Execution game in a Markovian environment*

Masamitsu OHNISHI$‡ and Makoto SHIMOSHIMIZU¶

$ Faculty of Political Science and Economics, Yamato University†

‡Center for Mathematical Modeling and Data Science, Osaka University‡

¶Graduate School of Management, Tokyo Metropolitan University§

Abstract

This paper examines an execution game model in a Markovian environment. We focus on how two risk–averse large traders execute a large volume of a risky asset to maximize the expected utility of each large trader from the terminal wealth over a finite horizon. The price impact caused by each large trader and the Markovian environment are assumed to affect the market and execution price. A formulation as a Markov game model enables us to solve this problem. We obtain an equilibrium execution strategy and its associated value function under a Markov perfect equilibrium via the backward induction method of dynamic programming.

1 Introduction

Developments in trading technology for algorithmic trading have attracted a growing body of research regarding execution problems. According to [21], although traders did not often use high–frequency trading (HFT) around 2000, HFTs have accounted for 20 percent of the total trading volume in the market since the mid–2000s (until 2019). The volume–weighted average price (VWAP) or time–weighted average price (TWAP) strategy was the mainstream of algorithmic trading in the early 2000s. However, liquidity–seeking algorithm usage has become more common since the mid–2000s (until 2019). These facts underscore the importance of analyzing algorithmic trading that large traders have heavily used for more than a decade.

With the above fact in mind, we examine an execution problem for two large traders. In particular, our model sheds light on the effect of a Markovian environment on an "equilibrium" execution strategy for the large traders. We can interpret the Markovian environment in several ways. An example would be to consider the price impact caused by (random) aggregate orders of small traders, such as [32], [13], and [33]. The so–called "order book imbalance," as investigated in [38] and [27], is also significant in the analysis of execution problems and market microstructure. The order book imbalance is of great interest in recent literature. [34] conducts an experimental analysis and shows that public information available for all traders is well incorporated into market prices. This finding supports the model taking into account the order book imbalance since much information about the imbalance is available for all traders.

---

*This research was supported in part by the Research Institute for Mathematical Sciences, a Joint Usage/Research Center, Kyoto University in 2022.

†Address: Suita City, Osaka, 564–0082, Japan
E-mail address: onishi.masamitsu@yamato-u.ac.jp

‡Address: Toyonaka City, Osaka, 560–8531, Japan
E-mail address: ohnishi@econ.osaka-u.ac.jp

§Address: Chiyota–ku, Tokyo, 100–0005, Japan
E-mail address: shimosshi1q94@gmail.com

The model we analyze in this paper is closely related to [27], which investigates the influence of (Markovian) microstructure signal on the optimal execution strategy. Our model can be seen as an extension of this paper from the following viewpoint. Firstly, [27] considers an optimal execution problem for a single large trader, although our model focuses on an interaction between two large traders. To this end, our problem is formulated as a Markov game model. Secondarily, we incorporate the effect of permanent impact on the "fundamental price" we define in the sequel, as opposed to [27].[1] Besides, the most important difference between their model and our model is that we incorporate both transient price impact and risk–averse property of the large trader into the market model for deriving an equilibrium execution. Their analysis shows that a cost minimization problem with a risk–averse term under the influence of market microstructure signal admits at most one optimal execution strategy.[2] However, they only derive an explicit optimal execution strategy under the following situations with the effect of the market microstructure signal: (i) no term representing the risk aversion with transient price impact; (ii) no transient price impact with risk–aversion term. Our formulation enables us to derive an equilibrium execution strategy considering both transient price impact and a risk–averse term under the existence of a Markovian environment.

The model of the "fundamental price" in this paper is also different from seminal papers which stem from [26] and study the field of market microstructure. The so-called "order flow imbalance" is also a key ingredient of price fluctuation, as [9] empirically shows. One of the related works is [2], which studies the effect of dynamic order flow imbalance on an optimal execution. Their model also considers an endogenous impact on the order flow dynamics caused by a large trader as well as a trading horizon under a cost minimization framework. Their analysis is worth mentioning, although the model does not derive the optimal execution strategy explicitly in this setting.

We derive an equilibrium execution strategy at a Markov perfect equilibrium with the effect of a Markovian environment. Large traders are assumed to have a Constant Absolute Risk–Averse (CARA) Von Neumann-Morgenstern (vN–M) type utility. Our analysis prevails that the transient price impact and what we call the residual effect of past price impact and a Markovian environment described by an AR (1)–type normal distributed random variable affect the execution strategy. The derivation method is similar to [32].

The organization of this paper is as follows. Section 2 summarizes related literature. In section 3, we describe a market model where two large traders have large impacts on their execution price due to their large volumes of orders. An effect of a Markovian signal on a traded asset is embedded in the model. We describe the methodology to formulate this model as a Markov game model. Applying the backward induction method of dynamic programming then allows us to obtain an explicit equilibrium execution strategy at a Markov perfect equilibrium as an affine function of four state variables: the remaining execution volume of each large trader, the residual effect of past price impact caused by both large traders, and the last Markovian environment. The proof for the main theorem is shown in the appendix.

2 Related literature

2.1 Optimal execution problem

In the last (two) decades there has been considerable interest in optimal execution problems for a single large trader among academic researchers and practitioners. The first investigation into the optimal execution strategy is conducted by [3] in a discrete–time framework. They find that the optimal strategy becomes a basket of equally divided trading volumes. [1] subsequently extends and constructs their model incorporating both the execution cost caused by a large trader and the

[1] In [27], they call the underlying asset price "unaffected price," which (partially) corresponds to the "fundamental price" we define in our model setting. Our assumption that the permanent price impact would have an impact makes the difference of what we define as the fundamental price.

[2] They show that the objective function representing the cost with risk–averse term for the large trader is strictly convex with respect to the trading speed.

degree of risk–aversion of the large trader. The formulation of their model makes the analysis entail a mean–variance approach. In addition, [35] addresses an optimal execution strategy for a risk–averse large trader with CARA–type utility maximization. They show that the optimal execution strategy for such a large trader becomes deterministic. Another approach for an optimal trade execution has been put forward by [16] and [17]. They incorporate a predictable return into the cost minimization model with (quadratic) transaction cost which can be seen as a price impact in an infinite discrete– and continuous–time framework, respectively.

Much work on the optimal execution strategy has been carried out as we mentioned in the previous paragraph. However, there are still some points that need careful consideration. Firstly, a pitfall with much of the literature on the optimal execution problem including the above research is the lack of a transient part of the price impact. As [4] empirically demonstrates, the price impact dissipates over the trading window. Thus one should take a transient price impact into account. [15] extends the model considered in [4] for a continuous–time framework. [30] subsequently formulates the model from a viewpoint of a limit order book (LOB) dynamics with transient price impact. [32], [13], and [33] study the optimal execution problem with a generalized transient price impact model assuming that aggregate orders posed by small traders also cause price impact. All of these studies highlight the importance of transient price impact being embedded in the analysis of an execution problem by showing that transient price impact does affect the optimal execution strategy. Our formulation of the transient price impact model bears a close resemblance to our previous studies [25], [31], [32] and [13].

We can consider the price impact model with the effect of aggregate trading volumes posed by small traders on the market (and therefore execution) price. [6] and [7] include the price impact caused by order flow (or small traders) under a cost minimization problem for a large trader and derive the optimal execution strategy and the optimal VWAP execution strategy, respectively. Notwithstanding an insightful analysis, both studies, however, offer no explanation for a utility maximization problem. [25], [31], and [32] analyze a utility maximization problem for a large trader with a generalized price impact model (which incorporates the price impact caused by small traders) and derive the optimal execution strategy. These researches show that aggregate orders posed by small traders affect the optimal execution strategy for the large trader through the transient price impact. Moreover, [13] further if aggregate orders posed by small traders have a Markovian dependence, then a "statistical" arbitrage for a large trader exists.

2.2 Execution game

The situation in a real marketplace leads to a game–theoretic formulation, which is the second aspect one should take care of. Since multiple large traders affect the market price they execute with each other, the so–called market impact game model, which can describe a much more complicated financial market, might be more acceptable from a viewpoint of practitioners. [36] and [28], which are motivated by [37], investigate a market impact game model with a transient price impact for one risky asset. These studies then derive an equilibrium at a Nash equilibrium for a cost minimization problem as well as a utility maximization problem. [10] subsequently extend their model to a multiple risky asset one and derive an equilibrium execution strategy. The strategies obtained in these studies are all static and deterministic. However, an execution strategy should be constructed in a dynamic class even if the trading window is very short (e.g., one day or a few days). Thus, in [32] and [31], they address a market game model with transient price impact and derive an equilibrium execution strategy in a dynamic and non–deterministic class. Other researches, [19] or [5] for example, analyze an execution game model via a mean–field game approach, though their model does not take into account a transient price impact. The method to formulate the problem we focus on in this paper is reminiscent of the one used in [31] and [32].

2.3 Execution problem for multiple assets

Another direction of optimal and equilibrium execution problems is an execution problem of multiple (risky) assets. [8], for instance, studies the optimal execution strategy for multiple risky assets considering temporary and permanent price impact. As mentioned above, we should model a transient price impact as well as temporary and permanent price impacts. They also show the way one can incorporate the information that a large trader does not trade. [40] investigates the cross–impact of multiple risky assets in a transient price impact model via a close examination of order book dynamics. They show that a large trader can increase his/her expected utility by execution of other assets even when there is no obligation to buy/sell multiple risky assets (or when they are going to buy/sell only a single risky asset). [33] addresses a pair–trade execution problem for a single large trader and shows that buy and sell orders for each risky asset posed by small traders affect the optimal pair–trade execution volume for both risky assets. For other research, see, e.g.,[10].

2.4 Microstructure effect on market price

The following summary is based on [38]. The order book dynamics attract widespread interest among academic researchers and practitioners from theoretical and empirical points of view. In particular, how we should consider an underlying (or a fundamental) price of a risky asset is undergoing a revolution in the light of empirical analysis. These represent a price without any price impact caused by (large) order submissions. We may recognize the so–called *mid–price* as the underlying price. The mean of best–bid and best–ask accounts for the mid-price:

$$M := \frac{1}{2}\left(P^a + P^b\right), \tag{2.1}$$

where P^a and P^b are respectively the best–bid and best–ask. Another feature that may be of interest to practitioners is the weighted mid–price defined as follows:

$$W := wP^a + (1-w)P^b, \tag{2.2}$$

where the weight w is the order book imbalance defined by the total volume at the best bid Q^b and the total volume at the best ask Q^a:

$$w := \frac{Q^b}{Q^b + Q^a}. \tag{2.3}$$

Although both features make a certain sense in terms of being easily obtained from market data, empirical studies have shown that they have some shortcomings. [38] thus define another notion of what he called the micro–price. The micro–price incorporates the effect of mid–price M, order book imbalance I, and the bid–ask spread $S := P^b - P^a$ into the underlying price. In mathematical form, we can write the dependence as follows:

$$P^{micro} := M + g(I, S), \tag{2.4}$$

using a function g. The method to estimate the function g is explained in [38]. We will follow this spirit and construct what we call the “fundamental price” in the model setting.

