

Leveraging IS and TC: Optimal order execution subject to reference strategies

Xue Cheng¹, Peng Guo¹, and Tai-Ho Wang²

¹LMEQF, Department of Financial Mathematics, School of Mathematical Sciences, Peking University, Beijing 100871, China.

²Department of Mathematics, Baruch College, CUNY, 1 Bernard Baruch Way, New York, NY 10010, USA

March 5, 2025

Abstract

The paper addresses the problem of meta order execution from a broker-dealer's point of view in Almgren-Chriss model under execution risk. A broker-dealer agency is authorized to execute an order of trading on some client's behalf. The strategies that the agent is allowed to deploy is subject to a benchmark, referred to as the reference strategy, regulated by the client. We formulate the broker's problem as a utility maximization problem in which the broker seeks to maximize his utility of excess profit-and-loss at the execution horizon, of which optimal feedback strategies are obtained in closed form. In the absence of execution risk, the optimal strategies subject to reference strategies are deterministic. We establish an affine structure among the trading trajectories under optimal strategies subject to general reference strategies using implementation shortfall (IS) and target close (TC) orders as basis. Furthermore, an approximation theorem is proposed to show that with small error, general reference strategies can be approximated by piece-wise constant ones, of which the optimal strategy is piece-wise linear combination between IS and TC orders. We conclude the paper with numerical experiments illustrating the trading trajectories as well as histograms of terminal wealth and utility at investment horizon under optimal strategies versus those under TWAP strategies.

Keywords: Optimal execution; Price impact; Execution risk; Utility maximization; Reference strategy; Affine structure; Implementation shortfall order; Target close order

Contents

1	Introduction	3
2	Model setup	5
2.1	Price impact model	5
2.2	Reference strategy	6
3	Optimal execution as utility maximizing	7
3.1	Optimal execution with risk aversion	7

4 Zero execution risk	10
4.1 General reference strategy	11
4.2 Implementation shortfall (IS) order and target close (TC) order	12
4.3 Endpoints-only reference strategy	14
4.4 Piece-wise constant reference strategy	16
5 Numerical examples	18
5.1 Sample trading trajectories	18
5.2 Performance analysis and stress test	22
6 Conclusion	25
A Appendix	25

1 Introduction

In the financial industry, large position holders such as pension funds or investment banks for various reasons are required to trade in or trade out from their current position to an updated target, possibly subject to a given execution horizon which may vary from days to weeks. The net holdings to be adjusted between the current and the target positions are usually too large to be simply dumped to the market without a priori deliberately assessing the trade’s market impact. Untamed price impact by trading may result in significant transaction cost, potentially turned into a substantial loss. Such an execution risk requires to be properly managed and controlled; otherwise it would eventually influence the position holder’s overall profit and loss (P&L). A common practice is to delegate the execution to the firm’s order execution department or outsource to a broker-dealer agency.

The large position holders, while delegating their tasks to an agency for execution, may have in mind their own preferred strategies or benchmarks that they would like the delegated agent to closely track along. For instance, *implementation shortfall* (IS) orders [1] are frequently employed by managers for the purpose of short-term alpha pursuit. IS orders are constructed with a pre-trade benchmark price in mind, aiming at executing orders at an average price that remains relatively close to the market price at the beginning of the trade. Managers can often use arrival prices to help measure total trading costs: the closer the execution price is to the arrival price, the lower the associated costs (see CFA-level-II [2]). On the contrary, *target close* (TC) orders [3], often deployed by index-fund managers for the purpose of minimizing fund risk and tracking error, are formulated with a post-trade benchmark price in order to secure a price in average that remains relatively close to the closing price. This is often important for mutual fund managers who manage funds that only calculate NAV once daily at closing. *Volume weighted average price* (VWAP) and *Time weighted average price* (TWAP) strategies are benchmarks specified for trading sequences with constant trading rate in wall clock time (TWAP) and in volume time (VWAP) respectively. These algorithmic trading strategies are examples of benchmarks that may be imposed as trading constraints to the agent who is missioned to trade in or trade out the position. We shall delve into this type of order execution problems subject to a pre-specified benchmark strategy, which we refer to as the *reference strategy* (RS for short), as a stochastic control problem and determine their corresponding optimal strategies in feedback form.

The pioneering works in [4], [1], [5] and [6] are among the first to deal with the problem of order execution under price impact. Since its introduction to the order execution problem, numerous progresses and extensions on the classical Almgren-Chriss framework have been made extensively. For instance, [7] and [8] introduced the notion of transient impact to account for the dissipation of price impacts from the past trades. [9] discusses the objective in the optimization problem, while [10] extends the classical mean-variance framework first considered in [4] to encompass general risk measures as penalty for risk aversion. [11], [12], [13] solve the optimal execution problem in relation to a VWAP benchmark, while [14] asserts the use of the arrival price as a benchmark within the Almgren-Chriss framework. In fact, the arrival price, also known as “pre-trade benchmark”, appears to be the most commonly used benchmark in academic papers. The closing price is the value of a security at its last transaction during a trading session [3] [15]. Average price benchmarks may also appear in accelerated share repurchase (ASR) contracts [16][17], which can be regarded as optimal execution problems with optimal stopping. Above all, [18] introduces the concept of execution risk, that

pre-scheduled orders may not be fully executed, of which the empirical evidence is confirmed by [19] that the inventory processes of traders invariably contain a Brownian motion term, which contradicts the common assumption adopted in most optimal execution models that the inventory process is absolute continuous. Another perspective on introducing the noise term into the inventory process is that when a centralized trading desk aggregates order flows within a financial institution, stochastic order flow arises [20] [21]. The aforementioned papers are by no means meant for an exhausting list in literature on this line of active research.

In the current paper, we consider the order execution problem from a broker-dealer's point of view. Assume that a broker is delegated to reallocate a client's holdings of a certain stock under the Almgren-Chriss model with execution risk. The broker is regulated by his client to track a benchmark strategy, i.e., the reference strategy, to the client's preference. The broker's incentive of executing client's order in this circumstance is to maximize his own expected P&L excess to that of the reference strategy, marked-to-market. To account for risk aversion, we recast the broker's order execution problem as a utility maximization problem and, in certain cases, are able to solve the problem in closed form. In particular, when execution risk vanishes, or becomes negligible, we show that there exists an "affine structure" among the optimal strategies induced from various reference strategies. This algebraic structure is supposed to help the broker for concocting and understanding optimal strategies subject to client's general reference strategies. We argue that the framework is highly versatile in the sense that it encompasses commonly deployed execution strategies such as IS, TC as well as TWAP and VWAP orders as special cases for benchmarking. As per the algebraic structure, it follows that, for any given continuous reference strategy which can be approximated by a piece-wise constant function, we show that its corresponding optimal strategy can also be approximated by those induced from the piece-wise constant strategies.

The rest of the paper is organized as follows. In Section 2, we lay out the price impact model of Almgren-Chriss under execution risk and incorporate reference strategies into the problem of order execution. Section 3 presents the optimal feedback control of the order execution problem in Theorem 1 as one of the main results in the paper. Section 4 focuses and provides detailed discussions on the optimal strategies when execution risk vanishes. Reference strategies and their associated optimal strategies considered in Section 4 include IS and TC orders as well as piece-wise constant strategies. The emphasis is put on an "affine structure" among the trading trajectories induced by general reference strategies using unit IS and unit TC orders as a basis. Numerical examples illustrating the trading trajectories and the performance analysis under the optimal and TWAP strategies are shown and discussed in Section 5. For the sake of smooth reading, technical proofs of all the theorems, propositions and lemmas are postponed and collected till the end of the paper as an appendix in Section A.

Throughout the paper, $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a complete probability space equipped with a filtration describing the information structure $\mathbb{F} := \{\mathcal{F}_t\}_{t \in [0, T]}$, where t is the time variable and $T > 0$ the fixed finite liquidation horizon. Let $\{W_t, Z_t\}_{t \in [0, T]}$ be a two-dimensional Brownian motion with constant correlation ρ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The filtration \mathbb{F} is generated by the trajectories of the above Brownian motion, completed with all \mathbb{P} -null measure sets of \mathcal{F} .

2 Model setup

2.1 Price impact model

Assume a broker is delegated to reallocate a client's holdings of a certain stock from x_0 shares to A shares within a given horizon T . Let x_t be the number of shares the broker holds at time $t \in [0, T]$ during the reallocation process and \tilde{S}_t the transacted price at time t . The price dynamic is assumed to follow the Almgren–Chriss model [4] [1] [5]. In the Almgren–Chriss framework, the transacted price \tilde{S}_t consists of the fair price S_t and a slippage. The fair price S_t is driven by the SDE

$$dS_t = \mu dt + \gamma dx_t + \sigma dW_t,$$

or equivalently,

$$S_t = S_0 + \mu t + \gamma(x_t - x_0) + \sigma W_t,$$

where μt describes the tendency of the stock and $(W_t)_{t \in [0, T]}$ is a standard Brownian motion. The term $\gamma(x_t - x_0)$, for $\gamma \geq 0$, is usually referred to as the *permanent impact*. Penalized by a price slippage, the transacted price is thus given by

$$\tilde{S}_t = S_t - \eta v_t,$$

where v_t denotes the broker's intended trading rate at the instant t . The slippage $-\eta v_t$, for $\eta \geq 0$, is also referred to as the *temporary impact*. We remark that in the original setting of Almgren–Chriss and its extensions, the trading rate v_t plays a dual role. On the one hand, it serves as the realized trading rate per the relationship $v_t = -\dot{x}_t$ between x_t and v_t . On the other hand, it is regarded as the control variable in the problem of optimal execution. These two seemingly distinct roles coincide if all the scheduled orders are guaranteed fully executed. However, in reality it is well noticed among practitioners that, while executing a sequence of pre-scheduled orders, the orders in the sequence may not be fully executed, resulting in an uncontrollable realized order flow. This introduces an additional risk to the order execution problem, the *execution risk*. To account for this uncertainty, we introduce a noise component driven by a correlated Brownian motion Z_t into the dynamic of the position x_t as

$$dx_t = -v_t dt + m(v_t) dZ_t.$$

The diffusion term $m(v)$ characterizes the magnitude of execution risk. It is worth to reiterate that in a recent work in [19], the authors showed that the presence of a Brownian component in the broker's inventory during reallocation process is statistically significant. Moreover, because of this execution risk, the broker is no longer guaranteed to achieve his intended position at the terminal time T , which gives rise to an additional opportunity cost. As a result, the broker is obligated to take a final block trade at time T at a worse price. Overall, the realised P&L at the horizon T is given by

$$\Pi := (x_T - A)S_T - \beta(x_T - A)^2 + \int_0^T (-\tilde{S}_t) dx_t,$$

where the term $\beta(x_T - A)^2$, for $\beta \geq 0$, penalizes the discrepancy from a final block trade.

The following proposition shows that the P&L can be written as an Itô process.

Proposition 1. *The realised P&L can be rewritten as*

$$\begin{aligned}\Pi = & (x_0 - A)S_0 - \beta(x_T - A)^2 + \int_0^T (\mu(x_t - A) + \rho\sigma m(v_t) - \eta v_t^2 - \gamma v_t(x_t - A)) dt \\ & + \int_0^T \sigma(x_t - A)dW_t + \int_0^T (\eta v_t + \gamma(x_t - A))m(v_t)dZ_t.\end{aligned}$$

2.2 Reference strategy

A common practice in order execution brokerage is that clients may come forward to brokers with their own preferred strategies for benchmarking. These benchmark strategies can be either strategies suggested by elite investors or commonly used ones such as TWAP strategies. We shall refer to these pre-specified benchmark strategies as *reference strategies*. We consider the reference strategies that can be represented by a deterministic function $(R_t)_{t \in [0, T]}$ with $R_0 \equiv x_0$ and $R_T \equiv A$ hereafter. Moreover, we assume that R_t is a differentiable a.e. function. The broker's incentive of executing client's order is thus to maximize his own expected P&L excess to that of the reference strategy, marked-to-market. Specifically, by disregarding the price impact and slippages incurred from order execution, the stock price S_t^0 reads

$$S_t^0 = S_0 + \mu t + \sigma W_t.$$

The marked-to-market P&L Π^R for the reference strategy R is evaluated as

$$\begin{aligned}\Pi^R := & \int_0^T (-S_t^0) dR_t \\ = & -(S_t^0 R_t)|_0^T + \int_0^T R_t(\mu dt + \sigma dW_t) \\ = & -AS_T^0 + x_0 S_0 + \int_0^T \mu R_t dt + \int_0^T \sigma R_t dW_t \\ = & (x_0 - A)S_0 + \int_0^T \mu(R_t - A)dt + \int_0^T \sigma(R_t - A)dW_t.\end{aligned}$$

Hence, the broker's excess P&L $\tilde{\Pi}$, defined by the difference between Π and Π^R , is given by

$$\begin{aligned}\tilde{\Pi} := & \Pi - \Pi^R \\ = & -\beta(x_T - A)^2 + \int_0^T (\mu(x_t - R_t) + \rho\sigma m(v_t) - \eta v_t^2 - \gamma v_t(x_t - A)) dt \\ & + \int_0^T \sigma(x_t - R_t)dW_t + \int_0^T (\eta v_t + \gamma(x_t - A))m(v_t)dZ_t.\end{aligned}\tag{1}$$

The broker's goal is thus to maximize his expected excess P&L in a risk aversion manner which we recast as a utility maximization problem in the section that follows.

3 Optimal execution as utility maximizing

In this section, we recast the problem of order execution as a utility maximization problem as follows. Recall that (Ω, \mathcal{F}, P) denotes a complete probability space equipped with a filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfying the usual conditions. We assume that all random variables and stochastic processes are defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The set of all real-valued progressively measurable processes are denote by \mathcal{M} , while the collection of admissible controls \mathcal{A} is set as

$$\mathcal{A} := \left\{ v \in \mathcal{M} : \int_0^T \mathbb{E}[v_t^2] dt < \infty \right\}.$$

to meet the assumptions given in [22] and [23], which assure the solvability of both the risk-neutral problem and the risk-aversion one. The broker's problem is to determine an optimal admissible strategy v^* that maximizes his expected exponential utility at the horizon T , i.e.,

$$\sup_{v \in \mathcal{A}} \mathbb{E} \left[u \left(\tilde{\Pi} \right) \right], \quad (2)$$

where the utility function $u(x) := \frac{1}{\theta}(1 - e^{-\theta x})$, $\theta > 0$ represents the broker's preference following CARA (Constant Absolute Risk Aversion) preference and $\theta > 0$ is the risk-aversion parameter. Note that as $\theta \downarrow 0$, the utility function $u(x) = \lim_{\theta \downarrow 0} \frac{1}{\theta}(1 - e^{-\theta x}) = x$ reflects a risk-neutral preference. When the utility function is selected as $u(x) = x$, we solve the stochastic control problem and give the close form of the value function $\mathbb{E} \left[\tilde{\Pi} \right] \Big|_{v=v^{**}} = \sup_{v \in \mathcal{A}} \mathbb{E} \left[\tilde{\Pi} \right] < \infty$, where $v^{**} \in \mathcal{A}$ is the optimal feedback control in the risk-neutral case, see in Appendix A.2. Moreover, when $u(x) = \frac{1}{\theta}(1 - e^{-\theta x})$, $u'(x) = e^{-\theta x} > 0$, $u''(x) = -\theta e^{-\theta x} < 0$, so $u(\cdot)$ is a increasing concave function. For any admissible control \tilde{v} , by Jensen's inequality,

$$\mathbb{E} \left[u \left(\tilde{\Pi} \right) \right] \Big|_{v=\tilde{v}} \leq u \left(\mathbb{E} \left[\tilde{\Pi} \right] \Big|_{v=\tilde{v}} \right) \leq u \left(\mathbb{E} \left[\tilde{\Pi} \right] \Big|_{v=v^{**}} \right) < \infty,$$

which shows the finiteness of the utility maximizing problem. In the following, we solve the utility maximization problem (2) and present solutions in closed form in the cases where the execution risk $m(v)$ is either a constant or zero. The case of zero execution risk is postponed and will be discussed in more details in Section 4.

