Optimal market making

Olivier Guéant

📅 Published online: 02 Aug 2017.

✎ Submit your article to this journal ↗

🔍 View related articles ↗

⬛ View Crossmark data ↗

Routledge
Taylor & Francis Group

Check for updates

Optimal market making

Olivier Guéant[*]

Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, Paris, France

ABSTRACT

Market makers provide liquidity to other market participants: they propose prices at which they stand ready to buy and sell a wide variety of assets. They face a complex optimization problem with both static and dynamic components. They need indeed to propose bid and offer/ask prices in an optimal way for making money out of the difference between these two prices (their bid–ask spread). Since they seldom buy and sell simultaneously, and therefore hold long and/or short inventories, they also need to mitigate the risk associated with price changes and subsequently skew their quotes dynamically. In this paper, (i) we propose a general modelling framework which generalizes (and reconciles) the various modelling approaches proposed in the literature since the publication of the seminal paper 'High-frequency trading in a limit order book' by Avellaneda and Stoikov, (ii) we prove new general results on the existence and the characterization of optimal market making strategies, (iii) we obtain new closed-form approximations for the optimal quotes, (iv) we extend the modelling framework to the case of multi-asset market making and we obtain general closed-form approximations for the optimal quotes of a multi-asset market maker, and (v) we show how the model can be used in practice in the specific (and original) case of two credit indices.

1. Introduction

What is a market maker? In a nutshell, it is a liquidity provider. However, it is complex to give a precise definition because the exact role of market makers depends on the considered market. Furthermore, the very definition of a market maker has been blurred in recent years, because of the electronification of most markets and because of the emergence of high-frequency trading in many of them.

On most order-driven markets, such as many stock markets, there are nowadays several kinds of market makers. First, there are 'official' market makers (actually, market making companies): these market makers have usually signed an agreement with a given exchange, or with a given company, for maintaining fair and orderly markets. The designated market makers (DMM) on the NYSE, which succeeded the market specialists, are examples of such 'official' market makers. They often have contractual obligations, such as participating to the opening and closing auctions

CONTACT Olivier Guéant ✉ olivier.gueant@univ-paris1.fr 🖃 Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, 106 Boulevard de l'Hôpital, Paris, France

*The author was initially affiliated with ENSAE and CREST (3 avenue Pierre Larousse, 92245 Malakoff Cedex, France).

and/or quoting with a reasonable bid–ask spread – for example, the DMMs must quote at the National Best Bid and Offer a specified percentage of the time. In addition to these 'official' market makers, other market participants in the stock markets, in particular some high-frequency traders, are often regarded as market makers (Menkveld calls them the new market makers in Menkveld 2013) because they are almost continuously present on both sides of the limit order books. They are acting as liquidity providers even though they have no obligation to do so: they just try to make money out of their high-frequency market making strategies. The electronification of most order-driven markets makes it possible for trading firms to act as liquidity providers, hence a blurring of the definition of 'market maker'.

On quote-driven markets, such as the corporate bond markets, the market makers are the dealers (these markets are often also called 'dealer markets'). These dealers provide liquidity to the other market participants (the 'clients') by quoting bid and offer prices on a regular basis. However, their exact behaviour depends on the considered market. On some markets, dealers' quotes are firm quotes, whereas on other markets the quotes are streamed only for information (and for a specific size/notional) and become binding when dealers answer specific requests.

In this paper, we consider that a market maker is somebody (or in fact an algorithm) who proposes prices at which he/she/it stands ready to buy or sell one or several assets. In particular, we do not consider any contractual constraint, and we assume that all quotes are firm quotes (for a given fixed size). The problem we consider is the determination of the optimal quotes a market maker should propose at the bid and the offer to make money while mitigating inventory risk.

This problem is a complex one from a quantitative viewpoint with both static and dynamic components. Market makers face indeed a classical static trade-off: high margin and low volume vs. low margin and high volume. A market maker who quotes a large spread (with no skew) trades rarely, but each transaction leads to a large Mark-to-Market (MtM) gain. Conversely, a market maker quoting a narrow spread (with no skew) trades often, but each transaction leads to a small MtM gain. In addition to this static trade-off, market makers face a dynamic problem: they must adapt their quotes dynamically to reduce their exposure to price changes. For instance, a single-asset market maker with a long inventory should price conservatively on the bid side and aggressively on the ask side, because he wants to reduce his probability to buy and increase his probability to sell. Symmetrically, if he has a short inventory, then he should price aggressively on the bid side and conservatively on the ask side.

Like in almost all the mathematical literature on market making, we consider the problem of a single market maker in a simplified way: (i) market prices[1] are modelled by stochastic processes assumed to be exogenous to the market maker's behaviour[2] and (ii) the probability that the market maker buys (respectively sells) a security at the bid (respectively offer) price he quotes depends on the distance between the quoted price and the market price of that security – this is the classical Avellaneda–Stoikov modelling framework – see Avellaneda and Stoikov (2008). In particular, the competition between market makers is not explicitly modelled.

Since the publication of the seminal paper 'High-frequency trading in a limit order book' by Avellaneda and Stoikov (see Avellaneda and Stoikov 2008), market making has

been one of the important research topics in quantitative finance.[3] Therefore, many models have been proposed to address the problem faced by market makers. Guéant, Lehalle, and Fernandez-Tapia (2013) considered a variant of the model proposed by Avellaneda and Stoikov and showed that the four-dimensional Hamilton–Jacobi–Bellman (HJB) equation arising from the model could be simplified into a linear system of ordinary differential equations (ODEs) when a specific change of variables is used.[4] Guéant, Lehalle, and Fernandez-Tapia (2013) also contains the Guéant–Lehalle–Fernandez-Tapia formulas which are closed-form approximations of the optimal quotes of a single-asset market maker. These approximation formulas are used in practice by major banks in Europe and Asia for market making in (illiquid) quote-driven markets or for market making in some order-driven markets (in the specific case of a small tick size).

In the aforementioned papers, the objective function of the market maker is the expected CARA utility[5] of his P&L (sometimes with a penalty for the terminal inventory). Other models have been proposed in the literature with different objective functions. In their paper on market making with general price dynamics, Fodra and Labadie (2012) considered, in addition to the expected CARA utility case, the risk-neutral case and the risk-neutral case with a penalization on the terminal inventory. In a few papers, with various co-authors, and in their recent book (Cartea, Jaimungal, and Penalva 2015) with Penalva, Cartea and Jaimungal considered as an objective function the expected value of the P&L minus a running penalty on the inventory – see for instance Cartea, Donnelly, and Jaimungal (2013), Cartea and Jaimungal (2013), and Cartea, Jaimungal, and Ricci (2014).

The numerous researchers involved in market making modelling have also included many features in their models. Cartea and Jaimungal, with their co-authors, have proposed models with price impact, the possibility to consider short-term alpha, the existence of an adverse selection effect,[6] etc. Recently, new models have emerged to deal with ambiguity aversion: see for instance Cartea, Donnelly, and Jaimungal (2013) and Nyström, Ould Aly, and Zhang (2014) – see also the PhD dissertation of Donnelly (Donnelly 2014).

For strange reasons,[7] academic researchers have mainly focused on stock markets, which are certainly the least relevant markets to apply most of the models they have proposed.[8] In this paper, we have instead in mind the case of a market maker in a quote-driven market, or in an order-driven market if the tick sizes of the securities are small.

The academic literature on market making is also mainly focused on the case of a market maker operating on a single asset. However, in practice, almost all market makers are in charge of a list of securities. For a market maker in charge of several correlated assets, applying an independent market making strategy to each asset is suboptimal in terms of risk management. It is therefore of the utmost importance to build a model accounting for the correlation structure of the security price moves, especially in the case of corporate bonds where there are often dozens of bonds issued by the same company (which are therefore highly correlated).

In this paper, we consider a modelling framework à la Avellaneda–Stoikov with general intensity functions, instead of the exponential intensity functions of most models (see Section 2). We show that the four-variable HJB equation arising from the

various optimization criteria used in the literature can be transformed into a simple system of ODEs (see Section 3). This somehow reconciles the different approaches used in the literature and enables to understand the subtle differences between the various criteria used in the literature. In particular it helps understanding what it means to be averse to price risk and to non-execution risk. We then show in Section 4 how to find closed-form approximations for the optimal quotes. These approximations generalize the Guéant–Lehalle–Fernandez-Tapia formulas to the case of general intensity functions and to the case of the different optimization criteria used in the market making literature. In Section 5, we consider a problem that is very rarely dealt with in the academic literature in spite of its importance for practitioners: multi-asset market making. We show that many results obtained in the one-asset case can be generalized to our multi-asset market making model. In particular, we obtain for the first time in this paper closed-form approximations for the optimal quotes of a multi-asset market maker. This result is an important breakthrough for practitioners because most market makers are in charge of dozens of assets (or even hundreds of assets when the market maker is in fact an algorithm) and often reluctant to solve very large systems of nonlinear differential equations. In Section 6, we apply our findings to the case of two highly correlated credit indices: CDX.NA.IG (CDX North America Investment Grade) and CDX.NA.HY (CDX North America High Yield).

2. Modelling framework and notations

2.1. *Notations*

Let us fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ satisfying the usual conditions. We assume that all stochastic processes are defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$.

We consider in this section (and in the following two sections) a market maker in charge of a single asset. The reference price of this asset[9] is modelled by a process $(S_t)_t$ with the dynamics

$$dS_t = \sigma dW_t, \quad S_0 \text{ given}, \tag{2.1}$$

where $(W_t)_t$ is a standard Brownian motion adapted to the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$.

This market maker proposes bid and ask quotes to buy and sell the asset. These bid and ask quotes are modelled by two stochastic processes, respectively, denoted by $(S_t^b)_t$ and $(S_t^a)_t$.

Transactions occur at random times corresponding to the arrival times of agents willing to buy or sell the asset. The distribution of the trade times depends obviously on the liquidity of the asset and on the bid and ask prices quoted by the market maker. We denote by $(N_t^b)_t$ and $(N_t^a)_t$ the two point processes modelling the number of transactions at the bid and at the ask, respectively. We assume that assets are traded Δ by Δ, that is, that the quantities traded do not vary across trades.

The inventory of the market maker, modelled by the process $(q_t)_t$, has therefore the dynamics

$$dq_t = \Delta dN_t^b - \Delta dN_t^a, \quad q_0 \text{ given}. \tag{2.2}$$

We assume that the processes $(N_t^b)_t$ and $(N_t^a)_t$ are independent of the Brownian motion $(W_t)_t$. We denote by $(\lambda_t^b)_t$ and $(\lambda_t^a)_t$ the intensity processes of $(N_t^b)_t$ and $(N_t^a)_t$, respectively. As in the classical Avellaneda–Stoikov model (see Avellaneda and Stoikov 2008), we assume that the intensity processes are functions of the difference between the reference price and the prices quoted by the market maker. In addition, we assume that the market maker stops proposing a bid (respectively ask) quote when his position is above (respectively below) a given threshold Q (respectively $-Q$).[10] Formally, we assume that $(\lambda_t^b)_t$ and $(\lambda_t^a)_t$ verify

$$\lambda_t^b = \Lambda^b(\delta_t^b)1_{q_{t-}<Q} \quad \text{and} \quad \lambda_t^a = \Lambda^a(\delta_t^a)1_{q_{t-}>-Q}, \tag{2.3}$$

where

$$\delta_t^b = S_t - S_t^b \quad \text{and} \quad \delta_t^a = S_t^a - S_t,$$

and where Λ^b and Λ^a are two functions satisfying the following hypotheses:[11]

- Λ^b and Λ^a are twice continuously differentiable,
- Λ^b and Λ^a are decreasing, with $\forall \delta \in \mathbb{R}$, $\Lambda^{b\prime}(\delta) < 0$ and $\Lambda^{a\prime}(\delta) < 0$,
- $\lim_{\delta \to +\infty} \Lambda^b(\delta) = \lim_{\delta \to +\infty} \Lambda^a(\delta) = 0$,
- $\sup_\delta \frac{\Lambda^b(\delta)\Lambda^{b\prime\prime}(\delta)}{\left(\Lambda^{b\prime}(\delta)\right)^2} < 2$ and $\sup_\delta \frac{\Lambda^a(\delta)\Lambda^{a\prime\prime}(\delta)}{\left(\Lambda^{a\prime}(\delta)\right)^2} < 2$.

Finally, the process $(X_t)_t$ models the market maker's cash account. Given our modelling framework, $(X_t)_t$ has the dynamics

$$\begin{aligned}
dX_t &= S_t^a \Delta dN_t^a - S_t^b \Delta dN_t^b \\
&= (S_t + \delta_t^a)\Delta dN_t^a - (S_t - \delta_t^b)\Delta dN_t^b.
\end{aligned} \tag{2.4}$$

2.2. *The two classical optimization problems*

In the previous paragraphs, we have defined the three processes at the heart of most market making models: the reference price process $(S_t)_t$, the inventory process $(q_t)_t$ and the cash process $(X_t)_t$. We now need to define the problem faced by the market maker. Following the model proposed by Avellaneda and Stoikov in 2008, one can consider, as in Guéant, Lehalle, and Fernandez-Tapia (2013), that the market maker maximizes the expected value of a CARA utility function (with risk aversion parameter $\gamma > 0$) applied to the MtM value of the portfolio at a given date T. This MtM value at time T is basically $X_T + q_T S_T$, or $X_T + q_T S_T - \ell(|q_T|)$ if we add a liquidity premium for the remaining inventory (whatever its sign) – ℓ is a non-decreasing and convex function from $\mathbb{R}_+$ to $\mathbb{R}_+$. In this general framework, the goal of the market maker is to maximize

$$\mathbb{E}[-\exp(-\gamma(X_T + q_T S_T - \ell(|q_T|)))], \qquad \text{(Model A)}$$

over $(\delta_t^b)_t \in \mathcal{A}$ and $(\delta_t^a)_t \in \mathcal{A}$, where the set of admissible controls $\mathcal{A}$ is simply the set of predictable processes bounded from below.

Alternatively, one can consider that the market maker maximizes the expected value of the MtM value of the portfolio at date T, but that holding an inventory is penalized

over the time interval $[0, T]$. This is typically what is done by Cartea, Jaimungal and their co-authors (see the recent book Cartea, Jaimungal, and Penalva (2015) for several examples). In that kind of model, the goal of the market maker is to maximize an expression of the form

$$\mathbb{E}\left[X_T + q_T S_T - \ell(|q_T|) - \frac{1}{2}\gamma\sigma^2 \int_0^T q_t^2 dt\right], \qquad \text{(Model B)}$$

over $(\delta_t^b)_t \in \mathcal{A}$ and $(\delta_t^a)_t \in \mathcal{A}$.

3. Towards a single system of ODEs for characterizing the optimal quotes

Both Model A and Model B can be solved using the classical tools of stochastic optimal control. In particular, we show that, in both models, finding the value function (and the optimal bid and ask quotes) boils down to solving a tridiagonal system of ODEs and that the equations associated with Model A and Model B are part of the same family of ODEs.

