cogent mathematics

*Corresponding author: Andrea Macrina, Department of Mathematics, University College London, London, UK; Department of Actuarial Science, University of Cape Town, Cape Town, South Africa
E-mail: a.macrina@ucl.ac.uk

APPLIED & INTERDISCIPLINARY MATHEMATICS | RESEARCH ARTICLE

Inventory management in customised liquidity pools

M. Alessandra Crisafi[1] and Andrea Macrina[1,2]*

Abstract: We consider "customised liquidity pools" (CLP), which are trading venues that offer over-the-counter brokerage and dealer services to selected market participants. The dealer activity, whereby two-sided liquidity is offered to a limited pool of clients, shares in common similarities with the market-making problem. The arrival flow of client orders is assumed random. The CLP offers a stream of two-way prices to its clients, which are functions of the amount traded by the client and the CLP holding. The concern is inventory risk, which increases for critically small or large numbers of held positions. The CLP controls its inventory by choosing the size and the skew of its spread, so to encourage, e.g. buy orders instead of sell orders. Furthermore, it can submit limit orders to standard exchanges, of which execution is uncertain, and market orders, which are expensive. In either case, the CLP risks information leakage, which is discouraged by a penalty for trading in standard exchanges. We numerically solve a double-obstacle impulse-control problem associated with the optimal management of the inventory. We observe that the CLP skews its spread before resorting to limit orders, and ultimately crosses the spread in the standard exchange. We learn that it is optimal to post limit orders deeper in the book when the inventory is relatively small and to progressively move towards the best price as the inventory increases. The CLP adjusts its pricing-hedging strategy according to its profit and losses (P&L) target, a feature we analyse by considering various degrees of the CLP's risk appetite.

ABOUT THE AUTHORS

Ms M. Alessandra Crisafi is a PhD candidate in the UCL Department of Mathematics. She holds an MPhil in Economics from the University of Cambridge and an MSc in Financial Mathematics from King's College London. Her main research interests are focused on algorithmic and high-frequency trading, and optimal stochastic control.

Dr Andrea Macrina is a Senior Lecturer in Mathematics and the Director of the Financial Mathematics MSc Programme in the UCL Department of Mathematics. In 2013, he was appointed to Adjunct Associate Professor in the Department of Actuarial Science, University of Cape Town. Dr Macrina holds a PhD in Mathematics from King's College, University of London, and an MSc in Physics from the University of Bern, Switzerland. His research in Financial and Insurance Mathematics includes asset pricing and real-time risk management, market information models, interest-rate modelling and inflation-linked pricing, optimal strategies in electronic markets, data analytics, and climate risk securitization.

PUBLIC INTEREST STATEMENT

In electronic trading, financial transactions are executed by computers at high speed. The computerised trading strategies are the result of cutting-edge algorithms, which are based on mathematical models. Such algorithms take current market conditions as inputs and produce the trading decisions, which are the solutions to stochastic optimal control problems. The proposed mathematical models may be utilised in the financial industry by firms offering liquidity to their clients in the form of bid and ask quotes. The present research work also provides the optimal management of the inventory held by the firm offering customised trade solutions. In doing so, the firm ensures that it does not expose itself to an unsafe risk level, which may arise if the trading books are too unbalanced. The optimal strategies, given a variety of market conditions, are elucidated by a detailed analysis of the models and are further explained in a number of graphs.

cogent·oa

cogent · mathematics

1. Introduction

Market makers are liquidity providers. They set bid and ask quotes and trade with impatient investors who seek to immediately buy or sell a certain quantity of a financial asset. A portion of the market-maker's P&L derives from the spread charged to the clients. On the other hand, holding a non-zero inventory carries an intrinsic risk associated with the unpredictable changes to which an asset price is subject. This risk is further increased by a potential information asymmetry due to which a market maker trades in the wrong direction.

We consider a financial entity that offers broker-dealer services to its clients. Such an entity may be a small financial shop, an individual trading desk of large institutions, as well as large firms which provide liquidity to a selected pool of clients. Among these, we find, e.g. investment banks, hedge funds and high-frequency traders. In the industry, such financial entity is sometimes referred to as *customised liquidity pool*[1] (CLP). These bespoke liquidity pools share some characteristics with so-called dark pools, while *additionally* providing the dealer service. As such CLPs may be viewed as "grey pools", that is, a kind of hybrid between a dark pool and a "lit" pool. A "lit" pool is a transparent venue where prices and trading orders are publicly displayed, while dark pools are venues that do not display prices and guarantee anonymity to the selected group of customers which trade within them. CLPs typically offer two-way prices to their clients while preserving their anonymity. Clients can compare prices from various dealers through an internal graphic user interface (GUI). applications, but there is no centralised liquidity pool[2] that displays those prices, which are streamed directly to the clients in conditions of market opacity and confidentiality. CLPs thus offer for all practical purposes a market-making service, though they have no *obligation* to offer two-sided liquidity at any time. In the remainder of the paper, we use the terms "dealer activity" and "market making" interchangeably.

In this paper we consider the optimal management of a CLP inventory. In our situation the CLP (i) offers liquidity to its clients (who may be both, buyers and sellers) and (ii) may post limit and market orders in a standard exchange ("lit" exchange, centralised liquidity pool) to control the level of its inventory. We emphasise that the work presented here has no particular asset class in mind, since the CLP may specialise in, e.g. stocks, commodities and foreign exchange trading. When we refer to standard exchanges, we consider centralised exchange platforms, which in principle are accessible to any market participant and which follow the price-priority rule, with first-in-first-out (FIFO) applied for limit orders posted at the same price.

A good summary of the algorithmic and high-frequency trading literature is provided in the books by Cartea, Jaimungal, & Penalva (2016) and Guéant (2016). Previous work on market making includes Amihud and Mendelson (1980) who, based on Garman (1976), relate the bid-ask prices to the share holding of a risk-neutral agent. They find a relationship between the optimal quotes and the distance from the "preferred" inventory position. Stoll (1978) considers a two-period model in which a risk-averse agent supplies liquidity and maximises his expected utility. Ho and Stoll (1981) use the dynamic programming principle (DPP) to obtain the optimal quotes which maximise the terminal wealth in a single-dealer market.

The recent evolution in financial markets, arisen with algorithmic and high-frequency trading, has shifted the optimal market-making problem to an order-driven market environment where optimal quotes and trading strategies are computed and submitted by electronic machines. For example Avellaneda and Stoikov (2008) treat a market-making problem in a limit order book (LOB). They

consider the maximisation of the agent's expected utility and consider both, the finite and the infinite-time cases. They model the arrival of buy and sell orders by Poisson processes and the dynamics of the mid-price by an arithmetic Brownian motion. They find and numerically solve the Hamilton-Jacobi-Bellman (HJB) PDE. This type of problem has been investigated elsewhere, too. The works by Cartea and Jaimungal (2013) on risk metrics and by Cartea, Donnelly, and Jaimungal (2014) consider ambiguity and self-exciting processes, respectively. Guéant, Lehalle, and Tapia (2013) deal with the inventory risk and reduce a complex optimisation problem to a system of ODEs. Guilbaud and Pham (2013) consider a market maker who continuously submits limit orders at the best quoted prices and resorts to market orders when the inventory becomes too large. They numerically solve a finite-time impulse-control problem and find the optimal order sizes and quotes to be posted in the standard exchange.

The model presented in this paper is inspired by Guilbaud and Pham (2013). In the present work we formulate a double-obstacle impulse-control problem and we use the viscosity theory to characterise the solution of the associated system of quasi variational inequalities (QVIs). We refer the reader to Crandall, Ishii, and Lions (1992), Fleming and Soner (2006) and Pham (2009) for a comprehensive treatment of viscosity solutions.

This paper is organised as follows. After the introduction, in Section 2, we present the trading strategies of the CLP for the purpose of inventory management. In Section 3, the optimisation problem is introduced and, by making use of the DPP, we derive the HJB equation. Section 4 is devoted to a numerical example of the CLP strategies, and in Section we summarise the results and research contributions of this work. All propositions, necessary to characterise the value function by the unique viscosity solution of the system of QVIs, and their proofs can be found in the Section 6.

The main contributions of the present work can be summarised as follows. First, the study of the inventory management of a firm, which can hedge via limit and market order, is a novelty. Second, the model presented in this work is intuitive, suitable for calibration and the results are in line with real-world practice. Additional (though more technical) innovations include the possibility of making the CLP's prices a function of the size traded by clients. Finally, we construct a flexible setup which can be adapted to various needs, e.g. different ways of calculating revenues, presence of transaction fees, different market-spread properties, etc.

2. Inventory management

We consider a CLP that supplies liquidity to its clients by providing bid and ask quotes. The pool trades with buyers and sellers by being their counterparty. It executes incoming buy and sell orders by its clients over a finite period of time $0 \leq u \leq T < \infty$ and may resort to the centralised exchange platform if its inventory becomes critically small or large. The CLP derives a profit by choosing an optimal spread at which it is willing to offer liquidity to its clients. In particular, we assume that the CLP price is aligned with the standard exchange mid-price S_u, for which we need a model. We fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_u\}_{0 \leq u \leq T}, \mathbb{P})$ satisfying the usual conditions.

Along the lines of Guilbaud and Pham (2013), we model the LOB bid-ask half spread by a continuous-time Markov chain $\{k_u\}_{0 \leq u \leq T}$ with a discrete state space $\mathbb{K} := \{k_0, k_1, k_2, \ldots, k_n\}$, where $k_0 < k_1 < k_2 < \cdots < k_n$ are set so to reflect the granularity of the standard-exchange prices. In particular, we let $k_j - k_{j-1} = \bar{k}$, which means that the LOB prices are at least $2\bar{k}$ units apart. The chain is generated by $\{Q\} = (r_{ij})$ such that $\mathbb{P}[k_{u+du} = k_j | k_u = k_i] = r_{ij}\, du + o(du)$ and $\mathbb{P}[k_{u+du} = k_i | k_u = k_i] = 1 + r_{ii}\, du + o(du)$, with $r_{ij} \geq 0$ for all $j \neq i$ and $r_{ii} = -\sum_{j \neq i} r_{ij}$. Such a choice has interesting financial justifications. Because the spread is a measure of the market liquidity, lower states in $\mathbb{K}$ (e.g. k_0) may be associated with periods of higher liquidity compared to higher states (e.g. k_n). Also, by correctly choosing the transition probabilities, one can postulate the existence of a "normal" level of the spread—for each particular class of assets—from which deviations are unlikely to happen. Furthermore, the transition from one state to another can be associated with the submission of both, aggressive and passive orders by market participants. We remark that when we write

k_u, we mean the level of the spread at time u, whereas with the notation $k \in \mathbb{K}$ we specify a particular state k_j, for $j = 0, 1, 2, \ldots, n$.

We define the LOB mid-price by

$$\mathrm{d}S_u = \bar{k}^{\pm}_{M_{u^-}} \, \mathrm{d}M_u, \tag{2.1}$$

where $\{M_u\}_{0 \leq u \leq T}$ is a Poisson process with intensity λ^m and $\bar{k}^{\pm}_i$ is a collection of i.i.d. random variables valued in $\{-\bar{k}, \bar{k}\}$. At $u \in [0, T]$, the best LOB bid and ask prices are given by $S^b_u = S_u - k_u$ and $S^a_u = S_u + k_u$, respectively.

We choose such simple dynamics in Equation (2.1)—i.e. the mid price can only move half a tick almost surely—because our intention is to focus on the market-making activity of the CLP, disregarding any potential market-view it may have, or sudden price moves. At each time $u \in [0, T]$, we consider three options for the inventory management: (i) the CLP's order flow may be controlled by accordingly choosing the prices offered to its clients, (ii) a limit order—of which execution is uncertain—is posted to the standard exchange, or (iii) a costly market order is submitted to the standard exchange.

Our goal is thus to obtain the critical levels of the inventory for which it is optimal to (i) skew the CLP prices, (ii) submit limit orders, and (iii) submit market orders. Such critical levels derive from the impulse-part of the optimisation problem presented in Section 3 and determine the hedging strategy of the CLP. In particular, if either of such levels is exceeded by the CLP's inventory, the CLP starts to post orders in the standard exchange so to reduce the magnitude of its inventory.

