or sell) during a short period $[0, T]$ a quantity Q_T of traded assets and we look for the optimal distance where he has to post his order to minimize the execution cost.

We are typically at a smaller time scale than in usual optimal liquidation frameworks. In fact order posting strategies derived from the viewpoint presented below can be "plugged" into any larger scale strategy. We are modeling the market impact of an aggressive order using a penalization function $\kappa \cdot \Phi(Q)$, where Q is the size of the market order.

If a stochastic algorithm approach has been already proposed in [20] by the authors for optimal spatial split of orders across different Dark Pools, here the purpose is not to control fractions of the size of orders, but to adjust successive posting prices to converge to an optimal price. Qualitatively, this framework can be used as soon as a trader wants to trade a given quantity Q_T over a given time interval $[0, T]$ with no firm constraint on its trading rate between 0 and T. It is typically the case for small *Implementation Shortfall* benchmarked orders. The trader can post his order very close to the *"fair price"* $(S_t)_{t \in [0,T]}$ (which can be seen as the fundamental price, the mid price of the available trading venues or any other reference price). In this case he will be exposed to the risk to trade too fast at a "bad price" and being adversely selected. Conversely he can post it far away from the fair price; in that case he will be exposed to never obtain a transaction for the whole quantity Q_T, but only for a part of it (say the positive part of $Q_T - N_T$, where N_T is the quantity that the trading flow allowed him to trade). He will then have to consume aggressively liquidity with the remaining quantity, disturbing the market and paying not only the current market price S_T, but also a market impact (say $S_T \Phi(Q_T - N_T)$ where Φ is a market impact penalization function).

The approach presented here follows the mechanism of a "learning trader". He will try to guess the optimal posting distance to the fair price achieving the balance between being too demanding in price and too impatient, by successive trials, errors and corrections. The optimal recursive procedure derived from our framework gives the best price adjustment to apply to an order on given stopping time (reassessment dates) given the observed past on the market. We provide proofs of the convergence of the procedure and of its optimality.

To this end, we model the execution process of orders by a Poisson process $(N_t^{(\delta)})_{0 \leq t \leq T}$ which intensity $\Lambda_T(\delta, S)$ depends on the fair price $(S_t)_{t \geq 0}$ and the distance of order submission δ. The execution cost results from the sum of the price of the executed quantity and a penalization function depending on the remaining quantity to be executed at the end of the period $[0, T]$. This penalty $\kappa \cdot \Phi(Q)$ models the over-cost induced by crossing the spread and the resulting market impact of this execution. The aim is to find the optimal distance $\delta^* \in [0, \delta_{\max}]$, where $\delta_{\max}$ is the depth of the limit order book, which minimizes the execution cost. In practice, the prices are constrained to be on a "tick size grid" (see [25]), instead of being on the real line. We will follow the approach of papers on market making [14, 4, 10] assuming the the tick size is small enough to not change the dynamics

of the price and of the bid-ask spread [22]. Nevertheless, a "rounding effect" would not change the nature of our results, since any projection (*e.g.* to nearest-neighbor) is a non-decreasing transform and the co-monotony principle still holds (by slightly adapting the proofs). This leads to an optimization problem under constraints which we solve by using a recursive stochastic procedure with projection (this particular class of algorithm is studied in [18] and [19]). We prove the *a.s.* convergence of the constrained algorithm under additional assumptions on the execution cost function. From a practical point of view, it is not easy to check the conditions on the cost function. So we give criteria on the model parameters which ensure the viability of the algorithm which relies on the co-monotony principle assumed to be satisfied by the "fair price" process $(S_t)_{t \in [0,T]}$. This principle will be detailed in the Appendix Section B. We conclude this paper by some numerical experiments with simulated and real data. We consider the Poisson intensity presented in [4] and use a Brownian motion to model the fair price dynamics. We plot the cost function and its derivative and show the convergence of the algorithm to its target δ^*.

The paper is organized as follows: in Section 2, we first propose a model for the execution process of posted orders, then we define a penalized execution cost function (including the market impact at the terminal execution date). Then we devise the stochastic recursive procedure under constraint to solve the resulting optimization problem in terms of optimal posting distance on the limit order book. We state the main convergence result and provide operating criteria that ensure this convergence, based on a co-monotony principle for one dimensional diffusions. Section 3 establishes the representations as expectations of the cost function and its derivatives which allow to define the mean function of the algorithm. Section 4 presents the convergence criteria (which ensure that the optimization is well-posed) derived from the principle of co-monotony established in Section B of the appendix. Finally Section 5 illustrates with numerical experiments the convergence of the recursive procedure towards its target.

Notations. $\bullet$ $(x)_+ = \max\{x, 0\}$ denotes the positive part of x, $\lfloor x \rfloor = \max\{k \in \mathbb{N} : k \leq x\}$ $[\![0, x]\!] :=$ $\{y \in [0,1]^d : 0 \leq y \leq x\} = \prod_{i=1}^d [0, x_i]$, $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$.

$\bullet \langle \cdot \mid \cdot \rangle$ denotes the canonical inner product on $\mathbb{R}^d$.

$\bullet \overset{(\mathbb{R}^d)}{\Longrightarrow}$ denotes the weak convergence on $\mathbb{R}^d$ and $\overset{\mathcal{L}}{\longrightarrow}$ denotes the convergence in distribution.

$\bullet \mathcal{C}([0,T], A) := \{f : [0,T] \to A \text{ continuous}\}$ (equipped with the supnorm topology) and $\mathbb{D}([0,T], A) :=$ $\{f : [0,T] \to A \text{ càdlàg}\}$ (equipped with the Skorokhod topology when necessary see [15]) where càdlàg means right continuous with left limits and $A = \mathbb{R}^q, \mathbb{R}^q_+$, etc. They are equipped with the standard Borel σ-field $\sigma(\alpha \mapsto \alpha(t), t \in [0,T])$.

$\bullet \mathbb{P}\text{-esssup} f = \inf\{a \in \mathbb{R} : \mathbb{P}(\{x : f(x) > a\}) = 0\}$, $\|\alpha\|_\infty = \sup_{t \in [0,T]} |\alpha(t)|$, $\alpha \in \mathbb{D}([0,T], \mathbb{R})$ and f'_ℓ denotes the left derivative of f.

2 Design of the execution procedure and main results

2.1 Modeling and design of the algorithm

We focus our work on the problem of optimal trading with limit orders on one security without needing to model the limit order book dynamics. We only model the execution flow which reaches the price where the limit order is posted with a general price dynamics $(S_t)_{t \in [0,T]}$ since we intend to use real data. However there will be two frameworks for the price dynamics: either $(S_t)_{t \in [0,T]}$ is a process bounded by a constant L (which is obviously an unusual assumption but not unrealistic on a short time scale see Section 2.2.2) or $(S_t)_{t \in [0,T]}$ is ruled by a Brownian diffusion model (see Section 2.2.1).

We consider on a short period T, say a dozen of seconds, a Poisson process modeling the execution of posted passive buy orders on the market

$$\left(N_t^{(\delta)}\right)_{0\leq t\leq T} \quad \text{with intensity} \quad \Lambda_T(\delta, S) := \int_0^T \lambda(S_t - (S_0 - \delta))dt \qquad (2.1)$$

where $0 \leq \delta \leq \delta_{\max}$ with $\delta_{\max}$ the depth of the order book, $\delta_{\max} \in (0, S_0)$, and $(S_t)_{t\geq 0}$ is a stochastic process modeling the dynamics of the "fair price" of a security stock (from an economic point of view). In practice one may consider that S_t represents the best opposite price at time t.

One way to build $N^{(\delta)}$ is to set

$$N_t^{(\delta)} = \widetilde{N}_{\int_0^T \lambda(S_t - (S_0 - \delta))dt}$$

where $\widetilde{N}$ is a Poisson process with intensity 1 independent of the price $(S_t)_{t\in[0,T]}$.

This representation underlines the fact that for one given trajectory of the price S_t, the intensity of the Point process N is decreasing with δ: in fact the above representation for $N^{(\delta)}$ is even pathwise consistent in the sense that if $0 < \delta < \delta'$ then

$$\mathbb{P}\text{-}a.s. \quad \left(\forall t \in [0, T], \quad N_t^{(\delta)} \leq N_t^{(\delta')}\right).$$

The natural question how to account for the actually real possibility of simultaneously placing limit orders at different prices may be much harder to handle and is not studied in this paper. In particular, due to interacting impact features, it would need a more sophisticated approach then simply considering $(N^{\delta(k)})_{1\leq k\leq K}$, processes as above with $\delta(1) < \delta(2) < \ldots < \delta(K)$ (with the same N).

We assume that the function λ is defined on $[-S_0, +\infty)$ as a finite non-increasing convex function. Its specification will rely on parametric or non parametric statistical estimation based on former obtained transactions (see Figure 1 below and Section 5). At time $t = 0$, buy orders are posted in the limit order book at price $S_0 - \delta$. Between t and $t + \Delta t$, the probability for such an order to be executed is $\lambda(S_t - (S_0 - \delta))\Delta t$ where $S_t - (S_0 - \delta)$ is the distance to the current fair price of our posted order at time t. The further the order is at time t, the lower is the probability for this order to be executed since λ is decreasing on $[-S_0, +\infty)$. Empirical tests strongly confirm this kind of relationship with a convex function λ (even close to an exponential shape, see Figure 1). Over the period $[0, T]$, we aim
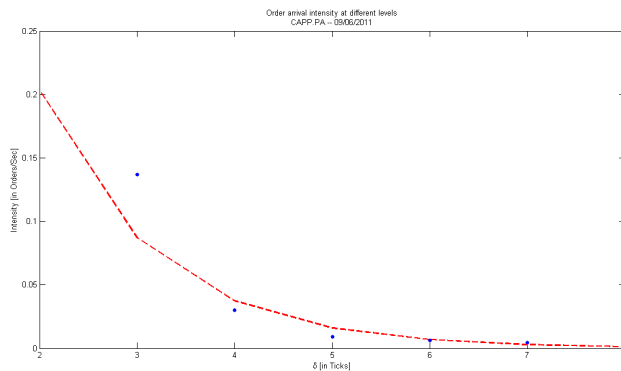
Figure 1: Empirical probabilities of execution (blue stars) and its fit with an exponential law (red dotted line) with respect to the distance to the "fair price".

at executing a portfolio of size $Q_T \in \mathbb{N}$ invested in the asset S. The execution cost for a distance δ is $\mathbb{E}\left[(S_0 - \delta)\left(Q_T \wedge N_T^{(\delta)}\right)\right]$. We add to this execution cost a penalization depending on the remaining

quantity to be executed, namely at the end of the period T, we want to have Q_T assets in the portfolio, so we buy the remaining quantity $\left(Q_T - N_T^{(\delta)}\right)_+$ at price S_T.

At this stage, we introduce a *market impact penalization function* $\Phi : \mathbb{R} \mapsto \mathbb{R}_+$, non-decreasing and convex, with $\Phi(0) = 0$ to model the additional cost of the execution of the remaining quantity (including the market impact). Then the resulting cost of execution on a period $[0, T]$ reads

$$C(\delta) := \mathbb{E}\left[(S_0 - \delta)\left(Q_T \wedge N_T^{(\delta)}\right) + \kappa\, S_T\, \Phi\left(\left(Q_T - N_T^{(\delta)}\right)_+\right)\right] \tag{2.2}$$

where $\kappa > 0$ is a free tuning parameter. When $\Phi(Q) = 1$, we just consider that we buy the remaining quantity at the end price S_T. Introducing a market impact penalization function $\Phi(x) = (1 + \eta(x))x$, where $\eta \geq 0$, $\eta \not\equiv 0$, models the market impact induced by the execution of $\left(Q_T - N_T^{(\delta)}\right)_+$ at time T whereas we neglect the market impact of the execution process via limit orders over $[0, T)$. Our aim is then to minimize this cost by choosing the distance to post at, namely to solve the following optimization problem

$$\min_{0 \leq \delta \leq \delta_{\max}} C(\delta). \tag{2.3}$$

Our strategy to solve numerically (2.3) using a large enough dataset is to take advantage of the representation of C and its first two derivatives as expectations to devise a recursive stochastic algorithm, namely a stochastic gradient procedure, to find the minimum of the (penalized) cost function (see below). Furthermore we will show that under natural assumptions on the quantity Q_T to be executed and on the parameter κ, the function C is twice differentiable, *strictly convex on* $[0, \delta_{\max}]$ *with* $C'(0) < 0$. Consequently,

$$\operatorname{argmin}_{\delta \in [0, \delta_{\max}]} C(\delta) = \{\delta^*\}, \quad \delta^* \in (0, \delta_{\max}]$$

and

$$\delta^* = \delta_{\max} \quad \text{iff} \quad C \text{ is non-increasing on } [0, \delta_{\max}].$$

Criteria involving κ and based on both the risky asset S and the trading process especially the execution intensity λ, are established further on in Proposition 4.1 and Proposition 4.2. We specify representations as expectations of the function C and its derivatives C' and C''. In particular we will exhibit a Borel functional

$$H : [0, \delta_{\max}] \times \mathbb{D}\left([0, T], \mathbb{R}\right) \longrightarrow \mathbb{R}$$

such that

$$\forall \delta \in [0, \delta_{\max}], \quad C'(\delta) = \mathbb{E}\left[H\left(\delta, (S_t)_{t \in [0,T]}\right)\right].$$

The functional H has an explicit form given in Proposition 3.2, Equations (3.14) or (3.16), involving integrals over $[0, T]$ of the intensity $\lambda(S_t - S_0 + \delta)$ of the Poisson process $(N_t^{(\delta)})_{t \in [0,T]}$. In particular, any quantity $H\left(\delta, (S_t)_{t \in [0,T]}\right)$ can be simulated, up to a natural time discretization, either from a true dataset (of past executed orders) or from the stepwise constant discretization scheme of a formerly calibrated diffusion process modeling $(S_t)_{t \in [0,T]}$ (see below). This will lead us to replace for practical implementations the continuous time process $(S_t)_{t \in [0,T]}$ over $[0, T]$, either by a discrete time sample, *i.e.* a finite dimensional $\mathbb{R}^{m+1}$-valued random vector $(S_{t_i})_{0 \leq i \leq m}$ (where $t_0 = 0$ and $t_m = T$) or by a time discretization scheme with step $\frac{T}{m}$ (typically the Euler scheme when $(S_t)_{t \in [0,T]}$ is a diffusion).

A THEORETICAL STOCHASTIC LEARNING PROCEDURE: Based on this representation (3.14) of C', we can formally devise a recursive stochastic gradient descent *a.s.* converging toward δ^*. However to make it consistent, we need to introduce constrain so that it lives in $[0, \delta_{\max}]$. In the classical literature on

Stochastic Approximation Theory (see [18] and [19]) this amounts to consider a variant *with projection on the "order book depth interval"* $[0, \delta_{\max}]$, namely

$$\delta_{n+1} = \mathrm{Proj}_{[0,\delta_{\max}]}\left(\delta_n - \gamma_{n+1} H\left(\delta_n, \left(S_t^{(n+1)}\right)_{t \in [0,T]}\right)\right), \ n \geq 0, \ \delta_0 \in (0, \delta_{\max}), \quad (2.4)$$

where

- $\mathrm{Proj}_{[0,\delta_{\max}]}$ denotes the projection on the (nonempty closed convex) $[0, \delta_{\max}]$,

- $(\gamma_n)_{n \geq 1}$ is a positive step sequence satisfying (at least) the minimal *decreasing step assumption* $\sum_{n \geq 1} \gamma_n = +\infty$ and $\gamma_n \to 0$.

- the sequence $\left\{(S_t^{(n)})_{t \in [0,T]}, n \geq 0\right\}$, is the "innovation" sequence of the procedure : ideally it is either a sequence of simulable independent copies of $(S_t)_{t \in [0,T]}$ or a sequence sharing some ergodic (or averaging) properties with respect to the distribution of $(S_t)_{t \in [0,T]}$.

The case of independent copies can be understood as a framework where the dynamics of S is typically a Brownian diffusion solution to an stochastic differential equation, which has been calibrated beforehand on a dataset in order to be simulated on a computer. The case of ergodic copies corresponds to a dataset which is directly plugged into the procedure *i.e.* $S_t^{(n)} = S_{t-n\Delta t}, \ t \in [0,T], \ n \geq 0$, where $\Delta t > 0$ is a fixed shift parameter. To make this second approach consistent, we need to make the assumption that at least within a laps of a few minutes, the dynamics of the asset S (starting in the past) is *stationary* and shares *e.g.* mixing properties.

THE RESULTING IMPLEMENTABLE PROCEDURE: In practice, the above procedure cannot be implemented since the full path $(S_t(\omega))_{t \in [0,T]}$ of a continuous process cannot be simulated nor a functional $H(\delta, (S_t(\omega))_{t \in [0,T]})$ of such a path can be computed. So we are led in practice to replace the "copies" $S^{(n)}$ by copies of a time discretization of step, say $\Delta t = \frac{T}{m}, \ (m \in \mathbb{N}^*)$. The time discretizations are formally defined in continuous time as follows

$$\bar{S}_t = \bar{S}_{t_i}, \ t \in [t_i, t_{i+1}), \ i = 0, \ldots, m-1 \ \text{ with } t_i = \frac{iT}{m}, \ i = 0, \ldots, m,$$

where $(\bar{S}_{t_i})_{0 \leq i \leq m} = (S_{t_i})_{0 \leq k \leq m}$ if $(S_{t_i})_{0 \leq i \leq m}$ can be simulated (see *e.g.* [6] for 1D-Brownian diffusions processes). The sequence $(\bar{S}_{t_i})_{0 \leq i \leq m}$ can also be a time discretization scheme (at times t_i) of $(S_t)_{t \in [0,T]}$, typically an Euler scheme with step $\frac{T}{m}$.

Then, with an obvious abuse of notation for the function H, we can write the *implementable procedure* as follows:

$$\delta_{n+1} = \mathrm{Proj}_{[0,\delta_{\max}]}\left(\delta_n - \gamma_{n+1} H\left(\delta_n, \left(\bar{S}_{t_i}^{(n+1)}\right)_{0 \leq i \leq m}\right)\right), \ n \geq 0, \ \delta_0 \in [0, \delta_{\max}] \quad (2.5)$$

where $\left(\bar{S}_{t_i}^{(n)}\right)_{0 \leq i \leq m}$ are copies of $(\bar{S}_{t_i})_{0 \leq i \leq m}$ either independent or sharing "ergodic" properties, namely some averaging properties in the sense of [21]. In the first case, one will think about simulated data after a calibration process and in the second case to a direct implementation of a historical high frequency data base of best opposite prices of the asset S (with *e.g.* $\bar{S}_{t_i}^{(n)} = S_{t_i - n\frac{T}{m}}$).

2.2 Main convergence results

The following theorems give *a.s.* convergence results for the stochastic procedure (2.4): the first one for i.i.d. sequences and the second one for "averaging" sequences (see [21]).

2.2.1 I.i.d. simulated data from a formerly calibrated model

In this section, we consider that the innovation process $\{(\bar{S}_{t_i}^{(n)})_{0 \leq i \leq m}, n \geq 0\}$ comes from a diffusion model beforehand *calibrated on real data* which can be simulated at time t_i, $0 \leq i \leq m$, either exactly or via a stepwise constant time discretization scheme.

Theorem 2.1. *(a)* THEORETICAL PROCEDURE. *Assume that C is strictly convex $[0, \delta_{\max}]$ with $C'(0) < 0$. Let $(S_t^{(n)})_{t \in [0,T]}$, $n \geq 1$, be a sequence of i.i.d. copies of $(S_t)_{t \in [0,T]}$. Furthermore, assume that the decreasing step sequence satisfies the standard "decreasing step assumption"*

$$\sum_{n \geq 1} \gamma_n = +\infty \quad and \quad \sum_{n \geq 1} \gamma_n^2 < +\infty. \tag{2.6}$$

Then the recursive procedure defined by (2.4) converges a.s. towards its target $\delta^* = \mathrm{argmin}_{\delta \in [0, \delta_{\max}]} C(\delta)$:

$$\delta_n \xrightarrow{a.s.} \delta^*.$$

(b) IMPLEMENTABLE PROCEDURE. *Assume the cost function $\bar{C}$ related to the discretization scheme $(\bar{S}_t)_{t \in [0,T]}$ is strictly convex $[0, \delta_{\max}]$ with $\bar{C}'(0) < 0$ and the step sequence satisfies the "decreasing step" assumption. Let $(\bar{S}_{t_i}^{(n)})_{0 \leq i \leq m}$, $n \geq 1$, be a sequence of i.i.d. copies of $(\bar{S}_{t_i})_{0 \leq i \leq m}$, then the recursive procedure defined by (2.5) converges a.s. towards its target $\bar{\delta}^* = \mathrm{argmin}_{\delta \in [0, \delta_{\max}]} \bar{C}(\delta)$.*

This theorem is a straightforward application of the classical *a.s.* convergence for constrained stochastic algorithms (see Appendix A). In particular, the fact that in the original theorem the innovation process takes values in a finite dimensional space $\mathbb{R}^q$ plays no role in the proof.