3.1. Dimensionality of the problem: a reduction from 4 to 2

The HJB equation associated with Model A is given by

$$0 = -\partial_t u(t, x, q, S) - \frac{1}{2}\sigma^2 \partial_{SS}^2 u(t, x, q, S) \tag{3.1}$$

$$- 1_{q < Q} \sup_{\delta^b} \Lambda^b(\delta^b)\left[u(t, x - \Delta S + \Delta\delta^b, q + \Delta, S) - u(t, x, q, S)\right]$$

$$- 1_{q > -Q} \sup_{\delta^a} \Lambda^a(\delta^a)[u(t, x + \Delta S + \Delta\delta^a, q - \Delta, S) - u(t, x, q, S)],$$

for $q \in \mathcal{Q} = \{-Q, -Q + \Delta, \ldots, Q - \Delta, Q\}$, and $(t, S, x) \in [0, T] \times \mathbb{R}^2$, with the terminal condition

$$u(T, x, q, S) = -\exp(-\gamma(x + qS - \ell(|q|))). \tag{3.2}$$

If one uses the ansatz

$$u(t, x, q, S) = -\exp(-\gamma(x + qS + \theta(t, q))), \tag{3.3}$$

then Equation (3.1) becomes

$$0 = -\partial_t \theta(t, q) + \frac{1}{2}\gamma\sigma^2 q^2 \tag{3.4}$$

$$-1_{q<Q} \sup_{\delta^b} \frac{\Lambda^b(\delta^b)}{\gamma}\left(1 - \exp\left(-\gamma\left(\Delta\delta^b + \theta(t, q + \Delta) - \theta(t, q)\right)\right)\right)$$

$$-1_{q>-Q} \sup_{\delta^a} \frac{\Lambda^a(\delta^a)}{\gamma}\left(1 - \exp\left(-\gamma\left(\Delta\delta^a + \theta(t, q - \Delta) - \theta(t, q)\right)\right)\right),$$

for $q \in \mathcal{Q}$, and $t \in [0, T]$, and the terminal condition (3.2) becomes $\theta(T, q) = -\ell(|q|)$. The HJB equation associated with Model B is given by

$$0 = -\partial_t u(t, x, q, S) + \tfrac{1}{2}\gamma\sigma^2 q^2 - \tfrac{1}{2}\sigma^2 \partial_{SS}^2 u(t, x, q, S) \tag{3.5}$$

$$-1_{q<Q} \sup_{\delta^b} \Lambda^b(\delta^b) \Big[u(t, x - \Delta S + \Delta\delta^b, q + \Delta, S) - u(t, x, q, S) \Big]$$

$$- 1_{q>-Q} \sup_{\delta^a} \Lambda^a(\delta^a) \Big[u(t, x + \Delta S + \Delta\delta^a, q - \Delta, S) - u(t, x, q, S) \Big],$$

for $q \in \mathcal{Q}$, and $(t, S, x) \in [0, T] \times \mathbb{R}^2$, with the terminal condition

$$u(T, x, q, S) = x + qS - \ell(|q|). \tag{3.6}$$

If one uses the ansatz

$$u(t, x, q, S) = x + qS + \theta(t, q), \tag{3.7}$$

then Equation (3.5) becomes

$$0 = -\partial_t \theta(t, q) + \tfrac{1}{2}\gamma\sigma^2 q^2 \tag{3.8}$$

$$-1_{q<Q} \sup_{\delta^b} \Lambda^b(\delta^b) \Big(\Delta\delta^b + \theta(t, q + \Delta) - \theta(t, q) \Big)$$

$$- 1_{q>-Q} \sup_{\delta^a} \Lambda^a(\delta^a) \Big(\Delta\delta^a + \theta(t, q - \Delta) - \theta(t, q) \Big),$$

for $q \in \mathcal{Q}$, and $t \in [0, T]$, and the terminal condition (3.10) becomes $\theta(T, q) = -\ell(|q|)$. Equations (3.4) and (3.8) are in fact two systems of ODEs which belong to the same family. If we introduce for $\xi > 0$ the functions

$$H_\xi^b(p) = \sup_\delta \frac{\Lambda^b(\delta)}{\xi} (1 - \exp(-\xi\Delta(\delta - p)))$$

and

$$H_\xi^a(p) = \sup_\delta \frac{\Lambda^a(\delta)}{\xi} (1 - \exp(-\xi\Delta(\delta - p))),$$

and the limit functions (for $\xi = 0$)

$$H_0^b(p) = \Delta \sup_\delta \Lambda^b(\delta)(\delta - p)$$

and

$$H_0^a(p) = \Delta \sup_\delta \Lambda^a(\delta)(\delta - p),$$

then we can indeed consider the general equation

$$0 = -\partial_t \theta(t, q) + \tfrac{1}{2}\gamma\sigma^2 q^2 \tag{3.9}$$

$$-1_{q<Q} H_\xi^b \left(\frac{\theta(t, q) - \theta(t, q + \Delta)}{\Delta} \right) - 1_{q>-Q} H_\xi^a \left(\frac{\theta(t, q) - \theta(t, q - \Delta)}{\Delta} \right),$$

for $q \in \mathcal{Q}$, and $t \in [0, T]$, with the terminal condition

$$\theta(T, q) = -\ell(|q|). \tag{3.10}$$

Equation (3.4) corresponds to Equation (3.9) for $\xi = \gamma$, while Equation (3.8) corresponds to Equation (3.9) for $\xi = 0$.

3.2. Existence and uniqueness of a solution θ

In the following paragraphs, we prove, for all $\xi \geq 0$, that there exists a unique solution θ to Equation (3.9) with terminal condition (3.10).

Let us start with a lemma on H_{ξ}^{b} and H_{ξ}^{a}.

Lemma 3.1. $\forall \xi \geq 0$, H_{ξ}^{b} and H_{ξ}^{a} are two decreasing functions of class C^2.

The supremum in the definition of $H_{\xi}^{b}(p)$ is attained at a unique $\tilde{\delta}_{\xi}^{b*}(p)$ characterized by

$$p = \tilde{\delta}_{\xi}^{b*}(p) - \frac{1}{\xi\Delta}\log\left(1 - \xi\Delta\frac{\Lambda^{b}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b'}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}\right), \quad \text{if } \xi > 0,$$

and

$$p = \tilde{\delta}_{\xi}^{b*}(p) + \frac{\Lambda^{b}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b'}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}, \quad \text{if } \xi = 0,$$

or equivalently by

$$\tilde{\delta}_{\xi}^{b*}(p) = \Lambda^{b-1}\left(\xi H_{\xi}^{b}(p) - \frac{H_{\xi}^{b'}(p)}{\Delta}\right). \tag{3.11}$$

Similarly, the supremum in the definition of $H_{\xi}^{a}(p)$ is attained at a unique $\tilde{\delta}_{\xi}^{a*}(p)$ characterized by

$$p = \tilde{\delta}_{\xi}^{a*}(p) - \frac{1}{\xi\Delta}\log\left(1 - \xi\Delta\frac{\Lambda^{a}\left(\tilde{\delta}_{\xi}^{a*}(p)\right)}{\Lambda^{a'}\left(\tilde{\delta}_{\xi}^{a*}(p)\right)}\right), \quad \text{if } \xi > 0,$$

and

$$p = \tilde{\delta}_{\xi}^{a*}(p) + \frac{\Lambda^{a}\left(\tilde{\delta}_{\xi}^{a*}(p)\right)}{\Lambda^{a'}\left(\tilde{\delta}_{\xi}^{a*}(p)\right)}, \quad \text{if } \xi = 0,$$

or equivalently by

$$\tilde{\delta}_{\xi}^{a*}(p) = \Lambda^{a-1}\left(\xi H_{\xi}^{a}(p) - \frac{H_{\xi}^{a\prime}(p)}{\Delta}\right). \tag{3.12}$$

Furthermore, the functions $p \mapsto \tilde{\delta}_{\xi}^{b*}(p)$ and $p \mapsto \tilde{\delta}_{\xi}^{a*}(p)$ are C^1 and increasing.

Proof.
We prove the results for the bid side. The proof is similar for the ask side.

Let us start with $\xi > 0$.

$\forall p \in \mathbb{R}$, let us define $g_p : \delta \mapsto \frac{\Lambda^b(\delta)}{\xi}(1 - \exp(-\xi\Delta(\delta - p)))$.

g_p is a function of class C^1, positive for $\delta \in (p, +\infty)$ and non-positive otherwise. Because $g_p(p) = 0$ and $\lim_{\delta \to +\infty} g_p(\delta) = 0$, the supremum of g_p is attained at, at least, one point $\tilde{\delta}_{\xi}^{b*}(p) \in (p, +\infty)$. The first-order condition characterizing the suprema of g_p is

$$\frac{\Lambda^{b\prime}(\delta)}{\xi}(1 - \exp(-\xi\Delta(\delta - p))) + \Delta\Lambda^b(\delta)\exp(-\xi\Delta(\delta - p)) = 0.$$

By rearranging the terms, we obtain

$$p = \delta - \frac{1}{\xi\Delta}\log\left(1 - \xi\Delta\frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}\right).$$

Because $\Lambda^b(\delta)\Lambda^{b\prime\prime}(\delta) < 2\left(\Lambda^{b\prime}(\delta)\right)^2$, the function

$$j : \delta \mapsto \delta - \frac{1}{\xi\Delta}\log\left(1 - \xi\Delta\frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}\right)$$

is increasing[12] and there is therefore a unique maximizer $\tilde{\delta}_{\xi}^{b*}(p)$ of g_p, characterized by

$$p = \tilde{\delta}_{\xi}^{b*}(p) - \frac{1}{\xi\Delta}\log\left(1 - \xi\Delta\frac{\Lambda^b\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b\prime}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}\right).$$

Moreover, by the implicit function theorem, $p \mapsto \tilde{\delta}_{\xi}^{b*}(p)$ is a function of class C^1 which verifies

$$\tilde{\delta}_{\xi}^{b*\prime}(p) = \frac{1}{j'\left(\tilde{\delta}_{\xi}^{b*}(p)\right)} = \frac{1 - \xi\Delta\dfrac{\Lambda^b\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b\prime}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}}{2 - \dfrac{\Lambda^b\left(\tilde{\delta}_{\xi}^{b*}(p)\right)\Lambda^{b\prime\prime}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b\prime}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)^2} - \xi\Delta\dfrac{\Lambda^b\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}{\Lambda^{b\prime}\left(\tilde{\delta}_{\xi}^{b*}(p)\right)}} > 0.$$

In particular, $p \mapsto \tilde{\delta}_\xi^{b*}(p)$ is increasing.

Moreover, the function H_ξ^b is of class C^2, with

$$H_\xi^{b'}(p) = -\Lambda^b(\tilde{\delta}_\xi^{b*}(p))\Delta \exp\left(-\xi\Delta\left(\tilde{\delta}_\xi^{b*}(p) - p\right)\right)$$

and

$$H_\xi^{b''}(p) = \left(\tilde{\delta}_\xi^{b*'}(p)\Lambda^{b'}(\tilde{\delta}_\xi^{b*}(p)) + \xi\Delta\Lambda^b(\tilde{\delta}_\xi^{b*}(p))\left(\tilde{\delta}_\xi^{b*'}(p) - 1\right)\right)\Delta \exp\left(-\xi\Delta\left(\tilde{\delta}_\xi^{b*}(p) - p\right)\right).$$

In particular, H_ξ^b is decreasing.

We also see, by using the definition of H_ξ^b, that

$$\tilde{\delta}_\xi^{b*}(p) = \Lambda^{b-1}\left(\xi H_\xi^b(p) - \frac{H_\xi^{b'}(p)}{\Delta}\right).$$

In the $\xi = 0$ case, we define $\forall p \in \mathbb{R},\ h_p : \delta \mapsto \Delta\Lambda^b(\delta)(\delta - p)$.

h_p is a function of class C^1, positive for $\delta \in (p, +\infty)$ and non-positive otherwise. By using the same reasoning as in footnote 11, we see that there is a unique maximizer $\tilde{\delta}_0^{b*}(p)$ of h_p, characterized by

$$p = \tilde{\delta}_0^{b*}(p) + \frac{\Lambda^b\left(\tilde{\delta}_0^{b*}(p)\right)}{\Lambda^{b'}\left(\tilde{\delta}_0^{b*}(p)\right)}.$$

As earlier, by the implicit function theorem, $p \mapsto \delta_0^{b*}(p)$ is a function of class C^1 which verifies

$$\tilde{\delta}_0^{b*'}(p) = \frac{1}{2 - \dfrac{\Lambda^b(\tilde{\delta}_0^{b*}(p))\Lambda^{b''}(\tilde{\delta}_0^{b*}(p))}{\Lambda^{b'}(\tilde{\delta}_0^{b*}(p))^2}} > 0.$$

In particular, $p \mapsto \tilde{\delta}_0^{b*}(p)$ is increasing.

Moreover, the function H_0^b is of class C^2, with

$$H_0^{b'}(p) = -\Delta\Lambda^b(\tilde{\delta}_0^{b*}(p))$$

and

$$H_0^{b''}(p) = -\Delta\tilde{\delta}_0^{b*'}(p)\Lambda^{b'}(\tilde{\delta}_0^{b*}(p)).$$

In particular, H_0^b is decreasing and we have

$$\tilde{\delta}_0^{b*}(p) = \Lambda^{b-1}\left(-\frac{H_0^{b'}(p)}{\Delta}\right).$$

This proves the lemma.   □

We now prove a comparison principle for Equation (3.9) which gives a priori bounds that will enable us to prove the existence of a solution to Equation (3.9) with terminal condition (3.10).

Lemma 3.2. *Let* $\tau \in [0, T)$.

Let $\underline{\theta} : [\tau, T] \times \mathcal{Q} \to \mathbb{R}$ *be a* C^1 *function with respect to time satisfying the subsolution property, that is,*

$$\forall q \in \mathcal{Q}, \quad \underline{\theta}(T, q) \leq -\ell(|q|)$$

and $\forall (t, q) \in [\tau, T) \times \mathcal{Q}$,

$$-\partial_t \underline{\theta}(t, q) + \frac{1}{2} \gamma \sigma^2 q^2 - 1_{q < Q} H_\xi^b \left(\frac{\underline{\theta}(t, q) - \underline{\theta}(t, q + \Delta)}{\Delta} \right)$$
$$- 1_{q > -Q} H_\xi^a \left(\frac{\underline{\theta}(t, q) - \underline{\theta}(t, q - \Delta)}{\Delta} \right) \leq 0.$$

Let $\bar{\theta} : [\tau, T] \times \mathcal{Q} \to \mathbb{R}$ *be a* C^1 *function with respect to time satisfying the supersolution property, that is,*

$$\forall q \in \mathcal{Q}, \quad \bar{\theta}(T, q) \geq -\ell(|q|)$$

and $\forall (t, q) \in [\tau, T) \times \mathcal{Q}$,

$$-\partial_t \bar{\theta}(t, q) + \frac{1}{2} \gamma \sigma^2 q^2 - 1_{q < Q} H_\xi^b \left(\frac{\bar{\theta}(t, q) - \bar{\theta}(t, q + \Delta)}{\Delta} \right)$$
$$- 1_{q > -Q} H_\xi^a \left(\frac{\bar{\theta}(t, q) - \bar{\theta}(t, q - \Delta)}{\Delta} \right) \geq 0.$$

Then
$$\bar{\theta} \geq \underline{\theta}.$$

Proof.

Let $\varepsilon > 0$.

Let us consider a couple $(t_\varepsilon^*, q_\varepsilon^*)$ such that
$$\underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \varepsilon(T - t_\varepsilon^*) = \sup_{(t,q) \in [\tau, T] \times \mathcal{Q}} \underline{\theta}(t, q) - \bar{\theta}(t, q) - \varepsilon(T - t).$$

If $t_\varepsilon^* \neq T$, then

$$\partial_t \underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \partial_t \bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) + \varepsilon \leq 0.$$

Now, by using the definition of the functions $\underline{\theta}$ and $\bar{\theta}$, the above inequality gives

$$1_{q_\varepsilon^* < Q} H_\xi^b \left(\frac{\bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \bar{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) + 1_{q_\varepsilon^* > -Q} H_\xi^a \left(\frac{\bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \bar{\theta}(t_\varepsilon^*, q_\varepsilon^* - \Delta)}{\Delta} \right)$$

$$-1_{q_\varepsilon^* < Q} H_\xi^b \left(\frac{\underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \underline{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) - 1_{q_\varepsilon^* > -Q} H_\xi^a \left(\frac{\underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \underline{\theta}(t_\varepsilon^*, q_\varepsilon^* - \Delta)}{\Delta} \right) \leq -\varepsilon.$$

But, by definition of $(t_\varepsilon^*, q_\varepsilon^*)$, since H_ξ^b and H_ξ^a are decreasing functions, we have

$$1_{q_\varepsilon^* < Q} \left(H_\xi^b \left(\frac{\bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \bar{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) - H_\xi^b \left(\frac{\underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \underline{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) \right) \geq 0$$

and

$$1_{q_\varepsilon^* > -Q} \left(H_\xi^a \left(\frac{\bar{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \bar{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) - H_\xi^a \left(\frac{\underline{\theta}(t_\varepsilon^*, q_\varepsilon^*) - \underline{\theta}(t_\varepsilon^*, q_\varepsilon^* + \Delta)}{\Delta} \right) \right) \geq 0.$$

This leads to $0 \leq -\varepsilon$. By contradiction, we must have $t_\varepsilon^* = T$.

Therefore,

$$\sup_{(t,q) \in [\tau, T] \times \mathcal{Q}} \underline{\theta}(t, q) - \bar{\theta}(t, q) - \varepsilon(T - t) = \underline{\theta}(T, q_\varepsilon^*) - \bar{\theta}(T, q_\varepsilon^*) \leq 0.$$

As a consequence, $\forall (t, q) \in [\tau, T) \times \mathcal{Q}$,

$$\underline{\theta}(t, q) - \bar{\theta}(t, q) \leq \varepsilon T.$$

By sending ε to 0, we obtain $\underline{\theta} \leq \bar{\theta}$. □

Let us now come to the existence and uniqueness of a solution to Equation (3.9) with terminal condition (3.10).

