



**HAL**  
open science

## **Bid-ask spread modelling, a perturbation approach**

Thomas Lim, Vathana Ly Vath, Jean-Michel Sahut, Simone Scotti

► **To cite this version:**

Thomas Lim, Vathana Ly Vath, Jean-Michel Sahut, Simone Scotti. Bid-ask spread modelling, a perturbation approach. 2011. hal-00679042

**HAL Id: hal-00679042**

**<https://hal.science/hal-00679042>**

Preprint submitted on 26 Mar 2012

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

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

# Bid-ask spread modelling, a perturbation approach

Thomas Lim <sup>\*</sup>    Vathana Ly Vath<sup>†‡</sup>    Jean-Michel Sahut<sup>§</sup>    Simone Scotti<sup>¶</sup>

March 26, 2012

## Abstract

Our objective is to study liquidity risk, in particular the so-called “bid-ask spread”, as a by-product of market uncertainties. “Bid-ask spread”, and more generally “limit order books” describe the existence of different sell and buy prices, which we explain by using different risk aversions of market participants. The risky asset follows a diffusion process governed by a Brownian motion which is uncertain. We use the error theory with Dirichlet forms to formalize the notion of uncertainty on the Brownian motion. This uncertainty generates noises on the trajectories of the underlying asset and we use these noises to expound the presence of bid-ask spreads. In addition, we prove that these noises also have direct impacts on the mid-price of the risky asset. We further enrich our studies with the resolution of an optimal liquidation problem under these liquidity uncertainties and market impacts. To complete our analysis, some numerical results will be provided.

**Keywords :** Error theory using Dirichlet Forms, order books, bid-ask spread, optimal control problem, liquidity.

**JEL Classification :** C61.

**MSC Classification (2000) :** 60G15, 60H10, 60G45, 91G80.

---

<sup>\*</sup>ENSIIE and Université d’Evry, Laboratoire d’Analyse et Probabilités, France, [lim@ensiie.fr](mailto:lim@ensiie.fr)

<sup>†</sup>ENSIIE and Université d’Evry, Laboratoire d’Analyse et Probabilités, France, [lyvath@ensiie.fr](mailto:lyvath@ensiie.fr)

<sup>‡</sup>This research benefitted from the support of the “Chaire Risque de Crédit”, Fédération Bancaire Française

<sup>§</sup>Geneva School of Business Administration, [jean-michel.sahut@hesge.ch](mailto:jean-michel.sahut@hesge.ch)

<sup>¶</sup>Université Paris Diderot, Laboratoire Probabilités et Modèles Aléatoires, France, [simone.scotti@paris7.jussieu.fr](mailto:simone.scotti@paris7.jussieu.fr)

## Introduction

A stock index is defined as an aggregate value produced by combining several stocks or other investment vehicles together and expressing their total values against a base value from a specific date. Stock Market indexes, for instance, are intended to represent an entire stock market and thus track the market's changes over time. They are generally used as benchmark to measure the relative performance of any given investment portfolio. But they are not directly tradable assets. An increasing role of the asset management industry is to provide investors with investment tools capable of replicating a wide range of indices. These investment tools are often known as trackers or Exchange Traded Funds (ETF). To put it into perspective, investments in ETF have significantly increased over the last decade.

Let us now consider a process  $X$  representing a stock market index, i.e. a non-tradable asset, which is governed by a given SDE (stochastic differential equation). We also consider the process  $S$  representing an ETF on the non-tradable index  $X$ .

Our main objective in this paper is to address two important questions arising from the trading of the ETF. The first question concerns the price value and its trajectory behavior compared to that of the index. Beyond this first question, we will, in particular, address the microstructure problem of the transaction market and attempt to model the financial and economic rationale behind the existence of the bid-ask spread. The second question is linked to the practical study of optimal liquidation strategy in the context of non-liquidity.

Managing a tracker fund is not as simplistic as it first looks. Indeed, the impossibility to trade continuously, the lack of liquidity in the transaction of the index's components, liquidity and transaction costs, etc, imply that the values  $S$  of the index tracker fund and the index  $X$  may be highly correlated but not identical, i.e. this phenomenon is called tracking errors.

By value of the tracker fund, we mean the "fundamental value" obtained by summing parts of the tracker fund. Due to tracking errors, which are non-observable, the "theoretical" value of the tracker fund is itself a non-observable process, i.e. at each time, it is a random variable. Furthermore, should we take into account the liquidity problem, the "value" of the tracker fund should be the liquidation value of the assets owned in the fund. The liquidation value is very different from the last-transaction based price or the mid-price. However, from the investors' point of view, the price value of the tracker fund is also and especially its quoted prices. Indeed, the ETF is itself a trading asset. As such, to invest in (disinvest from) the tracker fund, one has to go through the market and make appropriate buy (sell) order transaction. Its price therefore depends on the dynamics of the overall buy and sell orders.

To sum up, the dynamics of the price value  $S$  of the ETF has two driving factors: a fundamental factor and a trading factor. The fundamental factor considers the tracker fund as the aggregate value produced by combining its components while the trading factor deals with the microstructure of market transactions, i.e. when the liquidity risk and cost are factored in.

In addressing our first question on the dynamics of the price value  $S$  and its trajectory, we will in particular study how the non-observability of the fundamental values of the tracker fund explains the existence of the limit order book and more specifically the bid-ask spread. Indeed, at any fixed time, the price value  $S$  is not observable but it is a random variable where its law or at least its mean and variance may be characterized. In the context of this randomness and non-observability, the existence of different risk aversions of market participants directly explains the presence of different types of orders, i.e. limit and market

orders and in particular the presence of many buy and sell limit order prices forming the limit order book. In order to focus on the modelling of the bid-ask spread, we consider a “representative” price-setter market participant who places all limit orders, in particular the best buy and best sell limit orders, i.e. the prices and the number of shares he is willing to buy and sell. Prior to setting the buy and sell orders, the representative agent obtains the distribution of possible asset values from the market information but has no possibility to observe the assets’ realized values. A rational decision is to send a buy (sell) limit order with a lower (higher) price with respect to the assets mean value such that their difference justifies the risk taken.

In the mathematical finance literature, there are several approaches modelling liquidity risk. One approach is to explain liquidity risk by the presence of an insider [2], [17]. Another approach is the market manipulation literature where prices are assumed to depend directly on the trading strategies, see [3] and [10]. A third approach is to consider liquidity risk in terms of the difference between the bid and ask prices, i.e. the existence of a bid-ask spread. Transaction costs [16] are one way to model the bid-ask spread. In [18], liquidity risk is expressed by the presence of transaction costs and market manipulation. Another way is to directly model the bid-ask spread or the order books and explain its presence as being intrinsic to the financial markets, driven by trades between different market participants.

Bid-ask spread and order books play a crucial role in many financial problems such as unwinding large block of shares for large investors and hedging strategy of options for traders. These problems were investigated by the likes of [1] and [20]. Due to the complexity of the study, state process representing the underlying price is assumed to have simple dynamics such as Bachelier’s dynamics [20], or is assumed to be a martingale process [1].

However, in the above models, liquidity risk is considered a posteriori. In other words, assuming the existence of liquidity risk, different approaches are used to replicate its effects. To our knowledge, few studies in the fields of mathematical finance have attempted to model the financial and economic rationales behind the existence of the bid-ask spread. This is precisely the objective in this paper: study and explain, in the context of the ETF trading, the existence of the bid-ask spread, and more generally limit order book, as a by-product of market uncertainties.

The mathematical formulation of such problems relies on the specification of a coherent framework to describe the remaining randomness on prices. In our study, the asset value depends on two random sources: the first one describes the evolution of the asset mean value while the second delineates the shape of asset (sell-buy) prices at a given fixed time. The coupling of the two probability spaces, with its respective filtration, requires complex tools and represents the principal drawback of this kind of approach. Therefore, we choose a different strategy based on error theory using Dirichlet forms formalism developed in [8] and [9]. The advantages of this approach are inherent to its elasticity and powerful tools.

Such an approach provides us with a perfect knowledge of the bid-ask spread component of the order book, i.e. the best/highest bid price and the best/lowest ask price of the order book placed by the “representative” price-setter market participant. Once our bid-ask spread model obtained, as in [20] or [1], we investigate an optimal liquidation problem for a large portfolio.

The article is organized as follows. In Section 1, we introduce the economic model for bid-ask spread and we present the analysis of prices variance and bias. In Section 2, we study an optimal liquidation problem associated with the bid-ask spread model developed in the previous section. And finally, in Section 3, we provide some numerical results.

# 1 The model

In this section, we aim at modelling the dynamics of the ask and bid prices of an ETF fund. Our objective is to define an asset price model that considers the bid-ask spread as an inherent part of asset price evolution.

We consider a process  $X$  representing an observable but non-tradable benchmark index, e.g. an industrial sector index, which is governed by a given stochastic differential equation (SDE). We also consider a process  $S$  representing an ETF on this index. The first question arising from the trading of the asset  $S$  concerns its price value. In particular, we address the way the “representative agent” sets his best bid and ask price and the resulting dynamics of those two prices. It is clear that the price of the asset  $S$  is generally very closed to the tracked index  $X$ .