2.2.2 Direct implementation on a historical high frequency dataset sharing averaging properties

In this framework we will focus on the time discretized procedure *i.e.* on $(\bar{S}_t)_{t \in [0,T]}$ rather than on $(S_t)_{t \in [0,T]}$ itself. Keep in mind that, when directly implementing a high frequency dataset, then

$$\bar{S}_t = S_{t_i} \ t \in [t_i, t_{i+1}), i = 0, \ldots, m \ \text{ and } \bar{S}_T = S_T.$$

and that the sequence $(\bar{S}_{t_i}^{(n)})_{0 \leq i \leq m}$, $n \geq 1$, is usually obtained by shifting the data as follows: if $\Delta t > 0$ denotes a fixed time shift parameter such that $t_i - t_{i-1} = \Delta t = \frac{T}{m}$, we set

$$\forall t \in [0, T], \quad \bar{S}_{t_i}^{(n)} = \bar{S}_{t_i - n\Delta t} = \bar{S}_{t_{i-n}}.$$

We will assume that the sequence $(\bar{S}_{t_i}^{(n)})_{0 \leq i \leq m}$ shares an averaging property with respect to a distribution ν as developed in [21]. The definition is recalled below.

Definition 2.1. *Let $m \in \mathbb{N}$ and ν be a probability measure on $([0, L]^{m+1}, \mathcal{B}or([0, L]^{m+1}))$. A $[0, L]^{m+1}$-valued sequence $(\xi_n)_{n \geq 1}$ is ν-averaging if*

$$\frac{1}{n} \sum_{k=1}^{n} \delta_{\xi_k} \overset{(\mathbb{R}^{m+1})}{\Longrightarrow} \nu \quad as \ n \to \infty.$$

Then $(\xi_n)_{n\geq 1}$ satisfies

$$D_n^*(\xi) := \sup_{x \in [0,L]^{m+1}} \left| \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{[\![0,x]\!]}(\xi_k) - \nu([\![0,x]\!]) \right| \longrightarrow 0 \quad \text{as } n \to \infty,$$

where $D_n^*(\xi)$ is called the discrepancy at the origin or star discrepancy.

The resulting execution cost function $\bar{C}$ is defined by (2.2) where S is replaced by $(\bar{S}_t)_{t\in[0,T]}$ whose distribution is entirely characterized by the distribution ν. In some sense this function $\bar{C}$ is *the best possible approximation of the true execution function C* that we can get from the high frequency database.

In this setting, we apply the previous results to the price sequence $\{(\bar{S}_{t_i}^{(n)})_{0\leq i\leq m}, n \geq 0\}$, *i.e.* we set for every $n \geq 1$, $\xi_n = (\bar{S}_{t_i}^{(n)})_{0\leq i\leq m}$. In particular we will make the assumption that the dataset is bounded by a real number $L \in (0, +\infty)$ so that $\xi_n \in [0,L]^{m+1}$ for every $n \geq 1$. Moreover, we will need to prove the existence of a *pathwise* Lyapunov function, which means in this one dimensional setting that $H(\cdot, ((s_{t_i})_{0\leq i\leq m})$ is non-decreasing for every $(s_{t_i})_{0\leq i\leq m} \in \mathbb{R}_+^{m+1}$, $n \geq 1$.

Theorem 2.2. IMPLEMENTABLE PROCEDURE. *Let $\lambda(x) = Ae^{-kx}$, $A > 0$, $k > 0$. Assume $(\bar{S}^{(n)})_{n\geq 1}$ is an $[0,L]^{m+1}$-valued ν-averaging sequence where ν is a probability measure on $(\mathbb{R}^{m+1}, \mathcal{B}or(\mathbb{R}^{m+1}))$. Assume that the execution cost function C is strictly convex over $[0, \delta_{\max}]$ with $\bar{C}'(0) < 0$ and $\bar{C}'(\delta_{\max}) > 0$. Finally assume that the step sequence $(\gamma_n)_{n\geq 1}$ is a positive non-increasing sequence satisfying*

$$\sum_{n\geq 1} \gamma_n = +\infty, \quad nD_n^*(\bar{S})\gamma_n \xrightarrow[n\to\infty]{} 0, \quad \text{and} \quad \sum_{n\geq 1} nD_n^*(\bar{S}) \max\left(\gamma_n^2, |\Delta\gamma_{n+1}|\right) < +\infty. \tag{2.7}$$

Furthermore (having in mind that $\bar{S}_0 = S_0$), assume that

$$Q_T \geq 2T\lambda(-\bar{S}_0) \quad \text{and} \quad \kappa \leq \frac{1 + k(\bar{S}_0 - \delta_{\max})}{k \left\|\bar{S}\right\|_\infty (\Phi(Q_T) - \Phi(Q_T - 1))} \tag{2.8}$$

Then the recursive procedure defined by (2.5) converges a.s. towards its target $\bar{\delta}^* = \operatorname{argmin}_{\delta\in[0,\delta_{\max}]} \bar{C}(\delta)$:

$$\delta_n \xrightarrow{a.s.} \bar{\delta}^*.$$

Proof. We apply Theorem 2.1 Section 2 of [21] in a QMC framework similar to Appendix A. First we have that $H(\bar{\delta}^*, \cdot) \in L^1(\mathbb{P})$. Note that $\bar{\delta}^* \in (0, \delta_{\max})$ since $\bar{C}'(0) < 0$ and $\bar{C}'(\delta_{\max}) > 0$ so we can extend $\bar{C}'$ as a convex function on the whole real line. Moreover, by using the proof of Proposition 4.1(b), we prove that if $Q_T \geq 2T\lambda(-\bar{S}_0)$ and (2.8) is satisfied, then H is non-decreasing in δ so that H satisfies the strict pathwise Lyapunov assumption with $L(\delta) = \frac{1}{2}|\delta - \bar{\delta}^*|^2$, namely

$$\forall \delta \in \mathbb{R}\setminus\{\delta^*\}, \ \forall y \in \mathbb{R}^{m+1}, \quad \langle H(\delta, y) - H(\bar{\delta}^*, y) \,|\, \delta - \bar{\delta}^* \rangle > 0.$$

It remains to check the averaging rate assumption for $H(\bar{\delta}^*, \cdot)$: as $(s_i)_{1\leq i\leq m} \mapsto H(\bar{\delta}^*, (s_i)_{1\leq i\leq m})$ is a non-decreasing function (for $(s_i)_{1\leq i\leq m} \leq (s_i')_{1\leq i\leq m}$, *i.e.* $\forall 1 \leq i \leq m$, $s_i \leq s_i'$, *i.e.* $[\![0,s]\!] \subset [\![0,s']\!]$), then $H(\bar{\delta}^*, \cdot)$ has finite variation and by using the Koksma-Hlawka Inequality, we get

$$\left| \frac{1}{n} \sum_{k=1}^n H(\bar{\delta}^*, \bar{S}^{(k)}) - \int_{[0,L]^{m+1}} H(\bar{\delta}^*, s)\nu(ds) \right| \leq (H(\bar{\delta}^*, L) - H(\bar{\delta}^*, 0))D_n^*(\bar{S}),$$

so that $H(\bar{\delta}^*, \cdot)$ is ν-averaging at rate $\varepsilon_n = D_n^*(\bar{S})$. Finally, Theorem 2.1 of Section 2 from [21] yields

$$\delta_n \xrightarrow{a.s.} \bar{\delta}^*. \qquad \square$$

8

Practical comments of the needed bounds.

- The constraint $Q_T \geq 2T\lambda(-S_0)$ is structural: it only involves parameters of the model and the asked quantity Q_T. It means that Q_T does have some chances not to be fully executed before the end of a slice of duration T (*i.e.* the intensity of trades obtained very far away from the current price is smaller than $Q_T/2$).

- The criterion involving the free parameter κ is two-folded depending on the modeling of the "market impact".

  - *The market impact does not depend on the remaining quantity to be traded* (*i.e.* when $\Phi = \mathrm{id}$ or $\eta \equiv 0$ which implies that $\Phi(Q_T) - \Phi(Q_T - 1) = 1$). This setting is appropriate for executing very small quantities or trading very liquid assets (like equity futures). Then the criterion on the free parameter κ reads

    $$\kappa \leq \frac{\bar{S}_0 - \delta_{\max}}{\left\|\bar{S}\right\|_\infty} + \frac{1}{k\left\|\bar{S}\right\|_\infty}.$$

    It states that in this case, the constant premium to pay for the remaining quantity (*i.e.* κ, in basis points) has to be lower than the price range inside which we wish to trade (*i.e.* $(S_0 - \delta_{\max})/\left\|S\right\|_\infty$) plus a *margin*, namely $1/(k\left\|S\right\|_\infty)$. It can be seen as a *symmetry argument*: a model where one cannot imagine buying at a lower price than a given threshold, one cannot either accept to pay (on the other side) more market impact than this very threshold.

  - *The market impact of the remaining quantity is a function of the quantity.* The interpretation is very similar to the previous one. In this case a quantity homogeneous to the market impact (*i.e.* $\kappa \cdot (\Phi(Q_T) - \Phi(Q_T - 1))$ in basis points) should not exceed $(\bar{S}_0 - \delta_{\max})/\left\|\bar{S}\right\|_\infty + 1/(k\left\|\bar{S}\right\|_\infty)$. Here again it is a symmetry argument: the trader cannot consider to pay more market impact for almost one share (*i.e.* $\Phi(Q_T) - \Phi(Q_T - 1)$ plays the role of $\Phi(1)$ but "taken around Q_T") than his reasonable trading range $((\bar{S}_0 - \delta_{\max})/\left\|\bar{S}\right\|_\infty$ in basis points again), plus a margin. Looking more carefully at this margin: $1/(k\left\|\bar{S}\right\|_\infty)$ (the same as in the constant market impact case), it can be read that it is in basis points, and means that if one considers large intensities of fill rates (*i.e.* k is small) then the trader can be lazy on the market impact constraint (because his margin, proportional to $1/k$, is large in such a case), mainly because he will have to pay market impact not that often. If on the contrary, k is large (*i.e.* he will have remaining quantities) then he needs to really fulfill his constraint on market impact.

The needed bounds to obtain the convergence of δ_n toward δ^* are thus not only fulfilled naturally. They also emphasis the consistency of the model and the mechanisms of the proofs.

2.3 Criteria for the convexity and monotony at the origin

In this section, we look for simple criteria involving the parameter κ, that imply the requested assumption on the execution cost function C (or $\bar{C}$). One important feature of this section is that we will never need to really specify the process S. For notational convenience we will drop the $\bar{C}$ notation.

Checking that the assumptions on the function C (*i.e.* C convex with $C'(0) < 0$) in Theorem 2.1 are satisfied on $[0, \delta_{\max}]$ is a nontrivial task: in fact, as emphasized further on in Figures 2 and 7 in Section 5, the function C in (2.2) is never convex on the whole non-negative real line, so we need reasonably simple criteria involving the market impact function Φ, Q_T and the parameter κ and others

quantities related to the asset dynamics which ensure that the required conditions are fulfilled by the function C. These criteria should take the form of upper bounds on the free parameter κ.

Their original form, typically those derived by simply writing $C'(0) < 0$ and $C''(0) \geq 0$, are not really operating since they involve ratios of expectations of functionals combining both the dynamics of the asset S and the execution parameters in a highly nonlinear way. A large part of this paper is devoted to establish simpler criteria (although slightly more conservative) when $(S_t)_{[0,T]}$ is a continuous process satisfying a functional co-monotony principle.

We still need an additional assumption, this time on the function λ. Roughly speaking we need that the functional Λ depends on the distance parameter δ *essentially exponentially* in the following sense (which depends on S):

$$0 < \underline{k}_1 := \mathbb{P}\text{-}\operatorname*{essinf}_{\delta\in[0,\delta_{\max}]} \left(-\frac{\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)}{\Lambda_T(\delta,S)} \right) \leq \overline{k}_1 := \mathbb{P}\text{-}\operatorname*{esssup}_{\delta\in[0,\delta_{\max}]} \left(-\frac{\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)}{\Lambda_T(\delta,S)} \right) < +\infty, \quad (2.9)$$

$$0 < \underline{k}_2 := \mathbb{P}\text{-}\operatorname*{essinf}_{\delta\in[0,\delta_{\max}]} \left(-\frac{\frac{\partial^2}{\partial\delta^2}\Lambda_T(\delta,S)}{\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)} \right) \leq \overline{k}_2 := \mathbb{P}\text{-}\operatorname*{esssup}_{\delta\in[0,\delta_{\max}]} \left(-\frac{\frac{\partial^2}{\partial\delta^2}\Lambda_T(\delta,S)}{\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)} \right) < +\infty. \quad (2.10)$$

Note that the above assumption implies

$$\overline{k}_0 := \mathbb{P}\text{-esssup}\left(-\frac{\frac{\partial}{\partial\delta}\Lambda_T(0,S)}{\Lambda_T(0,S)} \right) \geq \underline{k}_0 := \mathbb{P}\text{-essinf}\left(-\frac{\frac{\partial}{\partial\delta}\Lambda_T(0,S)}{\Lambda_T(0,S)} \right) \geq \underline{k}_1 > 0. \quad (2.11)$$

Although this assumption is stated on the functional Λ (and subsequently depends on S), this is mainly an assumption on the intensity function λ. In particular, the above assumptions are satisfied by intensity functions λ of the form

$$\lambda_k(x) = e^{-kx}, \quad x \in \mathbb{R}, \quad k \in (0,+\infty).$$

For these functions λ_k, one checks that $\underline{k}_0 = \overline{k}_0 = \underline{k}_1 = \overline{k}_1 = \underline{k}_2 = \overline{k}_2 = k$.

The key to establish the criteria is the *functional co-monotony principle* that we establish in the Appendix B for a wide class of diffusions and their associated time discretization schemes.

A Borel functional $F : \mathbb{D}([0,T],\mathbb{R}) \to \mathbb{R}$ is non-decreasing if

$$\forall\alpha,\beta \in \mathbb{D}([0,T],\mathbb{R}), \quad (\forall t\in[0,T],\, \alpha(t)\leq\beta(t)) \Longrightarrow F(\alpha)\leq F(\beta).$$

It is monotonic if F or $-F$ is non-decreasing. Two functionals F and G on $\mathbb{D}([0,T],\mathbb{R})$ are co-monotonic if they are monotonic with the same monotony.

A functional F has polynomial growth if

$$\exists r > 0 \quad \text{s.t.} \quad \forall\alpha \in \mathbb{D}([0,T],\mathbb{R}), \quad |F(\alpha)| \leq K\left(1 + \|\alpha\|_\infty^r\right).$$

Definition 2.2. *A stepwise constant càdlàg (resp. continuous) process $(S_t)_{t\in[0,T]}$ satisfies a functional co-monotony principle if for every pair $F,G : \mathbb{D}([0,T],\mathbb{R}) \to \mathbb{R}$ of co-monotonic Borel functionals (resp. continuous at every $\alpha \in \mathcal{C}([0,T],\mathbb{R})$ for the sup-norm) with polynomial growth such that $F(S)$, $G(S)$ and $F(S)G(S) \in L^1$, one has*

$$\mathbb{E}\left[F\big((S_t)_{t\in[0,T]}\big)G\big((S_t)_{t\in[0,T]}\big) \right] \geq \mathbb{E}\left[F\big((S_t)_{t\in[0,T]}\big) \right]\mathbb{E}\left[G\big((S_t)_{t\in[0,T]}\big) \right]. \quad (2.12)$$

We will use in the proofs below this principle for the price process $(S_t)_{t\in[0,T]}$ for monotonic functionals with opposite monotony. Thus Inequality in (2.12) is reversed (all we have to do is to replace F by $-F$). This co-monotony principle is established and proved in Appendix B for a wide class of Brownian diffusions, their discrete time samples and their Euler schemes.

The main dynamics in which we are interested are the following

10

1. An "admissible" Brownian diffusion in the sense od Definition B.2 satisfies a functional co-monotony principle. So does its Euler scheme with step $\frac{T}{m}$, at least for m large enough, say $m \geq m_{b,\sigma}$, where $m_{b,\sigma}$ only depends on (the Lipschitz coefficients of) the drift b and the diffusion coefficient σ of the SDE that $(S_t)_{t \in [0,T]}$ is solution to.

2. Finally any discrete time sample $(S_{t_i})_{0 \leq i \leq m}$ of a continuous process satisfying a functional co-monotony principle also satisfies a co-monotony principle for continuous functions $f, g : \mathbb{R}^{m+1} \to \mathbb{R}$ monotonic in each of their variables with the same monotony such that $f((S_{t_i})_{0 \leq i \leq m})$, $g((S_{t_i})_{0 \leq i \leq m})$, $fg((S_{t_i})_{0 \leq i \leq m}) \in L^1$.

3. If $(S_{t_i})_{0 \leq i \leq m}$ is a Markov chain whose transitions

$$P_i g(x) = \mathbb{E}\left[g(S_{t_{i+1}} \mid S_{t_i} = x\right], \quad i = 0, \ldots, m,$$

preserve monotony (g non-decreasing implies $P_i g$ is non-decreasing), then $(S_{t_i})_{0 \leq i \leq m}$ also satisfies a co-monotony principle in the same sense as for discrete time sample.

The proof of 1. is postponed in Appendix B (Theorem B.2). Admissible diffusions include standard Brownian motion, geometrical Brownian motion and most models of dynamics traded assets (see Section B.2.3).

Claim 2. follows by associating to a function f the functional $F(\alpha) = f((\alpha(t_i))_{0 \leq i \leq m})$.

For claim 3., we refer to [23] Proposition 2.1.

Theorem 2.3. *Assume that $(S_t)_{t \in [0,T]}$ satisfies a functional co-monotony principle. Assume that the function λ is essentially exponential in the sense of (2.9) and (2.10) above. Then the following monotony and convexity criteria hold true.*

(a) MONOTONY AT THE ORIGIN: *The derivative $C'(0) < 0$ as soon as*

$$Q_T \geq 2T\lambda(-S_0) \quad \text{and} \quad \kappa \leq \frac{1 + \underline{k}_0 S_0}{\overline{k}_0 \mathbb{E}\left[S_T\right](\Phi(Q_T) - \Phi(Q_T - 1))}.$$

(b) CONVEXITY. *Let $\rho_Q \in \left(0, 1 - \frac{\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\big|_{\mu = T\lambda(-S_0)}\right)$. If $\Phi \neq \mathrm{id}$, assume that Φ satisfies*

$$\forall x \in [1, Q_T - 1], \quad \Phi(x) - \Phi(x-1) \leq \rho_Q(\Phi(x+1) - \Phi(x)).$$

If $$Q_T \geq \left(2 \vee \left(1 + \frac{\overline{k}_1^2}{\underline{k}_1 \underline{k}_2}\right)\right) T\lambda(-S_0) \quad \text{and} \quad \kappa \leq \frac{2\underline{k}_1}{\overline{k}_1 \overline{k}_2 \mathbb{E}\left[S_T\right]\Phi'_\ell(Q_T)},$$

then $C''(\delta) \geq 0$, $\delta \in [0, \delta_{\max}]$, so that C is convex on $[0, \delta_{\max}]$.

