Decentralised Finance and Automated Market Making: Execution and Speculation

Álvaro Cartea^{a,b}, Fayçal Drissi^{a,c}, Marcello Monga^{a,b}

^a Oxford-Man Institute of Quantitative Finance, Oxford, UK
 ^b Mathematical Institute, University of Oxford, Oxford, UK
 ^c Université Paris 1 Panthéon-Sorbonne, Paris, France

Abstract

Automated market makers (AMMs) are a new prototype of trading venues which are revolutionising the way market participants interact. At present, the majority of AMMs are constant function market makers (CFMMs) where a deterministic trading function determines how markets are cleared. A distinctive characteristic of CFMMs is that execution costs are given by a closed-form function of price, liquidity, and transaction size. This gives rise to a new class of trading problems. We focus on constant product market makers and show how to optimally trade a large position in an asset and how to execute statistical arbitrages based on market signals. We employ stochastic optimal control tools to devise two strategies. One strategy is based on the dynamics of prices in competing venues and assumes constant liquidity in the AMM. The other strategy assumes that AMM prices are efficient and liquidity is stochastic. We use Uniswap v3 data to study price, liquidity, and trading cost dynamics, and to motivate the models. Finally, we perform consecutive runs of in-sample estimation of model parameters and out-of-sample liquidation and arbitrage strategies to showcase the performance of the strategies.

Keywords: Decentralised Finance, Automated Market Making, Algorithmic Trading, Statistical Arbitrage, Predictive Signals, Market Impact, Adaptive Strategies, Smart Contracts.

1. Introduction

Decentralised Finance (DeFi) is a collective term for blockchain-based financial services that do not rely on intermediaries such as brokers or banks. New powerful technologies are the engine behind the remarkable growth of DeFi, which is changing the financial landscape and is in

We are grateful to Tarek Abou Zeid, Olivier Guéant, Sebastian Jaimungal, Anthony Ledford, and Andre Rzym for insightful comments. We are also grateful to seminar participants at Oxford–ETH 2022 Workshop, and the Oxford Victoria Seminar. MM acknowledges financial support from the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EP/S023925/1).

direct competition with many traditional stakeholders. Within DeFi, automated market makers (AMMs) are a new paradigm in the design of trading venues and are revolutionising the way market participants provide and take liquidity. Currently, AMMs are mainly known as exchanges for cryptocurrencies; however, the core concepts of AMMs go beyond the cryptocurrency sector and they are poised to challenge traditional electronic exchanges in all asset classes.

At present, the majority of AMMs are constant **function** market makers (CFMMs). In CFMMs, a trading function and a set of rules determine how liquidity takers (LTs) and liquidity providers (LPs) interact, and how markets are cleared. The trading function is deterministic and known to all market participants. CFMMs display pools of liquidity for pairs of assets, where the relative prices between the two assets are determined by their quantities in the pool as prescribed by the trading function. The trading function establishes the link between liquidity and prices, so LTs can compute the execution costs of their trades as a function of the trade size. A key difference between CFMMs and limit order books (LOBs) is that execution costs in CFMMs are given by a closed-form formula where the convexity implied by the trading function plays a key role. As in traditional markets that operate an LOB, the larger the size of an order, the higher are the execution costs.

Within CFMMs, we focus on constant **product** market makers (CPMMs), which are the most popular type of CFMM and where the trading function uses the product of the quantities of each asset in the pool to determine clearing prices. In this paper, we solve the problem of an investor who wishes to trade in a CPMM to execute a large position in an asset and to execute statistical arbitrages based on market signals. We formulate the trading problem as a stochastic control problem where the investor controls the speed at which she sends liquidity taking orders. Key to the performance of the investor's strategies, is to balance price risk and execution costs. In CPMMs, the execution costs of the trading function are inversely proportional to the depth of the pool and proportional to a non-linear transformation of the relative prices of the two assets in the pool.

Despite very high levels of activity for many of the pairs traded in CPMMs, price formation currently occurs in the LOBs of alternative electronic markets. In one version of our model, we assume that LTs in a CPMM inform their decisions with the prices in the CPMM and those from an alternative venue, and assume that the depth of the pool does not vary during the investor's trading horizon. In this setup, we derive a versatile trading strategy which can be used to focus on the execution of a large order or on statistical arbitrages.

When the focus is to exchange a large position in one asset for another asset, both of which are provided as a pair in the CPMM, the strategy relies on two components. One component is as in the

traditional execution strategies (e.g., TWAP-like or Almgren-Chriss), and the second component is an arbitrage that takes advantage of short-lived discrepancies in the prices of the CPMM and those in the alternative venue. Instead, if the objective is speculation, the strategy relies more on the statistical arbitrage component to take advantage of the lead-follow relationship between the prices in the CPMM and those in the alternative venue.

In anticipation of the growth of AMMs, another version of our model assumes that prices in the CFMMs are efficient, so the discrepancies between the CFMM and LOB prices are not economically significant. The increase in the efficiency of prices in CFMMs will be a result of an increase in the activity of LPs and LTs, which will also result in more changes in the depth of the pool of the CPMM. Thus, we propose another model where the depth of the pool is stochastic and we solve the investor's execution problem for large orders.

We use Uniswap data for CPMMs that trade pairs of cryptocurrencies to study the empirical properties of this particular AMM, and to illustrate the performance of the proposed liquidation and speculative strategies. The efficient prices are those from Binance where traders interact through a traditional price-time priority LOB. To showcase the performance of our strategies, we use in-sample data to estimate model parameters and out-of-sample data to execute the strategies in 'real time' as an investor would have done. In our analysis, we use rolling time windows of a few hours starting 1 July 2021 and ending 5 May 2022 to obtain the distribution of the financial performance of the strategies. We look at two pairs of assets, one that is heavily traded and one that is not as frequently traded. We show the superior performance of our execution strategy over TWAP and over a strategy that would have executed the whole inventory in one trade at the beginning of the trading horizon. Finally, we show that there are profitable opportunities to execute statistical arbitrages in Uniswap when the strategy is informed by Binance prices.

Early work on optimal execution in traditional electronic markets is in Bertsimas and Lo (1998) and Almgren and Chriss (2000); see also Cartea et al. (2015) and Guéant (2016). Our work is related to the optimal execution literature that uses market signals to improve performance of strategies. The work of Cartea and Jaimungal (2016) derives closed-form solutions for execution problems with market signals (e.g., order flow); see also Bechler and Ludkovski (2015), Cartea et al. (2018a), Lehalle and Neuman (2019), Neuman and Voß (2020), Forde et al. (2022), and Belak et al. (2018). Further extensions to this approach involve, among others, multi-asset frameworks such as those in Cartea et al. (2018b) and Bergault et al. (2022). Algorithmic trading in AMMs is subject to latency, and we refer the reader to the work in Cartea and Sánchez-Betancourt (2021a,b); Cartea et al. (2021) for extensions of algorithmic trading models with delay.

Optimal execution in AMMs considers new market impact models that use endogenous dy-

namics of price and liquidity. Early work on dynamic market impact models is in Almgren (2012) and Gatheral and Schied (2013); see also Graewe et al. (2018), which uses a price-dependent impact function, Cheridito and Sepin (2014), which uses stochastic volatility and temporary price impact, and Fouque et al. (2021), where the authors introduce fast-mean reverting stochastic price impact.

To the best of our knowledge, this is the first paper to solve optimal trading for LTs in CPMMs. Early work on microstructural aspects of AMMs can be found in Lipton and Treccani (2021) and Chiu and Koeppl (2019). For LPs, the work in Cartea et al. (2022) provides an analysis of liquidity provision in CFMMs and CPMMs with concentrated liquidity. In particular, they introduce predictable loss (PL) as a comprehensive metric to compute the losses incurred by LPs in these venues. They show that PL is a result of two sources of loss; the convexity cost as a result of liquidity taking activity and the properties of the trading function, and the opportunity cost as a result of locking the LPs' assets in the pool. Finally, they provide an optimal liquidity provision strategy in closed-form in the case of CPMMs with concentrated liquidity (CL), which is shown to have a superior performance to that of the LPs in the Uniswap v3 pool they consider.

The remainder of this paper is organised as follows. Section 2 discusses how CFMMs operate and uses Uniswap v3 data to study price, liquidity, and trading cost dynamics in CPMMs. Section 3 solves the optimal execution problem when the pool has constant depth during the execution window and price formation is in an alternative trading venue. Section 4 solves the optimal execution problem when the pool depth is stochastic and price formation takes place in the AMM. Finally, Section 5 showcases the performance of liquidation and statistical arbitrage strategies.

2. Automated market making

In this section, we discuss how CFMMs operate and how they differ from electronic markets where traders interact through an LOB. In particular, we describe the interactions of market participants with a CFMM that is in charge of a pair of assets. We use transaction data from Uniswap v3 to study the activity of market participants, the dynamics of liquidity, and implicit transaction costs.

2.1. Description

AMMs are hard-coded and immutable programs running on a network. They provide a venue to trade pairs of assets X and Y, where the liquidity of the pool consists of x units of X and y units of Y. The exchange rate of the pool is the price of Y in terms of the price of X, and it is

determined by the quantities x and y.¹ Two types of market participants interact in an AMM: LPs deposit their assets in the pool and LTs trade directly with the pool.

Here, we consider a CFMM without CL, in charge of a single pool for the pair of assets X and Y. CFMMs are characterised by a deterministic trading function f(x,y) that determines the rules of engagement among participants in the pool. For instance, the trading function of the CPMM is $f(x,y)=x\times y$. Other types of CFMMs are the constant sum market maker with f(x,y)=x+y; the constant mean market maker with $f(x,y)=w_x\,x+w_y\,y$, where $w_x,w_y>0$ and $w_x+w_y=1$; and the hybrid function market maker, which uses combinations of trading functions.

In peer-to-peer networks, participants invoke the code of the AMM smart contract to instruct market operations. LPs send messages with instructions to deposit or withdraw liquidity, and LTs send messages to exchange one asset for the other. To provide liquidity, an LP instructs the AMM with the quantities in assets X and Y to be deposited in a specific pool. On the other hand, LTs indicate to the AMM the pool and the quantity of the asset to be exchanged. The available liquidity in the pool and the trading function of the AMM determine the exchange rate received by the LT. For each trade, LTs pay the AMM a transaction fee, which is distributed amongst LPs in the same proportion as their contributions to the pool.²

The trading function f(x,y) is increasing in x and y, and it ties the state of the pool before and after an LT transaction is executed. Throughout, the signs of Δx and Δy are the same. If $\Delta y > 0$, the LT sells asset Y, and if $\Delta y < 0$, the LT buys asset Y; for simplicity, we assume zero fees.³ The condition

$$f(x - \Delta x, y + \Delta y) = f(x, y) = \kappa^2$$

determines the quantity Δx that the agent receives (pays) when exchanging $\Delta y > 0$ ($\Delta y < 0$). The trading function keeps the quantity κ^2 constant before and after a trade is executed. We write $f(x,y) = \kappa^2$ as $x = \varphi(y)$ for an appropriate function φ that depends on κ ; we refer to φ as the level function, and assume it is convex.⁴

If an LT wishes to sell Δy of asset Y, she receives $\Delta x = \Delta y \times \tilde{Z}(\Delta y)$ of asset X in exchange. Here, $\tilde{Z}(\Delta y)$, with units X/Y, is the exchange rate received by the agent when trading a quantity

¹Some AMMs also display pools with more than two assets.

²See Heimbach et al. (2021) and Cartea et al. (2022) for an analysis on how LPs profit from their activity.

³To take into account the fee for an LT transaction, one applies a discount to the quantity Δy before calculations are carried out.

⁴The convexity of the level function is by design. One can show that a no-arbitrage condition leads to the necessary convexity of the level function; see Cartea et al. (2022).

 Δy of asset Y. Therefore,

$$x - \Delta x = \varphi(y + \Delta y) \implies \varphi(y) - \Delta y \, \tilde{Z}(\Delta y) = \varphi(y + \Delta y)$$

SO

$$\tilde{Z}(\Delta y) = \frac{\varphi(y) - \varphi(y + \Delta y)}{\Delta y}, \qquad (1)$$

and for an infinitesimal quantity Δy we write

$$Z = -\varphi'(y). (2)$$

We refer to Z as the *instantaneous rate* of the AMM, which is equivalent to the midprice in an LOB. The instantaneous rate Z is a reference exchange rate – the difference between its value and the execution rate is similar to the difference between the LOB midprice and the average price obtained by a liquidity taking order that crosses the spread and walks the book when it is filled.

The trading function f(x,y) is increasing in the pool quantities x and y. Thus, when LP activity increases (decreases) the size of the pool, the value of κ increases (decreases). We refer to κ as the depth of the pool. A distinctive characteristic of AMMs is that liquidity provision changes the depth of the pool, but it does not change the instantaneous rate. For example, in a CPMM, the instantaneous rate is the ratio of the quantities supplied in the pool, i.e.,

$$Z = \frac{x}{y},$$

and when an LP deposits quantities Δx and Δy in the pool, the pair $(\Delta x, \Delta y)$ must satisfy

$$\frac{x}{y} = \frac{x + \Delta x}{y + \Delta y} = Z, \tag{3}$$

and the value of κ changes from $\sqrt{x \times y}$ to $\sqrt{(x + \Delta x)(y + \Delta y)}$. For (3) to hold, there exists ρ such that $\Delta x = \rho x$ and $\Delta y = \rho y$, i.e., liquidity provision and removal by LPs in a CPMM is performed in fractions of the pool quantities x and y, and the depth changes from κ to $(1 + \rho) \kappa$.

The level function φ is decreasing, thus the rate in (1) received by an LT deteriorates as the size of the trade increases.⁵ The formulas (1) and (2) encode all the information needed by an LT to interact with an AMM. For a trade of size Δy , the distance between the instantaneous rate Z

⁵The trading function f is increasing in x and y and $\partial_y f(x,y) = \partial_y f(\varphi(y),y) = 0$ so $\varphi'(y) = -\frac{\partial_y f(x,y)}{\partial_x f(x,y)} < 0$.

and the execution rate $\tilde{Z}(\Delta y)$ is the *unitary execution cost* of the AMM, specifically

Unitary execution cost
$$= |Z - \tilde{Z}(\Delta y)|$$
. (4)

Another distinctive characteristic of AMMs is that liquidity taking activity may change the instantaneous rate Z, but does not change the depth of the pool. Furthermore, AMMs and LOBs differ in a number of other aspects including accessibility and the way LPs are compensated. By design, AMMs are permissionless, so anyone can participate in the market.⁶ LPs in AMMs are compensated in two ways. One, they are rewarded with the proceeds from fees that LTs pay for every trade. Two, the execution costs in (4) incurred by LTs due to the convexity of the level function stay in the pool. In some AMMs, the fee proceeds are put back in the pool, so the number of assets owned by LPs increases. In other AMMs, in particular those with a CL feature, fees are accumulated in a separate account and are earned by LPs when they withdraw their liquidity from the pool.