## 1.1 Theoretical analysis of path sensitivity and approximation

We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and the process  $X$  is governing by the following SDE.

$$(1.1) \quad dX_t = r X_t dt + X_t \sigma(t, X_t) dW_t ,$$

where  $r$  is the drift and  $\sigma$  a function on  $\mathbb{R}^+ \times \mathbb{R}$  that verifies the following assumptions.

### Assumption H-1 (Underlying diffusion)

1. SDE (1.1) admits a unique strong solution, denoted  $X$ , such that  $X$  is square-integrable and does not explode in finite time with probability 1.
2. The solution  $X$  of SDE (1.1) is always positive.
3.  $f(x) = x \sigma(t, x)$  is twice derivable function in  $x$  and the derivatives are Lipschitz and bounded.

These assumptions cover a large class of stochastic models in finance. In particular, Assumption H-1 is satisfied by log-normal diffusion and a large part of local volatility models. For Constant Elasticity of Variance model, we may refer to [11].

To simplify our notations, we denote the first and second derivatives of  $x \sigma(t, x)$ :

$$\begin{aligned} \zeta(t, x) &= \sigma(t, x) + x \frac{\partial \sigma}{\partial x}(t, x) , \\ \eta(t, x) &= 2 \frac{\partial \sigma}{\partial x}(t, x) + x \frac{\partial^2 \sigma}{\partial x^2}(t, x) . \end{aligned}$$

We assume that the price process  $S$  of the ETF follows the same SDE as the one of the index  $X$ , but its Brownian motion is perturbed by the problem of replication. The process  $S$  is therefore assumed to follow the following SDE

$$(1.2) \quad dS_t = r S_t dt + \sigma(t, S_t) S_t dB_t ,$$

where  $B$  is a Brownian motion, which is almost explained by  $W$  but characterized by a small uncertainty. We assume that

$$(1.3) \quad B_t = \sqrt{e^{-\epsilon}} W_t + \sqrt{1 - e^{-\epsilon}} \widehat{W}_t ,$$

where  $\epsilon$  is a small parameter and  $\widehat{W}$  is a Brownian motion, independent w.r.t. the filtration  $(\mathcal{F}_t)_{t \geq 0}$  where  $\mathcal{F}_t = \sigma\{W_s, 0 \leq s \leq t\}$ , that resume all hedging errors.

The two Brownian motions,  $W$  and  $\widehat{W}$ , play different roles.  $W$  describes the market information, that is progressively known through the index value. Therefore, at time  $t$  the information  $\mathcal{F}_t$  is known, whereas the information  $\mathcal{G}_t = \sigma\{\widehat{W}_s, 0 \leq s \leq t\}$  is unknown or unobservable. To deal with this problem, we choose to follow error theory approach using Dirichlet forms developed on [8] and [9]. We fix an error structure  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{D}, \Gamma)$ , where  $(\Omega, \mathcal{F}, \mathbb{P})$  is the Wiener space in which the Brownian motion  $B$  lives, while  $\Gamma$  is an Ornstein-Uhlenbeck carré du champ operator with constant weight  $\theta$  (see Section 3 in [9]). Using this theory, formula (1.3), known as Mehler formula, is automatically justified (see Section VI.2. in [8]).

Error theory enables us to find a limited expansion of the law of the price of illiquid asset due to the noise on Brownian motion. In particular, we have the following results.

**Theorem 1.1 (Law of illiquid asset price)**

*Under Assumption H-1 and a Gaussian approximation, the uncertainty on Brownian motion is transmitted to the process  $S$ , which represents the illiquid asset price. Then, any realization  $\bar{\omega}$  of process  $X$ , at time  $t$ , fixes a  $(\mathcal{F}_t - \text{conditional})$  random variable  $S_t(\bar{\omega})$  described by*

$$(1.4) \quad S_t(\bar{\omega}, \widehat{\omega}) = X_t(\bar{\omega}) + \epsilon \mathcal{A}[S]_t(\bar{\omega}) + \sqrt{\epsilon \Gamma[S]_t(\bar{\omega})} \widetilde{\mathcal{N}}(\widehat{\omega}),$$

where  $\widetilde{\mathcal{N}}$  is a centered reduced Gaussian random variable independent w.r.t.  $\mathcal{F}_t$ , while  $\Gamma[S]_t(\bar{\omega})$  and  $\mathcal{A}[S]_t(\bar{\omega})$  are given by

$$(1.5) \quad \begin{cases} \Gamma[S]_t = \theta M_t^2 \int_0^t \frac{X_s^2 \sigma^2(s, X_s)}{M_s^2} ds + \Gamma[S]_0 M_t^2, \\ \mathcal{A}[S]_t = M_t \int_0^t \frac{\eta(s, X_s) \Gamma[S]_s - \theta X_s \sigma(s, X_s)}{2 M_s} [dW_s - \zeta(s, X_s) ds], \\ M_t = \mathcal{E} \left\{ \int_0^t \zeta(s, X_s) dW_s + \int_0^t r ds \right\}_t, \end{cases}$$

where  $\mathcal{E}$  denotes the Doleans-Dade exponential.

The proof of this theorem is mainly based on the truncated expansion in error theory using Dirichlet forms, see [8] and [9]. The two following lemmas form the main backbone of the proof. Indeed, they give the expression of the variance  $\Gamma[S]$  and the bias  $\mathcal{A}[S]$ .

**Lemma 1.1 (Variance due to Brownian motion)**

*Let  $X$  be the solution of SDE (1.1) and assume that Assumption H-1 holds. Then, the uncertainty effect on process  $S$  satisfies the following SDE*

$$(1.6) \quad d\Gamma[S]_t = 2\zeta(t, X_t) \Gamma[S]_t dW_t + [2r + \zeta^2(t, X_t)] \Gamma[S]_t dt + \theta \sigma^2(t, X_t) X_t^2 dt.$$

Moreover,  $\Gamma[S]$  has the following closed form solution

$$\Gamma[S]_t = \theta M_t^2 \int_0^t \frac{X_s^2 \sigma^2(s, X_s)}{M_s^2} ds + \Gamma[S]_0 M_t^2.$$

**Lemma 1.2 (Bias due to Brownian motion)**

Let  $X$  be the solution of SDE (1.1) and assume that Assumption H-1 holds. Then, the bias effect on process  $S$  satisfies the following SDE

$$(1.7) \quad d\mathcal{A}[S]_t = r\mathcal{A}[S]_t dt + \left[ \zeta(t, X_t)\mathcal{A}[S]_t + \frac{1}{2}\eta(t, X_t)\Gamma[S]_t - \frac{\theta}{2}\sigma(t, X_t)X_t \right] dW_t.$$

Moreover,  $\mathcal{A}[S]$  has the following closed form solution

$$\mathcal{A}[S]_t = M_t \int_0^t \frac{\eta(s, X_s)\Gamma[S]_s - \theta X_s \sigma(s, X_s)}{2M_s} [dW_s - \zeta(s, X_s) ds].$$

The proof of these lemmas are postponed in Appendix A.

**Remark 1.1 (Closed forms)**

Equations (1.5) show an interesting property of processes  $\Gamma[S]$  and  $\mathcal{A}[S]$ , it is easy to check that the law of  $(\Gamma[S]_t, \mathcal{A}[S]_t)_{t \geq 0}$  is completely explicit given the law of  $(X_t, W_t)_{t \geq 0}$ . Therefore, Equations (1.5) are closed forms in the sense of involving only algebraic operations and stochastic integrals.

**Remark 1.2 (Black Scholes case)**

In the particular case of  $\sigma$  constant, i.e. in the Black Scholes model, Equations (1.5) are simplified with  $\Gamma[S]_t$  proportional to the square of  $X_t$  and  $\mathcal{A}[S]_t$  proportional to  $X_t$ .

Moreover, we have the following corollary:

**Corollary 1.1 (Equilibrium price)**

The equilibrium price, i.e. the mean of the price distribution, is given by

$$(1.8) \quad S_t^M = \mathbb{E}[S_t | \mathcal{F}_t] = X_t + \epsilon \mathcal{A}[S]_t.$$

The equilibrium price is therefore different from  $X_t$ . In particular, this shift exists in Black Scholes framework. However, in this case, this shift is proportional to  $X_t$ , so it is possible to include it into the starting point  $S_0^M$ . This shift can explain tracking errors usually observed on ETF-markets, for instance see [14].

Finally as a corollary of the two previous lemmas, we have the following Markov property.

**Corollary 1.2 (Markov property)**

The triplet  $\tilde{X} = (X, \Gamma[S], \mathcal{A}[S])$  is a Markovian process if and only if  $X$  is Markovian.

This assertion is a direct consequence of the fact that  $\Gamma[S]$  verifies SDE (1.6) which only depends on the process  $X$ , and finally  $\mathcal{A}[S]$  follows SDE (1.7) which depends on both  $X$  and  $\Gamma[S]$ .

## 1.2 Bid-ask model

Theorem 1.1 gives us the law of the illiquid asset price given the value of the benchmark/index. In this subsection, we explain how this approach can be used to define bid and ask prices and suggest a model that reproduces it.