(c) *The same inequalities hold for the Euler scheme of $(S_t)_{t \in [0,T]}$ with step $\frac{T}{m}$, for $m \geq m_{b,\sigma}$, or for any $\mathbb{R}^{m+1}$-time discretization sequence $(S_{t_i})_{0 \leq i \leq m}$ which satisfies a co-monotony principle.*

Remark. These conditions on the model parameters are conservative. Indeed, "sharper" criteria can be given whose bounds involve ratios of expectation which can be evaluated only by Monte Carlo simulations:

$$C'(0) < 0 \iff 0 < \kappa < b_2,$$

where $$b_2 = \frac{\mathbb{E}\left[-Q_T \mathbb{P}^{(0)}\left(N^\mu > Q_T\right) + \left(S_0 \frac{\partial}{\partial \delta}\Lambda_T(0, S) - \Lambda_T(0, S)\right)\mathbb{P}^{(0)}\left(N^\mu \leq Q_T - 1\right)\right]}{\mathbb{E}\left[S_T \frac{\partial}{\partial \delta}\Lambda_T(0, S)\varphi^{(0)}(\mu)\right]}$$

11

and
$$C \text{ is convex on } [0, \delta_{\max}] \iff 0 < \kappa < \min_{\delta \in \mathcal{D}_+} \frac{A(\delta)}{B(\delta)}$$

where
$$\begin{aligned}
A(\delta) &= \mathbb{E}\left[\left((S_0 - \delta)\frac{\partial^2}{\partial \delta^2}\Lambda_T(\delta, S) - 2\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \leq Q_T - 1\right)\right. \\
&\qquad \left. - (S_0 - \delta)\left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)^2 \mathbb{P}^{(\delta)}\left(N^\mu = Q_T - 1\right)\right],
\end{aligned}$$

$$B(\delta) = \mathbb{E}\left[S_T\left(\frac{\partial^2}{\partial \delta^2}\Lambda_T(\delta, S)\varphi^{(\delta)}(\mu) - \left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)^2\psi^{(\delta)}(\mu)\right)\right] \quad \text{and} \quad \mathcal{D}_+ = \{\delta \in [0, \delta_{\max}] \mid B(\delta) > 0\}.$$

3 Representations as expectations of C and its derivatives

First we briefly recall for convenience few basic facts on Poisson distributed variables that will be needed to compute the cost function C and its derivatives C' and C'' (proofs are left to the reader).

Proposition 3.1. (Classical formulas). *Let $(N^\mu)_{\mu > 0}$ be a family of Poisson distributed random variables with parameter $\mu > 0$.*

(i) *For every function $f : \mathbb{N} \to \mathbb{R}_+$ such that $\log f(n) = O(n)$,*

$$\frac{d}{d\mu}\mathbb{E}\left[f(N^\mu)\right] = \mathbb{E}\left[f(N^\mu + 1) - f(N^\mu)\right] = \mathbb{E}\left[f(N^\mu)\left(\frac{N^\mu}{\mu} - 1\right)\right].$$

In particular, for any $k \in \mathbb{N}$, $\frac{d}{d\mu}\mathbb{P}\left(N^\mu \leq k\right) = -\mathbb{P}\left(N^\mu = k\right)$.

For any $k \in \mathbb{N}^*$,

(ii) $\mathbb{E}\left[k \wedge N^\mu\right] = k\mathbb{P}\left(N^\mu > k\right) + \mu\mathbb{P}\left(N^\mu \leq k - 1\right)$ *and* $\frac{d}{d\mu}\mathbb{E}\left[k \wedge N^\mu\right] = \mathbb{P}\left(N^\mu \leq k - 1\right)$,

(iii) $\mathbb{E}\left[(k - N^\mu)_+\right] = k\mathbb{P}\left(N^\mu \leq k\right) - \mu\mathbb{P}\left(N^\mu \leq k - 1\right)$,

(iv) $k\mathbb{P}\left(N^\mu = k\right) = \mu\mathbb{P}\left(N^\mu = k - 1\right)$.

To compute the cost function (or its gradient), it is convenient to proceed a pre-conditioning with respect to $\mathcal{F}_T^S := \sigma\left(S_t, 0 \leq t \leq T\right)$. We come down to compute the above quantity when $N^{(\delta)}$ is replaced by a standard Poisson random variable with parameter μ denoted N^μ. Therefore we have

$$\begin{aligned}
C(\delta) &= \mathbb{E}\left[(S_0 - \delta)\left(Q_T \wedge N_T^{(\delta)}\right) + \kappa S_T \Phi\left(\left(Q_T - N_T^{(\delta)}\right)_+\right)\right] \\
&= \mathbb{E}\left[(S_0 - \delta)\mathbb{E}\left[\left(Q_T \wedge N_T^{(\delta)}\right) \mid \mathcal{F}_T^S\right] + \kappa S_T\mathbb{E}\left[\Phi\left(\left(Q_T - N_T^{(\delta)}\right)_+\right) \mid \mathcal{F}_T^S\right]\right] \\
&= \mathbb{E}\left[(S_0 - \delta)\mathbb{E}\left[Q_T \wedge N^\mu\right]_{|\mu = \Lambda_T(\delta, S)} + \kappa S_T\mathbb{E}\left[\Phi\left((Q_T - N^\mu)_+\right)\right]_{|\mu = \Lambda_T(\delta, S)}\right] \\
&= \mathbb{E}\left[\widetilde{C}\left(\delta, \Lambda_T(\delta, S), (S_t)_{0 \leq t \leq T}\right)\right], \tag{3.13}
\end{aligned}$$

where for every $x \in \mathcal{C}([0, T], \mathbb{R}_+)$ and every $\mu \in \mathbb{R}_+$,

$$\widetilde{C}\left(\delta, \mu, x\right) = (x_0 - \delta)\left(Q_T\mathbb{P}\left(N^\mu > Q_T\right) + \mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right) + \kappa x_T\mathbb{E}\left[\Phi\left((Q_T - N^\mu)_+\right)\right].$$

We introduce some notations for reading convenience: we set

$$\mathbb{P}^{(\delta)}\left(N^\mu > Q_T\right) = \mathbb{P}\left(N^\mu > Q_T\right)_{|\mu = \Lambda_T(\delta, S)} \quad \text{and} \quad \mathbb{E}^{(\delta)}\left[f(\mu)\right] = \mathbb{E}\left[f(\mu)\right]_{|\mu = \Lambda_T(\delta, S)}.$$

Now we are in position to compute the first and second derivatives of the cost function C.

Proposition 3.2. *(a) If* $\Phi \neq$ id, *then* $C'(\delta) = \mathbb{E}\left[H(\delta, S)\right]$ *with*

$$H(\delta, S) = -Q_T \mathbb{P}^{(\delta)}\left(N^\mu > Q_T\right) + \left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)(S_0 - \delta) - \Lambda_T(\delta, S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \leq Q_T - 1\right)$$

$$-\kappa S_T \frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\varphi^{(\delta)}(\mu) \tag{3.14}$$

where $\varphi^{(\delta)}(\mu) = \mathbb{E}^{(\delta)}\left[\left(\Phi\left(Q_T - N^\mu\right) - \Phi\left(Q_T - N^\mu - 1\right)\right)\mathbb{1}_{\{N^\mu \leq Q_T - 1\}}\right]$ *and*

$$C''(\delta) = \mathbb{E}\left[\left((S_0 - \delta)\frac{\partial^2}{\partial \delta^2}\Lambda_T(\delta, S) - 2\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \leq Q_T - 1\right) - \kappa S_T \frac{\partial^2}{\partial \delta^2}\Lambda_T(\delta, S)\varphi^{(\delta)}(\mu)\right.$$

$$\left. -(S_0 - \delta)\left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)^2 \mathbb{P}^{(\delta)}\left(N^\mu = Q_T - 1\right) + \kappa S_T\left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)^2 \psi^{(\delta)}(\mu)\right] \tag{3.15}$$

where $\psi^{(\delta)}(\mu) = \mathbb{E}^{(\delta)}\left[\Phi\left((Q_T - N^\mu - 2)_+\right) - 2\Phi\left((Q_T - N^\mu - 1)_+\right) + \Phi\left((Q_T - N^\mu)_+\right)\right]$.

(b) If $\Phi =$ id, *then* $C'(\delta) = \mathbb{E}\left[H(\delta, S)\right]$ *with*

$$H(\delta, S) = -Q_T \mathbb{P}^{(\delta)}\left(N^\mu > Q_T\right) + \left((S_0 - \delta - \kappa S_T)\frac{\partial}{\partial \delta}\Lambda_T(\delta, S) - \Lambda_T(\delta, S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \leq Q_T - 1\right) \tag{3.16}$$

and $\quad C''(\delta) = \mathbb{E}\left[\left((S_0 - \delta - \kappa S_T)\frac{\partial^2}{\partial \delta^2}\Lambda_T(\delta, S) - 2\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \leq Q_T - 1\right)\right.$

$$\left. -(S_0 - \delta - \kappa S_T)\left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right)^2 \mathbb{P}^{(\delta)}\left(N^\mu = Q_T - 1\right)\right] \tag{3.17}$$

Proof. Interchanging derivation and expectation in the representation (3.13) implies

$$C'(\delta) = \mathbb{E}\left[\frac{\partial}{\partial \delta}\widetilde{C}\left(\delta, \Lambda_T(\delta, S), (S_t)_{0 \leq t \leq T}\right) + \frac{\partial}{\partial \mu}\widetilde{C}\left(\delta, \Lambda_T(\delta, S), (S_t)_{0 \leq t \leq T}\right)\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\right].$$

(a) We come down to compute the partial derivatives of $\widetilde{C}(\delta, \mu, x)$.

$$\frac{\partial \widetilde{C}}{\partial \delta}(\delta, \mu, x) = -\mathbb{E}\left[(Q_T \wedge N^\mu)\right] = -Q_T \mathbb{P}\left(N^\mu > Q_T\right) - \mu \mathbb{P}\left(N^\mu \leq Q_T - 1\right),$$

$$\frac{\partial \widetilde{C}}{\partial \mu}(\delta, \mu, x) = -(x_0 - \delta)\frac{\partial}{\partial \mu}\mathbb{E}\left[(Q_T - N^\mu)_+\right] + \kappa x_T \frac{\partial}{\partial \mu}\mathbb{E}\left[\Phi\left((Q_T - N^\mu)_+\right)\right].$$

We have $\quad \dfrac{\partial}{\partial \mu}\mathbb{E}\left[(Q_T - N^\mu)_+\right] = -Q_T \mathbb{P}\left(N^\mu = Q_T\right) - \mathbb{P}\left(N^\mu \leq Q_T - 1\right) + \mu \mathbb{P}\left(N^\mu = Q_T - 1\right)$

$$= -\mathbb{P}\left(N^\mu \leq Q_T - 1\right) \quad \text{thanks to } (iv) \text{ in Proposition 3.1}$$

and $\quad \dfrac{\partial}{\partial \mu}\mathbb{E}\left[\Phi\left((Q_T - N^\mu)_+\right)\right] = \mathbb{E}\left[\Phi\left((Q_T - N^\mu - 1)_+\right) - \Phi\left((Q_T - N^\mu)_+\right)\right]$

$$= \mathbb{E}\left[\left(\Phi\left(Q_T - N^\mu - 1\right) - \Phi\left(Q_T - N^\mu\right)\right)\mathbb{1}_{\{N^\mu \leq Q_T - 1\}}\right] := -\varphi(\mu)$$

owing to (v) in Proposition 3.1. Therefore

$$\frac{\partial \widetilde{C}}{\partial \mu}(\delta, \mu, x) = (x_0 - \delta)\mathbb{P}\left(N^\mu \leq Q_T - 1\right) - \kappa x_T \varphi(\mu).$$

13

Consequently

$$
\begin{aligned}
C'(\delta) &= \mathbb{E}\left[-Q_T \mathbb{P}^{(\delta)}\left(N^\mu > Q_T\right) + \left(\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)(S_0-\delta) - \Lambda_T(\delta,S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \le Q_T-1\right)\right. \\
&\qquad \left. - \kappa S_T \frac{\partial}{\partial\delta}\Lambda_T(\delta,S)\varphi^{(\delta)}(\mu)\right] \\
&= \mathbb{E}\left[\widehat{C}\left(\delta, \Lambda_T(\delta,S), \frac{\partial}{\partial\delta}\Lambda_T(\delta,S), (S_t)_{0\le t\le T}\right)\right],
\end{aligned}
\tag{3.18}
$$

where $\varphi^{(\delta)}(\mu) := \varphi(\mu)_{|\mu=\Lambda_T(\delta,S)}$ and for every $x \in \mathcal{C}([0,T],\mathbb{R}_+)$ and every $\mu,\nu \in \mathbb{R}_+$,

$$
\widehat{C}(\delta,\mu,\nu,x) = -Q_T\mathbb{P}\left(N^\mu > Q_T\right) + (\nu(x_0-\delta)-\mu)\mathbb{P}\left(N^\mu \le Q_T-1\right) - \kappa x_T\nu\varphi(\mu).
$$

Interchanging derivation and expectation in the representation (3.18) implies

$$
\begin{aligned}
C''(\delta) &= \mathbb{E}\left[\frac{\partial}{\partial\delta}\widehat{C}\left(\delta, \Lambda_T(\delta,S), \frac{\partial}{\partial\delta}\Lambda_T(\delta,S), (S_t)_{0\le t\le T}\right)\right. \\
&\qquad +\frac{\partial}{\partial\mu}\widehat{C}\left(\delta, \Lambda_T(\delta,S), \frac{\partial}{\partial\delta}\Lambda_T(\delta,S), (S_t)_{0\le t\le T}\right)\frac{\partial}{\partial\delta}\Lambda_T(\delta,S) \\
&\qquad \left. +\frac{\partial}{\partial\nu}\widehat{C}\left(\delta, \Lambda_T(\delta,S), \frac{\partial}{\partial\delta}\Lambda_T(\delta,S), (S_t)_{0\le t\le T}\right)\frac{\partial^2}{\partial\delta^2}\Lambda_T(\delta,S)\right].
\end{aligned}
$$

We deal now with the partial derivatives of $\widehat{C}(\delta,\mu,\nu,x)$.

$$
\begin{aligned}
\frac{\partial\widehat{C}}{\partial\delta}(\delta,\mu,\nu,x) &= -\nu\mathbb{P}\left(N^\mu \le Q_T-1\right), \\
\frac{\partial\widehat{C}}{\partial\mu}(\delta,\mu,\nu,x) &= -\mathbb{P}\left(N^\mu \le Q_T-1\right) - (x_0-\delta)\nu\mathbb{P}\left(N^\mu = Q_T-1\right) + \kappa x_T\nu\psi(\mu), \\
\frac{\partial\widehat{C}}{\partial\nu}(\delta,\mu,\nu,x) &= (x_0-\delta)\mathbb{P}\left(N^\mu \le Q_T-1\right) - \kappa x_T\varphi(\mu).
\end{aligned}
$$

Consequently

$$
\begin{aligned}
C''(\delta) &= \mathbb{E}\left[\left((S_0-\delta)\frac{\partial^2}{\partial\delta^2}\Lambda_T(\delta,S) - 2\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \le Q_T-1\right) - \kappa S_T\frac{\partial^2}{\partial\delta^2}\Lambda_T(\delta,S)\varphi^{(\delta)}(\mu)\right. \\
&\qquad \left. -(S_0-\delta)\left(\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)\right)^2\mathbb{P}^{(\delta)}\left(N^\mu = Q_T-1\right) + \kappa S_T\left(\frac{\partial}{\partial\delta}\Lambda_T(\delta,S)\right)^2\psi^{(\delta)}(\mu)\right].
\end{aligned}
$$

(b) If $\Phi = \mathrm{id}$ so that $\frac{\partial}{\partial\mu}\mathbb{E}\left[\Phi\left((Q_T-N^\mu)_+\right)\right] = -\mathbb{P}\left(N^\mu \le Q_T-1\right)$. Therefore

$$
\frac{\partial\widetilde{C}}{\partial\delta}(\delta,\mu,x) = -Q_T\mathbb{P}\left(N^\mu > Q_T\right) - \mu\mathbb{P}\left(N^\mu \le Q_T-1\right) \quad \text{and} \quad \frac{\partial\widetilde{C}}{\partial\mu}(\delta,\mu,x) = (x_0-\delta-\kappa x_T)\mathbb{P}\left(N^\mu \le Q_T-1\right).
$$

Consequently

$$
\begin{aligned}
C'(\delta) &= \mathbb{E}\left[-Q_T\mathbb{P}^{(\delta)}\left(N^\mu > Q_T\right) + \left((S_0-\delta-\kappa S_T)\frac{\partial}{\partial\delta}\Lambda_T(\delta,S) - \Lambda_T(\delta,S)\right)\mathbb{P}^{(\delta)}\left(N^\mu \le Q_T-1\right)\right] \\
&= \mathbb{E}\left[\widehat{C}\left(\delta, \Lambda_T(\delta,S), \frac{\partial}{\partial\delta}\Lambda_T(\delta,S), (S_t)_{0\le t\le T}\right)\right],
\end{aligned}
\tag{3.19}
$$

14

where for every $x \in \mathcal{C}([0,T], \mathbb{R}_+)$ and every $\mu, \nu \in \mathbb{R}_+$,

$$\widehat{C}(\delta, \mu, \nu, x) = -Q_T \mathbb{P}(N^\mu > Q_T) + ((x_0 - \delta - \kappa x_T)\nu - \mu) \mathbb{P}(N^\mu \le Q_T - 1).$$

Interchanging derivation and expectation in the the representation (3.19) implies

$$
\begin{aligned}
C''(\delta) &= \mathbb{E}\left[\frac{\partial}{\partial \delta} \widehat{C}\left(\delta, \Lambda_T(\delta, S), \frac{\partial}{\partial \delta} \Lambda_T(\delta, S), (S_t)_{0 \le t \le T}\right) \right.\\
&\quad + \frac{\partial}{\partial \mu} \widehat{C}\left(\delta, \Lambda_T(\delta, S), \frac{\partial}{\partial \delta} \Lambda_T(\delta, S), (S_t)_{0 \le t \le T}\right) \frac{\partial}{\partial \delta} \Lambda_T(\delta, S) \\
&\quad + \left. \frac{\partial}{\partial \nu} \widehat{C}\left(\delta, \Lambda_T(\delta, S), \frac{\partial}{\partial \delta} \Lambda_T(\delta, S), (S_t)_{0 \le t \le T}\right) \frac{\partial^2}{\partial \delta^2} \Lambda_T(\delta, S) \right].
\end{aligned}
$$

We come down to compute the partial derivatives of $\widehat{C}\left(\delta, \Lambda_T(\delta, S), \frac{\partial}{\partial \delta} \Lambda_T(\delta, S), (S_t)_{0 \le t \le T}\right)$.

$$\frac{\partial \widehat{C}}{\partial \delta}(\delta, \mu, \nu, x) = -\nu \mathbb{P}(N^\mu \le Q_T - 1), \quad \frac{\partial \widehat{C}}{\partial \nu}(\delta, \mu, \nu, x) = (x_0 - \delta - \kappa x_T) \mathbb{P}(N^\mu \le Q_T - 1),$$

$$\frac{\partial \widehat{C}}{\partial \mu}(\delta, \mu, \nu, x) = -\mathbb{P}(N^\mu \le Q_T - 1) - (x_0 - \delta - \kappa x_T) \frac{\partial}{\partial \delta} \Lambda_T(\delta, S) \mathbb{P}(N^\mu = Q_T - 1).$$

Consequently

$$
\begin{aligned}
C''(\delta) &= \mathbb{E}\left[\left((S_0 - \delta - \kappa S_T) \frac{\partial^2}{\partial \delta^2} \Lambda_T(\delta, S) - 2 \frac{\partial}{\partial \delta} \Lambda_T(\delta, S)\right) \mathbb{P}^{(\delta)}(N^\mu \le Q_T - 1) \right.\\
&\quad \left. - (S_0 - \delta - \kappa S_T) \left(\frac{\partial}{\partial \delta} \Lambda_T(\delta, S)\right)^2 \mathbb{P}^{(\delta)}(N^\mu = Q_T - 1) \right]. \qquad \square
\end{aligned}
$$

4 Convexity and monotony criteria for the cost function C

To ensure that the optimization problem is well-posed, namely that the cost function C has a minimum on $[0, \delta_{\max}]$, we need some additional assumptions: the cost function C must satisfy $C'(0) < 0$. This leads to define bounds for the parameter κ and this section is devoted to give sufficient condition on κ to ensure that this two properties are satisfied. The computations of the bounds given below rely on the co-monotony principle introduced in the Appendix Section B.

4.1 Criteria for local and global monotony

The proposition below gives bounds for the parameter κ which ensure that the execution cost function C has a minimum. The aim of this subsection is to obtain sufficient bounds, easy to compute, namely depending only of the model parameters.