3 Execution game model

In a discrete time framework $t \in \{1, \dots, T, T+1\}\,(,T \in \mathbb{Z}_{++} := \{1, 2, \dots\})$, we assume that two large traders, denoted by $i \in \{1, 2\}$, purchase one risky asset in a trading market. It is also supposed that each large trader has a CARA vN-M (or negative exponential) utility function with the absolute risk aversion parameter $\gamma^i > 0$ for $i \in \{1, 2\}$.

3.1 Market

We consider the situation that each large trader must purchase $\mathfrak{Q}^i(\in \mathbb{R})$ volume of one risky asset by the time $T+1$. In the sequel, $q_t^i(\in \mathbb{R})$ stands for the large amount of orders submitted by the large trader $i \in \{1,2\}$ at time $t \in \{1,\ldots,T\}$.[3] We denote by $\overline{Q}_t^i$ the remained execution volume of the risky asset for the large trader $i \in \{1,2\}$, i.e., the number of shares remained to purchase by the large trader at time $t \in \{1,\ldots,T,T+1\}$. So we have

$$\overline{Q}_{t+1}^i = \overline{Q}_t^i - q_t^i, \tag{3.1}$$

with the initial and terminal conditions: $\overline{Q}_1^i = \mathfrak{Q}^i \in \mathbb{R}$; $\overline{Q}_{T+1}^i = 0 \in \mathbb{R}$, for each large trader $i \in \{1,2\}$. In the sequel of this paper, the buy–trade and sell–trade of a large trader are supposed to induce the same (instantaneous) linear price impact.[4]

The market price (or quoted price) of the risky asset at time $t \in \{1,\ldots,T,T+1\}$ is P_t. Then, the execution price of the asset becomes $\widehat{P}_t$ since the large traders submit a large number of orders, influencing the asset price at which they execute the transaction. In the rest of this paper, we assume that submitting one unit of (large) order at time $t \in \{1,\ldots,T\}$ causes the instantaneous price impact denoted as $\lambda_t(>0)$.

We subsequently define the residual effect of past price impact caused by both large traders at time $t \in \{1,\ldots,T\}$, represented by R_t. It characterizes the discounted sum of past transient price impact. Many existing researches, conducted from both theoretical and empirical viewpoints, highlight the significance of the transient nature of price impacts (e.g., [4], [15], and [30]). By means of the following exponential function $G\colon \mathbb{R} \to \mathbb{R}_{++} := (0,\infty)$:

$$G(t) := \mathrm{e}^{-\rho t}, \tag{3.2}$$

where ρ $(\in [0,\infty))$ stands for the deterministic resilience speed, we formulate the residual effect of the past orders posed by both large traders.

Remark 3.1 (Extension of deterministic resilience speed)**.** We can extend the exponential decay kernel model. The time dependency for the resilience speed, i.e., ρ_t, is consistent with empirical analysis. However, we conduct the following analysis without assuming the time dependency for the resilience speed since the assumption does not lead to any illuminating results.

Then the dynamics of the residual effect of past price impact are defined as follows:

$$\begin{aligned}
R_1 &= 0;\\
R_{t+1} &:= \sum_{k=1}^{t} \alpha_k \lambda_k \left(q_k^1 + q_k^2\right) \mathrm{e}^{-\rho((t+1)-k)}\\
&= \mathrm{e}^{-\rho} \sum_{k=1}^{t-1} \alpha_k \lambda_k \left(q_k^1 + q_k^2\right) \mathrm{e}^{-\rho(t-k)} + a_t \lambda_t \left(q_t^1 + q_t^2\right) \mathrm{e}^{-\rho}\\
&= \mathrm{e}^{-\rho} \left[R_t + \alpha_t \lambda_t \left(q_t^1 + q_t^2\right)\right], \quad t = 1,\ldots,T,
\end{aligned} \tag{3.3}$$

where $\alpha_t \in [0,1]$ represents the linear price impact coefficients representing the temporary price impacts. Eq. (3.3) indicates that R_t has a Markov property in this settings, which stems from the assumption of the exponential decay kernel.

Furthermore, we define a sequence of independent random variables ϵ_t at time $t \in \{1,\ldots,T\}$ as the effect of the public news/information about the economic situation between t and $t+1$ since

[3] For each large trader $i \in \{1,2\}$, the positive q_t^i for $t \in \{1,\ldots,T\}$ stand for the acquisition and negative q_t^i the liquidation of the risky asset. This setting allows us to establish a similar setup for a selling problem of large traders.

[4] This assumption is justified by some empirical studies, for example, [6] and [7].

some public news or information affect the price. ϵ_t for $t \in \{ 1, \ldots, T \}$ are assumed to follow a normal distribution with mean $\mu_t^\epsilon \in \mathbb{R}$ and variance $(\sigma_t^\epsilon)^2 \in \mathbb{R}_{++}$, i.e.,

$$\epsilon_t \sim N\left(\mu_t^\epsilon, (\sigma_t^\epsilon)^2\right), \quad t = 1, \ldots, T. \tag{3.4}$$

In the sequel, we assume that $\mu_t^\epsilon = 0$ for all $t \in \{ 1, \ldots, T \}$.

We here focus on the dynamics of the "fundamental price" at time $t \in \{ 1, \ldots, T \}$, denoted by P_t^f. The fact that the residual effect of the past price impact dissipates over the course of the trading horizon allows us to define $P_t - R_t$ as the fundamental price of the risky asset, i.e.,

$$P_t^f := P_t - R_t. \tag{3.5}$$

We assume that the linear permanent price impact is represented by

$$\beta_t \lambda_t \left(q_t^1 + q_t^2\right), \tag{3.6}$$

where $\beta_t \in [0, 1]$. Here the additional factor that affects the fundamental price is assumed to affect the fundamental price. The *Markovian environment*, denoted by $\mathcal{I}_t$, directly influences the fundamental price of the risky asset. The distribution of $\mathcal{I}_t$ is assumed to have a *Markovian dependence* as follows:

$$\begin{aligned} \mathcal{I}_0 &= 0; \\ \mathcal{I}_{t+1}|_{\mathcal{I}_t} &\sim N\left(a_{t+1}^{\mathcal{I}} - b_{t+1}^{\mathcal{I}}\mathcal{I}_t, \left(\sigma_{t+1}^{\mathcal{I}}\right)^2\right). \end{aligned} \tag{3.7}$$

Note that $a_t^{\mathcal{I}}, b_t^{\mathcal{I}}$, and $\left(\sigma_t^{\mathcal{I}}\right)$ are deterministic functions of time t. We can rewrite the dynamics of $\mathcal{I}_t$ as follows:

$$\begin{aligned} \mathcal{I}_0 &= 0; \\ \mathcal{I}_{t+1} &= (a_{t+1}^{\mathcal{I}} - b_{t+1}^{\mathcal{I}}\mathcal{I}_t) + \sigma_{t+1}^{\mathcal{I}}\omega_{t+1}, \quad t = 0, \ldots, T-1, \end{aligned} \tag{3.8}$$

where $\omega_t \sim N(0, 1)$ for all $t \in \{ 1, \ldots, T \}$.

Remark 3.2 (Implication of Markovian environment)**.** The interpretation of a Markovian environment is various and needs to be carefully mentioned. We can consider the price impact caused by *aggregate orders of small traders* as the Markovian environment. [6] and [7], for instance, analyze the effect of order flows on the optimal execution strategy under the existence of temporary and permanent price impacts. [12], [13], and [33] also investigate the case that aggregate orders posed by small traders follow a normal distribution and have a Markovian dependence in a transient price impact as well as temporary and permanent price impacts. These studies show that, under this setting, the small traders' orders directly affect the optimal execution strategy for a single large trader. Another example is the so–called *order book imbalance*. [38] defines a notion of micro–price as an extension of mid–price or weighted mid–price and shows the importance of incorporating order book imbalance into the formulation of market price dynamics. [27] investigates an optimal execution strategy focusing on the effect of order book imbalance (or what they call a *marketmicrostructure signal*) and shows that the signal does influence the optimal execution strategy. From these viewpoints, we can consider the Markovian environment as an extension of these models.

Remark 3.3 (Property of Markovian environment)**.** Eq. (3.7) and (3.8) take the same form as the aggregate orders posed by small traders in [13]. The classification in terms of various conditions for $a_{t+1}^{\mathcal{I}}$ and $b_{t+1}^{\mathcal{I}}$ are the same as and thus detailed in the paper.

Here we make the following assumptions.

Assumption 3.1 (Correlation between two stochastic processes $\mathcal{I}_t$ and ϵ_t)**.** We assume that $\mathcal{I}_t$ and ϵ_t are *correlated* with correlation coefficient $\rho^{\mathcal{I},\epsilon} \in (-1,1)$ for each time $t \in \{1,\dots,T\}$. So we have

$$\begin{pmatrix} \mathcal{I}_{t+1} \\ \epsilon_{t+1} \end{pmatrix} \Bigg|_{\mathcal{I}_t} \sim N\left(\begin{pmatrix} a^{\mathcal{I}}_{t+1} - b^{\mathcal{I}}_{t+1}\mathcal{I}_t \\ \mu^{\epsilon}_{t+1} \end{pmatrix}, \begin{pmatrix} \left(\sigma^{\mathcal{I}}_{t+1}\right)^2 & \rho^{\mathcal{I},\epsilon}\sigma^{\mathcal{I}}_{t+1}\sigma^{\epsilon}_{t+1} \\ \rho^{\mathcal{I},\epsilon}\sigma^{\mathcal{I}}_{t+1}\sigma^{\epsilon}_{t} & (\sigma^{\epsilon}_{t+1})^2 \end{pmatrix} \right). \tag{3.9}$$

In addition, no other sequential dependencies between two stochastic sequences exist in the sequel.