2.1. Pricing strategy

We consider the processes δ^+_u, δ^-_u to be (i) the controls of the optimisation problem presented in Section 3 and (ii) part of the optimal strategy of the CLP (which also includes the impulse-part). We define by $S_u - \delta^-_u$ and $S_u + \delta^+_u$ the *base prices* and by $\delta^+_u + \delta^-_u$ the *base spread*. The CLP optimally chooses such processes, from which it derives the prices offered to its clients. In fact, the actual spread paid by the client to the CLP is also a function of the client's order size, see Equation (2.3). The CLP does not bet on the future price movements, and its earnings are a function of the volume traded since it charges the spread.

In order to maintain competitiveness with respect to standard exchanges and other liquidity providers, we assume that the CLP base spread does not exceed the standard exchange spread, i.e. the market spread, under normal circumstances (i.e. acceptable inventory levels). Nonetheless, it may be wider than the "lit"-exchange spread on the protect side (e.g. the sell side when the CLP has a negative inventory), while we assume that the aggress side (e.g. the buy side when the CLP has a positive inventory) shall not cross the mid-price level. This last assumption is justified since crossing the mid-price provides buy-side agents with very sensitive information regarding the CLP level of inventory.

We thus suppose that δ^+_u, δ^-_u are predictable processes valued in $[0, \bar{\delta}]$, where the upper bound is justified by practical considerations (N.B. an infinite spread results in an infinite price), as well as by the incentive for the CLP to maintain its quotes within a constant range.

The CLP flow of client orders can be affected by changing δ^+_u and δ^-_u. We might assume for example that the CLP has a positive inventory at time $u \in [0, T]$. It can make trading more attractive to buyers rather than sellers, so to rebalance its inventory level. In particular, by lowering δ^+_u and increasing δ^-_u, the CLP encourages buyers to place orders while sellers are discouraged to do so.

We shall now define the evolution of the CLP's inventory and the respective variations in its cash holdings. Every time a client submits a buy order, the inventory should decrease by an amount equal

to the size of the order. At the same time, the cash should increase proportionally to the order size and the charged price. The opposite holds when a sell order is submitted. We thus model the inventory in Equation (2.2) by a pure jump process where the negative jumps account for incoming buy orders and positive jumps model incoming sell orders. The cash process, modelled by Equation (2.3), varies accordingly.

Throughout the market-making activity (i.e. between hedging times), we define the CLP inventory process $\{X_u\}_{0 \leq u \leq T}$ by

$$dX_u = dJ_u^- - dJ_u^+, \tag{2.2}$$

where $J_u^\pm := \sum_{i=1}^{N_u^\pm} q_i^\pm$ and where the Cox processes $\{N_u^\pm\}$ have intensities $\lambda_\delta^\pm = \lambda^\pm(\delta_u^\pm)$. The random variables $q_i^\pm$ are i.i.d. with support $\mathbb{Q} := \{0, q_1, \ldots, q_N\}$ where these model the size of the trades executed by the CLP. At any time $u \in [0, T]$, we have $\text{sign}[X_u] = \{-1, 0, 1\}$ where we include short-selling for the case $\text{sign}[X_u] = -1$. We model the CLP cash process $\{Y_u\}_{0 \leq u \leq T}$ by

$$dY_u = f\left(u, S_{u^-}, \delta_u^+, q_{u^-}^+\right) dJ_u^+ - f\left(u, S_{u^-}, \delta_u^-, q_{u^-}^-\right) dJ_u^-, \tag{2.3}$$

where $q_{u^-}^\pm$ is shorthand notation for $q_{N_{u^-}^\pm}^\pm$. The function f allows the CLP to offer a stream of prices related to the size of the client's order. In particular the CLP offers tighter spreads for smaller sizes and wider spreads for larger sizes. The function f allows also for various ways to calculate P&L. Equations (2.2) and (2.3) are strongly coupled. For example, the arrival of a seller at time u increases the inventory X_{u^-} by $q_{u^-}^-$ and reduces the cash amount Y_{u^-} by $f\left(u, S_{u^-}, \delta_u^-, q_{u^-}^-\right) q_{u^-}^-$. The analogous holds for the arrival of buyers. We assume that the CLP accepts clients' orders at time u (i.e. it keeps streaming dealable prices to clients) only if the inventory at u^- lies within $[-\bar{X}, \bar{X}]$, where $\bar{X} > 0$ so that, if either boundary is reached, the CLP may only trade in one direction. Such an assumption is supported by the following financial interpretation: the CLP is subject to regulatory constraints (e.g. internal risk-management) which make it hard to hold or short-sell an amount of positions bigger than a fixed authorised quantity.

It is realistic to assume that the CLP is risk-averse and thus, we shall assume that it can reduce its inventory by submitting orders in the "lit" pool. Each of such orders gives an impulse to the system by modifying the inventory and cash holdings of the CLP. In what follows, we explain what actions the CLP may take and how such decisions impact the system.

The CLP can resort to the standard exchange to liquidate (respectively refill) part of its inventory; we assume that it cannot post speculative orders. This means that at time $u \in [0, T]$ a buy order can be posted only if $X_u < 0$ while a sell order can be posted if $X_u > 0$. We also assume that—if the boundaries are reached—the CLP may only trade in one direction.

2.2. Limit orders strategy

The CLP can post a limit order by specifying a quantity η and a limit price $S \pm (k + \kappa)$, where κ is the optimal distance from the best price, at which it wants to buy or sell. We only consider immediate-or-cancel orders and we model their execution percentages by a Amihud and Mendelson (1980)-valued sequence of i.i.d. random variables z_i, of which cumulative distribution function heavily depends on the limit-price chosen by the CLP. In particular, if a limit order is posted at time τ_i^ℓ, for $i = 1, 2, \ldots$, then it impacts the inventory and the cash processes as follows:

$$X_{\tau_i^\ell} = \Gamma\left(\eta_i, X_{\tau_i^\ell -}, z_i\right), \qquad Y_{\tau_i^\ell} = \chi\left(\eta_i, Y_{\tau_i^\ell -}, z_i, S_{\tau_i^\ell -}, k_{\tau_i^\ell -}, \kappa_i\right). \tag{2.4}$$

Since the CLP cannot post speculative orders in the "lit" pool, it must hold $\left|\Gamma\left(\eta_i, X_{\tau_i^\ell -}, z_i\right)\right| \leq \left|X_{\tau_i^\ell -}\right|$. For example a limit buy order ($\eta > 0$), which if executed increases the inventory, can only be posted if the CLP holds a negative inventory, and vice versa. The cash process changes accordingly. We state

these assumptions rigorously in the Section 6. We let $\mathcal{T}_T$ be the set of stopping times not greater than T, and $\mathcal{N} := \left[\min\left(0, -X_{\tau_i^\ell-}\right), \max\left(0, -X_{\tau_i^\ell-}\right)\right]$ be the set of all admissible control actions. A limit-order strategy is a collection of stopping times and actions $L = (\tau_i^\ell, \eta_i, \kappa_i)_{i \geq 1} \in \mathcal{T}_T \times \mathcal{N} \times \mathbb{K}^\ell$, where the elements in $\mathbb{K}^\ell \subset \mathbb{K}$ reflect the price-granularity of the exchange.

2.3. Market orders strategy

Alternatively, the CLP can submit a market order, which (i) is more expensive and (ii) benefits of sure execution as it is matched with existing limit orders. A market order of size ξ_i posted at a time τ_i^m impacts the inventory and cash processes as follows:

$$X_{\tau_i^m} = \Lambda\left(\xi_i, X_{\tau_i^m-}\right), \quad Y_{\tau_i^m} = c\left(\xi_i, Y_{\tau_i^m-}, S_{\tau_i^m-}, k_{\tau_i^m-}\right), \tag{2.5}$$

where $\left|\Lambda\left(\xi_i, X_{\tau_i^m-}\right)\right| \leq \left|X_{\tau_i^m-}\right|$. A market-order strategy is a collection of stopping times and actions $M = (\tau_i^m, \xi_i)_{i \geq 1} \in \mathcal{T}_T \times \mathcal{X}$, where $\mathcal{X} := \left[\min\left(0, -X_{\tau_i^m-}\right), \max\left(0, -X_{\tau_i^m-}\right)\right]$.

The level of generality in the limit-order and the market-order impulses offers the flexibility to include various features. For example, there are different methods to compute the P&L. Furthermore, one may like to account for the fees paid to the exchange for using their services and for possible liquidity rebates for limit orders.

By solving the optimisation problem presented in the next section, the CLP finds—at each time— the critical levels of inventory for which it is optimal to start hedging by means of limit and market orders.

3. Optimisation problem and viscosity solution

We consider the problem of maximising expected terminal cash subject to a terminal penalty for holding a non-zero inventory by the terminal date. In defining the objective function, we are led by Guilbaud and Pham (2013). We define the value function by

$$V(t, x, y, s; k) = \sup_{D, L, M} \mathbb{E}\left[U(X_T, Y_T, S_T, k_T) + \int_t^T g(u, X_u) \, du - \sum_{t \leq \tau_i^m < T} \epsilon_m - \sum_{t \leq \tau_i^\ell < T} \epsilon_l\right], \tag{3.1}$$

where $D := (\delta_u^+, \delta_u^-)_{u \geq t}$, the function U is the utility derived from the cash and inventory holdings at time T, and g is a running penalty for the risk of holding the inventory. In the summations of Equation (3.1), we include the penalties ϵ_m and ϵ_ℓ for submitting market and limit orders in the standard exchange, where $\epsilon_m > \epsilon_l > 0$. Throughout the paper we have the vector of state variables $x := [x, y, s] \in \mathcal{O} \times \mathbb{R}_+ =: \mathcal{S}$ and we let $\mathcal{T}_T$ be the set of all stopping times less than T. Equation (3.1) satisfies the DPP (see Fleming & Soner, 2006). That is, for all $\tau \in \mathcal{T}_T$, we have

$$V(t, \boldsymbol{x}; k) = \sup_{D, L, M} \mathbb{E}\left[\int_t^\tau g(u, X_u) \, du - \sum_{t \leq \tau_i^m < \tau} \epsilon_m - \sum_{t \leq \tau_i^\ell < \tau} \epsilon_l + V(\tau, \boldsymbol{X}_\tau; k_\tau)\right]. \tag{3.2}$$