Proposition 4.1. *Assume that* $(S_t)_{t \in [0,T]}$ *satisfies a co-monotony principle.*

(a) MONOTONY AT THE ORIGIN. *$C'(0) < 0$ as soon as $Q_T \ge 2T\lambda(-S_0)$, $\underline{k}_0 = \mathrm{infess}\left(-\frac{\frac{\partial}{\partial \delta} \Lambda_T(0,S)}{\Lambda_T(0,S)}\right) > 0$ and*

$$\kappa \le \frac{1 + \underline{k}_0 S_0}{\overline{k}_0 \mathbb{E}[S_T] (\Phi(Q_T) - \Phi(Q_T - 1))}.$$

15

In particular, when $\Phi \equiv \mathrm{id}$, the condition reduces to

$$\kappa \le \frac{1 + \underline{k}_0 S_0}{\overline{k}_0 \mathbb{E}\left[S_T\right]}.$$

(b) GLOBAL MONOTONY (EXPONENTIAL INTENSITY). *Assume that* $s^* := \|\sup_{t \in [0,T]} S_t\|_{L^\infty} < +\infty$. *If* $\lambda(x) = A e^{-kx}$, $A > 0$, $k > 0$, $Q_T \ge 2T\lambda(-S_0)$ *and*

$$\kappa \le \frac{1 + k(S_0 - \delta_{\max})}{k\, s^*} \;\; \text{if } \Phi \ne \mathrm{id}, \quad \kappa \le \frac{1 + k(S_0 - \delta_{\max})}{k\, s^*(\Phi(Q_T) - \Phi(Q_T - 1))} \;\; \text{if } \Phi = \mathrm{id}, \tag{4.20}$$

then $H(\cdot, (y_i)_{0 \le i \le m})$ *is non-decreasing on* $[0, \delta_{\max}]$ *for every* $(y_i)_{0 \le i \le m} \in [0, s^*]^{m+1}$.

(c) *The same inequality holds for the Euler scheme of* $(S_t)_{t \in [0,T]}$ *with step* $\frac{T}{m}$, *for* $m \ge m_{b,\sigma}$, *or for any* $\mathbb{R}^{m+1}$-*time discretization sequence* $(S_{t_i})_{0 \le i \le m}$ *which satisfies a co-monotony principle (see Theorem* 2.3-(c) *and Corollary* B.1$)$.

To prove this result, we need to establish the monotony of several functions of μ which appear in the expression of C'.

Lemma 4.1. (i) *The function* $\mu \longmapsto \mu \mathbb{P}\left(N^\mu \le Q\right)$ *is non-decreasing on* $\left[0, \left\lfloor \frac{Q+1}{2} \right\rfloor\right]$.

(ii) *The function* $\mu \longmapsto \Theta(Q, \mu) := \mathbb{E}\left[\Phi(Q - N^\mu) - \Phi(Q - N^\mu - 1) \mid N^\mu \le Q - 1\right]$ *satisfies* $\Theta(Q_T, \mu) \le \Theta(Q_T, 0) = \Phi(Q) - \Phi(Q - 1)$ *for all* μ.

Proof of Lemma 4.1. (i) We have $\frac{d}{d\mu}\left(\mu \mathbb{P}\left(N^\mu \le Q\right)\right) = \mathbb{P}\left(N^\mu \le Q\right) - \mu \mathbb{P}\left(N^\mu = Q\right)$. Consequently

$$\frac{d}{d\mu}\left(\mu \mathbb{P}\left(N^\mu \le Q\right)\right) \ge 0 \quad \text{iff} \quad \sum_{k=0}^{Q} \frac{\mu^k}{k!} \ge \mu \frac{\mu^Q}{Q!}.$$

But $k \longmapsto \frac{\mu^k}{k!}$ is non-decreasing on $\{0, 1, \ldots, \lfloor \mu \rfloor\}$ and non-increasing on $\{\lfloor \mu \rfloor, \ldots\}$.

Hence $\sum_{k=0}^{Q} \frac{\mu^k}{k!} \ge \sum_{k=\lfloor \mu \rfloor}^{Q} \frac{\mu^Q}{Q!} = (Q - \lfloor \mu \rfloor)\frac{\mu^Q}{Q!}$, so that $\sum_{k=0}^{Q} \frac{\mu^k}{k!} \ge \mu \frac{\mu^Q}{Q!}$ as soon as $Q \ge 2\lfloor \mu \rfloor + 1$.

(ii) The function Φ is non-decreasing, non-negative and convex with $\Phi(0) = 0$. If we look at the representation of $\mu \mapsto N^\mu$ by

$$N^\mu(\omega) = \max\left\{n \in \mathbb{N} \mid \prod_{i=1}^{n} U_i(\omega) > e^{-\mu}\right\},$$

where U_i are i.i.d. uniformly distributed random variables on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, then $\mu \mapsto N^\mu$ is clearly non-decreasing, so $\mu \mapsto Q - N^\mu$ is non-increasing and $\mu \mapsto \varphi(\mu) = \Phi(Q - N^\mu) - \Phi(Q - N^\mu - 1)$ too (because of the convexity of Φ). $\qquad \square$

Remark. If $\mu \in (0, 1)$, then $\mu \longmapsto \mu \mathbb{P}\left(N^\mu \le Q\right)$ is always non-decreasing. If $\mu \in [1, 2)$, then the function $\mu \longmapsto \mu \mathbb{P}\left(N^\mu = 0\right) = \mu e^{-\mu}$ is not always non-decreasing clearly, but only on $[0, 1]$.

Proof of Proposition 4.1. (a) In our problem the intensity parameter $\mu = \int_0^T \lambda(S_t - S_0 + \delta)dt$ is continuous, non-increasing to zero when δ tends to $+\infty$ and bounded by assumption $(\lambda(-S_0) < +\infty)$. Hence $\mu \in [0, \lambda(-S_0)T]$.

16

(i) From (3.18), we have for $\delta = 0$,

$$\widehat{C}\left(0, \Lambda_T(0,S), \frac{\partial}{\partial\delta}\Lambda_T(0,S), (S_t)_{0\leq t\leq T}\right)$$

$$\leq \left(S_0\frac{\partial}{\partial\delta}\Lambda_T(0,S) - \Lambda_T(0,S)\right)\mathbb{P}^{(0)}\left(N^\mu \leq Q_T - 1\right) - \kappa S_T\frac{\partial}{\partial\delta}\Lambda_T(0,S)\varphi^{(0)}(\mu)$$

$$= \left(\frac{\partial}{\partial\delta}\Lambda_T(0,S)\left(S_0 - \kappa S_T\Theta(Q_T,\Lambda_T(0,S))\right) - \Lambda_T(0,S)\right)\mathbb{P}^{(0)}\left(N^\mu \leq Q_T - 1\right),$$

because $-Q_T\mathbb{P}\left(N^\mu > Q_T\right)_{|\mu=\Lambda_T(0,S)} < 0$ and small if Q_T is large. Let $\underline{k}_0$ and $\overline{k}_0$ be defined by (2.11). We have that $\underline{k}_0 > 0$ a.s. and $\overline{k}_0 > 0$ a.s. by assumption, i.e. $\frac{\partial}{\partial\delta}\Lambda_T(0,S) \leq -\underline{k}_0\Lambda_T(0,S)$ a.s. and $\frac{\partial}{\partial\delta}\Lambda_T(0,S) \geq -\overline{k}_0\Lambda_T(0,S)$ a.s.. Then

$$\widehat{C}\left(0, \Lambda_T(0,S), \frac{\partial}{\partial\delta}\Lambda_T(0,S), (S_t)_{0\leq t\leq T}\right)$$

$$\leq -\left(1 + \underline{k}_0 S_0 - \kappa S_T\overline{k}_0\Theta(Q_T, \Lambda_T(0,S))\right)\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}.$$

Now, by Lemma 4.1, $$\Theta(Q_T, \mu) \leq \Theta(Q_T, 0) = \Phi(Q) - \Phi(Q-1) = \varphi(0).$$

Therefore $\quad\widehat{C}\left(0, \Lambda_T(0,S), \frac{\partial}{\partial\delta}\Lambda_T(0,S), (S_t)_{0\leq t\leq T}\right)$

$$\leq -\left(1 + \underline{k}_0 S_0 - \kappa S_T\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right)\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}.$$

By Lemma 4.1, if $Q_T \geq 2T\lambda(-S_0)$, then $\mu \longmapsto \mu\mathbb{P}\left(N^\mu \leq Q - 1\right)$ is non-decreasing. Moreover, the functionals

$$F : \mathbb{D}([0,T,\mathbb{R}) \rightarrow \mathbb{R}$$
$$\alpha \mapsto \Lambda_T(0,\alpha) = \int_0^T \lambda(\alpha(t) - S_0)dt$$

is non-increasing and

$$\alpha(T) \longmapsto \left(-1 - \underline{k}_0 S_0 + \kappa\alpha(T)\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right) \quad \text{is non-decreasing.}$$

Both are continuous (with respect to the sup-norm) at any $\alpha \in \mathcal{C}([0,T],\mathbb{R})$. Conserquently, we obtain by the functional co-monotony principle

$$\mathbb{E}\left[\left(-1 - \underline{k}_0 S_0 + \kappa S_T\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right)\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}\right]$$

$$\leq \mathbb{E}\left[-1 - \underline{k}_0 S_0 + \kappa S_T\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right]\mathbb{E}\left[\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}\right].$$

In turn, this implies

$$C'(0) \leq \mathbb{E}\left[-1 - \underline{k}_0 S_0 + \kappa S_T\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right]\mathbb{E}\left[\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}\right].$$

As $\mathbb{E}\left[\left(\mu\mathbb{P}\left(N^\mu \leq Q_T - 1\right)\right)_{|\mu=\Lambda_T(0,S)}\right] \geq 0$, then $C'(0) \leq 0$ as soon as

$$\mathbb{E}\left[-1 - \underline{k}_0 S_0 + \kappa S_T\overline{k}_0\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)\right] \leq 0,$$

i.e. $$\kappa \leq \frac{1 + \underline{k}_0 S_0}{\overline{k}_0\mathbb{E}\left[S_T\right]\left(\Phi(Q_T) - \Phi(Q_T - 1)\right)}.$$

(b) From (3.18), the form of λ and (a), we get that for every $\delta \in [0, \delta_{\max}]$, $S = (S_i)_{1 \le i \le m} \in \mathbb{R}^{m+1}$,

$$H(\delta, S) = -Q_T \mathbb{P}^{(\delta)}(N^\mu > Q_T) + f(\delta, S)\left(\mu \mathbb{P}(N^\mu \le Q_T - 1)\right)_{|\mu = \Lambda_T(\delta, S)},$$

where $f(\delta, S) = -1 - k(S_0 - \delta - \kappa S_T \Theta(Q_T - 1, \Lambda_T(\delta, S)))$ if $\Phi \ne \mathrm{id}$ and $f(\delta, S) = -1 - k(S_0 - \delta - \kappa S_T)$ if $\Phi = \mathrm{id}$. Since $\delta \mapsto -Q_T \mathbb{P}^{(\delta)}(N^\mu > Q_T)$ is non-decreasing, $\delta \mapsto \left(\mu \mathbb{P}(N^\mu \le Q_T - 1)\right)_{|\mu = \Lambda_T(\delta, S)}$ is non-increasing and non positive owing to Lemma 4.1 (i) and $\delta \mapsto f(\delta, S)$ is non-decreasing owing to Lemma 4.1 (ii), $\delta \mapsto H(\delta, S)$ is non-decreasing if $f(\delta, S) > 0$, $\delta \in [0, \delta_{\max}]$, $S \in \mathbb{R}^{m+1}$, which leads to (4.20). $\qquad \square$

4.2 Sufficient condition for the convexity condition

We give below sufficient condition to insure the convexity of the execution cost function C, namely conservative upper bound for the free parameter κ.

Proposition 4.2. *Assume that $(S_t)_{t \in [0,T]}$ satisfies a co-monotony principle.*

(i) If $\Phi \ne \mathrm{id}$, assume that there exists $\rho_Q \in \left(0, 1 - \frac{\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \le Q_T - 1)}_{|\mu = T\lambda(-S_0)}\right)$ such that

$$\forall x \in [1, Q_T - 1], \quad \Phi(x) - \Phi(x-1) \le \rho_Q(\Phi(x+1) - \Phi(x)).$$

If
$$Q_T \ge \left(2 \vee \left(1 + \frac{\overline{k}_1^2}{\underline{k}_1 \underline{k}_2}\right)\right) T\lambda(-S_0) \qquad and \qquad \kappa \le \frac{2\underline{k}_1}{\overline{k}_1 \overline{k}_2 \mathbb{E}[S_T] \Phi'_\ell(Q_T)},$$

where Φ'_ℓ is the left derivative of the convex function Φ, then $C''(\delta) \ge 0$, $\delta \in [0, \delta_{\max}]$, so that C is convex on $[0, \delta_{\max}]$.

(ii) When $\Phi = \mathrm{id}$, the bound on κ reads

$$\kappa \le \frac{2\underline{k}_1}{\overline{k}_1 \overline{k}_2 \mathbb{E}[S_T]}.$$

Remark. Note that if we replace the diffusion process $(S_t)_{t \in [0,T]}$ by its stepwise constant Euler scheme or by a discrete time sample $(S_{t_i})_{0 \le i \le m}$, we have the same criteria owing to Corollary B.1 (see also Theorem 2.3-(c)).

To prove the above Proposition, we need the following results

Lemma 4.2. *If $\mu \le Q - 1$, then $\mu \mapsto \frac{\mathbb{P}(N^\mu = Q - 1)}{\mathbb{P}(N^\mu \le Q - 1)}$ is non-decreasing.*

Proof of Lemma 4.2.

$$\frac{d}{d\mu} \frac{\mathbb{P}(N^\mu = Q - 1)}{\mathbb{P}(N^\mu \le Q - 1)} = \frac{\mathbb{P}(N^\mu = Q - 2) - \mathbb{P}(N^\mu = Q - 1)}{\mathbb{P}(N^\mu \le Q - 1)} + \frac{\mathbb{P}(N^\mu = Q - 1)^2}{\mathbb{P}(N^\mu \le Q - 1)^2}$$

$$= \frac{\mathbb{P}(N^\mu = Q - 1)\left(\mathbb{P}(N^\mu \le Q - 1)\left(\frac{Q-1}{\mu} - 1\right) + \mathbb{P}(N^\mu = Q - 1)\right)}{\mathbb{P}(N^\mu \le Q - 1)^2}$$

$$\ge 0 \quad \text{if} \quad \mu \le Q - 1. \qquad \square$$

Lemma 4.3. *Assume that $\Phi \ne \mathrm{id}$. If there exists $\rho_Q \in \left(0, 1 - \frac{\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \le Q_T - 1)}_{|\mu = T\lambda(-S_0)}\right)$ such that*

$$\forall x \in [1, Q_T - 1], \quad \Phi(x) - \Phi(x-1) \le \rho_Q(\Phi(x+1) - \Phi(x)),$$

then $\mu \mapsto \frac{\varphi(\mu)}{\mathbb{P}(N^\mu \le Q_T - 1)}$ is non-increasing where $\varphi(\mu) = \Phi(Q - N^\mu) - \Phi(Q - N^\mu - 1)$.

Remark. If $\Phi = \mathrm{id}$, then $\mu \mapsto \frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)} \equiv 1$, therefore we do not need the previous lemmas.

Proof of Lemma 4.3. We have

$$\frac{d}{d\mu} \frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)} = \frac{\mathbb{P}(N^\mu = Q_T - 1)\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)^2} - \frac{\psi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)} \leq 0$$

where $\psi(\mu) = \Phi\left((Q_T - N^\mu - 2)_+\right) - 2\Phi\left((Q_T - N^\mu - 1)_+\right) + \Phi\left((Q_T - N^\mu)_+\right)$

iff
$$-\psi(\mu)\mathbb{P}(N^\mu \leq Q_T - 1) + \mathbb{P}(N^\mu = Q_T - 1)\varphi(\mu) \leq 0$$

iff $\quad \mathbb{P}(N^\mu \leq Q_T - 1)\mathbb{E}\left[\Phi\left((Q_T - N^\mu - 1)_+\right) - \Phi\left((Q_T - N^\mu - 2)_+\right)\right] \leq \mathbb{P}(N^\mu \leq Q_T - 2)\varphi(\mu).$

But

$$
\begin{aligned}
\Phi\left((Q_T - N^\mu - 1)_+\right) - &\ \Phi\left((Q_T - N^\mu - 2)_+\right) \\
\leq &\ \rho_Q \left(\Phi\left((Q_T - N^\mu)_+\right) - \Phi\left((Q_T - N^\mu - 1)_+\right)\right) \mathbb{1}_{\{N^\mu \leq Q_T - 2\}} \\
= &\ \rho_Q \left(\Phi\left((Q_T - N^\mu)_+\right) - \Phi\left((Q_T - N^\mu - 1)_+\right) - (\Phi(1) - \Phi(0))\,\mathbb{1}_{\{N^\mu = Q_T - 1\}}\right) \\
\leq &\ \rho_Q \left(\Phi\left((Q_T - N^\mu)_+\right) - \Phi\left((Q_T - N^\mu - 1)_+\right)\right)
\end{aligned}
$$

since
$$(\Phi(1) - \Phi(0))\,\mathbb{1}_{\{N^\mu = Q_T - 1\}} \geq 0 \quad a.s.$$

Consequently

$$
\begin{aligned}
&\ \mathbb{P}(N^\mu \leq Q_T - 1)\mathbb{E}\left[\Phi\left((Q_T - N^\mu - 1)_+\right) - \Phi\left((Q_T - N^\mu - 2)_+\right)\right] - \mathbb{P}(N^\mu \leq Q_T - 2)\varphi(\mu) \\
\leq &\ (\rho_Q \mathbb{P}(N^\mu \leq Q_T - 1) - \mathbb{P}(N^\mu \leq Q_T - 2))\varphi(\mu) \\
= &\ \left(\rho_Q - \left(1 - \frac{\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\right)\right)\mathbb{P}(N^\mu \leq Q_T - 1)\varphi(\mu) \leq 0 \quad \text{if} \quad \rho_Q \leq 1 - \frac{\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \leq Q_T - 1)}. \quad \square
\end{aligned}
$$

Proof of Proposition 4.2. By using the notation (2.9)-(2.10), we obtain the following bound for the second derivative of the cost function

$$
\begin{aligned}
C''(\delta) \geq &\ \mathbb{E}\Big[2\underline{k}_1 \Lambda_T(\delta, S)\mathbb{P}^{(\delta)}(N^\mu \leq Q_T - 1) \\
&\ + (S_0 - \delta)\underline{k}_1 \underline{k}_2 \Lambda_T(\delta, S)\left(\mathbb{P}^{(\delta)}(N^\mu \leq Q_T - 1) - \frac{\overline{k}_1^2}{\underline{k}_1 \underline{k}_2}\Lambda_T(\delta, S)\mathbb{P}^{(\delta)}(N^\mu = Q_T - 1)\right) \\
&\ - \kappa S_T \Lambda_T(\delta, S)\left(\overline{k}_1 \overline{k}_2 \phi^{(\delta)}(\mu) - \underline{k}_1^2 \Lambda_T(\delta, S)\psi^{(\delta)}(\mu)\right)\Big].
\end{aligned}
$$

By adapting the result of Lemma 4.1, we obtain, if $Q_T \geq \left(1 + \frac{\overline{k}_1^2}{\underline{k}_1 \underline{k}_2}\right)T\lambda(-S_0)$, that

$$\mathbb{E}\left[\left(\mathbb{P}(N^\mu \leq Q_T - 1) - \frac{\overline{k}_1^2}{\underline{k}_1 \underline{k}_2}\mu\mathbb{P}(N^\mu = Q_T - 1)\right)_{|\mu=\Lambda_T(\delta, S)}\right] \geq 0$$

and by convexity of the penalty function Φ, we have $\psi^{(\delta)}(\mu) \geq 0$ $a.s.$ Then we obtain the following upper bound for κ,

$$\kappa \leq \frac{2\underline{k}_1 \mathbb{E}\left[(\mu\mathbb{P}(N^\mu \leq Q_T - 1))_{|\mu=\Lambda_T(\delta, S)}\right]}{\overline{k}_1 \overline{k}_2 \mathbb{E}\left[S_T \Lambda_T(\delta, S)\varphi^{(\delta)}(\mu)\right]}. \tag{4.21}$$