We consider the presence of many agents on the market, all informed about the economic evolution of the benchmark price  $X$ , but without money-market intelligence about the residual information drawn by the perturbation, i.e. the independent Brownian motion  $\widehat{W}$ . We assume that all agents are risk adverse and can estimate the distribution of the

uncertainty of the illiquid asset price, at any fixed time  $t$ , given by Theorem 1.1. We now consider uniquely price-setter agents or liquidity providers who place limit orders as opposed to market orders placed by price-taker agents, liquidity takers. Indeed, their aggregated limit orders constitute an order book and therefore the bid-ask spread. It stands to reason that, at any given time  $t$ , there exists a price-setter agent with minimal risk aversion with respect to other agents. This agent accepts to buy the asset at a price  $S_t^B$  higher than the prices proposed by the other agents. Thus, the price proposed by this agent is the bid price and it is denoted by  $S_t^B$ . This price is completely defined by the law of the illiquid asset and the risk aversion of this agent. A symmetric analysis generates the ask price  $S_t^A$ .

Let us assume, for the sake of simplicity, that there exists a representative price-setter agent who always submits the best buy and sell prices, which we respectively define as best bid price  $S_t^B$  and best ask price  $S_t^A$ . Indeed, we assume that he accepts to buy the illiquid asset at a price  $S_t^B$  such that the risk of overvaluing of this asset is equal to a supportable risk probability  $\chi_B$ . Therefore he takes the risk against the expected earnings, see Figure 1. In conclusion,  $S_t^B$  is the  $\chi_B$ -quantile of the illiquid asset price distribution given by the uncertainty on the Brownian motion, see Theorem 1.1.

The definition of ask price is symmetric, i.e.  $S_t^A$  is the  $(1 - \chi_A)$ -quantile of the illiquid asset price distribution given by the uncertainty on the Brownian motion. It is clear that  $\chi_A + \chi_B < 1$ , since the representative agent is risk-averse.

**Definition 1.1 (Static bid and ask prices)**

Let  $\chi_B$  and  $\chi_A$  with  $\chi_A + \chi_B < 1$  be risks taken by the “representative price setter” in respectively overvaluing and undervaluing the illiquid asset at a given time  $t$ . The corresponding bid  $S_t^B$  and ask  $S_t^A$  prices are defined as follow

$$\begin{aligned} S_t^B &= X_t + \epsilon \mathcal{A}[S]_t + \sqrt{\epsilon \Gamma[S]_t} \tilde{\mathcal{N}}^{-1}(\chi_B) , \\ S_t^A &= X_t + \epsilon \mathcal{A}[S]_t + \sqrt{\epsilon \Gamma[S]_t} \tilde{\mathcal{N}}^{-1}(1 - \chi_A) . \end{aligned}$$

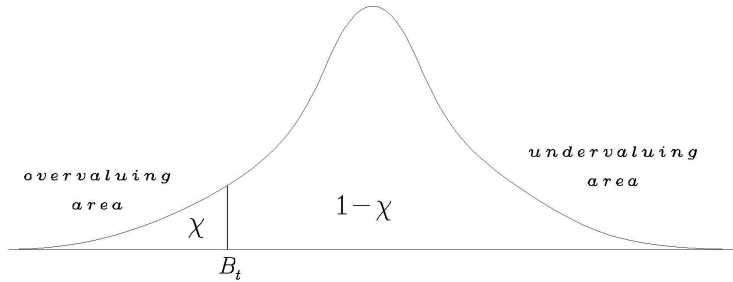


Figure 1: Bid price definition, defined by a risk probability  $\chi$ .

Since the law of residual uncertainty is always Gaussian, the definition of the bearable risk is equivalent to the definition of the trader utility function. For sake of simplicity, we fix the same supportable risk for sell  $S_t^B$  and buy  $S_t^A$  prices, i.e.  $\chi_B = \chi_A = \chi$ . In Figure 2, we study an example and consider the CEV model. Given a trajectory of  $X$ , we compute the evolution of the bid price  $S^B$ , the mid-price  $S^M$  and the ask price  $S^A$ .



Evolution of prices

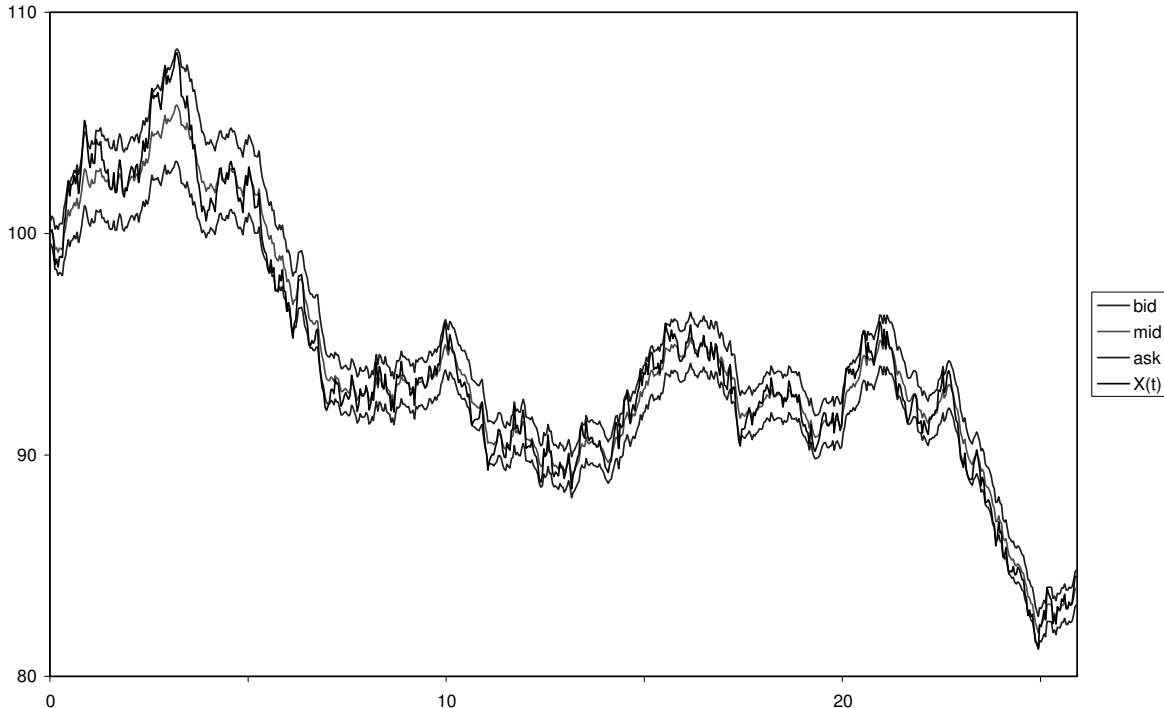


Figure 2: An example with CEV model: given a single path of the process  $X$ , in black, we can compute explicitly the mid-price, in light grey, and the bid and the ask prices using a standard deviation, in dark grey.

**Remark 1.3**

The trajectories of  $X$  and  $S^M$  are different. This is due to the fact that  $X$  is not linear w.r.t.  $W$  in CEV diffusion, see Section 2.3, hence the error introduces a bias, see Corollary 1.1.

In order to define a bid-ask model, we have to choose a dynamic for this risk aversion, since a static risk aversion is very restrictive. The dynamic risk aversion is not only justified by the very nature of the “representative price-setter agent” but also by market order flow from price-taker agents.

We now turn to the choice of a dynamics of the bid-ask spread. In the economic literature, bid-ask spread depends mainly on two factors: the value of the stock and the trading volume, see [22]. In particular, the bid-ask spread converges to zero (resp. infinity) when the asset price goes to zero (resp. infinity). However the relative spread<sup>1</sup> converges to a strictly positive constant (resp. zero) when the asset price goes to zero (resp. infinity). This effect can be explained endogenously with the evolution of the variance  $\Gamma[S]$ , see Section 2.3 for an analysis in CEV case.

The trading volume equally plays a leading role. If we analyze two assets with almost the same price but with different trading volumes, we notice that the lower the trading

<sup>1</sup>The relative spread is defined as the ratio between bid-ask spread and the asset.

volume, i.e the more illiquid the asset is, the larger bid-ask spread, see for instance [24]. An economic explanation is that the traders accept higher risks if they can easily close their positions, which is possible with the presence of many counterparts. Historical data shows that average trading volumes are mean-reverting on medium term, the bid-ask spread shows the same behavior.

In order to satisfy the above behavior, we use a continuous time finite state markov chain process  $L$  representing market liquidity, all state values of  $L$  are positive. We assume that the markov chain is irreducible, so that we can say that  $L$  is a mean reverting process. Finally, we suppose that  $(L_t)_{t \geq 0}$  is independent with respect to  $(\mathcal{F}_t)_{t \geq 0}$ . We consider the following model of the bid-ask spread.

**Definition 1.2 (Bid and ask model)**

At any time  $t$ , given the value of the benchmark  $X_t(\bar{w})$ , the bid and ask prices are given by

$$(1.9) \quad \begin{cases} S_t^A = X_t + \epsilon \mathcal{A}[S]_t + \sqrt{\epsilon \Gamma[S]_t} L_t , \\ S_t^B = X_t + \epsilon \mathcal{A}[S]_t - \sqrt{\epsilon \Gamma[S]_t} L_t . \end{cases}$$

**Remark 1.4**

The choice of this model is justified by the following properties.