Theorem 3.1. *There exists a unique function $\theta : [0, T] \times \mathcal{Q} \to \mathbb{R}$, C^1 in time, solution of Equation (3.9) with terminal condition (3.10).*

Proof.
Equation (3.9) with terminal condition (3.10) can be regarded as a backward Cauchy problem. Since H_ξ^b and H_ξ^a are functions of class C^1, by Cauchy–Lipschitz, there exists $\tau \in [0, T)$ and a function $\theta : (\tau, T] \times \mathcal{Q} \to \mathbb{R}$, C^1 in time, solution of Equation (3.9) on $(\tau, T]$ with terminal condition (3.10).

It is straightforward to verify that $\forall q \in \mathcal{Q}, t \in (\tau, T] \mapsto \theta(t, q) + \frac{1}{2} \gamma \sigma^2 q^2 (T - t)$ is a decreasing function. Therefore, the only reason why there would not be a global solution on $[0, T]$ is because $\sup_{q \in \mathcal{Q}} \theta(t, q)$ blows up at $\tau > 0$. However, by using Lemma 3.2, we know that

$$\bar{\theta}(t, q) = (H_\xi^b(0) + H_\xi^a(0))(T - t)$$

defines a supersolution of Equation (3.9) with terminal condition (3.10), and therefore that $\sup_{q \in \mathcal{Q}} \theta(t, q) \leq (H_\xi^b(0) + H_\xi^a(0))(T - t)$ cannot blow up in finite time.

The conclusion is that θ is in fact defined on $[0, T] \times \mathcal{Q}$. Uniqueness comes then from the Cauchy–Lipschitz theorem. □

The existence (and uniqueness) of a function θ solution of Equation (3.9) with terminal condition (3.10) enables us to find a solution to the HJB equation associated with Model A or Model B. We will use a verification argument in the next subsection in order to prove that the solution to the HJB equation we obtain by this way is indeed the value function of the stochastic optimal control problem under consideration. However, before that, a remark needs to be made on θ and on Equation (3.9) in the specific case – often (not to say almost always) used in the academic literature – of exponential intensities.

If we have $\Lambda^b(\delta) = \Lambda^a(\delta) = Ae^{-k\delta} =: \Lambda(\delta)$, then we obtain (by straightforward computations)

$$H_\xi(p) := H_\xi^b(p) = H_\xi^a(p) = \frac{A\Delta}{k} C_\xi \exp(-kp),$$

where

$$C_\xi = \begin{cases} \left(1 + \frac{\xi\Delta}{k}\right)^{-\frac{k}{\xi\Delta} - 1} & \text{if } \xi > 0 \\ e^{-1} & \text{if } \xi = 0. \end{cases}$$

By using Equation (3.9), the function

$$v : (t, q) \in [0, T] \times \mathcal{Q} \mapsto v_q(t) = \exp\left(\frac{k}{\Delta} \theta(t, q)\right)$$

is solution of the linear system of ODEs

$$\forall q \in \mathcal{Q}, \forall t \in [0, T], -\partial_t v_q(t) + \frac{1}{2\Delta} k\gamma\sigma^2 q^2 v_q(t) \tag{3.13}$$
$$-AC_\xi \left(1_{q < Q} v_{q+\Delta}(t) + 1_{q > -Q} v_{q-\Delta}(t)\right) = 0,$$

with terminal condition $\forall q \in \mathcal{Q}, v_q(T) = \exp\left(-\frac{k}{\Delta} \ell(|q|)\right)$.

3.3. *Verification argument*

We are now ready to solve the stochastic optimal control problems associated with Model A and Model B (the proofs follow the classical steps of the verification approach and are left to the reader).

Theorem 3.2 *Let us consider the solution θ of Equation (3.9) with terminal condition (3.10) for $\xi = \gamma$.*

Then, $u : (t, x, q, S) \mapsto - \exp(-\gamma(x + qS + \theta(t, q)))$ *defines a solution to Equation (3.1) with terminal condition (3.2), and*

$$u(t, x, q, S) = \sup_{(\delta_s^b)_{s \geq t}, (\delta_s^a)_{s \geq t} \in \mathcal{A}(t)} \mathbb{E}\left[- \exp\left(-\gamma\left(X_T^{t,x,\delta^b,\delta^a} + q_T^{t,q,\delta^b,\delta^a} S_T^{t,S} - \ell\left(\left| q_T^{t,q,\delta^b,\delta^a} \right| \right) \right) \right) \right],$$

where $\mathcal{A}(t)$ *is the set of predictable processes on* $[t, T]$, *bounded from below and where*
$$dS_s^{t,S} = \sigma dW_s, \qquad S_t^{t,S} = S,$$

$$dX_s^{t,x,\delta^b,\delta^a} = (S_s + \delta_s^a)\Delta dN_s^a - (S_s - \delta_s^b)\Delta dN_s^b, \qquad X_t^{t,x,\delta^b,\delta^a} = x,$$

$$dq_s^{t,q,\delta^b,\delta^a} = \Delta dN_s^b - \Delta dN_s^a, \qquad q_t^{t,q,\delta^b,\delta^a} = q,$$

where the point processes N^b *and* N^a *have stochastic intensity* $(\lambda_s^b)_s$ *and* $(\lambda_s^a)_s$ *given by* $\lambda_s^b = \Lambda^b(\delta_s^b)1_{q_{s-} < Q}$ *and* $\lambda_s^a = \Lambda^a(\delta_s^a)1_{q_{s-} > -Q}$.

The optimal bid and ask quotes $S_t^b = S_t - \delta_t^{b*}$ *(for* $q_{t-} < Q$) *and* $S_t^a = S_t + \delta_t^{a*}$ *(for* $q_{t-} > -Q$) *are characterized by*

$$\delta_t^{b*} = \tilde{\delta}_\gamma^{b*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} + \Delta)}{\Delta} \right) \text{ and}$$

$$\delta_t^{a*} = \tilde{\delta}_\gamma^{a*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} - \Delta)}{\Delta} \right), \tag{3.14}$$

where the functions $\tilde{\delta}_\gamma^{b*}(\cdot)$ *and* $\tilde{\delta}_\gamma^{a*}(\cdot)$ *are defined in Equations (3.11) and (3.12).*

Theorem 3.3. *Let us consider the solution* θ *of Equation (3.9) with terminal condition (3.10) for* $\xi = 0$.

Then, $u : (t, x, q, S) \mapsto x + qS + \theta(t, q)$ *defines a solution to Equation (3.5) with terminal condition (10), and*

$$u(t, x, q, S) =$$
$$\sup_{(\delta_s^b)_{s \geq t}, (\delta_s^a)_{s \geq t} \in \mathcal{A}(t)} \mathbb{E}\left[X_T^{t,x,\delta^b,\delta^a} + q_T^{t,q,\delta^b,\delta^a} S_T^{t,S} - \ell\left(\left| q_T^{t,q,\delta^b,\delta^a} \right| \right) - \frac{1}{2}\gamma\sigma^2 \int_t^T q_s^{t,q,\delta^b,\delta^a 2} ds \right],$$

where $\mathcal{A}(t)$ *is the set of predictable processes on* $[t, T]$, *bounded from below and where*

$$dS_s^{t,S} = \sigma dW_s, \qquad S_t^{t,S} = S,$$

$$dX_s^{t,x,\delta^b,\delta^a} = (S_s + \delta_s^a)\Delta dN_s^a - (S_s - \delta_s^b)\Delta dN_s^b, \qquad X_t^{t,x,\delta^b,\delta^a} = x,$$

$$dq_s^{t,q,\delta^b,\delta^a} = \Delta dN_s^b - \Delta dN_s^a, \qquad q_t^{t,q,\delta^b,\delta^a} = q,$$

where the point processes N^b and N^a have stochastic intensity $(\lambda_s^b)_s$ and $(\lambda_s^a)_s$ given by $\lambda_s^b = \Lambda^b(\delta_s^b)1_{q_{s-} < Q}$ and $\lambda_s^a = \Lambda^a(\delta_s^a)1_{q_{s-} > -Q}$.

The optimal bid and ask quotes $S_t^b = S_t - \delta_t^{b*}$ (for $q_{t-}<Q$) and $S_t^a = S_t + \delta_t^{a*}$ (for $q_{t-}> -Q$) are given by

$$\delta_t^{b*} = \tilde{\delta}_0^{b*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} + \Delta)}{\Delta}\right) \text{ and}$$

$$\delta_t^{a*} = \tilde{\delta}_0^{a*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} - \Delta)}{\Delta}\right),$$

(3.15)

where the functions $\tilde{\delta}_0^{b*}(\cdot)$ and $\tilde{\delta}_0^{a*}(\cdot)$ are defined in Equations (3.11) and (3.12).

3.4. Comments on the results

In both Model A and Model B, the dynamic optimization problem faced by the market maker was initially characterized by a HJB equation with four variables: the time t, and three state variables (the cash x, the inventory q and the reference price S). Computing a numerical approximation for the solution of a four-dimensional HJB equation such as Equation (3.1) or Equation (3.5) is always time-consuming. Therefore, the results obtained in Theorem 3.2 and Theorem 3.3 are very useful: they state that the optimal quotes of a market maker in both Model A and Model B can in fact be computed by solving a tridiagonal system of nonlinear ODEs. This corresponds to a reduction of the dimensionality of the problem from 4 to 2. Furthermore, the systems of nonlinear ODEs are similar for Model A and Model B: they correspond to Equation (3.9) – with terminal condition (3.10) – with $\xi = \gamma$ for Model A and with $\xi = 0$ for Model B.

The objective functions of Model A and Model B lead to similar equations, but it is interesting to understand the differences between the two modelling approaches. In fact, the penalization term

$$\frac{1}{2}\gamma\sigma^2 \int_0^T q_t^2 dt$$

in Model B leads to the term $\frac{1}{2}\gamma\sigma^2 q^2$ in the ODEs characterizing θ (when $\xi = 0$), and this term arises also in the ODE associated with Model A (when $\xi = \gamma$) because of the market maker's aversion to price risk. However, in Model A, the market maker is not only averse to price risk but also to the risk of not finding a counterparty to trade with – this is what we call non-execution risk. There is indeed a source of risk coming from the process $(W_t)_t$, and another source of risk coming from the processes $(N_t^b)_t$ and $(N_t^a)_t$, and risk aversion in Model A applies to both kinds of risk. In other words, things work as if the market maker of Model A was risk averse to both kinds of risk, while the market maker of Model B is only averse to the risk associated with price changes. In particular, the parameter ξ can be regarded as some form of risk aversion parameter applying to non-execution risk only: it is equal to γ in the case of Model A, and equal to 0 in the case of Model B.

4. Closed-form and almost-closed-form approximations

In Guéant, Lehalle, and Fernandez-Tapia (2013), the authors show in the specific case where $\Lambda^b(\delta) = \Lambda^a(\delta) = Ae^{-k\delta} =: \Lambda(\delta)$ that there is an asymptotic regime far from T for the optimal quotes in Model A.[13] In other words, far from the terminal time T, the optimal quotes in Guéant, Lehalle, and Fernandez-Tapia (2013) are well approximated by functions that only depend on the inventory q – and not on the time variable t. In practice, in markets (such as most dealer-driven OTC markets) for which there is no natural terminal time T, this result is not surprising – even, somehow, reassuring – and only the asymptotic formula should be used. Furthermore, Guéant, Lehalle, and Fernandez-Tapia (2013) proposed closed-form approximations for the asymptotic values of the optimal quotes. In this section, we propose new approximation formulas which generalize those obtained in Guéant, Lehalle, and Fernandez-Tapia (2013) to a more general set of intensity functions and to both Model A and Model B (only Model A was considered in Guéant, Lehalle, and Fernandez-Tapia (2013)). These more general approximations are based on heuristic arguments, and we will see in the numerical experiments of Section 6 when they are (or are not) satisfactory.

4.1. Approximation with an elliptic partial differential equation

To compute the optimal quotes given in Equations (3.14) and (3.15), the first step consists in computing the function θ solution of the system of ODEs (3.9), with terminal condition (3.10). In order to approximate the optimal quotes, we first approximate, therefore, the function θ.

To carry out our reasoning, we suppose that the intensity functions Λ^b and Λ^a are identical (equal to Λ) and that $H_\xi := H_\xi^b = H_\xi^a$ verifies $H_\xi''(0) > 0$.[14]

Our heuristic reasoning consists in replacing the function $\theta : [0, T] \times \mathcal{Q} \to \mathbb{R}$ by a function $\tilde{\theta} : [0, T] \times \mathbb{R} \to \mathbb{R}$ and to replace the system of ODEs (3.9) characterizing θ, i.e.,

$$0 = -\partial_t \theta(t, q) + \frac{1}{2}\gamma\sigma^2 q^2 - 1_{q < Q} H_\xi^b\left(\frac{\theta(t, q) - \theta(t, q + \Delta)}{\Delta}\right)$$

$$- 1_{q > -Q} H_\xi^a\left(\frac{\theta(t, q) - \theta(t, q - \Delta)}{\Delta}\right)$$

by the PDE

$$0 = -\partial_t \tilde{\theta}(t, q) + \frac{1}{2}\gamma\sigma^2 q^2 - 2H_\xi(0)$$

$$- H_\xi''(0)(\partial_q \tilde{\theta}(t, q))^2 + \Delta H_\xi'(0)\partial_{qq}^2 \tilde{\theta}(t, q). \tag{4.1}$$

This PDE comes from an expansion in ϵ of the expression[15]

$$0 = -\partial_t \tilde{\theta}(t, q) + \frac{1}{2}\gamma\sigma^2 q^2$$

$$- H_\xi\left(\frac{\tilde{\theta}(t, q) - \tilde{\theta}(t, q + \epsilon\Delta)}{\Delta}\right) - H_\xi\left(\frac{\tilde{\theta}(t, q) - \tilde{\theta}(t, q - \epsilon\Delta)}{\Delta}\right),$$

applied to $\epsilon = 1$.[16]

We have indeed

$$H_\xi \left(\frac{\tilde{\theta}(t,q) - \tilde{\theta}(t,q+\epsilon\Delta)}{\Delta} \right) + H_\xi \left(\frac{\tilde{\theta}(t,q) - \tilde{\theta}(t,q-\epsilon\Delta)}{\Delta} \right)$$

$$= H_\xi \left(-\epsilon\partial_q\tilde{\theta}(t,q) - \frac{1}{2}\epsilon^2\Delta\partial^2_{qq}\tilde{\theta}(t,q) + o(\epsilon^2) \right)$$

$$+ H_\xi \left(\epsilon\partial_q\tilde{\theta}(t,q) - \frac{1}{2}\epsilon^2\Delta\partial^2_{qq}\tilde{\theta}(t,q) + o(\epsilon^2) \right)$$

$$= 2H_\xi(0) - \epsilon^2\Delta H'_\xi(0)\partial^2_{qq}\tilde{\theta}(t,q) + \epsilon^2 H''_\xi(0)\left(\partial_q\tilde{\theta}(t,q) \right)^2 + o(\epsilon^2).$$

By considering

$$\tilde{\nu}(t,q) = \exp \left(-\frac{H''_\xi(0)}{\Delta H''_\xi(0)}\tilde{\theta}(t,q) \right),$$

the nonlinear PDE (4.1) becomes the linear PDE[17]

$$0 = \partial_t\tilde{\nu}(t,q) - \frac{H''_\xi(0)}{\Delta H'_\xi(0)}\left(2H_\xi(0) - \frac{1}{2}\gamma\sigma^2 q^2 \right)\tilde{\nu}(t,q) - \Delta H'_\xi(0)\partial^2_{qq}\tilde{\nu}(t,q), \quad (4.2)$$

and the terminal condition relevant with our problem is

$$\tilde{\nu}(T,q) = \exp \left(\frac{H''_\xi(0)}{\Delta H'_\xi(0)}\ell(|q|) \right).$$

Equation (4.1) is a linear PDE and it can be studied using basic tools of spectral theory. Our goal is to study the asymptotic behaviour of $\tilde{\nu}(t,q)$ when T tends to infinity and to use the formulas obtained in this asymptotic regime in order to approximate successively $\tilde{\nu}$, $\tilde{\theta}$, θ, and ultimately the optimal quotes $(\delta^{b*}_t)_t$ and $(\delta^{a*}_t)_t$.

4.2. Generalization of the Guéant–Lehalle–Fernandez-Tapia's formulas

By classical spectral theory,[18] we know that

$$\tilde{\nu}(t,q) \sim_{T\to+\infty} (\tilde{\nu}(T,\cdot), \tilde{f}^0)\tilde{f}^0(q)\exp(\nu(T-t)),$$

where ν and $\tilde{f}^0$ are, respectively, the minimum and a minimizer of the functional

$$\tilde{f} \in \{\tilde{g} \in H^1(\mathbb{R}), \| \tilde{g} \|_{L^2(\mathbb{R})} = 1\} \mapsto \int_{-\infty}^{\infty} \left(\alpha x^2\tilde{f}(x)^2 + \eta\tilde{f}'(x)^2 \right)dx,$$

with

$$\alpha = -\frac{1}{2}\frac{H_\xi''(0)}{\Delta H_\xi'(0)}\gamma\sigma^2 \text{ and } \eta = -\Delta H_\xi'(0),$$

and where $(\cdot, \cdot)$ designates the scalar product in $L^2(\mathbb{R})$.