By Lemma 4.3, $\mu \mapsto \frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}$ is non-increasing and by Lemma 4.1 $\mu \mapsto \mu \mathbb{P}(N^\mu \leq Q_T - 1)$ is non-decreasing for $Q_T \geq 2\lfloor \mu \rfloor - 1$. Furthermore $\alpha \mapsto \Lambda_T(\delta, \alpha)$ is non-increasing. By applying the functional co-monotony principle, we then have, for $Q_T \geq 2T\lambda(-S_0)$, that

$$\mathbb{E}\left[S_T \Lambda_T(\delta, S)\varphi^{(\delta)}(\mu)\right] \leq \mathbb{E}\left[(\mu \mathbb{P}(N^\mu \leq Q_T - 1))_{|\mu = \Lambda_T(\delta, S)}\right] \mathbb{E}\left[\left(\frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\right)_{|\mu = \Lambda_T(\delta, S)}\right].$$

Therefore (4.21) will be satisfied as soon as

$$\kappa \leq \frac{2\underline{k}_1}{\overline{k}_1 \overline{k}_2 \mathbb{E}\left[S_T \left(\frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\right)_{|\mu = \Lambda_T(\delta, S)}\right]}.$$

Since by Lemma 4.3, $\mu \mapsto \frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}$ is non-increasing, we get

$$\left(\frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\right)_{|\mu = \Lambda_T(\delta, S)} \leq \left(\frac{\varphi(\mu)}{\mathbb{P}(N^\mu \leq Q_T - 1)}\right)_{|\mu = 0} = \Phi(Q_+) - \Phi((Q - 1)_+) \leq \Phi'_\ell(Q),$$

which finally yields the announced (more stringent) criterion

$$\kappa \leq \frac{2\underline{k}_1}{\overline{k}_1 \overline{k}_2 \mathbb{E}[S_T]\Phi'(Q)}. \qquad \square$$

Remark. As $\delta \in [0, \delta_{\max}]$, then $(S_0 - \delta) \in [S_0 - \delta_{\max}, S_0]$ and

$$(S_0 - \delta)\underline{k}_1\underline{k}_2\Lambda_T(\delta, S)\left(\mathbb{P}^{(\delta)}(N^\mu \leq Q_T - 1) - \frac{\overline{k}_1^2}{\underline{k}_1\underline{k}_2}\Lambda_T(\delta, S)\mathbb{P}^{(\delta)}(N^\mu = Q_T - 1)\right)$$

$$\geq (S_0 - \delta_{\max})\underline{k}_1\underline{k}_2\Lambda_T(\delta, S)\left(\mathbb{P}^{(\delta)}(N^\mu \leq Q_T - 1) - \frac{\overline{k}_1^2}{\underline{k}_1\underline{k}_2}\Lambda_T(\delta, S)\mathbb{P}^{(\delta)}(N^\mu = Q_T - 1)\right)$$

$$= (S_0 - \delta_{\max})\underline{k}_1\underline{k}_2\left[\mu\left(\mathbb{P}(N^\mu \leq Q_T - 1) - \frac{\overline{k}_1^2}{\underline{k}_1\underline{k}_2}\mu\mathbb{P}(N^\mu = Q_T - 1)\right)\right]_{|\mu = \Lambda_T(\delta, S)}.$$

Unfortunately we cannot use the functional co-monotony principle to improve the bound because, for $Q \geq \left(2 \vee \left(1 + \frac{\overline{k}_1^2}{\underline{k}_1\underline{k}_2}\right)\right)T\lambda(-S_0)$, the function $\mu \mapsto \mu\mathbb{P}(N^\mu \leq Q_T - 1)$ is non-decreasing and $\mu \mapsto 1 - \frac{\mu\mathbb{P}(N^\mu = Q_T - 1)}{\mathbb{P}(N^\mu \leq Q_T - 1)}$ is non-increasing, and we need to obtain a lower bound for this expression but co-monotony naturally yields an upper bound.

5 Numerical experiments

In this section, we present numerical results with simulated and real data. We first present the chosen model for the price dynamic and the penalization function. Within the numerical examples, we are modeling the optimal behavior of a "learning trader" reassessing the price of his passive order every 5 units of time (can be seconds or minutes) in the order books to adapt to the characteristics of the market (fair price moves S_t and order flow dynamics N_t). During each period of 5 seconds, the trader posts her order of size Q_5 in the book at a distance δ of the best opposite price (δ lower than the best ask for a buy order), and waits 5 seconds. If the order is not completely filled after these 5 seconds (say at time T), the trader cancels the remaining quantity $(Q_5 - N_5)_+$ and buys it using a market order

at price S_T plus a market impact; she will buy at $\kappa S_T \left(1 + \eta((Q_5 - N_5)_+)\right)$. Then she can reproduce the experiment choosing another value for the distance to the best opposite δ.

The reassessment procedure used here is the one of formula (2.4) using the expectation representation of C' given by Property 3.2 to provide the proper form for the function H.

Then we plot the cost function and its derivative for a trivial penalization function $\Phi = Id$ ($\eta \equiv 0$) and for a nontrivial one. We conclude by giving the results of the recursive procedure in both cases, either on simulated data or on a real data obtained by "replaying" the market.

5.1 Simulated data

We assume that
$$dS_t = \sigma dW_t, \quad S_0 = s_0 \quad \text{and} \quad \Lambda_T(\delta, S) = A \int_0^T e^{-k(S_t - S_0 + \delta)} dt$$

where $(W_t)_{t \geq 0}$ is a standard Brownian motion and $\sigma, A, k > 0$ (this means that $\lambda(x) = A e^{-kx}$). We denote here by $(\bar{S}_t)_{t \geq 0}$ the Euler scheme with step $\frac{T}{m}$ of $(S_t)_{t \in [0,T]}$ defined by

$$\bar{S}_{k+1} := \bar{S}_k + \sigma \sqrt{\frac{T}{m}} Z_{k+1}, \quad \bar{S}_0 = s_0, \quad Z_{k+1} \sim \mathcal{N}(0,1), \quad k \geq 0,$$

and we approximate $\Lambda_T(\delta, S)$ by $\bar{\Lambda}_T(\delta, S) = A \frac{T}{m} \sum_{k=0}^m e^{-k(\bar{S}_k - S_0 + \delta)}$. The market impact penalization function is $\Phi(x) = (1 + \eta(x))x$ with $\eta(x) = A' e^{k'x}$. Now we present the cost function and its derivative for the following parameters:

- parameters of the asset dynamics: $s_0 = 100$ and $\sigma = 0.01$,

- parameters of the intensity of the execution process: $A = 5$ and $k = 1$,

- parameters of the execution: $T = 5$ and $Q = 10$,

- parameters of the penalization function: $\kappa = 1$, $A' = 0.1$ and $k' = 0.05$.

We use $m = 20$ for the Euler scheme and $M = 10000$ Monte Carlo simulations.

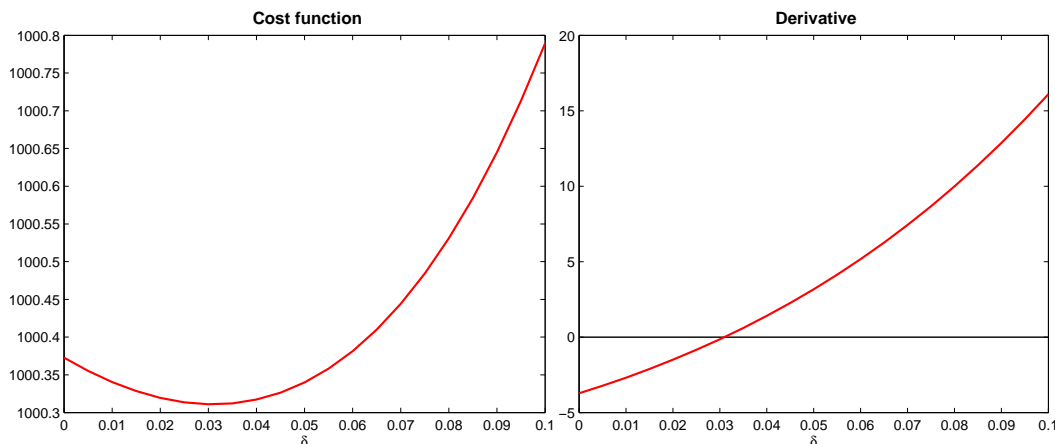**_Setting 1_** ($\eta \not\equiv 0$)

Figure 2: $\eta \not\equiv 0$: $T = 5$, $A = 5$, $k = 1$, $s_0 = 100$, $\sigma = 0.01$, $Q = 10$, $\kappa = 6$, $A' = 1$, $k' = 0.01$, $m = 20$ and $M = 10000$.

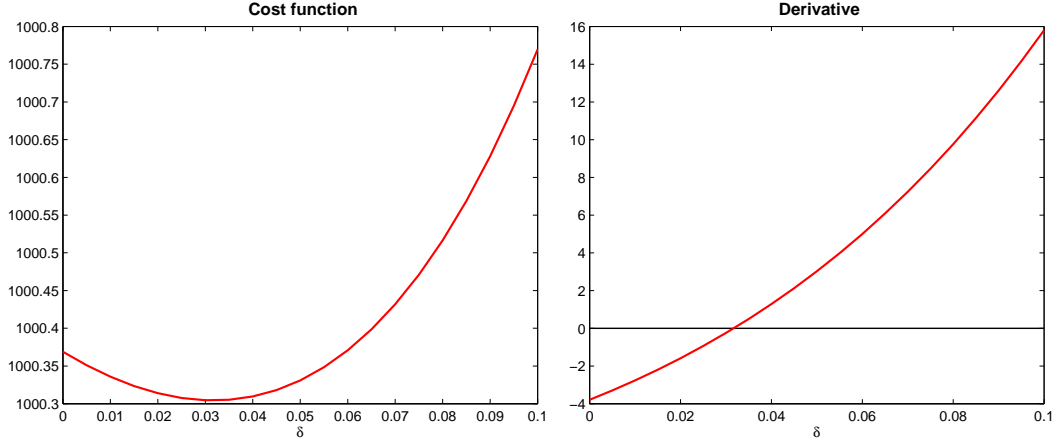<u>**Setting 2**</u> $(\eta \equiv 0)$

Figure 3: $\eta \equiv 0$: $T = 5$, $A = 5$, $k = 1$, $s_0 = 100$, $\sigma = 0.01$, $Q = 10$, $\kappa = 12$, $m = 20$ and $M = 10000$.

Remark. When one looks at the cost functions in Figures 2 and 3 (left), one may think that it would be simpler to compute the cost functions to derive the minimum and the associated optimal distance. But the computing time of the costs is about 100 seconds which is too large compared to the length of the posting period $T = 5$ seconds, whereas that of the stochastic recursive procedure is about 1 second.

Now we present the results of the stochastic recursive procedure for the two settings with

$$n = 100 \quad \text{and} \quad \gamma_n = \frac{1}{100n}.$$
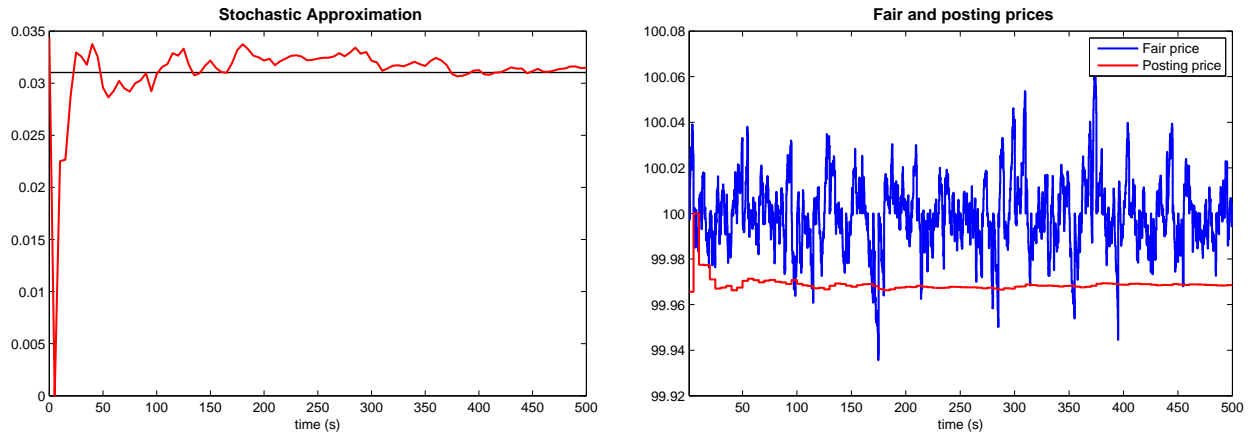
<u>**Setting 1**</u> $(\eta \not\equiv 0)$

Figure 4: $\eta \not\equiv 0$: $T = 5$, $A = 5$, $k = 1$, $s_0 = 100$, $\sigma = 0.01$, $Q = 10$, $\kappa = 6$, $A' = 1$, $k' = 0.01$, $m = 20$ and $n = 100$
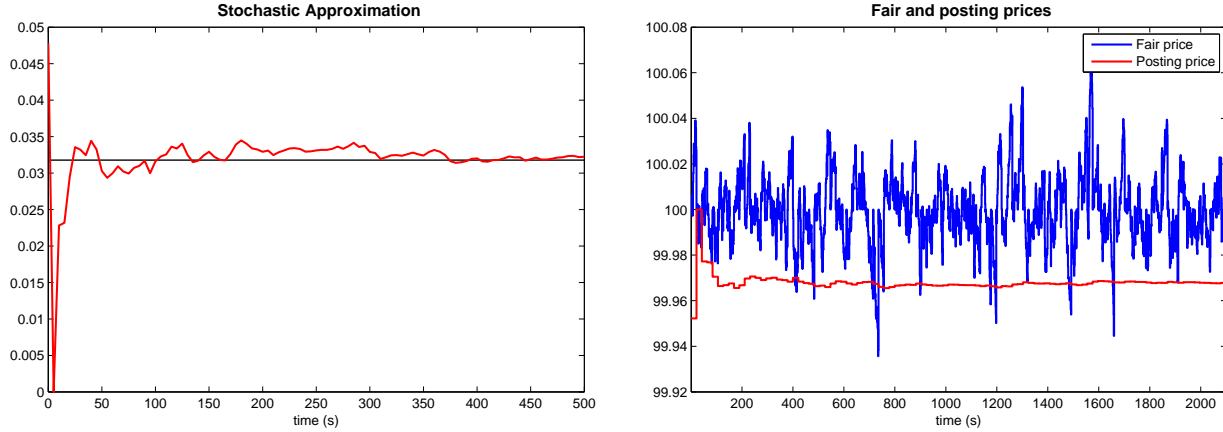
Figure 5: $\eta \equiv 0$: $T = 5$, $A = 5$, $k = 1$, $s_0 = 100$, $\sigma = 0.01$, $Q = 10$, $\kappa = 12$, $m = 20$ and $n = 100$.

5.2 Market data

The self-adaptiveness nature of this recurrence procedure allows to implement it on real data, even if they are not exactly fulfilling the model assumptions. In the numerical example of this section, the trader reassess his order using the previously exposed recurrence procedure on real data on which the parameters of the models (k, A, κ, k', A') have been beforehand fitted.

As market data, we use the bid prices of Accor SA (ACCP.PA) of 11/11/2010 for the fair price process $(S_t)_{t \in [0,T]}$. We divide the day into periods of 15 trades which will denote the steps of the stochastic procedure. Let N_{cycles} be the number of these periods. For every $n \in N_{\text{cycles}}$, we have a sequence of bid prices $(S_{t_i}^{(n)})_{1 \leq i \leq 15}$ and we approximate the jump intensity of the Poisson process $\Lambda_{T^n}(\delta, S)$, where $T^n = \sum_{i=1}^{15} t_i$, by

$$\forall n \in \{1, \ldots, N_{\text{cycles}}\}, \quad \Lambda_{T^n}(\delta, S) = A \sum_{i=2}^{15} e^{-k(S_{t_i}^{(n)} - S_{t_1} + \delta)}(t_i - t_{i-1}).$$

The empirical mean of the intensity function

$$\bar{\Lambda}(\delta, S) = \frac{1}{N_{\text{cycles}}} \sum_{n=1}^{N_{\text{cycles}}} \Lambda_{T^n}(\delta, S)$$
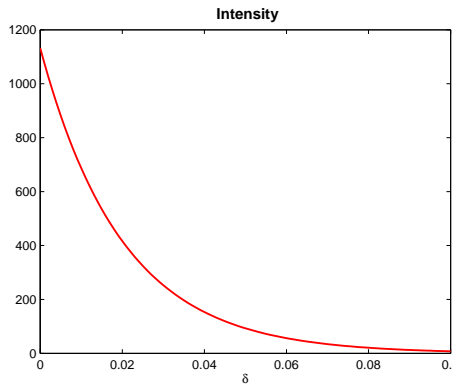
is plotted on Figure 6.

Figure 6: Fit of the exponential model on real data (Accor SA (ACCP.PA) 11/11/2010): $A = 1/50$, $k = 50$ and $N_{\text{cycles}} = 220$.

23

The penalization function has the following form

$$\Phi(x) = (1 + \eta(x))x \quad \text{with} \quad \eta(x) = A'e^{k'x}.$$

Now we present the cost function and its derivative for the following parameter values: $A = 1/50$, $k = 50$, $Q = 100$, $A' = 0.001$ and $k' = 0.0005$.
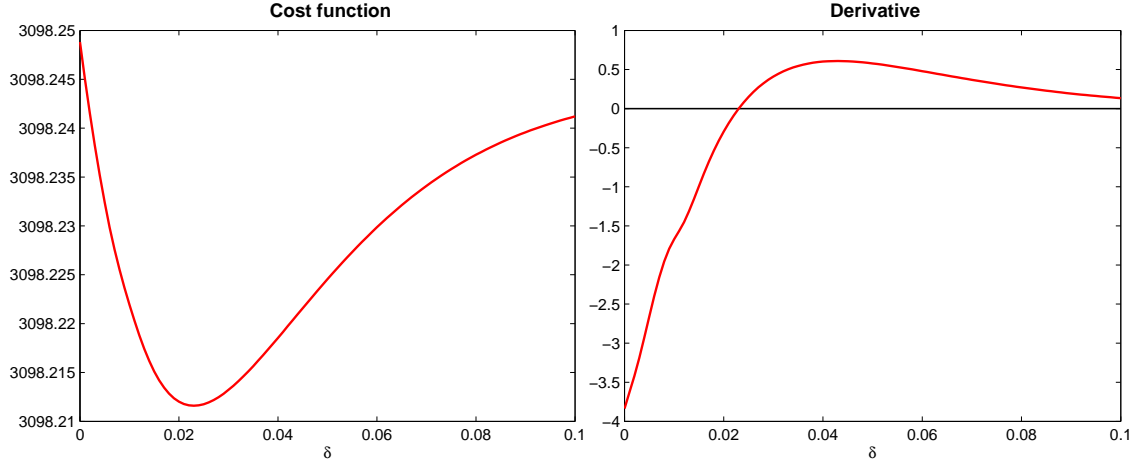
Setting1 ($\eta \not\equiv 0$)

Figure 7: $\eta \not\equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1$, $A' = 0.001$, $k' = 0.0005$ and $N_{\text{cycles}} = 220$.
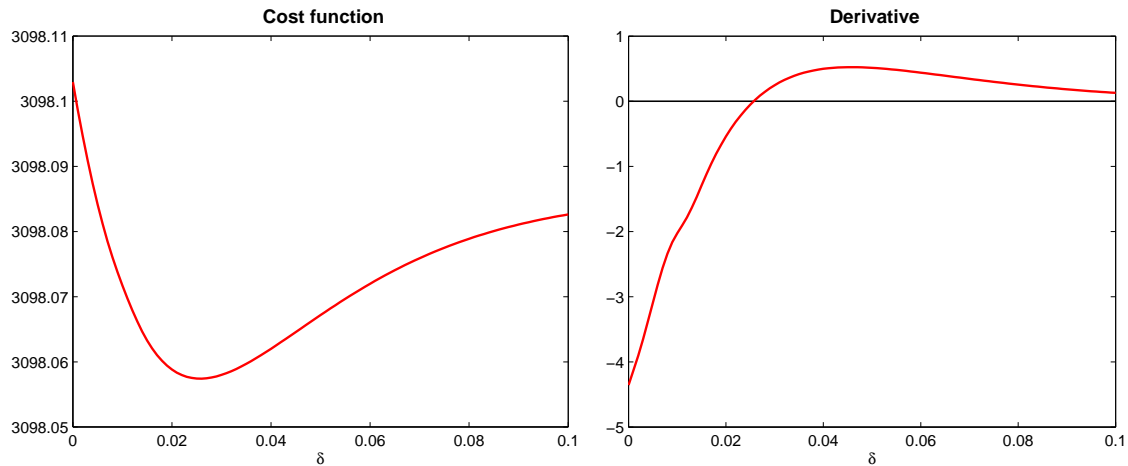
Setting2 ($\eta \equiv 0$)

Figure 8: $\eta \equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1.001$ and $N_{\text{cycles}} = 220$.