- *Positivity*: The ask price is always higher than the bid price.
- *Closed forms*: In our model, all terms, excepted the underlying  $X$ , have an explicit form. The law of  $X$  is the unique law that we have to estimate numerically. This computation can be easily performed using a Monte-Carlo method.
- *Error tracking*: The mid-price  $S^M$  is different to the benchmark one, since a systematic bias exists. The two prices are relatively closed given a small parameter  $\epsilon$ .
- *Separation*: In our model, the bid-ask spread is explained by two independent factors. The first factor concerns the sensitivity of the benchmark/index level path w.r.t. the Brownian motion  $W$ . In an economic point of view, it corresponds to the sensitivity w.r.t. “market” information. The second one is risk aversion of market participants which mainly depends on trading volumes.
- *Mean reverting*: If the underlying value is relatively stable, the bid-ask spread shows a mean reverting behavior.
- *Bid-ask spread tails*: Given the evolution of the benchmark, the law of the bid-ask spread is lognormal, so extremely wide or small spreads are possible but with a very low probability.

## 2 Optimal liquidation portfolio problem

### 2.1 The economic motivations and the objective functions

The above bid-ask spread model, as defined in Definition 1.2, highlights the market imperfections due to liquidity risk. Thus, a natural but challenging problem for both professionals and academics in finance to solve is the optimal portfolio liquidation problem. Let us consider a price-taker investor who decides to close his position over a finite horizon, he has to define a trading strategy which maximizes his terminal portfolio value. Since the attempt to sell the whole block of shares causes, generally, a lack of balance between supply and demand, it results in an average selling price well below the best pre-order bid price. In

practice, large orders are generally split into a number of consecutive small orders to reduce the overall price impact.

Let us therefore investigate a problem of an investor seeking to liquidate  $N$  shares of stock over a finite time horizon  $T$ . To solve this problem, we consider a discrete framework by assuming that trading occurs only at discrete times  $t_0 < t_1 < \dots < t_n = T$ . A strategy decision  $\pi$  for the investor is a sequence  $(\pi_i)_{0 \leq i \leq n}$  valued in  $[0, N]$  where  $\pi_i$  is  $\mathcal{F}_{t_i}$ -measurable and represents the number of shares to be sold at time  $t_i$ . We define an admissible strategy as being a strategy  $\pi$  such that  $\sum_{i=0}^n \pi_i = N$ . As such we define the set of admissible strategies  $\mathcal{A}(t_i, p)$  as

$$(2.1) \quad \mathcal{A}(t_i, p) = \left\{ \pi = \{\pi_i, \dots, \pi_n\}, \pi_j \geq 0 \quad \forall j \in \{i, \dots, n\} \text{ and } \sum_{j=i}^n \pi_j = p \right\}.$$

**Price impact.** In addition to the existence of the bid-ask spread as evidenced in the previous section, we equally take into account a lack of market depth by assuming that marginal selling prices are non-increasing. Indeed, there is no infinite liquidity at either the best bid price nor at the best ask price. For that purpose, we introduce an impact function  $g$  which indicates the average price obtained at a sell market order. More precisely, when an investor submits a sell order of  $x$  number of shares through a single sell order at time  $t$ , the obtained average price  $\bar{S}_t^B(x)$  is assumed

$$(2.2) \quad \bar{S}_t^B(x) = S_t^B g(x),$$

where the function  $g$  verifies the following assumption.

**Assumption H-2 (Trading impact function)**

1.  $g$  is a continuous positive deterministic function independent to  $S$ .
2.  $g$  is non-increasing.
3.  $h(x) = x g(x)$  is strictly non-decreasing and concave.

**Remark 2.5**

1. We assume that trading impact is temporary when trading occurs. Only the price takers who place market price orders (at best selling prices) pay the liquidity costs. After the trades, the order book is refilled with limit orders from other market participants [1].
2.  $g(x)$  corresponds to the ratio between the average stock price received following the sale of  $x$  shares at market price order and the best bid price. This average price obviously decreases with the number of traded shares.
3. The marginal price  $[(x + \delta x)g(x + \delta x) - xg(x)]S^B$  should be non-negative and non-increasing. Therefore, the function  $h(x) = x g(x)$  must be non-decreasing and concave. The concavity comes from the shape of the order book, which displays a maximum around the best bid price, see [22].

**Objective function.** The objective of the investors is to maximize their net present value (wealth) from the sales of the stock shares in holding. To fully describe our state process, we should take into account not only the processes  $X$  and  $L$  but also  $\Gamma[S]$  and  $\mathcal{A}[S]$ . As such, the state process to consider is  $Z = (\tilde{X}, L)$ , where  $\tilde{X}$  is defined as in Corollary 1.2. At any initial time  $t_i$  and any state value  $(z, p)$  of the variables  $(Z_{t_i}, P_{t_i})$ , with  $P_{t_i}$  the number of

stock shares that we initially have at time  $t_i$ , we define our reward function for any strategy  $\pi \in \mathcal{A}(t_i, p)$  by

$$J(i, z, p, \pi) = \mathbb{E} \left[ \sum_{j=i}^n e^{-\rho(t_j - t_i)} \pi_j S_{t_j}^B g(\pi_j) \middle| \mathcal{F}_{t_i} \right],$$

where  $\rho$  represents the interest rate.

The objective of the investor is to maximize this reward function over all admissible strategies. We therefore introduce the following value function

$$(2.3) \quad v(i, z, p) = \sup_{\pi \in \mathcal{A}(t_i, p)} (J(i, z, p, \pi)).$$

For an initial state  $(i, z, p)$ ,  $\hat{\pi} \in \mathcal{A}(t_i, p)$  is called an optimal strategy if

$$v(i, z, p) = J(i, z, p, \hat{\pi}).$$

In the sequel, we restrict the set of admissible strategies  $\mathcal{A}(t_i, p)$  to Markov strategies subset of  $\mathcal{A}(t_i, p)$ , which is denoted  $\bar{\mathcal{A}}(t_i, p)$  (that is possible from Proposition 8.1 of [5]).

## 2.2 Theoretical solution of the optimization problem

We now prove the existence of a solution to our optimization problem (2.3) and its uniqueness. Given an initial  $N$  stock shares of the risky assets, our objective is to prove that an optimal strategy in liquidating our portfolio exists in  $\mathcal{A}(t_0, N)$  and this one is unique. For notation convenience, we shall denote  $Z_i$  (resp.  $S_i, P_i$ ) for  $Z_{t_i}$  (resp.  $S_{t_i}, P_{t_i}$ ). Using the dynamic programming principle, we have:

### Theorem 2.1 (Existence)

*Under Assumptions H-1 and H-2, there exists an optimal policy  $\hat{\pi} = (\hat{\pi}_0, \dots, \hat{\pi}_n)$  to the optimization problem, such that  $\hat{\pi} \in \bar{\mathcal{A}}(t_0, N)$ . This optimal strategy is given by the argmax in the following programming equation*

$$(2.4) \quad \begin{cases} v(i, z, p) = \operatorname{ess\,sup}_{0 \leq \pi_i \leq p} \left\{ \pi_i s_i^B g(\pi_i) + \mathbb{E} \left[ e^{-\rho(t_{i+1} - t_i)} v(i+1, Z_{i+1}^{i,z}, p - \pi_i) \middle| \mathcal{F}_{t_i} \right] \right\}, \\ v(n, z, p) = p s_n^B g(p), \end{cases}$$

where  $s_i^B$  is defined by the components of the variable  $z_i$  (see the definition of  $S^B$  in (1.9)).

**Proof.** This is an immediate application of Proposition 8.5 of [5]. From (1.9), we have

$$\mathbb{E} [S_t^B] = \mathbb{E} [X_t] + \epsilon \mathbb{E} [\mathcal{A}[S]_t] - \mathbb{E} \left[ \sqrt{\epsilon \Gamma[S]_t} \right] \mathbb{E} [L_t].$$

From Assumption H-1,  $X$  does not explode in finite time and  $\sigma(t, x)$ ,  $\zeta(t, x)$  and  $\eta(t, x)$  are Lipschitz and bounded, thus  $\mathbb{E} [\mathcal{A}[S]_t] < \infty$  and  $\mathbb{E} [\sqrt{\epsilon \Gamma[S]_t}] < \infty$ . Since the process  $L$  is a finite state markov chain, it is clear that  $\mathbb{E} [L_t] < \infty$  and as the process  $X$  is square integrable, we also have that  $\mathbb{E} [X_t] < \infty$ . Therefore  $\mathbb{E} [S_t^B] < \infty$  and it enables us to check assumptions  $(F^+)$  and  $(F^-)$  in Proposition 8.5 of [5]. Then, it remains to prove that the supremum in relation (2.4) is attained. This immediately follows from the continuity of  $v(i+1, z, p)$  with respect to  $p$ , which is the case thanks to Assumption H-2.  $\square$

We now turn to the uniqueness property of the optimal strategy.