In particular,

$$\tilde{f}^0(q) \propto \exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right).$$

From

$$\tilde{v}(t,q) \sim_{T\to+\infty} C\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)\exp(\nu(T-t)),$$

where C is a constant, independent of (t, q), we deduce:

$$\tilde{\theta}(t,q) + \frac{\Delta H_\xi'(0)}{H_\xi''(0)}\nu(T-t) \to_{T\to+\infty} -\frac{\Delta H_\xi'(0)}{H_\xi''(0)}\left(\log(C) - \frac{1}{2}\sqrt{\frac{\alpha}{\eta}}\,q^2\right),$$

that is,

$$\tilde{\theta}(t,q) + \frac{\Delta H_\xi'(0)}{H_\xi''(0)}\nu(T-t) \to_{T\to+\infty} -\frac{\Delta H_\xi'(0)}{H_\xi''(0)}\left(\log(C) - \frac{1}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}q^2\right).$$

As a consequence, we consider the approximations

$$\frac{\theta(t,q) - \theta(t,q+\Delta)}{\Delta} \simeq \frac{2q+\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}$$

and

$$\frac{\theta(t,q) - \theta(t,q-\Delta)}{\Delta} \simeq -\frac{2q-\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}.$$

These approximations are independent of t and of the final penalty function ℓ. They can be plugged into Equations (3.14) and (3.15) to obtain the general approximation formulas

$$\delta_t^{b*} \simeq \delta_{\text{approx}}^{b*}(q_{t-}) := \tilde{\delta}_\xi^*\left(\frac{2q_{t-}+\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}\right) \qquad (4.3)$$

and

$$\delta_t^{a*} \simeq \delta_{\text{approx}}^{a*}(q_{t-}) := \tilde{\delta}_\xi^*\left(-\frac{2q_{t-}-\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}\right), \qquad (4.4)$$

where

$$\tilde{\delta}_{\xi}^{*}(p) = \Lambda^{-1}\left(\xi H_{\xi}(p) - \frac{H_{\xi}'(p)}{\Delta}\right). \tag{4.5}$$

In particular, if $\Lambda(\delta) = Ae^{-k\delta}$, then

$$\tilde{\delta}_{\xi}^{*}(p) = \begin{cases} p + \frac{1}{\xi\Delta}\log\left(1 + \frac{\xi\Delta}{k}\right) & \text{if } \xi > 0 \\ p + \frac{1}{k} & \text{if } \xi = 0, \end{cases}$$

and we obtain

$$\delta_{\text{approx}}^{b*}(q) = \begin{cases} \dfrac{1}{\xi\Delta}\log\left(1 + \dfrac{\xi\Delta}{k}\right) + \dfrac{2q+\Delta}{2}\sqrt{\dfrac{\gamma\sigma^2}{2A\Delta k}\left(1 + \dfrac{\xi\Delta}{k}\right)^{\frac{k}{\xi\Delta}+1}} & \text{if } \xi > 0 \\ \dfrac{1}{k} + \dfrac{2q+\Delta}{2}\sqrt{\dfrac{\gamma\sigma^2 e}{2A\Delta k}} & \text{if } \xi = 0. \end{cases} \tag{4.6}$$

and

$$\delta_{\text{approx}}^{a*}(q) = \begin{cases} \dfrac{1}{\xi\Delta}\log\left(1 + \dfrac{\xi\Delta}{k}\right) - \dfrac{2q-\Delta}{2}\sqrt{\dfrac{\gamma\sigma^2}{2A\Delta k}\left(1 + \dfrac{\xi\Delta}{k}\right)^{\frac{k}{\xi\Delta}+1}} & \text{if } \xi > 0 \\ \dfrac{1}{k} - \dfrac{2q-\Delta}{2}\sqrt{\dfrac{\gamma\sigma^2 e}{2A\Delta k}} & \text{if } \xi = 0. \end{cases} \tag{4.7}$$

In particular, we recover, in the specific case where $\Delta = 1$ and $\xi = \gamma$, the Guéant–Lehalle–Fernandez-Tapia's formula of (2013) and Guéant (2016) often used in the industry.

4.3. Comments on the approximations

The approximations obtained earlier deserve a few comments. First, in the general case (i.e., even when the intensity function Λ is not exponential), the approximations are almost in closed form, in the sense that they are only functions of the parameters and of transforms of Λ. In practice, one simply needs to compute Λ^{-1}, H_ξ, H_ξ' and H_ξ'', in order to compute the approximations (4.3) and (4.4). Second, the above approximations enable to better understand the optimal strategy of a market maker and the role played by the different parameters. In particular, they enable to better understand the different types of risk faced by a market maker.

By using Equations (4.3) and (4.4), we see that

$$\frac{d}{dq}\delta_{\text{approx}}^{b*}(q) = \sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}\,\tilde{\delta}_{\xi}^{*'}\left(\frac{2q+\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H_\xi''(0)}}\right) > 0$$

and

$$\frac{d}{dq} \delta^{a*}_{\text{approx}}(q) = -\sqrt{\frac{\gamma\sigma^2}{2H''_\xi(0)}} \tilde{\delta}^{*\prime}_\xi \left(-\frac{2q-\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H''_\xi(0)}} \right) < 0.$$

This means that a market maker proposes lower prices at the bid and at the ask when his inventory increases, and conversely, higher prices at the bid and at the ask when his inventory decreases. In particular, a market maker with a positive or negative inventory always skews his bid and ask prices in order to increase his chance to go back to a flat position.

In the particular case of exponential intensities, it is interesting to notice that the approximation of the bid–ask spread is independent of q, and the skew is linear in q:

$$\delta^{b*}_{\text{approx}}(q) + \delta^{a*}_{\text{approx}}(q) = \begin{cases} \frac{2}{\xi\Delta}\log\left(1 + \frac{\xi\Delta}{k}\right) + \Delta\sqrt{\frac{\gamma\sigma^2}{2A\Delta k}\left(1 + \frac{\xi\Delta}{k}\right)^{\frac{k}{\xi\Delta}+1}} & \text{if } \xi > 0 \\ \frac{2}{k} + \Delta\sqrt{\frac{\gamma\sigma^2 e}{2A\Delta k}} & \text{if } \xi = 0, \end{cases}$$

(4.8)

$$\delta^{b*}_{\text{approx}}(q) - \delta^{a*}_{\text{approx}}(q) = \begin{cases} 2q\sqrt{\frac{\gamma\sigma^2}{2A\Delta k}\left(1 + \frac{\xi\Delta}{k}\right)^{\frac{k}{\xi\Delta}+1}} & \text{if } \xi > 0 \\ 2q\sqrt{\frac{\gamma\sigma^2 e}{2A\Delta k}} & \text{if } \xi = 0. \end{cases}$$

(4.9)

As far as volatility is concerned, we have

$$\frac{d}{d\sigma} \delta^{b*}_{\text{approx}}(q) = \frac{2q+\Delta}{2}\sqrt{\frac{\gamma}{2H''_\xi(0)}}\, \tilde{\delta}^{*\prime}_\xi \left(\frac{2q+\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H''_\xi(0)}} \right)$$

and

$$\frac{d}{d\sigma} \delta^{a*}_{\text{approx}}(q) = -\frac{2q-\Delta}{2}\sqrt{\frac{\gamma}{2H''_\xi(0)}}\, \tilde{\delta}^{*\prime}_\xi \left(-\frac{2q-\Delta}{2}\sqrt{\frac{\gamma\sigma^2}{2H''_\xi(0)}} \right).$$

Therefore we have three cases:

- If $q = 0$, then $\frac{d}{d\sigma}\delta^{b*}_{\text{approx}}(q) = \frac{d}{d\sigma}\delta^{a*}_{\text{approx}}(q) > 0$. In other words, an increase in volatility leads to an increase in the bid–ask spread, symmetric around the reference price (no skew).
- If $q \geq \Delta$, then $\frac{d}{d\sigma}\delta^{b*}_{\text{approx}}(q) > 0$ and $\frac{d}{d\sigma}\delta^{a*}_{\text{approx}}(q) < 0$. In other words, an increase in volatility leads to lower bid and ask prices: it increases the skew in absolute value, *ceteris paribus*.
- If $q \leq -\Delta$, then $\frac{d}{d\sigma}\delta^{b*}_{\text{approx}}(q) < 0$ and $\frac{d}{d\sigma}\delta^{a*}_{\text{approx}}(q) > 0$. In other words, an increase in volatility leads to higher bid and ask prices: it increases the skew in absolute value, *ceteris paribus*.

In the particular case of exponential intensities, it is interesting to notice that the bid–ask spread is approximated by an affine function of σ, and the skew by a linear function of σ (see Equations (4.8) and (4.9)).

As far as liquidity is concerned, if we replace Λ by $\beta\Lambda$, for $\beta > 0$, then we see that H_ξ is replaced by βH_ξ and that $\tilde{\delta}_\xi$ is unchanged (see Equation (4.5)). Therefore, we see from Equations (4.3) and (4.4) that replacing Λ by $\beta\Lambda$ is equivalent to replacing σ^2 by $\frac{\sigma^2}{\beta}$. In other words, an increase in liquidity is equivalent to a decrease in volatility and, conversely, a decrease in liquidity has the same effects as an increase in volatility.

As far as risk aversion is concerned, the differences between Model A and Model B help to clarify the different roles played by γ.

In the case of Model B, where $\xi = 0$, we see from Equations (4.3) and (4.4) that an increase in γ is equivalent to an increase in σ^2. In particular, an increase in γ increases the bid–ask spread and increases the skew in absolute value. This is expected, since γ, in Model B, penalizes positive and negative inventory.

In the case of Model A, the situation is different, but the introduction of the variable ξ helps to understand what is at stake. As already mentioned, everything works as if ξ was a risk aversion parameter for non-execution risk and γ a risk aversion parameter for price risk. To analyse the different effects, we consider the specific case of exponential intensities. We see in Equation (4.8) that the approximation of the bid–ask spread is made of two parts:

(1) $\frac{2}{\xi\Delta}\log\left(1 + \frac{\xi\Delta}{k}\right)$, which is decreasing in ξ. This term is related to the static risk faced by a market maker, associated with transaction uncertainty only. When $\xi = \gamma$ increases, a market maker reduces his bid–ask spread to lower the uncertainty with respect to transactions.

(2) $\Delta\sqrt{\frac{\gamma\sigma^2}{2A\Delta k}\left(1 + \frac{\xi\Delta}{k}\right)^{\frac{k}{\xi\Delta}+1}}$, which is increasing in γ and $\xi = \gamma$. This term, that only appears with volatility, is related to the dynamic risk faced by a market maker. This risk is complex and definitely more subtle than the classical risk that the price moves. In fact, both ξ and γ appear in the formula because the risk faced by a market maker is actually the risk that the price moves adversely without him being able to unwind his position rapidly enough (because of trade uncertainty). The higher the risk aversion to this combination of price risk and non-execution risk, the larger the bid–ask spread, because a market maker wants to avoid holding large inventories (in absolute value).

As far as the skew is concerned, only the second effect matters. This is confirmed by Equation (4.8), and we see that the skew in absolute value is increasing with γ and $\xi = \gamma$.

Comparative statics is always interesting to understand the role played by the different parameters involved in a model. Here, we have carried out comparative statics on almost-closed-form and closed-form approximations and not on the original optimal bid and ask quotes, which can only be computed numerically. We will see in Section 6 the differences between the actual optimal bid and ask quotes and the approximations proposed in this section.[19]

5. Multi-asset market making strategies

In most papers of the academic literature on market making, only single-asset market making is tackled. In practice, however, market makers are often in charge of a book of several assets. An evident case is the one of corporate bonds, since there are usually dozens of bonds issued by the same company, and the same market maker is in charge of all these bonds. As a consequence, optimal quotes for a specific bond should not depend on the market maker's inventory in that bond but instead on the risk profile of the whole bond portfolio with respect to the issuer. In particular, when a market maker has a short inventory in an asset and an almost equivalent long inventory in another asset, highly correlated with the first, there may be no reason for him to skew his bid and ask quotes on these two assets, contrary to what single-asset market making models would suggest. In this section, we generalize our market making model to the multi-asset case. In particular, we obtain closed-form approximations for the optimal quotes of a multi-asset market maker.

5.1. *Modelling framework and notations*

We consider a market maker in charge of d assets. For $i \in \{1, \ldots, d\}$, the reference price of asset i is modelled by a process $(S_t^i)_t$ with the following dynamics

$$dS_t^i = \sigma^i dW_t^i, \quad S_0^i \quad \text{given}, \tag{5.1}$$

where $(W_t^1, \ldots, W_t^d)_t$ is a d-dimensional Brownian motion adapted to the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$, with non-singular correlation matrix. We denote by $\Sigma = (\rho^{i,j} \sigma^i \sigma^j)_{1 \le i,j \le d}$, the variance–covariance matrix associated with the process $(S_t)_t = (S_t^1, \ldots, S_t^d)_t$.

This market maker proposes bid and ask quotes to buy and sell the d assets. These bid and ask quotes are modelled by $2d$ stochastic processes, respectively, denoted by $(S_t^{1,b})_t, \ldots, (S_t^{d,b})_t$ and $(S_t^{1,a})_t, \ldots, (S_t^{d,a})_t$.

As in the single-asset case, we denote by $(N_t^{i,b})_t$ and $(N_t^{i,a})_t$, for each $i \in \{1, \ldots, d\}$, the two point processes modelling the number of transactions at the bid and at the ask, respectively, for asset i. We assume that the asset i is traded Δ^i units by Δ^i units.

The inventory of the market maker, modelled by the d-dimensional process $(q_t)_t = (q_t^1, \ldots, q_t^d)_t$, has therefore the following dynamics:

$$\forall i \in \{1, \ldots, d\}, dq_t^i = \Delta^i dN_t^{i,b} - \Delta^i dN_t^{i,a}, \quad q_0^i \quad \text{given}. \tag{5.2}$$

We assume that the processes $(N_t^{1,b}, \ldots, N_t^{d,b})_t$ and $(N_t^{1,a}, \ldots, N_t^{d,a})_t$ are independent of the Brownian motion $(W_t^1, \ldots, W_t^d)_t$. For $i \in \{1, \ldots, d\}$, we denote by $(\lambda_t^{i,b})_t$ and $(\lambda_t^{i,a})_t$ the intensity processes of $(N_t^{i,b})_t$ and $(N_t^{i,a})_t$, respectively. We assume that $(\lambda_t^{i,b})_t$ and $(\lambda_t^{i,a})_t$ verify

$$\lambda_t^{i,b} = \Lambda^{i,b}(\delta_t^{i,b})1_{q_{t_-}^i < Q^i} \quad \text{and} \quad \lambda_t^{i,a} = \Lambda^{i,a}(\delta_t^{i,a})1_{q_{t_-}^i > -Q^i}, \tag{5.3}$$

where

$$\delta_t^{i,b} = S_t^i - S_t^{i,b} \quad \text{and} \quad \delta_t^{i,a} = S_t^{i,a} - S_t^i,$$

and where $\Lambda^{i,b}$ and $\Lambda^{i,a}$ are two functions satisfying the following hypotheses:

- $\Lambda^{i,b}$ and $\Lambda^{i,a}$ are twice continuously differentiable,
- $\Lambda^{i,b}$ and $\Lambda^{i,a}$ are decreasing, with $\forall \delta \in \mathbb{R}$, $\Lambda^{i,b'}(\delta) < 0$ and $\Lambda^{i,a'}(\delta) < 0$,
- $\lim_{\delta \to +\infty} \Lambda^{i,b}(\delta) = \lim_{\delta \to +\infty} \Lambda^{i,a}(\delta) = 0$,
- $\sup_\delta \frac{\Lambda^{i,b}(\delta)\Lambda^{i,b''}(\delta)}{\left(\Lambda^{i,b'}(\delta)\right)^2} < 2 \quad \text{and} \quad \sup_\delta \frac{\Lambda^{i,a}(\delta)\Lambda^{i,a''}(\delta)}{\left(\Lambda^{i,a'}(\delta)\right)^2} < 2.$

Finally, the process $(X_t)_t$ modelling the market maker's cash account has the dynamics

$$dX_t = \sum_{i=1}^d S_t^{i,a}\Delta^i dN_t^{i,a} - S_t^{i,b}\Delta^i dN_t^{i,b}$$

$$= \sum_{i=1}^d (S_t^i + \delta_t^{i,a})\Delta^i dN_t^{i,a} - (S_t^i - \delta_t^{i,b})\Delta^i dN_t^{i,b}. \tag{5.4}$$

In the d-dimensional generalization of Model A, the problem consists in maximizing

$$\mathbb{E}\left[-\exp\left(-\gamma\left(X_T + \sum_{i=1}^d q_T^i S_T^i - \ell_d(q_T^1, \ldots, q_T^d)\right)\right)\right], \tag{Model A}$$

over $(\delta_t^{1,b}, \ldots, \delta_t^{d,b})_t \in \mathcal{A}^d$ and $(\delta_t^{1,a}, \ldots, \delta_t^{d,a})_t \in \mathcal{A}^d$, where ℓ_d is a penalty function.