Now we present the results of the stochastic recursive procedure for two cases. To smoothen the behavior of the stochastic algorithm, we use Ruppert and Poliak's averaging principle (see [8]). In short, this principle is two-folded:

– *Phase 1*: Implement the original zero search procedure with $\gamma_n = \frac{\gamma_1}{n^\rho}$, $\frac{1}{2} < \rho < 1$, $\gamma_1 > 0$,

– *Phase 2*: Compute the arithmetic mean at each step n of all the past values of the procedure, namely

$$\bar{\delta}_n = \frac{1}{n+1}\sum_{k=0}^{n} \delta_k, \ n \geq 1.$$

24

It has been shown by several authors that this procedure under appropriate assumptions is ruled by a CLT having a minimal asymptotic variance (among recursive procedures).
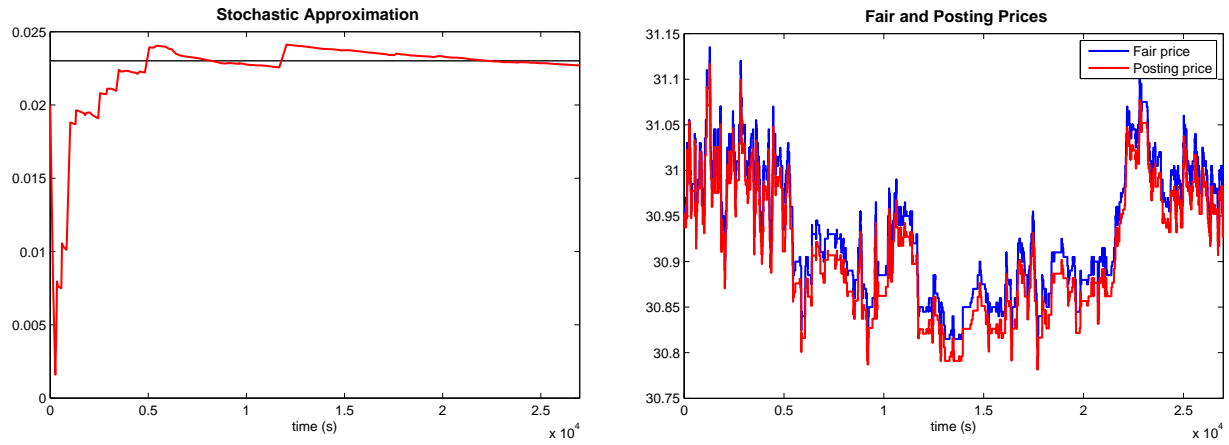
Setting 1 $(\eta \not\equiv 0)$

Figure 9: $\eta \not\equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1$, $A' = 0.001$, $k' = 0.0005$ and $N_{\text{cycles}} = 220$. Crude algorithm with $\gamma_n = \frac{1}{550n}$.
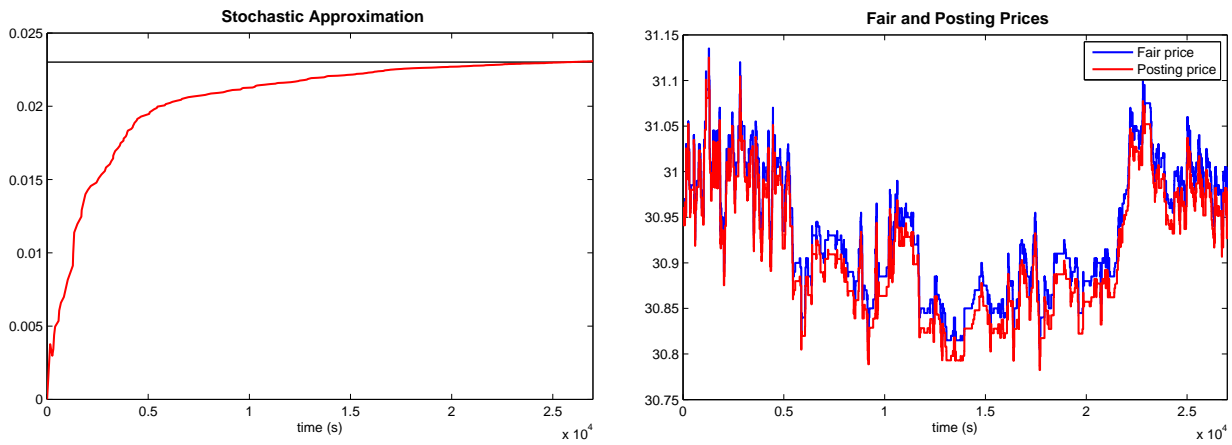

Figure 10: $\eta \not\equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1$, $A' = 0.001$, $k' = 0.0005$ and $N_{\text{cycles}} = 220$. Ruppert and Poliak's averaging algorithm with $\gamma_n = \frac{1}{550n^{0.95}}$.
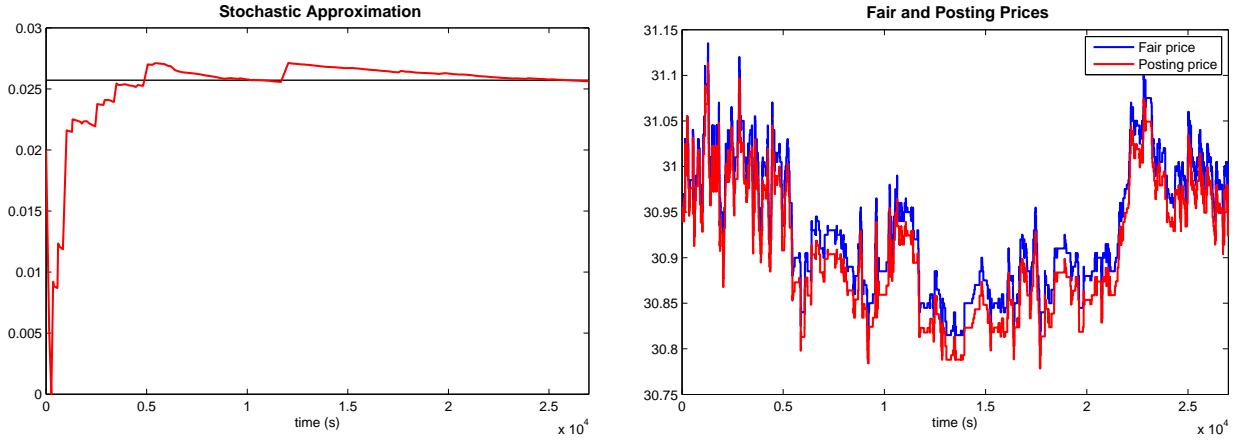
Setting 2 $(\eta \equiv 0)$

Figure 11: $\eta \equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1.001$, $\gamma_n = \frac{1}{450n}$ and $N_{\text{cycles}} = 220$.
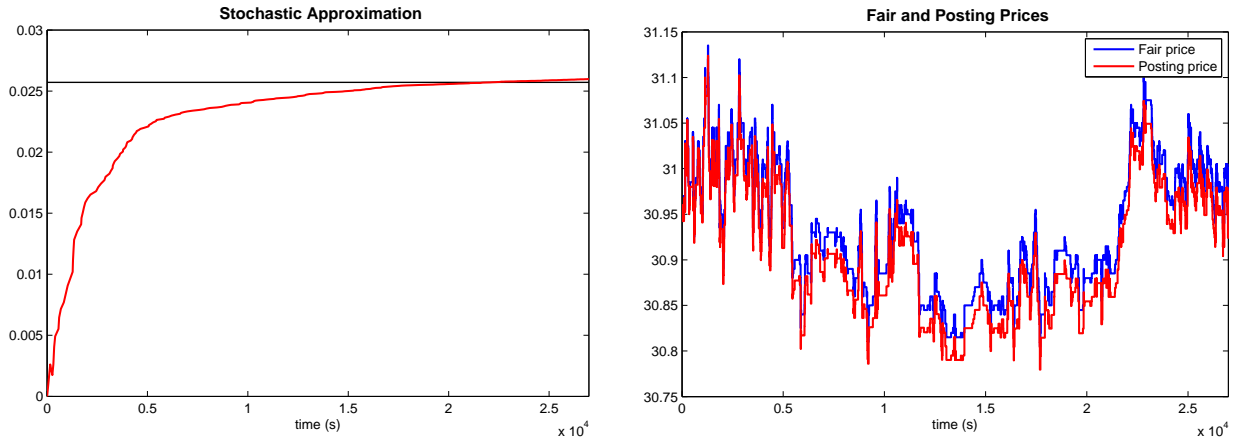

Figure 12: $\eta \not\equiv 0$: $A = 1/50$, $k = 50$, $Q = 100$, $\kappa = 1.001$ and $N_{\text{cycles}} = 220$. Ruppert and Poliak's averaging algorithm with $\gamma_n = \frac{1}{450n^{0.95}}$.

We see on Figures 10 (for $\eta \not\equiv 0$) and 12 (for $\eta \equiv 0$) that the recursive procedures converge toward their respective targets, namely the minimum of the execution cost functions presented in Figures 7 left (for $\eta \not\equiv 0$) and 8 left (for $\eta \equiv 0$).

5.3 Comparisons and comment on the efficiency of the reassessment rules

Practically, the stochastic algorithm defined by the dynamics of equation (2.4) has to be read as a reassessment rule to apply to the distance to the reference price (which can be taken as the best opposite) at a given frequency (that can be expressed in calendar time or business time, *i.e.* number of trades or traded quantities): each cycle (of given 5 seconds), the trading algorithm updates δ_n and modifies the price of its limit order to be the best opposite price minus δ_n (for a buy order).

The change from δ_{n-1} to δ_n is computed using the expression (3.14) for H, and a typical choice for the step is $\gamma_n = \gamma_0/n^\rho$ $(0 < \rho < 1)$.

Comparing the results obtained on simulated data (subsection 5.1) to the ones obtained on real data (subsection 5.2), the first point to notice is that the convergence is ten times slower on real data than on simulated ones (less than 10 minutes on real data *vs* less than one minute on simulated ones).

The second element is that once the algorithm succeeded in being close to the optimal value δ^*, it oscillates around it without any need to wait again from one to ten minutes to converge. It means

that using a heuristic rule to choose δ_0 rather than a random value can practically avoid "paying" this first step.

Another important element to highlight is that the posting price dynamics on simulated data are really different than the ones on real data. On simulated data, the price evolution diagrams are unrealistic for any trader. It mainly comes from the fact that the price does not behave like a Brownian motion at a very short time scale. On real data, the posting price adapts to it, being also far more realistic by following the last prices more closely than in a simulated environment.

It shows the adaptiveness property of a stochastic algorithm, that may need more time to catch a reasonable value for δ_n, but then keep very close to it, despite the price moves that are less smooth than a classical diffusion.

Going back to the expression of H, it must be qualitatively said that:

- the first term $-Q_T \mathbb{P}^{(\delta)} (N^\mu > Q_T)$ will push the price away when the order has been completely filled before the end of the reload period T;

- the second term $\left(\frac{\partial}{\partial \delta}\Lambda_T(\delta, S)(S_0 - \delta) - \Lambda_T(\delta, S)\right) \mathbb{P}^{(\delta)} (N^\mu \leq Q_T - 1)$ will attract the price to the best opposite as far as the price move between S_t and S_{t+T} does not change the intensity of the "order flow" $\Lambda(\delta_n, S_{t+T})$ it will be exposed to;

- the last component of the reassessment policy $-\kappa S_T \frac{\partial}{\partial \delta}\Lambda_T(\delta, S)\varphi^{(\delta)}(\mu)$ attracts the price to the best opposite when the "price impact cost" (*i.e.* coming into the equation via $\kappa\varphi^{(\delta)}$, which can bee seen as the first derivative of the market impact $\kappa\Phi$) to pay when the expected fill rate has not been obtain, is too high.

This qualitative interpretation of the update rule obtained by rigorous derivation of criteria (2.2) shows how to estimate and optimally mix these three natural effects that a trader would like to see into any price reassessment policy. It shows the value of a stochastic algorithm approach can bring to optimal trading at every time scale: (1) clearly links an objective criteria to the reassessment policies, (2) exposing the assumptions needed to guarantee its convergence. It is clearly the risk-control role that one can expect from a formalized approach of trading.

6 Conclusion

This paper presents a rigorous proof of convergence of a reassessment scheme for a trading tactic aiming at capturing liquidity "*around the book*". It implements a *learning by trading* approach, validated inside our class of model (which is quite general):

- the distance to a reference point (in practice it can be the best opposite, the mid point, or any efficient price estimate) is fixed at δ_n during few seconds or market trades,

- the tactic observes the market feedback resulting from the combination of the natural diffusion of the reference price and a point process filling my order with an intensity depending on my instantaneous distance to the reference price (which varies),

- our formal results give the optimal way to adjust the distance δ_{n+1} to the reference price, given then marginal variations of the different market components which can be anticipated to an increase or decrease of posting distance.

The robustness of the approach is not only guaranteed theoretically (provided that the market impact is in a realistic range, as commented on section 2.2.1), but also confirmed and emphasized by tests carried out on real data (partially reproduced in Section 5.2). This benchmark shows that even if the real data behave differently from Monte-Carlo generated scenarios (see Section 5.1), the convergence still occurs.

This paper strongly suggests that such iterative trading procedures, very often used by practitioners because of the way it can be efficiently fitted *on line* to real time data and providing optimal reassessment rules. This has to be compared to a stochastic control approach which needs to be calibrated on using data on longer time frames inducing somehow an "averaging" effect of the instantaneous liquidity effects that can occur on the real markets.

With the recent modifications of market microstructure following fragmentation of markets and emergence of high frequency trading, it is clear that algorithmic traders will need to devise more reactive and short term tactics. Covering this aspect of trading, this paper opens the door to other researches (like multi trading pools and multi asset reassessment of limit prices), and to applications for practitioners. Traders can use such scheme as sub-tactics of a brokerage algorithm, of an high frequency market making mechanism, or of any intraday arbitrage automated process.

Appendix

A   Convergence theorem for constrained algorithms

The aim is to determine an element of the set $\{\theta \in \Theta \ : \ h(\theta) = \mathbb{E}\left[H(\theta, Y)\right] = 0\}$ (zeros of h in Θ) where $\Theta \subset \mathbb{R}^d$ is a closed convex set, $h : \mathbb{R}^d \to \mathbb{R}^d$ and $H : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$. For $\theta_0 \in \Theta$, we consider the $\mathbb{R}^d$-valued sequence $(\theta_n)_{n \geq 0}$ defined by

$$\theta_{n+1} = \mathrm{Proj}_\Theta \left(\theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})\right), \tag{A.22}$$

where $(Y_n)_{n \geq 1}$ is an i.i.d. sequence with the same law as Y, $(\gamma_n)_{n \geq 1}$ is a positive sequence of real numbers and Proj_Θ denotes the Euclidean projection on Θ. The recursive procedure (A.22) can be rewritten as follows

$$\theta_{n+1} = \theta_n - \gamma_{n+1} h(\theta_n) - \gamma_{n+1} \Delta M_{n+1} + \gamma_{n+1} p_{n+1}, \tag{A.23}$$

where $\Delta M_{n+1} = H(\theta_n, Y_{n+1}) - h(\theta_n)$ is a martingale increment and

$$p_{n+1} = \frac{1}{\gamma_{n+1}} \mathrm{Proj}_\Theta \left(\theta_n - \gamma_{n+1} H(\theta_n, Y_{n+1})\right) - \frac{\theta_n}{\gamma_{n+1}} + H(\theta_n, Y_{n+1}).$$

Theorem A.1. *(see [18] and [19]) Let $(\theta_n)_{n \geq 0}$ be the sequence defined by (A.23). Assume that there exists a unique $\theta^* \in \Theta$ such that $h(\theta^*) = 0$ and that the mean function satisfies on Θ the following mean-reverting property, namely*

$$\forall \theta \neq \theta^* \in \Theta, \quad \langle h(\theta) \,|\, \theta - \theta^* \rangle > 0. \tag{A.24}$$

Assume that the gain parameter sequence $(\gamma_n)_{n \geq 1}$ satisfies

$$\sum_{n \geq 1} \gamma_n = +\infty \quad and \quad \sum_{n \geq 1} \gamma_n^2 < +\infty. \tag{A.25}$$

If the function H satisfies

$$\exists\, K > 0 \ \ such \ that \ \ \forall \theta \in \Theta, \quad \mathbb{E}\left[|H(\theta, Y)|^2\right] \leq K(1 + |\theta|^2), \tag{A.26}$$

then

$$\theta_n \xrightarrow[n \to +\infty]{a.s.} \theta^*.$$

Remark. If Θ is bounded (A.26) reads $\sup_{\theta \in \Theta} \mathbb{E}\left[|H(\theta, Y)|^2\right] < +\infty$, which is always satisfied if Θ is compact and $\theta \mapsto \mathbb{E}\left[|H(\theta, Y)|^2\right]$ is continuous.

B   Functional co-monotony principle for a class of one-dimensional diffusions

In this section, we present the principle of co-monotony, first for random vectors taking values in a nonempty interval I, then for one-dimensional diffusions lying in I.

B.1   Case of random variables and random vectors

First we recall a classical result for random variables.