3.1 Optimal execution with risk aversion

When the execution risk is constant, i.e., $m(v) \equiv m_0$ for some fixed constant m_0 , the utility maximizing problem reduces to the following Stochastic Linear Exponential Quadratic (SLEQ) control problem [23] [24]

$$\left\{ \begin{array}{l} \sup_{v \in \mathcal{A}} \mathbb{E} \left[u \left(\tilde{\Pi} \right) \right] = \sup_{v \in \mathcal{A}} \mathbb{E} \left[\frac{1}{\theta} \left(1 - \exp \left(\theta \beta (x_T - A)^2 - \int_0^T \theta (\mu(x_t - R_t) + \rho \sigma m_0 - \eta v_t^2 - \gamma v_t(x_t - A)) dt \right. \right. \right. \\ \left. \left. \left. - \int_0^T \theta \sigma (x_t - R_t) dW_t - \int_0^T \theta (\eta v_t + \gamma (x_t - A)) m_0 dZ_t \right) \right) \right], \\ \text{s.t. } x_t = x_0 - \int_0^t v_s ds + m_0 Z_t. \end{array} \right. \quad (3)$$

Define the value function of the problem by

$$V(t, x) := \sup_{v \in \mathcal{A}} \mathbb{E} \left[\frac{1}{\theta} \left(1 - \exp \left(\theta \beta (x_T - A)^2 - \int_t^T \theta (\mu(x_s - R_s) + \rho \sigma m_0 - \eta v_s^2 - \gamma v_s (x_s - A)) ds \right. \right. \right. \\ \left. \left. \left. - \int_t^T \theta \sigma (x_s - R_s) dW_s - \int_t^T \theta (\eta v_s + \gamma (x_s - A)) m_0 dZ_s \right) \right) \middle| x_t = x \right], \quad (4)$$

where $\mathcal{A}_t := \{v \text{ is progressively measurable in } [t, T] \text{ and } \int_t^T \mathbb{E}[v_s^2] ds < \infty\}$. The optimal feedback control can be obtained in closed form. We summarize the result in the following theorem.

Theorem 1. *The value function (4) of the utility maximization problem has the closed form expression*

$$V(t, x) = \frac{1}{\theta} \left(1 - \exp \left\{ \left(b_2(t) + \frac{\theta \gamma}{2} \right) (x - A)^2 + b_1(t)(x - A) + b_0(t) \right\} \right).$$

where the parameters H , l_1 , and l_3 are given by

$$\begin{cases} H = \theta \eta + \frac{1}{2} \theta^2 \eta^2 m_0^2, \\ l_3 = \frac{1}{2} m_0 \eta \theta^2 \rho \sigma, \\ l_1 = \frac{\theta^2 \sigma^2}{2} H. \end{cases}$$

The time dependent function b_2 is given in closed form by

$$b_2(t) := \sqrt{l_1} \coth \left(A_0 + \frac{\sqrt{l_1}}{H} (T - t) \right) - l_3, \quad \text{where } A_0 := \coth^{-1} \left(\frac{l_3 + \frac{\theta}{2}(2\beta - \gamma)}{\sqrt{l_1}} \right),$$

b_1 is the solution to the following terminal value problem

$$b'_1(t) = \theta \mu + \frac{1}{H} b_1(t) (b_2(t) + l_3) + \frac{\eta \theta^2 \sigma}{H} (R_t - A) \left(\theta \sigma + \frac{1}{2} m_0^2 \theta^2 \eta \sigma (1 - \rho^2) - \rho m_0 b_2(t) \right), \quad b_1(T) = 0,$$

and

$$b_0(t) := \int_t^T \left[\frac{l_1 - l_3^2}{H} (R_s - A)^2 + \left(\theta \mu + \frac{l_3}{H} b_1 \right) (R_s - A) + \left(-\frac{b_1^2}{4H} + m_0^2 \left(b_2 + \frac{\theta \gamma}{2} \right) - \rho m_0 \theta \sigma \right) \right] ds.$$

Moreover, the optimal feedback control $v^* \in \mathcal{A}$ of the utility maximization problem (2) is given by

$$v_t^* = \frac{(1 + \theta \eta m_0^2) b_2(t)}{H} \cdot (x_t - A) + \frac{1 + \theta \eta m_0^2}{2H} \cdot b_1(t) + \frac{l_3}{H} \cdot (R_t - x_t), \quad (5)$$

It's worth mentioning that this theorem does not contain much mathematical breakthrough, instead, we regard it as a fundamental result to introduce valuable results in section 4, where the affine structure is what we find the most interesting.

Remark 1. *The optimal feedback control v_t^* in (5) consists of three parts:*

- The first part $\frac{(1 + \theta \eta m_0^2) b_2(t)}{H} \cdot (x_t - A)$ has the same sign as $x_t - A$. Without loss of generality, we assume that $x_0 > A$, when $x_t \geq A$, which implies the reallocation process has not finished, so the broker should

continue to sell the stock. Conversely, when $x_t < A$, which implies the strategy is over-shooting, so the broker should buy some shares of the stock back.

- The second part $\frac{1+\theta\eta m_0^2}{2H} \cdot b_1$, which is a pre-specified deterministic function in time t , does not depend on the state variable x_t .
- The third part $\frac{l_3}{H} \cdot (R_t - x_t)$ has the same sign as $\rho(R_t - x_t)$, where ρ is the correlation between the stock price process and the execution risk. The discussions in [25] state the fact that the sign of ρ depends on the order type adopted by the trader due to adverse selection: when the trader uses market orders, the correlation ρ is negative while ρ is positive whenever trading with limit orders. In this paper, we regard ρ as a market parameter reflecting whether the market is trader-friendly or not.
 - When $\rho \geq 0$, the stock price and execution risk tend to increase or decrease together: when the stock price increases, execution risk makes traders buy more or sell less than the amount they submit, which is benefit to the traders; when the stock price decreases, execution risk forces traders to buy less or sell more, which is also benefit to the traders. In such trader-friendly market, like the figure (1a), the broker should keep away from the reference strategy R_t because an overly conservative strategy is not necessary.
 - On the contrary, $\rho < 0$ implies a trader-harmful market: when the stock price increases, execution risk makes traders buy less or sell more than the amount they submit, which is harm to the traders; when the stock price decreases, execution risk forces traders to buy more or sell less, which is also harm to the traders. In such trader-harmful market, like the figure (1b), the broker should keep close to the reference strategy R_t to attain lower risk.

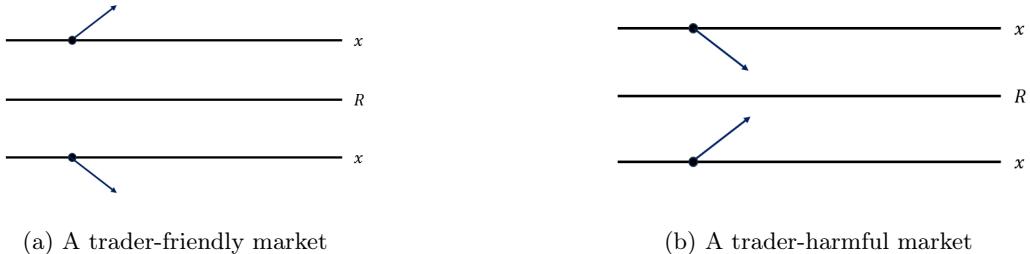


Figure 1: Market parameter ρ

Remark 2. By applying the optimal strategy (5), one obtains the optimal liquidation trajectory x_t^* as follows. For simplicity, take the parameters $A = \rho = 0$, $R_t \equiv 0$ and $\mu = 0$, x_t^* satisfies the SDE

$$dx_t^* = - \left(\frac{(1 + \theta\eta m_0^2)b_2}{H} x_t^* + \frac{(1 + \theta\eta m_0^2)b_1}{2H} \right) dt + m_0 dZ_t,$$

as $\beta \rightarrow +\infty$, x_t^* has the limit

$$x_t^{**} = \left(\frac{\sinh \left(\frac{\sqrt{l_1}}{H}(T-t) \right)}{\sinh \left(\frac{\sqrt{l_1}}{H}T \right)} \right)^{\frac{2H}{\theta\eta}-1} x_0 + m_0 \int_0^t \left(\frac{\sinh \left(\frac{\sqrt{l_1}}{H}(T-s) \right)}{\sinh \left(\frac{\sqrt{l_1}}{H}(T-s) \right)} \right)^{\frac{2H}{\theta\eta}-1} dZ_s.$$

Furthermore, when $m_0 \rightarrow 0$, $H \rightarrow \theta\eta$, the limit of the expression above is

$$\lim_{m_0 \rightarrow 0} x_t^{**} = \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)} x_0,$$

where $\kappa = \sqrt{\frac{\theta\sigma^2}{2\eta}}$. It recovers the classical implementation shortfall (IS) strategy (see in section 4.2), which implies the fact that our strategy can be regarded as a version of adaptive IS strategy. Furthermore, we may find that

$$\lim_{m_0 \rightarrow 0} \frac{x_t^{**} - \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)} x_0}{m_0} = \int_0^t \frac{\sinh(\kappa(T-t))}{\sinh(\kappa(T-s))} dZ_s,$$

which is an Ornstein-Uhlenbeck bridge [26]. Therefore, when m_0 is small, one may find an approximation to the optimal strategy

$$x_t^{**} \approx \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)} x_0 + m_0 \int_0^t \frac{\sinh(\kappa(T-t))}{\sinh(\kappa(T-s))} dZ_s,$$

which is a classical IS strategy plus an Ornstein-Uhlenbeck bridge, which reduces the strategy's inventory risk. Figure (2) shows the inventory risk of the optimal strategy and the classical IS strategy, which shows the fact that our optimal strategy has lower inventory risk than the classical IS strategy, especially at the end of the trading process.

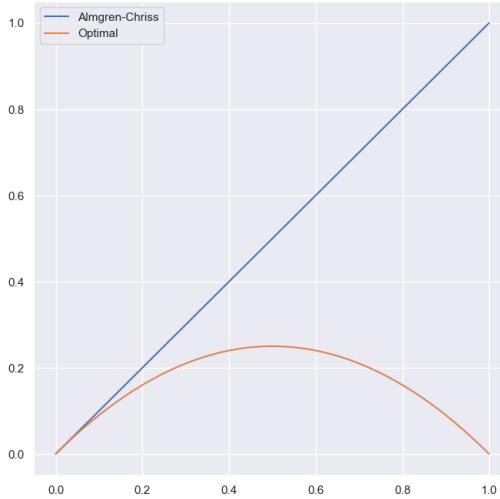


Figure 2: Inventory risk of the optimal strategy and the classical IS strategy

To summary the section, it should be emphasized that despite the optimal feedback control behaves linearly in the states carries a certain meaning which we discuss in Remark 1, however, the case which we care most is where the execution risk does not exist, which is a limit case of Theorem 1, discussed in the next section.

4 Zero execution risk

For the purpose of better understanding as well as visualizing the optimal strategy obtained in (5), we present in this section the trading trajectory x_t under optimal strategy in the case when $m_0 = 0$ for various reference strategies. In Section 4.1, we give the optimal strategy for general reference strategies. In Section 4.2, we present two classic trading strategies: IS order and TC order, both widely used in the market, and show that

they are special cases of our model. Next, we consider the case when the reference strategy is an endpoints-only one in Section 4.3, of which the optimal strategy has a heuristic form. Last but not least, piece-wise constant reference strategies are studied in Section 4.4.

Notice that in this case the setting of the problem reduces to that of the Almgren-Chriss but subject to a reference strategy. Furthermore, we shall set the final penalty parameter $\beta \rightarrow +\infty$, indicating that a final block trade at terminal time T is strictly prohibited.

4.1 General reference strategy

For a given generic reference strategy R_t , the following theorem shows a representation for the trading trajectory x_t under optimal control v^* .

Theorem 2. *Let $m(v) \equiv 0$ and $\beta \rightarrow +\infty$. The trajectory x_t under the solution to the optimization problem (2) is given by*

$$x_t = \frac{\kappa}{\sinh \kappa T} \int_0^T \left(R_s \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ + \left(x_0 - \frac{\mu}{\theta \sigma^2} \right) \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + A + \left(-A + \frac{\mu}{\theta \sigma^2} \right) \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T},$$

with the tuning parameter $\kappa := \sqrt{\frac{\theta \sigma^2}{2\eta}}$.

A few remarks on the trajectory x_t shown in Theorem 2 are in order.

Remark 3. *The parameter κ can be regarded as a tuning parameter between the TWAP and the reference strategies if $\mu = 0$. Specifically, in the limit as κ approaches zero, the optimal strategy converges to the TWAP strategy, i.e., for any $t \in [0, T]$, we have*

$$\lim_{\kappa \rightarrow 0} x_t = A + (x_0 - A) \cdot \left(1 - \frac{t}{T} \right).$$

On the other extreme as $\kappa \rightarrow \infty$, the optimal trajectory converges to the reference strategy R_t itself

$$\lim_{\kappa \rightarrow \infty} x_t = R_t$$

for any $t \in [0, T]$.

Remark 4. *The optimal trajectory may also be written as an “affine transformation” as*

$$x_t = \frac{\mu}{\theta \sigma^2} + a_t \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + b_t \cdot \frac{\sinh \kappa t}{\sinh \kappa T},$$

where

$$a_t = x_0 - \frac{\mu}{\theta \sigma^2} + \int_0^t \kappa R_s \sinh \kappa s ds, \\ b_t = A - \frac{\mu}{\theta \sigma^2} + \int_t^T \kappa R_s \sinh \kappa(T-s) ds.$$

A more detailed discussion on this affine transformation can be found in Section 4.4.

Next, we further specialize the reference strategies and present their corresponding trajectories in the following sections.

4.2 Implementation shortfall (IS) order and target close (TC) order

Assume $\mu = 0$ and the broker is asked to take as the reference strategy a block trade of size $|A - x_0|$ at $t = 0$, i.e.,

$$R_t^{\text{IS}} = \begin{cases} x_0, & t = 0, \\ A, & 0 < t \leq T. \end{cases}$$

The corresponding marked-to-market P&L of the reference strategy R^{IS} is given by

$$\begin{aligned} \Pi^{\text{IS}} := \int_0^T (-S_t^{\text{IS}}) dR_t^{\text{IS}} &= -(x_0 - A) \cdot \frac{S_{0-}^{\text{IS}} + S_{0+}^{\text{IS}}}{2} \\ &= -(x_0 - A)S_0 + \frac{\gamma}{2}(x_0 - A)^2, \end{aligned}$$

which is equivalent to choosing the initial price S_0 as reference price. This model is referred to as an *implementation shortfall model* in [27], of which the optimal strategy is

$$x_t = A + (x_0 - A) \cdot \text{IS}_t, \quad (6)$$

where

$$\text{IS}_t := \frac{\sinh \kappa(T-t)}{\sinh \kappa T}. \quad (7)$$

We shall refer to the strategy IS_t as the *unit implementation shortfall (IS) order*. Hence, the trajectory in (6) can be interpreted as executing $x_0 - A$ unit IS orders if a broker is missioned to make transition of his position from x_0 to A shares. We remark that, as we will show in Section 4.4, the unit IS orders are used as one of the building blocks for the construction of optimal trajectories for more general reference strategies.