By definition of ϵ_t, we define the dynamics of the fundamental price $P^f_t := P_t - R_t$ with Markovian environment and the permanent price impact as follows:

$$\begin{aligned} P^f_{t+1} &:= P^f_t + \beta_t\lambda_t\left(q^1_t + q^2_t\right) + \mathcal{I}_t + \epsilon_t \\ &\quad (= P_{t+1} - R_{t+1}) \\ &\quad = P_t - R_t + \beta_t\lambda_t\left(q^1_t + q^2_t\right) + \mathcal{I}_t + \epsilon_t, \quad t = 1,\dots,T. \end{aligned} \tag{3.10}$$

Remark 3.4 (Implication of Eq. (3.10))**.** The above relationships indicate that the permanent price impact caused by large traders and the public news or information about an economic situation is assumed to affect the fundamental price. This assumption also reveals that the permanent price impact may give a non–zero trend to the fundamental price. For a more detailed discussion, see [32].

According to Eq. (3.3) and (3.10), the dynamics of market price are described as

$$\begin{aligned} P_{t+1} &= P_t + (R_{t+1} - R_t) + \beta_t\lambda_t\left(q^1_t + q^2_t\right) + \mathcal{I}_t + \epsilon_t \\ &= P_t - (1 - \mathrm{e}^{-\rho})R_t + \left(\alpha_t\mathrm{e}^{-\rho} + \beta_t\right)\lambda_t\left(q^1_t + q^2_t\right) + \mathcal{I}_t + \epsilon_t, \quad t = 1,\dots,T. \end{aligned} \tag{3.11}$$

We here consider the following assumption in the rest of this paper.

Assumption 3.2. For $\alpha_t \in [0,1]$, $\beta_t \in [0,1]$ and $\rho \in [0,\infty)$, the relationships

$$\alpha_t\mathrm{e}^{-\rho} + \beta_t < 1 \tag{3.12}$$

holds for all $t \in \{1,\dots,T\}$.

The implication for Eq. (3.12) is that the friction of permanent and transient price impact at time $t \in \{1,\dots,T\}$ is (strictly) less than the price impact caused by both large traders. This assumption is plausible from the perspective of limit order book dynamics.

Remark 3.5. In this context,

$$\beta_t\lambda_t\left(q^1_t + q^2_t\right); \quad \alpha_t\lambda_t\left(q^1_t + q^2_t\right); \quad \mathrm{e}^{-\rho}\alpha_t\lambda_t\left(q^1_t + q^2_t\right), \tag{3.13}$$

represent the permanent impact, temporary impact, and transient impact at time $t \in \{1,\dots,T\}$, respectively. Moreover, if $\rho \to \infty$, the residual effect of past price impact becomes zero for all $t \in \{1,\dots,T\}$ since $R_1 = 0$ and from Eq. (3.3)

$$\lim_{\rho\to\infty} R_{t+1} = \lim_{\rho\to\infty} \mathrm{e}^{-\rho}\left[R_t + \alpha_t\lambda_t\left(q^1_t + q^2_t\right)\right] = 0, \quad t = 1,\dots,T, \tag{3.14}$$

and therefore,

$$P_{t+1} = P_t + \beta_t\lambda_t\left(q^1_t + q^2_t\right) + \mathcal{I}_t + \epsilon_t, \quad t = 1,\dots,T, \tag{3.15}$$

that is, we have a permanent impact model (with an effect of the Markovian environment).

From the definition of the execution price, the wealth process for each large trader $i \in \{1,2\}$, denoted by W^i_t ($\in \mathbb{R}$), evolves as follows:

$$W^i_{t+1} = W^i_t - \widehat{P}_t q^i_t = W^i_t - \left\{P_t + \lambda_t\left(q^1_t + q^2_t\right)\right\} q^i_t, \quad t = 1,\dots,T. \tag{3.16}$$

3.2 Formulation as a Markov game

In a discrete–time window $t \in \{ 1, \ldots, T, T+1 \}$, we define the state of the decision process at time $t \in \{ 1, \ldots, T, T+1 \}$ as 7–tuple and denote it as

$$\boldsymbol{s}_t = \left(W_t^1, W_t^2, P_t, \overline{Q}_t^1, \overline{Q}_t^2, R_t, \mathcal{I}_{t-1} \right) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} =: S. \tag{3.17}$$

For $t \in \{ 1, \ldots, T \}$, an allowable action chosen at state $\boldsymbol{s}_t$ is an execution volume $q_t^i \in \mathbb{R} =: A^i$ so that the set A^i of admissible actions is independent of the current state $\boldsymbol{s}_t$.

When an action q_t^i is chosen in a state $\boldsymbol{s}_t$ at time $t \in \{ 1, \ldots, T \}$, a transition to a next state

$$\boldsymbol{s}_{t+1} = (W_{t+1}^1, W_{t+1}^2, P_{t+1}, \overline{Q}_{t+1}^1, \overline{Q}_{t+1}^2, R_{t+1}, \mathcal{I}_t) \in S \tag{3.18}$$

occurs according to the law of motion which we have precisely described in the previous subsection. We symbolically describe the transition by a (Borel measurable) system dynamics function $\mathbf{h}_t$ (: $S \times A \times A \times (\mathbb{R} \times \mathbb{R}) \longrightarrow S$):

$$\boldsymbol{s}_{t+1} = \boldsymbol{h}_t(\boldsymbol{s}_t, q_t^1, q_t^2, (\omega_t, \epsilon_t)), \quad t = 1, \ldots, T. \tag{3.19}$$

A utility payoff (or reward) arises only in a terminal state $\boldsymbol{s}_{T+1}$ at the end of horizon $T+1$ as

$$g_{T+1}^i(\boldsymbol{s}_{T+1}) := \begin{cases} -\exp\{-\gamma^i W_{T+1}^i\} & \text{if } \overline{Q}_{T+1}^i = 0; \\ -\infty & \text{if } \overline{Q}_{T+1}^i \neq 0, \end{cases} \tag{3.20}$$

where $\gamma^i > 0$ represents the risk aversion parameter of the large trader $i \in \{ 1, 2 \}$. The term $-\infty$ means a hard constraint enforcing the large trader to execute all of the remaining volume $\overline{Q}_T^i$ at the maturity T, that is, $q_T^i = \overline{Q}_T^i$. The types of large traders could be defined by

$$(W^i, \mathfrak{Q}^i, \gamma^i), i = 1, 2, \tag{3.21}$$

and these are assumed to be their common knowledge.[5]

If we define a (history–independent) one–stage decision rule f_t at time $t \in \{ 1, \ldots, T \}$ by a Borel measurable map from a state $\boldsymbol{s}_t \in S = \mathbb{R}^5$ to an action

$$q_t^i = f_t^i(\boldsymbol{s}_t) \in A = \mathbb{R}, \tag{3.22}$$

then a Markov execution strategy π is defined as a sequence of one–stage decision rules

$$\pi^i := (f_1^i, \ldots, f_t^i, \ldots, f_T^i). \tag{3.23}$$

We denote the set of all Markov execution strategies as Π_{M}. Further, for $t \in \{ 1, \ldots, T \}$, we define the sub–execution strategy after time t of a Markov execution strategy $\pi \in \Pi_{\mathrm{M}}$ as

$$\pi_t^i := (f_t^i, \ldots, f_T^i), \tag{3.24}$$

and the entire set of π_t^i as $\Pi_{\mathrm{M},t}^i$.

[5]In a real market, large traders have little access to this information of the counterpart. We can, however, consider a plausible explanation for the assumption of Eq. (3.21) from the viewpoint of a game–theoretic analysis. In this model, our focus is placed on how the existence of the other large trader influences the execution strategy in comparison with a single large trader's (optimal) execution problem. We formulate this Markov game model as a dynamic game of complete information. Therefore, the above (hypothesized) definition and assumption associated with the notion of common knowledge are legitimate so that the solution concept of a Nash equilibrium in a non-cooperative game is (rationally or ideally) applicable in this model. The formulation of a generalized model as a dynamic game of incomplete information requires further intricate analysis, which is left for our future research.

By definition (3.20), the value function under a pair of execution strategies (π^1,π^2) becomes an expected utility payoff arising from the terminal wealth W^i_{T+1} of the large trader $i \in \{1,2\}$ with the absolute risk aversion parameter $\gamma^i \in \mathbb{R}_{++}$:

$$\begin{aligned} V^i_1(\pi^1,\pi^2)\big[\boldsymbol{s}_1\big] &= \mathbb{E}^{(\pi^1,\pi^2)}_1\Big[g^i_{T+1}(\boldsymbol{s}_{T+1})\Big|\boldsymbol{s}_1\Big] \\ &= \mathbb{E}^{\pi}_1\Big[-\exp\big\{-\gamma^i W^i_{T+1}\big\}\cdot \mathbb{1}_{\big\{\overline{Q}^i_{T+1}=0\big\}} + (-\infty)\cdot \mathbb{1}_{\big\{\overline{Q}^i_{T+1}\neq 0\big\}}\Big|\boldsymbol{s}_1\Big]. \end{aligned} \tag{3.25}$$

Then, for $t \in \{1,\dots,T,T+1\}$ and $\boldsymbol{s}_t \in S$, we further let

$$\begin{aligned} V^i_t(\pi^1_t,\pi^2_t)\big[\boldsymbol{s}_t\big] &= \mathbb{E}^{(\pi^1_t,\pi^2_t)}_t\Big[g^i_{T+1}(\boldsymbol{s}_{T+1})\Big|\boldsymbol{s}_t\Big] \\ &= \mathbb{E}^{(\pi^1_t,\pi^2_t)}_t\Big[-\exp\big\{-\gamma^i W^i_{T+1}\big\}\cdot \mathbb{1}_{\big\{\overline{Q}^i_{T+1}=0\big\}} + (-\infty)\cdot \mathbb{1}_{\big\{\overline{Q}^i_{T+1}\neq 0\big\}}\Big|\boldsymbol{s}_t\Big], \end{aligned} \tag{3.26}$$

be the expected utility payoff at time t under the strategy π. Note that the expression of the conditional expectation, $\mathbb{E}^{(\pi^1,\pi^2)}_1$ in (3.25) and $\mathbb{E}^{(\pi^1_t,\pi^2_t)}_t$ in (3.25), implies the dependence of the probability laws on the strategy profiles, (π^1,π^2) and (π^1_t,π^2_t), respectively. Also, $\mathbb{1}_A$ stands for the indicator function of an event A.