2.2. Data analysis

2.2.1. Data description

Uniswap v3 is considered the cornerstone of DeFi and is currently the most liquid AMM. It is a CPMM with trading function

$$f(x,y) = x \times y = \kappa^2, \tag{5}$$

so the level function is

$$\varphi(y) = \frac{\kappa^2}{y} \, .$$

When an LT trades Δy , the execution rate (1) is

$$\tilde{Z}(\Delta y) = \frac{1}{\Delta y} \left(\frac{\kappa^2}{y} - \frac{\kappa^2}{y + \Delta y} \right) ,$$

and the instantaneous rate (2) is

$$Z = -\varphi'(y) = \frac{\kappa^2}{y^2} .$$

To further study the characteristics of CPMMs and motivate our framework, we use transaction data from Uniswap v3 and the traditional LOB-based exchange Binance. Specifically, we look at

⁶This is true for AMMs running on permissionless peer-to-peer networks. If the same AMMs were to be implemented by a central authority for fiat currencies or stocks, then one expects stricter participation rules.

the two pairs ETH/USDC and ETH/DAI. The ticker ETH represents the cryptocurrency *Ether*, which is the native cryptocurrency of the Ethereum blockchain. The ticker USDC represents *USD coin*, a cryptocurrency fully backed by U.S. Dollars (USD); and DAI represents the cryptocurrency *Dai*, which tracks parity with the U.S. Dollar. For the pool ETH/USDC, the unit of depth κ is $\sqrt{\text{ETH} \cdot \text{USDC}}$, of x is USDC, of y is ETH, and the instantaneous rate, the execution rate, and the unitary execution cost are all in USDC/ETH; similarly for the pool ETH/DAI. For ease of reading, in the remainder of this work we omit the units of κ .

We analyse transaction data of the most liquid pools for the pairs ETH/USDC and ETH/DAI. Transaction information from decentralised exchanges is public. Between 5 May 2021 and 5 May 2022, there are 1,757,181 LT transactions and 42,403 LP transactions for the pair ETH/USDC, and 101,538 LT transactions and 12,142 LP transactions for the pair ETH/DAI. Figures 1 and 2 show the historical daily volumes of transactions in terms of USD value and in terms of the number of transactions for ETH/USDC and ETH/DAI, respectively. These pools are considered an alternative to LOB-based trading venues such as Binance, which is the most liquid and active venue for both pairs.

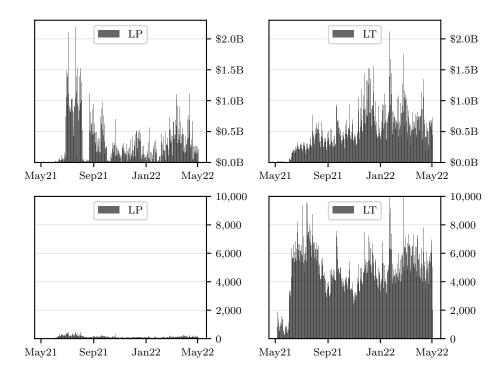


Figure 1: Daily transactions for the ETH/USDC pool (B is shorthand for billions), between 5 May 2021 to 5 May 2022. **Top left**: LP transaction volume in USD. **Top right**: LT transaction volume in USD. **Bottom left**: number of LP transactions. **Bottom right**: number of LT transactions.

Next, we study liquidity provision and taking activity within the pools and examine their key

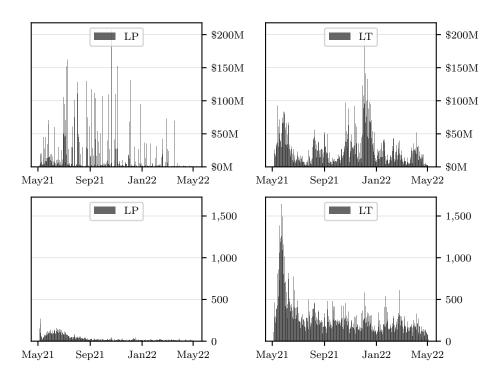


Figure 2: Daily transactions for the ETH/DAI pool (M is shorthand for millions), between 5 May 2021 to 5 May 2022. **Top left**: LP transaction volume in USD. **Top right**: LT transaction volume in USD. **Bottom left**: number of LP transactions.

characteristics, all of which we use to frame the agent's execution problem.

2.2.2. Rates and liquidity dynamics

Compared with most CPMMs, Uniswap v3 operates with the called concentrated liquidity feature. In CPMMs without this feature, each LP owns a percentage of the pool, and the fees paid by LTs are distributed to LPs in the same proportion; thus, LPs provide liquidity at all feasible rates. In contrast, LPs in Uniswap v3 specify the range of rates where they supply liquidity. For example, LPs can target a range around the instantaneous rate to earn more fees than those who provide liquidity at rates far from the instantaneous rate. In practice, the continuous space of possible rates is discretised and subdivided in rate intervals whose boundaries are called *ticks*. Two consecutive ticks define a *tick range* and the rate can take values in this range with increments set by the AMM. LPs designate two ticks between which they wish to provide liquidity. Therefore, with CL, the pool is characterised by the available quantities in tick ranges. Finally, in Uniswap v3, fees paid by LTs are distributed among the LPs who had provided liquidity in a range that

⁷In Uniswap v3, ticks are specific rates that are used as the boundaries of an LP transaction. In contrast, ticks in LOBs represent the smallest price increment.

included the rate at which liquidity was taken.

In CPMMs without CL, the value of κ is the same for all the tick ranges and recall that the value of κ can only change when LPs deposit or withdraw liquidity from the pool. On the other hand, due to the CL feature of Uniswap v3, the depth of the pool may be different when the instantaneous rate crosses the boundary of a tick because liquidity is distributed unevenly across rates in the pool. Therefore, in pools with CL, the level function is

$$\varphi(y) = \frac{1}{y} \sum_{i=1}^{N} \kappa_i^2 \, \mathbb{1}_{Z \in [Z_i, Z_{i+1})},$$

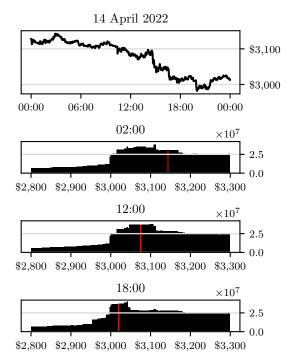
where N is the number of ticks, Z is the instantaneous rate, Z_i and Z_{i+1} are two consecutive ticks, κ_i is the depth of the liquidity available in the tick range $[Z_i, Z_{i+1}]$, and $\mathbbm{1}$ is the indicator function. Thus, to estimate the execution cost incurred when trading, one must track the distribution of liquidity across ranges of rates. Finally, Note that in the extreme case where all liquidity is withdrawn from a range of rates around the current rate Z, this effectively constitutes a change in the instantaneous rate of the pool – the next LT transaction will start at a rate different from Z.

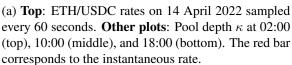
To analyse liquidity provision and consumption activity, we reconstruct the supply and demand of liquidity since the inception of Uniswap v3 in May 2021. For example, Figure 3 shows the amount of liquidity, given by the depth $\sqrt{x_i \times y_i} = \kappa_i$, available at each rate range at 02:00, 12:00, and 18:00, on 14 April 2022 and 15 April 2022 for the ETH/USDC pool.

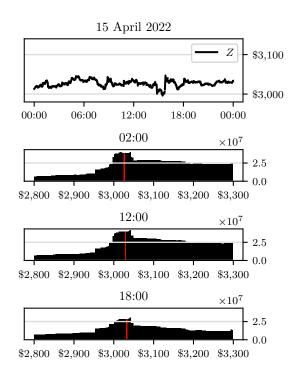
The bars in the panels of Figure 3 represent a tick range. Whenever the instantaneous rate is in a given range, the liquidity of that range defines the depth κ used by the trading function, and hence the execution rate (1). When the volume of an LT transaction is large enough to make the instantaneous rate cross a tick where the level of liquidity changes, the AMM treats it as multiple trades, each with a different κ_i ; similar to a market order walking the levels of an LOB. Tracking the instantaneous rate, and the liquidity around it, is critical for an agent interacting with an AMM. In Figure 3, most of the liquidity is concentrated around the instantaneous rate and the depth is the same over a large range around the rate. The behaviour of LPs is different when volatility is high or low. When Z is volatile, such as on 14 April on the left panel of Figure 3, LPs provide liquidity over a wider range around the rate to earn fees that anticipate large swings in Z. On the other hand, when markets are less volatile, such as on 15 April in the right panel of Figure 3, LPs concentrate their liquidity provision more tightly around the rate Z.

Next, we use transaction data from the LOB-based exchange Binance for the same pairs to compare the dynamics in the two trading venues. Figure 4 shows the Binance quoted rate and the

⁸Here, κ_i is the sum of all the liquidity provided in a range containing the tick range $[Z_i, Z_{i+1})$.





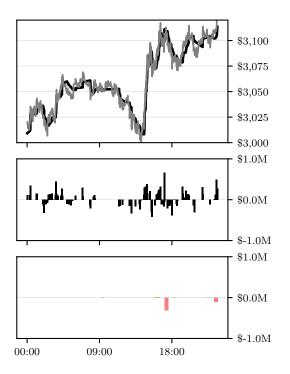


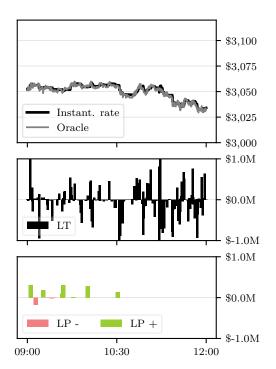
(b) **Top**: ETH/USDC rates on 15 April 2022 sampled every 60 seconds. **Other plots**: Pool depth κ at 02:00 (top), 10:00 (middle), and 18:00 (bottom). The red bar corresponds to the instantaneous rate.

Figure 3: Rate and LP dynamics in AMMs.

AMM rate, and liquidity provision activity for both pools between 09:00 and 12:00 on 13 April 2022.

Figure 4 shows that Binance rates are more volatile and lead the rates in the AMM. This is not by design, it is a consequence of the higher liquidity in Binance. Currently, it is crucial to consider the rate from a more liquid venue when trading in an AMM. In what follows, we call the leading exchange rate from another trading venue the *oracle*, which in this case is the Binance quoted rate. Similar to Figures 1 and 2, Figure 4 shows that there is more LT than LP activity measured by the frequency of instructions and the size of the orders. During the periods with little trading activity in the ETH/DAI pool, the oracle rate plays a central role to attract LT activity in the AMM whenever the difference between the instantaneous rate and the oracle rate is significant (recall that only liquidity taking trades can change the rate of the pool). The widening of the difference between the two rates triggers LT activity which drives the two exchange rates to converge; i.e., arbitrageurs keep markets in check.





(a) **Top**: ETH/DAI instantaneous rate Z and oracle S between April 13 00:00 and April 13, 2022 12:00. **Middle**: LT Transaction size. **Bottom**: LP transaction size.

(b) **Top**: ETH/USDC Instantaneous rate Z and oracle S between April 13 09:00 and April 13, 2022 12:00. **Middle**: LT Transaction size. **Bottom**: LP transaction size.

Figure 4: Rate and LP dynamics in AMMs.

2.2.3. Convexity and execution costs

Here, we analyse execution costs implied by the pool reserves and the trading function (5). Figure 5 shows the historical unitary execution costs as a function of the transaction size Δy and the liquidity for all transactions between 1 March and 31 March 2022 in the ETH/USDC pool. The figure shows that for the same transaction size, it is cheaper to trade in a more liquid pool; clearly, as liquidity increases (higher value of κ), execution costs decrease. The figure shows two unitary execution cost curves, i.e., $\Delta y \mapsto |Z - (\varphi(y + \Delta y) - \varphi(y))|$, corresponding to the same depth $\kappa = 3 \times 10^7$. One curve assumes Z = 2,975 or y = 500,000, and the other assumes Z = 3,600 or y = 550,000.

Next, we analyse the geometry of the constant product trading function to understand how unitary execution costs relate to the depth of the pool and the instantaneous rate Z. Figure 6 shows the level function φ for $\kappa=2,500,000$. Every point on the curve gives possible values for x and y that result in the same pool depth. The point O corresponds to the current pool quantities x and y. The slope of the tangent at that point gives the current instantaneous rate $Z=-\varphi'(y)$.

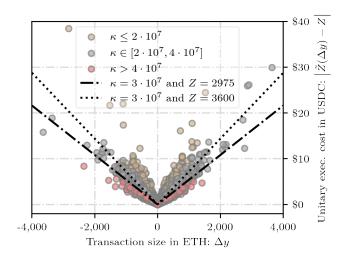


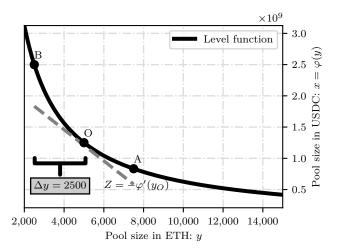
Figure 5: Unitary execution costs for all transactions between 1 March and 31 March, 2022 in the ETH/USDC pool. Negative transaction sizes corresponds to buying ETH, and positive sizes to selling ETH. The unitary execution cost is defined in (4) and is given for the following ranges of the depth $\kappa : [0, 2 \times 10^7], [2 \times 10^7, 4 \times 10^7], \text{ and } [4 \times 10^7, +\infty].$

A change from O to A in Figure 6 is the result of an LT selling 2,500 ETH. A change from O to B is the result of an LT buying 2,500 ETH. The new rates, after these transactions, are given by the slopes at the new points A and B, respectively. When an LT sells $\Delta y = 2,500$ ETH, the unitary execution rate $\Delta x/\Delta y$ is given by the slope of the line (OA). Similarly, the slope of the line (OB) gives the unitary rate for buying Δy . On the other hand, the unitary execution cost is given by the difference between the slope of the lines and the slope of the tangent at point O; the magnitude of this difference depends on the curvature of φ in the neighbourhood of O. This curvature is proportional to the convexity of the level function and can be approximated by the second-order Taylor polynomial $\frac{1}{2} \varphi''(y_0) \Delta y^2$. A higher degree of convexity, i.e., more curvature around point O, does not change the slope of the tangent at point O, but changes the slopes of the lines (OA) and (OB). The convexity of the level function is given by

$$\varphi''(y) = \frac{2 \kappa^2}{y^3} = \frac{2 Z^{3/2}}{\kappa}.$$

Therefore, the execution rate obtained for buying or selling Δy is always less advantageous than the instantaneous rate Z because the level function φ is convex. Clearly, as the depth κ increases, the convexity of the level function is less pronounced. For orders of "small size" one can approximate the unitary execution cost in (4) with

$$|Z - \tilde{Z}(\Delta y)| = \frac{1}{\kappa} Z^{3/2} |\Delta y|, \qquad (6)$$



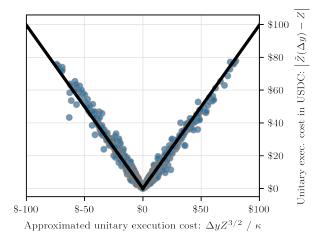


Figure 6: Geometry of the constant product trading function. The figure shows the function $\varphi(y) = x$ where x is the quantity of USDC, and y is the quantity of ETH in the pool.