This is an optimal double-obstacle impulse control problem. We define the non-local operators $\mathcal{L}$ and $\mathcal{M}$, for limit and market orders respectively, by

$$\mathcal{L}V(t, \boldsymbol{x}; k) = \sup_{\eta \in \mathcal{N}, \kappa \in \mathbb{K}^\ell} \mathbb{E}^{(z)}\left[V(t, \Gamma(\eta, x, z), \chi(\eta, y, z, s, k, \kappa), s; k)\right] - \epsilon_l, \tag{3.3}$$

where the expectation is taken with respect to the random variable z, and

$$\mathcal{M}V(t, \boldsymbol{x}; k) = \sup_{\xi \in \mathcal{X}} V(t, \Lambda(\xi, x), c(\xi, y, s, k), s; k) - \epsilon_m. \tag{3.4}$$

cogent··mathematics

We introduce the operator $\bar{\mathcal{A}}$ defined by

$$
\begin{aligned}
\bar{\mathcal{A}}(t,\boldsymbol{x},k,p,\phi,\delta^+,\delta^-) = {}& p + \sum_{k'\neq k} r_{kk'}\big[\phi(t,\boldsymbol{x};k') - \phi(t,\boldsymbol{x};k)\big] \\
& + \lambda^m \mathbb{E}^{(\bar{k}^\pm)}\big[\big(\phi\big(t,x,y,s+\bar{k}^\pm;k\big) - \phi(t,\boldsymbol{x};k)\big)\big] \\
& + \lambda_\delta^+ \mathbb{E}^{(q^+)}\big[\big(\phi\big(t,x-q^+,y+f(t,s,\delta^+,q^+)q^+,s;k\big) - \phi\big(t,\boldsymbol{x};k\big)\big)\big]\mathbb{1}_{\{x\geq -\bar{X}\}} \\
& + \lambda_\delta^- \mathbb{E}^{(q^-)}\big[\big(\phi\big(t,x+q^-,y-f(t,s,\delta^-,q^-)q^-,s;k\big) - \phi\big(t,\boldsymbol{x};k\big)\big)\big]\mathbb{1}_{\{x\leq \bar{X}\}},
\end{aligned}
\tag{3.5}
$$

where the expectations are taken with respect to the random variables $\bar{k}^\pm$, q^+ and q^-, respectively. Furthermore, we set $\mathcal{A}(t,\boldsymbol{x},k,p,\phi) = \sup_{\delta^\pm\in[0,\bar{\delta}]} \bar{\mathcal{A}}(t,\boldsymbol{x},k,p,\phi,\delta^+,\delta^-)$. The value function $V(t,\boldsymbol{x};k)$ satisfies the QVI system

$$
\min\big\{-g(t,x) - \mathcal{A}(t,\boldsymbol{x},k,\partial_t V,V);(V-\mathcal{M}V)(t,\boldsymbol{x};k);(V-\mathcal{L}V)(t,\boldsymbol{x};k)\big\} = 0,
\tag{3.6}
$$

on $[0,T]\times\mathcal{S}\times\mathbb{K}$. Equation (3.6) can be interpreted as follows: if $V-\mathcal{M}V > 0$ and $V-\mathcal{L}V > 0$, then the value function cannot be improved by an impulse and thus no orders are submitted to the standard exchange. As soon as $V-\mathcal{M}V < 0$ or $V-\mathcal{L}V < 0$, the value function is set to $V-\mathcal{M}V = 0$ or $V-\mathcal{L}V = 0$ and an impulse takes place. In the event $V-\mathcal{M}V < 0$ and $V-\mathcal{L}V < 0$, the value function is set to $V - \max\{\mathcal{M}V,\mathcal{L}V\} = 0$.

We thus consider intervention times (τ_i^ℓ and τ_i^m) and impulses (η_i and ξ_i) by which the CLP can control the evolution of the state variables X_u and Y_u. For this purpose, we define the continuation region (CR), the limit orders impulse region (LI) and the market orders impulse region (MI) by

$$
\begin{aligned}
\text{CR:} &= \big\{(u,\boldsymbol{x},k)\in[0,T]\times\mathcal{S}\times\mathbb{K}: V > \mathcal{L}V \text{ and } V > \mathcal{M}V\big\}, \\
\text{LI:} &= \big\{(u,\boldsymbol{x},k)\in[0,T]\times\mathcal{S}\times\mathbb{K}: \mathcal{L}V = V \text{ and } \mathcal{L}V > \mathcal{M}V\big\}, \\
\text{MI:} &= \big\{(u,\boldsymbol{x},k)\in[0,T]\times\mathcal{S}\times\mathbb{K}: \mathcal{M}V = V \text{ and } \mathcal{M}V > \mathcal{L}V\big\}.
\end{aligned}
\tag{3.7}
$$

The system of QVIs introduced in Equation (3.6) is highly non-linear and somewhat similar to the one studied in Guilbaud and Pham (2013), although in the present model two distinct impulses can take place. Some dimension reduction is possible if the mid price is assumed to be a martingale (see e.g. Cartea & Jaimungal, 2013; Guilbaud & Pham, 2013, 2015), since the optimal strategy in feedback-form will only be a function of the inventory. We make this assumption also in the numerical section that follows, while keeping the general model presented in this section at a higher degree of generality.

A possible structure of the intensities function has been extensively studied in, e.g. Cartea et al. (2014). In particular, they set $\lambda_\delta^\pm = \bar{\lambda}\,e^{-\zeta\delta^\pm}$, for $\bar{\lambda}, \zeta > 0$. The exponential decay of the intensity model describes an increasing likelihood of receiving clients' orders for lower values of the base spread charged. With such a structure, first-order conditions for optimality can be applied to Equation (3.6) to find the optimal base spread to be charged.

While keeping the model general, in the numerical example we take a slightly different approach (more similar to Guilbaud & Pham 2013, 2015). We assume that the CLP can only display three different types of quotes, i.e. skewed quotes (up or down) and non-skewed quotes. Thanks to this choice, we can show the time and inventory levels for which skewing is optimal, as opposed to the case of having processes valued in a closed subset of $\mathbb{R}_+$.

In Section 4 we provide some explicit examples of the model and we find numerically the optimal strategy by means of the solving algorithm proposed in Guilbaud and Pham (2013, 2015).

The existence and uniqueness of the viscosity solution (to which the numerical scheme converges) are shown in the Section 6.

4. Numerical solutions

In this section we explore some explicit examples of the impulse-control problem presented in this work and we analyse in detail their properties. While it would be desirable to ground the analysis on CLPs' real data, we note that such sensitive information is strictly private and not shared outside firms.

Throughout this section, we assume that the mid-price is modelled by Equation (2.1) and that limit orders in the "lit" exchange cannot be partially filled. This is not a strong assumption as long as we consider unit-sized orders posted in the standard exchange. In fact, the majority of the times, posting unit-sized orders in a "lit" market is a means to avoid consistent price slippage. We assume that clients pay the "adjusted base spread" to the CLP, where in the latter the size of the client's order is taken into account. We thus assume that the inventory and the cash processes evolve according to

$$dX_u = dJ_u^- - dJ_u^+,\tag{4.1}$$

$$dY_u = \left(S_{u^-} + \delta_u^+(1+c)^{q^+}\right)dJ_u^+ - \left(S_{u^-} - \delta_u^-(1+c)^{q^-}\right)dJ_u^-,\tag{4.2}$$

where $0 < c < 1$ is a pre-determined fixed constant. The convexity of the non-linear function $(1+c)^{q^+}$ reflects the fact that for the CLP it is more expensive to hedge a larger order due to the price slippage in the exchange. Thus, the CLP will charge more for each unit of a larger order.

Next, we introduce the possibility of submitting unit-sized market and limit orders in the standard exchange. At each time t_i, the CLP checks whether it is more convenient to (i) execute trades in the CLP only, (ii) submit a market order such that

$$X_{t_i} = X_{t_i^-} + \xi_i, \quad Y_{t_i} = Y_{t_i^-} - \xi_i\left(S_{t_i} + \xi_i k_i\right),\tag{4.3}$$

where $\xi_i \in \mathcal{X} := \{-\mathbb{1}_{\{x>0\}}, \mathbb{1}_{\{x<0\}}\}$, or (iii) submit a limit order such that

$$X_{t_i} = X_{t_i^-} + \eta_i z_i, \quad Y_{t_i} = Y_{t_i^-} - \eta_i\left(S_{t_i} - (k_i + \kappa_i)\eta_i\right)z_i,\tag{4.4}$$

where $\eta_i \in \mathcal{N} = \{-\mathbb{1}_{\{x>0\}}, \mathbb{1}_{\{x<0\}}\}$ and z_i are i.i.d. random variables supported in $\{0, 1\}$.

We start by presenting a toy model in Section 4.1. Our intention is to look first at a reduced set of controls so to gain a better understanding of the behaviour of the optimal strategy that one should expect. Reducing the number of variables that play a role in the model and *progressively* removing the simplifying assumptions we make is, in our eyes, a good way to understand the structure of the model and the properties of the solution.

4.1. Toy model

Mathematical models designed for financial applications can sometimes be rather abstract and might have a degree of complexity that make real-world applications difficult. In the toy model that follows, we reduce the mathematical complexity of the model presented in Section 3 considerably in order to allow for a good understanding of the modelling approach we propose for the optimal inventory management in CLPs. We first lead through the basic features of the toy model and then, step by step, add structure to the toy model in Sections 4.2, 4.3 and 4.4 in order to capture features of real-world CLPs. This way, the more realistic model will be based on a solid understanding gained by analysing the more basic toy model.

For the time being, we suppose for the toy model that

(a) the CLP "base quotes" $\delta_u^+ = \delta^+$ and $\delta_u^- = \delta^-$ are constant and fixed *a priori*,

(b) limit orders can only be posted at the best bid-ask prices (i.e. $\kappa = 0$),

(c) $g(u, x) = 0$ and $U(x, y, s, k) = y + x(s - \alpha x)$,

(d) the bid-ask half spread is constant and set to k.

Assumptions (a) and (b) reduce the set of controls since here it is assumed that the CLP does not skew its prices and that it cannot post limit orders deep in the LOB. In assumption (c), $g(u, x) = 0$ means that the CLP is risk-neutral and hence does not consider the inventory size when making decisions. Moreover, given the choice of the function U, it seeks to maximise the terminal cash amount subject to a penalty, modelled by the parameter $\alpha \geq 0$, for holding a non-zero inventory at the terminal date T. Assumption (d) further reduces the dimensionality of the optimisation problem as we now deal with a single QVI rather than a system of QVIs. While we consider this model as rather unrealistic given the restrictive assumptions, it still gives a first insight into the type of optimal policy, see Figures 1 and 2. The QVI associated with the simplified control problem then consists of a PIDE and two impulse parts—the CLP market-making activity, the market orders and the limit orders, respectively. We have:

$$
\min \Big\{ -\partial_t V - \lambda^m \mathbb{E}^{(\bar{k}^{\pm})} \big[V(t, x, y, s + \bar{k}^{\pm}; k) - V(t, \mathbf{x}; k) \big]
$$

$$
- \lambda^+ \mathbb{E}^{(q^+)} \Big[V(t, x - q^+, y + (s + \delta^+(1 + c)^{q^+})q^+, s) - V(t, x, y, s) \Big] \mathbb{1}_{\{x \geq -\bar{X}\}}
$$

$$
- \lambda^- \mathbb{E}^{(q^-)} \Big[V(t, x + q^-, y - (s - \delta^-(1 + c)^{q^-})q^-, s) - V(t, x, y, s) \Big] \mathbb{1}_{\{x \leq \bar{X}\}};
$$

$$
V(t, x, y, s) - \sup_{\xi \in \mathcal{X}} \Big[V\big(t, x + \xi, y - \xi(s + \xi k), s\big) - \epsilon_m \Big];
$$

$$
V(t, x, y, s) - \sup_{\eta \in \mathcal{N}} \mathbb{E}^{(z)} \Big[V\big(t, x + \eta z, y - z\eta(s - \eta k), s\big) - \epsilon_l \Big] \Big\} = 0.
$$

(4.5)

Given the form of the terminal condition, we consider the ansatz $V(t, x, y, s) = y + xs + h(t, x)$. Under analogous conditions, Cartea and Jaimungal (2013) and Cartea, Jaimungal, and Ricci (2014) make use of the same ansatz while reducing the problem dimensions. By substituting the ansatz in (4.5) we obtain the QVI for $h(t, x)$:

$$
\min \Big\{ -\partial_t h(t, x) - \lambda^+ \Big(\mathbb{E}^{(q^+)} \big[\delta^+(1 + c)^{q^+})q^+ \big] + h(t, x - q^+) - h(t, x) \Big) \mathbb{1}_{\{x \geq -\bar{X}\}}
$$

$$
- \lambda^- \big(\mathbb{E}^{(q^-)} \big[\delta^-(1 + c)^{q^-})q^- \big] + h(t, x + q^-) - h(t, x) \big) \mathbb{1}_{\{x \leq \bar{X}\}} - \lambda^m x \mathbb{E}^{(\bar{k}^{\pm})} \big[\bar{k}^{\pm} \big];
$$

$$
h(t, x) - \sup_{\xi \in \mathcal{X}} \big[-k - \epsilon_m + h(t, x + \xi) \big]; h(t, x) - \sup_{\eta \in \mathcal{N}} \mathbb{E}^{(z)} \big[kz - \epsilon_l + h(t, x + \eta z) \big] \Big\} = 0.
$$

(4.6)