In the d-dimensional generalization of Model B, the problem consists instead in maximizing

$$\mathbb{E}\left[X_T + \sum_{i=1}^d q_T^i S_T^i - \ell_d(q_T^1, \ldots, q_T^d) - \frac{1}{2}\gamma\int_0^T \sum_{i=1}^d \sum_{j=1}^d \rho^{i,j}\sigma^i\sigma^j q_t^i q_t^j dt\right], \tag{Model B}$$

over $(\delta_t^{1,b}, \ldots, \delta_t^{d,b})_t \in \mathcal{A}^d$ and $(\delta_t^{1,a}, \ldots, \delta_t^{d,b})_t \in \mathcal{A}^d$.

5.2. *Towards a general system of ODEs*

For solving the two stochastic optimal control problems of Model A and Model B, we use similar changes of variables as in Section 3. In particular, we show that finding the value function (and the optimal bid and ask quotes) in both models boils down to solving a system of ODEs and that, as in the single-asset case, the

equations associated with Model A and Model B are part of the same family of ODEs.

The HJB equation associated with Model A is given by[20]

$$0 = -\partial_t u(t,x,q,S) - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\rho^{i,j}\sigma^i\sigma^j\partial^2_{S^iS^j}u(t,x,q,S) \tag{5.5}$$

$$-\sum_{i=1}^{d}1_{q^i < Q^i}\sup_{\delta^{i,b}}\Lambda^{i,b}(\delta^{i,b})\left[u(t,x-\Delta^iS^i+\Delta^i\delta^{i,b},q+\Delta^ie^i,S)-u(t,x,q,S)\right]$$

$$-\sum_{i=1}^{d}1_{q^i > -Q^i}\sup_{\delta^{i,a}}\Lambda^{i,a}(\delta^{i,a})\left[u(t,x+\Delta^iS^i+\Delta^i\delta^{i,a},q-\Delta^ie^i,S)-u(t,x,q,S)\right],$$

for $\forall i \in \{1,\ldots,d\}, q^i \in \mathcal{Q}^i = \{-Q^i, -Q^i + \Delta^i, \ldots, Q^i - \Delta^i, Q^i\}$, and $(t,S,x) \in [0,T]\times \mathbb{R}^d \times \mathbb{R}$, with the terminal condition

$$u(T,x,q,S) = -\exp\left(-\gamma\left(x+\sum_{i=1}^{d}q^iS^i-\ell_d(q^1,\ldots,q^d)\right)\right). \tag{5.6}$$

If one uses the ansatz

$$u(t,x,q,S) = -\exp\left(-\gamma\left(x+\sum_{i=1}^{d}q^iS^i+\theta(t,q)\right)\right), \tag{5.7}$$

then Equation (5.5) becomes

$$0 = -\partial_t\theta(t,q) + \frac{1}{2}\gamma\sum_{i=1}^{d}\sum_{j=1}^{d}\rho^{i,j}\sigma^i\sigma^jq^iq^j$$

$$-\sum_{i=1}^{d}1_{q^i < Q^i}\sup_{\delta^{i,b}}\frac{\Lambda^{i,b}(\delta^{i,b})}{\gamma}\left(1-\exp\left(-\gamma\left(\Delta^i\delta^{i,b}+\theta(t,q+\Delta^ie^i)-\theta(t,q)\right)\right)\right)$$

$$-\sum_{i=1}^{d}1_{q^i > -Q^i}\sup_{\delta^{i,a}}\frac{\Lambda^{i,a}(\delta^{i,a})}{\gamma}\left(1-\exp\left(-\gamma\left(\Delta^i\delta^{i,a}+\theta(t,q-\Delta^ie^i)-\theta(t,q)\right)\right)\right),$$

$$\tag{5.8}$$

for $\forall i \in \{1,\ldots,d\}, q^i \in \mathcal{Q}^i$, and $t \in [0,T]$, and the terminal condition (5.6) becomes $\theta(T,q) = -\ell_d(q^1,\ldots,q^d)$.

The HJB equation associated with Model B is given by

$$
\begin{aligned}
0 = & -\partial_t u(t, x, q, S) + \frac{1}{2}\gamma \sum_{i=1}^{d}\sum_{j=1}^{d}\rho^{i,j}\sigma^i\sigma^j q^i q^j \\
& -\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\rho^{i,j}\sigma^i\sigma^j \partial^2_{S^i S^j} u(t, x, q, S) \\
& -\sum_{i=1}^{d} 1_{q^i < Q^i} \sup_{\delta^{i,b}} \Lambda^{i,b}(\delta^{i,b})\Big[u(t, x - \Delta^i S^i + \Delta^i \delta^{i,b}, q + \Delta^i e^i, S) - u(t, x, q, S)\Big] \\
& -\sum_{i=1}^{d} 1_{q^i > -Q^i} \sup_{\delta^{i,a}} \Lambda^{i,a}(\delta^{i,a})\Big[u(t, x + \Delta^i S^i + \Delta^i \delta^{i,a}, q - \Delta^i e^i, S) - u(t, x, q, S)\Big],
\end{aligned}
\tag{5.9}
$$

for $\forall i \in \{1, \ldots, d\}, q^i \in \mathcal{Q}^i$ and $(t, S, x) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$, with the terminal condition

$$
u(T, x, q, S) = x + \sum_{i=1}^{d} q^i S^i - \ell_d(q^1, \ldots, q^d).
\tag{5.10}
$$

If one uses the ansatz

$$
u(t, x, q, S) = x + \sum_{i=1}^{d} q^i S^i + \theta(t, q),
\tag{5.11}
$$

then Equation (5.9) becomes

$$
\begin{aligned}
0 = & -\partial_t \theta(t, q) + \frac{1}{2}\gamma \sum_{i=1}^{d}\sum_{j=1}^{d}\rho^{i,j}\sigma^i\sigma^j q^i q^j \\
& -\sum_{i=1}^{d} 1_{q^i < Q^i} \sup_{\delta^{i,b}} \Lambda^{i,b}(\delta^{i,b})\Big(\Delta^i \delta^{i,b} + \theta(t, q + \Delta^i e^i) - \theta(t, q)\Big) \\
& -\sum_{i=1}^{d} 1_{q^i > -Q^i} \sup_{\delta^{i,a}} \Lambda^{i,a}(\delta^{i,a})\big(\Delta^i \delta^{i,a} + \theta(t, q - \Delta^i e^i) - \theta(t, q)\big),
\end{aligned}
\tag{5.12}
$$

for $\forall i \in \{1, \ldots, d\}, q^i \in \mathcal{Q}^i$, and $t \in [0, T]$, and the terminal condition (5.10) becomes $\theta(T, q) = -\ell_d(q^1, \ldots, q^d)$.

As in the single-asset case, Equations (5.8) and (5.12) are in fact two systems of ODEs which belong to the same family. Let us introduce for $\xi > 0$ the functions

$$
H^{i,b}_\xi(p) = \sup_\delta \frac{\Lambda^{i,b}(\delta)}{\xi}\left(1 - \exp\left(-\xi\Delta^i(\delta - p)\right)\right)
$$

and

$$
H^{i,a}_\xi(p) = \sup_\delta \frac{\Lambda^{i,a}(\delta)}{\xi}\left(1 - \exp\left(-\xi\Delta^i(\delta - p)\right)\right),
$$

and the limit functions (for $\xi = 0$)

$$H_0^{i,b}(p) = \Delta^i \sup_\delta \Lambda^{i,b}(\delta)(\delta - p),$$

and

$$H_0^{i,a}(p) = \Delta^i \sup_\delta \Lambda^{i,a}(\delta)(\delta - p).$$

Then, we can consider the general equation

$$
\begin{aligned}
0 = {} & -\partial_t \theta(t,q) + \tfrac{1}{2}\gamma \sum_{i=1}^d \sum_{j=1}^d \rho^{i,j}\sigma^i\sigma^j q^i q^j \\
& - \sum_{i=1}^d \mathbf{1}_{q^i < Q^i} H_\xi^{i,b}\left(\tfrac{\theta(t,q)-\theta(t,q+\Delta^i e^i)}{\Delta^i}\right) - \sum_{i=1}^d \mathbf{1}_{q^i > -Q^i} H_\xi^{i,a}\left(\tfrac{\theta(t,q)-\theta(t,q-\Delta^i e^i)}{\Delta^i}\right),
\end{aligned}
\tag{5.13}
$$

for $\forall i \in \{1,\ldots,d\}, q^i \in \mathcal{Q}^i$, and $t \in [0,T]$, with the terminal condition

$$\theta(T,q) = -\ell_d(q^1,\ldots,q^d). \tag{5.14}$$

Equation (5.8) corresponds to Equation (5.13) for $\xi = \gamma$ while Equation (5.12) corresponds to Equation (5.13) for $\xi = 0$.

5.3. *Solution of the market making problem*

In order to characterize the optimal quotes in our multi-asset market making model, we proceed as in the single-asset case. In particular, we start by proving that there exists a solution of Equation (5.13) with terminal condition (5.14).

Theorem 5.1. *There exists a unique function $\theta : [0,T] \times \prod_{i=1}^d \mathcal{Q}^i \to \mathbb{R}$, C^1 in time, solution of Equation (4.1) with terminal condition (4.2).*

Proof.
Equation (5.13) with terminal condition (5.14) is a backward Cauchy problem. Because the functions $H_\xi^{i,b}$ and $H_\xi^{i,a}$ are functions of class C^1 for all $i \in \{1,\ldots,d\}$, the Cauchy–Lipschitz theorem applies, and there exists $\tau \in [0,T)$ and a function $\theta : (\tau, T] \times \prod_{i=1}^d \mathcal{Q}^i \to \mathbb{R}$, C^1 in time, solution of Equation (5.13) on $(\tau, T]$ with terminal condition (5.14).

$\forall q \in \prod_{i=1}^d \mathcal{Q}^i$, the function $t \in (\tau, T] \mapsto \theta(t,q) + \tfrac{1}{2}\gamma\sum_{i=1}^d\sum_{j=1}^d \rho^{i,j}\sigma^i\sigma^j q^i q^j(T-t)$ is a decreasing function. Therefore, the only reason why there would not be a global solution on $[0,T]$ is because $\sup_{q\in\prod_{i=1}^d \mathcal{Q}^i}\theta(t,q)$ blows up at $\tau > 0$. However, by using a comparison principle similar to that of Lemma 3.2, we easily see that

$$\sup_{q\in\prod_{i=1}^d \mathcal{Q}^i} \theta(t,q) \le \sum_{i=1}^d (H_\xi^{i,b}(0) + H_\xi^{i,a}(0))(T-t).$$

Therefore, $\sup_{q \in \prod_{i=1}^{d} \mathcal{Q}^i} \theta(t, q)$ cannot blow up in finite time, and θ is in fact defined on $[0, T] \times \prod_{i=1}^{d} \mathcal{Q}^i$.

Uniqueness comes then for the Cauchy–Lipschitz theorem. □

We are now ready to state the two theorems characterizing the optimal quotes in Model A and Model B. The proofs of these results are based on verification arguments and are (*mutatis mutandis*) identical to those in the single-asset case.

Let us start with Model A.

Theorem 5.2. *Let us consider the solution θ of Equation (5.13) with terminal condition (5.14) for $\xi = \gamma$.*

Then, $u : (t, x, q, S) \mapsto -\exp(-\gamma(x + \sum_{i=1}^{d} q^i S^i + \theta(t, q)))$ defines a solution to Equation (5.5) with terminal condition (5.6), and

$$u(t, x, q, S) = \sup_{(\delta_s^{1,b}, \ldots, \delta_s^{d,b})_{s \geq t}, (\delta_s^{1,a}, \ldots, \delta_s^{d,a})_{s \geq t} \in \mathcal{A}(t)^d} \mathbb{E}\left[-\exp\left(-\gamma\left(X_T^{t,x,\delta^b,\delta^a} + \sum_{i=1}^{d} q_T^{i,t,q^i,\delta^b,\delta^a} S_T^{i,t,S^i} \right. \right. \right.$$

$$\left. \left. \left. - \ell_d(q_T^{1,t,q^1,\delta^b,\delta^a}, \ldots, q_T^{d,t,q^d,\delta^b,\delta^a}) \right) \right) \right],$$

where

$$\forall i \in \{1, \ldots, d\}, \quad dS_s^{i,t,S^i} = \sigma^i dW_s^i, \qquad S_t^{i,t,S^i} = S^i,$$

$$dX_s^{t,x,\delta^b,\delta^a} = \sum_{i=1}^{d} (S_s^i + \delta_s^{i,a}) \Delta^i dN_s^{i,a} - (S_s^i - \delta_s^{i,b}) \Delta^i dN_s^{i,b}, \qquad X_t^{t,x,\delta^b,\delta^a} = x,$$

$$\forall i \in \{1, \ldots, d\}, \quad dq_s^{i,t,q^i,\delta^b,\delta^a} = \Delta^i dN_s^{i,b} - \Delta^i dN_s^{i,a}, \qquad q_t^{i,t,q^i,\delta^b,\delta^a} = q^i,$$

and where $\forall i \in \{1, \ldots, d\}$, the point processes $N^{i,b}$ and $N^{i,a}$ have stochastic intensity $(\lambda_s^{i,b})_s$ and $(\lambda_s^{i,a})_s$ given by $\lambda_s^{i,b} = \Lambda^{i,b}(\delta_s^{i,b}) 1_{q_{s_-}^i < Q^i}$ and $\lambda_s^{i,a} = \Lambda^{i,a}(\delta_s^{i,a}) 1_{q_{s_-}^i > -Q^i}$.

For $i \in \{1, \ldots, d\}$, the optimal bid and ask quotes $S_t^{i,b} = S_t^i - \delta_t^{i,b*}$ (for $q_{t_-}^i < Q^i$) and $S_t^{i,a} = S_t^i + \delta_t^{i,a*}$ (for $q_{t_-}^i > -Q^i$) are characterized by

$$\delta_t^{i,b*} = \tilde{\delta}_\gamma^{i,b*}\left(\frac{\theta(t, q_{t_-}) - \theta(t, q_{t_-} + \Delta^i e^i)}{\Delta^i} \right) \text{ and}$$

$$\delta_t^{i,a*} = \tilde{\delta}_\gamma^{i,a*}\left(\frac{\theta(t, q_{t_-}) - \theta(t, q_{t_-} - \Delta^i e^i)}{\Delta^i} \right), \tag{5.15}$$

where the functions $\tilde{\delta}_\gamma^{i,b*}(\cdot)$ and $\tilde{\delta}_\gamma^{i,a*}(\cdot)$ are defined by

$$\tilde{\delta}_\gamma^{i,b*}(p) = \Lambda^{i,b-1}\left(\gamma H_\gamma^{i,b}(p) - \frac{H_\gamma^{i,b'}(p)}{\Delta^i}\right) \text{ and}$$

$$\tilde{\delta}_\gamma^{i,a*}(p) = \Lambda^{i,a-1}\left(\gamma H_\gamma^{i,a}(p) - \frac{H_\gamma^{i,a'}(p)}{\Delta^i}\right).$$

For model B, the result is the following.