Figure 7: Scatter plot of transaction costs and the approximation (6) for all transactions in the ETH/USDC pool between 5 May 2021 and 5 May 2022.

or equivalently approximate the execution rate with

$$\tilde{Z}(\Delta y) = Z - \frac{1}{\kappa} Z^{3/2} \, \Delta y \, .$$

To show the extent to which the convexity of the level function captures the execution cost, we use transaction data from the ETH/USDC pool from inception on 5 May 2021 until 5 May 2022, approximately 1,757,000 LT transactions, to compare unitary execution costs incurred by LTs with our approximation in (6). Figure 7 shows a scatter plot of the observed unitary execution costs from transaction data and the approximation $(Z^{3/2}/\kappa) \Delta y$. Recall that negative values of Δy are buy orders and positive values are sell orders. The figure shows that the expression in (6) is a good approximation, and that the convexity of the level function φ can be used to model the unitary execution cost incurred when interacting in CPMMs.

In an optimal trading framework, the investor controls the speed ν_t at which she sends orders to the AMM. Now, assume that the investor trades the quantity $\Delta y_t = \nu_t \, \Delta t$, where Δt is a fixed time-step that determines the investor's frequency of trading and ν_t is fixed during the time interval. The execution rate for Δy is $\tilde{Z}(\Delta y) = Z - Z^{3/2} \, \nu_t \, \Delta t \, / \, \kappa$. Thus, to reflect the investor's pace of trading we write the investor's execution rate as

$$\tilde{Z}_t = Z_t - \frac{\eta}{\kappa} Z_t^{3/2} \nu_t , \qquad (7)$$

where the parameter η scales the execution costs according to the investor's trading frequency.

Due to the convexity of the level function, it is sub-optimal to execute large orders in one trade. An optimal trading framework, similar to those developed for traditional LOB-based markets, should balance the trade-off between execution costs and rate risk. The execution cost in CFMMs is similar to the cost of "walking the book" when trading in LOBs, sometimes referred to as the temporary price impact. The difference between the temporary price impact and the CFMM execution cost is that in the CFMM we have a deterministic closed-form expression for the execution cost as a function of the depth and the rate, both of which are estimated by LTs to estimate execution costs. On the other hand, in LOBs, traders usually rely on historical data analysis and assumptions to obtain an estimate of the execution costs. In LOBs, it is generally assumed that temporary impact is a linear function of the speed of trading where the slope of the function is assumed to be fixed; see Cartea et al. (2015) and Guéant (2016).

As in LOBs, there is uncertainty in execution costs received by LTs. On blockchains, rate and liquidity updates occur at the block validation frequency. For instance, in the *Ethereum* network hosting the most popular AMMs, a block is validated every 13 seconds, on average. Hence, the most up-to-date information available is the previous block's rate and liquidity data. Transactions are grouped and executed in a block which is externally validated. Within the block, the transactions form a queue that determines the priority with which they are executed. Market participants can pay *tips* to gain priority in the block. Finally, every transaction sent on the network pays *gas fees*. Gas is a unit that measures computational effort of network transactions, and all market participants pay gas fees to the network – fees are proportional to the algorithmic complexity of the functions a trader invokes in the AMM. Consequently, there is randomness regarding the exact execution costs.

Below, in Section 3, we propose a trading model (Model I) in AMMs where execution costs are given by the convexity of the level function, where the depth κ does not change over the agent's trading window, and where it is assumed that there is an oracle rate. Model I introduces a method that uses piecewise constant execution costs to obtain a closed-form approximation strategy that adjusts to the convexity of the level function. We use Model I in our performance analysis of Section 5. In the future, an oracle rate may become less relevant because activity in AMMs will increase and significant rate discrepancies will very seldom appear. Thus, in Section 4 we propose Model II for liquid AMMs where the depth κ of the pool is stochastic and where an oracle rate is redundant.

3. Model I: optimal trading with an oracle rate

We consider the problem of an investor who trades in a CPMM and wishes to exchange a large position in asset Y into asset X or to execute a statistical arbitrage in the pair. In both cases, the investor uses rate information from the pool in the CPMM and from another more liquid exchange in which the oracle rate S is the price of Y in terms of that of X. The investor acknowledges that rate formation does not take place in the AMMs.

The depth κ of the pool is assumed to be constant during the execution window [0,T], where T>0. Despite intraday changes in the depth of the pool, this assumption does not have a material effect on the investor's execution problem. Currently, the activity and liquidity in AMMs is such that the depth κ is constant for periods of time which are longer than the trading horizon T considered by the investor.

In the future, activity in AMMs is expected to increase, and so will the informational content of the rates implied by the pool of AMMs. An increase in activity of the pool would affect our modelling choices in two ways. One, the innovations in depth κ will occur more often, so the value of κ cannot be regarded as constant throughout the execution window. Two, the oracle rate becomes redundant because the rates in the AMMs become efficient, so they incorporate all the information available to market participants – i.e., the discrepancies with rates and prices in other trading venues are negligible and economically insignificant. Therefore, below in Section 4 we propose Model II to solve the investor's execution problem with stochastic dynamics for the depth κ and without an external oracle rate.

We proceed with the setup of Model I, for which we fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]})$ satisfying the usual conditions, where \mathbb{F} is the natural filtration generated by the collection of observable stochastic processes that we define below.

The investor must liquidate a position \tilde{y}_0 in asset Y over the period of time [0,T], and her wealth is valued in terms of asset X. We introduce the progressively measurable oracle rate process $(S_t)_{t\in[0,T]}$ with dynamics

$$dS_t = \sigma S_t dW_t, \tag{8}$$

where the volatility parameter σ is a nonnegative constant, and $(W_t)_{t \in [0,T]}$ is a standard Brownian motion.

As discussed above, when an LT sends a trade to the AMM, the rate impact received by the order is encoded in the trading function f(x, y). As in traditional models for optimal execution

⁹Clearly, the investor cannot know when a change in liquidity of the pool occurs. Thus, there is the possibility that within the trading window of the execution programme the value of κ changes.

(see e.g., Cartea et al. (2015) and Guéant (2016)), the rate impact received by the investor's trade is a function of the trading speed. In AMMs, rate impact is a function of the trading speed and of the rate Z and the depth κ . Specifically, we write the difference between the execution rate \tilde{Z} and the instantaneous rate Z as in (7), i.e.,

$$\tilde{Z} - Z = -\frac{\eta}{\kappa} Z^{3/2} \nu. \tag{9}$$

The term on the right-hand side of (9) is the rate impact function of the AMM, which is determined by the convexity of the level function φ , see (6), and depends on the instantaneous rate Z_t of the pool at the time the liquidity taking order is executed. The key difference between the functional form of the execution costs in (9) and those in the equity LOB literature is that in general, the price impact functions proposed for LOB models do not depend on the price of the asset, see Cartea et al. (2015).

Thus, when an LT trades at speed ν , the quantity of Y swapped at every instant in time is given by νdt , so the dynamics of the investor's holdings in asset X are given by

$$d\tilde{x}_t = \tilde{Z}_t \,\nu_t \,dt = \left(Z_t - \frac{\eta}{\kappa} \,Z_t^{3/2} \,\nu_t\right) \,\nu_t \,dt \,. \tag{10}$$

Here, the execution cost is stochastic and its dynamics are known; see Barger and Lorig (2019) and Fouque et al. (2021) for similar frameworks where the price impact of trading is stochastic.

In the absence of market frictions, continuous arbitrage between the oracle rate and the rate quoted in the AMM would make rates converge so that $S_t = Z_t$ at any time t. However, exchanges and AMMs are not frictionless, and the oracle rate is the most efficient rate (i.e., reflects all the information available) so we consider the following dynamics for $(Z_t)_{t \in [0,T]}$:

$$dZ_t = \beta \left(S_t - Z_t \right) dt + \gamma Z_t dB_t , \qquad (11)$$

where $\beta>0$ is the mean-reverting parameter, $\gamma>0$ quantifies the dispersion induced by the trading flow that drives the AMM instantaneous rate Z_t away from the oracle rate S_t , and $(B_t)_{t\in[0,T]}$ is a standard Brownian motion independent of $(W_t)_{t\in[0,T]}$.

3.1. Optimal trading strategy: execution and statistical arbitrage

The investor trades at the speed $(\nu_t)_{t\in[0,T]}$, so her inventory $(\tilde{y}_t)_{t\in[0,T]}$ evolves as

$$d\tilde{y}_t = -\nu_t \, dt \,, \tag{12}$$

where, for simplicity, trading fees are zero. We do not restrict the speed in (12) to be positive; if $\nu > 0$ the investor sells the asset and if $\nu < 0$ the investor buys the asset. When the initial inventory is $\tilde{y}_0 > 0$ (resp. $\tilde{y}_0 = 0$) the investor executes a liquidation (resp. speculation) programme.

The investor maximises her expected terminal wealth in units of X while penalising inventory in Y. The set of admissible strategies is

$$\mathcal{A}_t = \left\{ (\nu_s)_{s \in [t,T]}, \ \mathbb{R}\text{-valued}, \ \mathbb{F}\text{-adapted}, \text{ and } \int_t^T |\nu_s|^2 \, ds < +\infty, \ \mathbb{P}\text{-a.s.} \right\}. \tag{13}$$

We write $\mathcal{A} := \mathcal{A}_0$ and let $\nu \in \mathcal{A}$. The performance criterion of the investor, who trades at speed ν , is the function $u^{\nu} : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ given by

$$u^{\nu}(t, \tilde{x}, \tilde{y}, Z, S) = \mathbb{E}_{t, \tilde{x}, \tilde{y}, Z, S} \left[\tilde{x}_{T}^{\nu} + \tilde{y}_{T}^{\nu} Z_{T} - \alpha \left(\tilde{y}_{T}^{\nu} \right)^{2} - \phi \int_{t}^{T} \left(\tilde{y}_{s}^{\nu} \right)^{2} ds \right], \tag{14}$$

and the investor's value function $u:[0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}_{++}\times\mathbb{R}_{++}\to\mathbb{R}$ is

$$u(t, \tilde{x}, \tilde{y}, Z, S) = \sup_{\nu \in \mathcal{A}} u^{\nu} (t, \tilde{x}, \tilde{y}, Z, S) .$$

The first term on the right-hand side of (14) represents the investor's holdings in asset X at the end of the trading window. The second term represents the investor's earnings from liquidating her remaining inventory at the final time T at rate Z_T . The third term is the 'cost' of liquidating the final inventory \tilde{y}_T at time T. Finally, the last term on the right-hand side of (14) represents a running inventory penalty where the parameter $\phi \geq 0$ quantifies the urgency of the investor to liquidate inventory; the units of ϕ are such that the penalty is in units of X.

In Appendix C, we use the tools of stochastic optimal control to study the optimisation problem in (14). The functional form of the execution costs leads to a system of PDEs, one of which is a semilinear PDE which we cannot solve in closed-form; see (C.9). The optimal trading speed in feedback form is a function of the solution to the semilinear PDE, see (C.6), and one can compute the optimal trading speed with a numerical scheme. However, in our case, the numerical scheme is computationally expensive because the semilinear PDE requires a thin grid and a linearisation iterative method at each time step to transform the nonlinear problem into a sequence of linear problems.

We refer to the optimal strategy obtained with a numerical scheme as the numerical approximation strategy. Clearly, the numerical approximation strategy takes too long both to compute and to implement by the LT in real time. In practice, the profitability of execution and statistical

arbitrage strategies relies on computing the strategy and instructing the AMM within very short periods of time (e.g., milliseconds). Thus, because speed is paramount for LTs, below we derive a strategy in closed-form that can be deployed by the LT in real time. We refer to this strategy as the closed-form approximation strategy. Finally, to assess the precision of the closed-form approximation strategy, Subsection 3.4 compares it with the numerical approximation strategy (C.6) derived in Appendix C.

3.2. Constant impact parameter strategy

To obtain a trading scheme that can be implemented by the LT in real time, we first derive a strategy where the impact parameter of the execution cost is deterministic. We use this strategy as the building block for the LT's closed-form approximation strategy we derive below. Accordingly, we write the execution cost in (9) as

$$\tilde{Z} - Z = -\eta \zeta \nu$$
,

where $\zeta>0$ is the impact parameter and recall that the value of η depends on the investor's trading frequency. With fixed executions costs, the investor can derive a closed-form optimal trading strategy $\left(\nu_t^{\star,\zeta}\right)_{t\in[0,T]}$ for a given value of the parameter ζ .

For each ζ the set of admissible strategies is

$$\mathcal{A}_t^{\zeta} = \left\{ (\nu_s)_{s \in [t,T]}, \ \mathbb{R}\text{-valued}, \ \mathbb{F}\text{-adapted, and} \ \int_t^T |\nu_s|^2 \, ds < +\infty, \ \ \mathbb{P}\text{-a.s.} \right\},$$

and we write $\mathcal{A}^\zeta:=\mathcal{A}_0^\zeta$. We consider an optimal trading problem in which the investor trades at speed $\left(\nu_t^\zeta\right)_{t\in[0,T]}$, so the inventory $\left(\tilde{y}_t^\zeta\right)_{t\in[0,T]}$ evolves as

$$d\tilde{y}_t^{\zeta} = -\nu_t^{\zeta} dt \,,$$

where we do not restrict the speed to be positive, and recall that trading fees are zero.