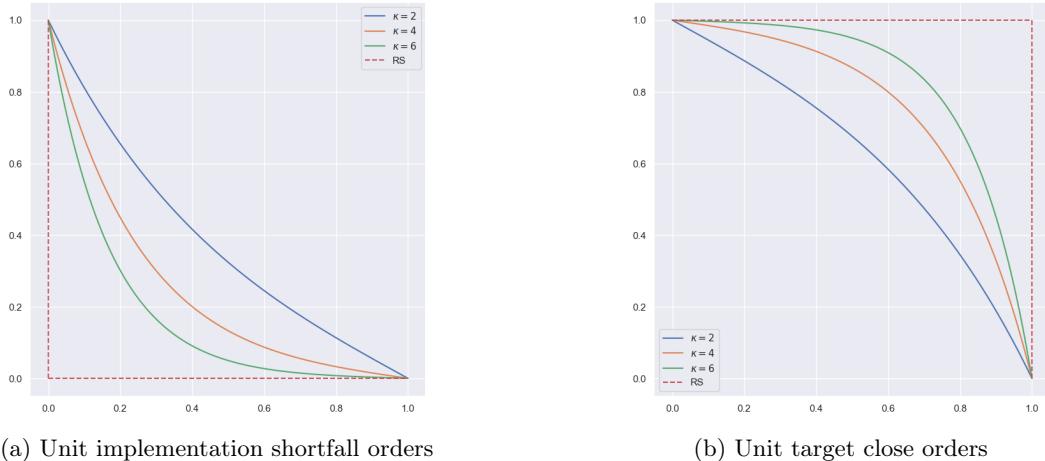


Figure 3: Basic strategies in varying parameter κ .

Figure (3a) shows examples of the unit implementation shortfall orders under various values of κ . Notice that, since $\kappa = \sqrt{\frac{\theta\sigma^2}{2\eta}}$, as θ, σ increase or η decreases, thus κ increases, the trajectories suggest the broker

trade faster at the beginning and more slowly for the rest of the reallocation process. Indeed, the larger the κ , the faster the trading at the beginning. The financial rationale is as follows. Recall that θ is the coefficient of relative risk aversion for the utility function and σ is the volatility of stock price. Hence, if either of the parameters increases, the broker is more concerned with the price risk than the execution risk, he tends toward trading faster at the beginning and more slowly thereafter. Also, since η is the coefficient of slippage, as it decreases, the slippage decreases as well, resulting in lower cost from faster transactions. Therefore, again to mitigate the concern of price risk during reallocation process, the broker should also trade faster at the beginning as well.

Finally, we remark that, as $\kappa \rightarrow 0$, the unit IS order (7) converges to the **TWAP** (Time-Weight-Average-Price) strategy: for any $t \in [0, T]$,

$$\lim_{\kappa \rightarrow 0} \text{IS}_t = \lim_{\kappa \rightarrow 0} \frac{\sinh \kappa(T-t)}{\sinh \kappa T} = \frac{T-t}{T}.$$

On the flip side, as $\kappa \rightarrow +\infty$, the optimal trajectory (7) converges to a block trade at time 0:

$$\lim_{\kappa \rightarrow \infty} \text{IS}_t = \lim_{\kappa \rightarrow \infty} \frac{\sinh \kappa(T-t)}{\sinh \kappa T} = \lim_{\kappa \rightarrow \infty} e^{-\kappa t} \cdot \frac{1 - e^{-2\kappa(T-t)}}{1 - e^{-2\kappa T}} = \begin{cases} 1, & t = 0, \\ 0, & 0 < t \leq T. \end{cases}$$

Assume again that $\mu = 0$. Here the broker is given as the reference strategy to take only a block trade of size $|A - x_0|$ at terminal time T . That is,

$$R_t^{\text{TC}} = \begin{cases} x_0, & 0 \leq t < T, \\ A, & t = T. \end{cases}$$

The marked-to-market P&L Π^{TC} for this reference strategy is given by

$$\begin{aligned} \Pi^{\text{TC}} &= \int_0^T (-S_t^{\text{TC}}) dR_t^{\text{TC}} = -(x_0 - A) \cdot \frac{S_{T-}^{\text{TC}} - S_{T+}^{\text{TC}}}{2} \\ &= -(x_0 - A) S_{T+}^{\text{TC}} - \frac{\gamma}{2} (x_0 - A)^2, \end{aligned}$$

which is equivalent to choosing the stock price S_T at terminal time as the reference price. This model is referred to as an *target close model* [27][3], of which the optimal strategy is given by

$$x_t = A + (x_0 - A) \cdot \text{TC}_t, \quad (8)$$

where

$$\text{TC}_t := \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T}. \quad (9)$$

We refer to the strategy TC_t in (9) as the *unit target close (TC) order*. Thus, similar to the case of implementation shortfall order in Section 4.2, the trajectory in (8) can be interpreted as executing $x_0 - A$ unit TC orders if a broker is missioned to make transition of his position from x_0 to A shares. Likewise as for the unit IS orders, we show in Section 4.4 that the unit TC orders are used as the other building blocks for the

construction of optimal trajectories for more general reference strategies.

Figure (3b) shows examples of plots for unit target close orders in different values of κ . We observe that the trading trajectories are concave in time as opposed to convex in time as that of the unit IS orders. Also, as κ increases, the optimal trajectory moves towards the reference strategy and towards the TWAP trajectory as κ decreases. In fact, similar to the case for implementation shortfall, in the limit as $\kappa \rightarrow 0$, the unit TC order (9) converges to the TWAP strategy: for any $t \in [0, T]$,

$$\lim_{\kappa \rightarrow 0} \text{TC}_t = \lim_{\kappa \rightarrow 0} \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T} = 1 - \frac{t}{T};$$

whereas in the other extreme as $\kappa \rightarrow +\infty$, the optimal trajectory converges to a block trade at terminal time T :

$$\lim_{\kappa \rightarrow \infty} \text{TC}_t = \lim_{\kappa \rightarrow \infty} \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T} = 1 - \lim_{\kappa \rightarrow \infty} e^{-\kappa(t-T)} \cdot \frac{1 - e^{-\kappa t}}{1 - e^{-\kappa T}} = \begin{cases} 0, & 0 \leq t < T, \\ 1, & t = T. \end{cases}$$

4.3 Endpoints-only reference strategy

In this section, we consider the following reference strategy R_t , which we called the *endpoints-only*.

$$R_t = \begin{cases} x_0, & t = 0, \\ R, & 0 < t < T, \\ A, & t = T. \end{cases} \quad (10)$$

The strategy is simply to hold R shares during the entire reallocation process, except at the initial and terminal times.,

The following lemma will prove itself useful in the determination of the optimal strategy subject to the reference strategy (10).

Lemma 1. *We have that*

$$\begin{aligned} & \frac{\kappa}{\sinh \kappa T} \int_0^T \left(\sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ &= \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T} - \frac{\sinh \kappa(T - t)}{\sinh \kappa T} \\ &= \text{TC}_t - \text{IS}_t. \end{aligned}$$

In other words, the given integral can be written as the difference between a unit TC and a unit IS order.

We summarize the result for optimal strategies subject to (10) in the proposition that follows.

Proposition 2. *With the reference strategy given in (10), the optimal trajectory is given by an affine combination of a unit IS order and a unit TC order as*

$$x_t = \left(x_0 - \frac{\mu}{\theta \sigma^2} - R \right) \cdot \text{IS}_t + \left(-A + \frac{\mu}{\theta \sigma^2} + R \right) \cdot \text{TC}_t + A. \quad (11)$$

Note that when $\mu = 0$ and $R = x_0$, the optimal strategy reduces to the following TC strategy as in (8)

$$x_t = (x_0 - A) \cdot TC_t + A;$$

whereas when $\mu = 0$ and $R = A$, the optimal strategy reduces to the following IS strategy as in (6)

$$x_t = (x_0 - A) \cdot IS_t + A.$$

Note that for any given reference level R , the optimal trajectory in (11) at any point in time is an affine combination of a unit IS order and a unit TC order. The coefficients represent a trade-off between the two strategies. For example, if R is closer to the initial position x_0 and further away from the target position A , the strategy in (11) suggests we weigh in more on the unit TC order and less on the unit IS order; conversely, one should then weigh in more on the unit IS order. Figure (4) shows the plots of optimal trajectories in different reference levels R , assuming $x_0 = 1$ and $A = 0$. Again, we end up with a unit TC order when $R = x_0$ and a unit IS order when $R = A$.

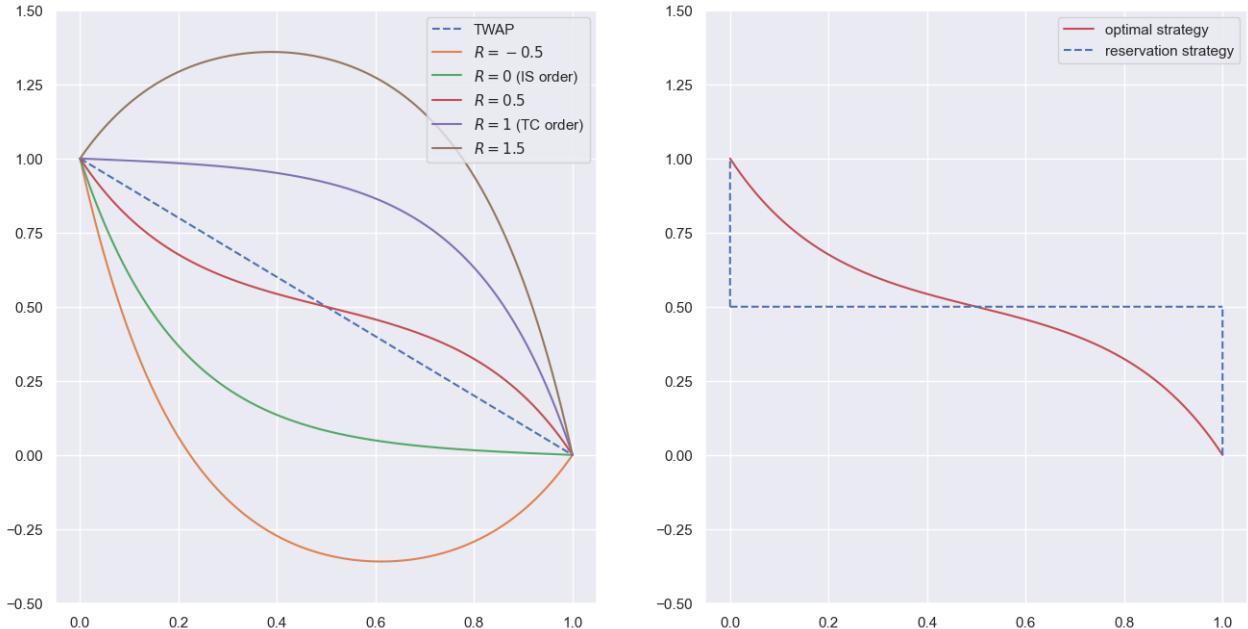


Figure 4: Optimal strategies in different reference levels R .

As shown in Figure 4, when $R > x_0 = 1$ the optimal strategy suggest overshoot the target, meaning buying more than needed then sell back later and quicker when time approaches the terminal time. On the other hand, if $R < A = 0$, the optimal strategy suggests undershoots the target. We provide conditions on the parameters under which an overshooting or an undershooting occur in the following proposition.

Proposition 3. *Without loss of generality, assume $x_0 > A$. Let x_t be the optimal strategy given in (11). We have that, if $R + \frac{\mu}{\theta\sigma^2} > x_0$, then x_t is concave in t and it becomes convex in t if $R + \frac{\mu}{\theta\sigma^2} < A$. It follows that x_t overshoots if and only if*

$$R + \frac{\mu}{\theta\sigma^2} > x_0 + \frac{1}{\cosh \kappa T - 1} \cdot (x_0 - A)$$

and it undershoots if and only if

$$R + \frac{\mu}{\theta\sigma^2} < A - \frac{1}{\cosh \kappa T - 1} \cdot (x_0 - A).$$

It appears that, when μ is nonzero, the terms in the optimal trajectory (11) that have μ involved always come in the form of $\frac{\mu}{\theta\sigma^2}$. We provide a financial rationale of this combination of parameters as follows. Note that the combination resembles the Merton ratio in Merton's optimal portfolio problem in the market consisting of two assets: one risky and the other risk-less. The agent is risk averse with power utility $u(x) = x^\gamma$ and the dynamic of risky asset is governed by a geometric Brownian motion with expected return μ and volatility σ . The risk-less asset is assumed accruing zero interest. In this setting, the Merton ratio, which represents the percentage of wealth invested in the risky asset, is given by $\frac{\mu}{(1-\gamma)\sigma^2}$. In our setting, the broker is with exponential utility $u(x) = \frac{1}{\theta}(1 - e^{-\theta x})$ and the stock price S_t is assumed following an arithmetic Brownian motion with drift μ and volatility σ , i.e.,

$$dS_t = \mu dt + \sigma dW_t.$$

As in the derivation of the Merton ratio in Merton's problem, one may show that the "Merton ratio" [28] in this case is given by $\frac{\mu}{\theta\sigma^2}$ except that this ratio does not represent the percentage of wealth invested in the risky asset as in the original Merton's problem. Indeed, this ratio represents the number of shares to be held in the optimal portfolio.

4.4 Piece-wise constant reference strategy

We consider in this section the reference strategies that are piece-wise constant across time. The consideration is that, to possibly account for real time market environments during the entire reallocation process, the broker's strategies may be subject to interval TWAP or VWAP strategies. For example, in an intraday trading activity, it is documented that the market trades much more actively at the times close to opening and closing than that in the middle of common trading days. Also, in the case where trading horizon across multiple days, the broker is most likely treating each trading day individually. In these regards, a piece-wise constant reference strategy is considered plausible.

Theoretically, as will be demonstrated, optimal trajectories subject to piece-wise constant reference strategies exhibit an elegant algebraic structure, derived from the unit IS and unit TC orders discussed in Sections 4.2. By applying standard approximation and limiting processes for integrable functions, this structure offers practical and insightful guidance for developing optimal strategies under general integrable reference strategies. Concretely, a generic piece-wise constant reference strategy R_t is defined as

$$R_t = (x_0 - R^{(1)}) \mathbb{1}_{\{t=0\}} + \sum_{k=1}^n R^{(k)} \mathbb{1}_{\left\{ \frac{(k-1)T}{n} \leq t < \frac{kT}{n} \right\}} + A \mathbb{1}_{\{t=T\}}, \quad (12)$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function, n is the number of periods, and the $R^{(k)}$'s are fixed constants. The following proposition summarizes the optimal trajectory under the piece-wise constant reference strategy in (12).