Proposition B.1. *Let $f, g : I \subset \mathbb{R} \to \mathbb{R}$ be two monotonic functions with same monotony. Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to I$ be a real valued random variable such that $f(X), g(X) \in L^2(\mathbb{P})$. Then*

$$\mathrm{Cov}(f(X), g(X)) \geq 0.$$

Proof. Let X, Y be two independent random variables defined on the same probability space with the same distribution $\mathbb{P}_X$. Then

$$(f(X) - f(Y))(g(X) - g(Y)) \geq 0$$

hence its expectation is non-negative too. Consequently

$$\mathbb{E}\left[f(X)g(X)\right] - \mathbb{E}\left[f(X)g(Y)\right] - \mathbb{E}\left[f(Y)g(X)\right] + \mathbb{E}\left[f(Y)g(Y)\right] \geq 0$$

so, using that $Y \overset{(d)}{=} X$ and Y, X are independent, yields

$$2\mathbb{E}\left[f(X)g(X)\right] \geq \mathbb{E}\left[f(X)\right]\mathbb{E}\left[g(Y)\right] + \mathbb{E}\left[f(Y)\right]\mathbb{E}\left[g(X)\right] = 2\mathbb{E}\left[f(X)\right]\mathbb{E}\left[g(X)\right]$$

that is $\mathrm{Cov}(f(X), g(X)) \geq 0$. $\qquad\qquad\square$

Proposition B.2. *Let $F, G : \mathbb{R}^d \to \mathbb{R}$ be two monotonic functions with same monotony in each of their variables, i.e. for every $i \in \{1, \ldots, d\}$, $x_i \longmapsto F(x_1, \ldots, x_i, \ldots, x_d)$ and $x_i \longmapsto G(x_1, \ldots, x_i, \ldots, x_d)$ are monotonic with the same monotony which may depend on i (but does not depend on $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$). Let $X_1, \ldots, X_d$ be independent real valued random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that $F(X_1, \ldots, X_d), G(X_1, \ldots, X_d) \in L^2(\mathbb{P})$. Then*

$$\mathrm{Cov}\left(F(X_1, \ldots, X_d), G(X_1, \ldots, X_d)\right) \geq 0.$$

Proof. The proof of the above proposition is made by induction on d. The case $d = 1$ is given by Proposition B.1. We give here the proof for $d = 2$ for notational convenience, but the general case of dimension d follows likewise. By the monotonic assumption on F and G, we have for every $x_2 \in \mathbb{R}$, if $X_1' \overset{d}{=} X_1$ with X_1', X_1 independent, that

$$\left(F(X_1, x_2) - F(X_1', x_2)\right)\left(G(X_1, x_2) - G(X_1', x_2)\right) \geq 0.$$

This implies that (see Proposition B.1)

$$\mathrm{Cov}\left(F(X_1, x_2)G(X_1, x_2)\right) \geq 0.$$

If X_1 and X_2 are independent, using Fubini's Theorem and what precedes, we have

$$
\begin{aligned}
\mathbb{E}\left[F(X_1, X_2)G(X_1, X_2)\right] &= \int_{\mathbb{R}} \mathbb{P}_{X_2}(dx_2)\mathbb{E}\left[F(X_1, x_2)G(X_1, x_2)\right] \\
&\geq \int_{\mathbb{R}} \mathbb{P}_{X_2}(dx_2)\mathbb{E}\left[F(X_1, x_2)\right]\mathbb{E}\left[G(X_1, x_2)\right].
\end{aligned}
$$

By setting $\varphi(x_2) = \mathbb{E}\left[F(X_1, x_2)\right]$ and $\psi(x_2) = \mathbb{E}\left[G(X_1, x_2)\right]$ and using the monotonic assumptions on F and G, we have that φ and ψ are monotonic with the same monotony so that

$$\int_{\mathbb{R}} \mathbb{P}_{X_2}(dx_2)\mathbb{E}\left[F(X_1, x_2)\right]\mathbb{E}\left[G(X_1, x_2)\right] = \mathbb{E}\left[\varphi(X_2)\psi(X_2)\right] \geq \mathbb{E}\left[\varphi(X_2)\right]\mathbb{E}\left[\psi(X_2)\right].$$

Combining these above two inequalities finally yields $\text{Cov}\left(F(X_1, X_2)G(X_1, X_2)\right) \geq 0$. $\qquad\square$

B.2 Case of (one-dimensional) diffusions

This framework corresponds to the infinite dimensional case and we can not apply straightforwardly the result of Proposition B.1: indeed, if we define the following natural order relation on $\mathbb{D}([0, T], \mathbb{R})$

$$\forall \alpha_1, \alpha_2 \in \mathbb{D}([0, T], \mathbb{R}), \quad \alpha_1 \leq \alpha_2 \Longleftrightarrow (\forall t \in [0, T], \ \alpha_1(t) \leq \alpha_2(t)),$$

this order is partial which makes the formal proof of Proposition B.1 collapse. To establish a co-monotony principle for diffusions, we proceed in two steps: first, we use the Lamperti transform to "force" the diffusion coefficient to be equal to 1 and we establish the co-monotony principle for this class of diffusions. Then by the inverse Lamperti transform, we go back to the original process.

In this section, we first present our framework in more details. Then we recall some weak convergence results for diffusion with diffusion coefficient equal to 1. Afterwards we present the Lamperti transform and we conclude by the general co-monotony principle.

Let I be a nonempty open interval of $\mathbb{R}$. One considers a real-valued Brownian diffusion process

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \in I, \quad t \in [0, T], \tag{B.27}$$

where $b, \sigma : [0, T] \times I \to \mathbb{R}$ are Borel functions with at most linear growth such that the above Equation (B.27) admits at least one (weak) solution over $[0, T]$ and W is a Brownian motion defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that the diffusion X *a.s.* does not explode and lives in the interval I. This implies assumptions on the function b and σ especially in the neighborhood (in I) of the endpoints of I that we will not detail here. At a finite endpoint of I, these assumptions are strongly connected with the Feller classification for which we refer to [17] (with $\sigma(t, \cdot) > 0$ for every $t \in [0, T]$). We will simply make the classical linear growth assumption on b and σ (which prevents explosion at a finite time) that will be used for different purpose in what follows.

To "remove" the diffusion coefficient of the diffusion X, we will introduce the so-called *Lamperti transform* which requires additional assumptions on the drift b and the diffusion coefficient σ, namely

$$(\mathcal{A}_{b,\sigma}) \equiv \begin{cases} (i) & \sigma \in \mathcal{C}^1([0, T] \times I, \mathbb{R}), \\[2mm] (ii) & \forall (t, x) \in [0, T] \times I, \quad |b(t, x)| \leq C(1 + |x|) \quad \text{and} \quad 0 < \sigma(t, x) \leq C(1 + |x|), \\[2mm] (iii) & \forall x \in I, \quad \int_{(-\infty, x] \cap I} \frac{d\xi}{\sigma(t, \xi)} = \int_{[x, +\infty) \cap I} \frac{d\xi}{\sigma(t, \xi)} = +\infty \end{cases} \tag{B.28}$$

Remark. Condition (iii) clearly does not depend on $x \in I$. Furthermore, if $I = \mathbb{R}$, (iii) follows from (ii) since $\frac{1}{\sigma(t, \xi)} \geq \frac{1}{C}\frac{1}{1+|\xi|}$.

Before passing to a short background on the Lamperti transform which will lead to the new diffusion deduced from (B.27) whose diffusion coefficient is equal to 1, we need to recall (and adapt) some background on solution and discretization of such SDE.

B.2.1 Background on diffusions with $\sigma \equiv 1$ (weak solution, discretization).

The following proposition gives a condition on the drift for the existence and the uniqueness of a weak solution of a SDE when $\sigma \equiv 1$ (see [16] Proposition 3.6, Chap. 5, p. 303 and Corollary 3.11, Chap. 5, p. 305).

Proposition B.3. *Consider the stochastic differential equation*

$$dY_t = \beta(t, Y_t)dt + dW_t, \quad t \in [0, T], \tag{B.29}$$

where T is a fixed positive number, W is a one-dimensional Brownian motion and $\beta : [0, T] \times \mathbb{R} \to \mathbb{R}$ is a Borel-measurable function satisfying

$$|\beta(t, y)| \leq K(1 + |y|), \quad t \in [0, T], \quad y \in \mathbb{R}, \quad K > 0.$$

For any probability measure ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, equation (B.29) has a weak solution with initial distribution ν.

If, furthermore, the drift term β satisfies one of the following conditions:

(i) β is bounded on $[0, T] \times \mathbb{R}$,

(ii) β is continuous, locally Lipschitz in $y \in \mathbb{R}$ uniformly in $t \in [0, T]$,

then this weak solution is unique (in fact (ii) is a strong uniqueness assumption).

Now we introduce the stepwise constant (Brownian) Euler scheme $\bar{Y}^m = \left(\bar{Y}_{\frac{kT}{m}}\right)_{0 \leq k \leq m}$ with step $\frac{T}{m}$ of the process $Y = (Y_t)_{t \in [0,T]}$ defined by (B.29). It is defined by

$$\bar{Y}_{t_{k+1}^m} = \bar{Y}_{t_k^m} + \beta(t_k^m, \bar{Y}_{t_k^m})\frac{T}{m} + \sqrt{\frac{T}{m}}U_{k+1}, \quad \bar{Y}_0 = Y_0 = y_0, \quad k = 0, \ldots, m-1, \tag{B.30}$$

where $t_k^m = \frac{kT}{m}$, $k = 0, \ldots, m$, and $(U_k)_{0 \leq k \leq m}$ denotes a sequence of i.i.d. $\mathcal{N}(0, 1)$-distributed random variables given by

$$U_k = \sqrt{\frac{m}{T}}\left(W_{t_k^m} - W_{t_{k-1}^m}\right), \qquad k = 1, \ldots, m.$$

The following theorem gives a weak convergence result for the stepwise constant Euler scheme (B.30). Its proof is a straightforward consequence of the functional limit theorems for semi-martingales (to be precise Theorem 3.39, Chap. IX, p. 551 in [15]).

Theorem B.1. *Let $\beta : [0, T] \times \mathbb{R} \to \mathbb{R}$ be a continuous function satisfying*

$$\exists K > 0, \quad |\beta(t, y)| \leq K(1 + |y|), \quad t \in [0, T], \quad y \in \mathbb{R}.$$

Assume that the weak solution of equation (B.29) is unique. Then, the stepwise constant Euler scheme of (B.29) with step $\frac{T}{m}$ satisfies

$$\bar{Y}^m \xrightarrow{\mathcal{L}} Y \quad \text{for the Skorokhod topology as } m \to \infty.$$

In particular, for every functional $F : \mathbb{D}([0, T], \mathbb{R}) \to \mathbb{R}$ $\mathbb{P}_Y$-a.s. continuous at $\alpha \in \mathcal{C}([0, T], \mathbb{R})$, with polynomial growth, we have

$$\mathbb{E}F(\bar{Y}^m) \xrightarrow[m \to \infty]{} \mathbb{E}F(Y)$$

(by uniform integrability since $\sup_{t \in [0,T]} |\bar{Y}_t^m| \in \bigcap_{p > 0} L^p$).

B.2.2 Background on the Lamperti transform

We will introduce a new diffusion $Y_t := L(t, X_t)$ which will satisfy a new SDE whose diffusion coefficient will be constant equal to 1. This function L defined on $[0, T] \times I$ is known in the literature as the Lamperti transform. It is defined for every $(t, x) \in [0, T] \times I$ by

$$L(t, x) := \int_{x_1}^{x} \frac{d\xi}{\sigma(t, \xi)} \tag{B.31}$$

where x_1 is an arbitrary fixed value lying in I. The Lamperti transform clearly depends on the choice of x_1 in I but not its properties of interest. First, under $(\mathcal{A}_{b,\sigma})$-(i)-(ii), $L \in \mathcal{C}^{1,2}([0, T] \times I)$ with

$$\frac{\partial L}{\partial t}(t, x) = -\int_{x_1}^{x} \frac{1}{\sigma^2(t, \xi)} \frac{\partial \sigma}{\partial t}(t, \xi) d\xi, \quad \frac{\partial L}{\partial x}(t, x) = \frac{1}{\sigma(t, x)} > 0 \quad \text{and} \quad \frac{\partial^2 L}{\partial x^2}(t, x) = -\frac{1}{\sigma^2(t, x)} \frac{\partial \sigma}{\partial x}(t, x).$$

Let $t \in [0, T]$, $L(t, \cdot)$ is an increasing $\mathcal{C}^2$-diffeomorphism from I onto $\mathbb{R} = L(t, I)$ (the last claim follows from $(\mathcal{A}_{b,\sigma})$-(iii)). Its inverse will be denoted $L^{-1}(t, \cdot)$.

Notice that, $(t, y) \mapsto L^{-1}(t, y)$ is continuous on $[0, T] \times I$ since both sets

$$\left\{ (t, y) \in [0, T] \times I \ : \ L^{-1}(t, y) \leq c \right\} = \left\{ (t, y) \in [0, T] \times \mathbb{R} \ : \ L(t, c) \geq y \right\}$$

and

$$\left\{ (t, y) \in [0, T] \times I \ : \ L^{-1}(t, y) \geq c \right\} = \left\{ (t, y) \in [0, T] \times \mathbb{R} \ : \ L(t, c) \leq y \right\}$$

are both closed for every $c \in \mathbb{R}$. Therefore, if $(\mathcal{A}_{b,\sigma})$ holds, the function $\beta : [0, T] \times I \mapsto \mathbb{R}$ defined by

$$\beta(t, y) := \left(\frac{b}{\sigma} - \int_{x_1}^{\cdot} \frac{1}{\sigma^2(t, \xi)} \frac{\partial \sigma}{\partial t}(t, \xi) d\xi - \frac{1}{2} \frac{\partial \sigma}{\partial x} \right)(t, L^{-1}(t, y)) \tag{B.32}$$

is a Borel function, continuous as soon as b is.

Now, we set
$$\forall t \in [0, T], \quad Y_t := L(t, X_t).$$

Itô formula straightforwardly yields

$$dY_t = \beta(t, Y_t)dt + dW_t, \quad Y_0 = L(0, x_0) =: y_0 \in \mathbb{R}. \tag{B.33}$$

Remarks. • In the homogeneous case, which is the most important case for our applications,

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}, \quad t \in [0, T], \tag{B.34}$$

we have

$$L(t, x) = L(x) := \int_{x_1}^{x} \frac{d\xi}{\sigma(\xi)}.$$

Then by setting $Y_t := L(X_t)$, we obtain

$$dY_t = \beta(Y_t)dt + dW_t, \quad Y_0 = L(x_0) =: y_0 \quad \text{with} \quad \beta := \left(\frac{b}{\sigma} - \frac{\sigma'}{2} \right) \circ L^{-1}.$$

Note that β is bounded as soon as $\frac{b}{\sigma} - \frac{\sigma'}{2}$ is.

• If the partial derivative b'_x exists on $[0, T] \times I$, one easily checks, using $(L^{-1})'_y(t, y) = \sigma(t, L^{-1}(t, y))$, that for every $(t, y) \in [0, T] \times I$,

$$\beta'_y(t, y) = \left(b'_x - \frac{b\sigma'_x + \sigma'_t}{\sigma} - \frac{\sigma\sigma''_{x^2}}{2} \right)(t, L^{-1}(t, y)). \tag{B.35}$$

As a consequence, as soon as the function

$$b'_x - \frac{b\sigma'_x + \sigma'_t}{\sigma} - \frac{\sigma\sigma''_{x^2}}{2} \begin{cases} \text{is bounded on } [0,T] \times I, \text{ then } \beta \text{ satisfies the linear growth} \\ \text{Lipschitz assumption} \\ \geq 0, \text{ then } \beta \text{ is non-decreasing.} \end{cases} \tag{B.36}$$

Definition B.1. *The* functional Lamperti transform, *denoted* Λ, *is a functional from* $\mathcal{C}([0,T],I)$ *to* $\mathcal{C}([0,T],\mathbb{R})$ *defined by*

$$\forall\,\alpha \in \mathcal{C}([0,T],I), \quad \Lambda(\alpha) = L(\cdot,\alpha(\cdot)).$$

Proposition B.4. *If the diffusion coefficient* σ *satisfies* $(\mathcal{A}_{b,\sigma})$, *the functional Lamperti transform is an homeomorphism from* $\mathcal{C}([0,T],I)$ *onto* $\mathcal{C}([0,T],\mathbb{R})$.

Proof. Let $\alpha \in \mathcal{C}([0,T],I)$. Since σ is bounded away from 0 on the compact set $[0,T] \times \alpha([0,T])$, standard arguments based on Lebesgue domination theorem, imply that $\Lambda(\alpha) \in \mathcal{C}([0,T],\mathbb{R})$. Conversely, as $L(t,\cdot) : I \to \mathbb{R}$ is an homeomorphism for every $t \in [0,T]$, Λ admits an inverse defined by

$$\forall\,\xi \in \mathcal{C}([0,T],\mathbb{R}), \quad \Lambda^{-1}(\xi) := \big(t \mapsto L^{-1}(t,\xi(t))\big) \in \mathcal{C}([0,T],I).$$

Let U_K denote the topology of the convergence on compact sets of I on $\mathcal{C}([0,T],I)$.

$\triangleright$ U_K-*Continuity of* Λ *on* $[0,T] \times I$: If $\alpha_n \xrightarrow{U_K} \alpha_\infty$, the set $K = [0,T] \times \bigcup_{n \in \overline{\mathbb{N}}} \alpha_n([0,T])$ is a compact set included in I. Hence σ is bounded away from 0 on K so that

$$\forall\,t \in [0,T], \quad |L(t,\alpha_n(t)) - L(t,\alpha_\infty(t))| \leq \frac{1}{\inf_K \sigma}|\alpha_n(t) - \alpha_\infty(t)|$$

i.e.

$$\|\Lambda(\alpha_n) - \Lambda(\alpha_\infty)\|_\infty \leq \frac{1}{\inf_K \sigma}\|\alpha_n - \alpha_\infty\|_\infty.$$

$\triangleright$ U_K-*Continuity of* Λ^{-1} *on* $[0,T] \times I$: by using $(\mathcal{A}_{b,\sigma})$-(ii), we have for a fixed $t \in [0,T]$,

$$\forall\,x,x' \in I, \quad |L(t,x) - L(t,x')| \geq \frac{1}{C}\int_{x\wedge x'}^{x\vee x'} \frac{d\xi}{1+|\xi|} = \frac{1}{C}\big|\Phi(x) - \Phi(x')\big|,$$

where $\Phi(z) = \text{sign}(z)\log(1+|z|)$. Thus,

$$\forall\,y,y' \in \mathbb{R}, \quad \big|\Phi(L^{-1}(t,y)) - \Phi(L^{-1}(t,y'))\big| \leq C\,|y - y'|.$$

Let $(\xi_n)_{n\geq 1}$ be a sequence of functions of $\mathbb{D}([0,T],\mathbb{R})$ such that $\xi_n \xrightarrow[n\to+\infty]{U} \xi \in \mathcal{C}([0,T],\mathbb{R})$. Then, for every $t \in [0,T]$ and $n \geq 1$,

$$\big|\Phi(L^{-1}(t,\xi_n(t))) - \Phi(L^{-1}(t,0))\big| \leq C\,|\xi_n(t)| \leq C\,(\|\xi_n(t) - \xi\| + \|\xi\|) + |\Phi(x_0)| \leq C',$$

since $L^{-1}(t,0) = x_0$. Consequently, for every $t \in [0,T]$ and every $n \geq 1$, $L^{-1}(t,\xi_n(t)) \in K' := \Phi^{-1}([-C',C'])$. The set K' is compact (because the function Φ is continuous and proper ($\lim_{|z|\to\infty}|\Phi(z)| = +\infty$)). As $\inf_{K'} \Phi' > 0$, we deduce that there exists $\eta_0 > 0$ such that

$$\forall\,x,y \in I, \quad |\Phi(x) - \Phi(y)| > \eta_0|x - y|,$$

i.e.

$$\forall\,t \in [0,T], \ \forall\,u,v \in L(t,I), \quad \big|L^{-1}(t,u) - L^{-1}(t,v)\big| \leq C''\,|u - v|, \quad C'' > 0.$$

Hence, one concludes that

$$\|\Lambda^{-1}(\xi_n) - \Lambda^{-1}(\xi_\infty)\|_\infty \leq C''\|\xi_n - \xi_\infty)\|_\infty. \qquad \square$$

B.2.3   Functional co-monotony principle for diffusion

Definition B.2. *The diffusion process (B.27) is* admissible *if* $(\mathcal{A}_{b,\sigma})$ *holds and*

(i) *for every starting value $x_0 \in I$, (B.27) has a unique weak solution which lives in I up to $t = +\infty$ (see Proposition B.3 for a criteria),*

(ii) *the function β defined by*

$$\beta(t,y) := \left(\frac{b}{\sigma} - \int_{x_1}^{\cdot} \frac{1}{\sigma^2(t,\xi)} \frac{\partial \sigma}{\partial t}(t,\xi) d\xi - \frac{1}{2} \frac{\partial \sigma}{\partial x} \right)(t, L^{-1}(t,y)),$$

is continuous on $[0,T] \times \mathbb{R}$, non-decreasing in y for every $t \in [0,T]$ or Lipschitz in y uniformly in $t \in [0,T]$, and satisfies

$$\exists K > 0 \ \ such \ that \ \ |\beta(t,y)| \leq K(1 + |y|), \ t \in [0,T], \ y \in \mathbb{R}.$$

Definition B.3. *Let $F : \mathbb{D}([0,T], \mathbb{R}) \to \mathbb{R}$ be a functional.*

(i) *The functional F is non-decreasing (resp. non-increasing) on $\mathbb{D}([0,T], \mathbb{R})$ if*

$$\forall \alpha_1, \alpha_2 \in \mathbb{D}([0,T], \mathbb{R}), \quad (\forall t \in [0,T], \ \alpha_1(t) \leq \alpha_2(t)) \Rightarrow F(\alpha_1) \leq F(\alpha_2) \ (resp. \ F(\alpha_1) \geq F(\alpha_2)).$$

(ii) *The functional F is continuous at $\alpha \in \mathcal{C}([0,T], \mathbb{R})$ if*

$$\forall \alpha_m \in \mathbb{D}([0,T], \mathbb{R}), \quad \alpha_m \xrightarrow{U} \alpha \in \mathcal{C}([0,T], \mathbb{R}), \quad F(\alpha_m) \to F(\alpha).$$

where U denotes the uniform convergence of functions on $[0,T]$. The functional F is C-continuous if it is continuous at every $\alpha \in \mathcal{C}([0,T], \mathbb{R})$.