Theorem 5.3. *Let us consider the solution θ of Equation (5.13) with terminal condition (5.14) for $\xi = 0$.*

Then, $u : (t, x, q, S) \mapsto x + \sum_{i=1}^d q^i S^i + \theta(t, q)$ defines a solution to Equation (5.9) with terminal condition (38), and

$$u(t, x, q, S) = \sup_{(\delta_s^{1,b}, \ldots, \delta_s^{d,b})_{s \geq t}, (\delta_s^{1,a}, \ldots, \delta_s^{d,a})_{s \geq t} \in \mathcal{A}(t)^d} \mathbb{E}\left[X_T^{t,x,\delta^b,\delta^a} + \sum_{i=1}^d q_T^{i,t,q^i,\delta^b,\delta^a} S_T^{i,t,S^i}\right.$$
$$\left. - \ell_d(q_T^{1,t,q^1,\delta^b,\delta^a}, \ldots, q_T^{d,t,q^d,\delta^b,\delta^a}) - \frac{1}{2}\gamma \int_0^T \sum_{i=1}^d \sum_{j=1}^d \rho^{i,j} \sigma^i \sigma^j q_t^{i,t,q^i,\delta^b,\delta^a} q_t^{j,t,q^j,\delta^b,\delta^a} dt\right],$$

where:

$$\forall i \in \{1, \ldots, d\}, \quad dS_s^{i,t,S^i} = \sigma^i dW_s^i, \qquad S_t^{i,t,S^i} = S^i,$$

$$dX_s^{t,x,\delta^b,\delta^a} = \sum_{i=1}^d (S_s^i + \delta_s^{i,a})\Delta^i dN_s^{i,a} - (S_s^i - \delta_s^{i,b})\Delta^i dN_s^{i,b}, \qquad X_t^{t,x,\delta^b,\delta^a} = x,$$

$$\forall i \in \{1, \ldots, d\}, \quad dq_s^{i,t,q^i,\delta^b,\delta^a} = \Delta^i dN_s^{i,b} - \Delta^i dN_s^{i,a}, \qquad q_t^{i,t,q^i,\delta^b,\delta^a} = q^i,$$

and where $\forall i \in \{1, \ldots, d\}$, the point processes $N^{i,b}$ and $N^{i,a}$ have stochastic intensity $(\lambda_s^{i,b})_s$ and $(\lambda_s^{i,a})_s$, given by $\lambda_s^{i,b} = \Lambda^{i,b}(\delta_s^{i,b})1_{q_{s-}^i < Q^i}$ and $\lambda_s^{i,a} = \Lambda^{i,a}(\delta_s^{i,a})1_{q_{s-}^i > -Q^i}$.

For $i \in \{1, \ldots, d\}$, the optimal bid and ask quotes $S_t^{i,b} = S_t^i - \delta_t^{i,b*}$ (for $q_{t-}^i < Q^i$) and $S_t^{i,a} = S_t^i + \delta_t^{i,a*}$ (for $q_{t-}^i > -Q^i$) are characterized by

$$\delta_t^{i,b*} = \tilde{\delta}_0^{i,b*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} + \Delta^i e^i)}{\Delta^i}\right) \text{ and}$$

$$\delta_t^{i,a*} = \tilde{\delta}_0^{i,a*}\left(\frac{\theta(t, q_{t-}) - \theta(t, q_{t-} - \Delta^i e^i)}{\Delta^i}\right), \tag{5.16}$$

where the functions $\tilde{\delta}_0^{i,b*}(\cdot)$ and $\tilde{\delta}_0^{i,a*}(\cdot)$ are defined by

$$\tilde{\delta}_0^{i,b*}(p) = \Lambda^{i,b^{-1}}\left(-\frac{H_0^{i,b'}(p)}{\Delta^i}\right) \ and \ \tilde{\delta}_0^{i,a*}(p) = \Lambda^{i,a^{-1}}\left(-\frac{H_0^{i,a'}(p)}{\Delta^i}\right).$$

5.4. *About closed-form approximations*

In the single-asset case, closed-form approximations were obtained in Section 4, in the special case where $\Lambda^b = \Lambda^a =: \Lambda$ and $H_\xi''(0) > 0$.[21] In the multi-asset case, if we assume that $\forall i \in \{1, \ldots, d\}, \Lambda^{i,b} = \Lambda^{i,a} =: \Lambda^i$, and $H_\xi^{i''}(0) > 0$, then it is natural to wonder whether the same techniques can be used in order to obtain closed-form approximations.

The answer is in fact that the change of variables used to derive closed-form approximations does not work in general in dimension higher than 1. However, the idea of transforming Equation (5.13) into a multidimensional equivalent of Equation (4.1) enables to obtain results, without using the Hopf–Cole transform – that is, without relying on a multidimensional equivalent of Equation (4.2).

Following the same reasoning as in Section 4, we can indeed introduce the PDE

$$0 = -\partial_t \tilde{\theta}(t, q) + \frac{1}{2}\gamma \sum_{i=1}^d \sum_{j=1}^d \rho^{i,j}\sigma^i\sigma^j q^i q^j - 2 \sum_{i=1}^d H_\xi^i(0)$$

$$- \sum_{i=1}^d H_\xi^{i''}(0)(\partial_{q^i}\tilde{\theta}(t, q))^2 + \Delta^i H_\xi^{i'}(0)\partial_{q^i q^i}^2 \tilde{\theta}(t, q), \tag{5.17}$$

with final condition $\tilde{\theta}(T, q_1, \ldots, q_d) = -\ell_d(q_1, \ldots, q_d)$.

In the case where $\ell_d(q_1, \ldots, q_d) = \sum_{i=1}^d \sum_{j=1}^d a^{i,j} q^i q^j$ with $(a^{i,j})_{i,j}$ a symmetric positive matrix, it is easy to see that Equation (5.17) can be solved in closed form by using the ansatz

$$\tilde{\theta}(t, q) = \theta_0(t) - q'\theta_2(t)q,$$

where $\theta_2(t)$ is a $d \times d$ symmetric matrix (see the companion paper Evangelista, Guéant, and Vieira (2017)).

In particular, we show in Evangelista, Guéant, and Vieira (2017) that $\theta_2(t)$ verifies:

$$\theta_2(t) \to_{T \to +\infty} \frac{1}{2}\sqrt{\frac{\gamma}{2}}\Gamma,$$

where[22]

$$\Gamma = D^{-\frac{1}{2}}\left(D^{\frac{1}{2}}\Sigma D^{\frac{1}{2}}\right)^{\frac{1}{2}}D^{-\frac{1}{2}}, \qquad D = \begin{pmatrix} H_\xi^{1''}(0) & 0 & \ldots & 0 \\ 0 & H_\xi^{2''}(0) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & H_\xi^{d''}(0) \end{pmatrix}.$$

As a consequence, we can consider the approximations

$$\frac{\theta(t,q) - \theta(t, q + \Delta^i e^i)}{\Delta^i} \simeq \sqrt{\frac{\gamma}{2}}\left(\Gamma^{ii}\frac{2q^i + \Delta^i}{2} + \sum_{1 \leq j \leq d, j \neq i}\Gamma^{ij}q^j\right)$$

and

$$\frac{\theta(t,q) - \theta(t, q - \Delta^i e^i)}{\Delta^i} \simeq \sqrt{\frac{\gamma}{2}}\left(\Gamma^{ii}\frac{2q^i - \Delta^i}{2} + \sum_{1 \leq j \leq d, j \neq i}\Gamma^{ij}q^j\right).$$

These approximations can be plugged into Equations (5.15) and (5.16) to obtain the general approximation formulas

$$\delta_t^{i,b*} \simeq \delta_{\text{approx}}^{i,b*}(q_{t-}) := \tilde{\delta}_\xi^{i*}\left(\sqrt{\frac{\gamma}{2}}\left(\Gamma^{ii}\frac{2q_{t-}^i + \Delta^i}{2} + \sum_{1 \leq j \leq d, j \neq i}\Gamma^{ij}q_{t-}^j\right)\right) \tag{5.18}$$

and

$$\delta_t^{i,a*} \simeq \delta_{\text{approx}}^{i,a*}(q_{t-}) := \tilde{\delta}_\xi^{i*}\left(-\sqrt{\frac{\gamma}{2}}\left(\Gamma^{ii}\frac{2q_{t-}^i - \Delta^i}{2} + \sum_{1 \leq j \leq d, j \neq i}\Gamma^{ij}q_{t-}^j\right)\right), \tag{5.19}$$

where

$$\tilde{\delta}_\xi^{i*}(p) = \Lambda^{i-1}\left(\xi H_\xi^i(p) - \frac{H_\xi^{i'}(p)}{\Delta^i}\right). \tag{5.20}$$

These approximation formulas are interesting because we see the cross-effects coming from the non-diagonal terms of the matrix Γ.

6. Application: the case of two credit indices

In this section, we apply our single-asset and multi-asset market making models, along with the associated closed-form approximations, to the case of two credit (or CDS) indices: the IG index CDX.NA.IG and the HY index CDX.NA.HY. We consider a market maker who is in charge of proposing bid and ask quotes for these two indices, and we will assume throughout this section that this market maker is only concerned with spread risk and not with default risk – this hypothesis is always made by practitioners for market making fixed income and credit instruments.

Without going into the details of these indices,[23] we need to specify their main financial characteristics. Basically, for the IG index, the protection buyer pays quarterly (at fixed dates in order to ease compensation) a coupon corresponding to an annualized rate of 100 bps and pays upfront an amount (positive or negative) corresponding to an upfront rate (positive or negative) determined by the market. In practice, for market making, the upfront rate is the relevant variable because a round trip on the index leads to a PnL corresponding to the difference between upfront rates (times the notional of the transaction). However, in practice, this index is quoted in spread – this spread being computed using a basic CDS model. For the HY index, the protection buyer pays quarterly (at fixed dates) a coupon corresponding to an annualized rate of 500 bps and pays upfront an amount (positive or negative) corresponding to an upfront rate (positive or negative) determined by the market. Unlike the IG index, the HY index is quoted in upfront rate, or more precisely as $100(1 - \text{upfront rate})$. It is also noteworthy that, in practice, buying the IG index means buying protection, whereas buying the HY index means selling protection. For simplifying the exposition, we will consider that buying always means buying protection, and that the index quotes are the upfront rates. The conversion of our numerical results into market standard quotes can easily be carried out by using a basic CDS model.

In order to apply our models to these credit indices, we need first to estimate the value of the different parameters. This has been done thanks to the data provided by BNP Paribas in the framework of the Research Initiative 'Nouveaux traitements pour les données lacunaires issues des activités de crédit', which is financed by BNP Paribas under the aegis of the Europlace Institute of Finance. For estimating the volatilty and correlation parameters σ^{IG}, σ^{HY}, and ρ, mid-prices (prices here are upfront rates) have been considered. For the intensity functions, exponential intensities have been considered and the parameters A^{IG}, k^{IG}, A^{HY}, and k^{HY} have been estimated with classical likelihood maximization techniques using real quotes posted by the bank and the trades occurring between the bank and other market participants.[24]

If we consider that the two theoretical assets correspond to \$1 of each index respectively, then the value of the parameters are the following (figures are rounded):

	IG index		HY index
$\sigma \, (\$.s^{-\frac{1}{2}})$	$\sigma^{IG} = 5.83 \cdot 10^{-6}$		$\sigma^{HY} = 2.15 \cdot 10^{-5}$
ρ		$\rho = 0.9$	
$A \, (s^{-1})$	$A^{IG} = 9.10 \cdot 10^{-4}$		$A^{HY} = 1.06 \cdot 10^{-3}$
$k \, (\$^{-1})$	$k^{IG} = 1.79 \cdot 10^{4}$		$k^{HY} = 5.47 \cdot 10^{3}$

Coming now to the size of orders, we consider orders of size $\Delta^{IG} = \$50$ million for the IG index, and orders of size $\Delta^{HY} = \$10$ million for the HY index.

As far as risk aversion is concerned, we consider a reference value $\gamma = 6 \times 10^{-5}\$^{-1}$.

Regarding risk limits, we consider that $\frac{Q^{IG}}{\Delta^{IG}} = \frac{Q^{HY}}{\Delta^{HY}} = 4$.

Finally, we always consider a final time $T = 7200\,s$, corresponding to 2 h. We will see indeed on the examples below that the asymptotic regime is reached very rapidly, in far less than 2 h.

We can consider first the case of the IG index alone. We approximated the solution θ of the systems of ODEs (3.9) by using an implicit scheme and a Newton's method at each time step to deal with the nonlinearity. Then we obtained the feedback control function

$$(t, q^{IG}) \mapsto (\delta^{IG,b}(t, q^{IG}), \delta^{IG,a}(t, q^{IG})),$$

which gives the optimal bid and ask quotes[25] at time t when $q^{IG}_{t^-} = q^{IG}$.

We see in Figure 1 that the asymptotic regime is reached after less than 1 h.

In Figures 2 and 3, we plot the initial (i.e., asymptotic) values of the bid and ask quotes, obtained with Model A, for the IG index, when it is considered on a stand-alone basis. We see that the market maker quotes conservatively at the bid and aggressively at the ask when he is long, and conversely that he quotes conservatively at the ask and aggressively at the bid when he is short.

We also see that the closed-form approximations are satisfactory for small values of the inventory (in absolute value), but more questionable for larger values. In particular, the optimal quotes are not affine functions of the inventory as the closed-form approximations suggest.

The difference between actual values, obtained through the numerical approximation of the solution of a system of ODEs, and closed-form approximations can also be seen in Figures 4 and 5, which represent the bid–ask spread and the skew of a market maker quoting optimally. The bid–ask spread is indeed not constant, and the skew is not linear on our example.