Figure 1. Optimal boundaries.
We set $\delta^{\pm} = 0.5$, $\lambda^{\pm} = 0.5$, $k = 2$, $\alpha = 2$, $p = 0.9$, $\epsilon_m = 9$, $\epsilon_{\ell} = 3$.
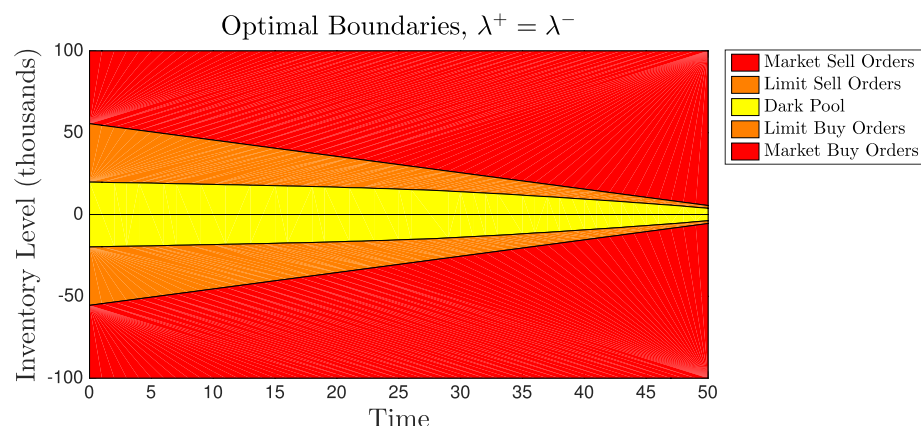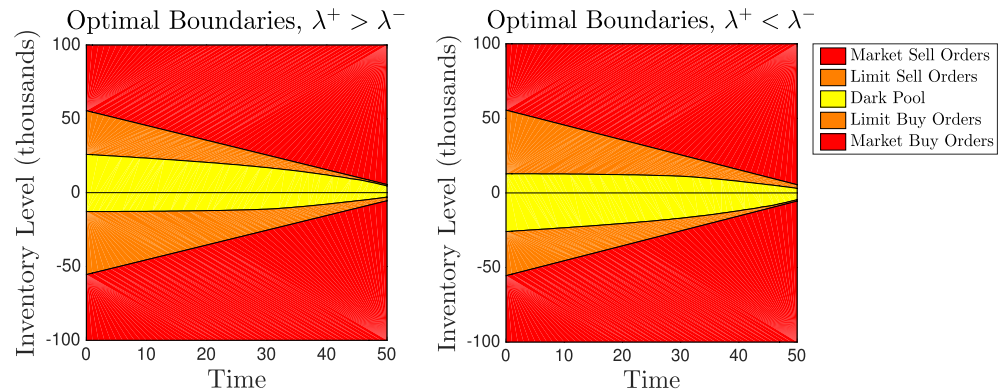
Figure 2. Optimal boundary for the double-obstacle problem. Most of the parameters are as in Figure 1. In the left panel we set $\lambda^+ = 0.4$ and $\lambda^- = 0.6$. In the right panel we set $\lambda^+ = 0.6$ and $\lambda^- = 0.4$.

In Figure 1 we show the numerical results of Equation (4.6) for $\lambda^+ = \lambda^-$. The yellow region corresponds to the market-making region (continuation region). The CLP provides liquidity until either one of the boundaries is surpassed. When the orange region is reached, the CLP starts hedging by submitting limit orders to the "lit" pool (sell orders if the inventory is in the top orange region and buy orders otherwise). If the inventory level lies in the red regions, the optimal strategy is to cross the spread and post aggressive orders in the "lit" exchange.

The symmetry around the zero inventory level in Figure 1 is due to the particular choice made for the arrival flow of buyers and sellers, i.e. $\lambda^+ = \lambda^-$. In Figure 2 we lose such a symmetry and see a shift upwards (resp. downwards) of the optimal boundaries when the arrival intensity of sellers λ^- decreases (resp. the arrival intensity of buyers λ^+ decreases).

Before relaxing some of the assumptions made in this section, we briefly comment on the role played by the other parameters of the model. Higher values of the terminal penalty for holding the inventory α produce a shrinkage of the continuation and limit-order regions plotted in Figure 1 while higher values of the penalties ϵ_m and ϵ_ℓ widen the aforementioned regions. A higher value of the CLP base spread $(\delta^+ + \delta^-)$ widens the continuation region while higher values of p (the probability of limit-order execution) increase the limit orders thresholds. Ultimately, the optimal strategy is a trade-off between the earnings derived from the market-making activity, the penalty for holding the inventory and the hedging costs sustained to reduce the position on the asset. The hedging part of the model (i.e. the two impulses) is predominant when dealing with a risk-averse CLP, which is shown in Section 4.2.

4.2. Optimal CLP base spread

The model presented in Section 4.1 is limited in that, e.g. it does not allow the CLP to skew its prices to control the incoming flows of buyers and sellers. We remove this assumption here and we let the CLP choose between three possible scenarios. (i) It can choose not to skew the prices and thus to let $\delta^+ = \delta^-$ and $\lambda^+ = \lambda^-$ (which we call "no skew"), (ii) it may skew the prices downward such that $\delta^+ < \delta^-$ and $\lambda^+ > \lambda^-$ (which we call "left" or "downward" skew), and (iii) it may skew the prices upward such that $\delta^+ > \delta^-$ and $\lambda^+ < \lambda^-$ (which we call "right" or "upwar" skew). The upward skew penalises buyers over sellers while the opposite holds for the downward skew. Mathematically, the above reduces to replacing assumption (a) with

(a') the CLP optimally chooses the prices $(\delta^+, \delta^-) \in \mathcal{D}:= \left\{(\delta_n^+, \delta_n^-), (\delta_r^+, \delta_r^-), (\delta_l^+, \delta_l^-)\right\}$ for the no skew, the right skew and the left skew scenarios, respectively. Associated to such prices, we assume arrival intensities of the form $(\lambda^+(\delta^+), \lambda^-(\delta^-)) \in \mathcal{L}:= \left\{(\lambda_n^+(\delta_n^+), \lambda_n^-(\delta_n^-)), (\lambda_r^+(\delta_r^+), \lambda_r^-(\delta_r^-)), (\lambda_l^+(\delta_l^+), \lambda_l^-(\delta_l^-))\right\}.$

We further relax assumption (c) and substitute it with

(c') $g(u, x) = -\bar{\phi}x^2$, where $\bar{\phi} \geq 0$. That is, the CLP is subject to a quadratic running penalty for holding a non-zero inventory. This manifests the CLP's risk-adversity since a large inventory reduces the value function of the pool. In Cartea et al. (2014) it is shown that such a choice may also be linked to both, the variance of the portfolio and the model's ambiguity.

In what follows, we indicate by $\lambda^{\pm}(\delta^{\pm})$ the dependence of the arrival flows of buyers and sellers on the prices offered by the CLP. As mentioned before, a higher price is reflected in a lower arrival of the respective side of clients (i.e. $\delta^+ < \delta^- \Rightarrow \lambda^+ > \lambda^-$ and $\delta^+ > \delta^- \Rightarrow \lambda^+ < \lambda^-$). The QVI now reads:

$$
\min \left\{ \bar{\phi}x^2 - \sup_{(\delta^+, \delta^-) \in \mathcal{D}} \left[\lambda^+(\delta^+) \left(\mathbb{E}^{(q^+)} \left[\delta^+(1+c)^{q^+})q^+ \right] + h(t, x-q^+) - h(t,x) \right) \mathbb{1}_{\{x \geq -\bar{X}\}} \right. \right.
$$
$$
\left. + \lambda^-(\delta^-) \left(\mathbb{E}^{(q^-)} \left[\delta^-(1+c)^{q^-})q^- \right] + h(t, x+q^-) - h(t,x) \right) \mathbb{1}_{\{x \leq \bar{X}\}} \right]
$$
$$
- \partial_t h(t,x) - \lambda^m x \mathbb{E}^{(\bar{k}^{\pm})} \left[\bar{k}^{\pm} \right]; h(t,x) - \sup_{\xi \in \mathcal{X}} \left[-k - \epsilon_m + h(t, x+\xi) \right];
$$
$$
\left. h(t,x) - \sup_{\eta \in \mathcal{N}} \mathbb{E}^{(z)} \left[kz - \epsilon_l + h(t, x+\eta z) \right] \right\} = 0.
$$

(4.7)

As before, we numerically solve (4.7) to find the optimal CLP pricing and hedging strategy. We use an explicit finite-difference scheme to solve Equation (4.7) backward in time. At each time step, we plug in the three alternatives (namely $\delta_n^{\pm}, \delta_r^{\pm}, \delta_l^{\pm}$) and store the combination which maximises the value function V. Such an optimal combination depends, at each time, on the inventory level of the CLP and the time-to-maturity. We thus expect to find additional boundaries to the one obtained in Section 4.1 since we do now have different pricing options in the CLP.

In Figure 3 we see that when the inventory is relatively small, it is optimal not to skew the prices so to receive on average an equal number of buy and sell orders (here we assume that no-skewing implies equal arrival intensities of buyers and sellers, which may not be the case in particular market conditions, e.g. new information is available to a number of clients which are incentivised to trade in the same direction). If the inventory increases (resp. decreases), the CLP employs a left/downward (resp. right/upward) skew so to encourage buyers (resp. sellers). As in the toy model, there are inventory levels for which it is optimal to resort to the standard exchange. The critical inventory level at which the CLP begins placing orders in the standard exchange falls as the terminal liquidation date is approached. We plot the optimal boundaries for the case of moderate risk aversion (right panel) and high risk aversion (right panel). We notice that the hedging activity of the CLP is highly correlated to its degree of risk aversion. Indeed, for high values of $\bar{\phi}$, the CLP starts posting orders to the standard exchange for smaller inventory values and vice versa.

Figure 3. Optimal boundary for the double-obstacle problem. We set $\mathcal{D}:= \{(0.5, 0.5), (0.3, 0.6), (0.6, 0.3)\}$ and $\mathcal{L}:= \{(0.5, 0.5), (0.6, 0.4), (0.4, 0.6)\}$. **In the left panel we set** $\bar{\phi} = 0.01$. **In the right panel we set** $\bar{\phi} = 0.0001$.
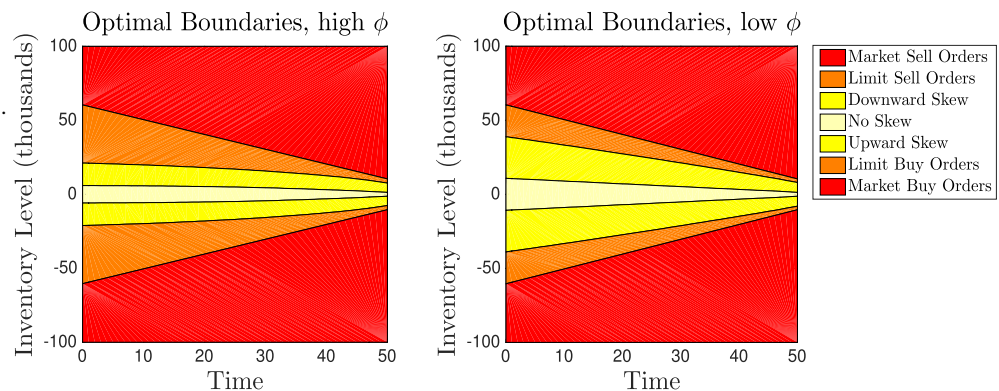
4.3. Optimal posting to a standard exchange

In the previous section we have included the option of skewing the prices that the CLP offers, but the model is still limited in that the hedging in the standard exchange by means of limit orders can only be done at the top of the book. Here we modify this assumption and we allow the CLP to also post at the second best and third best prices.

In this section we still keep the "lit" exchange half-spread k constant. This assumption will be removed in the next and last numerical example. According to our LOB model, we defined the minimum price tick by $2\bar{k}$, and thus allow the CLP to post limit sell orders at prices $s + k, s + k + 2\bar{k}$ and $s + k + 4\bar{k}$, while limit buy orders can be posted at prices $s - k, s - k - 2\bar{k}$ and $s - k - 4\bar{k}$. According to the notation used in Equation (4.4), we modify assumption (b) and replace it with

(b') the limit price at which the CLP posts in the standard exchange can be optimally chosen between $\kappa \in \mathbb{K}^\ell := \{0, 2\bar{k}, 4\bar{k}\}$.