What we seek here is an equilibrium execution strategy for large traders. First, we consider the definition of a Nash equilibrium in this model as follows.

Definition 3.1 (Nash Equilibrium)**.** $(\pi^{1*},\pi^{2*}) \in \Pi^1_{\mathrm{M}} \times \Pi^2_{\mathrm{M}}$ is a *Nash equilibrium* starting from a fixed initial state $\boldsymbol{s}_1$ if and only if

$$V^1_1(\pi^{1*},\pi^{2*})\big[\boldsymbol{s}_1\big] \geq V^1_1(\pi^1,\pi^{2*})\big[\boldsymbol{s}_1\big], \quad \forall \pi^1 \in \Pi^1_{\mathrm{M}}; \tag{3.27}$$

$$V^2_1(\pi^{1*},\pi^{2*})\big[\boldsymbol{s}_1\big] \geq V^2_1(\pi^{1*},\pi^2)\big[\boldsymbol{s}_1\big], \quad \forall \pi^2 \in \Pi^2_{\mathrm{M}}. \tag{3.28}$$

We can define a refinement of the Nash equilibrium of this model as the notion of a Markov perfect equilibrium:

Definition 3.2 (Markov Perfect Equilibrium)**.** $(\pi^{1*},\pi^{2*}) \in \Pi^1_{\mathrm{M}} \times \Pi^2_{\mathrm{M}}$ is a *Markov perfect equilibrium* if and only if

$$V^1_t(\pi^{1*}_t,\pi^{2*}_t)\big[\boldsymbol{s}_t\big] \geq V^1_t(\pi^1_t,\pi^{2*}_t)\big[\boldsymbol{s}_t\big], \quad \forall \pi^1_t \in \Pi^1_{\mathrm{M},t}, \quad \forall \boldsymbol{s}_t \in S, \quad \forall t = 1,\dots,T; \tag{3.29}$$

$$V^2_t(\pi^{1*}_t,\pi^{2*}_t)\big[\boldsymbol{s}_t\big] \geq V^1_t(\pi^{1*}_t,\pi^2_t)\big[\boldsymbol{s}_t\big], \quad \forall \pi^2_t \in \Pi^2_{\mathrm{M},t}, \quad \forall \boldsymbol{s}_t \in S, \quad \forall t = 1,\dots,T. \tag{3.30}$$

Based on the following *One Stage [Step, Shot] Deviation Principle*, we obtain an equilibrium execution strategy at a Markov perfect equilibrium by backward induction procedure of dynamic programming from time T to 1.

$$\begin{aligned} V^1_t(\pi^{1*}_t,\pi^{2*}_t)\big[\boldsymbol{s}_t\big] &= \sup_{q^1_t \in \mathbb{R}} \mathbb{E}\Big[V^1_{t+1}(\pi^{1*}_{t+1},\pi^{2*}_{t+1})\big[\boldsymbol{h}_t(\boldsymbol{s}_t,(q^1_t,f^{2*}_t(\boldsymbol{s}_t)),(\omega_t,\epsilon_t))\big]\Big|\boldsymbol{s}_t\Big] \\ &= \mathbb{E}\Big[V^1_{t+1}(\pi^{1*}_{t+1},\pi^{2*}_{t+1})\big[\boldsymbol{h}_t(\boldsymbol{s}_t,(f^{1*}_t(\boldsymbol{s}_t),f^{2*}_t(\boldsymbol{s}_t)),(\omega_t,\epsilon_t))\big]\Big|\boldsymbol{s}_t\Big]; \end{aligned} \tag{3.31}$$

$$\begin{aligned} V^2_t(\pi^{1*}_t,\pi^{2*}_t)\big[\boldsymbol{s}_t\big] &= \sup_{q^2_t \in \mathbb{R}} \mathbb{E}\Big[V^2_{t+1}(\pi^{1*}_{t+1},\pi^{2*}_{t+1})\big[\boldsymbol{h}_t(\boldsymbol{s}_t,(f^{1*}_t(\boldsymbol{s}_t),q^2_t),(\omega_t,\epsilon_t))\big]\Big|\boldsymbol{s}_t\Big] \\ &= \mathbb{E}\Big[V^2_{t+1}(\pi^{1*}_{t+1},\pi^{2*}_{t+1})\big[\boldsymbol{h}_t(\boldsymbol{s}_t,(f^{1*}_t(\boldsymbol{s}_t),f^{2*}_t(\boldsymbol{s}_t)),(\omega_t,\epsilon_t))\big]\Big|\boldsymbol{s}_t\Big]. \end{aligned} \tag{3.32}$$

3.3 Equilibrium execution under a Markov perfect equilibrium

Theorem 3.1 (Equilibrium execution under a Markov perfect equilibrium)**.** There exists a Markov perfect equilibrium at which the following properties hold for each large trader $i \in \{1,2\}$:

1. The execution volume at the Markov perfect equilibrium for the large trader $i \in \{1, 2\}$ at time $t \in \{1, \dots, T\}$, denoted as q_t^{i*}, becomes an *affine function* of the Markovian environment *at time* $t-1$, $\mathcal{I}_{t-1}$, the remaining execution volume of each large trader, $\overline{Q}_t^i$ and $\overline{Q}_t^j$ $(i \neq j, i, j \in \{1, 2\})$, and the cumulative residual effect, R_t, that is,
$$q_t^{i*} = a_t^i + b_t^i \overline{Q}_t^i + c_t^i \overline{Q}_t^j + d_t^i R_t + e_t^i \mathcal{I}_{t-1}, \quad t = 1, \dots, T. \tag{3.33}$$
The dynamics of $a_t^i, b_t^i, c_t^i, d_t^i, e_t^i$ for $t \in \{1, \dots, T, T+1\}$ are deterministic functions of time t which are dependent on the problem parameters and can be computed backwardly in time t from maturity T.
2. The value function $V_t^i(\pi^1, \pi^2)[\boldsymbol{s}_t]$ at time $t \in \{1, \dots, T, T+1\}$ for each large trader $i \in \{1, 2\}$ is represented as a functional form as follows:
$$\begin{aligned} & V_t^i(\pi_t^1, \pi_t^2)\left[W_t^1, W_t^2, P_t, \overline{Q}_t^1, \overline{Q}_t^2, R_t, \mathcal{I}_{t-1}\right] \\ & = -\exp\Big\{ -\gamma \Big[W_t^i - P_t^\top \overline{Q}_t^i + G_t^{1i} \left(\overline{Q}_t^i\right)^2 + G_t^{2i} \overline{Q}_t^i + H_t^{1i} \overline{Q}_t R_t \\ & \quad + H_t^{2i} R_t^2 + H_t^{3i} R_t + I_t^{1i} \overline{Q}_t^i \overline{Q}_t^j + I_t^{2i} \overline{Q}_t^j R_t + I_t^{3i} \left(\overline{Q}_t^j\right)^2 + I_t^{4i} \overline{Q}_t^j \\ & \quad + J_t^{1i} \overline{Q}_t^i \mathcal{I}_{t-1} + J_t^{2i} R_t \mathcal{I}_{t-1} + J_t^{3i} \overline{Q}_t^j \mathcal{I}_{t-1} + J_t^{4i} \mathcal{I}_{t-1}^2 + J_t^{5i} \mathcal{I}_{t-1} + Z_t^i \Big] \Big\}, \end{aligned} \tag{3.34}$$
where $G_t^{1i}, G_t^{2i}, H_t^{1i}, H_t^{2i}, H_t^{3i}, I_t^{1i}, I_t^{2i}, I_t^{3i}, I_t^{4i}, J_t^{1i}, J_t^{2i}, J_t^{3i}, J_t^{4i}, J_t^{5i}, Z_t^i$ for $t \in \{1, \dots, T, T+1\}$ are deterministic functions of time t which are dependent on the problem parameters and can be computed backwardly in time t from maturity T.

Proof. See Appendix A □

As the above theorem shows, the equilibrium execution volume q_t^{i*} at the Markov perfect equilibrium for $t \in \{1, \dots, T\}$ depends on the state $\boldsymbol{s}_t = (W_t^1, W_t^2, P_t, \overline{Q}_t^1, \overline{Q}_t^2, R_t, \mathcal{I}_{t-1}) \in S$ of the decision process through the Markovian environment at the previous time, $\mathcal{I}_{t-1}$, in addition to the remaining execution volume of each large trader, $\overline{Q}_t^i$ for $i \in \{1, 2\}$, and the cumulative residual effect, R_t, and not through the wealth of each large trader, W_t^i for $i \in \{1, 2\}$, or market price P_t. Furthermore, by the definition of the Markovian environment, the equilibrium execution volume q_t^{i*} for $t \in \{1, \dots, T\}$ includes a nondeterministic term through $\mathcal{I}_{t-1}$.

Corollary 3.1 (Deterministicness of the equilibrium execution strategy)**.** If the Markovian environment for $t \in \{1, \dots, T\}$ are deterministic, the equilibrium execution volumes q_t^{i*} at time $t \in \{1, \dots, T\}$ for each large trader also become deterministic functions of time.

Remark 3.6 (In the case without transient price impact)**.** If we consider only temporary and permanent price impact, the optimal execution volume for the large trader at time $t \in \{1, \dots, T\}$ becomes
$$q_t^{i*} = a_t^i + b_t^i \overline{Q}_t^i + c_t^i \overline{Q}_t^j + d_t^i \mathcal{I}_{t-1}. \tag{3.35}$$
In this case, the Markovian environment affects the optimal execution volume of the large trader. However, if we further assume that $\mathcal{I}_t$ is an independent random sequence and follows a normal distribution as follows:
$$\mathcal{I}_t \sim N_{\mathbb{R}}\left(\mu_t^{\mathcal{I}}, \sigma_t^{\mathcal{I}}\right), \tag{3.36}$$
then the equilibrium execution volume of each large trader at time $t \in \{1, \dots, T\}$ takes the form as follows:
$$q_t^{i*} = a_t^i + b_t^i \overline{Q}_t^i + c_t^j \overline{Q}_t^j, \tag{3.37}$$
meaning that the Markovian environment does not affect the equilibrium execution strategy, even if we incorporate the effect of the environment on the fundamental or market price.

4 Conclusion

This paper examines an execution game under a transient price impact with a Markovian environment. We then derive an equilibrium execution strategy and its associated value function at a Markov perfect equilibrium and show that the Markovian environment directly affects the execution strategy.