Proposition 4. When $\mu = m_0 = 0, \beta \rightarrow +\infty$, denote $R^{(0)} := x_0$ and $R^{(n+1)} := A$, the solution to the previous optimization problem under the piece-wise constant reference strategy given in (12) is of the following form, for $1 \leq k \leq n$,

$$x_t = \left\{ a_k + \left(R^{(k)} - a_k \right) \text{TC}_{t-\frac{(k-1)T}{n}} + \left(a_{k-1} - R^{(k)} \right) \text{IS}_{t-\frac{(k-1)T}{n}} \right\} \mathbb{1}_{\left\{ \frac{(k-1)T}{n} \leq t < \frac{kT}{n} \right\}}, \quad (13)$$

with $\kappa := \sqrt{\frac{\theta\sigma^2}{2\eta}}$,

$$\text{IS}_t := \frac{\sinh \kappa \left(\frac{T}{n} - t \right)}{\sinh \kappa \frac{T}{n}}, \quad \text{TC}_t := \frac{\sinh \kappa \frac{T}{n} - \sinh \kappa t}{\sinh \kappa \frac{T}{n}},$$

and a_k is denoted as a weighted average of $R^{(i)}$'s:

$$a_k := \frac{\sinh \kappa \left(T - \frac{kT}{n} \right)}{\sinh \kappa T} \cdot \sum_{i=0}^k b_i R^{(i)} + \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \sum_{i=k+1}^{n+1} b_{n-i+1} R^{(i)}, \quad 1 \leq k \leq n-1,$$

where

$$b_i = \begin{cases} 1, & i = 0, \\ \cosh \left(\kappa \frac{iT}{n} \right) - \cosh \left(\kappa \frac{(i-1)T}{n} \right), & 1 \leq i \leq n-1. \end{cases}$$

Notice that the IS_t and TC_t in Proposition 4 are indeed respectively the unit IS order given in (7) and the unit TC order in (9). Proposition 4 essentially states that, for a given piece-wise constant reference strategy, its corresponding optimal solution in each sub-interval is given by an affine transformation of a unit IS order and a unit TC order. In this sense we may regard the unit IS orders and the unit TC orders obtained in Sections 4.2 as the building blocks or bases for optimal trajectories under piece-wise constant reference strategies. The algebraic structure of the optimal trajectories is thus depicted by the affine space generated by the unit IS and unit TC orders. What are left to be determined are the coefficients that are sub-interval dependent. In practice, one should firstly calculate a_k 's by linear combinations of $R^{(i)}$'s, then optimal strategy is obtained by connecting a_k 's using unit IS and unit TC orders. As an example, Figure (5) shows the plots of a piece-wise constant reference strategy (in dotted red) its optimal trajectory (in blue) with three sub-intervals $n = 3$, the other parameters are given by $T = 3, \kappa = 5, R^{(1)} = \frac{15}{2}, R^{(2)} = \frac{9}{2}, R^{(3)} = \frac{3}{2}, x_0 = 9, A = 0$. To conclude the

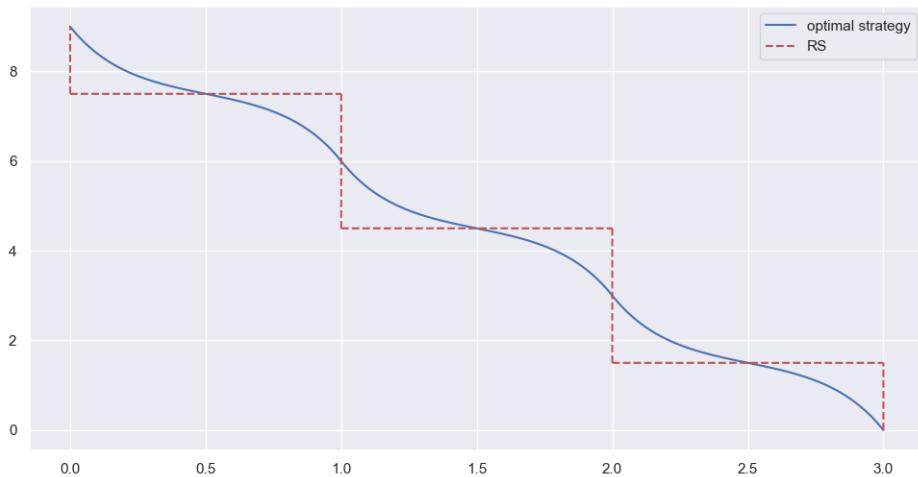


Figure 5: A three-period piece-wise constant reference strategy and its corresponding optimal trajectory.

section, we show in Theorem 3 a stability theorem for optimal strategies subject to general, smooth enough, reference strategies.

Theorem 3. *For any given reference strategies R_t and \tilde{R}_t satisfying*

$$\int_0^T \left(R_t - \tilde{R}_t \right)^2 dt < \varepsilon,$$

for some $\varepsilon > 0$. Let x_t and \tilde{x}_t be the corresponding optimal trajectories for R_t and \tilde{R}_t , respectively. We have that, for any $0 \leq t \leq T$,

$$|x_t - \tilde{x}_t| < \frac{1}{2} \sqrt{\kappa \varepsilon}.$$

The above theorem basically states the fact that when two reference strategies are close in the L^2 sense, their corresponding optimal strategies remain close in the sup norm sense. Note that for any $\varepsilon > 0$ and an almost everywhere continuous bounded RS R_t , there exists a piece-wise constant RS \tilde{R}_t such that

$$\int_0^T \left(R_t - \tilde{R}_t \right)^2 dt < \varepsilon,$$

It follows that the optimal trajectory in (13) for \tilde{R}_t , which is an affine transformation of unit IS and unit TC orders in each sub-interval, can be applied as an approximation to the optimal trajectory under reference strategy R_t , with error estimate given in Theorem 3.

5 Numerical examples

We conduct in this section numerical experiments on the implementation of optimal strategy obtained in Theorem 1 and stress testing the strategy against various parameters. Monte Carlo simulations are implemented to illustrate sample trading trajectories and the performances criteria of the optimal strategy versus those of a related TWAP strategy. The performance of the optimal and TWAP strategies are then stress tested against certain extreme parameters. We remark that the parameters chosen for the numerical examples in this section are for convenience only. In practice, parameters are supposedly calibrated to market data prior to implementation, causing possible issues associated with estimation risk.

5.1 Sample trading trajectories

We present in Figures (6) through (11) sample trading trajectories under the optimal and a related TWAP strategies in various parameters. Parameters chosen as base case are $(m_0, \eta, \rho, \theta, \sigma, \beta) = (0.05, 10, 0, 0.002, 200, 1000)$. Table 1 summarizes the parameters that vary, while the others are held fixed, in each case and their corresponding figures. The reference strategy R_t is selected as a block trade at $t = 0$, as shown in Section 4.2.

Figure (6) illustrates sample trading trajectories for the optimal and the TWAP strategies in different values of m_0 . We note that, since m_0 quantifies the execution risk, as m_0 increases both trajectories become more volatile and fluctuating.

Table 1: Sample trading trajectory in varying parameters

Figure number	Parameter	Case 1	Case 2	Case 3
Figure 6	m_0	0.00	0.05	0.10
Figure 7	η	5	10	20
Figure 8	ρ	-0.9	0	0.9
Figure 9	θ	0.001	0.002	0.004
Figure 10	σ	100	200	400
Figure 11	β	100	1000	10000

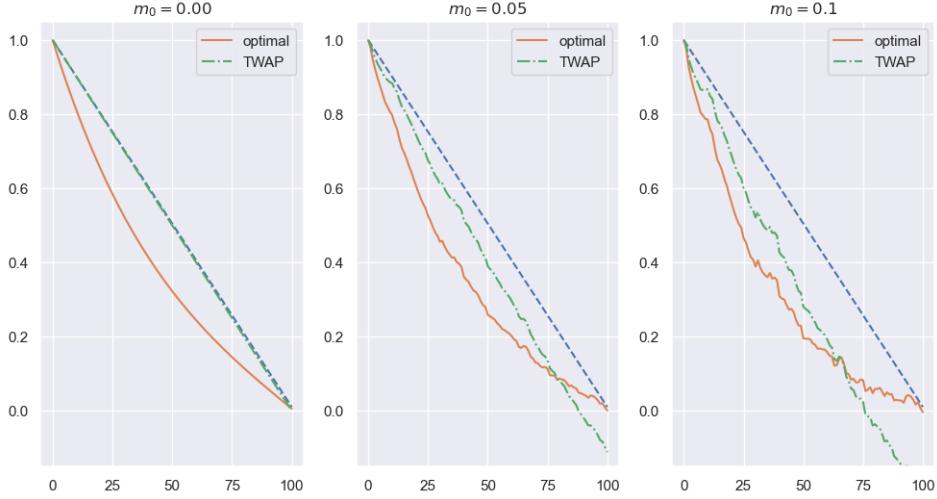


Figure 6: Sample trading trajectories in different m_0

Figure (7) exhibits sample trading trajectories in varying values of η . It is worth mentioning that, since η reflects the level of transaction costs, as the value of η increases, transaction cost becomes more significant in the determination of optimal strategies. Consequently, the optimal strategy tends to align more closely with the classical TWAP strategy. In other words, a higher η places greater emphasis on minimizing transaction costs, motivating the adoption of a strategy that mirrors the TWAP approach, which aims for consistent execution over time.

Figure (8) shows optimal strategies in the values of $\rho = -0.5, 0, 0.5$. Recall that ρ denotes the correlation between the stock price and the execution risk. We note that there seems no significant dependence of the optimal strategies on ρ within this set of parameters.

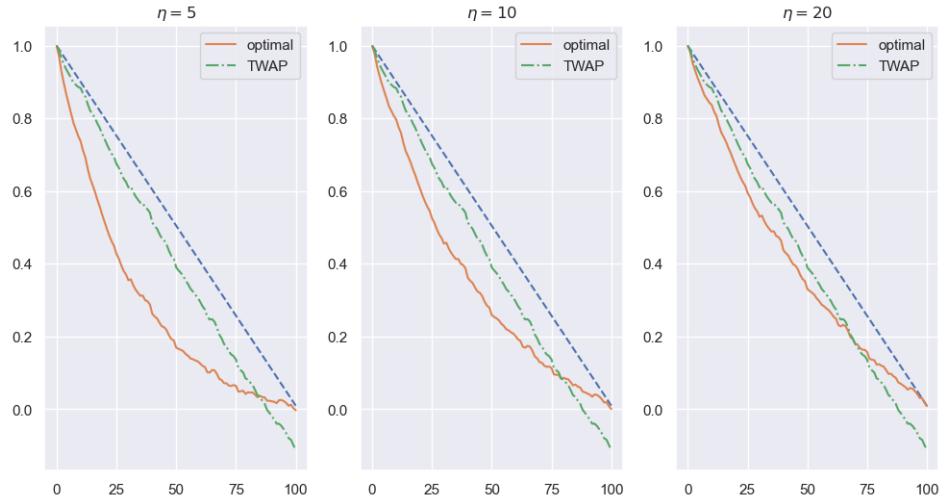


Figure 7: Sample trading trajectories in different η

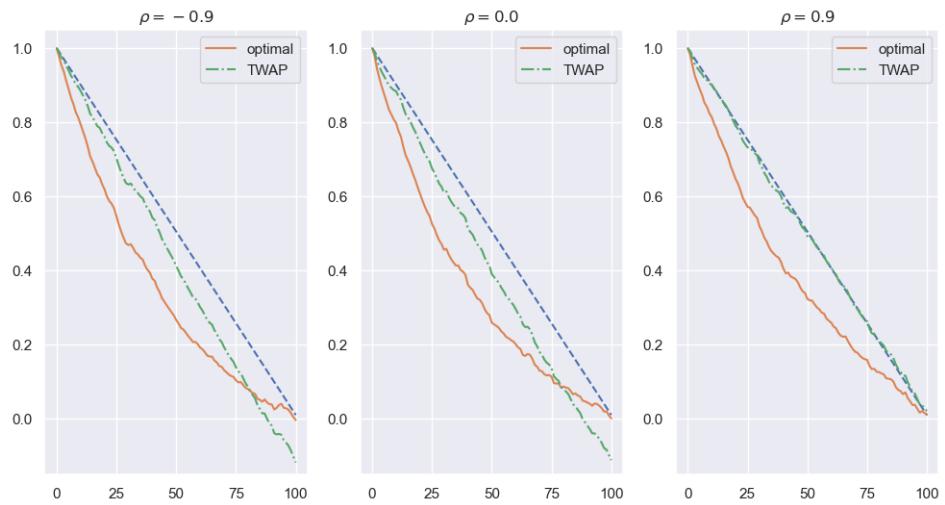


Figure 8: Sample trading trajectories in different ρ

Figure (9) illustrates the optimal strategies in the values of $\theta = 0.001, 0.002, 0.004$, which represents the broker's taste of risk aversion. The larger the value of θ , the more risk averse the broker tends to be. It follows that the sample trading trajectory under optimal strategy gradually converges towards the reference strategy as θ increases, indicating the broker's inclination to adopt a more conservative and risk-averse approach.

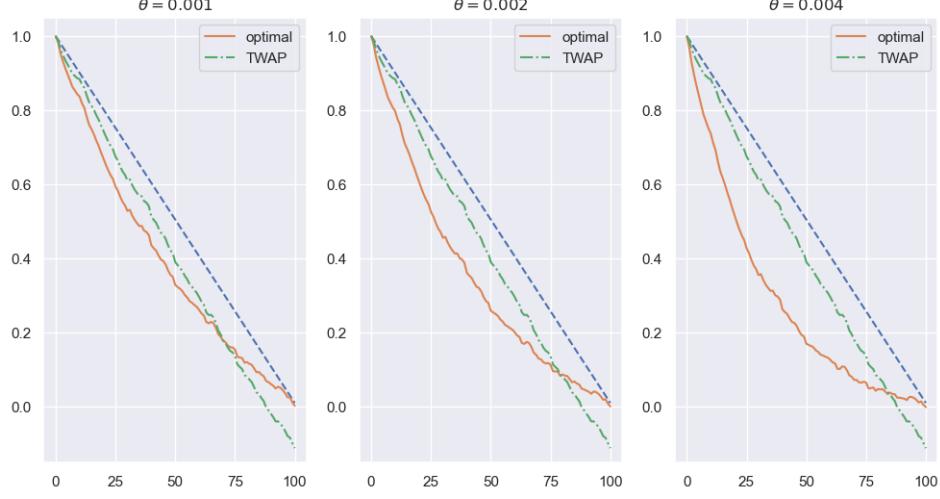


Figure 9: Sample trading trajectories in different θ

Figure (10) shows sample trading trajectories under optimal strategies with respect to the stock volatility $\sigma = 100, 200, 400$. As σ increases, the price risk increases. Thus, to mitigate the risk incurred from the price volatility during execution, the optimal strategy tracks more closely to the reference strategy.

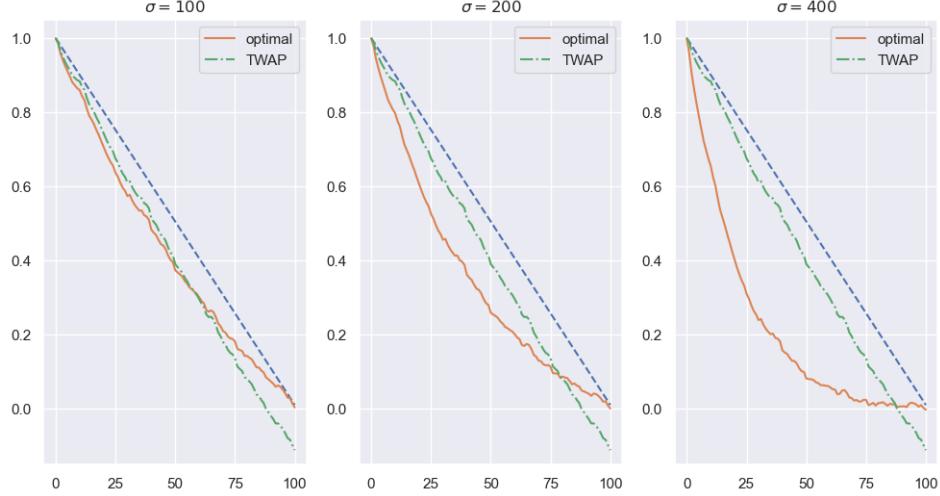


Figure 10: Sample trading trajectories in different σ

Figure (11) presents the optimal strategies in the values of $\beta = 100, 1000, 10000$, which quantifies the penalty of block trade at terminal time. Large β indicates that any remaining inventory at terminal time is unfavorable. Therefore, it is seen in the figure that the terminal inventory x_T^* of the sample trading trajectories under optimal strategy gradually approaches zero. It is thus fathomable that, as β approaches infinity, a final block trade at terminal time turns into strictly prohibited.