By posting deeper in the book, the CLP earns a higher spread if its order gets executed, while the filling probability of such order is reduced. In fact, for a deep limit order to be executed a market order that walks the book is needed, and the latter are quite rare. To reflect the fact that the filling-probability of a limit order depends on how far from the mid-price such an order is posted, we assume that $\mathbb{P}[z_i = 0] = \ell^\kappa(z_0)$ and $\mathbb{P}[z_i = 1] = \ell^\kappa(z_1) = 1 - \ell^\kappa(z_0)$. The associated QVI for the function $h(t, x)$ is now

$$\min\left\{ \bar{\phi}x^2 - \sup_{(\delta^+, \delta^-)\in D}\left[\lambda^+(\delta^+)\Big(\mathbb{E}^{(q^+)}\big[\delta^+(1+c)^{q^+})q^+\big] + h(t, x - q^+) - h(t, x)\Big)\mathbb{1}_{\{x \geq -\bar{x}\}}\right.\right.$$
$$+ \lambda^-(\delta^-)\Big(\mathbb{E}^{(q^-)}\big[\delta^-(1+c)^{q^-})q^-\big] + h(t, x + q^-) - h(t, x)\Big)\mathbb{1}_{\{x \leq \bar{x}\}}\Big]$$
$$\left. - \partial_t h(t, x) - \lambda^m x \mathbb{E}^{(\bar{k}^\pm)}\big[\bar{k}^\pm\big]; h(t, x) - \sup_{\xi \in \mathcal{X}}\big[-k - \epsilon_m + h(t, x + \xi)\big]; \right. \tag{4.8}$$

$$\left. h(t, x) - \sup_{\eta \in \mathcal{N}, \kappa \in \mathbb{K}^\ell} \mathbb{E}^{(z)}\big[(k + \kappa)z - \epsilon_l + h(t, x + \eta z)\big] \right\} = 0.$$

with terminal condition $h(T, x) = -\alpha x^2$.

In Figure 4, the optimal strategy obtained by solving (4.8) is shown. We notice that after skewing the prices, the CLP should start submitting limit orders deep in the book and progressively moves towards the top of the book. Shortly before the end of the trading period, it should resort to market orders. Again, we show the different strategies employed by a highly risk-averse and a moderately risk-averse CLP (in the left and right panels, respectively). In Figure 5 we show the different strategies for high and low values of the terminal-penalty parameter α. When the penalty for holding a non-zero inventory at T increases, we notice that the boundaries shrink dramatically, especially towards the end.

Figure 4. Optimal boundary for the double-obstacle problem. We set $\delta = 1$ and $\ell^{\kappa=0}(z_1) = 0.9$, $\ell^{\kappa=1}(z_1) = 0.8, \ell^{\kappa=2}(z_1) = 0.6$. In the left panel we set $\bar{\phi} = 0.01$. In the right panel we set $\bar{\phi} = 0.0001$.
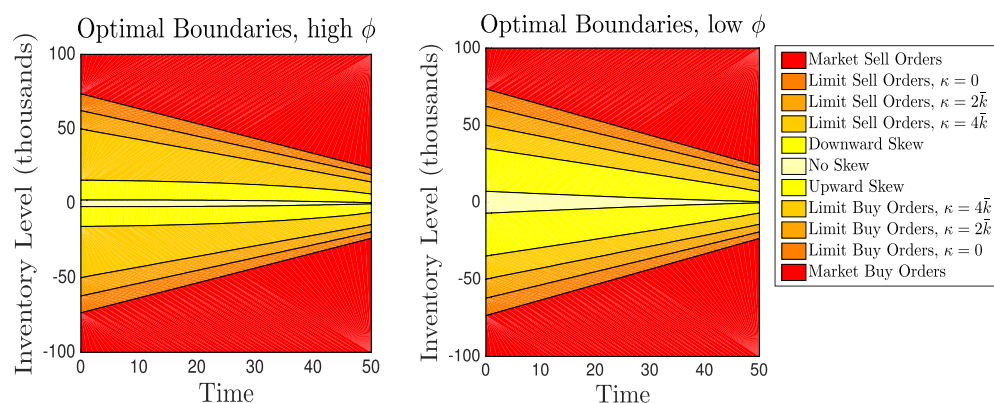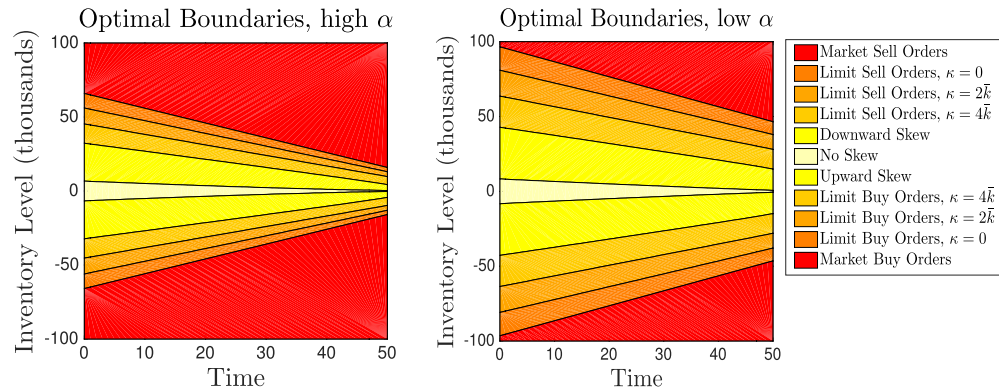
Figure 5. Optimal boundary for the double-obstacle problem. In the left panel we set $\alpha = 6$. In the right panel we set $\alpha = 0.5$.

We emphasise that if the terminal preferred inventory level were non-zero, it would produce a shift in the optimal boundaries by an equal amount.

4.4. Stochastic standard exchange spread

In the last simulation we relax assumption (d) and introduce a stochastic bid-ask spread. We let $\mathbb{K} := \{1, 2\}$, that is the market can be in a tight-spread regime and a wide-spread regime, depending on whether there is good or poor liquidity in the market, respectively. The generator matrix can be chosen in various ways, so to model the specific features of the market under consideration. For example, we could consider a "seasonal" pattern where transitions between regimes happen rarely and last for longer periods of time. Also, we could reproduce features similar to mean-reversion by choosing a "preferred" state and make any deviation from that state very unlikely while reversion to that state more likely. The matrices Q_1 and Q_2 are examples of seasonal and mean-reverting patterns, respectively:

$$Q_1 = \begin{pmatrix} -r & r \\ r & -r \end{pmatrix}, \quad Q_2 = \begin{pmatrix} -r_{low} & r_{low} \\ r_{high} & -r_{high} \end{pmatrix}, \tag{4.9}$$

where $r, r_{low}, r_{high} > 0$. In Q_1 the lower r, the rarer the transition between states happens. In Q_2, we choose $k = 1$ as the preferred state, and the higher r_{high}, compared to r_{low}, the higher is the transition rate to state 1.

The value function satisfies now a two-dimensional system of QVIs. We consider the ansatz $V(t, x, y, s; k) = y + xs + h_k(t, x)$, where the subscript k indicates that we refer to the regime $k \in \mathbb{K}$, and the function h satisfies the following system of QVIs:

$$\min\Bigg\{ - \sup_{(\delta^+, \delta^-) \in D} \Big[\lambda^+(\delta^+) \Big(\mathbb{E}^{(q^+)} \Big[\delta^+ (1 + c)^{q^+} \Big) q^+ \Big] + h_k(t, x - q^+) - h_k(t, x) \Big) \mathbb{1}_{\{x \geq -\bar{X}\}}$$

$$+ \lambda^-(\delta^-) \Big(\mathbb{E}^{(q^-)} \Big[\delta^- (1 + c)^{q^-} \Big) q^- \Big] + h_k(t, x + q^-) - h_k(t, x) \Big) \mathbb{1}_{\{x \leq \bar{X}\}} \Big]$$

$$\bar{\phi} x^2 - \partial_t h_k(t, x) - \lambda^m x \mathbb{E}^{(\bar{k}^\pm)} \big[\bar{k}^\pm \big] - \sum_{k' \neq k} r_{kk'} \big[h_{k'}(t, x) - h_k(t, x) \big];$$

$$h_k(t, x) - \sup_{\xi \in \mathcal{X}} \big[-k - \epsilon_m + h_k(t, x + \xi) \big];$$

$$h_k(t, x) - \sup_{\eta \in \mathcal{N}, \kappa \in \mathbb{K}^\ell} \mathbb{E}^{(z)} \big[(k + \kappa) z - \epsilon_l + h_k(t, x + \eta z) \big] \Bigg\} = 0. \tag{4.10}$$

In Figures 6 and 7 we plot the optimal strategy found by solving (4.10) for the case of seasonal and mean-reverting patterns, respectively. We first note that in the case of high-spread regime in Figure 6, the CLP starts (i) earlier to submit limit orders and (ii) later to submit market orders, compared to the case of low-spread regime. The same shape is observable in Figure 7, although it is more evident

cogent ·· mathematics

Figure 6. Optimal boundary for the double-obstacle problem. Seasonal pattern.
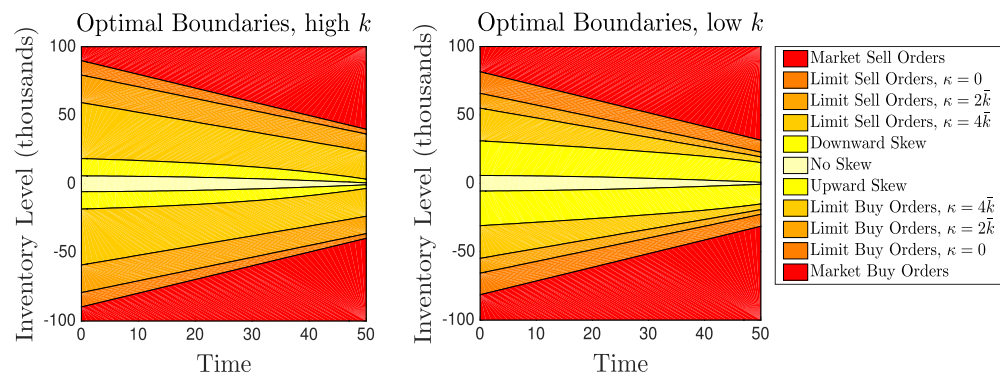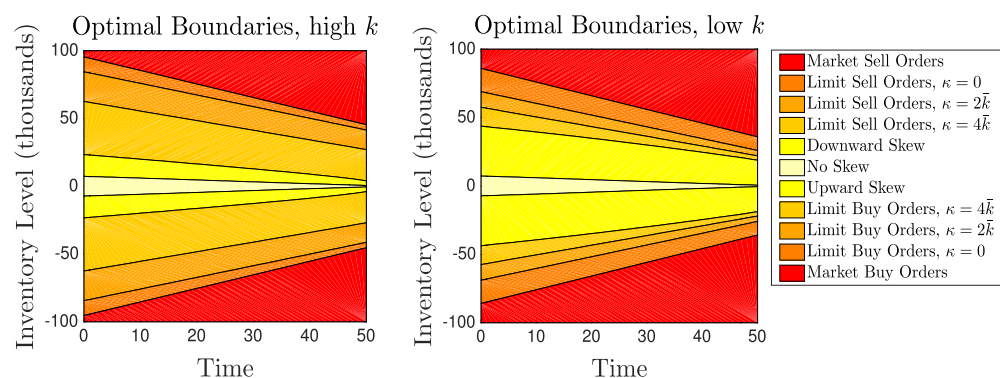

Figure 7. Optimal boundary for the double-obstacle problem. Mean-reverting pattern.

since the return to the state of low-spread regime (preferred state) is highly likely. This behaviour has the following financial interpretation: when the spread is high, limit orders are more remunerative (or, better, cheaper if we also consider the penalty for posting to the standard exchange), while market orders are more expensive and thus their submission is postponed.