One direction of future research is to consider an endogenous model for optimal or equilibrium execution problems. The submission of large orders by large traders may affect the subsequent orders posed by small traders in a real market. Thus incorporating the orders submitted by large traders into the modeling of aggregate orders posed by small traders endogenously deserve consideration. This model may enable us to investigate the interaction between large traders and small traders in detail.

Acknowledgements

The authors would like to thank Dr. Seiya Kuno for closely examining our preliminary draft and the comments at the 3rd Autumn Meeting of the Nippon Finance Association.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Japan Society for the Promotion of Science under KAKENHI [Grant Numbers 21H04399, 21K13325, 22K01573].

References

[1] Almgren, R., and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, **3**, 5–40.

[2] Bechler, K., and Ludkovski, M. (2015). Optimal execution with dynamic order flow imbalance. *SIAM Journal on Financial Mathematics*, **6**(1), 1123–1151.

[3] Bertsimas, D., and Lo, A. W. (1998). Optimal control of execution costs. *Journal of Financial Markets*, **1**(1), 1–50.

[4] Bouchaud, J. P., Gefen, Y., Potters, M., and Wyart, M. (2004). Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance*, **4**(2), 176–190.

[5] Casgrain, P., and Jaimungal, S. (2020). Mean‑field games with differing beliefs for algorithmic trading. *Mathematical Finance*, **30**(3), 995–1034.

[6] Cartea, Á., and Jaimungal, S. (2016a). Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, **10**(3), 339–364.

[7] Cartea, Á., and Jaimungal, S. (2016b). A closed-form execution strategy to target volume weighted average price. *SIAM Journal on Financial Mathematics*, **7**(1), 760–785.

[8] Cartea, Á., Gan, L., and Jaimungal, S. (2019). Trading co-integrated assets with price impact. *Mathematical Finance*, **29**(2), 542–567.

[9] Cont, R., Kukanov, A., and Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics*, **12**(1), 47–88.

[10] Cordoni, F., and Lillo, F. (2022). Instabilities in multi-asset and multi-agent market impact games. *Annals of Operations Research*, 1–35.

[11] Eisler, Z., Bouchaud, J. P., and Kockelkoren, J. (2012). The price impact of order book events: market orders, limit orders and cancellations. *Quantitative Finance*, **12**(9), 1395–1419.

[12] Fukasawa, M., Ohnishi, M., and Shimoshimizu, M. (2020). Optimal execution strategies with generalized price impact in a discrete–time setting. *RIMS Kokyuroku*, **2158**, 66–79.

[13] Fukasawa, M., Ohnishi, M., and Shimoshimizu, M. (2021a). Discrete–time optimal execution under a generalized price impact model with Markov exogenous orders. *International Journal of Theoretical and Applied Finance*, **24**(05), 2150025.

[14] Fukasawa, M., Ohnishi, M., and Shimoshimizu, M. (2021b). Optimal execution under a generalized price impact model with Markovian exogenous orders in a continuous–time setting. *RIMS Kokyuroku*, **2207**, 1–22.

[15] Gatheral, J. (2010). No-dynamic-arbitrage and market impact. *Quantitative Finance*, **10**(7), 749–759.

[16] Gârleanu, N., and Pedersen, L. H. (2013). Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, **68**(6), 2309–2340.

[17] Gârleanu, N., and Pedersen, L. H. (2016). Dynamic portfolio choice with frictions. *Journal of Economic Theory*, **165**, 487–516.

[18] Guéant, O. (2016). *The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making*. CRC Press.

[19] Huang, X., Jaimungal, S., and Nourian, M. (2019). Mean-field game strategies for optimal execution. *Applied Mathematical Finance*, 26(2), 153–185.

[20] Huberman, G., and Stanzl, W. (2004). Price manipulation and quasi-arbitrage. *Econometrica*, **72**(4), 1247–1275.

[21] Kissell, R. (2021). *Algorithmic Trading Methods: Applications Using Advanced Statistics, Optimization, and Machine Learning Techniques.* Academic Press.

[22] Kunou, S., and Ohnishi, M. (2010) Optimal execution strategies with price impact. *RIMS Kokyuroku*, **1675**, 234–247.

[23] Kuno, S., and Ohnishi, M. (2015). Optimal execution in illiquid market with the absence of price manipulation. *Journal of Mathematical Finance*, **5**(1), 1–15.

[24] Kuno, S., Ohnishi, M., and Shimizu, P. (2016). Optimal off-exchange execution with closing price. *Journal of Mathematical Finance*, **7**(1), 54–64.

[25] Kuno, S., Ohnishi, M., and Shimoshimizu, M. (2018) Optimal execution strategies with generalized price impact models. *RIMS Kokyuroku*, **2078**, 77–83.

[26] Kyle, A. S. (1985) Continuous auctions and insider trading. *Econometrica*, **53**, 1315–1335.

[27] Lehalle, C. A., and Neuman, E. (2019). Incorporating signals into optimal trading. *Finance and Stochastics*, **23**(2), 275–311.

[28] Luo, X., and Schied, A. (2020) Nash equilibrium for risk-averse investors in a market impact game with transient price impact. *Market Microstructure and Liquidity*, **5**(01), 2050001.

[29] Ma, G., Siu, C. C., and Zhu, S. P. (2019). Dynamic portfolio choice with return predictability and transaction costs. *European Journal of Operational Research*, **278**(3), 976–988.

[30] Obizhaeva, A. A., and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, **16**(1), 1–32.

[31] Ohnishi, M., and Shimoshimizu, M. (2019). Equilibrium execution strategy with generalized price impacts. *RIMS Kokyuroku*, **2111**, 84–106.

[32] Ohnishi, M., and Shimoshimizu, M. (2020). Optimal and equilibrium execution strategies with generalized price impact. *Quantitative Finance*, **20**(10), 1625–1644.

[33] Ohnishi, M., and Shimoshimizu, M. (2022). Optimal pair–trade execution with generalized cross–impact. *Asia-Pacific Financial Markets*, **29**, 253–289.

[34] Page, L., and Siemroth, C. (2021). How much information is incorporated in financial asset prices? Experimental Evidence. *Review of Financial Studies*, **34**(9), 4412-–4449.

[35] Schied, A., Schöneborn, T., and Tehranchi, M., Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance*, **17**, 471–489, 2010.

[36] Schied, A., and Zhang, T. (2019). A market impact game under transient price impact. *Mathematics of Operations Research*, **44**(1), 102–121.

[37] Schöneborn, T. (2008) Trade execution in illiquid markets: Optimal stochastic control and multi-agent equilibria. Doctoral dissertation, Technische Universität Berlin.

[38] Stoikov, S. (2018) The micro-price: a high-frequency estimator of future prices. *Quantitative Finance*, **18**(12), 1959–1966.

[39] Tong, Y. L. (2012). *The Multivariate Normal Distribution*. Springer Science and Business Media.

[40] Tsoukalas, G., Wang, J., and Giesecke, K. (2019). Dynamic portfolio execution. *Management Science*, **65**(5), 2015–2040.

Appendix

We here use the notation $\mathbb{R}^n$ to denote the set of all n–dimensional real–valued vectors and $\mathcal{M}_n(\mathbb{R})$ to denote the set of all $n \times n$ real–valued square matrices. For an $n \times m$ real–valued matrix (or vector) $\mathbf{A}$, we denote by $\mathbf{A}^\top$ the transpose of the matrix (or vector). Moreover, if a random variable $\boldsymbol{X}$ follows an n–dimensional normal distribution with mean $\boldsymbol{\mu}_{\boldsymbol{X}} \in \mathbb{R}^n$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{X}} \in \mathcal{M}_n(\mathbb{R})$, we denote it by $\boldsymbol{X} \sim N_{\mathbb{R}^n}(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\Sigma}_{\boldsymbol{X}})$.

A Distribution of sum of normally distributed random variables with correlation

Here we show, as a lemma, that any finite sum of normally distributed random variables with correlation also follows a normal distribution. Although the statement is straightforward, we note the result below for this paper to be self–contained.

Lemma A.1 (Distribution of sum of normally distributed random variables with correlation)**.** Define, for a set of random variables $X_1, X_2, \ldots, X_n$, $\mathbb{E}[X_i] := \mu^i$ and $\mathrm{Cov}[X_i, X_j] := \sigma^{ij}$. If an $\mathbb{R}^n$–valued random variable $\boldsymbol{X} := (X_1, X_2, \ldots, X_n)$ follows a normal distribution with mean $\boldsymbol{\mu}_{\boldsymbol{X}} \in \mathbb{R}^n$ and variance $\boldsymbol{\Sigma}_{\boldsymbol{X}} \in \mathcal{M}_n(\mathbb{R})$, i.e.,

$$\boldsymbol{X} \sim N_{\mathbb{R}^n}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\Sigma}_{\boldsymbol{X}}\right), \tag{A.1}$$

where

$$\boldsymbol{\mu}_{\boldsymbol{X}} := \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \in \mathbb{R}^n; \quad \boldsymbol{\Sigma}_{\boldsymbol{X}} := \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{nn} & \sigma_{nn} & \cdots & \sigma_{nn} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}), \tag{A.2}$$

then the following (one–dimensional) sum of the random variables each of which is multiplied by constants:

$$\boldsymbol{c}^\top \boldsymbol{X} := c_1 X_1 + c_2 X_2 + \cdots + c_n X_n, \tag{A.3}$$

where $\boldsymbol{c} := (c_1, \ldots, c_n)^\top \in \mathbb{R}^n$, also follows a normal distribution with mean $\sum_{i=1}^{n} c_i \mu_i \in \mathbb{R}$ and variance $\sum_{i,j=1}^{n} c_i c_j \sigma_{ij} \in \mathbb{R}_{++} := (0, \infty)$.

Proof. The characteristic function for the random variable $\boldsymbol{X}$, denoted by $\varphi(\boldsymbol{c})$, is given by

$$\varphi(\boldsymbol{c}) := \mathbb{E}\left[\exp\left\{\mathrm{i}\boldsymbol{c}^\top \boldsymbol{X}\right\}\right] = \exp\left\{\mathrm{i}\boldsymbol{c}^\top \boldsymbol{\mu}_{\boldsymbol{X}} - \frac{1}{2}\boldsymbol{c}^\top \boldsymbol{\Sigma}_{\boldsymbol{X}} \boldsymbol{c}\right\}, \tag{A.4}$$

where i is the imaginary number that satisfies $\mathrm{i}^2 = -1$. The characteristic function of a random variable uniquely determines its probabilistic law or distribution. Thus, Eq. (A.4) shows that the random variable $\sum_{i=1}^{n} c_i X_i$ follows a normal distribution with mean $\boldsymbol{c}^\top \boldsymbol{\mu}_{\boldsymbol{X}} = \sum_{i=1}^{n} c_i \mu^i$ and variance $\boldsymbol{c}^\top \boldsymbol{\Sigma}_{\boldsymbol{X}} \boldsymbol{c} = \sum_{i,j=1}^{n} c_i c_j \sigma_{ij}$. □

Remark A.1 (Definition as a multivariate normal distribution)**.** Some books state the result above as the definition for a random variable to follow a normal distribution (for example, [39]).