When an LT trades at speed ν^{ζ} , the quantity of Y swapped at every instant in time is given by $\nu^{\zeta} dt$, so the dynamics of the investor's holdings in asset X are given by

$$d\tilde{x}_t^{\zeta} = \tilde{Z}_t \, \nu_t^{\zeta} \, dt = \left(Z_t - \eta \, \zeta \, \nu_t^{\zeta} \right) \, \nu_t^{\zeta} \, dt \, .$$

Let $\nu^\zeta \in \mathcal{A}^\zeta$. The performance criterion of the investor, who trades at speed ν^ζ , is a function $u^{\nu^\zeta} \colon [0,T] \times \mathbb{R}^4 \to \mathbb{R}$ given by

$$u^{\nu^{\zeta}}(t, \tilde{x}, \tilde{y}, Z, S) = \mathbb{E}_{t, \tilde{x}, \tilde{y}, Z, S} \left[\tilde{x}_{T}^{\nu^{\zeta}} + \tilde{y}_{T}^{\nu^{\zeta}} Z_{T} - \alpha \left(\tilde{y}_{T}^{\nu^{\zeta}} \right)^{2} - \phi \int_{t}^{T} \left(\tilde{y}_{s}^{\nu^{\zeta}} \right)^{2} ds \right] ,$$

and the value function $u_{\zeta} \colon [0,T] \times \mathbb{R}^4 \to \mathbb{R}$ of the investor is given by

$$u_{\zeta}(t, \tilde{x}, \tilde{y}, Z, S) = \sup_{\nu^{\zeta} \in \mathcal{A}^{\zeta}} u^{\nu^{\zeta}}(t, \tilde{x}, \tilde{y}, Z, S) . \tag{15}$$

Next, we solve the above dynamic optimisation problem for each value of $\zeta > 0$ and derive the optimal liquidation strategy. The value function (15) is the unique classical solution to the Hamilton–Jacobi–Bellman (HJB) equation

$$0 = \partial_t w_{\zeta} - \phi \, \tilde{y}^2 + \beta \, (S - Z) \, \partial_Z w_{\zeta} + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} w_{\zeta} + \frac{1}{2} \, \sigma^2 \, S^2 \, \partial_{SS} w_{\zeta}$$

$$+ \sup_{\nu \in \mathbb{R}} \left(\left(\nu \, Z - \eta \, \zeta \, \nu^2 \right) \partial_{\tilde{x}} w_{\zeta} - \nu \, \partial_{\tilde{y}} w_{\zeta} \right) ,$$

$$(16)$$

with terminal condition

$$w_{\zeta}(T, \tilde{x}, \tilde{y}, Z, S) = \tilde{x} + \tilde{y} Z - \alpha \, \tilde{y}^2 \,. \tag{17}$$

The form of the terminal condition (17) suggests the ansatz

$$w_{\zeta}(t, \tilde{x}, \tilde{y}, Z, S) = \tilde{x} + \tilde{y} Z + \theta_{\zeta}(t, \tilde{y}, Z, S)$$

where the first two terms on the right-hand side represent the mark-to-market value of the investor's holdings at the instantaneous rate of the AMM. The last term represents the additional value obtained by the investor when following the optimal strategy. The ansatz is justified by the following proposition, for which a proof is straightforward.

Proposition 1 Assume there exists a function $\theta_{\zeta} \in C^{1,1,2,2}([0,T] \times \mathbb{R}^3)$ which solves

$$0 = \partial_t \theta_{\zeta} - \phi \, \tilde{y}^2 + \beta \, (S - Z) \, (\tilde{y} + \partial_Z \theta_{\zeta}) + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} \theta_{\zeta} + \frac{1}{2} \, \sigma^2 \, S^2 \, \partial_{SS} \theta_{\zeta}$$

$$+ \sup_{\nu \in \mathbb{R}} \left(-\eta \, \zeta \, \nu^2 - \nu \, \partial_{\tilde{y}} \theta_{\zeta} \right) ,$$

$$(18)$$

on $[0,T) \times \mathbb{R}^3$, with terminal condition

$$\theta_{\zeta}(T, \tilde{y}, Z, S) = -\alpha \, \tilde{y}^2 \,. \tag{19}$$

Then, the function $w_{\zeta} \colon [0,T] \times \mathbb{R}^4 \to \mathbb{R}$ defined by

$$w_{\zeta}(t, \tilde{x}, \tilde{y}, Z, S) = \tilde{x} + \tilde{y} Z + \theta_{\zeta}(t, \tilde{y}, Z, S) ,$$

is a solution to (16) on $[0,T) \times \mathbb{R}^4$, with terminal condition (17).

Next, solve the first order condition of the supremum term in (18) to obtain the investor's optimal trading speed in feedback form

$$\nu^{\zeta,\star} = -\frac{1}{2\,\eta\,\zeta}\,\partial_{\tilde{y}}\theta_{\zeta} \,. \tag{20}$$

Substitute (20) into (18) to write

$$\partial_t \theta_{\zeta} = -\phi \, \tilde{y}^2 + \beta \, (S - Z) \, (\tilde{y} + \partial_Z \theta_{\zeta}) + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} \theta_{\zeta} + \frac{1}{2} \, \sigma^2 \, S^2 \, \partial_{SS} \theta_{\zeta}$$

$$+ \frac{1}{4 \, \eta \, \zeta} \, (\partial_{\tilde{y}} \theta_{\zeta})^2 .$$
(21)

Finally, simplify (21) with the ansatz

$$\theta_{\zeta}(t, \tilde{y}, Z, S) = A_{\zeta}(t) \, \tilde{y}^2 + B_{\zeta}(t) \, Z \, \tilde{y} + C_{\zeta}(t) \, \tilde{y} \, S + D_{\zeta}(t) \, \tilde{y} + E_{\zeta}(t) \, Z^2$$

$$+ F_{\zeta}(t) \, S^2 + G_{\zeta}(t) \, Z \, S + H_{\zeta}(t) \, Z + I_{\zeta}(t) \, S + J_{\zeta}(t) \, ,$$

which is justified by the following proposition, for which a proof is straightforward.

Proposition 2 Assume there exist functions $A_{\zeta} \in C^{1}([0,T]), \ B_{\zeta} \in C^{1}([0,T]), \ C_{\zeta} \in C^{1}([0,T]), \ D_{\zeta} \in C^{1}([0,T]), \ E_{\zeta} \in C^{1}([0,T]), \ F_{\zeta} \in C^{1}([0,T]), \ G_{\zeta} \in C^{1}([0,T]), \ H_{\zeta} \in C^{1}([0,T]), \ I_{\zeta} \in C^{1}([0,T]), \ and \ J_{\zeta} \in C^{1}([0,T]) \text{ which solve the system of ODEs}$

$$\begin{cases} A'_{\zeta}(t) &= \phi - \frac{A_{\zeta}(t)^{2}}{\eta \zeta_{N}^{j}} , \\ B'_{\zeta}(t) &= \beta + \beta B_{\zeta}(t) - \frac{A_{\zeta}(t) B_{\zeta}(t)}{\eta \zeta} , \\ C'_{\zeta}(t) &= -\beta - \beta B_{\zeta}(t) - \frac{A_{\zeta}(t) C_{\zeta}(t)}{\eta \zeta} , \end{cases} \text{ and } \begin{cases} F'_{\zeta}(t) &= -\beta G_{\zeta}(t) - \sigma^{2} F_{\zeta}(t) - \frac{C_{\zeta}(t)^{2}}{4 \eta \zeta} , \\ G'_{\zeta}(t) &= -2 \beta E_{\zeta}(t) + \beta G_{\zeta}(t) - \frac{B_{\zeta}(t) C_{\zeta}(t)}{2 \eta \zeta} , \\ H'_{\zeta}(t) &= \beta H_{\zeta}(t) - \frac{B_{\zeta}(t) D_{\zeta}(t)}{2 \eta \zeta} , \\ I'_{\zeta}(t) &= -\beta H_{\zeta}(t) - \frac{C_{\zeta}(t) D_{\zeta}(t)}{2 \eta \zeta} , \\ I'_{\zeta}(t) &= -\beta H_{\zeta}(t) - \frac{C_{\zeta}(t) D_{\zeta}(t)}{2 \eta \zeta} , \end{cases}$$

$$(22)$$

on [0,T), with terminal conditions

$$A_{\zeta}(T) = -\alpha ,$$

$$B_{\zeta}(T) = C_{\zeta}(T) = D_{\zeta}(T) = E_{\zeta}(T) = F_{\zeta}(T) = G_{\zeta}(T) = H_{\zeta}(T) = I_{\zeta}(T) = J_{\zeta}(T) = 0 .$$

Then, the function $\theta_{\zeta}:[0,T]\times\mathbb{R}^3\to\mathbb{R}$ defined by

$$\theta_{\zeta}(t, \tilde{y}, Z, S) = A_{\zeta}(t) \, \tilde{y}^2 + B_{\zeta}(t) \, Z \, \tilde{y} + C_{\zeta}(t) \, \tilde{y} \, S + D_{\zeta}(t) \, \tilde{y} + E_{\zeta}(t) \, Z^2$$

$$+ F_{\zeta}(t) \, S^2 + G_{\zeta}(t) \, Z \, S + H_{\zeta}(t) \, Z + I_{\zeta}(t) \, S + J_{\zeta}(t) \, ,$$

solves (18) over $[0,T) \times \mathbb{R}^3$ with terminal condition (19).

The system of ODEs (22) can be solved sequentially. First, solve the ODE in A_{ζ} , and then solve for B_{ζ} , C_{ζ} , D_{ζ} , E_{ζ} , G_{ζ} , F_{ζ} , H_{ζ} , I_{ζ} , I_{ζ} , respectively. The system in (22) admits a unique solution, which is given by

$$\begin{cases} A_{\zeta}(t) &= \sqrt{\phi \eta \zeta} \tanh \left(\frac{\sqrt{\phi}}{\sqrt{\eta \zeta}} (T - t) + \operatorname{arctanh} \left(-\frac{\alpha}{\sqrt{\phi \eta \zeta}} \right) \right), \\ B_{\zeta}(t) &= \int_{0}^{t} \beta \exp \left(\int_{s}^{t} \left(\beta - \frac{1}{\eta \zeta} A_{\zeta}(u) \right) du \right) ds, \\ C_{\zeta}(t) &= -B_{\zeta}(t), \\ E_{\zeta}(t) &= -\int_{0}^{t} \exp \left(-(\gamma^{2} - 2\beta)(t - s) \right) \frac{1}{4\eta \zeta} B_{\zeta}(s)^{2} ds, \\ F_{\zeta}(t) &= -\int_{0}^{t} \exp \left(-\sigma^{2}(t - s) \right) \left(\beta G_{\zeta}(s) + \frac{1}{4\eta \zeta} C_{\zeta}(s)^{2} \right) ds, \\ G_{\zeta}(t) &= -\int_{0}^{t} \exp \left(\beta (t - s) \right) \left(2\beta E_{\zeta}(s) - \frac{1}{2\eta \zeta} B_{\zeta}(s)^{2} \right) ds, \\ H_{\zeta}(t) &= D_{\zeta}(t) = I_{\zeta}(t) = J_{\zeta}(t) = 0. \end{cases}$$

The optimal trading strategy in feedback form (20) over $[0,T] \times \mathbb{R}^3$ is now given by

$$\nu^{\zeta,\star}(t,\tilde{y},Z,S) = -\frac{1}{\eta\zeta}A_{\zeta}(t)\,\tilde{y} + \frac{1}{2\eta\zeta}B_{\zeta}(t)\,(S-Z). \tag{23}$$

The first term on the right-hand side of (23) is the optimal liquidation rate in the continuous Almgren-Chriss model. The second term is an arbitrage component; it accounts for the spread between the instantaneous rate Z and the oracle rate S.

3.3. The closed-form approximation strategy

Here, we use a family of closed-form strategies of the type in (23) to derive a piecewise-defined trading strategy which approximates the optimal trading speed in feedback form (C.11). Specifically, we partition the space of the rate Z into strips and define a piecewise strategy which uses a different impact parameter ζ in each different strip. Finally, we show that as the width of the strip becomes arbitrarily small, the piecewise strategy converges to the closed-form approximation strategy.

Let $\{Z_0, \ldots, Z_N\}$ be a partition of $[\underline{Z}, \overline{Z}]$, where $0 < \underline{Z} < \overline{Z}$, so that for each $N \in \mathbb{N}$ and $j \in \{0, \ldots, N\}$ we define

$$Z_j^N := \underline{Z} + \frac{j}{N} \left(\overline{Z} - \underline{Z} \right) \quad \text{and} \quad \zeta_j^N = \frac{1}{\kappa} \left(Z_j^N \right)^{3/2} .$$
 (24)

In the remainder of this section, we simplify the notation and use $\nu^{\star,j,N}$ instead of $\nu^{\zeta_j^N,\star}$ to denote the optimal trading strategy with impact parameter ζ_j^N .

Note that whenever Z is arbitrarily close to Z_j^N the impact parameter in (9) can be approximated by ζ_j^N . Thus, to construct the approximate trading strategy, we first define a strategy ν_N^* that uses the closed-form optimal trading speed $\nu^{\star,j,N}$ to approximate the optimal trading speed whenever the rate is close to ζ_j^N . We define the piecewise-defined trading ν_N^* : $[0,T] \times \mathbb{R}^3 \to \mathbb{R}$

$$\begin{split} \nu^{\star,N}\left(t,\tilde{y},Z,S\right) = & \ \nu^{\star,0,N}\left(t,\tilde{y},Z,S\right) \mathbbm{1}_{Z < Z_1^N} + \sum_{j=1}^{N-1} \nu^{\star,j,N}\left(t,\tilde{y},Z,S\right) \mathbbm{1}_{Z \in [Z_j^N,Z_{j+1}^N)} \\ & + \nu^{\star,N,N}\left(t,\tilde{y},Z,S\right) \mathbbm{1}_{Z > Z_N^N} \,. \end{split}$$

The strategy $\nu^{\star,N}$ (t,\tilde{y},Z,S) has first-type discontinuity points; specifically, it is discontinuous over $[0,T]\times\mathbb{R}\times\{Z_j^N\}\times\mathbb{R}$ for each $j\in\{1,\ldots,N\}$ because for each $\left(t,\tilde{y},Z_{j+1}^N,S\right)\in[0,T]\times\mathbb{R}^2$ we have $\nu^{\star,j,N}$ $\left(t,\tilde{y},Z_{j+1}^N,S\right)\neq\nu^{\star,j+1,N}$ $\left(t,\tilde{y},Z_{j+1}^N,S\right)$.