However, if we consider market conditions with less volatility, then the closed-form approximations are far better – see Figure 6 where we computed the optimal bid quotes (in Model A) for a value of σ^{IG} divided by 2 (the result is the same for the ask quotes). The quality of the approximations depends therefore strongly on the considered market
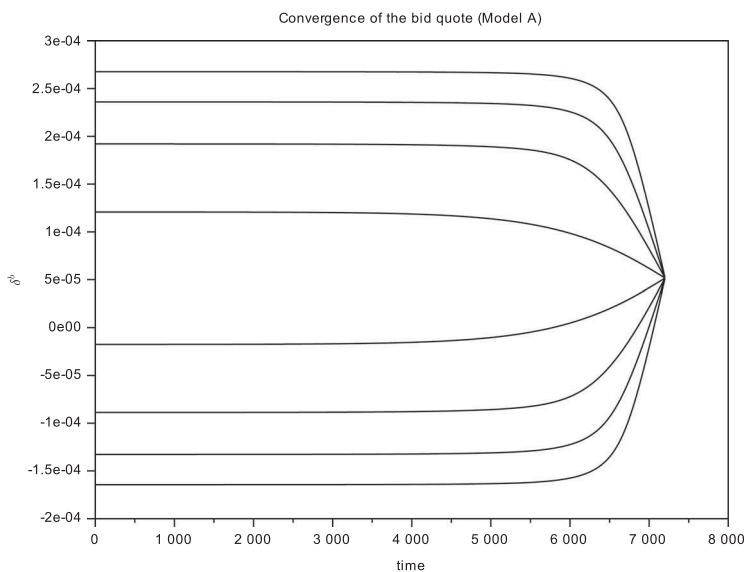
Figure 1. $t \mapsto \delta^{IG,b}(t, q^{IG})$ in Model A for the different values of q^{IG}.
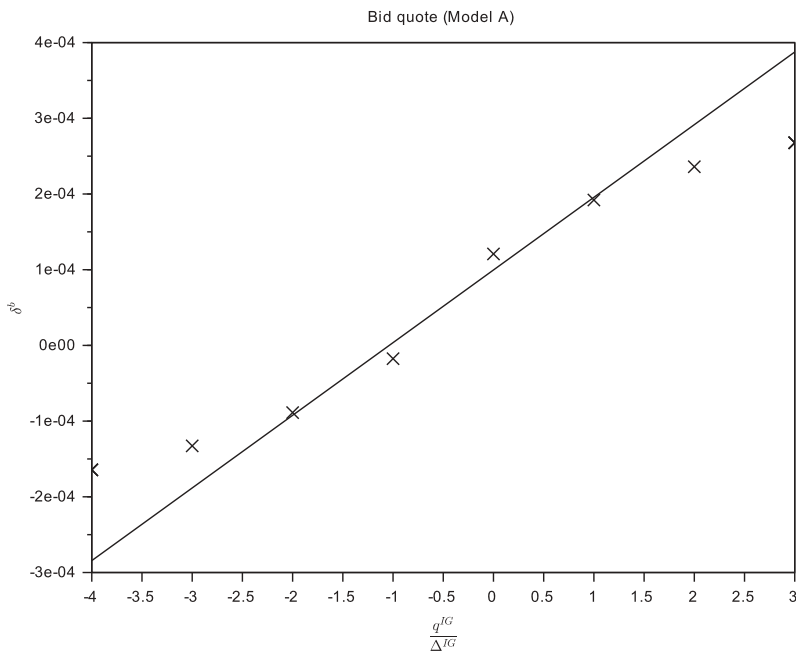
Figure 2. $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.6) – in the case of Model A.
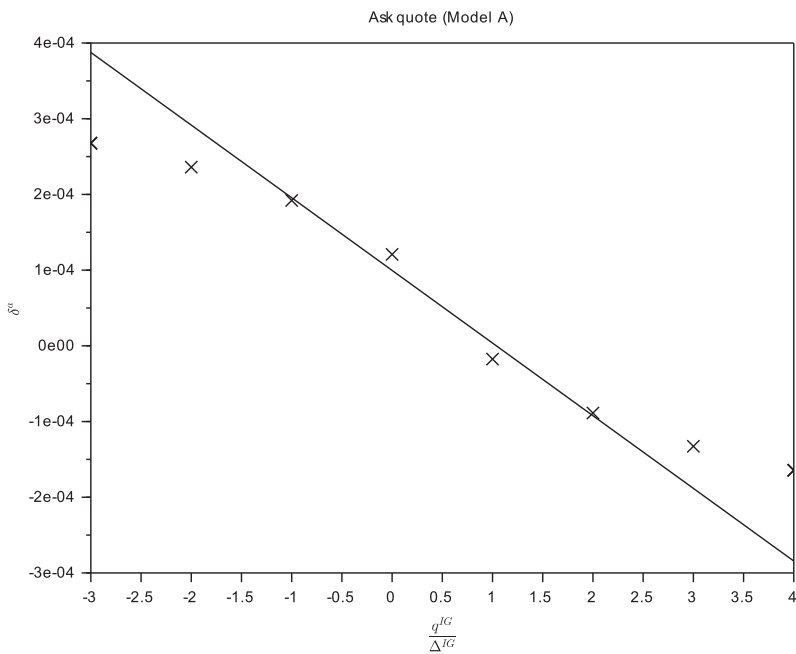
Figure 3. $q^{IG} \mapsto \delta^{IG,a}(0, q^{IG})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.7) – in the case of Model A.
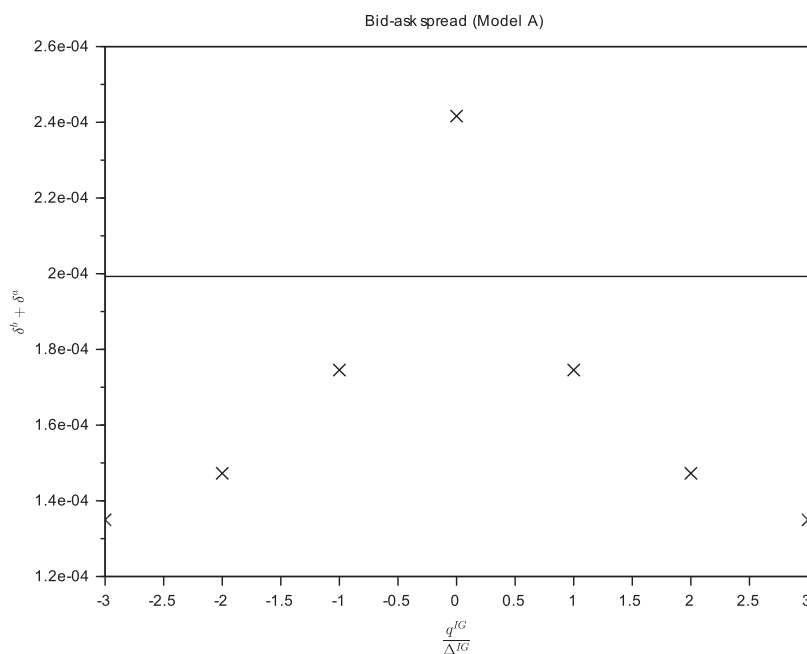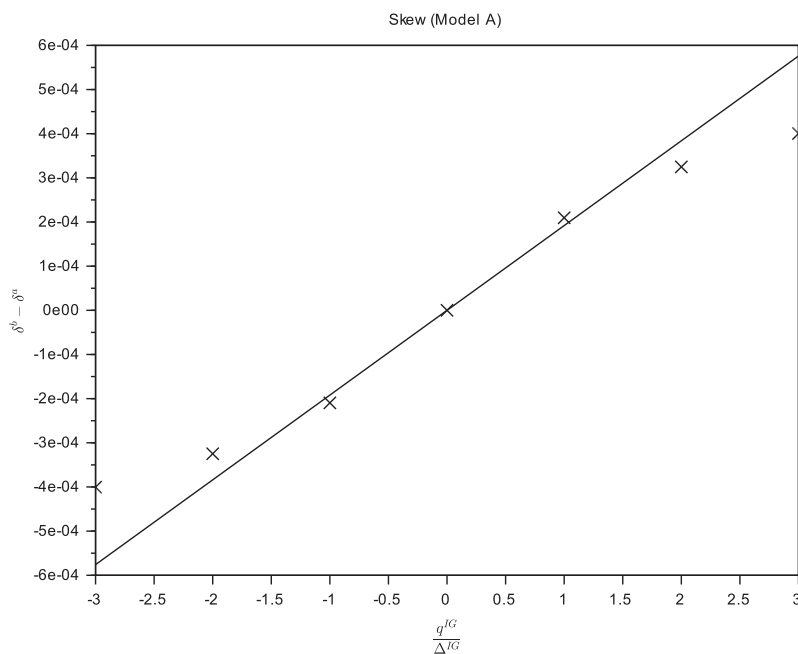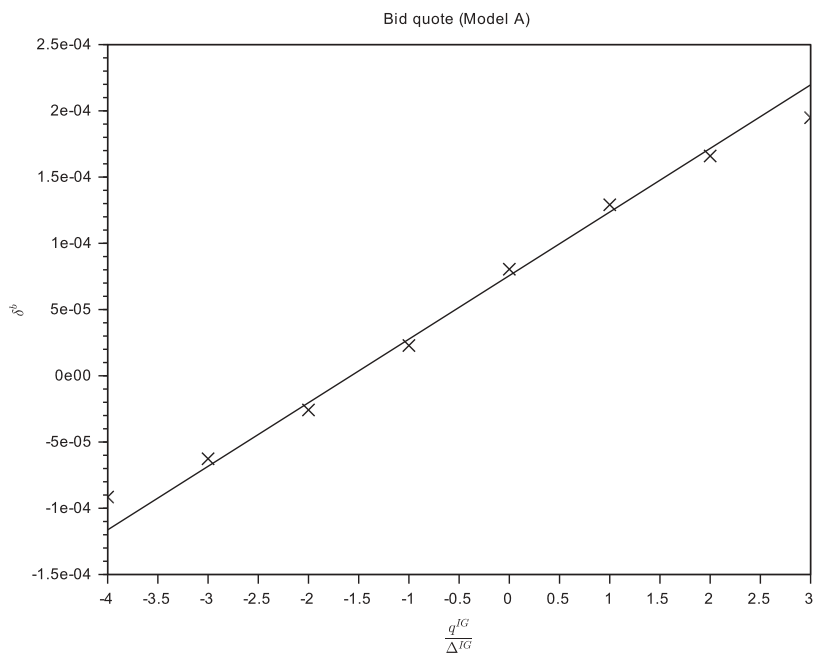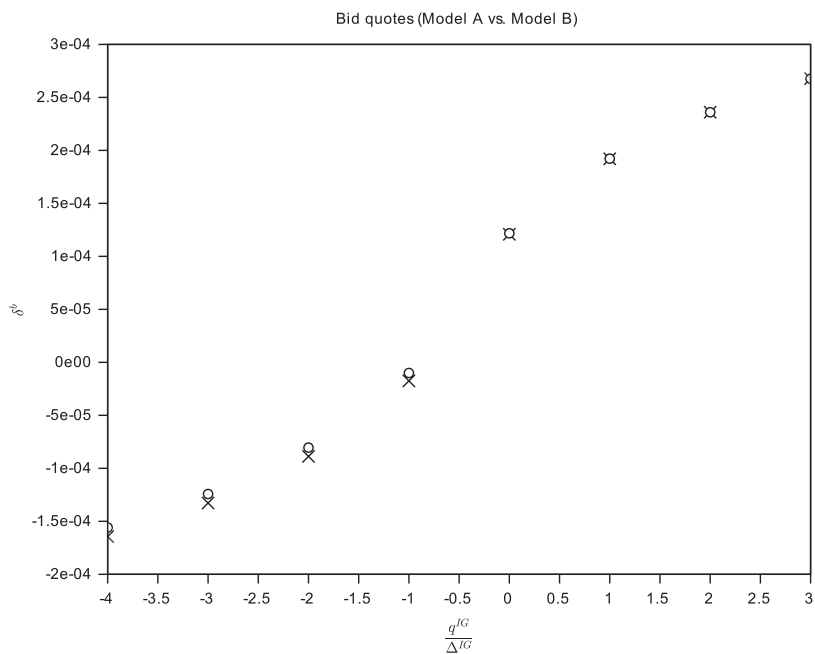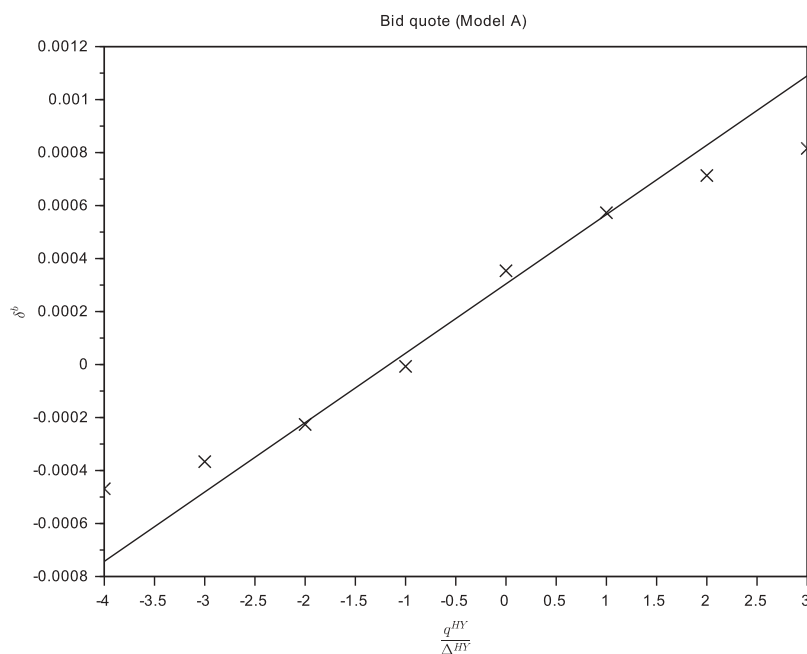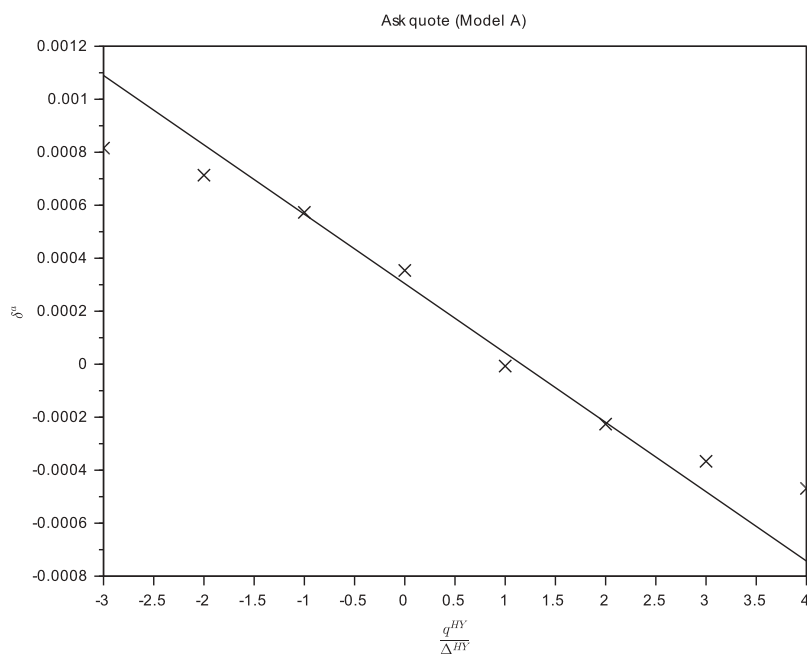
Figure 4. $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG}) + \delta^{IG,a}(0, q^{IG})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.8) – in the case of Model A.

Figure 5. $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG}) - \delta^{IG,a}(0, q^{IG})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.9) – in the case of Model A.

Figure 6. $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.6) – in the case of Model A, when σ^{IG} is reduced by half.

Figure 7. $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG})$ in Model A (crosses) and $q^{IG} \mapsto \delta^{IG,b}(0, q^{IG})$ in Model B (circles).

Figure 8. $q^{HY} \mapsto \delta^{HY,b}(0, q^{HY})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.6) – in the case of Model A.

Figure 9. $q^{HY} \mapsto \delta^{HY,a}(0, q^{HY})$ (crosses) and the associated closed-form approximations (line) obtained with Equation (4.7) – in the case of Model A.

and on the market context. Practitioners must subsequently understand in depth the trade-off between accuracy and computational time (especially when there are hundreds of assets) in order to choose between the two methods.

So far in this section, we have only considered optimal quotes in Model A. We see in Figure 7 that the differences between the two models are in fact very small. In other words, although Model B ignores part of the risk (or more precisely aversion to part of the risk), it constitutes a very interesting simplification of Model A.

Let us now come to the case of the HY index alone. Like for the IG index, we approximated the solution θ of the systems of ODEs (3.9) by using an implicit scheme and a Newton's method at each time step to deal with the nonlinearity. Then we obtained the feedback control function

$$(t, q^{HY}) \mapsto (\delta^{HY,b}(t, q^{HY}), \delta^{HY,a}(t, q^{HY}))$$

which gives the optimal bid and ask quotes at time t when $q_{t-}^{HY} = q^{HY}$.

In Figures 8 and 9, we plot the initial (i.e., asymptotic[26]) values of the bid and ask quotes, obtained with Model A, for the HY index, when it is considered on a stand-alone basis. As earlier, we see that the market maker quotes conservatively at the bid and aggressively at the ask when he is long, and conversely that he quotes conservatively at the ask and aggressively at the bid when he is short. We also see that the closed-form approximations are satisfactory only for small values of the inventory (in absolute value).

As far as the comparison between Model A and Model B is concerned, we see in Figure 10 that the differences between the two models are very small, as in the case of the IG index.
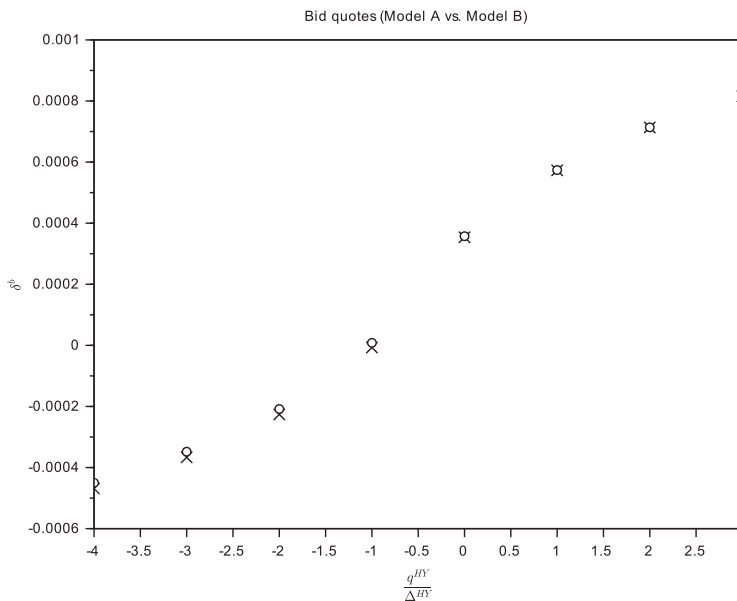

Figure 10. $q^{HY} \mapsto \delta^{HY,b}(0, q^{HY})$ in Model A (crosses) and $q^{HY} \mapsto \delta^{HY,b}(0, q^{HY})$ in Model B (circles).
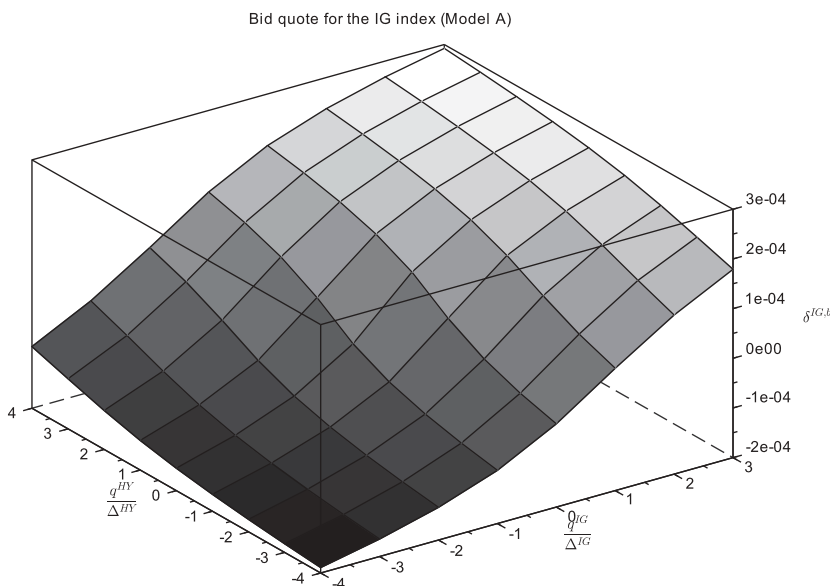
Bid quote for the IG index (Model A)

Figure 11. $(q^{IG}, q^{HY}) \mapsto \delta^{IG,b}(0, q^{IG}, q^{HY})$ – in the case of Model A.
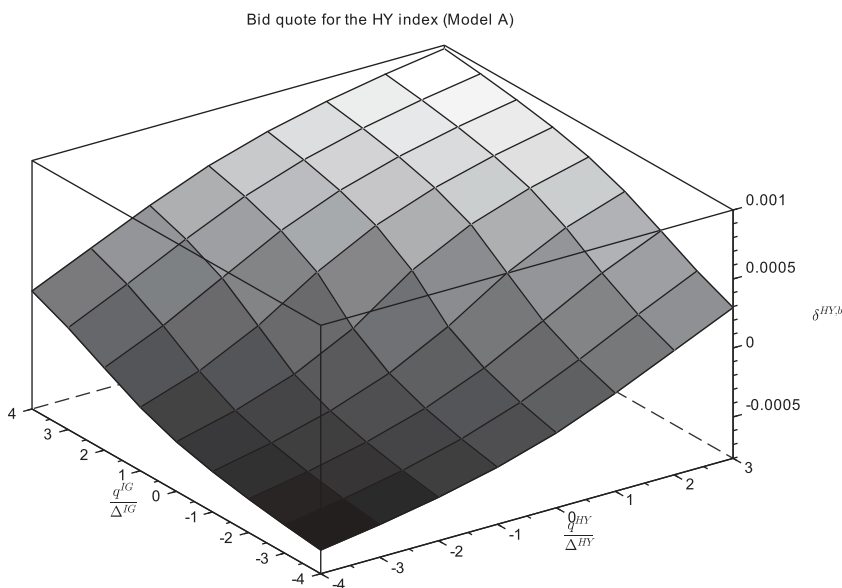
Bid quote for the HY index (Model A)

Figure 12. $(q^{IG}, q^{HY}) \mapsto \delta^{HY,b}(0, q^{IG}, q^{HY})$ – in the case of Model A.