5. Concluding remarks

In the present work we study an optimal market-making problem faced by a CLP. The CLP earns the optimally selected spread by trading with its clients. Market participants, who consider trading via CLPs, may be of the view that it is desirable to take advantage of favourable prices offered by the CLPs and to benefit from avoiding price impact—to which they would be otherwise exposed—especially if forced to submit market orders in a "lit" exchange.

A stream of two-way prices is offered to each client. Such prices are functions of the size traded by the client and the CLP holding. Throughout the activity the pool faces an inventory risk, which can be reduced (i) by controlling the width and the skew of the CLP spread, and (ii) by resorting to the standard exchange via both, market and limit orders. Internal CLP transactions are preferred since it protects from information leakage. Such a feature is modelled via a fixed penalty incurred by the CLP whenever it submits an order to the "lit" exchange.

As confirmed by the numerical results, the CLP will refrain from placing orders in a standard exchange as long as the size of the inventory is small. Whenever the optimal boundary is exceeded, the CLP resorts to the standard exchange by means of limit orders. A limit order is cheaper though its execution is uncertain. The CLP can choose the limit price; we find that the more the inventory grows, the closer to the mid-price the CLP will post. This is reasonable since the filling-intensity of limit orders depends on how far from the mid-price they are posted. If the inventory becomes critically large, market orders will be preferred instead, which are costly but benefit of sure execution. When the end of the market-making activity approaches (which, e.g. might be thought of as the end of the

cogent · mathematics

trading day) the market-orders region in the "lit" exchange widens while the CLP and limit-order regions in the standard exchange diminish. In fact, the market maker will incur in a higher penalty for holding a large inventory at the terminal date.

These conclusions are obtained by formulating and numerically solving a double obstacle standard stochastic and impulse control problem, for which we provide four numerical examples with increasing complexity. Compared to the state-of-the-art literature on "lit" pool market making (see e.g. Cartea et al., 2014; Guéant et al., 2013; Guilbaud & Pham, 2013, 2015), adding a CLP to the model substantially modifies the "standard" "lit" exchange market making problem for the following reasons: (i) the prices offered to clients can be functions of the size traded (which would not be possible in a classic LOB), and (ii) there are no such things as minimum tick size, minimum quantities or queues in the CLP that may be assumed, and (iii) the standard exchange is only utilised as a hedging venue, while two-way prices are offered to clients and are not available to general market participants. This work provides a rather flexible setup for the management of a CLP inventory. The pricing and the hedging strategies are both illustrated in detail.

6. Appendix

6.1. Viscosity solutions
We state all the standing assumptions which are introduced in the modelling setup presented in this paper.

(i) $f:[0, T] \times \mathbb{R}_+ \cup \{0\} \times [0, \bar{\delta}] \times \mathbb{Q} \to \mathbb{R}_+ \cup \{0\}$ satisfies Lipschitz continuity and the linear growth conditions

$$\left| f\left(t_1, s_1, d, q\right) - f\left(t_2, s_2, d, q\right) \right|^2 \le C \left(|t_1 - t_2|^2 + |x_1 - x_2|^2 + |s_1 - s_2|^2 \right)$$

$$\left| f\left(t_1, s_1, d, q\right) \right|^2 \le C \left(1 + |x_1|^2 + |s_1|^2 \right)$$

(6.1)

for all $t_1, t_2 \in [0, T], s_1, s_2 \in \mathbb{R} \cup \{0\}, q \in \mathbb{Q}$, and $d \in \mathbb{K}$;

(ii) $\Gamma:\mathcal{N} \times \left[-\bar{X} - q_N, \bar{X} + q_N\right] \times [0, 1] \to \mathbb{R}$ and $\chi:\mathcal{N} \times \mathbb{R} \times [0, 1] \times \mathbb{R} \times \mathbb{K} \times \mathbb{K}^\ell \to \mathbb{R}$ are Lipschitz continuous functions satisfying

$$|x|^2 - \mathbb{E}\left[|\Gamma(\eta, x, z)|^2\right] > 1, \quad |y|^2 - \mathbb{E}\left[|\chi(\eta, y, z, s, k, \kappa)|^2\right] > 1,$$

for all $\eta \in \mathcal{N}, z \in [0, 1]$ and $x, y, s \in \mathcal{S}$.

(iii) $\Lambda:\mathcal{X} \times \left[-\bar{X} - q_N, \bar{X} + q_N\right] \to \mathbb{R}$ and $c:\mathcal{X} \times \mathbb{R}^2 \times \mathbb{K} \to \mathbb{R}$ are Lipschitz continuous functions satisfying, for $M > 0$, the following properties:

$$|x|^2 - |\Lambda(\xi, x)|^2 > 1, \quad |y|^2 - |c(\xi, y, s, k)|^2 > 1$$

for all $\xi \in \mathcal{X}$ and $x, y, s \in \mathcal{S}$.

(iv) $g:[0, T] \times \left[-\bar{X} - q_N, \bar{X} + q_N\right] \to \mathbb{R}$ satisfies Lipschitz continuity

$$\left| g\left(t_1, x_1\right) - g\left(t_2, x_2\right) \right|^2 \le C \left(|t_1 - t_2|^2 + |x_1 - x_2|^2 \right)$$

(6.2)

for all $t_1, t_2 \in [0, T]$ and $x_1, x_2 \in \left[-\bar{X} - q_N, \bar{X} + q_N\right]$;

(v) $U:\mathcal{S} \times \mathbb{K} \to \mathbb{R}$ satisfies Lipschitz continuity and the linear growth conditions

$$\left| U(\boldsymbol{x}_1, k) - U(\boldsymbol{x}_2, k) \right|^2 \leq C \left(\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2 \right)$$

$$\left| U(\boldsymbol{x}_1, k) \right|^2 \leq C \left(1 + \|\boldsymbol{x}_1\|_2^2 \right), \tag{6.3}$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$ and $k \in \mathbb{K}$.

For each $k \in \mathbb{K}$, let us consider the upper and lower semi-continuous envelopes of the function $V(\cdot; k)$ defined by

$$V^*(t, \boldsymbol{x}; k) = \limsup_{\substack{t' \to t \\ \boldsymbol{x}' \to \boldsymbol{x}}} V(t', \boldsymbol{x}'; k), \quad V_*(t, \boldsymbol{x}; k) = \liminf_{\substack{t' \to t \\ \boldsymbol{x}' \to \boldsymbol{x}}} V(t', \boldsymbol{x}'; k). \tag{6.4}$$

Definition 6.1   A system of functions $V: [0, T) \times S \times \mathbb{K} \to \mathbb{R}$ is a viscosity subsolution, (resp. supersolution), of (3.6) if

$$\min \left\{ -g(\bar{t}, \bar{x}) - \mathcal{A}(\bar{t}, \bar{\boldsymbol{x}}, \bar{k}, \partial_t \phi, \phi); (V^* - \mathcal{M}V^*)(\bar{t}, \bar{\boldsymbol{x}}; \bar{k}); (V^* - \mathcal{L}V^*)(\bar{t}, \bar{\boldsymbol{x}}; \bar{k}) \right\} \leq 0, \tag{6.5}$$

$$(resp. \ \min \left\{ -g(\bar{t}, \bar{x}) - \mathcal{A}(\bar{t}, \bar{\boldsymbol{x}}, \bar{k}, \partial_t \phi, \phi); (V_* - \mathcal{M}V_*)(\bar{t}, \bar{\boldsymbol{x}}; \bar{k}); (V_* - \mathcal{L}V_*)(\bar{t}, \bar{\boldsymbol{x}}; \bar{k}) \right\} \geq 0), \tag{6.6}$$

where $\phi \in C^{1,0,0}([0, T) \times S \times \mathbb{K})$ is such that $V^*(t, \boldsymbol{x}, k) - \phi(t, \boldsymbol{x}; k)$ (resp. $V_*(t, \boldsymbol{x}, k) - \phi(t, \boldsymbol{x}; k)$) attains its maximum (resp. minimum) at $(\bar{t}, \bar{\boldsymbol{x}}, \bar{k}) \subset [t, T) \times S \times \mathbb{K}$.

Proposition 6.2   (Existence)   The system of functions $V(t, \boldsymbol{x}; k)$ is a viscosity solution of the QVI (3.6).

Proof   We use Definition 6.1 and we show that the system of functions $V(t, \boldsymbol{x}; k)$ is a viscosity solution by proving that it is both a supersolution and a subsolution.

We note that we have $V(T, \boldsymbol{x}; k) = U(\boldsymbol{x}; k)$ on $\{T\} \times S \times \mathbb{K}$, thus we need to prove the viscosity property only on $[t, T) \times S \times \mathbb{K}$. Results in, e.g. Vath, Mnif, and Pham ([2007](#)) ensure that $\mathcal{M}V_* \leq (\mathcal{M}V)_*$ and $\mathcal{L}V_* \leq (\mathcal{L}V)_*$. By definition of the value function, we have $V \geq \mathcal{M}V$ and $V \geq \mathcal{L}V$ for all $(u, \boldsymbol{x}) \in [t, T) \times S$. It follows that $V_* \geq (\mathcal{M}V)_* \geq \mathcal{M}V_*$ and $V_* \geq (\mathcal{L}V)_* \geq \mathcal{L}V_*$. That is, to prove the supersolution property, it suffices to show that

$$-g(\bar{t}, \bar{x}) - \mathcal{A}(\bar{t}, \bar{\boldsymbol{x}}, \bar{k}, \partial_t \phi, \phi) \geq 0. \tag{6.7}$$

Let $(V_* - \phi)(\bar{t}, \bar{\boldsymbol{x}}; \bar{k}) = 0$, where $(\bar{t}, \bar{\boldsymbol{x}}, \bar{k}) = \arg\min(V_* - \phi)(t, \boldsymbol{x}; k)$. By definition of V_*, there exists a sequence $(t_m, \boldsymbol{x}_m) \to (\bar{t}, \bar{\boldsymbol{x}})$ such that $V_*(t_m, \boldsymbol{x}_m; \bar{k}) \to V_*(\bar{t}, \bar{\boldsymbol{x}}; \bar{k})$ as $m \to \infty$. We define the stopping time

$$\theta_m = \inf \left\{ u > t_m \mid \boldsymbol{X}_{t_m \boldsymbol{x}_m}(u) \notin B_\eta(t_m, \boldsymbol{x}_m) \right\}, \tag{6.8}$$

where $B_\eta(t_m, \boldsymbol{x}_m)$ is the open ball of radius η centred in $(t_m, \boldsymbol{x}_m)$. We choose a strictly positive sequence $h_m \to 0$ and let the stopping time $\theta_m^* := \theta_m \wedge (t_m + h_m) \wedge \theta^* \wedge \tau_i$, where θ^* is the first time the regime switches from its initial value $\bar{k}$ and where τ_i is the first time an impulse takes place. By the dynamic programming principle and the definition of the function ϕ, we have for any admissible control strategy

$$V(t_m, \boldsymbol{x}_m; \bar{k}) \geq \mathbb{E}\left[\int_{t_m}^{\theta_m^*} g\left(u, X_{t_m, x_m}(u) \right) du + \phi\left(\theta_m^*, \boldsymbol{X}_{t_m, \boldsymbol{x}_m}(\theta_m^*); k_{\theta_m^*} \right) \right]. \tag{6.9}$$

An application of Itò's formula to ϕ between t_m and θ_m^* yields

$$(V_* - \phi)(t_m, \boldsymbol{x}_m; \bar{k}) \geq \mathbb{E}\left[\int_{t_m}^{\theta_m^*} \left[g\left(u, X_{t_m, x_m}(u) \right) + \bar{\mathcal{A}}\left(u, \boldsymbol{X}_{t_m, \boldsymbol{x}_m}(u), \bar{k}, \partial_t \phi, \phi, \delta_u^+, \delta_u^- \right) \right] du \right], \tag{6.10}$$

We can divide by $-h_m$, then let $m \to \infty$ and apply the the mean value theorem. Finally, the result follows from the arbitrariness of the control variable.