The theorem below shows how to partition $[\underline{Z}, \overline{Z}]$ to make the discontinuities in $\nu^{\star,N}$ (t, \tilde{y}, Z, S) arbitrarily small. Furthermore, when the distance between points in the partition becomes sufficiently small, the sequence of piecewise-defined optimal strategies $\{\nu^{\star,N}\}_{N\in\mathbb{N}}$ converges uniformly to a continuous closed-form approximation strategy which we use in our performance study of Section 5.

Theorem 1 For each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\max_{j=1,\dots,N} \left| \nu^{\star,j,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) - \nu^{\star,j+1,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) \right| < \varepsilon. \tag{25}$$

Furthermore, for each $N \in \mathbb{N}$, let $\tilde{\nu}^{\star,N} := \nu^{\star,N}|_{[0,T] \times \mathbb{R} \times [\underline{Z},\overline{Z}] \times \mathbb{R}}$. Then, the sequence $\{\tilde{\nu}^{\star,N}\}$ converges to $\tilde{\nu}^{\star}$ uniformly in $[0,T] \times \mathbb{R} \times [\underline{Z},\overline{Z}] \times \mathbb{R}$, where

$$\tilde{\nu}^{\star}(t, \tilde{y}, Z, S) = -\frac{\kappa}{\eta} Z^{-3/2} A(t, Z) \, \tilde{y} + \frac{\kappa}{2 \, \eta} Z^{-3/2} B(t, Z) \, (S - Z) \,, \tag{26}$$

and

$$A(t,Z) = \sqrt{\frac{\phi \eta Z^{3/2}}{\kappa}} \tanh \left(\frac{\sqrt{\phi \kappa}}{\sqrt{\eta Z^{3/2}}} t + \operatorname{arctanh} \left(-\frac{\alpha \sqrt{\kappa}}{\sqrt{\phi \eta Z^{3/2}}} \right) \right) ,$$

$$B(t,Z) = \int_{t}^{T} \beta \exp \left(\int_{s}^{t} \left(\beta - \frac{\kappa}{\eta Z^{3/2}} A(u,Z) \right) du \right) ds .$$
(27)

For a proof see Appendix A.

3.4. Comparison with the numerical approximation strategy

In this subsection, we use a Euler scheme to compute the numerical approximation strategy in (C.9), where execution costs are not piecewise constant. We compare this numerical solution with the closed-form approximation strategy in (26).

The numerical approximation strategy uses a three-dimensional grid; one dimension is time, one is the rate Z, and one is the oracle rate S. We consider a wide interval for the values of Z and S, and use a Neumann boundary condition for both space variables (the derivatives are zero at the boundaries) because the value function should not vary significantly for values of the rates Z and S that are considerably far from the observed rates. For the semilinear PDE in (C.9), we use the Picard iterative method to linearise the problem at every time level of the grid. More precisely, at each time level, we use Picard iterations to linearise the term that is quadratic in the unknown solution, which is replaced by the product of the unknown solution and the most recent computed solution obtained at the previous time level. The iterative approximation process is carried on until the difference between two consecutive numerical solutions is smaller than a given threshold.

Figure 8 compares the numerical approximation strategy (C.11) and the closed-form approximation strategy (26) for different values of the inventory, the rate Z, and the oracle rate S. Figure 8 indicates that both the closed-form and the numerical approximation strategies are significantly close and both capture the same financial effects. In particular, the strategies clearly depend on the spread S-Z and the inventory \tilde{y} . When the LT has zero inventory, the strategy is mostly speculative because the strategy buys asset Y when the oracle rate S is above the rate Z, and sells otherwise; recall that the investor buys asset Y when the trading speed ν is positive, and sells otherwise. When the LT has a positive (negative) inventory, the optimal strategy buys (sells) asset Y only when S is significantly higher (lower) than Z.

Figure 8 shows that the absolute difference between both strategies increases as the difference between the rate Z and S increases, and the difference is minimal when the rates are equal. In practice, by arbitrage, the rates Z and S are aligned by LTs in the pool, so differences between both rates are small. For instance, with the market data we use in Section 5, the average absolute difference between both rates is 2 USD (0.07%) in the liquid pool and 10 USD (0.37%) in the

illiquid pool we consider. Finally, it is unclear whether the difference between the optimal speed obtained with a numerical scheme and the closed-form approximation strategy shown in the lower panels of Figure 8 stem from the numerical approximation or from the closed-form approximation strategy, but in either case, the difference has no bearing on the results we discuss in Section 5.

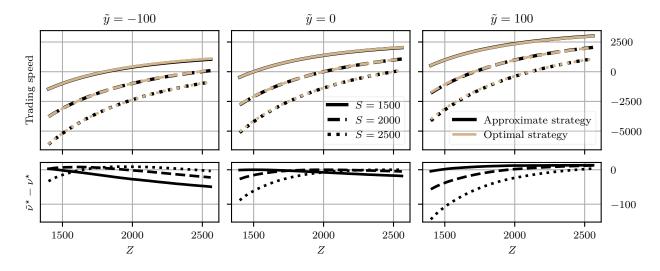


Figure 8: Comparison of the optimal speed (C.11) obtained with a numerical scheme and the closed-form approximation strategy in (27). Model parameters: $T=0.1,\,\sigma=0.03,\,\gamma=0.02,\,\beta=1,\,\alpha=5,\,\kappa=10^7,\,\eta=1,\,\phi=10^{-5},$ and $Z_0=S_0=2000$. Inventory is $\tilde{y}=-100$ (left panel), $\tilde{y}=0$ (middle panel), and $\tilde{y}=100$ (right panel).

4. Model II: optimal execution without oracle rate

In this section, we consider the problem of an investor who wants to exchange a large position in asset Y into asset X in a CPMM. The key differences between Model II and Model I of Section 3 are that in Model II the AMM rates are efficient, so the investor does not use an oracle rate from another venue, and the depth of the pool is stochastic.

The investor must liquidate a large position in asset Y in a CPMM over a period of time [0,T], where T>0. Asset Y and the wealth of the investor are valued in terms of asset X. The CPMM is very active and with concentrated liquidity, so the depth κ exhibits very frequent and random updates. The instantaneous rate $(Z_t)_{t\in[0,T]}$ evolves as

$$dZ_t = \gamma Z_t dB_t$$
,

and the dynamics of the depth process $(\kappa_t)_{t \in [0,T]}$ are given by

$$d\kappa_t = \varsigma \, \kappa_t \, dL_t \; .$$

Here, $(B_t)_{t\in[0,T]}$ and $(L_t)_{t\in[0,T]}$ are independent standard Brownian motions and ς is the volatility of the depth κ . The cash process $(\tilde{x}_t)_{t\in[0,T]}$ satisfies

$$d\tilde{x}_t = \left(Z_t - \eta \frac{Z_t^{3/2}}{\kappa_t} \nu_t \right) \nu_t dt.$$

4.1. Performance criterion and value function

For each $(t, \tilde{x}, \tilde{y}, Z, \kappa) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$, and for each admissible control $\nu \in \mathcal{A}$ the performance criterion of the investor is given by

$$u^{\nu}(t,\tilde{x},\tilde{y},Z,\kappa) = \mathbb{E}_{t,\tilde{x},\tilde{y},Z,\kappa} \left[\tilde{x}_{T}^{\nu} + \tilde{y}_{T}^{\nu} Z_{T} - \alpha \left(\tilde{y}_{T}^{\nu} \right)^{2} - \phi \int_{t}^{T} \left(\tilde{y}_{s}^{\nu} \right)^{2} ds \right] , \qquad (28)$$

and the value function is

$$u(t, \tilde{x}, \tilde{y}, Z, \kappa) = \sup_{\nu \in \mathcal{A}} u^{\nu}(t, \tilde{x}, \tilde{y}, Z, \kappa) . \tag{29}$$

4.2. The dynamic programming equation

The value function (29) is the unique classical solution to the HJB equation

$$0 = \partial_t w - \phi \, \tilde{y}^2 + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} w + \frac{1}{2} \, \varsigma^2 \, \kappa^2 \, \partial_{\kappa \kappa} w + \sup_{\nu \in \mathbb{R}} \left(\left(\nu \, Z - \eta \, \frac{Z^{3/2}}{\kappa} \, \nu^2 \right) \partial_{\tilde{x}} w - \nu \, \partial_{\tilde{y}} w \right) 30$$

with terminal condition

$$w(T, \tilde{x}, \tilde{y}, Z, \kappa) = \tilde{x} + \tilde{y}Z - \alpha \tilde{y}^2. \tag{31}$$

The terminal condition (31) suggests the ansatz

$$w(t, \tilde{x}, \tilde{y}, Z, \kappa) = \tilde{x} + \tilde{y} Z + \theta(t, \tilde{y}, Z, \kappa),$$

which we justify by the following proposition, for which a proof is straightforward.

Proposition 3 Assume there exists a function $\theta \in C^{1,1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++})$ that solves

$$0 = \partial_t \theta - \phi \,\tilde{y}^2 + \frac{1}{2} \,\gamma^2 \,Z^2 \,\partial_{ZZ}\theta + \frac{1}{2} \,\varsigma^2 \,\kappa^2 \,\partial_{\kappa\kappa}\theta + \sup_{\nu \in \mathbb{R}} \left(-\eta \,\frac{Z^{3/2}}{\kappa} \,\nu^2 - \nu \,\partial_{\tilde{y}}\theta \right) \,, \tag{32}$$

on $[0,T) \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$, with terminal condition

$$\theta(T, \tilde{y}, Z, \kappa) = -\alpha \, \tilde{y}^2 \,. \tag{33}$$

Then, the function $w: [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ defined by

$$w(t, \tilde{x}, \tilde{y}, Z, \kappa) = \tilde{x} + \tilde{y}Z + \theta(t, \tilde{y}, Z, \kappa)$$
(34)

is a solution to (30) on $[0,T) \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ with terminal condition (31).

Next, solve the first order condition in (32) to obtain the investor's trading speed in feedback form

$$\nu^* = -\frac{\kappa}{2\,\eta}\,\partial_{\tilde{y}}\theta\,Z^{-3/2}\,.$$
(35)

Substitute (35) into (32) to write

$$0 = \partial_t \theta - \phi \,\tilde{y}^2 + \frac{1}{2} \,\gamma^2 \,Z^2 \,\partial_{ZZ}\theta + \frac{1}{2} \,\varsigma^2 \,\kappa^2 \,\partial_{\kappa\kappa}\theta + \frac{\kappa}{4 \,\eta} \,\left(\partial_{\tilde{y}}\theta\right)^2 \,Z^{-3/2} \,. \tag{36}$$

Finally, simplify (36) with the ansatz

$$\theta(t, \tilde{y}, Z, \kappa) = \theta_0(t, Z, \kappa) + \theta_1(t, Z, \kappa) \,\tilde{y} + \theta_2(t, Z, \kappa) \,\tilde{y}^2 \,, \tag{37}$$

which is justified by the following proposition, for which a proof is straightforward.

Proposition 4 Assume there exist functions $\theta_0 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, $\theta_1 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, and $\theta_2 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$ which solve the system of PDEs

$$\begin{cases}
0 = \partial_t \theta_2 - \phi + \frac{1}{2} \gamma^2 Z^2 \partial_{ZZ} \theta_2 + \frac{1}{2} \varsigma^2 \kappa^2 \partial_{\kappa\kappa} \theta_2 + \frac{\kappa}{\eta} \theta_2^2 Z^{-3/2}, \\
0 = \partial_t \theta_1 + \frac{1}{2} \gamma^2 Z^2 \partial_{ZZ} \theta_1 + \frac{1}{2} \varsigma^2 \kappa^2 \partial_{\kappa\kappa} \theta_1 + \frac{\kappa}{\eta} \theta_1 \theta_2 Z^{-3/2}, \\
0 = \partial_t \theta_0 + \frac{1}{2} \gamma^2 Z^2 \partial_{ZZ} \theta_0 + \frac{1}{2} \varsigma^2 \kappa^2 \partial_{\kappa\kappa} \theta_0 + \frac{\kappa}{4 \eta} \theta_1^2 Z^{-3/2},
\end{cases} (38)$$

on $[0,T) \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ with terminal conditions

$$\theta_2(T, Z, \kappa) = -\alpha$$
, $\theta_1(T, Z, \kappa) = 0$, and $\theta_0(T, Z, \kappa) = 0$. (39)

Then, the function $\theta \colon [0,T] \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ defined by

$$\theta(t, \tilde{y}, Z, \kappa) = \theta_0(t, Z, \kappa) + \theta_1(t, Z, \kappa) \, \tilde{y} + \theta_2(t, Z, \kappa) \, \tilde{y}^2 \,,$$

solves (36) over $[0,T) \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ with terminal condition (33).

The optimal strategy in feedback form (35) is now given by

$$\nu^* = -\frac{\kappa}{2\eta} \left(2\,\theta_2\,\tilde{y} + \theta_1 \right) \, Z^{-3/2} \,. \tag{40}$$

The system of PDEs in (38) can be solved sequentially as follows. Solve the first PDE in the system to obtain θ_2 . Substitute θ_2 is the second and third equations of the system so the PDEs in θ_1 and θ_0 become linear. We cannot solve the semilinear PDE in θ_2 in closed-form, and providing an existence result is out of the scope of this work. However, Theorem 2 provides *a priori* estimates for θ_2 and suggests the system of PDEs is well-behaved. The proof is omitted as it uses the same arguments as those in the proof of Theorem 3.

Theorem 2 Let $\theta_0 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, $\theta_1 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, and $\theta_2 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$ be a solution to the system of PDEs (38) with terminal condition (39). Define θ in (37) and define w in (34). Assume that for all $(t, \tilde{x}, \tilde{y}, Z, \kappa) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ we have

$$u^{\nu}(t, \tilde{x}, \tilde{y}, Z, \kappa) \leq w(t, \tilde{x}, \tilde{y}, Z, \kappa),$$

where u^{ν} is defined in (28), and that equality is obtained for the optimal control $(\nu^{\star})_{t \in [0,T]}$ in feedback form in (40). Then θ_2 has the following bounds

$$-\alpha - \phi\left(T - t\right) \le \theta_2\left(t, \tilde{x}, \tilde{y}, Z, \kappa\right) \le 0, \quad \forall \left(t, \tilde{x}, \tilde{y}, Z, \kappa\right) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}.$$

5. Performance of strategies

We study the performance of two versions of the closed-form approximation strategy in (26) and (27) of Model I; see Section 3. One version focuses on liquidating a large position in one asset and the other uses the lead-follow relationship between the oracle and AMM rates to execute a statistical arbitrage. We use the Uniswap v3 data of Section 2 for the liquid pool ETH/USDC and the illiquid pool ETH/DAI. We account for AMM and gas fees and assume that the orders sent by the investor do not impact the dynamics of the pools.