B Proof of Theorem 3.1

We derive an equilibrium execution volume q_t^{i*} at the Markov perfect equilibrium for each large trader $i \in \{1, 2\}$ and time $t \in \{1, \dots, T\}$ by backward induction method of dynamic programming from time $t = T$ via the following steps.

Step 1 From the assumption that each large trader must unwind all the remainder of his/her position at time $t = T$, we have

$$\overline{Q}^i_{T+1} = \overline{Q}^i_T - q^i_T = 0, \tag{B.1}$$

for $i \in \{1, 2\}$. Thus, $q^i_T = \overline{Q}^i_T$ holds. Then, for $t = T$, the value function for each large trader $i, j \in \{1, 2\}$ $(i \neq j)$ is:

$$\begin{aligned}
V^i_T(\pi^{1*}_T, \pi^{2*}_T)\big[\boldsymbol{s}_T\big] &= \sup_{q^i_T \in \mathbb{R}} \mathbb{E}\Big[V_{T+1}(\pi^{1*}_{T+1}, \pi^{2*}_{T+1})\big[\boldsymbol{s}_{T+1}\big]\Big|\boldsymbol{s}_T\Big] \\
&= \sup_{q^i_T \in \mathbb{R}} \mathbb{E}\Big[-\exp\left\{-\gamma^i W^i_{T+1}\right\}\Big|W^1_T, W^2_T, P_T, \overline{Q}^1_T, \overline{Q}^2_T, R_T, \mathcal{I}_{T-1}\Big] \\
&= \sup_{q^i_T \in \mathbb{R}} \mathbb{E}\Big[-\exp\Big\{-\gamma^i\big[W^i_T - \Big[P_T + \lambda_T\left(q^i_T + q^j_T\right)\Big]^\top q^i_T\big]\Big\}\Big|W^1_T, W^2_T, P_T, \overline{Q}^1_T, \overline{Q}^2_T, R_T, \mathcal{I}_{T-1}\Big] \\
&= -\exp\Big\{-\gamma^i\big[W^i_T - P^\top_T \overline{Q}^i_T - \lambda_T\left(\overline{Q}^i_T\right)^2 - \lambda_T \overline{Q}^i_T \overline{Q}^j_T\big]\Big\} \\
&= -\exp\Big\{-\gamma^i\big[W^i_T - P^\top_T \overline{Q}^i_T + G^{1i}_T\left(\overline{Q}^i_T\right)^2 + J^{1i}_T \overline{Q}^i_T \overline{Q}^j_T\big]\Big\},
\end{aligned} \tag{B.2}$$

where

$$G^{1i}_T := -\lambda_T (< 0); \tag{B.3}$$

$$J^{1i}_T := -\lambda_T (< 0). \tag{B.4}$$

Step 2 For $t = T-1$, the value functions, $V^i_{T-1}(\pi^{1*}_{T-1}, \pi^{2*}_{T-1})\big[\boldsymbol{s}_{T-1}\big]$ for each large trader $i \in \{1, 2\}$, satisfy the following functional equations:

$$\begin{aligned}
&V^i_{T-1}(\pi^{1*}_{T-1}, \pi^{2*}_{T-1})\big[\boldsymbol{s}_{T-1}\big] \\
&= \sup_{q_{T-1} \in \mathbb{R}} \mathbb{E}\Big[V^i_T(\pi^{1*}_T, \pi^{2*}_T)\big[\boldsymbol{s}_T\big]\Big|\boldsymbol{s}_{T-1}\Big] \\
&= \sup_{q^i_{T-1} \in \mathbb{R}} \mathbb{E}\Big[-\exp\Big\{-\gamma^i\big[W^i_T - P^\top_T \overline{Q}^i_T + G^{1i}_T (\overline{Q}^i_T)^2 + J^{1i}_T \overline{Q}^i_T \overline{Q}^j_T\big]\Big\}\Big|\boldsymbol{s}_{T-1}\Big] \\
&= \sup_{q^i_{T-1} \in \mathbb{R}} -\exp\Big\{-\gamma^i\Big[\left(-\lambda_{T-1} + \lambda_{T-1}\alpha^{T-1} + G^{1i}_T\right)(q^i_{T-1})^2 + \big[(-\lambda_{T-1}\alpha^{T-1} - 2G^{1i}_T)\overline{Q}^i_{T-1} \\
&\quad + (-J^{i1}_T)\overline{Q}^j_{T-1} + \{-(1 - \mathrm{e}^{-\rho})\}R_{T-1} + (-\lambda_{T-1} + \lambda_{T-1}\alpha^{T-1} + J^{1i}_T)q^j_{T-1}\big]q^i_{T-1} \\
&\quad + W^i_{T-1} - P_{T-1}\overline{Q}^i_{T-1} + G^{1i}_T(\overline{Q}^i_{T-1})^2 + (1 - \mathrm{e}^{-\rho})\overline{Q}^i_{T-1}R_{T-1} + J^{1i}_{T-1}\overline{Q}^i_{T-1}\overline{Q}^j_{T-1} \\
&\quad + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)\overline{Q}^i_{T-1}q^j_{T-1}\Big]\Big\} \\
&\quad \times \mathbb{E}\Big[\exp\Big\{\gamma^i(\overline{Q}^i_{T-1} - q^i_{T-1})(\mathcal{I}_{T-1} + \epsilon_{T-1})\Big\}\Big|\boldsymbol{s}_{T-1}\Big],
\end{aligned} \tag{B.5}$$

where $\alpha^{T-1} := \alpha_{T-1}\mathrm{e}^{-\rho} + \beta_{T-1}$. As for the expectation term in Eq. (B.5), we have

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big\{\gamma^i(\overline{Q}^i_{T-1} - q^i_{T-1})(\mathcal{I}_{T-1} + \epsilon_{T-1})\Big\}\Big|\boldsymbol{s}_{T-1}\Big] \\
&= \exp\Big\{\gamma^i(\overline{Q}^i_{T-1} - q^i_{T-1})(a^{\mathcal{I}}_{T-1} - b^{\mathcal{I}}_{T-1}\mathcal{I}_{T-2}) + \frac{1}{2}(\overline{Q}^i_{T-1} - q^i_{T-1})^2 \Sigma^{\mathcal{I},\epsilon}_{T-1}\Big\},
\end{aligned} \tag{B.6}$$

where

$$\Sigma^{\mathcal{I},\epsilon}_{T-1} := \mathbb{V}\Big[\mathcal{I}_{T-1} + \epsilon_{T-1}\Big|\boldsymbol{s}_{T-1}\Big] := (\sigma^{\mathcal{I}}_{T-1})^2 + (\sigma^{\epsilon}_{T-1})^2 + 2\rho^{\mathcal{I},\epsilon}\sigma^{\mathcal{I}}_{T-1}\sigma^{\epsilon}_{T-1}, \tag{B.7}$$

according to the lemma shown in Appendix A. Thus, substituting Eq. (B.6) into Eq. (B.5) and rearranging results in

$$\begin{aligned} &V^i_{T-1}(\pi^{1*}_{T-1}, \pi^{2*}_{T-1})\big[\boldsymbol{s}_{T-1}\big] \\ &= \sup_{q^i_{T-1}\in\mathbb{R}} -\exp\Big\{-\gamma^i\Big[-A^i_{T-1}(q^i_{T-1})^2 + \big[B^i_{T-1}\overline{Q}^i_{T-1} + C^i_{T-1}\overline{Q}^j_{T-1} + D^i_{T-1}R_{T-1} + F^i_{T-1}\mathcal{I}_{T-2} + M^i_{T-1} \\ &\quad + N^i_{T-1}q^j_{T-1}\big]q^i_{T-1} + W^i_{T-1} - P_{T-1}\overline{Q}^i_{T-1} + \left(G^{1i}_{T-1} - \frac{1}{2}\gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1}\right)\left(\overline{Q}^i_{T-1}\right)^2 + (-a^{\mathcal{I}}_{T-1})\overline{Q}^i_{T-1} \\ &\quad + (1-\mathrm{e}^{-\rho})R_{T-1}\overline{Q}^i_{T-1} + J^{1i}_T\overline{Q}^i_{T-1}\overline{Q}^j_{T-1} + b^{\mathcal{I}}_{T-1}\overline{Q}^i_{T-1}\mathcal{I}_{t-2} + \big(-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T\big)\,\overline{Q}^i_{T-1}q^j_{T-1}\Big]\Big\}, \end{aligned} \tag{B.8}$$

with the following relations:

$$\begin{cases} A^i_{T-1} := \lambda_{T-1} - \lambda_{T-1}\alpha^{T-1} - G^{1i}_T + \frac{1}{2}\gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1}(>0); \\ B^i_{T-1} := -\lambda_{T-1}\alpha^{T-1} - 2G^{1i}_T + \gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1}; \\ C^i_{T-1} := -J^{1i}_T; \\ D^i_{T-1} := -(1-\mathrm{e}^{-\rho}); \\ F^i_{T-1} := -b^{\mathcal{I}}_{T-1}; \\ M^i_{T-1} := a^{\mathcal{I}}_{T-1}; \\ N^i_{T-1} := -\lambda_{T-1} + \lambda_{T-1}\alpha^{T-1} + J^{1i}_T. \end{cases} \tag{B.9}$$