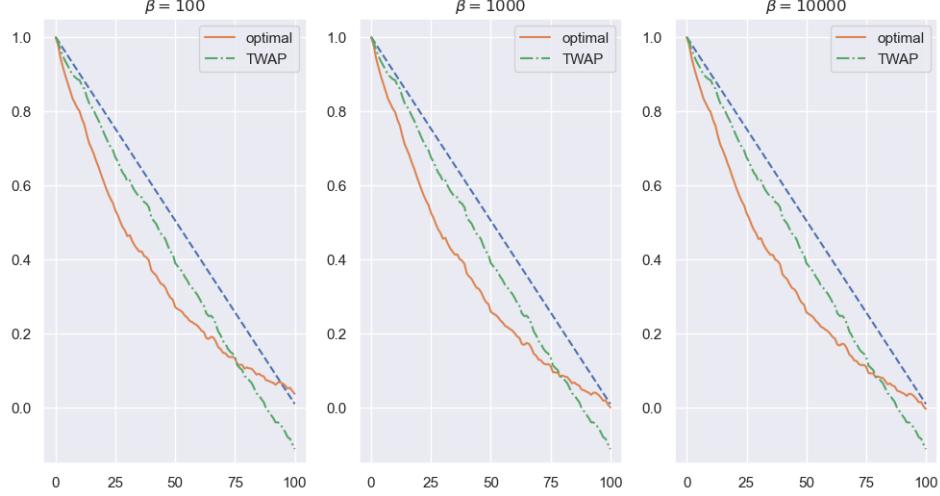


Figure 11: Sample trading trajectories in different β

5.2 Performance analysis and stress test

To demonstrate the performance of the optimal strategy against TWAP, we implement the strategies using Monte Carlo simulations and present the resulting histograms and boxplots of the terminal wealths as well as the utilities at the investment horizon. This performance analysis serves to illustrate the optimality and robustness of the optimal strategy.

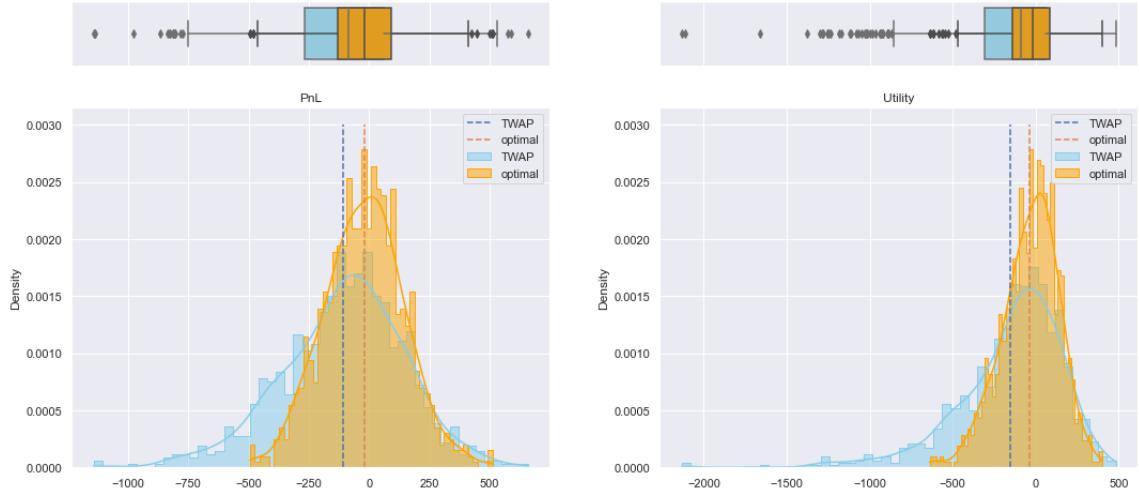


Figure 12: Histograms and boxplots of terminal wealths and utilities under the optimal and the TWAP strategies. Terminal wealth on the left, utility on the right. Vertical dashed lines indicate sample means.

Figure (12) exhibits the histograms and boxplots of the terminal P&L (on the left) and the utility (on the left) at investment horizon for the optimal and the TWAP strategies. We observe that the optimal

Table 2: Stress testing parameters

	m_0	η	ρ	θ	σ	β
Baseline	0.05	10	0	0.002	200	1e6
Scenario 1 (large σ)	0.05	10	0	0.002	2000	1e6
Scenario 2 (large β)	0.05	10	0	0.002	200	1e7
Scenario 3 (large m_0)	5	10	0	0.002	200	1e6
Scenario 4 (small η)	0.05	0.1	0	0.002	200	1e6

strategy not only consistently achieves higher average terminal wealth and utility than those under TWAP but also bears lower variability. Furthermore, an intriguing observation is made regarding the tail ends of the distribution. Specifically, the TWAP strategy demonstrates a greater density in these tail regions, indicating a higher likelihood of incurring significant losses when compared to our optimal strategy. This discrepancy in density at the tail end reinforces the notion that the TWAP strategy may accentuate the potential for unfavorable outcomes, thus reflecting a subpar approach to risk management.

Finally, we stress test the optimal and the TWAP strategies against extreme parameters. While the remaining parameters are held the same as in the benchmark case, the stock volatility σ is scaled up by ten folds in Scenario 1, the terminal penalty coefficient β up by ten folds in Scenario 2, the execution risk m_0 up by a hundred folds in Scenario 3, and lastly in Scenario 4 down by 0.1 hundred folds the temporary cost η . Table 2 summarizes the parameters that are stress tested in each scenario and Figure 13 shows the histograms and boxplots of P&Ls at the terminal time under the optimal and the TWAP strategies.

Apparently, in all the scenarios the histograms under optimal strategies are concentrated around zero whereas those under TWAP strategies are much more widely spreading. Also, the sample means under optimal strategies are higher than those of TWAP strategies. We conclude that, even under extreme situations, the optimal strategies outperform TWAP strategies not only in higher averaged P&L but also with lower volatility risk, which shows the robustness of our results.

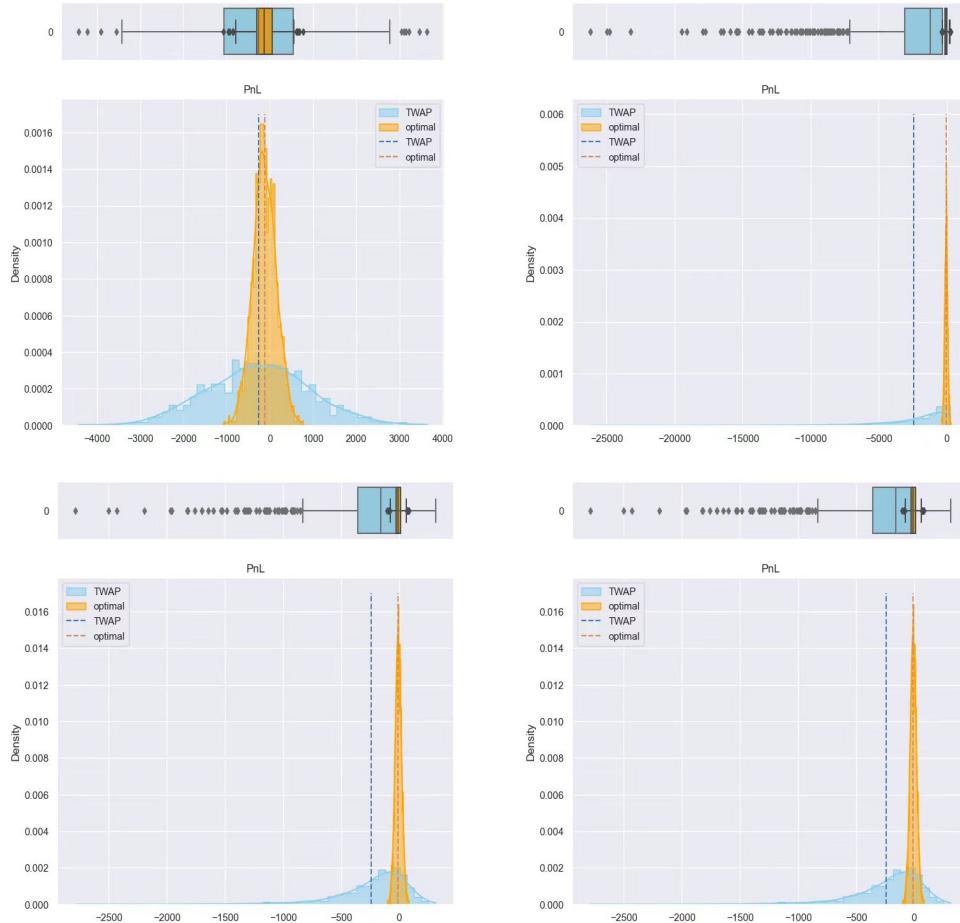


Figure 13: Histograms and boxplots of terminal P&L under stress tests. Scenarios 1 on the top left, 2 on the top right, 3 on the bottom left, and 4 on the bottom right. Vertical dashed lines indicate sample means.

6 Conclusion

In this article, we showed how the broker's order execution problem under execution risk subject to client's reference strategy can be recast as a utility maximization problem. The optimal solution to the utility maximization problem was obtained by solving an associated Riccati differential equation and was presented in feedback form. When execution risk vanishes, the optimal strategies become deterministic and were given in closed form. We showed that trading trajectories subject to the unit IS and the unit TC orders form a basis of an affine structure for trading trajectories under optimal strategies subject to general piece-wise constant reference strategies. As for general continuous reference strategies, we proved an approximation and stability theorem for the trading trajectory under the corresponding optimal strategy via the trajectories from piece-wise constant reference strategies.

Numerical experiments showed that the optimal strategies not only achieve higher expected values but also bear lower variabilities as opposed to those under TWAP strategies. Numerical performance analysis and stress tests confirmed that optimal strategies outperform TWAP strategies, by a wide margin in certain cases. Possible extensions of the framework considered in the paper include (a) a stochastic component to the price evolution reflecting the overall market activity such macroeconomics factors or news; (b) the price impact, rather than permanent, being transient with proper decay kernel. We leave these considerations to a future study.

Acknowledgement

This research is supported by the National Natural Science Foundation of China (Grants No.11971040).

A Appendix

A.1 Proof to Proposition 1

By Itô's formula,

$$\begin{aligned} (x_T - A)S_T = & (x_0 - A)S_0 + \int_0^T S_t dx_t + \int_0^T (\mu(x_t - A) + \rho\sigma m(v_t) - \gamma v_t(x_t - A)) dt \\ & + \int_0^T \sigma(x_t - A)dW_t + \int_0^T \gamma(x_t - A)m(v_t)dZ_t. \end{aligned}$$

Furthermore, note that $\tilde{S}_t = S_t + \eta v_t$,

$$\int_0^T (-\tilde{S}_t) dx_t = - \int_0^T S_t dx_t - \int_0^T \eta v_t^2 dt + \int_0^T \eta v_t m(v_t) dZ_t,$$

in summary,

$$\begin{aligned}\Pi &= (x_T - A)S_T - \beta(x_T - A)^2 + \int_0^T (-\tilde{S}_t) dx_t \\ &= (x_0 - A)S_0 - \beta(x_T - A)^2 + \int_0^T (\mu(x_t - A) + \rho\sigma m(v_t) - \eta v_t^2 - \gamma v_t(x_t - A)) dt \\ &\quad + \int_0^T \sigma(x_t - A)dW_t + \int_0^T (\eta v_t + \gamma(x_t - A))m(v_t)dZ_t.\end{aligned}$$

□

A.2 Risk-Neutral Case of Theorem 1

In Section (3) we recast the optimal execution problem as a utility maximization problem with the CARA preference $u(x) = \frac{1}{\theta}(1 - e^{-\theta x})$ with parameter $\theta > 0$, which reflects the broker's risk aversion. As $\theta \rightarrow 0$, the utility function $u(x) = \lim_{\theta \rightarrow 0} \frac{1}{\theta}(1 - e^{-\theta x}) = x$ describes a risk-neutral preference. With the admissible set \mathcal{A} and the excess P&L $\tilde{\Pi}$ given in Section (2.2) and Section (3), the risk-neutral expected utility is given by

$$\begin{aligned}\mathbb{E}[u(\tilde{\Pi})] &= \mathbb{E}[\tilde{\Pi}] = \mathbb{E}\left[-\beta(x_T - A)^2 + \int_0^T (\mu(x_t - R_t) + \rho\sigma m(v_t) - \eta v_t^2 - \gamma v_t(x_t - A)) dt\right. \\ &\quad \left.+ \int_0^T \sigma(x_t - R_t)dW_t + \int_0^T (\eta v_t + \gamma(x_t - A))m(v_t)dZ_t\right] \\ &= \mathbb{E}\left[-\beta(x_T - A)^2 + \int_0^T (\mu(x_t - R_t) + \rho\sigma m(v_t) - \eta v_t^2 - \gamma v_t(x_t - A)) dt\right].\end{aligned}$$

When the execution risk is constant (i.e. $m(\cdot) \equiv m_0$ for some $m_0 \geq 0$), the utility maximization problem reduces to the following Stochastic Linear Quadratic (SLQ) control problem [22]:

$$\begin{cases} \sup_{v \in \mathcal{A}} \mathbb{E}\left[-\beta(x_T - A)^2 + \int_0^T (\mu(x_t - R_t) + \rho\sigma m_0 - \eta v_t^2 - \gamma v_t(x_t - A)) dt\right], \\ \text{s.t. } x_t = x_0 - \int_0^t v_s ds + m_0 Z_t. \end{cases}$$

Define the value function $V(t, x)$ as

$$V(t, x) = \sup_{v \in \mathcal{A}_t} \mathbb{E}\left[-\beta(x_T - A)^2 + \int_t^T (\mu(x_s - R_s) + \rho\sigma m_0 - \eta v_s^2 - \gamma v_s(x_s - A)) ds \middle| x_t = x\right],$$

where $\mathcal{A}_t := \{v \text{ is progressively measurable in } [t, T] \text{ and } \int_t^T \mathbb{E}[v_s^2] ds < \infty\}$. Note that our problem satisfies assumption **(L1)** given in Section 3.1 of [22] due to the boundedness of the coefficients, so the problem is solvable, and by the dynamic programming principle, the value function V is unique satisfying the Hamilton-Jacobi-Bellman (HJB hereafter) equation

$$V_t + \frac{1}{2}m_0^2 V_{xx} + \sup_{v \in \mathbb{R}} \{-\eta v^2 - v(\gamma(x - A) + V_x)\} + \mu(x - R_t) + \rho\sigma m_0 = 0$$

with terminal condition $v(T, x) = -\beta(x - A)^2$. With first order criterion, the optimal feedback control is given by $v^* = -\frac{1}{2\eta}(\gamma(x - A) + V_x)$, hence the HJB equation reduces to

$$V_t + \frac{1}{2}m_0^2V_{xx} + \frac{1}{4\eta}(\gamma(x - A) + V_x)^2 + \mu(x - R_t) + \rho\sigma m_0 = 0, \quad v(T, x) = -\beta(x - A)^2.$$

Assume the Ansatz for value function V

$$V(t, x) = a(t)x^2 + b(t)x + c(t).$$

Plug the Ansatz into the HJB equation yields the Riccati ODE system

$$\begin{cases} a'(t) + \frac{1}{4\eta}(2a(t) + \gamma)^2 = 0, & a(T) = -\beta, \\ b'(t) + \frac{1}{2\eta}(2a(t) + \gamma)(b(t) - \gamma A) + \mu = 0, & b(T) = 2\beta A, \\ c'(t) + m_0^2a(t) + \frac{1}{4\eta}(b(t) - \gamma A)^2 - \mu R_t + \rho\sigma m_0 = 0, & c(T) = -\beta A^2. \end{cases}$$