We use in-sample data to estimate model parameters and use out-of-sample data to execute the strategies. For in-sample data, we use a window of 24 hours prior to the start of the trading programme for both pools. For out-of-sample data, we use windows of 2 and 12 hours when the investor trades in the liquid and illiquid pools, respectively. To measure performance, we use rolling time windows for estimation and execution, between 1 July 2021 and 5 May 2022, to carry out this procedure. Specifically, after every execution programme, we shift both windows by 2

and 12 hours for the liquid and illiquid pools, respectively, and repeat the same procedure, i.e., estimate parameters with in-sample data and trade with out-of-sample data. We remark that we do not simulate rates, we use those of the AMM and Binance, and execution costs are those the trades would have received. In total, we run 3,635 and 607 execution programmes for ETH/USDC and ETH/DAI, respectively.

We proceed as follows. Subsection 5.1 describes how parameter estimates are obtained and showcases the performance of the liquidation strategy. Subsection 5.2 discusses the use of model I for statistical arbitrage, and showcases the performance.

5.1. Liquidation strategy

We describe how to estimate the in-sample model parameters for every run of the liquidation strategy. For rate dynamics, the investor performs OLS regressions on the discretised versions of (8) and (11):

$$\Delta \log S_t = -\frac{\sigma^2}{2} \Delta t + \sigma \sqrt{\Delta t} v_t,$$

$$\Delta \log Z_t = -\frac{\gamma^2}{2} \Delta t + \beta \left(\frac{S_t - Z_t}{Z_t}\right) \Delta t + \gamma \sqrt{\Delta t} \epsilon_t,$$
(41)

where $\{\epsilon_t, \upsilon_t\}$ are error terms. Here, the size of the time-step Δt is the frequency of the liquidity taking orders (from all LTs) that arrive in the pool during the estimation period.

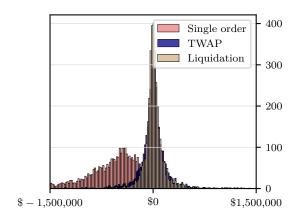
For the liquidation strategy, we target a participation rate of 50% of the observed hourly volume to set the initial inventory, which is liquidated by the investor over the trading window at the same frequency as the observed average trading frequency over the in-sample estimation period. The investor's trading frequency determines the value of the parameter η in (7), which scales the execution costs.

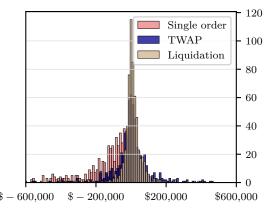
The value of the other model parameters are as follows. The value of the parameter κ in the trading speed (26) is the last observed depth of the pool before the start of the execution. The value of the running inventory parameter ϕ is kept constant for all runs. The value of the terminal penalty parameter α is arbitrarily large to enforce full liquidation of outstanding inventory by the end of the trading horizon. Table 1 shows the parameter values we use for all strategy runs. Finally, as a more detailed example, Appendix B describes parameter estimation and performance for a specific run of the liquidation strategy.

We benchmark the performance of the liquidation strategy with two strategies: TWAP, which consists in trading at a constant rate; and a single order execution strategy, which consists in executing the entire order at the beginning of the execution window. The market rates at the time

	ETH/USDC	ETH/DAI
\overline{T}	0.083 days	0.5 days
ϕ	$0.005~\mathrm{USDC}\cdot\mathrm{ETH}^{-2}$	$0.05~\mathrm{DAI}\cdot\mathrm{ETH}^{-2}$
α	$10 \text{ USDC} \cdot \text{ETH}^{-2}$	$10\mathrm{DAI}\cdot\mathrm{ETH}^{-2}$

Table 1: Values of the liquidation model parameters.





- (a) Distribution of gross PnL for all strategies for ETH/USDC (3,635 executions).
- (b) Distribution of gross PnL of strategies for ETH/DAI (607 executions).

Figure 9: PnL distribution.

of trading are used to compute the execution costs for all strategies. Gas fees are 5 USD per transaction, regardless of transaction size. On the other hand, AMM fees depend on transaction size, and here we impute a 0.01% fee to the value of every transaction. Figure 9 depicts the distribution of the gross PnL for each execution programme, which is given by $\tilde{x}_T + \tilde{y}_T Z_T - \tilde{y}_0 Z_0$. Tables 2 and 3 show the average and standard deviation of the gross PnLs, the number of transactions, the gas fees, and the AMM fees. ¹⁰

	Gross avg. PnL	Std. dev.	Avg. num. trades	Avg. fees
Single order	-956,298	1,963,014	1	2,538
TWAP	3998	217,001	439	10,529
Liquidation	27,185	288,518	439	11,885

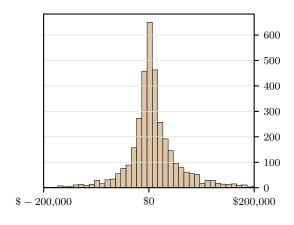
Table 2: Performance and fees for ETH/USDC (3,635 executions). The Average PnL does not include fees

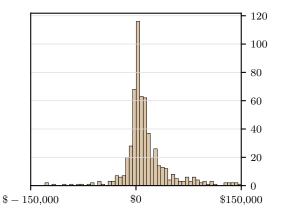
	Gross avg. PnL	Std. dev.	Avg. num. trades	Avg. fees
Single order	-233,390	428,688	1	634
TWAP	1,875	170,008	108	1,217
Liquidation	12,240	63,605	108	1,782

Table 3: Performance and fees for ETH/DAI (607 executions). The Average PnL does not include fees

Figure 9 and Tables 2 and 3 show that liquidating all the inventory in one trade is sub-optimal

¹⁰Gross PnL, as opposed to net PnL, is computed without the AMM fees and gas fees paid by the investor.





- (a) Distribution of gross PnL for ETH/USDC.
- (b) Distribution of gross PnL for ETH/DAI.

Figure 10: Statistical arbitrage PnL distribution.

	Gross avg. PnL	Std. dev.	Avg. num. trades	Avg. fees
Speculative				
strategy	22,693	190,789	439	7,111

	Gross avg. PnL	Std. dev.	Avg. num. trades	Avg. fees
Speculative				
strategy	20,886	54,043	108	2,082

Table 4: Performance and fees for the pair ETH/USDC (3,635 executions). The Average PnL does not include fees

Table 5: Performance and fees for the pair ETH/DAI (607 executions). The Average PnL does not include fees

compared with the other strategies due to the high execution costs of the large order. In both cases, our model outperforms TWAP in terms of the ratio between performance, net of fees, and risk measured by the standard deviation. Key to the outperformance is that the liquidation strategy uses the rates in Binance as a trading signal.

5.2. Speculative strategy

We consider the same setup as before, i.e., the in-sample estimation and out-of-sample execution. Here, the investor arbitrages the AMM. To this end, the investor starts with zero inventory in Y and sets the values of ϕ to $0.001~\rm USDC \cdot ETH^{-2}$ and $0.01~\rm DAI \cdot ETH^{-2}$ for the liquid and illiquid pools, respectively. The strategy profits from the oracle rate as a predictive signal. Figure 10 depicts the distribution of the gross PnL for each run and Tables 4 and 5 show the average and standard deviation of the gross PnLs, the number of transactions, and the estimated AMM and gas fees. Note that the speculative strategy is more profitable in the illiquid pool due to a larger discrepancies between the oracle and the instantaneous pool rate, leading to more arbitrage opportunities.

6. Conclusions

In this work, we used Uniswap v3 data to analyse rate, liquidity, and execution costs of CP-MMs. We proposed two models for optimal trading in CPMMs. In the first model, we assumed that prices are formed in an alternative venue and the liquidity provided in the CPMM remains constant for relevant periods of time. In the second model, price formation takes place in the pool and the liquidity provided in the pool is stochastic. Finally, we used in-sample estimation of model parameters and out-of-sample market data to test the performance of the strategies, so our results do not rely on simulations. We showed that our strategy considerably outperforms TWAP and a strategy that consists in sending a single large order. We also showed that there are significant arbitrage opportunities between Binance and AMM rates.

Appendix A. Proof of Theorem 1

Recall that for each fixed values of N and j

$$\nu^{\star,j,N}(t,\tilde{y},Z,S) = -\frac{1}{\eta \,\zeta_N^j} A_{j,N}(t) \,\tilde{y} + \frac{1}{2 \,\eta \,\zeta_N^j} B_{j,N}(t) (S-Z) \,,$$

where

$$\zeta_N^j \coloneqq \frac{1}{\kappa} \left(Z_j^N \right)^{3/2} \,,$$

and

$$\begin{split} A_{j,N}(t) &\coloneqq A_{\zeta_j^N}(t) = \sqrt{\phi\,\eta\,\zeta_N^j} \tanh\left(\frac{\sqrt{\phi}}{\sqrt{\eta\,\zeta_N^j}}t + \operatorname{arctanh}\left(-\frac{\alpha}{\sqrt{\phi\,\eta\,\zeta_N^j}}\right)\right)\;,\\ B_{j,N}(t) &\coloneqq B_{\zeta_j^N}(t) = \int_0^t \beta \exp\left(\int_s^t \left(\beta - \frac{1}{\eta\,\zeta_N^j}A_{j,N}(u)\right)du\right)ds\;. \end{split}$$

Moreover, recall that

$$A(t,Z) = \sqrt{\frac{\phi \eta Z^{3/2}}{\kappa}} \tanh \left(\frac{\sqrt{\phi \kappa}}{\sqrt{\eta Z^{3/2}}} t + \operatorname{arctanh} \left(-\frac{\alpha \sqrt{\kappa}}{\sqrt{\phi \eta Z^{3/2}}} \right) \right) ,$$

$$B(t,Z) = \int_0^t \beta \exp \left(\int_s^t \left(\beta - \frac{\kappa}{\eta Z^{3/2}} A(u,Z) \right) du \right) ds .$$

To prove (25), take (t, \tilde{y}, S) and write

$$\begin{split} \left| \nu^{\star,j,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) - \nu^{\star,j+1,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) \right| \\ &= \left| -\frac{1}{\eta \zeta_N^j} A_{j,N}(t) \, \tilde{y} + \frac{1}{\eta \zeta_N^{j+1}} A_{j+1,N}(t) \, \tilde{y} + (S - Z_{j+1}^N) \left(\frac{1}{2 \, \eta \, \zeta_N^j} B_{j,N}(t) - \frac{1}{2 \, \eta \, \zeta_N^{j+1}} B_{j+1,N}(t) \right) \right| \\ &\leq \frac{|\tilde{y}|}{\eta} \left| -\frac{1}{\zeta_N^j} A_{j,N}(t) + \frac{1}{\zeta_N^{j+1}} A_{j+1,N}(t) \right| + \frac{|S| + \overline{Z}}{\eta} \left| \frac{1}{2 \, \zeta_N^j} B_{j,N}(t) - \frac{1}{2 \, \zeta_N^{j+1}} B_{j+1,N}(t) \right| \\ &= \frac{|\tilde{y}|}{\eta} \left| -\frac{\kappa}{\left(Z_j^N\right)^{3/2}} A\left(t, Z_j^N \right) + \frac{\kappa}{\left(Z_{j+1}^N\right)^{3/2}} A\left(t, Z_{j+1}^N \right) \right| \\ &+ \frac{|S| + \overline{Z}}{2 \, \eta} \left| \frac{\kappa}{\left(Z_j^N\right)^{3/2}} B\left(t, Z_j^N \right) - \frac{\kappa}{\left(Z_{j+1}^N\right)^{3/2}} B\left(t, Z_{j+1}^N \right) \right|. \end{split}$$

Observe that for a fixed $t \in [0, T]$ the functions

$$Z\mapsto rac{\kappa}{Z^{3/2}}A\left(t,Z
ight) \ \ ext{and} \ \ Z\mapsto rac{\kappa}{Z^{3/2}}B\left(t,Z
ight)$$

are uniformly continuous on $[\underline{Z},\overline{Z}]$ because they are both compositions of continuous functions defined over a closed interval. By definition of the partition in (24), $|Z_j^N - Z_{j+1}^N| = 1/N$ so for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\max_{j=1,\dots,N} \left| \nu^{\star,j,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) - \nu^{\star,j+1,N} \left(t, \tilde{y}, Z_{j+1}^N, S \right) \right| \le \varepsilon.$$

To prove that $\{\tilde{\nu}^{\star,N}\}$ converges uniformly to $\tilde{\nu}^{\star}$ in $[0,T] \times \mathbb{R} \times \left[\underline{Z},\overline{Z}\right] \times \mathbb{R}$, take $(t,\tilde{y},Z,S) \in [0,T] \times \mathbb{R}^2 \times \left[\underline{Z},\overline{Z}\right] \times \mathbb{R}$, and $N \in \mathbb{N}$, and observe that there exists $\overline{j} \in \{1,\ldots,N\}$ such that $Z \in \left[Z_{\overline{j}}^N, Z_{\overline{j}+1}^N\right)$ and thus

$$\begin{split} \left| \tilde{\nu}_{N}^{*} \left(t, \tilde{y}, Z, S \right) - \tilde{\nu}^{*} \left(t, \tilde{y}, Z, S \right) \right| \\ &= \left| \tilde{\nu}_{\bar{j}, N}^{*} \left(t, \tilde{y}, Z, S \right) - \tilde{\nu}^{*} \left(t, \tilde{y}, Z, S \right) \right| \\ &= \left| -\frac{1}{\eta \zeta_{N}^{\bar{j}}} A_{\bar{j}, N}(t) \, \tilde{y} + \frac{\kappa}{\eta \, Z^{3/2}} A(t, Z) \, \tilde{y} + (S - Z) \left(\frac{1}{2 \, \eta \, \zeta_{N}^{\bar{j}}} B_{\bar{j}, N}(t) - \frac{\kappa}{2 \, \eta \, Z^{3/2}} B(t, Z) \right) \right| \\ &\leq \frac{\left| \tilde{y} \right| \, \kappa}{\eta \, Z^{3/2}} \left| -A \left(t, Z_{\bar{j}}^{N} \right) + A \left(t, Z \right) \right| + \frac{\left| S \right| + \overline{Z}}{2 \, \eta \, Z^{3/2}} \, \kappa \, \left| B \left(t, Z_{\bar{j}}^{N} \right) - B \left(t, Z \right) \right| \, . \end{split}$$

The uniform convergence of $\{\tilde{\nu}_N^*\}$ to $\tilde{\nu}^*$ follows from the uniform continuity of A(t,Z) and B(t,Z) on $[0,T]\times \left[\underline{Z},\overline{Z}\right]$.