We observe that if $V^* \leq \mathcal{M}V^*$ or $V^* \leq \mathcal{L}V^*$, the subsolution property is immediately satisfied. We assume therefore that $V^* > \mathcal{M}V^*$ and $V^* > \mathcal{L}V^*$; we then need to show that

$$-g(\bar{t},\bar{x}) - \mathcal{A}\left(\bar{t},\bar{x},\bar{k},\partial_t\phi,\phi\right) \leq 0. \tag{6.11}$$

By continuity of the mapping in (6.11), we assume on the contrary that there exists a $\epsilon_1 > 0$ and an $\epsilon_2 > 0$ such that $-g(\bar{t},\bar{x}) - \mathcal{A}(\bar{t},\bar{x},\bar{k},\partial_t\phi,\phi) \geq \epsilon_1$, for all $X_{\bar{t},\bar{x}}(u) \in B_{\epsilon_2}(\bar{t},\bar{x})$. We take the sequences $h_m \to 0$ and $(t_m, \boldsymbol{x}_m) \to (\bar{t},\bar{x})$ valued in $B_{\epsilon_2}(\bar{t},\bar{x})$ and we define the stopping times θ_m by (6.8) with $\eta < \epsilon_2$ and $\theta_m^* := \theta_m \wedge (t_m + h_m) \wedge \theta^* \wedge \tau_i$. By Itô's formula and the dynamic programming principle, there exists an admissible control strategy $(\delta_u^{+,*}, \delta_u^{-,*})$ for which

$$\gamma_m - \frac{\epsilon_1 h_m}{2} \leq \mathbb{E}\left[\int_{t_m}^{\theta_m^*} \left[g\left(u, X_{t_m, x_m}(u)\right) + \bar{\mathcal{A}}\left(u, X_{t_m, x_m}(u), \bar{k}, \partial_t\phi, \phi, \delta_u^{+,*}, \delta_u^{-,*}\right)\right] du\right], \tag{6.12}$$

where $\gamma_m = (V^* - \phi)(t_m, \boldsymbol{x}_m; \bar{k})$. Dividing by $-h_m$, we find that

$$0 \geq \frac{\gamma_m}{h_m} - \frac{\epsilon_1}{2} + \frac{\epsilon_1}{h_m}\mathbb{E}[\theta_m^* - t_m]. \tag{6.13}$$

Since $\mathbb{E}[\theta_m^* - t_m]/h_m \to 1$ as $m \to \infty$, we get $\epsilon_1/2 \leq 0$, which contradicts $\epsilon_1 > 0$. □

Proposition 6.3 (Strong Comparison Principle) Let v and u be a supersolution and a subsolution respectively of the QVI (3.6). Then $u^* \leq v_*$ on $[0,T] \times S \times \mathbb{K}$.

Proof We write v and u in place of v_* and u^*, for simplicity. We first prove that there exists a ζ-strict supersolution, where $0 < \zeta < \epsilon_l$. We refer to, e.g. Seydel (2009) for technical details.

We consider the function $v^\zeta(t, \boldsymbol{x}; k) = v(t, \boldsymbol{x}; k) + \zeta\, e^{\beta(T-t)}\left(1 + ||\boldsymbol{x}||_{2p}^{2p}\right)$, where $\beta > 0$ and $p > 1$ are to be determined later. Then we have:

$$v^\zeta(t, \boldsymbol{x}; k) - \mathcal{M}v^\zeta(t, \boldsymbol{x}; k) \geq v(t, \boldsymbol{x}; k) + \zeta\, e^{\beta(T-t)}(1 + |x|^{2p} + |y|^{2p}) - \mathcal{M}v(t, \boldsymbol{x}; k)$$

$$- \sup_{\xi \in \mathcal{X}}\left[\zeta\, e^{\beta(T-t)}\left(1 + |\Lambda(\xi, x)|^{2p}\right)\right] - \sup_{\xi \in \mathcal{X}}\left[\zeta\, e^{\beta(T-t)}\left(1 + |c(\xi, y, s, k)|^{2p}\right)\right] + \epsilon_m$$

$$\geq \zeta\, e^{\beta(T-t)}\left[|x|^{2p} - \sup_{\xi \in \mathcal{X}}\left(|\Lambda(\xi, x)|^{2p}\right)\right] + \zeta\, e^{\beta(T-t)}\left[|y|^{2p} - \sup_{\xi \in \mathcal{X}}\left(|c(\xi, y, s, k)|^{2p}\right)\right] + \epsilon_m$$

$$> \zeta, \tag{6.14}$$

where the second-to-last inequality follows form the supersolution property of the function v, while the last inequality follows from assumption (iii) and the fact that $|a| > |b| \Rightarrow |a|^p > |b|^p \,\forall\, p > 1$. Analogously, we have:

$$v^\zeta(t, \boldsymbol{x}; k) - \mathcal{L}v^\zeta(t, \boldsymbol{x}; k) \geq \zeta\, e^{\beta(T-t)}\left[|x|^{2p} - \sup_{\eta \in \mathcal{N}, \kappa \in \mathbb{K}^\ell} \mathbb{E}^{(z)}\left[\left(|\Gamma(\eta, x, z)|^{2p}\right)\right]\right]$$

$$+ \zeta\, e^{\beta(T-t)}\left[|y|^{2p} - \sup_{\eta \in \mathcal{N}, \kappa \in \mathbb{K}^\ell} \mathbb{E}^{(z)}\left[\left(|\chi(\eta, y, z, s, k, \kappa)|^{2p}\right)\right]\right] + \epsilon_\ell$$

$$> \zeta.$$

Finally we take into consideration the PIDE part. We let ϕ^ζ be the test function for v^ζ. Then $\phi := \phi^\zeta - \zeta\, e^{\beta(T-t)}\left(1 + ||\boldsymbol{x}||_{2p}^{2p}\right)$ is the test function for v. We therefore have:

$$-g(t,x) - \mathcal{A}\left(t,\boldsymbol{x},k,\partial_t\phi^\zeta,v^\zeta\right) \geq -g(t,x) - \mathcal{A}\left(t,\boldsymbol{x},k,\partial_t\phi,v\right) + \beta\zeta\, e^{\beta(T-t)}\left(1 + \|\boldsymbol{x}\|_{2p}^{2p}\right)$$

$$+ \lambda^m \mathbb{E}^{(\bar{k}^\pm)}\left[\zeta\, e^{\beta(T-t)}\left(|s + \bar{k}^\pm|^{2p} - |s|^{2p}\right)\right]$$

$$- \sup_{\delta^+\in[0,\bar{\delta}]} \lambda_\delta^+ \mathbb{E}^{(q^+)}\left[\zeta\, e^{\beta(T-t)}\left(|x - q^+|^{2p} + |y + f(t,s,\delta^+,q^+)q^+|^{2p} - |x|^{2p} - |y|^{2p}\right)\right]$$

$$- \sup_{\delta^-\in[0,\bar{\delta}]} \lambda_\delta^- \mathbb{E}^{(q^-)}\left[\zeta\, e^{\beta(T-t)}\left(|x + q^-|^{2p} + |y - f(t,s,\delta^-,q^-)q^-|^{2p} - |x|^{2p} - |y|^{2p}\right)\right]$$

$$\geq \beta\zeta\, e^{\beta(T-t)}(1 + |x|^{2p}) - \sup_{\delta^+\in[0,\bar{\delta}]} \lambda_\delta^+ \mathbb{E}^{(q^+)}\left[\zeta\, e^{\beta(T-t)}\left(|x - q^+|^{2p} - |x|^{2p}\right)\right]$$

$$- \sup_{\delta^-\in[0,\bar{\delta}]} \lambda_\delta^- \mathbb{E}^{(q^-)}\left[\zeta\, e^{\beta(T-t)}\left(|x + q^-|^{2p} - |x|^{2p}\right)\right] \geq \zeta,$$

for β sufficiently large. Now we set

$$v_m = \left(1 - \frac{1}{m}\right)v + \frac{1}{m}v^\zeta, \quad u_m = \left(1 + \frac{1}{m}\right)u - \frac{1}{m}v^\zeta. \tag{6.15}$$

Using Definition 6.1, one can prove that

$$\min\left\{-g(t,x) - \mathcal{A}\left(t,\boldsymbol{x},k,\partial_t\phi_m,\phi_m\right); (v_m - \mathcal{M}v_m)(t,\boldsymbol{x};k); (v_m - \mathcal{L}v_m)(t,\boldsymbol{x};k)\right\} \geq \frac{\zeta}{m}. \tag{6.16}$$

where $\phi^m := \phi + \frac{1}{m}\zeta\, e^{\beta(T-t)}\left(1 + \|\boldsymbol{x}\|_{2p}^{2p}\right)$ is the test function for v_m and ϕ is the test function for v and

$$\min\left\{-g(t,x) - \mathcal{A}\left(t,\boldsymbol{x},k,\partial_t\varphi_m,\varphi_m\right); (u_m - \mathcal{M}u_m)(t,\boldsymbol{x};k); (u_m - \mathcal{L}u_m)(t,\boldsymbol{x};k)\right\} \leq -\frac{\zeta}{m}, \tag{6.17}$$

where $\varphi_m := \frac{m+1}{m}\varphi - \frac{1}{m}\phi - \frac{1}{m}\zeta e^{\beta(T-t)}\left(1 + \|\boldsymbol{x}\|_{2p}^{2p}\right)$ is the test function for u_m and φ is the test function for u. We further note that u and v are polynomially bounded (see e.g. Crisafi & Macrina, 2016, Proposition 6.3, for details). Thus, we have for each $k \in \mathbb{K}$

$$\lim_{\boldsymbol{x}\to\pm\infty}(u_m - v_m)(t,\boldsymbol{x};k) = \lim_{\boldsymbol{x}\to\pm\infty}\left(1 + \frac{1}{m}\right)(u - v)(t,\boldsymbol{x};k) - \frac{2}{m}\zeta\, e^{\beta(T-t)}\left(1 + \|\boldsymbol{x}\|_{2p}^{2p}\right) = -\infty, \tag{6.18}$$

where we set p larger than the degree of the bounding polynomial of u and v. Thus the supremum is attained in a bounded set. Since $u_m - v_m$ is upper semicontinuous, it attains a maximum over a compact set.

Next we show that, for all m large, we have

$$M := \max_{t,\boldsymbol{x},k}\left(u_m(t,\boldsymbol{x};k) - v_m(t,\boldsymbol{x};k)\right) \leq 0, \tag{6.19}$$

where $(\bar{t},\bar{\boldsymbol{x}},\bar{k}) = \arg\max(u_m(t,\boldsymbol{x};k) - v_m(t,\boldsymbol{x};k))$. We define the auxiliary function Ψ^ϵ by

$$\Psi^\epsilon\left(t_1,t_2,\boldsymbol{x}_1,\boldsymbol{x}_2;k\right) := u_m\left(t_1,\boldsymbol{x}_1;k\right) - v_m\left(t_2,\boldsymbol{x}_2;k\right) - \frac{1}{2\epsilon}\left(|t_1 - t_2|^2 + \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2\right), \tag{6.20}$$

For each $k \in \mathbb{K}$, Ψ^ϵ is upper semicontinuous and therefore it admits a maximum $M^{\epsilon,k}$ at $\left(t_1^{k,\epsilon},t_2^{k,\epsilon},\boldsymbol{x}_1^{k,\epsilon},\boldsymbol{x}_2^{k,\epsilon}\right)$. Let M^ϵ be defined by $M^\epsilon = \max_{k\in\mathbb{K}} M^{\epsilon,k}$, attained at the point $(t_1^\epsilon,t_2^\epsilon,\boldsymbol{x}_1^\epsilon,\boldsymbol{x}_2^\epsilon,k^\epsilon) \to (\bar{t},\bar{t},\bar{\boldsymbol{x}},\bar{\boldsymbol{x}},\bar{k})$ as $\epsilon \to 0$. Furthermore we have that $M^\epsilon \geq M$ and $M^\epsilon \to M$ as $\epsilon \to 0$. Let us assume on the contrary that $M^\epsilon > 0$.