**Theorem 2.2 (Uniqueness)**

Under Assumptions H-1 and H-2, there is at most one solution to optimization problem (2.3).

**Proof.** We first introduce the following function  $\vartheta$  defined for any  $x \leq y$  as

$$\vartheta(i, z, x, y) = x s_i^B g(x) + \mathbb{E} \left[ e^{-\rho \Delta_i} v(i+1, Z_{i+1}^{i,z}, y-x) \middle| \mathcal{F}_{t_i} \right], \quad i \in \{1, \dots, n-1\},$$

with  $\Delta_i := t_{i+1} - t_i$ .

We now prove by recurrence that  $\vartheta$  is concave w.r.t.  $(x, y)$  and the value function  $v$  is strictly concave w.r.t.  $p$  for all  $i \in \{0, \dots, n\}$ .

We first note that for  $i = n$ ,  $v(n, z, p)$  is strictly concave w.r.t.  $p$ , according to Assumption H-2. We can easily verify that  $\vartheta(n, z, x, y)$  is concave w.r.t.  $(x, y)$ . Assuming that for  $i+1$ ,  $v(i+1, z, p)$  and  $\vartheta(i+1, z, x, y)$  are respectively strictly concave w.r.t.  $p$  and to  $(x, y)$ , let prove that it is equally the case for  $i$ .

Let  $0 \leq \lambda \leq 1$ ,  $(x_1, y_1)$  and  $(x_2, y_2)$ , with  $0 \leq x_i \leq y_i \leq N$ , we have

$$\begin{aligned} \vartheta(i, z, \lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) &= (\lambda x_1 + (1-\lambda)x_2) s_i^B g(\lambda x_1 + (1-\lambda)x_2) \\ &\quad + \mathbb{E} \left[ e^{-\rho \Delta_i} v(i+1, Z_{i+1}^{i,z}, \lambda(y_1 - x_1) + (1-\lambda)(y_2 - x_2)) \middle| \mathcal{F}_{t_i} \right], \end{aligned}$$

since the first term is strictly concave, we obtain

$$(\lambda x_1 + (1-\lambda)x_2) s_i^B g(\lambda x_1 + (1-\lambda)x_2) > \lambda x_1 s_i^B g(x_1) + (1-\lambda)x_2 s_i^B g(x_2),$$

and by recurrence the second term is strictly concave, we obtain

$$\begin{aligned} v(i+1, Z_{i+1}^{i,z}, \lambda(y_1 - x_1) + (1-\lambda)(y_2 - x_2)) &> \lambda v(i+1, Z_{i+1}^{i,z}, y_1 - x_1) \\ &\quad + (1-\lambda) v(i+1, Z_{i+1}^{i,z}, y_2 - x_2). \end{aligned}$$

Taking the expectation, we get

$$\begin{aligned} \mathbb{E} \left[ e^{-\rho \Delta_i} v(i+1, Z_{i+1}^{i,z}, \lambda(y_1 - x_1) + (1-\lambda)(y_2 - x_2)) \middle| \mathcal{F}_{t_i} \right] &\geq \lambda \mathbb{E} \left[ e^{-\rho \Delta_i} v(i+1, Z_{i+1}^{i,z}, y_1 - x_1) \middle| \mathcal{F}_{t_i} \right] \\ &\quad + (1-\lambda) \mathbb{E} \left[ e^{-\rho \Delta_i} v(i+1, Z_{i+1}^{i,z}, y_2 - x_2) \middle| \mathcal{F}_{t_i} \right]. \end{aligned}$$

Thus  $\vartheta(i, z, x, y)$  is strictly concave w.r.t.  $(x, y)$  as the sum of two strictly concave functions.

We now prove that  $v(i, z, y)$  is strictly concave w.r.t.  $y$ . For any  $0 \leq x_1 \leq y_1$  and  $0 \leq x_2 \leq y_2$ , from the expression of  $v$  in (2.4), we have for all  $0 \leq \lambda \leq 1$

$$v(i, z, \lambda y_1 + (1-\lambda)y_2) \geq \vartheta(i, z, \lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2).$$

Since  $\vartheta(i, z, x, y)$  is strictly concave with respect to  $(x, y)$ , we get

$$v(i, z, \lambda y_1 + (1-\lambda)y_2) > \lambda \vartheta(i, z, x_1, y_1) + (1-\lambda) \vartheta(i, z, x_2, y_2).$$

The latter equality holds for any positive  $x_1 \leq y_1$  and  $x_2 \leq y_2$ . In particular since the supremum is attained (from Theorem 2.1), we can take  $x_1^*$  and  $x_2^*$  such that

$$\begin{aligned} v(i, z, y_1) &= \vartheta(i, z, x_1^*, y_1) = \sup_{0 \leq x \leq y_1} \vartheta(i, z, x, y_1) \\ v(i, z, y_2) &= \vartheta(i, z, x_2^*, y_2) = \sup_{0 \leq x \leq y_2} \vartheta(i, z, x, y_2) \end{aligned}$$

thus

$$v(i, z, \lambda y_1 + (1 - \lambda) y_2) > \lambda v(i, z, y_1) + (1 - \lambda) v(i, z, y_2) .$$

Hence,  $v(i, z, p)$  is strictly concave w.r.t.  $p$ . We have therefore proved the strict concavity of both functions. Using relation (2.4) and the above concavity property, we may obtain by iteration at most one solution to the optimization problem.  $\square$

### 2.3 Log-Normal and Constant Elasticity of Variance Diffusions

For numerical and implementation purpose, we now consider our two particular diffusion models, with the first being the log-normal diffusion, i.e.

$$dX_t = r X_t dt + \sigma X_t dW_t .$$

It is clear that this diffusion verifies Assumption H-1. In this case, we remark that the bias  $\mathcal{A}[S]_t$  and the variance  $\Gamma[S]_t$  become proportional respectively to  $X_t$  and  $X_t^2$ . As a result,

$$(2.5) \quad S_t^B = X_t [1 - \epsilon \mathbf{a} + \sqrt{\epsilon \gamma} L_t] ,$$

where  $\mathbf{a}$  and  $\gamma$  are constants, and the average price at which we sell a quantity  $\pi_i$  at time  $t_i$ , given by formula (2.2), is simplified and we have the following average price

$$(2.6) \quad \bar{S}_{t_i}^B(\pi_i) = X_{t_i} g(\pi_i) [1 - \epsilon \mathbf{a} + \sqrt{\epsilon \gamma} L_{t_i}] .$$

As such, we consider an extension of the Black Scholes model, which is the CEV model, see [11]. This model takes into account the heteroscedasticity of the asset returns and explains the down-sloping behavior of the implied volatility, see for instance [19].

#### Assumption H-3 (CEV diffusion)

The volatility function  $\sigma(t, X_t)$  is equal to  $\sigma X_t^\alpha$ , where  $\sigma$  is a positive constant and  $\alpha$  is constant and belongs to  $(-1, 1)$ . We also assume that  $X_t \geq \xi > 0$  for all  $t \in [0, T]$ . That is SDE (1.1) is replaced by the following SDE

$$(2.7) \quad dX_t = r X_t dt + \sigma X_t^{\alpha+1} dW_t .$$

The CEV diffusion, unfortunately, does not verify Assumption H-1. However, all previous results still hold and their proofs remain substantially the same with the main difference coming from some properties of CEV diffusion that can be found in [15].

Under Assumption H-3, we have also the following rewriting of Theorem 1.1.

#### Corollary 2.1 (Constant Elasticity of Variance model)

Under Assumption H-3, the result of Theorem 1.1 remains true and Equations (1.5) are replaced by

$$(2.8) \quad \left\{ \begin{array}{l} \Gamma[S]_t = \theta M_t^2 \int_0^t \frac{\sigma^2 X_s^{2\alpha+2}}{M_s^2} ds + \Gamma[S]_0 M_t^2 , \\ \mathcal{A}[S]_t = M_t \int_0^t \frac{\alpha(\alpha+1) X_s^{\alpha-1} \Gamma[S]_s - \theta \sigma X_s^{\alpha+1}}{2 M_s} [dW_s - \sigma(\alpha+1) X_s^\alpha ds] , \\ M_t = \mathcal{E} \left\{ \sigma(\alpha+1) \int_0^t X_s^\alpha dW_s + r t \right\} . \end{array} \right.$$

Moreover, the martingale parts of the Doob decomposition of  $\sqrt{\Gamma[S]_t}$  and  $\frac{\sqrt{\Gamma[S]_t}}{X_t}$  are respectively

$$(2.9) \quad \begin{cases} \sigma (\alpha + 1) X_t^\alpha \sqrt{\Gamma[S]_t} dW_t , \\ \sigma \alpha X_t^{\alpha-1} \sqrt{\Gamma[S]_t} dW_t . \end{cases}$$

**Proof.** The proof of the first part is just a simplification of relation (1.5) in the case of CEV diffusion. The second part is an easy application of Ito formula.  $\square$

### Remark 2.6

In particular, we may notice that in CEV case:

- The fluctuations of the absolute spread, which are proportional to  $\sqrt{\Gamma[S]}$ , are always positively correlated with the underlying  $X$ .
- The fluctuations of the relative spread, which are proportional to the ratio  $\frac{\sqrt{\Gamma[S]}}{X}$ , are negatively (resp. positively) correlated with the underlying  $X$  if the CEV-exponent  $\alpha$  is negative (resp. positive). The case usually treated in literature is when the CEV-exponent is smaller than one, for instance see [6]. Therefore, the relative bid-ask spread grows when the asset price falls, whereas the absolute bid-ask spread falls with the benchmark.