Note that, for all $x, B, C \in \mathbb{R}$ and all $\gamma, A \in \mathbb{R}_{++} := (0,\infty)$, two functions $c_1(x) := -\exp\{-\gamma x\}$ and $c_2(x) := -Ax^2 + Bx + C$ are strictly concave function, and therefore so is the composite function of the two, $K(x) := c_1 \circ c_2(x) = -\exp\big\{-\gamma(-Ax^2 + Bx + C)\big\}$. Thus, we obtain the execution volume attaining the supremum of Eq. (B.8) by completing the square of the following function:

$$\begin{aligned} K_{T-1}(q_{T-1}) &:= -A^i_{T-1}(q^i_{T-1})^2 + \big[B^i_{T-1}\overline{Q}^i_{T-1} + C^i_{T-1}\overline{Q}^j_{T-1} + D^i_{T-1}R_{T-1} + F^i_{T-1}\mathcal{I}_{T-2} + M^i_{T-1} \\ &\quad + N^i_{T-1}q^j_{T-1}\big]q^i_{T-1} + W^i_{T-1} - P_{T-1}\overline{Q}^i_{T-1} + \left(G^{1i}_{T-1} - \frac{1}{2}\gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1}\right)\left(\overline{Q}^i_{T-1}\right)^2 + (-a^{\mathcal{I}}_{T-1})\overline{Q}^i_{T-1} \\ &\quad + (1-\mathrm{e}^{-\rho})R_{T-1}\overline{Q}^i_{T-1} + J^{1i}_T\overline{Q}^i_{T-1}\overline{Q}^j_{T-1} + b^{\mathcal{I}}_{T-1}\overline{Q}^i_{T-1}\mathcal{I}_{t-2} + \big(-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T\big)\,\overline{Q}^i_{T-1}q^j_{T-1}. \end{aligned} \tag{B.10}$$

Then, the best response of large trader $i \in \{1, 2\}$ to the other large trader, denoted by $BR^i(q^j_{T-1})$, becomes

$$BR^i(q^j_{T-1}) = \frac{1}{2A^i_{T-1}}\left(B^i_{T-1}\overline{Q}^i_{T-1} + C^i_{T-1}\overline{Q}^j_{T-1} + D^i_{T-1}R_{T-1} + F^i_{T-1}\mathcal{I}_{T-2} + M^i_{T-1} + N^i_{T-1}q^j_{T-1}\right). \tag{B.11}$$

Thus, at the Markov perfect equilibrium, we have

$$\begin{cases} q^{i*}_{T-1} = \dfrac{1}{2A^i_{T-1}}\left(B^i_{T-1}\overline{Q}^i_{T-1} + C^i_{T-1}\overline{Q}^j_{T-1} + D^i_{T-1}R_{T-1} + F^i_{T-1}\mathcal{I}_{T-2} + M^i_{T-1} + N^i_{T-1}q^{j*}_{T-1}\right); \\ q^{j*}_{T-1} = \dfrac{1}{2A^j_{T-1}}\left(B^j_{T-1}\overline{Q}^j_{T-1} + C^j_{T-1}\overline{Q}^i_{T-1} + D^j_{T-1}R_{T-1} + F^j_{T-1}\mathcal{I}_{T-2} + M^j_{T-1} + N^j_{T-1}q^{i*}_{T-1}\right). \end{cases} \tag{B.12}$$

Solving the above simultaneous equations results in

$$\begin{aligned} q^{i*}_{T-1} &:= B^{i*}_{T-1}\overline{Q}^i_{T-1} + C^{i*}_{T-1}\overline{Q}^j_{T-1} + D^{i*}_{T-1}R_{T-1} + F^{i*}_{T-1}\mathcal{I}_{T-2} + M^{i*}_{T-1} \\ &\quad (=: a^i_{T-1} + b^i_{T-1}\overline{Q}^i_{T-1} + c^i_{T-1}\overline{Q}^j_{T-1} + d^i_{T-1}R_{T-1} + e^i_{T-1}\mathcal{I}_{T-2}), \end{aligned} \tag{B.13}$$

where

$$\left\{ \begin{array}{l} \zeta^i_{T-1} := 2A^i_{T-1} - \dfrac{N^i_{T-1}N^j_{T-1}}{2A^j_{T-1}}; \quad B^{i*}_{T-1} := \dfrac{1}{\zeta^i_{T-1}}\left(B^i_{T-1} + \dfrac{N^i_{T-1}C^j_{T-1}}{2A^j_{T-1}}\right); \\ C^{i*}_{T-1} := \dfrac{1}{\zeta^i_{T-1}}\left(C^i_{T-1} + \dfrac{N^i_{T-1}B^j_{T-1}}{2A^j_{T-1}}\right); \quad D^{i*}_{T-1} := \dfrac{1}{\zeta^i_{T-1}}\left(D^i_{T-1} + \dfrac{N^i_{T-1}D^j_{T-1}}{2A^j_{T-1}}\right); \\ F^{i*}_{T-1} := \dfrac{1}{\zeta^i_{T-1}}\left(F^i_{T-1} + \dfrac{N^i_{T-1}F^j_{T-1}}{2A^j_{T-1}}\right); \quad M^{i*}_{T-1} := \dfrac{1}{\zeta^i_{T-1}}\left(M^i_{T-1} + \dfrac{N^i_{T-1}M^j_{T-1}}{2A^j_{T-1}}\right). \end{array} \right. \tag{B.14}$$

for each $i \in \{1,2\}$. q^{1*}_{T-1} and q^{2*}_{T-1} are equilibrium execution volumes at the Markov perfect equilibrium for time $T-1$. Then the value function for large trader $i \in \{1,2\}$ at the Markov Perfect equilibrium $(\pi^{1*}, \pi^{2*}) \in \Pi^1_M \times \Pi^2_M$ becomes

$$\begin{aligned} & V^i_{T-1}(\pi^{1*}_{T-1}, \pi^{2*}_{T-1})\big[\boldsymbol{s}_{T-1}\big] \\ &= -\exp\Big\{ -\gamma^i\Big[W^i_{T-1} - P_{T-1}\overline{Q}^i_{T-1} + \left(G^{1i}_{T-1} - \frac{1}{2}\gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1}\right)\left(\overline{Q}^i_{T-1}\right)^2 + (-a^{\mathcal{I}}_{T-1} - \mu^{\epsilon}_{T-1})\overline{Q}^i_{T-1} \\ &\quad + (1-\mathrm{e}^{-\rho})R_{T-1}\overline{Q}^i_{T-1} + J^{1i}_T\overline{Q}^i_{T-1}\overline{Q}^j_{T-1} + b^{\mathcal{I}}_{T-1}\overline{Q}^i_{T-1}\mathcal{I}_{t-2} + \left(-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T\right)\overline{Q}^i_{T-1}q^{j*}_{T-1} \\ &\quad + \frac{1}{4A^i_{T-1}}(B^{i**}_{T-1}\overline{Q}^i_{T-1} + C^{i**}_{T-1}\overline{Q}^j_{T-1} + D^{i*}_{T-1}R_t + F^{i**}_{T-1}\mathcal{I}_{T-2} + L^{i**}_{T-1})^2\Big]\Big\} \\ &= -\exp\Big\{ -\gamma^i\Big[W^i_{T-1} - P_{T-1}\overline{Q}^i_{T-1} + G^{1i}_{T-1}(\overline{Q}^i_{T-1})^2 + G^{2i}_{T-1}\overline{Q}^i_{T-1} + H^{1i}_{T-1}\overline{Q}^i_{T-1}R_{T-1} \\ &\quad + H^{2i}_{T-1}R^2_{T-1} + H^{3i}_{T-1}R_{T-1} + J^{1i}_{T-1}\overline{Q}^i_{T-1}\overline{Q}^j_{T-1} + J^{2i}_{T-1}\overline{Q}^j_{T-1}R_{T-1} + J^{3i}_{T-1}(\overline{Q}^j_{T-1})^2 + J^{4i}_{T-1}\overline{Q}^j_{T-1} \\ &\quad + L^{1i}_{T-1}\overline{Q}^i_t\mathcal{I}_{T-2} + L^{2i}_{T-1}R_{T-1}\mathcal{I}_{T-2} + L^{3i}_{T-1}\overline{Q}^j_t\mathcal{I}_{T-2} + L^{4i}_{T-1}\mathcal{I}^2_{T-1} + L^{5i}_{T-1}\mathcal{I}_{T-2} + Z^i_{T-1}\Big]\Big\}, \end{aligned} \tag{B.15}$$

where

$$\left\{ \begin{array}{l} B^{i**}_{T-1} := B^i_{T-1} + N^i_{T-1}C^{j*}_{T-1}; \quad C^{i**}_{T-1} := C^i_{T-1} + N^i_{T-1}B^{j*}_{T-1}; \\ D^{i**}_{T-1} := D^i_{T-1} + N^i_{T-1}D^{j*}_{T-1}; \quad F^{i**}_{T-1} := F^i_{T-1} + N^i_{T-1}F^{j*}_{T-1}; \\ M^{i**}_{T-1} := M^i_{T-1} + N^i_{T-1}M^{j*}_{T-1}, \quad i,j = 1,2, \quad i \neq j, \end{array} \right. \tag{B.16}$$

and

$$\left\{ \begin{array}{l} G^{1i}_{T-1} := G^{1i}_T - \dfrac{1}{2}\gamma^i\Sigma^{\mathcal{I},\epsilon}_{T-1} + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)C^{j*}_{T-1} + \dfrac{(B^{i**}_{T-1})^2}{4A^i_{T-1}}; \\ G^{2i}_{T-1} := -a^{\mathcal{I}}_{T-1} + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)M^{j*}_{T-1} + \dfrac{B^{i**}_{T-1}M^{i**}_{T-1}}{2A^i_{T-1}}; \\ H^{1i}_{T-1} := (1-\mathrm{e}^{-\rho}) + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)D^{j*}_{T-1} + \dfrac{B^{i**}_{T-1}D^{i**}_{T-1}}{2A^i_{T-1}}; \\ H^{2i}_{T-1} := \dfrac{(D^{i**}_{T-1})^2}{4A^i_{T-1}}; \quad H^{3i}_{T-1} := \dfrac{D^{i**}_{T-1}M^{i**}_{T-1}}{2A^i_{T-1}}; \quad J^{1i}_{T-1} := J^{1i}_T + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)B^{j*}_{T-1} + \dfrac{B^{i**}_{T-1}C^{i**}_{T-1}}{2A^i_{T-1}}; \\ J^{2i}_{T-1} := \dfrac{C^{i**}_{T-1}D^{i**}_{T-1}}{2A^i_{T-1}}; \quad J^{3i}_{T-1} := \dfrac{(C^{i**}_{T-1})^2}{4A^i_{T-1}}; \quad J^{4i}_{T-1} := \dfrac{C^{i**}_{T-1}M^{i**}_{T-1}}{2A^i_{T-1}}; \\ L^{1i}_{T-1} := b^{\mathcal{I}}_{T-1} + (-\lambda_{T-1}\alpha^{T-1} - J^{1i}_T)F^{j*}_{T-1} + \dfrac{B^{i**}_{T-1}F^{i**}_{T-1}}{2A^i_{T-1}}; \quad L^{2i}_{T-1} := \dfrac{D^{i**}_{T-1}F^{i**}_{T-1}}{2A^i_{T-1}}; \\ L^{3i}_{T-1} := \dfrac{C^{i**}_{T-1}F^{i**}_{T-1}}{2A^i_{T-1}}; \quad L^{4i}_{T-1} := \dfrac{(F^{i**}_{T-1})^2}{4A^i_{T-1}}; \quad L^{5i}_{T-1} := \dfrac{F^{i**}_{T-1}M^{i**}_{T-1}}{2A^i_{T-1}}; \quad Z^i_{T-1} := \dfrac{(M^{i**}_{T-1})^2}{4A^i_{T-1}}. \end{array} \right.$$