We analyse the various cases. Let

$$(u_m - \mathcal{M}u_m)\left(t_1^\epsilon,\boldsymbol{x}_1^\epsilon;k^\epsilon\right) \leq 0. \tag{6.21}$$

By the supersolution property of v_m and by subtracting the two inequalities, we have

$$(u_m - \mathcal{M}u_m)(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) - (v_m - \mathcal{M}v_m)(t_2^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon) + \frac{\zeta}{m} \leq 0. \tag{6.22}$$

We can now develop a contradiction argument since

$$M = \lim_{\epsilon \to 0} u_m(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) - v_m(t_2^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon)$$

$$\leq \lim_{\epsilon \to 0} \mathcal{M}u_m(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) - \mathcal{M}v_m(t_2^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon) - \frac{\zeta}{m} \leq \lim_{\epsilon \to 0} M^\epsilon - \frac{\zeta}{m} = M - \frac{\zeta}{m}. \tag{6.23}$$

The second case arises when $(u_m - \mathcal{L}u_m)(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) \leq 0$. We follow the same procedure to show that

$$M = \lim_{\epsilon \to 0} u_m(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) - v_m(t_2^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon)$$

$$\leq \lim_{\epsilon \to 0} \mathcal{L}u_m(t_1^\epsilon, \mathbf{x}_1^\epsilon; k^\epsilon) - \mathcal{L}v_m(t_2^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon) - \frac{\zeta}{m} \leq \lim_{\epsilon \to 0} M^\epsilon - \frac{\zeta}{m} = M - \frac{\zeta}{m}. \tag{6.24}$$

Next we consider the PIDE part. We subtract the argument of the subsolution from the argument of the supersolution and obtain

$$g(t_1^\epsilon, x_1^\epsilon) + \mathcal{A}\left(t_1^\epsilon, \mathbf{x}_1^\epsilon, k^\epsilon, \frac{1}{\epsilon}(t_1^\epsilon - t_2^\epsilon), u_m\right) - g(t_2^\epsilon, x_2^\epsilon) - \mathcal{A}\left(t_2^\epsilon, \mathbf{x}_2^\epsilon, k^\epsilon, \frac{1}{\epsilon}(t_1^\epsilon - t_2^\epsilon), v_m\right) \geq \frac{\zeta}{m}.$$

Since we assumed $M^\epsilon > 0$, we choose a $\varrho > 0$ such that

$$0 < \varrho M^\epsilon = \varrho\left(u_m(t_1^\epsilon, x_1^\epsilon; k^\epsilon) - v_m(t_2^\epsilon, x_2^\epsilon; k^\epsilon) - \frac{1}{2\epsilon}\left(|t_1^\epsilon - t_2^\epsilon|^2 + \|\mathbf{x}_1^\epsilon - \mathbf{x}_2^\epsilon\|_2^2\right)\right)$$

$$\leq g(t_1^\epsilon, x_1^\epsilon) + \mathcal{A}\left(t_1^\epsilon, \mathbf{x}_1^\epsilon, k^\epsilon, \frac{1}{\epsilon}(t_1^\epsilon - t_2^\epsilon), u_m\right) - g(t_2^\epsilon, x_2^\epsilon) - \mathcal{A}\left(t_2^\epsilon, \mathbf{x}_2^\epsilon, k^\epsilon, \frac{1}{\epsilon}(t_1^\epsilon - t_2^\epsilon), v_m\right).$$

We can now analyse every component in detail. First we note that, due to Assumption (iv),

$$g(t_1^\epsilon, x_1^\epsilon) - g(t_2^\epsilon, x_2^\epsilon) \leq \left|g(t_1^\epsilon, x_1^\epsilon) - g(t_2^\epsilon, x_2^\epsilon)\right| \leq C\left(|t_1^\epsilon - t_2^\epsilon| + |x_1^\epsilon - x_2^\epsilon|\right) \to 0 \tag{6.25}$$

as $\epsilon \to 0$. For the remaining parts, let us rewrite

$$\mathcal{B}_s^k(t, \mathbf{x}, \psi) := \mathbb{E}^{(\bar{k}^\pm)}\left[\left(\psi(t, x, y, s + \bar{k}^\pm; k) - \psi(t, \mathbf{x}; k)\right)\right],$$

$$\mathcal{B}_+^k(t, \mathbf{x}, \psi) := \sup_{\delta^+ \in [0, \bar{\delta}]} \lambda_\delta^+ \mathbb{E}^{(q^+)}\left[\left(\psi(t, x - q^+, y + f(t, s, \delta^+, q^+)q^+, s; k) - \psi(t, \mathbf{x}; k)\right)\right],$$

$$\mathcal{B}_-^k(t, \mathbf{x}, \psi) := \sup_{\delta^- \in [0, \bar{\delta}]} \lambda_\delta^- \mathbb{E}^{(q^-)}\left[\left(\psi(t, x + q^-, y - f(t, s, \delta^-, q^-)q^-, s; k) - \psi(t, \mathbf{x}; k)\right)\right],$$

and

$$\mathcal{Q}\psi(t, \mathbf{x}; k) := \sum_{k' \neq k} r_{kk'}\left[\psi(t, \mathbf{x}; k') - \psi(t, \mathbf{x}; k)\right].$$

We analyse $\mathcal{B}_+^{k^\epsilon}(t_1^\epsilon, \mathbf{x}_1^\epsilon, u_m) - \mathcal{B}_+^{k^\epsilon}(t_2^\epsilon, \mathbf{x}_2^\epsilon, v_m)$ as the other integrals can be treated analogously. After some manipulations, we find:

$$\mathcal{B}_a^{k^\epsilon}(t_1^\epsilon, \mathbf{x}_1^\epsilon, u_m) - \mathcal{B}_a^{k^\epsilon}(t_2^\epsilon, \mathbf{x}_2^\epsilon, v_m)$$

$$\leq \sup_{\delta^+ \in [0, \bar{\delta}]} \lambda_\delta^+ \mathbb{E}^{(q^+)}\big[\Psi^\epsilon(t_1^\epsilon, t_2^\epsilon, x_1^\epsilon - q^+, x_2^\epsilon - q^+, y_1^\epsilon + f(t_1^\epsilon, s_1^\epsilon, \delta^+, q^+)q^+, y_2^\epsilon + f(t_2^\epsilon, s_2^\epsilon, \delta^+, q^+)q^+, s_1^\epsilon, s_2^\epsilon; k^\epsilon)$$

$$- \Psi^\epsilon(t_1, t_2, \mathbf{x}_1, \mathbf{x}_2; k) + \frac{1}{2\epsilon}\left(\left|y_1^\epsilon + f(t_1^\epsilon, s_1^\epsilon, \delta^+, q^+)q^+ - y_2^\epsilon - f(t_2^\epsilon, s_2^\epsilon, \delta^+, q^+)q^+\right|^2\right)$$

$$- |y_1^\epsilon - y_2^\epsilon|\big)\big].$$

Since the function Ψ^ϵ attains its maximum at $(t_1^\epsilon, t_2^\epsilon, \mathbf{x}_1^\epsilon, \mathbf{x}_2^\epsilon; k^\epsilon)$, we have

$$\mathcal{B}_+^{k^\epsilon}(t_1^\epsilon, \mathbf{x}_1^\epsilon, u_m) - \mathcal{B}_+^{k^\epsilon}(t_2^\epsilon, \mathbf{x}_2^\epsilon, v_m)$$

$$\leq \sup_{\delta^+ \in [0, \bar{\delta}]} \lambda_\delta^+ \mathbb{E}^{(q^+)}\left[\frac{1}{2\epsilon}\left(\left|y_1^\epsilon + f(t_1^\epsilon, s_1^\epsilon, \delta^+, q^+)q^+ - y_2^\epsilon - f(t_2^\epsilon, s_2^\epsilon, \delta^+, q^+)q^+\right|^2 - |y_1^\epsilon - y_2^\epsilon|^2\right)\right].$$

Since the right-hand-side of the above equality tends to zero as $\epsilon \to 0$, we have that

$$\lim_{\epsilon \to 0} \mathcal{B}_+^{k^\epsilon}\left(t_1^\epsilon, \boldsymbol{x}_1^\epsilon, u_m\right) - \mathcal{B}_+^{k^\epsilon}\left(t_2^\epsilon, \boldsymbol{x}_2^\epsilon, v_m\right) \leq 0. \tag{6.26}$$

Finally, we have

$$\mathcal{Q}\left(u_m\left(t_1^\epsilon, \boldsymbol{x}_1^\epsilon; k^\epsilon\right) - v_m\left(t_2^\epsilon, \boldsymbol{x}_2^\epsilon; k^\epsilon\right)\right) \leq 0 \tag{6.27}$$

since the maximum is attained at k^ϵ. Thus, by letting $\epsilon \to 0$, we get $\varrho M \leq 0$, which is a contradiction since $\varrho > 0$. Therefore $M \leq 0$. Furthermore, since we have proved that $u^* \leq v_*$, the value function is continuous as it is both upper and lower semicontinuous. □

Author details
M. Alessandra Crisafi[1]
E-mail: m.crisafi.12@ucl.ac.uk
Andrea Macrina[1,2]
E-mail: a.macrina@ucl.ac.uk

[1] Department of Mathematics, University College London, London, UK.

[2] Department of Actuarial Science, University of Cape Town, Cape Town, South Africa.

Notes:
1. The terminology "alternative liquidity pool" is also used.
2. In foreign exchange, examples of centralised liquidity pools include EBS, FXAll, Hotspot, Thomson Reuters, etc. and in equity we may mention LSE, NYSE, NASDAQ, etc.

References
Amihud, Y., & Mendelson, H. (1980). Dealership market. with inventory. *Journal of Financial Economics, 8*, 31–53.

Avellaneda, M., & Stoikov, S. (2008). High-frequency trading in a limit order book. *Quantitative Finance, 8*, 217–224.

Cartea, Á., Donnelly, R., & Jaimungal, S. (2014). *Algorithmic trading with model uncertainty*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2310645

Cartea, Á., & Jaimungal, S. (2013). Risk metrics and fine tuning of high frequency trading strategies. *Forthcoming in Mathematical Finance*. doi:10.1111/mafi.12023

Cartea, Á, Jaimungal, S., & Penalva, J. (2016). *Algorithmic and high-frequency trading*. Cambridge: Cambridge University Press.

Cartea, Á, Jaimungal, S., & Ricci, J. (2014). Buy low sell high: A high frequency trading perspective. *SIAM Journal of Financial Mathematics, 5*, 415–444.

Crandall, M. G., Ishii, H., & Lions, P. L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bulletin American Mathematical Society, 27*, 1–67.

Crisafi, M. A., & Macrina, A. (2016). Simultaneous trading in 'Lit' and dark pools. *International Journal of Theoretical and Applied Finance, 19*, 8, 1650055.

Fleming, W. H., & Soner, H. M. (2006). *Controlled Markov processes and viscosity solutions*. New York, NY: Springer.

Garman, M. B. (1976). Market microstructure. *Journal of Financial Economics, 3*, 257–275.

Guéant, O. (2016). *The financial mathematics of market liquidity: From optimal execution to market making*. Boca Raton, FL: Chapman and Hall/CRC Financial Mathematics Series.

Guéant, O., Lehalle, C. A., & Tapia, J. F. (2013). Dealing with the inventory risk: A solution to the market making problem. *Mathematics and Financial Economics, 7*, 477–507.

Guilbaud, F., & Pham, H. (2013). Optimal high-frequency trading with limit and market orders. *Quantitative Finance, 13*, 74–94.

Guilbaud, F., & Pham, H. (2015). Optimal high-frequency trading in a pro rata microstructure with predictive information. *Mathematical Finance, 25*, 545–575.

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

Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*. Heidelberg: Springer.

Seydel, R. C. (2009). Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions. *Stochastic Processes and their Applications, 119*, 3719–3748.

Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance, 33*, 1133–1151.

Vath, V. L., Mnif, M., & Pham, H. (2007). A model of optimal portfolio selection under liquidity risk and price impact. *Finance and Stochastics, 4*, 579–603.

cogent·mathematics

Cogent Mathematics (ISSN: 2331-1835) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com