This remark is important, since it is well-known on financial markets that the bid-ask spread converges to zero (resp. infinity) when the asset price goes to zero (resp. infinity). Instead, the relative spread grows when the asset price goes to zero and converges to zero when the asset price goes to infinity. The previous remark says that our model can explain endogenously this effect with the evolution of the variance  $\Gamma[S]$  if we suppose that the parameter  $\alpha$  is negative, i.e. the diffusion is sub-linear. This case is usually presented in literature as a way to explain why the BS model overprices in-the-money call options and underprices out-of-the-money ones, see for instance [11] and [19].

## 3 Numerical results

In this section, we provide some numerical results of the optimal strategy of the liquidation problem. For this purpose, we use two models.

1. The Black Scholes model with the drift equal to zero. This model is used as a basic reference, since we have closed form expressions.
2. A CEV model with the drift equal to zero and a negative CEV-exponential. This model is used to evaluate the impact of sensitivity w.r.t. the Brownian motion.

We consider the following trading impact function  $g$

$$(3.1) \quad g(x) = \exp(-\lambda x) ,$$

where the constant  $\lambda < \frac{1}{N}$ , with  $N$  the total number of shares to liquidate. It is rather clear that function  $g$  verifies Assumption H-2 on the interval  $[0, N]$ . We restrict our function to the set  $[0, N]$  given the fact that the optimal liquidation strategy abstains to buy shares at any time.

The optimal liquidation strategy is determined by using the dynamic programming equation (2.4).

The classical approach to solve this kind of problem is to discretize all processes and to reduce the computation on a finite probability space, see [4]. Thus, we discretize our processes using Monte-Carlo simulations for CEV one and closed forms for log-normal one.

### 3.1 Black Scholes case

We first consider the case when the underlying  $X$  follows SDE (2.7) with  $\alpha = 0$ , i.e. the Black Scholes model. The dynamic programming principle (2.4) gives us the optimal strategy to liquidate our portfolio. The strategy depends on three factors:

- the level of the underlying  $X$ ,
- the value of the liquidity process  $L$ ,
- the residual quantity of stocks that we still have to sell.

In Figures 3, 4, and 5, we present our numerical results in the Black Scholes case, in particular, the dependencies of the optimal selling strategy on the three factors mentioned above.

#### Remark 3.7 (Black Scholes case)

- We find that the optimal strategy is completely independent with respect to underlying  $X_t$  (see Figure 3). This result is coherent with the literature, see [1], and [20], given the fact that the spread is proportional to the underlying price in this particular case.
- The optimal strategy is almost linear with respect to the number of remaining stocks (see Figure 4). A slight concavity is equally worth noticing. This effect is explained by the presence of an exponential cost (3.1) on optimization problem (2.3), which breaks the linearity of the problem and prevents very large orders.
- Finally, the main result is that the optimal strategy decreases when the bid-ask spread increases and the dependence is almost linear till the spread is lower than its long term average (see Figure 5). When the spread is bigger than its mean, the optimal strategy is to keep all remaining stocks. This result is very interesting since it says that the optimal strategy depends mainly on the bid-ask spread and its equilibrium law. The optimal strategy can be resumed by we have to sell when the spread is small and to wait a better time when it is wide.



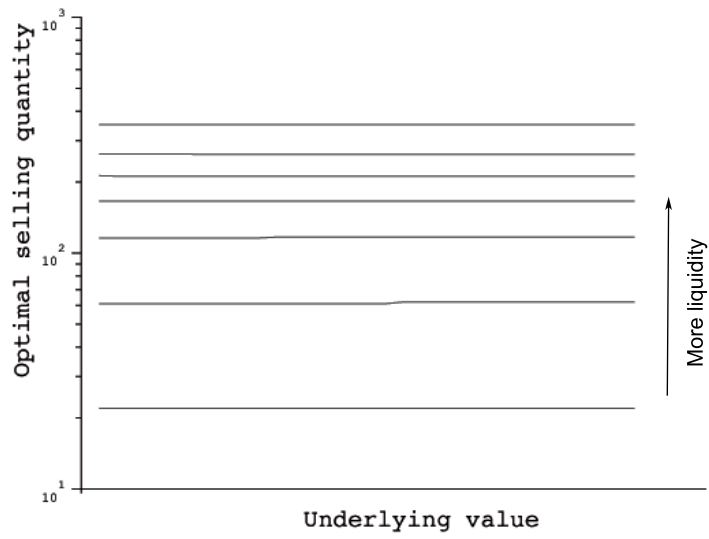


Figure 3: BS case: Optimal selling quantities as function of the underlying value for different values of liquidity process  $L_t$ .

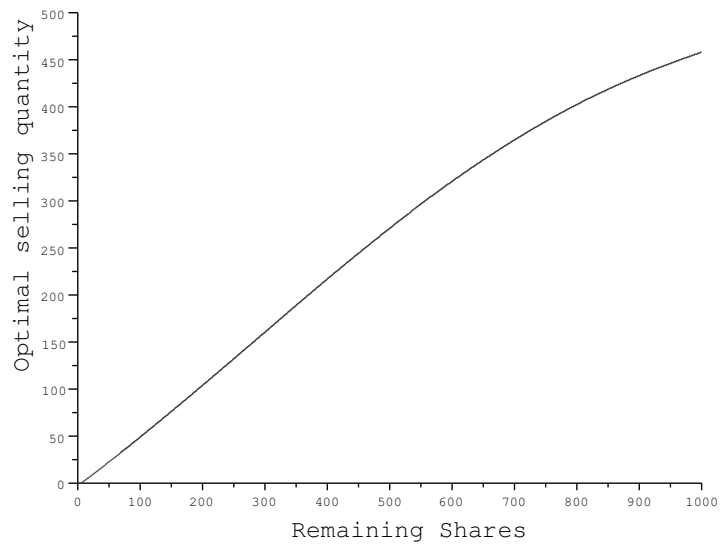


Figure 4: BS case: Optimal selling quantities as function of the remaining shares owned by the investor.

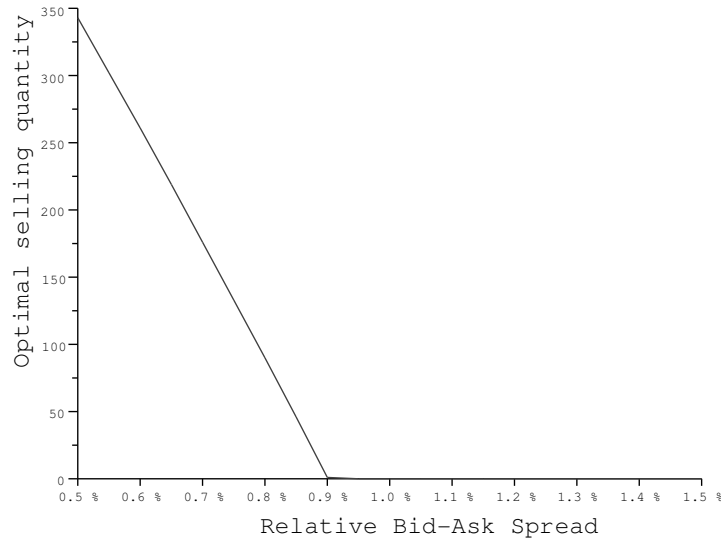


Figure 5: BS case: Optimal selling quantities as function of the relative bid-ask spread.

### 3.2 CEV case

We consider the case when the underlying follows SDE (2.7) with  $\alpha = -0.7$ . The numerical simulations show that the results found in the Black Scholes case remain true, except for the dependencies of the optimal selling strategies on underlying value (see Figure 6).

**Remark 3.8** (*CEV Case*)

1. The optimal selling strategy is completely unaffected by underlying price when the stock is highly liquid. However, when it is illiquid, the optimal selling strategy is positively correlated with the price of underlying asset.
2. We also have analyzed the impact of a change on the CEV exponential  $\alpha$ . When this parameter increases to zero, the dependency on the price of underlying asset is lessened.

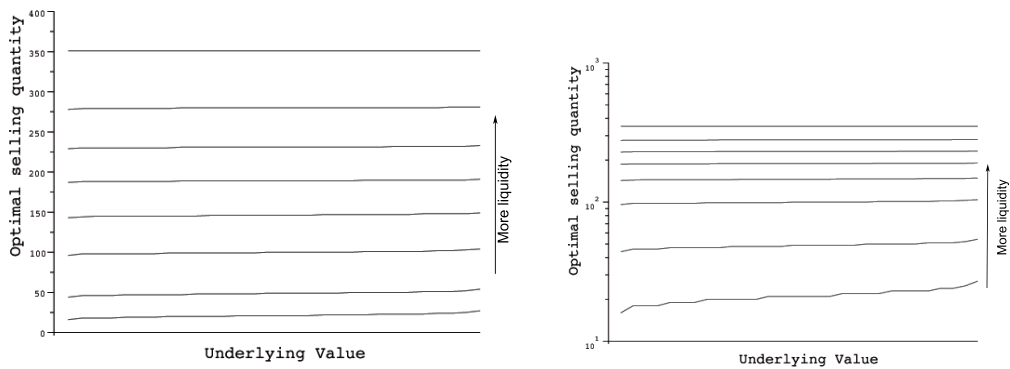


Figure 6: CEV case: Optimal selling quantities as function of the underlying value with linear scale (left) and logarithmic scale (right).