We can now consider the two indices together and look at the influence of correlation for the market making of several assets at the same time. We approximated the solution θ of the systems of ODEs (5.13) by using an implicit scheme and a Newton's method at each time step to deal with the nonlinearity.
Then we obtained the feedback control function
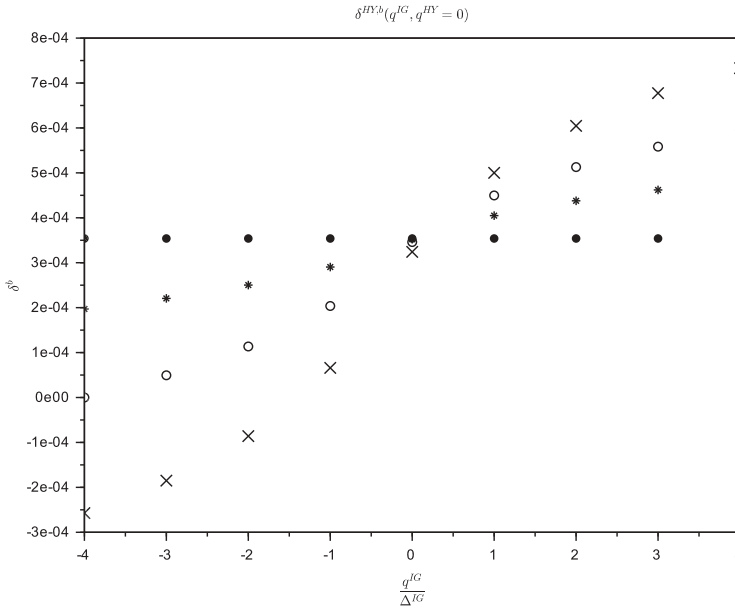
Figure 13. $q^{IG} \mapsto \delta^{HY,b}(0, q^{IG}, 0)$ in the case of Model A, for different values of ρ. $\rho = 0.9$ (crosses), $\rho = 0.6$ (circles), $\rho = 0.3$ (stars) and $\rho = 0$ (dots).

$$(t, q^{IG}, q^{HY}) \mapsto (\delta^{IG,b}(t, q^{IG}, q^{HY}), \delta^{IG,a}(t, q^{IG}, q^{HY}), \delta^{HY,b}(t, q^{IG}, q^{HY}), \delta^{HY,a}(t, q^{IG}, q^{HY})),$$

which gives the optimal bid and ask quotes at time t for the two indices when $q_{t-}^{IG} = q^{IG}$ and $q_{t-}^{HY} = q^{HY}$.

In Figures 11 and 12, we have plotted the optimal bid quotes for the two indices.[27] We see that the market maker's inventory on both indices influences his quotes. Because the correlation coefficient is positive, $(q^{IG}, q^{HY}) \mapsto \delta^{IG,b}(0, q^{IG}, q^{HY})$ and $(q^{IG}, q^{HY}) \mapsto \delta^{HY,b}(0, q^{IG}, q^{HY})$ are increasing in q^{IG} and q^{HY}.

To see the influence of correlation, we have also computed the optimal quotes for four values of the correlation parameter: $\rho \in \{0, 0.3, 0.6, 0.9\}$. Figure 13 represents, for these different values of ρ, the bid quote $\delta^{HY,b}(0, q^{IG}, 0)$ for the HY index, when the inventory with respect to the HY index is equal to 0, for different values of the inventory with respect to the IG index. We see that the correlation coefficient has a strong influence on the optimal quote: the more correlated the two assets, the more conservatively (respectively aggressively) the market maker should quote at the bid when he has a long (respectively short) inventory in the other asset.

7. Conclusion

In this paper, we considered a framework à la Avellaneda–Stoikov with general intensity functions, and we showed that for the different optimization criteria used in the literature, the dimensionality of the problem can be divided by 2. We also showed how to find closed-form

approximations for the optimal quotes, generalizing therefore the Guéant–Lehalle–Fernandez-Tapia formulas (used by many in the industry) to the two kinds of objective function used in the literature and to almost any intensity function. We also generalized our model to the multi-asset case and showed the importance of taking account of the correlation between assets. In particular, we have derived closed-form approximations for the optimal quotes of a multi-asset market maker, an important breakthrough for practitioners who sometimes cannot solve systems of dozens or hundreds of nonlinear ODEs. The simple applications to credit indices we considered confirm the importance of the multi-asset framework.

Financial support received for presenting the research results in an international conference

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

Notes

1. The exact nature of these market prices depends on the considered market. In the case of most order-driven markets (such as most stock markets), a market price may be a mid-price. It may also be a price based on the most recent transactions. In the case of the European corporate bond market, the Composite Bloomberg Bond Trader (CBBT) price is a composite price which may be regarded as a proxy for the market price of a bond. In the case of the US corporate bond market, a market price may also be built by using a mix

between TRACE data (in spite of the lag) and the CBBT prices. In all cases, the market prices involved in the model should be regarded as reference prices.

2. See Cartea, Jaimungal, and Penalva (2015) and Guéant, Lehalle, and Fernandez-Tapia (2013) for models with adverse selection effects.

3. Market making has always been an important topic for economists – see for instance the model of Grossman and Miller (1988). However, the dynamic approaches proposed by mathematicians have shed a new light on market making and make it possible to build algorithms for replacing human market makers. The main (old) economic paper really related to the mathematical literature on market making is the paper Ho and Stoll (1981) – see also Ho and Stoll (1983). It is noteworthy that this old paper by Ho and Stoll inspired Avellaneda and Stoikov when they wrote their seminal paper (Avellaneda and Stoikov 2008).

4. See also Guéant and Lehalle (2015).

5. CARA means constant absolute risk aversion. CARA utility functions are utility functions of the form $u(x) = -e^{-\gamma x}$ for $\gamma > 0$.

6. Adverse selection is also considered in Guéant, Lehalle, and Fernandez-Tapia (2013).

7. One reason is the interest for high-frequency trading. High-frequency trading is indeed often discussed for its influence on the price formation process of stocks. Another reason is that some market making models can be regarded as generalizations of optimal execution models built to solve problems coming from the cash-equity industry – see for instance Bayraktar and Ludkovski (2014), Guéant, Lehalle, and Fernandez-Tapia (2012) and Huitema (2012).

8. Models à la Avellaneda–Stoikov can hardly be applied to most stock markets for at least two reasons: (i) the discrete nature of prices (especially in the case of stocks with a large tick size) and (ii) the fact that the very nature of the limit order books, which are queuing systems with priorities and volumes, is not taken into account. One of the only market making models really well suited to stocks is the model proposed by Guilbaud and Pham (2013) – see also Guéant (2016) for a variant.

9. There may not be a proper market price (see the earlier discussion), hence the wording 'reference price'.

10. Q is assumed to be a multiple of Δ.

11. The first three hypotheses are natural. The fourth one is more technical. It ensures in particular that the functions $\pi^b : \delta \mapsto \delta \Lambda^b(\delta)$ and $\pi^a : \delta \mapsto \delta \Lambda^a(\delta)$, which are related to the instantaneous (expected) MtM PnL associated with each side, reach a maximum on $\mathbb{R}$ (in fact on $\mathbb{R}_+$). To see this (we focus on the bid side, but the proof is similar for the ask side), let us notice that

$$\pi^{b\prime}(\delta) = 0 \Leftrightarrow \delta + \frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)} = 0.$$

But $v^b : \delta \mapsto \delta + \frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}$ is a strictly increasing function with

$$\inf_\delta v^{b\prime}(\delta) = 2 - \sup_\delta \frac{\Lambda^b(\delta)\Lambda^{b\prime\prime}(\delta)}{\left(\Lambda^{b\prime}(\delta)\right)^2} > 0.$$

Therefore, the equation $v^b(\delta) = 0$ has a unique solution and it corresponds to a unique maximizer for π^b.

12. We have indeed

$$j'(\delta) = 1 + \frac{1 - \frac{\Lambda^b(\delta)\Lambda^{b\prime\prime}(\delta)}{\Lambda^{b\prime}(\delta)^2}}{1 - \xi\Delta\frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}} = \frac{2 - \frac{\Lambda^b(\delta)\Lambda^{b\prime\prime}(\delta)}{\Lambda^{b\prime}(\delta)^2} - \xi\Delta\frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}}{1 - \xi\Delta\frac{\Lambda^b(\delta)}{\Lambda^{b\prime}(\delta)}} > 0.$$

13. The authors of Guéant, Lehalle, and Fernandez-Tapia (2013) use the linear system of ODEs (3.13) in the case $\Delta = 1$ and $\xi = \gamma$.

14. The condition $H_\xi''(0) > 0$ is always verified when $\xi = 0$ (see the proof of Lemma 3.1). A sufficient condition in general is

$$\forall \delta \in \mathbb{R}, \quad \xi\Delta\frac{\Lambda(\delta)^2\Lambda''(\delta)}{\Lambda'(\delta)^3} < 1.$$

This condition (obtained by using the expression of H_ξ'' in the proof of Lemma 3.1) is verified for instance if Λ is convex (exponential intensities enter this category).

15. We remove here the boundaries associated with $-Q$ and Q.

16. Another way to see this expansion is to consider an expansion of order 2 in Δ (an expansion of order 1 would correspond, after rescaling H_ξ, to a fluid-limit regime where non-execution risk vanishes) combined with an approximation of H_ξ by using the first three terms of its Taylor expansion (in 0).

17. One can see the proximity with Equation (17).

18. The basic reasoning consists in proving that the operator $\tilde{v} \mapsto -\frac{1}{2}\frac{H_\xi''(0)}{\Delta H_\xi'(0)}\gamma\sigma^2 q^2\tilde{v} + \Delta H_\xi'(0)\partial_{qq}^2\tilde{v}$ is a positive self-adjoint operator with a compact inverse (see Chapter 6 of Brezis (2010) for more details). Therefore, this operator can be diagonalized in an orthonormal basis. Its minimum eigenvalue can be shown to be simple by using the same methodology as in Guéant, Lehalle, and Fernandez-Tapia (2013).

19. In particular, in the case of exponential intensities, the actual bid–ask spread is not independent of q.

20. We denote by $(e^1, \ldots, e^d)$ the canonical basis of $\mathbb{R}^d$.

21. Very recently, closed-form approximations have also been found in the case of asymmetric intensities – see Evangelista, Guéant, and Vieira (2017).

22. Square roots must be understood in the class of positive semi-definite matrices. Γ is well defined because D is positive and diagonal, and because Σ is positive semi-definite.

23. See www.markit.com for more details.

24. The period of estimation is the first semester of 2016.

25. In fact the difference between the reference price and the actual quote, as in the rest of the paper.

26. The convergence is as fast as in the IG case.

27. The results are similar, *mutatis mutandis*, for the ask quotes, and are not displayed.

References

Avellaneda, M., and S. Stoikov. 2008. "High-Frequency Trading in a Limit Order Book." *Quantitative Finance* 8 (3): 217–224. doi:10.1080/14697680701381228.

Bayraktar, E., and M. Ludkovski. 2014. "Liquidation in Limit Order Books with Controlled Intensity." *Mathematical Finance* 24 (4): 627–650. doi:10.1111/mafi.2014.24.issue-4.

Brezis, H. 2010. "Functional Analysis, Sobolev Spaces and Partial Differential Equations." *Springer Science & Business Media*.

Cartea, Á., R. Donnelly, and S. Jaimungal. 2013. "Algorithmic Trading with Model Uncertainty." SSRN 2310645.

Cartea, Á., and S. Jaimungal. 2013. "Risk Metrics and Fine Tuning of High Frequency Trading Strategies." *Mathematical Finance* 25 (3): 576–611. doi:10.1111/mafi.12023.

Cartea, Á., S. Jaimungal, and J. Ricci. 2014. "Buy Low, Sell High: A High Frequency Trading Perspective." *SIAM Journal on Financial Mathematics* 5 (1): 415–444. doi:10.1137/130911196.

Cartea, Á., S. Jaimungal, and J. Penalva. 2015. *Algorithmic and High-Frequency Trading*. Cambridge: Cambridge University Press.

Donnelly, R. 2014. "Ambiguity aversion in algorithmic and high frequency trading." PhD diss., University of Toronto.

Evangelista, D., O. Guéant, and D. Vieira. 2017. "New Closed-Form Approximations in Multi-Asset Market Making." *Preprint*.

Fodra, P., and M. Labadie. 2012. "High-Frequency Market-Making with Inventory Constraints and Directional Bets." *Arxiv Preprint Arxiv:1206.4810*.

Grossman, S., and M. Miller. 1988. "Liquidity and Market Structure." *The Journal of Finance* 43 (3): 617–633. doi:10.1111/j.1540-6261.1988.tb04594.x.

Guéant, O., C.-A. Lehalle, and J. Fernandez-Tapia. 2012. "Optimal Portfolio Liquidation with Limit Orders." *SIAM Journal on Financial Mathematics* 3 (1): 740–764. doi:10.1137/110850475.

Guéant, O., C.-A. Lehalle, and J. Fernandez-Tapia. 2013. "Dealing with the Inventory Risk: A Solution to the Market Making Problem." *Mathematics and Financial Economics* 7 (4): 477–507. doi:10.1007/s11579-012-0087-0.

Guéant, O., and C.-A. Lehalle. 2015. "General Intensity Shapes in Optimal Liquidation." *Mathematical Finance* 25 (3): 457–495. doi:10.1111/mafi.12052.

Guéant, O. 2016. *The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making*. CRC Press, Taylor and Francis.

Guilbaud, F., and H. Pham. 2013. "Optimal High-Frequency Trading with Limit and Market Orders." *Quantitative Finance* 13 (1): 79–94. doi:10.1080/14697688.2012.708779.

Ho, T., and H. Stoll. 1981. "Optimal Dealer Pricing under Transactions and Return Uncertainty." *Journal of Financial Economics* 9 (1): 47–73. doi:10.1016/0304-405X(81)90020-9.

Ho, T., and H. Stoll. 1983. "The Dynamics of Dealer Markets under Competition." *The Journal of Finance* 38 (4): 1053–1074. doi:10.1111/j.1540-6261.1983.tb02282.x.

Huitema, R. 2012. "Optimal Portfolio Execution Using Market and Limit Orders." working paper.

Menkveld, A. 2013. "High Frequency Trading and the New Market Makers." *Journal of Financial Markets* 16 (4): 712–740. doi:10.1016/j.finmar.2013.06.006.

Nyström, K., S. M. Ould Aly, and C. Zhang. 2014. "Market Making and Portfolio Liquidation under Uncertainty." *International Journal of Theoretical and Applied Finance* 17: 1450034. doi:10.1142/S0219024914500344.