Appendix B. Example for the liquidation strategy

In this appendix, we describe the parameters and strategy performance for a specific run of the liquidation strategy. Assume the investor will start trading at noon on 16 March 2022, so she uses the data between noon 15 March 2022 and noon 16 March to estimate model parameters.

For the 24 hours before noon 16 March 202, there are, on average, one liquidity taking order every 13 seconds in the liquid pool and one every 360 seconds in the illiquid pool; i.e., the time steps in the regressions (41) are $\Delta t = 13$ for ETH/USDC and $\Delta t = 360$ for ETH/DAI. Table B.6 shows parameter estimates.

	ETH/USDC	ETH/DAI
$\hat{\sigma}$	$0.045~{\rm day}^{-1/2}$	$0.053~{\rm day}^{-1/2}$
$\hat{\gamma}$	$0.034~{\rm day}^{-1/2}$	$0.027~{\rm day}^{-1/2}$
\hat{eta}	$657.9~{\rm day}^{-1}$	$14.78~{\rm day}^{-1}$

Table B.6: Parameter estimates for dynamics of Z and S with data between noon 15 March 2022 and noon 16 March 2022.

The parameter η of the execution costs in (7) is also set to $13\,\mathrm{seconds} = 17.3 \times 10^{-5}\,\mathrm{days}$ and $360\,\mathrm{seconds} = 41 \times 10^{-4}\,\mathrm{days}$ for the liquid and illiquid pool, respectively. The number of observed transactions in the in-sample data is approximately 238,039 ETH and 4,031 ETH in the liquid and illiquid pool, respectively. Thus the investors' target is to liquidate 14,877 and 1,007 units of ETH within 2 and 12 hours in the ETH/USDC and ETH/DAI pools, respectively. Table B.7 summarises all the parameters used to run our strategy.

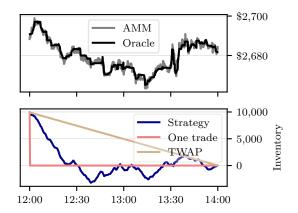
	ETH/USDC	ETH/DAI
κ_0	22,561,783	1,666,175
\tilde{y}_0	14,877 ETH	1,007 ETH
\tilde{S}_0	2,689.2 USDC	2,686.09 DAI
$ ilde{Z}_0$	2,690.77 USDC	2,694.04 DAI
η	17.3×10^{-5} days	$41 \times 10^{-4} \text{ days}$

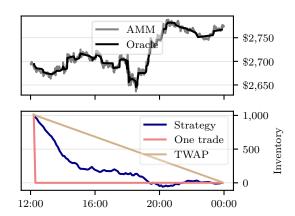
Table B.7: Values of model parameters.

Figure B.11 shows the instantaneous and oracle rates and the inventories of the strategies during the execution, for both ETH/USDC and ETH/DAI. Figure B.11 clearly showcases the difference between the strategies. In particular, the liquidation strategy is adaptive and trades on the difference between the two rates S and Z during the liquidation programme. Figure B.12 shows how the difference $S_t - Z_t$ drives the trading speed ν_t . The oracle rate is used as a predictive signal for future moves of the instantaneous rate.

Appendix C. Model I with stochastic convexity cost

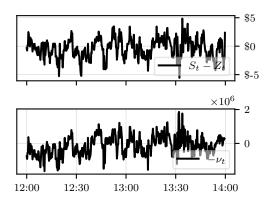
We consider the same problem as that of Section 3.1 where an investor exchanges a large position in asset Y into asset X in a CPMM or executes a statistical arbitrage over a period of time [0, T], where T > 0. She uses the rate Z in (11) from the pool and the rate S in 8 from



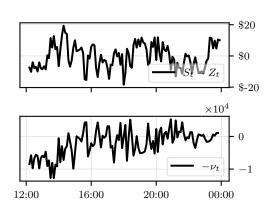


- (a) **Top**: Out-of-sample instantaneous and oracle rates for the pair ETH/USDC. **Bottom**: Inventory process \tilde{y} for the optimal, TWAP, and single order strategies.
- (b) **Top**: Out-of-sample instantaneous and oracle rates for the pair ETH/DAI. **Bottom**: Inventory process \tilde{y} for the optimal, TWAP, and single order strategies.

Figure B.11: Liquidation strategies starting at noon on 16 March 2022.



(a) **Top**: Difference between the oracle and the instantaneous rate for ETH/USDC. **Bottom**: $-\nu_t$.



(b) **Top**: Difference between the oracle and the instantaneous rate for ETH/DAI. **Bottom**: $-\nu_t$.

Figure B.12: Trading speed.

another more liquid exchange. The depth κ of the pool is constant and the investor liquidates a position $\tilde{y}_0 \in \mathbb{R}$ in asset Y. Her wealth is valued in terms of asset X and has dynamics (10), where the execution cost is stochastic and its dynamics are known. During the trading programme, the investor trades at the speed $(\nu_t)_{t \in [0,T]}$, so the inventory $(\tilde{y}_t)_{t \in [0,T]}$ evolves as in (12), where we do not restrict the speed to be positive, and for simplicity, trading fees are set to zero.

The investor maximises her expected terminal wealth in units of X while penalising inventory. The set of admissible strategies is defined in (13) and the investor's performance criterion is a function $u^{\nu} \colon [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ defined in (14). The value function $u \colon [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ of the investor is given by

$$u(t, \tilde{x}, \tilde{y}, Z, S) = \sup_{\nu \in A} \{ u^{\nu}(t, \tilde{x}, \tilde{y}, Z, S) \}.$$
 (C.1)

The value function solves the Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \partial_t w - \phi \, \tilde{y}^2 + \beta \, (S - Z) \, \partial_Z w + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} w + \frac{1}{2} \, \sigma^2 \, S^2 \, \partial_{SS} w$$

$$+ \sup_{\nu \in \mathbb{R}} \left(\left(\nu \, Z - \frac{\eta}{\kappa} \, Z^{3/2} \, \nu^2 \right) \partial_{\tilde{x}} w - \nu \, \partial_{\tilde{y}} w \right), \tag{C.2}$$

with terminal condition

$$w(T, \tilde{x}, \tilde{y}, Z, S) = \tilde{x} + \tilde{y}Z - \alpha \tilde{y}^2. \tag{C.3}$$

The form of the terminal condition (C.3) suggests the ansatz

$$w(T, \tilde{x}, \tilde{y}, Z, S) = \tilde{x} + \tilde{y}Z + \theta(t, \tilde{y}, Z, S). \tag{C.4}$$

Now plug (C.4) into (C.2) to obtain the PDE

$$0 = \partial_t \theta - \phi \, \tilde{y}^2 + \beta \, (S - Z) \, (\tilde{y} + \partial_Z \theta) + \frac{1}{2} \, \gamma^2 \, Z^2 \, \partial_{ZZ} \theta + \frac{1}{2} \, \sigma^2 \, S^2 \, \partial_{SS} \theta$$

$$+ \sup_{\nu \in \mathbb{R}} \left(-\frac{\eta}{\kappa} Z^{3/2} \, \nu^2 - \nu \, \partial_{\tilde{y}} \theta \right), \tag{C.5}$$

defined over $[0,T) \times \mathbb{R} \times \mathbb{R}^* \times \mathbb{R}$, with terminal condition

$$\theta(T, \tilde{y}, Z, S) = -\alpha \, \tilde{y}^2 \, .$$

The first two terms on the right-hand side of (C.4) are the mark-to-market value of the investor's holdings and the last term is the additional value that the investors obtains by following the optimal strategy. Next, solve the first order condition in (C.5) to obtain the optimal trading speed in feedback form

$$\nu^* = -\frac{\kappa}{2\eta} Z^{-3/2} \,\partial_{\tilde{y}}\theta \,\,, \tag{C.6}$$

and substitute (C.6) into (C.5) to write

$$\partial_t \theta = -\phi \,\tilde{y}^2 + \beta \,\left(S - Z\right) \,\left(\tilde{y} + \partial_Z \theta\right) + \frac{1}{2} \,\gamma^2 \,Z^2 \,\partial_{ZZ} \theta + \frac{1}{2} \,\sigma^2 \,S^2 \,\partial_{SS} \theta + \frac{\kappa}{4 \,\eta} \,Z^{-3/2} \,\partial_{\tilde{y}} \theta^2 \,. \tag{C.7}$$

Finally, we propose the ansatz

$$\theta(t, \tilde{y}, Z, S) = \theta_2(t, Z, S) \, \tilde{y}^2 + \theta_1(t, Z, S) \, \tilde{y} + \theta_0(t, Z, S) \,. \tag{C.8}$$

Now plug (C.8) into (C.7) to obtain the system of PDEs

$$\begin{cases} -\partial_{t}\theta_{2} = -\phi + \beta \, \left(S - Z \right) \, \partial_{Z}\theta_{2} + \frac{1}{2} \, \gamma^{2} \, Z^{2} \, \partial_{ZZ}\theta_{2} + \frac{1}{2} \, \sigma^{2} \, S^{2} \, \partial_{SS}\theta_{2} + \frac{\kappa}{\eta} \, Z^{-3/2} \, \theta_{2}^{2} \, , \\ -\partial_{t}\theta_{1} = \beta \, \left(S - Z \right) \left(1 + \partial_{Z}\theta_{1} \right) + \frac{1}{2} \, \gamma^{2} \, Z^{2} \, \partial_{ZZ}\theta_{1} + \frac{1}{2} \, \sigma^{2} \, S^{2} \, \partial_{SS}\theta_{1} + \frac{\kappa}{\eta} \, Z^{-3/2} \, \theta_{2} \, \theta_{1} \, , \\ -\partial_{t}\theta_{0} = \beta \, \left(S - Z \right) \, \partial_{Z}\theta_{0} + \frac{1}{2} \, \gamma^{2} \, Z^{2} \, \partial_{ZZ}\theta_{0} + \frac{1}{2} \, \sigma^{2} \, S^{2} \, \partial_{SS}\theta_{0} + \frac{\kappa}{4 \, \eta} \, Z^{-3/2} \, \theta_{1}^{2} \, , \end{cases}$$
(C.9)

defined over $[0,T) \times \mathbb{R}^* \times \mathbb{R}$, with terminal condition

$$\theta_2(T, Z, S) = -\alpha$$
, $\theta_1(T, Z, S) = 0$, $\theta_0(T, Z, S) = 0$. (C.10)

The optimal strategy in feedback form (C.6) is given by

$$\nu^* = -\frac{\kappa}{2\eta} Z^{-3/2} \left(2\theta_2 \,\tilde{y} + \theta_1 \right) . \tag{C.11}$$

The next result provides a-priori bounds for the solution to the system of PDEs (C.9).

Theorem 3 Assume there exist $\theta_0 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, $\theta_1 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$, and $\theta_2 \in C^{1,2,2}([0,T] \times \mathbb{R}_{++} \times \mathbb{R}_{++})$ solution to the system of PDEs (C.9) with terminal condition (C.10). Define θ in (C.8) and define w in (C.4). Assume that for all $(t,\tilde{x},\tilde{y},Z,\kappa) \in [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++}$ we have $u^{\nu}(t,\tilde{x},\tilde{y},Z,\kappa) \leq w(t,\tilde{x},\tilde{y},Z,\kappa)$, where u^{ν} is defined in (14), and that equality is obtained for the optimal control $(\nu^{\star})_{t \in [0,T]}$ in feedback form in (C.11). Then θ_0 , θ_1 , and

 θ_2 have the following bounds for all $(t, Z, S) \in [0, T] \times \mathbb{R}_{++} \times \mathbb{R}_{++}$:

$$\begin{cases} \theta_{0}(t, Z, S) \geq 0 \\ \theta_{0}(t, Z, S) \leq A(t) S^{2} + \frac{1}{2}B(t) S Z + C(t) Z^{2} - Z + \mathbb{E}[Z_{T}], \end{cases}$$

$$\begin{cases} \theta_{1}(t, Z, S) \geq -(\alpha + \phi(T - t)) - A(t) S^{2} - \frac{1}{2}B(t) S Z - C(t) Z^{2} \\ \theta_{1}(t, Z, S) \leq A(t) S^{2} + \frac{1}{2}B(t) S Z + C(t) Z^{2} + \alpha + \phi(T - t), \end{cases}$$

$$\begin{cases} \theta_{2}(t, Z, S) \geq -(\alpha + \phi(T - t)) \\ \theta_{2}(t, Z, S) \leq A(t) S^{2} + \frac{1}{2}B(t) S Z + C(t) Z^{2}, \end{cases}$$

where $\mathbb{E}[Z_T]$ is in (C.13) and A, B, and C are in (C.18).

Proof We provide a-priori bounds for the solutions θ_2 , θ_1 , θ_0 of (C.9).

<u>Lower bound for θ </u>. By definition of our control problem in (C.1) we know that for any control $\nu \in \mathcal{A}$, we have

$$u^{\underline{\nu}}(t, \tilde{x}, \tilde{y}, Z, S) \leq \sup_{\nu \in A} u^{\nu}(t, \tilde{x}, \tilde{y}, Z, S)$$
.