**Economic interpretation/explanation:** We may explain the effect mentioned in the first point of Remark 3.8 by the non-linear dependency of the bid-ask spread on the underlying value (see Corollary 2.1). Indeed, when the price of underlying asset falls, the relative bid-ask spread (in percentage of the asset price) increases, which in turn incites the investor to delay the selling or sell a smaller number of shares.

The below Figure 7 shows the shape of the selling region (above the curve) and the non-selling region (below the curve) at a given time and a given level of stock price.

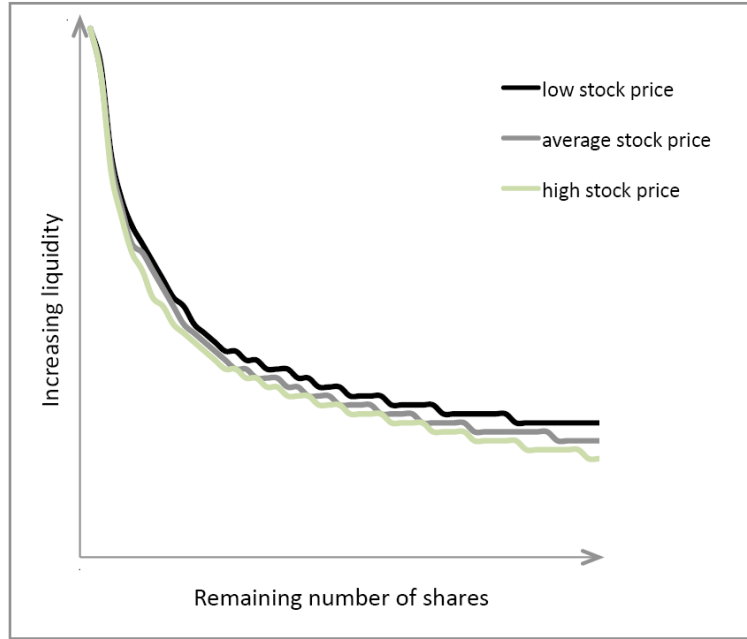


Figure 7: CEV case: An example of the selling regions as function of the remaining number of shares.

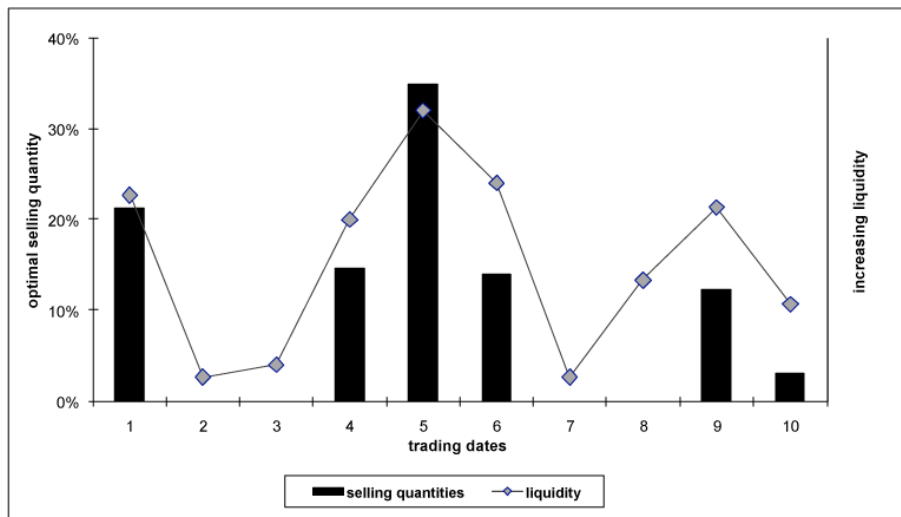


Figure 8: CEV case: An example of the optimal selling strategy given a liquidity trajectory.

## 4 Acknowledge

We are grateful to the anonymous referee for constructive suggestions and comments.

# Appendix

## A Proofs of Lemmas 1.1 and 1.2

We start with the proof of Lemma 1.1.

**Proof:** The proof is split into three steps.

**Step 1:** We compute the SDE satisfied by the sharp of  $S$ , see Section V.2 in [8] for the definition and more details:

$$dS_t^\# = r S_t^\# dt + \zeta(t, X_t) S_t^\# dW_t + \sqrt{\theta} \sigma(t, X_t) X_t d\widetilde{W}_t,$$

where  $\widetilde{W}$  is an independent Brownian motion defined in a probability space  $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$  copy of the original probability space. Assumption H-1 insures that the previous SDE admits a square integrable solution.

**Step 2:** We apply Ito formula to  $(S^\#)^2$  and take the expectation under the probability  $\widetilde{\mathbb{P}}$ , since one of the properties of the sharp operator is that  $\Gamma[S]_t = \widetilde{\mathbb{E}} \left[ (S_t^\#)^2 \right]$ , see Sections VI.2 and VII.4 in [8]. Therefore we find SDE (1.6).

**Step 3:** Finally, we prove that SDE (1.6) admits the closed form solution (1.5). Using the methods developed in Section 5 in [9] and noticing that the SDE verified by the sharp is linear, we may apply a variation of constant method, see for instance V.9 in [23], and obtain a closed form for  $S^\#$ . Then, we compute easily the expectation under  $\widetilde{\mathbb{P}}$  of the square of  $S^\#$ . Another possibility is to check that the first equation in (1.5) is solution to SDE (1.6).

□

We now turn to the proof of Lemma 1.2.

**Proof:** The proof is based on an  $L^2$ -convergence argument, see for instance [13], by using the fact that operators  $\Gamma[\cdot]$  and  $\mathcal{A}[\cdot]$  are closed, see [8].

We define a partition  $\{\tau_i\}_{i=0,\dots,n}$  of the interval  $[0, T]$ , where  $T$  is a sufficient large time. We approximate  $X$  with the following process

$$Z_t = \sum_{i=0}^{n-1} \sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i} (W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}) + r \sum_{i=0}^{n-1} Z_{\tau_i} (\tau_{i+1} \wedge t - \tau_i \wedge t).$$

It is clear that  $Z$  converges to  $X$  when the partition step goes to zero, thanks to hypotheses 1 and 3 of Assumption H-1. Then, we apply the bias operator on  $Z$  (see Section

6 in [9]), and we find

(A.1)

$$\begin{aligned}
\mathcal{A}[Z]_t &= \sum_{i=0}^{n-1} \left\{ \zeta(\tau_i, Z_{\tau_i}) \mathcal{A}[Z]_{\tau_i} (W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}) + r \mathcal{A}[Z]_{\tau_i} (\tau_{i+1} \wedge t - \tau_i \wedge t) \right\} \\
&+ \sum_{i=0}^{n-1} \left\{ \sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i} \mathcal{A} [W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}] + \zeta(\tau_i, Z_{\tau_i}) \tilde{\mathbb{E}} \left[ Z_{\tau_i}^{\#} (W_{\tau_{i+1} \wedge t}^{\#} - W_{\tau_i \wedge t}^{\#}) \right] \right\} \\
&+ \frac{1}{2} \sum_{i=0}^{n-1} \eta(\tau_i, Z_{\tau_i}) Z_{\tau_i} \Gamma [Z_{\tau_i}] (W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}) .
\end{aligned}$$

Then, we take the limit in SDE (A.1), when the step of the partition goes to zero and we have to prove that equation (A.1) converges to the integral form of equation (1.7). For sake of simplicity on notation, we compute all equations at the final time  $T$ . We start with the remark that  $Z$  converges to  $X$  in  $L^2$ -norm when the step partition goes to zero. It's also clear that  $\Gamma[Z]$  converges to  $\Gamma[S]$  due to the fact that  $\Gamma$  is a closed operator, see for instance [8], and the result of Lemma 1.1, this convergence is in  $L^1$ -norm but it is easy to check that it is true in  $L^2$ -norm too. Then, under Assumption H-1, we will prove that

$$(A.2) \quad \sum_{i=0}^{n-1} \eta(\tau_i, Z_{\tau_i}) \Gamma[Z]_{\tau_i} (W_{\tau_{i+1} \wedge T} - W_{\tau_i \wedge T}) \xrightarrow{L^2} \int_0^T \eta(t, X_t) \Gamma[S]_t dW_t .$$