By solving the ODE system, the deterministic functions a, b, c of time t are given by

$$\begin{aligned} a(t) &= -\frac{\eta}{T-t+\alpha} - \frac{\gamma}{2}, \\ b(t) &= \frac{\mu}{2}(T-t+\alpha) - \frac{\frac{1}{2}\mu\alpha^2 - 2A\eta}{T-t+\alpha} + \gamma A, \\ c(t) &= -m_0^2\eta \log \frac{T-t+\alpha}{\alpha} + \left(\rho\sigma m_0 - \frac{\gamma}{2}m_0^2\right)(T-t) \\ &\quad - \frac{\mu^2\alpha^3}{16\eta} \left(\frac{\alpha}{T-t+\alpha} + \frac{2(T-t+\alpha)}{\alpha} - \frac{(T-t+\alpha)^3}{3\alpha^3} - \frac{8}{3}\right) \\ &\quad - \frac{A^2\eta - \frac{1}{2}\mu A\alpha^2}{T-t+\alpha} - \frac{\gamma A^2 + \mu A(T-t+\alpha)}{2} + \int_t^T \mu(A - R_s)ds \end{aligned}$$

where $\alpha := \frac{2\eta}{2\beta-\gamma} > 0$. Note that $V(t, x) = a(t)x^2 + b(t)x + c(t)$ is continuous on $[0, T] \times \mathbb{R}$ and continuously differentiable on $(0, T) \times \mathbb{R}$ satisfying the HJB equation, therefore it must be the value function by the uniqueness of solutions to the HJB equation. Then the verification theorem gives rise to the optimal feedback control

$$v_t^{**} = \frac{x_t - A}{T-t+\alpha} - \frac{\mu}{4\eta} \left((T-t+\alpha) - \frac{\alpha^2}{T-t+\alpha} \right).$$

Due to the boundedness of $\frac{1}{T-t+\alpha}$ and $-\frac{A}{T-t+\alpha} - \frac{\mu}{4\eta((T-t+\alpha)-\frac{\alpha^2}{T-t+\alpha})}$, $v^{**} \in \mathcal{A}$, which implies that the optimal control v^{**} lies in the admissible set. Moreover, the maximal risk-neutral expected utility is given by

$$\begin{aligned} \sup_{v \in \mathcal{A}} \mathbb{E} \left[\tilde{\Pi} \right] &= \mathbb{E} \left[\tilde{\Pi} \right] \Big|_{v=v^{**}} = V(0, x_0) = a(0)x_0^2 + b(0)x_0 + c(0) \\ &= \left(-\frac{\eta}{T+\alpha} - \frac{\gamma}{2} \right) x_0^2 + \left(\frac{\mu}{2}(T+\alpha) - \frac{\frac{1}{2}\mu\alpha^2 - 2A\eta}{T+\alpha} + \gamma A \right) x_0 \\ &\quad - m_0^2\eta \log \frac{T+\alpha}{\alpha} + \left(\rho\sigma m_0 - \frac{\gamma}{2}m_0^2 \right) T - \frac{\mu^2\alpha^3}{16\eta} \left(\frac{\alpha}{T+\alpha} + \frac{2(T+\alpha)}{\alpha} - \frac{(T+\alpha)^3}{3\alpha^3} - \frac{8}{3} \right) \\ &\quad - \frac{A^2\eta - \frac{1}{2}\mu A\alpha^2}{T+\alpha} - \frac{\gamma A^2 + \mu A(T+\alpha)}{2} + \int_0^T \mu(A - R_s)ds < \infty. \end{aligned}$$

A.3 Proof to Theorem 1

Note that $x_t = x_0 - \int_0^t v_s ds + m_0 dZ_t$, by Itô's formula,

$$(x_T - A)^2 = (x_t - A)^2 + \int_t^T (-2(x_s - A)v_s + m_0^2) ds + \int_t^T (2m_0(x_s - A)) dZ_s.$$

Furthermore, note that standard Brownian motions W_t and Z_t have constant correlation ρ , denote standard Brownian motion $W_t^1 = \frac{Z_t - \rho W_t}{\sqrt{1-\rho^2}}$, then W_t^1 and W_t are independent. By extracting the constants and deterministic parts out of the expectation, the value function is given by

$$V(t, x) = \frac{1}{\theta} - \frac{1}{\theta} \cdot \exp \left(- \int_t^T \theta (\mu(A - R_s) + \rho \sigma m_0 - \beta m_0^2) ds + \theta \beta (x - A)^2 \right) \cdot V_0(t, x),$$

where we observe the equivalent optimization problem

$$\begin{aligned} V_0(t, x) &= \inf_{v \in \mathcal{A}_t} \mathbb{E}^{\mathbb{P}} \left[\exp \left\{ \int_t^T [-\theta \mu(x_s - A) + \theta \eta v_s^2 - \theta(2\beta - \gamma)v_s(x_s - A)] ds \right. \right. \\ &\quad + \int_t^T [(-\theta \sigma - \theta(\gamma - 2\beta)m_0\rho)(x_s - A) - \theta \eta m_0 \rho v_s + \theta \sigma(R_s - A)] dW_s \\ &\quad \left. \left. + \int_t^T [-\theta(\gamma - 2\beta)m_0 \sqrt{1 - \rho^2}(x_s - A) - \theta \eta m_0 \sqrt{1 - \rho^2} v_s] dW_s^1 \right\} \middle| x_t = x \right]. \\ &=: \inf_{v \in \mathcal{A}_t} \mathbb{E}^{\mathbb{P}} [\exp \{J(t, T)\} | x_t = x] \end{aligned}$$

Our stochastic control problem has a similar form as stochastic linear exponential quadratic (SLEQ) control problems given in [23] and [24], however, due to the constants and stochastic integrals appearing in our exponential function, these existing results can't be used directly. Instead, we follow the "completing the square" technique used in [23], which is used to transform the integrand into a square form by introducing a Radon-Nikodym derivative. In this paper, we eliminate the stochastic integrals and "complete the square" simultaneously by Girsanov's theorem. More precisely, we manage to find some probability measure $\mathbb{Q} \ll \mathbb{P}$ such that

$$\mathbb{E}^{\mathbb{P}} [\exp \{J(t, T)\} | x_t = x] = \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{L(t, T)} \cdot \exp \{J(t, T)\} \middle| x_t = x \right],$$

where $\frac{1}{L(t, T)} \cdot \exp \{J(t, T)\}$ can be transformed into a square form without Itô integrals. Concretely, suppose the Radon-Nikodym derivative is given by

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} &= L(t, T) := \exp \left\{ - \int_t^T \Psi_1(s, x_s, v_s) dW_s - \int_t^T \Psi_2(s, x_s, v_s) dW_s^1 \right. \\ &\quad \left. - \frac{1}{2} \int_t^T (\Psi_1^2(s, x_s, v_s) + \Psi_2^2(s, x_s, v_s)) ds \right\}, \end{aligned}$$

where Ψ_1 and Ψ_2 are L^2 -functions to be determined. At the same time, for some deterministic function $G_1(t)$ and $G_2(t)$ such that $G_1(T) = G_2(T) = 0$, by Ito's lemma,

$$\begin{aligned} & -G_1(t)(x_t - A)^2 - 2G_2(t)(x_t - A) \\ &= (G_1(T)(x_T - A)^2 + 2G_2(T)(x_T - A)) - (G_1(t)(x_t - A)^2 + 2G_2(t)(x_t - A)) \\ &= \int_t^T (G'_1(s)(x_s - A)^2 + 2G'_2(s)(x_s - A) - 2(G_1(s)(x_s - A) + G_2(s))v_s + m_0^2 G_1(s)) ds \\ &+ \int_t^T 2\rho m_0(G_1(s)(x_s - A) + G_2(s))dW_s + \int_t^T 2\sqrt{1 - \rho^2}m_0(G_1(s)(x_s - A) + G_2(s))dW_s^1. \end{aligned}$$

Hence,

$$\begin{aligned} & \frac{1}{L(t, T)} \cdot \exp \{J(t, T)\} \\ &= \exp \left\{ J(t, T) - \log L(t, T) + (-G_1(t)(x_t - A)^2 - 2G_2(t)(x_t - A)) + (G_1(t)(x_t - A)^2 + 2G_2(t)(x_t - A)) \right\} \\ &= \exp \left\{ \int_t^T \left[-\theta\mu(x_s - A) + \theta\eta v_s^2 - \theta(2\beta - \gamma)v_s(x_s - A) + \frac{1}{2}\Psi_1^2(s, x_s, v_s) + \frac{1}{2}\Psi_2^2(s, x_s, v_s) \right. \right. \\ &\quad \left. \left. + (G'_1(s)(x_s - A)^2 + 2G'_2(s)(x_s - A) - 2(G_1(s)(x_s - A) + G_2(s))v_s + m_0^2 G_1(s)) \right] ds \right. \\ &\quad \left. + \int_t^T \left[(-\theta\sigma - \theta(\gamma - 2\beta)m_0\rho)(x_s - A) - \theta\eta m_0\rho v_s + \theta\sigma(R_s - A) + \Psi_1(s, x_s, v_s) \right. \right. \\ &\quad \left. \left. + 2\rho m_0(G_1(s)(x_s - A) + G_2(s)) \right] dW_s \right. \\ &\quad \left. + \int_t^T \left[-\theta(\gamma - 2\beta)m_0\sqrt{1 - \rho^2}(x_s - A) - \theta\eta m_0\sqrt{1 - \rho^2}v_s + \Psi_2(s, x_s, v_s) \right. \right. \\ &\quad \left. \left. + 2\sqrt{1 - \rho^2}m_0(G_1(s)(x_s - A) + G_2(s)) \right] dW_s^1 \right. \\ &\quad \left. + (G_1(t)(x_t - A)^2 + 2G_2(t)(x_t - A)) \right\} \end{aligned}$$

By setting

$$\begin{cases} \Psi_1(s, x_s, v_s) = (\theta\sigma + m_0\rho(\theta(\gamma - 2\beta) - 2G_1(s))) - (x_s - A) + \theta\eta m_0\rho v_s - 2m_0\rho G_2(s) - \theta\sigma(R_s - A), \\ \Psi_2(s, x_s, v_s) = (\theta(\gamma - 2\beta) - 2G_1(s))m_0\sqrt{1 - \rho^2} - (x_s - A) + \theta\eta m_0\sqrt{1 - \rho^2}v_s - 2m_0\sqrt{1 - \rho^2}G_2(s), \end{cases}$$

the stochastic integrals $\int_t^T \cdots dW_s$ and $\int_t^T \cdots dW_s^1$ appear to be zero. Moreover,

$$\begin{aligned} & \frac{1}{L(t, T)} \cdot \exp \{J(t, T)\} \\ &= \exp \left\{ \int_t^T (A(s)v_s^2 + B(s)(x_s - A)v_s + C(s)(x_s - A)^2 + D(s)v_s + E(s)(x_s - A) + F(s)) dt, \right. \\ &\quad \left. + (G_1(t)(x_t - A)^2 + 2G_2(t)(x_t - A)) \right\} \end{aligned} \tag{14}$$

where the coefficients are given by

$$\left\{ \begin{array}{l} H \equiv A(s) := \theta\eta + \frac{1}{2}\theta^2\eta^2m_0^2 > 0, \\ B(s) = -\theta(2\beta - \gamma) - 2G_1(s) + m_0\eta\theta^2\rho\sigma + \theta\eta m_0^2(\theta(\gamma - 2\beta) - 2G_1(s)), \\ C(s) = G'_1(s) + \frac{1}{2}m_0^2(\theta(\gamma - 2\beta) - 2G_1(s))^2 + \frac{1}{2}\theta^2\sigma^2 + \theta\sigma\rho m_0(\theta(\gamma - 2\beta) - 2G_1(s)), \\ D(s) = -2G_2(s) - 2m_0^2\theta\eta G_2(s) - m_0\eta\theta^2\rho\sigma(R_s - A), \\ E(s) = -\theta\mu + 2G'_2(s) - 2m_0^2(\theta(\gamma - 2\beta) - 2G_1(s))G_2(s) - \theta^2\sigma^2(R_s - A) - 2m_0\rho\theta\sigma G_2(s) \\ \quad - m_0\rho\theta\sigma(\theta(\gamma - 2\beta) - 2G_1(s))(R_s - A) \\ F(s) = m_0^2G_1(s) + 2m_0^2G_2^2(s) + 2\rho m_0\theta\sigma(R_s - A)G_2(s) + \frac{1}{2}\theta^2\sigma^2(R_s - A)^2. \end{array} \right.$$

To make sure the integrand in (14) behaves as a square form (plus constants), we rewrite it as a quadratic function of v_s :

$$A(s)v_s^2 + (B(s)(x_s - A) + D(s))v_s + (C(s)(x_s - A)^2 + E(s)(x_s - A) + F_0(s)) + (F(s) - F_0(s)),$$

of which the discriminant is set to be zero:

$$0 = \Delta = (B(s)(x_s - A) + D(s))^2 - 4A(s)(C(s)(x_s - A)^2 + E(s)(x_s - A) + F_0(s)),$$

hence,

$$\left\{ \begin{array}{l} 0 = B^2(s) - 4A(s)C(s), \\ 0 = B(s)D(s) - 2A(s)E(s), \\ 0 = D^2(s) - 4A(s)F_0(s) \implies F_0(s) = \frac{D^2(s)}{4A(s)}. \end{array} \right.$$

Denote $b_2(s) := G_1(s) + \frac{\theta}{2}(2\beta - \gamma)$ and $b_1(s) := 2G_2(s)$, we may solve the ODEs $b_2(\cdot)$ and $b_1(\cdot)$ satisfy:

$$b'_2 = \frac{1}{H}((b_2 + l_3)^2 - l_1), \quad b_2(T) = \frac{\theta}{2}(2\beta - \gamma),$$

where $l_3 = \frac{1}{2}m_0\eta\theta^2\rho\sigma$ and $l_1 = \frac{\theta^2\sigma^2}{2}H$. We may solve that

$$b_2(t) = \sqrt{l_1} \coth \left(A_0 + \frac{\sqrt{l_1}}{H}(T-t) \right) - l_3,$$

where

$$A_0 := \coth^{-1} \left(\frac{l_3 + \frac{\theta}{2}(2\beta - \gamma)}{\sqrt{l_1}} \right).$$

Moreover, we have b_1 solves the ODE

$$b'_1 = \theta\mu + \frac{1}{H}b_1(b_2 + l_3) + \frac{\eta\theta^2\sigma}{H}(R_t - A) \left(\theta\sigma + \frac{1}{2}m_0^2\theta^2\eta\sigma(1 - \rho^2) - \rho m_0 b_2 \right), \quad b_1(T) = 0.$$

With b_1 and b_2 solved, we may rewrite the integrand in (14) as

$$\begin{aligned} & A(s)v_s^2 + B(s)(x_s - A)v_s + C(s)(x_s - A)^2 + D(s)v_s + E(s)(x_s - A) + F(s) \\ &= H(v_s - g_1(s)(x_s - A) - g_0(s))^2 + \left(F(s) - \frac{D^2(s)}{4H} \right) \\ &= H(v_s - g_1(s)(x_s - A) - g_0(s))^2 + \left(m_0^2 \left(b_2 - \frac{\theta}{2}(2\beta - \gamma) \right) + \frac{\theta^2\sigma^2}{2}(R_s - A)^2 - \frac{1}{4H} (b_1 - \rho m_0 \theta^2 \sigma \eta (R_s - A))^2 \right), \end{aligned}$$

where

$$g_1(t) := \frac{(1 + m_0^2 \theta \eta) b_2 - l_3}{H}, \quad g_0(t) := \frac{(1 + \theta \eta m_0^2) b_1 + 2l_3(R_u - A)}{2H}.$$