(iii) *The functional F has polynomial growth if there exists a positive real number $r > 0$ such that*

$$\forall \alpha \in \mathbb{D}([0,T], \mathbb{R}), \quad |F(\alpha)| \leq K \left(1 + \|\alpha\|_\infty^r \right). \tag{B.37}$$

Remark. *Any C-continuous functional in the above sense is in particular $\mathbb{P}_Z$-a.s. continuous for every process Z with continuous paths.*

Definition B.4. *A process $(X_t)_{t \in [0,T]}$ with* continuous *(resp. càdlàg stepwise constant) paths defined on $(\Omega, \mathcal{A}, \mathbb{P})$ satisfies a functional co-monotony principle if for every C-continuous functionals (resp. measurable functionals on $\mathbb{D}([0,T], \mathbb{R})$) F, G monotonic with the same monotony satisfying (B.37) such that $F(X)$, $G(X)$ and $F(X)G(X) \in L^1$, we have*

$$\mathrm{Cov}\left(F\big((X_t)_{t \in [0,T]} \big), G\big((X_t)_{t \in [0,T]} \big) \right) \geq 0.$$

The main result of this section is the following

Theorem B.2. *Assume that the real-valued diffusion process (B.27) is admissible (see Definition B.2). Then it satisfies a co-monotony principle.*

Corollary B.1. *Assume that the real-valued diffusion process (B.27) is admissible (see Defintion B.2).*
(a) Let $\left(\bar{X}_{t_k^m} \right)_{0 \leq k \leq m}$ be its stepwise constant Euler scheme with step $\frac{T}{m}$ ($t_k^m = \frac{kT}{m}$, $0 \leq k \leq m$). Then $\left(\bar{X}_{t_k^m} \right)_{0 \leq k \leq m}$ satisfies a co-monotony principle.
(b) Let $\left(\tilde{X}_{t_k} \right)_{0 \leq k \leq m}$ be a sample of discrete time observations of $(X_t)_{t \in [0,T]}$ for a subdivision $(t_k)_{0 \leq k \leq m}$ of $[0,T]$ ($0 = t_0 < \cdots < t_m = T$). Then $\left(\tilde{X}_{t_k} \right)_{0 \leq k \leq m}$ satisfies a co-monotony principle.

Remark. The proof of Corollary B.1 is contained in the proof of Theorem B.2. The only difference is that we do not need to transfer the co-monotony principle from the Euler scheme to the diffusion process.

Before passing to the proof of Theorem B.2, we need two lemmas: one is a key step to transfer co-monotony from the Euler scheme to the diffusion process, the other aims at transferring uniqueness property for weak solutions.

Lemma B.1. *For every $\alpha \in \mathbb{D}([0,T], \mathbb{R})$, set*

$$\alpha^{(m)} = \sum_{k=0}^{m-1} \alpha(t_k^m) \mathbb{1}_{[t_k^m, t_{k+1}^m)} + \alpha(T) \mathbb{1}_{\{T\}}, \quad m \geq 1, \tag{B.38}$$

with $t_k^m := \frac{kT}{m}$, $k = 0, \ldots, m$. Then $\alpha^{(m)} \xrightarrow{U} \alpha$ as $m \to \infty$.

If $F : \mathbb{D}([0,T], \mathbb{R}) \to \mathbb{R}$ is C-continuous and non-decreasing (resp. non-increasing), then the unique function $F_m : \mathbb{R}^{m+1} \to \mathbb{R}$ satisfying $F(\alpha^{(m)}) = F_m(\alpha(t_k^m), k = 0, \ldots, m)$ is continuous and non-decreasing (resp. non-increasing) in each of its variables. Furthermore, if F satisfies a polynomial growth assumption of the form

$$\forall \alpha \in \mathbb{D}([0,T], \mathbb{R}), \quad |F(\alpha)| \leq C(1 + \|\alpha\|_\infty^r)$$

then, for every $m \geq 1$,

$$|F_m(x_0, \ldots, x_m)| \leq C(1 + \max_{0 \leq k \leq m} |x_k|^r)$$

with the same real constant $C > 0$.

Lemma B.2. *Let (S, d), (T, δ) be two Polish spaces and let $\Phi : S \mapsto T$ be a continuous injective function. Let μ and μ' be two probability measures on $(S, \mathcal{B}or(S))$. If $\mu \circ \Phi^{-1} = \mu' \circ \Phi^{-1}$, then $\mu = \mu'$.*

Proof of Lemma B.2. For every Borel set A of S, $\mu(A) = \sup\{\mu(K), K \subset A, K \text{ compact}\}$. Let $A \in \mathcal{B}or(S)$ such that $\mu(A) \neq \mu'(A)$. Then there exists a compact set K of A such that $\mu(K) \neq \mu'(K)$. But $\Phi(K)$ is a compact set of S because Φ is continuous, so $\Phi^{-1}(\Phi(K))$ is a Borel set of S which contains K. As Φ is injective, $\Phi^{-1}(\Phi(K)) = K$. Therefore $\mu(\Phi^{-1}(\Phi(K))) \neq \mu'(\Phi^{-1}(\Phi(K)))$. We deduce that $\mu \circ \Phi^{-1} \neq \mu' \circ \Phi^{-1}$. $\qquad\square$

Proof of Theorem B.2. First we consider the Lamperti transform $(Y_t)_{t \geq 0}$ (see (B.31)) of the diffusion X solution to (B.29) with $X_0 = x_0 \in I$. Using the homeomorphism property of Λ and calling upon the above Lemma B.2 with Λ^{-1} and Λ, we see that existence and uniqueness assumptions on Equation (B.29) can be transferred to (B.33) since Λ is a one-to-one mapping between the solutions of these two SDE's.

To fulfill condition (ii) in Definition B.2, we need to introduce the smallest integer, denoted $m_{b,\sigma}$, such that $y \mapsto y + \frac{T}{m_{b,\sigma}} \beta(t, y)$ is non-decreasing in y for every $t \in [0, T]$. Its existence follows from $\mathcal{A}_{b,\sigma}$-(ii). Note that if β is non-decreasing in y for every $t \in [0, T]$, then $m_{b,\sigma} = 1$. Then we introduce the stepwise constant (Brownian) Euler scheme $\bar{Y}^m = (\bar{Y}_{\frac{kT}{m}})_{0 \leq k \leq m}$ with step $\frac{T}{m}$ (defined by (B.30)) of $Y = (Y_t)_{t \in [0,T]}$ with $m \geq m_{b,\sigma}$. It is clear by induction on k that there exists for every $k \in \{1, \ldots, m\}$ a function $\Theta_k : \mathbb{R}^{k+1} \to \mathbb{R}$ such that

$$\bar{Y}_{t_k^m} = \Theta_k(y_0, \Delta W_{t_1^m}, \ldots, \Delta W_{t_k^m})$$

where for $(y_0, z_1, \ldots, z_k) \in \mathbb{R}^{k+1}$,

$$
\begin{aligned}
\Theta_k(y_0, z_1, \ldots, z_k) &= \Theta_{k-1}(y_0, z_1, \ldots, z_{k-1}) + \beta(t_{k-1}^m, \Theta_{k-1}(y_0, z_1, \ldots, z_{k-1}))\frac{T}{m} + z_k \\
&= \left(\mathrm{id} + \beta(t_{k-1}^m, \cdot)\frac{T}{m}\right) \circ \Theta_{k-1}(y_0, z_1, \ldots, z_{k-1}) + z_k.
\end{aligned}
$$

Thus for every $i \in \{1, \ldots, k\}$, $z_i \mapsto \Theta_k(y_0, z_1, \ldots, z_i, \ldots, z_k)$ is non-decreasing because $y \mapsto \left(y + \beta(t_{k-1}^m, y)\frac{T}{m}\right)$ is non-decreasing for m large enough, say $m \geq m_{b,\sigma}$. We deduce that if $F_m : \mathbb{R}^{m+1} \to \mathbb{R}$ is non-decreasing in each variables, then, for every $i \in \{1, \ldots, k\}$,

$$
z_i \mapsto F_m\left(y_0, \Theta_1(y_0, z_1), \ldots, \Theta_m(y_0, z_1, \ldots, z_m)\right) \text{ is non-decreasing.}
$$

By the same reasoning, we deduce that for $G_m : \mathbb{R}^{m+1} \to \mathbb{R}$, non-increasing in each variables, we have for every $i \in \{1, \ldots, k\}$,

$$
z_i \mapsto G_m\left(y_0, \Theta_1(y_0, z_1), \ldots, \Theta_m(y_0, z_1, \ldots, z_m)\right) \text{ is non-increasing.}
$$

Let F_m and G_m be the functions defined on $\mathbb{R}^{m+1}$ associated to F and G respectively by Lemma B.1. As β has linear growth, Y and its Euler scheme have polynomial moments at any order $p > 0$. Then we can apply Proposition B.2 to deduce that

$$
\begin{aligned}
\mathbb{E}\left[FG\left(\bar{Y}^m\right)\right] &= \mathbb{E}\left[F_m\left(\left(\bar{Y}_{\frac{kT}{m}}\right)_{0 \leq k \leq m}\right) G_m\left(\left(\bar{Y}_{\frac{kT}{m}}\right)_{0 \leq k \leq m}\right)\right] \\
&\geq \mathbb{E}\left[F_m\left(\left(\bar{Y}_{\frac{kT}{m}}\right)_{0 \leq k \leq m}\right)\right] \mathbb{E}\left[G_m\left(\left(\bar{Y}_{\frac{kT}{m}}\right)_{0 \leq k \leq m}\right)\right] = \mathbb{E}\left[F\left(\bar{Y}^m\right)\right] \mathbb{E}\left[G\left(\bar{Y}^m\right)\right].
\end{aligned}
$$

Note that if F and G are C-continuous with polynomial growth, so is FG. We derive from Theorem B.1 that

$$
\mathbb{E}\left[FG\left(\bar{Y}^m\right)\right] \xrightarrow[m \to \infty]{} \mathbb{E}FG(Y), \quad \mathbb{E}\left[F\left(\bar{Y}^m\right)\right] \xrightarrow[m \to \infty]{} \mathbb{E}F(Y), \quad \mathbb{E}\left[G\left(\bar{Y}^m\right)\right] \xrightarrow[m \to \infty]{} \mathbb{E}G(Y),
$$

therefore

$$
\mathrm{Cov}\left(F(Y), G(Y)\right) \geq 0.
$$

To conclude the proof, we need to go back to the process X by using the inverse Lamperti transform. Indeed, for every $t \in [0, T]$, $X_t = L^{-1}(t, Y_t)$, where Y satisfies (B.33). Let $F : \mathbb{D}([0, T], \mathbb{R}) \to \mathbb{R}$ C-continuous. Set

$$
\forall \alpha \in \mathcal{C}([0, T], \mathbb{R}), \quad \widetilde{F}(\alpha) := F\left(\left(L^{-1}(t, \alpha_t)\right)_{t \in [0, T]}\right).
$$

Assume first that F and G are bounded. The functional $\widetilde{F}$ is C-continuous owing to Proposition B.4, non-decreasing (resp. non-increasing) since $L^{-1}(t, .)$ is for every $t \in [0, T]$ and is bounded. Consequently,

$$
\mathrm{Cov}\left(F(X), G(X)\right) = \mathrm{Cov}\left(\widetilde{F}(Y), \widetilde{G}(Y)\right) \geq 0.
$$

To conclude we approximate F and G in a robust way with respect to the "constraints", by a canonical truncation procedure, say

$$
F_M := \max\left((-M), \min\left(F, M\right)\right), \quad M \in \mathbb{N}.
$$

If F and G have polynomial growth, it is clear that $\mathrm{Cov}\left(F_M(X), G_M(X)\right) \to \mathrm{Cov}\left(F(X), G(X)\right)$ as $M \to \infty$. $\qquad\square$

Examples of admissible diffusions. • *The Bachelier model*: This simply means that $X_t = \mu t + \sigma W_t$, $\sigma > 0$, clearly fulfills the assumptions of Theorem B.2.

• *The Black-Scholes model*: The diffusion process X is a geometric Brownian motion, solution to the SDE

$$dX_t = rX_tdt + \vartheta X_tdW_t, \quad X_0 = x_0 > 0,$$

where $r \in \mathbb{R}$ and $\vartheta > 0$ are real numbers. The geometric Brownian motion lives in the open interval $I = (0, +\infty)$ and $\beta(y) = \frac{r}{\vartheta} - \frac{\vartheta}{2}$ is constant. One checks that $L(x) = \frac{1}{\sigma}\log\left(\frac{x}{x_1}\right)$ where $x_1 \in (0, +\infty)$ is fixed.

• *The Hull-White model*: It is an elementary improvement of the Black-Scholes model where $\vartheta : [0, T] \rightarrow (0, +\infty)$ is a deterministic positive function *i.e.* the diffusion process X is a geometric Brownian motion solution of the SDE

$$dX_t = rX_tdt + \vartheta(t)X_tdW_t, \quad X_0 = x_0 > 0.$$

Then, elementary stochastic calculus shows that

$$X_t = x_0 e^{rt - \frac{1}{2}\int_0^t \vartheta^2(s)ds + \int_0^t \vartheta(s)dW_s} = x_0 e^{rt - \frac{1}{2}\int_0^t \vartheta^2(s)ds + B_{\int_0^t \vartheta^2(s)ds}}$$

where $(B_u)_{u\geq 0}$ is a standard Brownian motion (the second equality follows form the Dambins-Dubins-Schwarz theorem).

Consequently $X_t = \varphi\left(t, B_{\int_0^t \vartheta^2(s)ds}\right)$ where the functional $\xi \mapsto \left(t \mapsto \varphi\left(t, \xi\left(\int_0^{\cdot} \vartheta^2(s)ds\right)\right)\right)$ defined on $\mathbb{D}([0, T_\vartheta], \mathbb{R})$, $T_\vartheta = \int_0^T \vartheta^2(t)dt$, is C-continuous on $\mathcal{C}([0, T_\vartheta], \mathbb{R})$. Hence for any C-continuous $\mathbb{R}$-functional on $\mathbb{D}([0, T], \mathbb{R})$, the $\mathbb{R}$-valued functional $\widetilde{F}$ defined by $\widetilde{F}(\xi) = F\left(\varphi\left(t, \xi\left(\int_0^{\cdot} \vartheta^2(s)ds\right)\right)\right)$ is C-continuous on $\mathbb{D}([0, T_\vartheta], \mathbb{R})$. Then, on can transfer the co-monotony property from B to X.

• *Local volatility model (elliptic case)*: More generally, it applies still with $I = (0, +\infty)$ to some usual extensions like the models with local volatility

$$dX_t = rX_tdt + \vartheta(X_t)X_tdW_t, \quad X_0 = x_0 > 0,$$

where $\vartheta : \mathbb{R} \rightarrow (\vartheta_0, +\infty)$, $\vartheta_0 > 0$, is a bounded, twice differentiable function satisfying $|\vartheta'(x)| \leq \frac{C}{1+|x|}$ and $|\vartheta''(x)| \leq \frac{C}{1+|x|^2}$, $x \in (0, +\infty)$.

In this case $I = (0, +\infty)$ and, $x_1 \in I$ being fixed, one has for every $x \in I$,

$$L(x) = \int_{x_1}^x \frac{d\xi}{\xi\vartheta(\xi)}$$

which clearly defines an increasing homeomorphism from I onto $\mathbb{R}$ since ϑ is bounded.

Furthermore, one easily derives from the explicit form (B.35) and the condition (B.36) that β is Lipschitz as soon as the function

$$x \mapsto rx\frac{\vartheta'}{\vartheta}(x) + \frac{x^2\vartheta\vartheta''(x)}{2} + x\vartheta\vartheta'(x) \text{ is bounded on } (0, \infty)$$

which easily follows from the assumptions made on ϑ.

Extension to other classes of diffusions and models. This general approach does not embody all situations: thus the true CEV model does not fulfill the above assumptions. The CEV model is a diffusion process X following the SDE

$$dX_t = rX_t dt + \vartheta X_t^\alpha dW_t, \quad X_0 = x_0,$$

where $\vartheta > 0$ and $0 < \alpha < 1$ are real numbers.

So this CEV model, for which $I = (0, +\infty)$, does not fulfill Definition $\mathcal{A}_{b,\sigma}$-(iii). As a consequence $L(t, I) \neq \mathbb{R}$ is an open interval (depending on the choice of x_1. To be precise, if $x_1 \in (0, +\infty)$ is fixed,

$$L(x) = \frac{1}{\vartheta(1-\alpha)}\big(x^{1-\alpha} - x_1^{1-\alpha}\big), \quad x \in (0, +\infty)$$

so that, if we set

$$J_{x_1} := L(I) = \Big(-\frac{x_1^{1-\alpha}}{\vartheta(1-\alpha)}, +\infty\Big),$$

L defines an homeomorphism from $I = (0, +\infty)$ onto J_{x_1}. Finally the function β defined by

$$\beta(y) = \frac{r}{\vartheta}\big(\vartheta(1-\alpha)y + x_1^{1-\alpha}\big) - \frac{\alpha\vartheta}{2}\frac{1}{(\vartheta(1-\alpha)y + x_1^{1-\alpha})}, \quad y \in J_{x_1}$$

is non-decreasing with linear growth at $+\infty$. Now, tracing the lines of the above proof, in particular establishing weak existence and uniqueness of the solution of the SDE (B.29) in that setting, leads to the same positive conclusion concerning the covariance inequalities for co-monotonic or anti-monotonic functionals.

References

[1] F. Abergel and A. Jedidi. A mathematical approach to order book modelling. In *Proceedings of the 5th Kolkata Econophysics conference, Springer (in press)*.

[2] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2):143–157, 2010.

[3] R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.

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

[5] E. Bayraktar and M. Ludkovski. Liquidation in limit order books with controlled intensity. *CoRR*, 2011.

[6] A. Beskos and G.O. Roberts. Exact simulation of diffusions. *Ann. appl. Prob.*, 15(4):2422–2444, 2005.

[7] B. Bouchard, N.-M. Dang, and C.-A. Lehalle. Optimal control of trading algorithms: a general impulse control approach. *SIAM J. Financial Math.*, 2:404–438, 2011.

[8] M. Duflo. *Algorithmes stochastiques*, volume 23 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996.

[9] T. Foucault, O. Kadan, and E. Kandel. Limit order book as a market for liquidity. Discussion Paper Series dp321, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem, January 2003.

[10] O. Guéant, J. Fernandez-Tapia, and C.-A. Lehalle. Dealing with the inventory risk. Technical report, 2011.

[11] O. Guéant, C.-A. Lehalle, and J. Razafinimanana. High Frequency Simulations of an Order Book: a Two-Scales Approach. In F. Abergel, B. K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-Driven Markets*, New Economic Windows. Springer, 2010.

[12] F. Guilbaud, M. Mnif, and H. Pham. Numerical methods for an optimal order execution problem. to appear in Journal of Computational Finance, 2010.

[13] F. Guilbaud and H. Pham. Optimal high-frequency trading with limit and market orders. 2012.

[14] T. Ho and H. R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9(1):47–73, 1981.

[15] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[16] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[17] S. Karlin and H. M. Taylor. *A second course in stochastic processes*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1981.

[18] H. J. Kushner and D. S. Clark. *Stochastic approximation methods for constrained and unconstrained systems*, volume 26 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1978.

[19] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[20] S. Laruelle, C.-A. Lehalle, and G. Pagès. Optimal Split of Orders Across Liquidity Pools: A Stochastic Algorithm Approach. *SIAM J. Financial Math.*, 2(1):1042–1076, 2011.

[21] S. Laruelle and G. Pagès. Stochastic approximation with averaging innovation applied to finance. *Monte Carlo Methods Appl.*, 18(1):1–51, 2012.

[22] J. McCulloch. A model of true spreads on limit order markets. 2011. Available at SSRN: http://ssrn.com/abstract=1815782.

[23] G. Pagès. A functional co-monotony principle with an application to peacoks. 2012. to appear in *Séminaire de probabilités*.

[24] S. Predoiu, G. Shaikhet, and S. Shreve. Optimal Execution of a General One-Sided Limit-Order Book. Technical report, Carnegie Mellon University, September 2010.

[25] C. Y. Robert and M. Rosenbaum. A new approach for the dynamics of ultra high frequency data: the model with uncertainty zones. *Journal of Financial Econometrics*, 9(2):344–366, 2011.