We separate the last integral using the partition  $(\tau_i)_{i=0, \dots, n}$  and we evaluate the difference in  $L^2$ -norm, so we find

$$\begin{aligned}
&\mathbb{E} \left[ \left\{ \eta(\tau_i, Z_{\tau_i}) \Gamma[Z]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) - \int_{\tau_i}^{\tau_{i+1}} \eta(t, X_t) \Gamma[S]_t dW_t \right\}^2 \right] \\
&= \mathbb{E} \left[ \left\{ \eta(\tau_i, Z_{\tau_i}) \Gamma[Z]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) - \eta(\tau_i, X_{\tau_i}) \Gamma[Z]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) \right. \right. \\
&\quad \left. \left. + \eta(\tau_i, X_{\tau_i}) \Gamma[Z]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) - \eta(\tau_i, X_{\tau_i}) \Gamma[S]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) \right. \right. \\
&\quad \left. \left. + \eta(\tau_i, X_{\tau_i}) \Gamma[S]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) - \int_{\tau_i}^{\tau_{i+1}} \eta(t, X_t) \Gamma[S]_t dW_t \right\}^2 \right] \\
&< \mathbb{E} \left[ \left\{ \eta(\tau_i, Z_{\tau_i}) - \eta(\tau_i, X_{\tau_i}) \right\} \Gamma[Z]_{\tau_i} (W_{\tau_{i+1}} - W_{\tau_i}) \right]^2 + \\
&\quad + \mathbb{E} \left[ \left\{ \eta(\tau_i, X_{\tau_i}) [\Gamma[Z]_{\tau_i} - \Gamma[S]_{\tau_i}] (W_{\tau_{i+1}} - W_{\tau_i}) \right\}^2 \right] + \\
&\quad + \mathbb{E} \left[ \left\{ \eta(\tau_i, X_{\tau_i}) \Gamma[S]_{\tau_i} (W_{\tau_i} - W_{\tau_{i-1}}) - \int_{\tau_i}^{\tau_{i+1}} \eta(t, X_t) \Gamma[S]_t dW_t \right\}^2 \right] .
\end{aligned}$$

The first expectation converges to zero thanks to the Lipschitz hypothesis on  $\eta(t, x)$  w.r.t.  $x$ . The second expectation goes to zero using the fact that  $\Gamma[Z]$  converges to  $\Gamma[S]$  in  $L^2$ -norm. The last expectation converges to zero when the partition step goes to zero in accordance with the definition of the stochastic integral. Then the limit (A.2) is proved.

Using the same arguments and Gronwall lemma, see for instance Chapter V in [23], we have

$$(A.3) \quad \sum_{i=0}^{n-1} \zeta(\tau_i, Z_{\tau_i}) \mathcal{A}[Z]_{\tau_i} (W_{\tau_{i+1} \wedge T} - W_{\tau_i \wedge T}) \xrightarrow{L^2} \int_0^T \zeta(t, X_t) \mathcal{A}[S]_t dW_t$$

and

$$(A.4) \quad \sum_{i=0}^{n-1} \mathcal{A}[Z]_{\tau_i} (\tau_{i+1} - \tau_i) \xrightarrow{L^2} \int_0^T \mathcal{A}[S]_t dt .$$

We study the third term in equation (A.1). We find

$$\sum_{i=0}^{n-1} \sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i} \mathcal{A} [W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}] = -\frac{\theta}{2} \sum_{i=0}^{n-1} \sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i} [W_{\tau_{i+1}} - W_{\tau_i}],$$

thanks to the chain rule of semigroup  $\mathcal{A}$ , see Section 3 in [9]. We also remark that  $\sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i}$  converges to  $\sigma(\tau_i, X_{\tau_i}) X_{\tau_i}$ , thanks to Assumption H-1. Using always the same arguments used to prove limit (A.2), we have therefore that

$$(A.5) \quad \sum_{i=0}^{n-1} \sigma(\tau_i, Z_{\tau_i}) Z_{\tau_i} \mathcal{A} [W_{\tau_{i+1} \wedge t} - W_{\tau_i \wedge t}] \xrightarrow{L^2} -\frac{\theta}{2} \int_0^T \sigma(t, X_t) X_t dW_t.$$

Finally, we analyze the term

$$\sum_{i=0}^{n-1} \zeta(\tau_i, Z_{\tau_i}) \tilde{\mathbb{E}} \left[ \left( Z_{\tau_i}^\# \right) \left( W_{\tau_{i+1} \wedge t}^\# - W_{\tau_i \wedge t}^\# \right) \right].$$

We introduce a conditional expectation w.r.t. the  $\sigma$ -algebra  $\tilde{\mathcal{F}}_{\tau_i} = \sigma\{(W_s^\#, W_u) \mid u, s < \tau_i\}$

$$\sum_{i=0}^{n-1} \tilde{\mathbb{E}} \left[ \tilde{\mathbb{E}} \left[ \left( Z_{\tau_i}^\# \right) \left( W_{\tau_{i+1} \wedge t}^\# - W_{\tau_i \wedge t}^\# \right) \mid \tilde{\mathcal{F}}_{\tau_i} \right] \right] = \sum_{i=0}^{n-1} \tilde{\mathbb{E}} \left[ \left( Z_{\tau_i}^\# \right) \tilde{\mathbb{E}} \left[ W_{\tau_{i+1}}^\# - W_{\tau_i}^\# \mid \tilde{\mathcal{F}}_{\tau_i} \right] \right] = 0,$$

using the fact that  $Z^\#$  is adapted to the filtration  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  and  $W^\#$  remains a Brownian motion w.r.t. filtration  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$ , since  $W^\#$  and  $W$  are independent. As a consequence, the fourth term in equation (A.1) is always equal to zero and we have proved the convergence of equation (A.1) to the integral form of equation (1.7). Now it is easy to check that the second equation in (1.5) solves SDE (1.7)

□

## References

- [1] Alfonsi, A., A. Schied and A. Schulz (2007): “Optimal execution strategies in limit order books with general shape functions”, preprint.
- [2] Back, K. (1992) : “Insider trading in continuous time", *Review of Financial Studies*, **5**, 387-409.
- [3] Bank, P. and D. Baum (2004): “Hedging and portfolio optimization in illiquid financial markets with a large trader”, *Mathematical Finance*, **14**, 1-18.
- [4] Bardou, O., S. Bouthemy and G. Pages (2009): “Optimal Quantization for the Pricing of Swing Options”, *Applied Mathematical Finance*, **16-2**, 183-217.
- [5] Bertsekas, D. and E. Shreve (1978): “Stochastic Optimal Control: The Discrete Time Case”, *Mathematics in Science and Engineering*, **139**.
- [6] Black, F. (1976): “Studies of Stock Price volatility Changes”, *Pr. Meeting American Stat. Association, Business and Economics Statistic Division*, 177-181.
- [7] Bouchaud, J.P., M. Mézard and M. Potters (2002): “Statistical properties of stock order books: empirical results and models”, preprint.

- [8] Bouleau, N. (2003): “Error Calculus for Finance and Physics”, De Gruyter, Berlin.
- [9] Bouleau, N. (2003): “Error Calculus and path sensitivity in financial models”, *Mathematical Finance*, **13-1**, 115-134.
- [10] Cetin, U., R. Jarrow and P. Protter (2004): “Liquidity risk and arbitrage pricing theory”, *Finance and Stochastics*, **8**, 311-341.
- [11] Cox, J. and S. Ross (1976): “The valuation of options for alternative stochastic processes”, *Journal of Financial Economics*, **3**, 145-166.
- [12] Cuoco, D. and J. Cvitanic (1998): “Optimal consumption choice for a large investor”, *Journal of Economic Dynamics and Control*, **22**, 401-436.
- [13] Da Prato, G. (2006): “An Introduction to Infinite-Dimensional Analysis”, Springer, Berlin.
- [14] Frino, A. and D. Gallagher (2001): “Tracking S&P 500 Index Funds”. *J. Portfolio Management*, **28**, 44-55.
- [15] Jeanblanc, M., M. Yor and M. Chesney (2009): “Mathematical Methods for Financial Markets”, Springer Finance.
- [16] Korn, R. (1998) : “Portfolio optimization with strictly positive transaction costs and impulse control”, *Finance and Stochastics*, **2**, 85-114.
- [17] Kyle, A. (1985) : “Continuous actions and insider trading”, *Econometrica*, **53**, 1315-1335.
- [18] Ly Vath, V., M. Mnif and H. Pham (2007): “A Model of Optimal Portfolio Selection under Liquidity Risk and Price Impact”, *Finance and Stochastics*, **11**, 51-90.
- [19] Macbeth, J. and L. Merville (1980): “Test of the Black Scholes and Cox Call Option Valuation Models”, *Journal of Finance*, **35** 285-301.
- [20] Obizhaeva, A. and J. Wang (2005): “Optimal Trading Strategy and Supply/Demand Dynamics”, preprint.
- [21] Parlour, C.A. (1998): “Price dynamics in limit order markets”, *Review of Financial Studies*, **11**, 789-816.
- [22] Potters, M. and J.P. Bouchaud (2003): “More statistical properties of order books and price impact”, *Physica A*, **243**, 133-140.
- [23] Protter, P. (1990): “Stochastic integration and differential equations”, Springer Verlag.
- [24] Wang, H.K. and J. Yau (2000): “Trading volume; bid-ask spread, and price volatility in futures markets”, *Journal of Futures markets*, **20**, 943-970.