Denote

$$c(s) := m_0^2 \left(b_2 - \frac{\theta}{2}(2\beta - \gamma) \right) + \frac{\theta^2\sigma^2}{2}(R_s - A)^2 - \frac{1}{4H} (b_1 - \rho m_0 \theta^2 \sigma \eta (R_s - A))^2, \quad 0 \leq s \leq T,$$

we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} [\exp \{ J(t, T) \} | x_t = x] = \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{L(t, T)} \cdot \exp \{ J(t, T) \} \middle| x_t = x \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\exp \left\{ \int_t^T H(v_s - g_1(s)(x_s - A) - g_0(s))^2 ds \right\} \right. \\ & \quad \cdot \exp \left\{ \left(b_2(t) - \frac{\theta}{2}(2\beta - \gamma) \right) (x_t - A)^2 + b_1(t)(x_t - A) + \int_t^T c(s) ds \right\} \left. \middle| x_t = x \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\exp \left\{ \int_t^T H(v_s - g_1(s)(x_s - A) - g_0(s))^2 ds \right\} \middle| x_t = x \right] \\ & \quad \cdot \exp \left\{ \left(b_2(t) - \frac{\theta}{2}(2\beta - \gamma) \right) (x - A)^2 + b_1(t)(x - A) + \int_t^T c(s) ds \right\} \\ &\geq \exp \left\{ \left(b_2(t) - \frac{\theta}{2}(2\beta - \gamma) \right) (x - A)^2 + b_1(t)(x - A) + \int_t^T c(s) ds \right\}, \end{aligned}$$

where the equation holds when $v_s = v_s^* := g_1(s)(x_s - A) + g_0(s)$. Note that g_1 and g_0 are bounded and continuous on $[0, T]$, so $v^* \in \mathcal{A}_t$, hence,

$$\begin{aligned} V_0(t, x) &= \inf_{v \in \mathcal{A}_t} \mathbb{E}^{\mathbb{P}} [\exp \{ J(t, T) \} | x_t = x] = \inf_{v \in \mathcal{A}_t} \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{L(t, T)} \cdot \exp \{ J(t, T) \} \middle| x_t = x \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{L(t, T)} \cdot \exp \{ J(t, T) \} \middle| x_t = x \right] \Big|_{v=v^*} \\ &= \exp \left\{ \left(b_2(t) - \frac{\theta}{2}(2\beta - \gamma) \right) (x - A)^2 + b_1(t)(x - A) + \int_t^T c(s) ds \right\}. \end{aligned}$$

Furthermore, the value function is given by

$$\begin{aligned}
V(t, x) &= \frac{1}{\theta} - \frac{1}{\theta} \cdot \exp \left\{ - \int_t^T \theta (\mu(A - R_s) + \rho\sigma m_0 - \beta m_0^2) ds + \theta\beta(x - A)^2 \right\} \cdot V_0(t, x) \\
&= \frac{1}{\theta} - \frac{1}{\theta} \cdot \exp \left\{ \left(b_2(t) + \frac{\theta\gamma}{2} \right) (x - A)^2 + b_1(t)(x - A) \right. \\
&\quad \left. + \int_t^T [c(s) - \theta\mu(A - R_s) - \theta\rho\sigma m_0 + \theta\beta m_0^2] ds \right\} \\
&= \frac{1}{\theta} \left(1 - \exp \left\{ \left(b_2(t) + \frac{\theta\gamma}{2} \right) (x - A)^2 + b_1(t)(x - A) + b_0(t) \right\} \right),
\end{aligned}$$

where

$$\begin{aligned}
b_0(t) &:= \int_t^T [c(s) - \theta\mu(A - R_s) - \theta\rho\sigma m_0 + \theta\beta m_0^2] ds \\
&= \int_t^T \left[\frac{l_1 - l_3^2}{H} (R_s - A)^2 + \left(\theta\mu + \frac{l_3}{H} b_1 \right) (R_s - A) + \left(-\frac{b_1^2}{4H} + m_0^2 \left(b_2 + \frac{\theta\gamma}{2} \right) - \rho m_0 \theta\sigma \right) \right] ds.
\end{aligned}$$

In summary, the optimal feedback control is given by

$$\begin{aligned}
v_t^* &= g_1(t)(x_t - A) + g_0(t) \\
&= \frac{(1 + m_0^2 \theta\eta) b_2 - l_3}{H} \cdot (x_t - A) + \frac{l_3}{H} \cdot (R_t - A) + \frac{(1 + m_0^2 \theta\eta) b_1}{2H}.
\end{aligned}$$

□

A.4 Proof to Theorem 2

When $m(v) = m_0 = 0$ and $\beta = 0$, $dx_t = -v_t dt$, according to Theorem 1, $H = \theta\eta$, $l_3 = 0$, $l_1 = \frac{\theta^3 \sigma^2 \eta}{2}$, $A_0 = 0$, and

$$b_2(t) = \sqrt{\frac{\theta^3 \sigma^2 \eta}{2}} \cdot \coth \left(\sqrt{\frac{\theta\sigma^2}{2\eta}} (T - t) \right).$$

Furthermore, b_1 solves the backward ODE

$$b'_1 = \theta\mu + \frac{1}{H} b_1 b_2 + \theta^2 \sigma^2 (R_t - A), \quad b_1(T) = 0.$$

Note that

$$\begin{aligned}
\left(b_1 \cdot \sinh \left(\sqrt{\frac{\theta\sigma^2}{2\eta}} (T - t) \right) \right)' &= \sinh \left(\sqrt{\frac{\theta\sigma^2}{2\eta}} (T - t) \right) \cdot b'_1 - b_1 \cdot \sqrt{\frac{\theta\sigma^2}{2\eta}} \cdot \cosh \left(\sqrt{\frac{\theta\sigma^2}{2\eta}} (T - t) \right) \\
&= \sinh \left(\sqrt{\frac{\theta\sigma^2}{2\eta}} (T - t) \right) \cdot (\theta\mu + \theta^2 \sigma^2 (R_t - A)).
\end{aligned}$$

Integrate on both sides,

$$\begin{aligned} -b_1(t) \cdot \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right) &= \int_t^T \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-s)\right) \cdot (\theta\mu + \theta^2\sigma^2(R_s - A)) \, ds. \\ \Rightarrow b_1(t) &= -\frac{1}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)} \cdot \int_t^T \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-s)\right) \cdot (\theta\mu + \theta^2\sigma^2(R_s - A)) \, ds. \end{aligned}$$

Note that now

$$v_t = -x'_t = -(x_t - A)' = \frac{b_2}{H}(x_t - A) + \frac{b_1}{2H},$$

hence,

$$\begin{aligned} \left(\frac{x_t - A}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)} \right)' &= \frac{(x_t - A)' \cdot \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right) + (x_t - A) \cdot \sqrt{\frac{\theta\sigma^2}{2\eta}} \cdot \cosh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)}{\sinh^2\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)} \\ &= \frac{\int_t^T \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-s)\right) \cdot (\theta\mu + \theta^2\sigma^2(R_s - A)) \, ds}{2H \sinh^2\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)}. \end{aligned}$$

Integrate on both sides then integrate by parts, we have

$$\begin{aligned} \frac{x_t - A}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)} - \frac{x_0 - A}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}T\right)} &= \int_0^t \int_s^T \frac{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-u)\right)}{2H \cdot \sinh^2\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-s)\right)} \cdot (\theta\mu + \theta^2\sigma^2(R_u - A)) \, du \, ds \\ &= \frac{1}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}T\right)} \cdot \int_0^t \frac{1}{\sqrt{2\theta^3\sigma^2\eta}} \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}s\right) \cdot (\theta\mu + \theta^2\sigma^2(R_s - A)) \, ds + \frac{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}t\right)}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}T\right)} \cdot \frac{1}{\sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-t)\right)} \\ &\quad \cdot \int_t^T \frac{1}{\sqrt{2\theta^3\sigma^2\eta}} \cdot \sinh\left(\sqrt{\frac{\theta\sigma^2}{2\eta}}(T-s)\right) \cdot (\theta\mu + \theta^2\sigma^2(R_s - A)) \, ds. \end{aligned}$$

Denote

$$\frac{\theta^2\sigma^2}{\sqrt{2\theta^3\sigma^2\eta}} = \sqrt{\frac{\theta\sigma^2}{2\eta}} =: \kappa,$$

we can solve that

$$\begin{aligned} x_t &= \frac{\kappa}{\sinh \kappa T} \int_0^T \left(R_s \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ &\quad + \left(x_0 - \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + A + \left(-A + \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T}. \end{aligned}$$

Furthermore,

$$\begin{aligned} x_t &= \frac{\mu}{\theta\sigma^2} + \left(x_0 - \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + \left(A - \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa t}{\sinh \kappa T} \\ &\quad + \frac{\sinh \kappa(T-t)}{\sinh \kappa T} \cdot \int_0^t \kappa \sinh \kappa s R_s \, ds + \frac{\sinh \kappa t}{\sinh \kappa T} \cdot \int_t^T \kappa \sinh \kappa(T-s) R_s \, ds \\ &= \frac{\mu}{\theta\sigma^2} + \frac{\sinh \kappa(T-t)}{\sinh \kappa T} \cdot \left(x_0 - \frac{\mu}{\theta\sigma^2} + \int_0^t \kappa \sinh \kappa s R_s \, ds \right) + \frac{\sinh \kappa t}{\sinh \kappa T} \cdot \left(A - \frac{\mu}{\theta\sigma^2} + \int_t^T \kappa \sinh \kappa(T-s) R_s \, ds \right) \end{aligned}$$

□

A.5 Proof to Lemma 1

$$\begin{aligned}
& \frac{\kappa}{\sinh \kappa T} \int_0^T \left(\sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\
&= \frac{\kappa}{\sinh \kappa T} \cdot \int_0^t \sinh \kappa s ds \cdot \sinh(\kappa(T - t)) + \frac{\kappa}{\sinh \kappa T} \cdot \int_t^T \sinh \kappa(T - s) ds \cdot \sinh \kappa t \\
&= \frac{\sinh(\kappa(T - t))}{\sinh \kappa T} \cdot (\cosh \kappa t - 1) + \frac{\sinh \kappa t}{\sinh \kappa T} \cdot (\cosh \kappa(T - t) - 1) \\
&= \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T} - \frac{\sinh \kappa(T - t)}{\sinh \kappa T} = \text{TC}_t - \text{IS}_t.
\end{aligned}$$

□

A.6 Proof to Proposition 2

By Theorem 2 and Lemma 1,

$$\begin{aligned}
x_t &= \frac{\kappa}{\sinh \kappa T} \int_0^T \left(R \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\
&\quad + \left(x_0 - \frac{\mu}{\theta \sigma^2} \right) \cdot \frac{\sinh \kappa(T - t)}{\sinh \kappa T} + A + \left(-A + \frac{\mu}{\theta \sigma^2} \right) \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T} \\
&= \left(x_0 - \frac{\mu}{\theta \sigma^2} - R \right) \cdot \text{IS}_t + \left(-A + \frac{\mu}{\theta \sigma^2} + R \right) \cdot \text{TC}_t + A.
\end{aligned}$$

□

A.7 Proof to Proposition 3

Note that $R + \frac{\mu}{\theta \sigma^2}$ appears as a whole, so WLOG, we may assume that $\mu \equiv 0$. When the reference strategy is R , the optimal strategy is given by

$$x_t = A + (R - A) \cdot \text{TC}_t + (x_0 - R) \cdot \text{IS}_t,$$

then

$$\begin{aligned}
v_t &= -x'_t = -(R - A) \cdot \text{TC}'_t - (x_0 - R) \cdot \text{IS}'_t \\
&= -(R - A) \cdot \frac{-\kappa \cosh \kappa t}{\sinh \kappa T} - (x_0 - R) \cdot \frac{-\kappa \cosh \kappa(T - t)}{\sinh \kappa T},
\end{aligned}$$

note that $0 \leq t \leq T$, then $\cosh \kappa t \geq 0$ and $\cosh \kappa(T - t) \geq 0$, so when $A \leq R \leq x_0$, $v_t \geq 0$, which represents that the optimal strategy does not contain overshooting. Note that

$$x''_t = -v'_t = (R - A) \cdot \frac{-\kappa^2 \sinh \kappa t}{\sinh \kappa T} + (x_0 - R) \cdot \frac{\kappa^2 \sinh \kappa(T - t)}{\sinh \kappa T},$$

so when $R > x_0 > A$, $x''_t = -v'_t \leq 0$, which implies that the optimal strategy is a concave function, so

$$\begin{aligned}
\min_{0 \leq t \leq T} v_t &= v_0 = -(R - A) \cdot \frac{-\kappa}{\sinh \kappa T} - (x_0 - R) \cdot \frac{-\kappa \cosh \kappa T}{\sinh \kappa T} \\
&= \frac{\kappa}{\sinh \kappa T} \cdot [(1 - \cosh \kappa T)R + (x_0 \cosh \kappa T - A)],
\end{aligned}$$

so when

$$R > \frac{x_0 \cosh \kappa T - A}{\cosh \kappa T - 1} = x_0 + \frac{1}{\cosh \kappa T - 1} \cdot (x_0 - A),$$

the optimal strategy contains overshooting, and when $x_0 < R \leq x_0 + \frac{1}{\cosh \kappa T - 1} \cdot (x_0 - A)$, the optimal strategy does not contain overshooting. Similarly, when $R < A < x_0$, $q_t'' = -v_t' \geq 0$, which implies that the optimal strategy is a convex function, so

$$\begin{aligned} \min_{0 \leq t \leq T} v_t &= v_T = -(R - A) \cdot \frac{-\kappa \cosh \kappa T}{\sinh \kappa T} - (x_0 - R) \cdot \frac{-\kappa}{\sinh \kappa T} \\ &= \frac{\kappa}{\sinh \kappa T} \cdot [(\cosh \kappa T - 1)R - (x_0 - A \cosh \kappa T)], \end{aligned}$$

so when

$$R < \frac{A \cosh \kappa T - x_0}{\cosh \kappa T - 1} = A - \frac{1}{\cosh \kappa T - 1} \cdot (x_0 - A),$$

the optimal strategy contains overshooting.