Use the definition of θ in (C.4) and the form of the performance criterion to write

$$\mathbb{E}\left[\tilde{x}_T^{\underline{\nu}} + \tilde{y}_T^{\underline{\nu}} Z_T - \alpha \left(\tilde{y}_T^{\underline{\nu}}\right)^2 - \phi \int_t^T \left(\tilde{y}_s^{\underline{\nu}}\right)^2 ds\right] \leq \tilde{x} + \tilde{y} Z + \theta(t, \tilde{y}, Z, S) .$$

Next, consider a sub-optimal strategy consisting in keeping a constant inventory with no trading, i.e., a strategy defined by $\underline{\nu}_s=0$ for each $s\geq t$. Then the inventory process is $(\tilde{y}^{\underline{\nu}}_s)_{s\geq t}=\tilde{y}$. Use the inequality above to write

$$\tilde{x} + \tilde{y}Z + \theta(t, \tilde{y}, Z, S) \ge \mathbb{E}\left[\tilde{x} + \tilde{y}Z - \tilde{y}(Z - Z_T) - \alpha \tilde{y}^2 - \phi \tilde{y}^2(T - t)\right],$$

and conclude that

$$\theta(t, \tilde{y}, Z, S) \ge -\tilde{y} Z - \tilde{y} \mathbb{E} \left[Z_T \right] - \alpha \, \tilde{y}^2 - \phi \, \tilde{y}^2 \left(T - t \right) > -\infty \,, \tag{C.12}$$

which provides a lower bound for θ . The last inequality in (C.12) is a consequence of

$$\mathbb{E}\begin{bmatrix} S_T \\ Z_T \end{bmatrix} = \begin{pmatrix} S \\ Z \end{pmatrix} e^{A(T-t)}, \quad \text{where} \quad A = \begin{pmatrix} 0 & 0 \\ -\theta & \theta \end{pmatrix},$$

which simplifies to

$$\mathbb{E}[Z_T] = Z e^{-\theta(T-t)} + \theta S \int_t^T e^{-\theta(T-s)} ds = Z e^{-\theta(T-t)} + S \left(1 - e^{-\theta(T-t)}\right). \tag{C.13}$$

Upper bound for θ . Use integration by parts to write, for all $(t, \tilde{x}, \tilde{y}, Z, S)$,

$$\tilde{x} + \tilde{y} Z + \theta(t, \tilde{y}, Z, S) = \sup_{\nu \in \mathcal{A}_{t}} \mathbb{E} \left[\tilde{x}_{T} + \tilde{y}_{T} Z - \alpha \, \tilde{y}_{T}^{2} - \phi \int_{t}^{T} \tilde{y}_{s}^{2} \, ds \right]$$

$$= \tilde{x} + \tilde{y} Z + \sup_{\nu \in \mathcal{A}_{t}} \mathbb{E} \left[\int_{t}^{T} \tilde{y}_{s} \, dZ_{s} - \frac{\eta}{\kappa} \int_{t}^{T} Z_{s}^{3/2} \, v_{s}^{2} \, ds - \alpha \, \tilde{y}_{T}^{2} - \phi \int_{t}^{T} \tilde{y}_{s}^{2} \, ds \right]$$

$$\leq \tilde{x} + \tilde{y} Z + \sup_{\nu \in \mathcal{A}_{t}} \mathbb{E} \left[\int_{t}^{T} \tilde{y}_{s} \, dZ_{s} - \phi \int_{t}^{T} \tilde{y}_{s}^{2} \, ds \right],$$
(C.14)

where the first equality is by definition of our control problem and θ , the second equality uses integration by parts, and the last inequality holds because the terms $\frac{\eta}{\kappa} \int_t^T Z_s^{3/2} v_s^2 \, ds$ and $\alpha \, \tilde{y}_T^2$ are always positive.

Let $t \in [0, T]$. Define the set

$$\mathcal{A}_t^M = \left\{ \left(\tilde{y}_s^M \right)_{s \in [t,T]}, \; \mathbb{R} - \text{valued}, \; \mathbb{F} - \text{adapted and such that} \; \int_t^T \left| \tilde{y}_u^M \right|^2 du < +\infty, \; \; \mathbb{P} - \text{a.s.} \right\}. \tag{C.15}$$

Clearly, one has

$$\sup_{\nu \in \mathcal{A}_t} \mathbb{E}\left[\int_t^T \tilde{y}_s \, dZ_s - \phi \int_t^T \tilde{y}_s^2 ds\right] \leq \sup_{\tilde{y}^M \in \mathcal{A}_t^M} \mathbb{E}\left[\int_t^T \tilde{y}_s^M \, dZ_s - \phi \int_t^T \left(\tilde{y}_s^M\right) \, ds\right].$$

The term on the right-hand side of the inequality is a classical Merton problem, where the rate Z follows the dynamics in (11). We introduce the wealth process $(\tilde{x}_t^M)_{t\in[0,T]}$ with dynamics

$$d\tilde{x}_t^M = \tilde{y}_t^M dZ_t,$$

In the Merton problem, the investor controls her inventory \tilde{y}^M , and her strategy solves the opti-

mization problem

$$\overline{u}(t, \tilde{x}, Z, S) = \sup_{\tilde{y} \in \mathcal{A}_t^M} \mathbb{E}\left[\tilde{x}_T^M - \phi \int_t^T \left(\tilde{y}_s^M\right)^2 ds\right] = \sup_{\tilde{y} \in \mathcal{A}_t^M} \mathbb{E}\left[\int_t^T \tilde{y}_s^M dZ_s - \phi \int_t^T \left(\tilde{y}_s^M\right)^2 ds\right].$$
(C.16)

We introduce the value function \overline{w} associated with the optimisation problem (C.16) and which solves the HJB

$$0 = \partial_t \overline{w} + \beta (S - Z) \partial_Z \overline{w} + \frac{1}{2} \gamma^2 Z^2 \partial_{ZZ} \overline{w} + \frac{1}{2} \sigma^2 S^2 \partial_{SS} \overline{w}$$
$$+ \sup_{\tilde{y}^M \in \mathbb{R}} \left(\beta (S - Z) \tilde{y}^M \partial_{\tilde{x}} \overline{w} + \frac{1}{2} \gamma^2 Z^2 (\tilde{y}^M)^2 \partial_{\tilde{x}\tilde{x}} \overline{w} + \gamma^2 Z^2 \tilde{y} \partial_{\tilde{x}Z} \overline{w} - \phi (\tilde{y}^M)^2 \right)$$

with terminal condition $\overline{w}\left(T,\tilde{x}^M,Z,S\right)=\tilde{x}^M.$ Use the ansatz $\overline{w}(t,\tilde{x}^M,Z,S)=\tilde{x}^M+\overline{\theta}(t,Z,S)$ to obtain the PDE

$$0 = \partial_t \overline{\theta} + \beta (S - Z) \partial_Z \overline{\theta} + \frac{1}{2} \gamma^2 Z^2 \partial_{ZZ} \overline{\theta} + \frac{1}{2} \sigma^2 S^2 \partial_{SS} \overline{\theta} + \frac{\beta^2 (S - Z)^2}{4\phi},$$

with terminal condition $\overline{\theta}(T,Z,S)=0$. The optimal control is

$$\tilde{y}^{M\star} = \frac{\beta \ (S - Z)}{2 \, \phi} \,. \tag{C.17}$$

Use the ansatz

$$\overline{\theta}(t, Z, S) = A(t) S^2 + \frac{1}{2} B(t) S Z + C(t) Z^2 + D(t) S + E(t) Z,$$

to obtain the following system of ODEs:

$$\begin{cases}
-A'(t) = \sigma^2 A(t) + \frac{\beta^2}{4\phi} + \frac{1}{2} \beta B(t), \\
-B'(t) = -\beta B(t) - \frac{\beta^2}{\phi} + 4 \beta C(t), \\
-C'(t) = (\gamma^2 - 2 \beta) C(t) + \frac{\beta^2}{4\phi}, \\
-D'(t) = E'(t) = \beta E(t),
\end{cases}$$

with terminal conditions A(T) = B(T) = C(T) = D(T) = E(T) = 0. The solution is

$$\begin{cases} A(t) = \int_{t}^{T} e^{\sigma^{2}(u-t)} \left(\frac{1}{2}\beta B(u) + \frac{\beta^{2}}{4\phi}\right) du, \\ B(t) = \int_{t}^{T} e^{-\beta(u-t)} (4\beta C(u) - \frac{\beta^{2}}{\phi}) du, \\ C(t) = \frac{-\beta^{2}}{4\phi(\gamma^{2}-2\beta)} \left(1 - e^{(\gamma^{2}-2\beta)(T-t)}\right), \\ D(t) = E(t) = 0. \end{cases}$$
(C.18)

Note that we obtain a classical solution to the Merton problem so standard results apply. In particular, one only needs to verify that the optimal control is admissible, which is clear from the optimal control (C.17) and the admissible set (C.15). Finally, write the inequalities in (C.12) and (C.14) as

$$A(t) S^{2} + \frac{1}{2} B(t) S Z + C(t) Z^{2} \ge \theta(t, \tilde{y}, Z, S) \ge -\tilde{y} \left(-Z + \mathbb{E}\left[Z_{T}\right]\right) - \left(\alpha + \phi \left(T - t\right)\right) \tilde{y}^{2}.$$
(C.19)

Lower and upper bounds for θ_0 , θ_1 , and θ_2 . From the à-priori bounds on θ given by the original control problem, we deduce à-priori bounds on the solutions θ_0 , θ_1 , and θ_2 of the system (C.9). Recall that

$$\theta(t, \tilde{y}, Z, S) = \theta_2(t, Z, S) \, \tilde{y}^2 + \theta_1(t, Z, S) \, \tilde{y} + \theta_0(t, Z, S) \,.$$

First, notice that the lower bound is in the form of a quadratic polynomial in \tilde{y} , so $\theta_2(t,Z,S) \ge -(\alpha+\phi(T-t))$. Next, for the upper bound, see that it holds for any value of the inventory \tilde{y} . In particular, observe that for $\tilde{y}=0$ we find that

$$A(t) S^{2} + \frac{1}{2} B(t) SZ + C(t) Z^{2} \ge \theta_{0}(t, Z, S) \ge 0.$$
 (C.20)

Moreover, if we let $\tilde{y} = 1$ and $\tilde{y} = -1$ in (C.19) one obtains

$$2A(t) S^2 + B(t) S Z + 2C(t) Z^2 > 2\theta_0(t, Z, S) + 2\theta_2(t, Z, S) > -2(\alpha + \phi(T - t))$$
.

Finally, use (C.20) to write

$$A(t) S^{2} + \frac{1}{2} B(t) S Z + C(t) Z^{2} \ge \theta_{2}(t, Z, S) \ge -(\alpha + \phi(T - t)).$$

Next, substitute $\tilde{y} = 1$ into (C.19) to write

$$\theta_{1}(t, Z, S) \geq -Z + \mathbb{E}[Z_{T}] - (\alpha + \phi(T - t)) - \theta_{2}(t, Z, S) - \theta_{0}(t, Z, S)$$

$$\geq -Z + \mathbb{E}[Z_{T}] - (\alpha + \phi(T - t)) - \left(A(t) S^{2} + \frac{1}{2}B(t) S Z + C(t) Z^{2}\right),$$

and

$$\theta_1(t, Z, S) \le A(t) S^2 + \frac{1}{2} B(t) S Z + C(t) Z^2 - \theta_2(t, Z, S) - \theta_0(t, Z, S)$$

$$\le A(t) S^2 + \frac{1}{2} B(t) S Z + C(t) Z^2 + \alpha + \phi (T - t) .$$

Thus, provided θ exists, then θ_0 , θ_1 , θ_2 have à-priori upper and lower bounds. The bounds are all at most linear in time, and quadratic in $\{S, Z\}$.

References

- Almgren, R., 2012. Optimal trading with stochastic liquidity and volatility. SIAM Journal on Financial Mathematics 3, 163–181.
- Almgren, R., Chriss, N., 2000. Optimal execution of portfolio transactions. Journal of Risk 3, 5–39.
- Barger, W., Lorig, M., 2019. Optimal liquidation under stochastic price impact. International Journal of Theoretical and Applied Finance 22, 1850059.
- Bechler, K., Ludkovski, M., 2015. Optimal execution with dynamic order flow imbalance. SIAM Journal on Financial Mathematics 6, 1123–1151. doi:10.1137/140992254.
- Belak, C., Muhle-Karbe, J., Ou, K., 2018. Optimal trading with general signals and liquidation in target zone models. arXiv preprint arXiv:1808.00515.
- Bergault, P., Drissi, F., Guéant, O., 2022. Multi-asset optimal execution and statistical arbitrage strategies under ornstein–uhlenbeck dynamics. SIAM Journal on Financial Mathematics 13, 353–390. doi:10.1137/21M1407756.
- Bertsimas, D., Lo, A., 1998. Optimal control of execution costs. Journal of Financial Markets 1, 1-50.
- Cartea, Á., Donnelly, R., Jaimungal, S., 2018a. Enhancing trading strategies with order book signals. Applied Mathematical Finance 25, 1–35.
- Cartea, Á., Drissi, F., Monga, M., 2022. Decentralised finance and automated market making: Predictable loss and optimal liquidity provision. Available at SSRN 4273989.
- Cartea, Á., Gan, L., Jaimungal, S., 2018b. Trading co-integrated assets with price impact. Mathematical Finance 29. doi:10.1111/mafi.12181.
- Cartea, Á., Jaimungal, S., 2016. Incorporating order-flow into optimal execution. Mathematics and Financial Economics 10, 339–364.
- Cartea, Á., Jaimungal, S., Penalva, J., 2015. Algorithmic and High-Frequency Trading. Cambridge University Press.
- Cartea, Á., Jaimungal, S., Sánchez-Betancourt, L., 2021. Latency and liquidity risk. International Journal of Theoretical and Applied Finance 24, 2150035.
- Cartea, Á., Sánchez-Betancourt, L., 2021a. Optimal execution with stochastic delay. Available at SSRN 3812324.
- Cartea, Á., Sánchez-Betancourt, L., 2021b. The shadow price of latency: Improving intraday fill ratios in foreign exchange markets. SIAM Journal on Financial Mathematics 12, 254–294.
- Cheridito, P., Sepin, T., 2014. Optimal trade execution under stochastic volatility and liquidity. Applied Mathematical Finance 21, 342–362.
- Chiu, J., Koeppl, T.V., 2019. Blockchain-based settlement for asset trading. The Review of Financial Studies 32, 1716–1753.
- Forde, M., Sánchez-Betancourt, L., Smith, B., 2022. Optimal trade execution for gaussian signals with power-law resilience. Quantitative Finance 22, 585–596.
- Fouque, J.P., Jaimungal, S., Saporito, Y.F., 2021. Optimal trading with signals and stochastic price impact. arXiv preprint arXiv:2101.10053.
- Gatheral, J., Schied, A., 2013. Dynamical models of market impact and algorithms for order execution. Handbook on Systemic Risk, Jean-Pierre Fouque, Joseph A. Langsam, eds., 579–599.
- Graewe, P., Horst, U., Séré, E., 2018. Smooth solutions to portfolio liquidation problems under price-sensitive market impact. Stochastic Processes and their Applications 128, 979–1006.

- Guéant, O., 2016. The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making. doi:10.1201/b21350.
- Heimbach, L., Wang, Y., Wattenhofer, R., 2021. Behavior of liquidity providers in decentralized exchanges. doi:10.48550/ARXIV.2105.13822.
- Lehalle, C.A., Neuman, E., 2019. Incorporating signals into optimal trading. Finance and Stochastics 23. doi:10. 1007/s00780-019-00382-7.
- Lipton, A., Treccani, A., 2021. Blockchain and Distributed Ledgers: Mathematics, Technology, and Economics. World Scientific.
- Neuman, E., Voß, M., 2020. Optimal signal-adaptive trading with temporary and transient price impact. doi:10.48550/ARXIV.2002.09549.