□

A.8 Proof to Proposition 4

By the results given in Theorem 2, the optimal strategy is a deterministic one. In other words, when $\mu = m_0 = 0$,

$$\sup_{v \in \mathcal{A}} \mathbb{E} [u(\tilde{\Pi})] = \sup_{v \in \mathcal{A}_{\text{det}}} \mathbb{E} [u(\tilde{\Pi})],$$

where $\mathcal{A}_{\text{det}} = \{v \in \mathcal{A} : v \text{ is deterministic}\}$. When $v \in \mathcal{A}_{\text{det}}$, $x_t = x_0 - \int_0^t v_s ds$ is also deterministic, hence,

$$\begin{aligned} \mathbb{E} [u(\tilde{\Pi})] &= \mathbb{E} \left[\frac{1}{\theta} \left(1 - \exp \left\{ \theta \beta (x_T - A)^2 - \int_0^T \theta (-\eta v_t^2 - \gamma v_t(x_t - A)) dt - \int_0^T \theta \sigma (x_t - R_t) dW_t \right\} \right) \right] \\ &= \frac{1}{\theta} - \frac{1}{\theta} \cdot \exp \left\{ \theta \beta (x_T - A)^2 - \int_0^T \left[\theta (-\eta v_t^2 - \gamma v_t(x_t - A)) - \frac{1}{2} \theta^2 \sigma^2 (x_t - R_t)^2 \right] dt \right\}. \end{aligned}$$

Subject to $x_T = A$ (or $\beta \rightarrow +\infty$), $\int_0^T v_t(x_t - A) dt = \frac{1}{2}(x_0 - A)^2$ is a constant, hence, maximizing the expected utility is equivalent to a typical variational problem

$$\begin{cases} \inf_{v \in \mathcal{A}_{\text{det}}} \int_0^T \left[\eta v_t^2 + \frac{\theta \sigma^2}{2} (x_t - R_t)^2 \right] dt \\ \text{s.t. } x_t = x_0 - \int_0^t v_s ds, \quad x_T = A, \end{cases}$$

Note that the objective can be divided into n parts, each with a constant RS:

$$\int_0^T \left(\eta v_t^2 + \frac{\gamma}{2} \sigma^2 (x_t - R_t)^2 \right) dt = \sum_{k=1}^n \int_{\frac{(k-1)T}{n}}^{\frac{kT}{n}} \left(\eta v_t^2 + \frac{\gamma}{2} \sigma^2 (x_t - R^{(k)})^2 \right) dt.$$

We divide the optimization problem into three steps:

- Find the optimal value of $x_{\frac{kT}{n}}^*$, $1 \leq k \leq n - 1$. We can directly use Theorem 2: suppose the optimal

strategy is x_t^* , note that $\mu = 0$,

$$\begin{aligned} x_t^* &= \frac{\kappa}{\sinh \kappa T} \int_0^T \left(R_s \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ &\quad + x_0 \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + A - A \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T}, \end{aligned}$$

where R is the given piece-wise constant RS. For $1 \leq k \leq n-1$,

$$\begin{aligned} x_{\frac{kT}{n}}^* &= x_0 \cdot \frac{\sinh \kappa(T - \frac{kT}{n})}{\sinh \kappa T} + A \cdot \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} + \frac{\kappa \sinh \kappa (T - \frac{kT}{n})}{\sinh \kappa T} \cdot \sum_{i=1}^k R^{(i)} \cdot \int_{\frac{(i-1)T}{n}}^{\frac{iT}{n}} \sinh(\kappa s) ds \\ &\quad + \frac{\kappa \sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \sum_{i=k+1}^n R^{(i)} \cdot \int_{\frac{(i-1)T}{n}}^{\frac{iT}{n}} \sinh(\kappa(T-s)) ds \\ &= x_0 \cdot \frac{\sinh \kappa(T - \frac{kT}{n})}{\sinh \kappa T} + A \cdot \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \\ &\quad + \frac{\sinh \kappa (T - \frac{kT}{n})}{\sinh \kappa T} \cdot \sum_{i=1}^k R^{(i)} \cdot \left(\cosh\left(\kappa \frac{iT}{n}\right) - \cosh\left(\kappa \frac{(i-1)T}{n}\right) \right) \\ &\quad + \frac{\kappa \sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \sum_{i=k+1}^n R^{(i)} \cdot \left(\cosh\left(\kappa \frac{(n-i+1)T}{n}\right) - \cosh\left(\kappa \frac{(n-i)T}{n}\right) \right). \end{aligned}$$

Denote $R^{(0)} := x_0$, $R^{(n+1)} := A$, and

$$b_i = \begin{cases} 1, & i = 0, \\ \cosh\left(\kappa \frac{iT}{n}\right) - \cosh\left(\kappa \frac{(i-1)T}{n}\right), & 1 \leq i \leq n-1, \end{cases}$$

we can rewrite

$$x_{\frac{kT}{n}}^* = \frac{\sinh \kappa (T - \frac{kT}{n})}{\sinh \kappa T} \cdot \sum_{i=0}^k b_i R^{(i)} + \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \sum_{i=k+1}^{n+1} b_{n-i+1} R^{(i)}, \quad 1 \leq k \leq n-1.$$

Furthermore, we denote the optimal values $a_k := x_{\frac{kT}{n}}^*, 1 \leq k \leq n-1$. Note that the sum of coefficients in a_k is

$$\begin{aligned} &\frac{\sinh \kappa (T - \frac{kT}{n})}{\sinh \kappa T} \cdot \sum_{i=0}^k b_i + \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \sum_{i=k+1}^{n+1} b_{n-i+1} \\ &= \frac{\sinh \kappa (T - \frac{kT}{n})}{\sinh \kappa T} \cdot \cosh \kappa \frac{kT}{n} + \frac{\sinh \kappa \frac{kT}{n}}{\sinh \kappa T} \cdot \cosh \kappa \frac{(n-k)T}{n} = 1, \end{aligned}$$

so a_k is a weighted average of $R^{(i)}$'s.

- Note that we already know the value of optimal strategy $x_{\frac{kT}{n}}$, so the optimization problem is equivalent to the following optimization problem with extra constraints:

$$\left\{ \begin{array}{l} \inf_{v \in \mathcal{A}_{\text{det}}} \int_0^T \left[\eta v_t^2 + \frac{\theta \sigma^2}{2} (x_t - R_t)^2 \right] dt = \inf_{v \in \mathcal{A}_{\text{det}}} \sum_{k=1}^n \int_{\frac{(k-1)T}{n}}^{\frac{kT}{n}} \left(\eta v_t^2 + \frac{\gamma}{2} \sigma^2 (x_t - R^{(k)})^2 \right) \\ \text{s.t. } x_t = x_0 - \int_0^t v_s ds, \quad x_T = A, \\ x_{\frac{kT}{n}} = a_k, \quad 1 \leq k \leq n-1. \end{array} \right.$$

Denote $a_0 = x_0$ and $a_n = A$, inheriting from the spirit of dynamic programming, we can further divide the optimization problem into n parts: for $1 \leq k \leq n$, where the k -th sub-problem is given by

$$\begin{cases} \inf_{v \in \mathcal{A}_{\text{det}}} \int_{\frac{(k-1)T}{n}}^{\frac{kT}{n}} \left(\eta v_t^2 + \frac{\theta\sigma^2}{2} (x_t - R^{(k)})^2 \right) dt \\ \text{s.t. } x_t = a_{k-1} - \int_{\frac{(k-1)T}{n}}^t v_s ds, \frac{(k-1)T}{n} \leq t \leq \frac{kT}{n}, x_{\frac{kT}{n}} = a_k, \end{cases}$$

which is a constant RS problem with time period $\left[\frac{(k-1)T}{n}, \frac{kT}{n} \right]$, of which the optimal strategy is given by

$$x_t^{(k)} = a_k + (R^{(k)} - a_k) \text{TC}_{t-\frac{(k-1)T}{n}} + (a_{k-1} - R^{(k)}) \text{IS}_{t-\frac{(k-1)T}{n}}, \frac{(k-1)T}{n} \leq t \leq \frac{kT}{n},$$

where $\kappa := \sqrt{\frac{\theta\sigma^2}{2\eta}}$ and

$$\text{IS}_t := \frac{\sinh \kappa \left(\frac{T}{n} - t \right)}{\sinh \kappa \frac{T}{n}}, \quad \text{TC}_t := \frac{\sinh \kappa \frac{T}{n} - \sinh \kappa t}{\sinh \kappa \frac{T}{n}}.$$

- Connect n optimal strategies to obtain the optimal strategy of the original optimization problem. Denote

$$x_t^{**} = \sum_{k=1}^n x_t^{(k)} \mathbb{1}_{\left\{ \frac{(k-1)T}{n} \leq t < \frac{kT}{n} \right\}} + A \mathbb{1}_{\{t=T\}},$$

then by the results of constant reference strategies, for any strategy x_t ,

$$\begin{aligned} \int_0^T \left[\eta v_t^2 + \frac{\theta\sigma^2}{2} (x_t - R_t)^2 \right] dt &= \sum_{k=1}^n \int_{\frac{(k-1)T}{n}}^{\frac{kT}{n}} \left(\eta v_t^2 + \frac{\gamma}{2} \sigma^2 (x_t - R^{(k)})^2 \right) dt \\ &\geq \sum_{k=1}^n \int_{\frac{(k-1)T}{n}}^{\frac{kT}{n}} \left(\eta (-\dot{x}_t^{(k)})^2 + \frac{\gamma}{2} \sigma^2 (x_t^{(k)} - R^{(k)})^2 \right) dt \\ &= \int_0^T \left[\eta (-\dot{x}_t^{**})^2 + \frac{\theta\sigma^2}{2} (x_t^{**} - R_t)^2 \right] dt, \end{aligned}$$

which implies that the optimal value is obtained when $x_t = x_t^{**}$, in other words, x_t^{**} is the optimal strategy.

□

A.9 Proof to Theorem 3

According to Theorem 2, the optimal strategies x_t and \tilde{x}_t can be given by

$$\begin{aligned} x_t &= \frac{\kappa}{\sinh \kappa T} \int_0^T \left(R_s \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ &\quad + \left(x_0 - \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + A + \left(-A + \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T}, \end{aligned}$$

and

$$\begin{aligned} \tilde{x}_t &= \frac{\kappa}{\sinh \kappa T} \int_0^T \left(\tilde{R}_s \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \\ &\quad + \left(x_0 - \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa(T-t)}{\sinh \kappa T} + A + \left(-A + \frac{\mu}{\theta\sigma^2} \right) \cdot \frac{\sinh \kappa T - \sinh \kappa t}{\sinh \kappa T}, \end{aligned}$$

where x_t, \tilde{x}_t are the optimal strategy with respect to reference strategies R_t and \tilde{R}_t respectively. Hence, by Jensen's inequality,

$$\begin{aligned}
(x_t - \tilde{x}_t)^2 &= \frac{\kappa^2}{\sinh^2 \kappa T} \left(\int_0^T \left((R_s^{(1)} - R_s^{(2)}) \sinh(\kappa \min(s, t)) \sinh(\kappa(T - \max(s, t))) \right) ds \right)^2 \\
&\leq \frac{\kappa^2}{\sinh^2 \kappa T} \int_0^T (R_s^{(1)} - R_s^{(2)})^2 ds \cdot \int_0^T \sinh^2(\kappa \min(s, t)) \sinh^2(\kappa(T - \max(s, t))) ds \\
&< \frac{\kappa^2}{\sinh^2 \kappa T} \cdot \varepsilon \cdot \frac{1}{4\kappa} (\sinh^2(\kappa t) (\sinh(2\kappa(T-t)) - 2\kappa(T-t)) + \sinh^2(\kappa(T-t)) (\sinh(2\kappa t) - 2\kappa t)) \\
&\leq \frac{\kappa^2}{\sinh^2 \kappa T} \cdot \varepsilon \cdot \frac{1}{2\kappa} \sinh^2 \left(\kappa \frac{T}{2} \right) (\sinh \kappa T - \kappa T) = \frac{\kappa (\sinh \kappa T - \kappa T)}{4 (\cosh \kappa T + 1)} \cdot \varepsilon < \frac{\kappa \varepsilon}{4}
\end{aligned}$$

which implies the fact that when $\varepsilon \rightarrow 0$, $\tilde{x}_t \rightarrow x_t$. So we may use the optimal strategy by a piece-wise constant RS to approximate the general RS, with small error

$$|x_t - \tilde{x}_t| < \frac{1}{2} \sqrt{\kappa \varepsilon}.$$

□

References

- [1] R. Almgren and N. Chriss, “Optimal execution of portfolio transactions,” *Journal of Risk*, vol. 3, pp. 5–40, 2001.
- [2] C. Institute, *CFA Program Curriculum 2018 Level II*. John Wiley & Sons, 2017.
- [3] O. Guéant, “Target close execution strategies.” [Online]. Available: <https://www.oliviergeant.com/uploads/4/3/0/9/4309511/targetclose.pdf>
- [4] R. Almgren and N. Chriss, “Value under liquidation,” *Risk*, vol. 12, no. 12, pp. 61–63, 1999.
- [5] R. F. Almgren, “Optimal execution with nonlinear impact functions and trading-enhanced risk,” *Applied Mathematical Finance*, vol. 10, no. 1, pp. 1–18, 2003.
- [6] D. Bertsimas and A. W. Lo, “Optimal control of execution costs,” *Journal of Financial Markets*, vol. 1, no. 1, pp. 1–50, 1998.
- [7] J. Gatheral, “No-dynamic-arbitrage and market impact,” *Quantitative finance*, vol. 10, no. 7, pp. 749–759, 2010.
- [8] J. Gatheral, A. Schied, and A. Slynko, “Transient linear price impact and fredholm integral equations,” *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 22, no. 3, pp. 445–474, 2012.
- [9] S. Tse, P. Forsyth, J. Kennedy, and H. Windcliff, “Comparison between the mean-variance optimal and the mean-quadratic-variation optimal trading strategies,” *Applied Mathematical Finance*, vol. 20, no. 5, pp. 415–449, 2013.
- [10] X. Cheng, M. Di Giacinto, and T.-H. Wang, “Optimal execution with dynamic risk adjustment,” *Journal of the Operational Research Society*, vol. 70, no. 10, pp. 1662–1677, 2019.
- [11] C. Frei and N. Westray, “Optimal execution of a vwap order: a stochastic control approach,” *Mathematical Finance*, vol. 25, no. 3, pp. 612–639, 2015.
- [12] Á. Cartea and S. Jaimungal, “A closed-form execution strategy to target volume weighted average price,” *SIAM Journal on Financial Mathematics*, vol. 7, no. 1, pp. 760–785, 2016.
- [13] O. Guéant and G. Royer, “Vwap execution and guaranteed vwap,” *SIAM Journal on Financial Mathematics*, vol. 5, no. 1, pp. 445–471, 2014.
- [14] R. Almgren and J. Lorenz, “Adaptive arrival price,” *Trading*, vol. 2007, no. 1, pp. 59–66, 2007.
- [15] C. Frei and N. Westray, “Optimal execution in hong kong given a market-on-close benchmark,” *Quantitative Finance*, vol. 18, no. 4, pp. 655–671, 2018.
- [16] O. Guéant, J. Pu, and G. Royer, “Accelerated share repurchase: pricing and execution strategy,” *International Journal of Theoretical and Applied Finance*, vol. 18, no. 03, p. 1550019, 2015.

- [17] L. Bargeron, M. Kulchania, and S. Thomas, “Accelerated share repurchases,” *Journal of Financial Economics*, vol. 101, no. 1, pp. 69–89, 2011.
- [18] X. Cheng, M. Di Giacinto, and T.-H. Wang, “Optimal execution with uncertain order fills in almgren–chriss framework,” *Quantitative Finance*, vol. 17, no. 1, pp. 55–69, 2017.
- [19] R. Carmona and L. Leal, “Optimal execution with quadratic variation inventories,” *SIAM Journal on Financial Mathematics*, vol. 14, no. 3, pp. 751–776, 2023.
- [20] M. Nutz, K. Webster, and L. Zhao, “Unwinding stochastic order flow: When to warehouse trades,” *arXiv preprint arXiv:2310.14144*, 2023.
- [21] Á. Cartea and L. Sánchez-Betancourt, “Brokers and informed traders: dealing with toxic flow and extracting trading signals,” *Available at SSRN 4265814*, 2022.
- [22] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*. Springer Science & Business Media, 2012, vol. 43.
- [23] T. E. Duncan, “Linear-exponential-quadratic gaussian control,” *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2910–2911, 2013.
- [24] A. E. Lim and X. Y. Zhou, “A new risk-sensitive maximum principle,” *IEEE transactions on automatic control*, vol. 50, no. 7, pp. 958–966, 2005.
- [25] R. Carmona and K. Webster, “The self-financing equation in limit order book markets,” *Finance and Stochastics*, vol. 23, no. 3, pp. 729–759, Jul. 2019.
- [26] A. Mazzolo, “Constraint Ornstein-Uhlenbeck bridges,” *Journal of Mathematical Physics*, vol. 58, no. 9, p. 093302, 09 2017.
- [27] O. Guéant, *The Financial Mathematics of Market Liquidity: From optimal execution to market making*. CRC Press, 2016, vol. 33.
- [28] R. C. Merton, “Optimum consumption and portfolio rules in a continuous-time model,” in *Stochastic optimization models in finance*. Elsevier, 1975, pp. 621–661.