Step 3 For $t+1 \in \{T-1, \ldots, 2\}$, we can assume from the above results that, at time $t+1$, the optimal value function has the following functional form:

$$
\begin{aligned}
&V^i_{t+1}(\pi^{1*}_{t+1}, \pi^{2*}_{t+1})\big[\boldsymbol{s}_{t+1}\big] \\
&= -\exp\Big\{ -\gamma\Big[W^i_{t+1} - P_{t+1}\overline{Q}^i_{t+1} + G^{1i}_{t+1}\left(\overline{Q}^i_{t+1}\right)^2 + G^{2i}_{t+1}\overline{Q}^i_{t+1} + H^{1i}_{t+1}\overline{Q}^i_{t+1}R_{t+1} \\
&\quad + H^{2i}_{t+1}R^2_{t+1} + H^{3i}_{t+1}R_{t+1} + J^{1i}_{t+1}\overline{Q}^i_{t+1}\overline{Q}^j_{t+1} + J^{2i}_t R_{t+1}\overline{Q}^j_{t+1} + J^{3i}_{t+1}\left(\overline{Q}^j_{t+1}\right)^2 + J^{4i}_{t+1}\overline{Q}^j_{t+1} \\
&\quad + L^{1i}_{t+1}\overline{Q}^i_{t+1}\mathcal{I}_t + L^{2i}_{t+1}R_{t+1}\mathcal{I}_t + L^{3i}_{t+1}\overline{Q}^j_{t+1}\mathcal{I}_t + L^{4i}_{t+1}\mathcal{I}^2_t + L^{5i}_{t+1}\mathcal{I}_t + Z^i_{t+1}\Big]\Big\}.
\end{aligned} \tag{B.17}
$$

Then, at time t,

$$
\begin{aligned}
&V^i_t(\pi^{1*}_t, \pi^{2*}_t)\big[\boldsymbol{s}_t\big] \\
&= \sup_{q^i_t\in\mathbb{R}} \mathbb{E}\Big[V^i_{t+1}(\pi^{1*}_{t+1}, \pi^{2*}_{t+1})\big[\boldsymbol{s}_{t+1}\big]\Big|\boldsymbol{s}_t\Big] \\
&= \sup_{q^i_t\in\mathbb{R}} -\exp\Big\{-\gamma^i\Big[-\big\{(1-\alpha^t)\lambda_t - G^{1i}_{t+1} + \alpha_t\lambda_t \mathrm{e}^{-\rho}H^{1i}_{t+1} + \alpha^2_t\lambda^2_t\mathrm{e}^{-2\rho}H^{2i}_{t+1}\big\}(q^i_t)^2 \\
&\quad + \big[\big(-\alpha^t\lambda_t - 2G^{1i}_{t+1} + \alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1}\big)\overline{Q}^i_t + \{-(1-\mathrm{e}^{-\rho}) - \mathrm{e}^{-\rho}H^{1i}_{t+1} + 2\alpha_t\lambda_t\mathrm{e}^{-2\rho}H^{2i}_{t+1}\}R_t \\
&\quad + \big(-J^{1i}_{t+1} + \alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1}\big)\overline{Q}^j_t + \big\{-(1-\alpha^t)\lambda_t - \alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1} + 2\alpha^2_t\lambda^2_t\mathrm{e}^{-2\rho}H^{2i}_{t+1} + J^{1i}_{t+1} \\
&\quad - \alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1}\big\}q^j_t + \big(-G^{2i}_{t+1} + \alpha_t\lambda_t\mathrm{e}^{-\rho}H^{3i}_{t+1}\big)\big]q^i_t \\
&\quad + W^i_t - P_t\overline{Q}^i_t + G^{1i}_{t+1}(\overline{Q}^i_t)^2 + G^{2i}_{t+1}\overline{Q}^i_t + \{(1-\mathrm{e}^{-\rho}) + \mathrm{e}^{-\rho}H^{1i}_{t+1}\}\overline{Q}^i_t R_t + \mathrm{e}^{-2\rho}H^{2i}_{t+1}R^2_t + \mathrm{e}^{-\rho}H^{3i}_{t+1}R_t \\
&\quad + J^{1i}_{t+1}\overline{Q}^i_t\overline{Q}^j_t + \mathrm{e}^{-\rho}J^{2i}_{t+1}R_t\overline{Q}^j_t + J^{3i}_{t+1}(\overline{Q}^j_t)^2 + J^{4i}_{t+1}\overline{Q}^j_t + Z^i_{t+1} \\
&\quad + \big(\alpha^2_t\lambda^2_t\mathrm{e}^{-2\rho}H^{2i}_{t+1} - \alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1} + J^{3i}_{t+1}\big)(q^j_t)^2 + \big[\big(-\alpha^t\lambda_t + \alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1} - J^{1i}_{t+1}\big)\overline{Q}^i_t \\
&\quad + \big(2\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{2i}_{t+1} - \mathrm{e}^{-\rho}J^{2i}_{t+1}\big)R_t + \big(\alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1} - 2J^{3i}_{t+1}\big)\overline{Q}^j_t + \big(\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{3i}_{t+1} - J^{4i}_{t+1}\big)\big]q^j_t\Big]\Big\} \\
&\times \mathbb{E}\Big[\exp\Big\{-\gamma^i\Big[L^{4i}_{t+1}\mathcal{I}^2_t + \big[(1 - L^{1i}_{t+1} + \alpha_t\lambda_t\mathrm{e}^{-\rho}L^{2i}_{t+1})q^i_t + (-1 + L^{1i}_{t+1})\overline{Q}^i_t + \mathrm{e}^{-\rho}L^{2i}_{t+1}R_t + L^{3i}_{t+1}\overline{Q}^j_t \\
&\quad + L^{5i}_{t+1} + \big(\alpha_t\lambda_t\mathrm{e}^{-\rho}L^{2i}_{t+1} - L^{3i}_{t+1}\big)q^j_t\big]\mathcal{I}_t - \big(\overline{Q}^i_t - q^i_t\big)\epsilon_t\Big\}\Big|\boldsymbol{s}_t\Big],
\end{aligned} \tag{B.18}
$$

where $\alpha^t := \alpha_t\mathrm{e}^{-\rho} + \beta_t$. Here we have the following result, which is a two–dimensional version of Lemma A.1 in [33]. We show the result here for this paper to be self–contained, although this result is not so difficult.

Lemma B.1. Suppose

$$
\begin{pmatrix} X \\ Y \end{pmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{B.19}
$$

where

$$
\boldsymbol{\mu} := \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \in \mathbb{R}^2; \quad \boldsymbol{\Sigma} := \begin{pmatrix} \sigma^2_X & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma^2_Y \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}), \tag{B.20}
$$

and $\rho \in (-1, 1)$ is the correlation coefficient between X and Y. Then, for any $a, b, c \in \mathbb{R}$, we have

$$
\mathbb{E}\Big[\exp\Big\{aX^2 + bX + cY\Big\}\Big] = \frac{\sqrt{|\tilde{\boldsymbol{\Sigma}}|}}{\sqrt{|\boldsymbol{\Sigma}|}}\exp\Big\{\frac{1}{2}(\boldsymbol{\mu}^*)^\top(\boldsymbol{\Sigma}^*)^{-1}\boldsymbol{\mu}^* - \frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\Big\}, \tag{B.21}
$$

where

$$
\tilde{\boldsymbol{\Sigma}} := \begin{pmatrix} \widetilde{\sigma_{11}} & \widetilde{\sigma_{12}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix} = \boldsymbol{\Sigma}^{-1}, \quad \boldsymbol{\Sigma}^* := \begin{pmatrix} \widetilde{\sigma_{11}} - 2a & \widetilde{\sigma_{12}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix}, \quad \boldsymbol{\mu}^* := \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \boldsymbol{s}, \tag{B.22}
$$

provided that $\boldsymbol{\Sigma}^*$ is positive definite (that is, $\boldsymbol{\Sigma}^*$ is invertible).

Proof. Define $\boldsymbol{b} := \begin{pmatrix} b \\ c \end{pmatrix} \in \mathbb{R}^2$ and denote the inverse of $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma}^{-1} =: \begin{pmatrix} \widetilde{\sigma_{11}} & \widetilde{\sigma_{11}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix}$. Then,

$$
\begin{aligned}
&\mathbb{E}\Big[\exp\Big\{aX^2 + bX + cY\Big\}\Big] \qquad \text{(B.23)}\\
&= \int_{\mathbb{R}^2} \exp\left\{\boldsymbol{x}^\top \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x}\right\} \frac{1}{2\pi|\boldsymbol{\Sigma}|} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \mathrm{d}\boldsymbol{x}\\
&= \frac{1}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_{\mathbb{R}^2} \exp\left\{\boldsymbol{x}^\top \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \boldsymbol{x} + \begin{pmatrix} b & c \end{pmatrix}^\top \boldsymbol{x} - \frac{1}{2}\boldsymbol{x}^\top \begin{pmatrix} \widetilde{\sigma_{11}} & \widetilde{\sigma_{11}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix} \boldsymbol{x} + \boldsymbol{\mu}^\top \begin{pmatrix} \widetilde{\sigma_{11}} & \widetilde{\sigma_{11}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix} \boldsymbol{x}\right.\\
&\qquad \left. - \frac{1}{2}\boldsymbol{\mu}^\top \begin{pmatrix} \widetilde{\sigma_{11}} & \widetilde{\sigma_{11}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix} \boldsymbol{\mu}\right\} \mathrm{d}\boldsymbol{x}\\
&= \frac{1}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_{\mathbb{R}^2} \exp\left\{-\frac{1}{2}\boldsymbol{x}^\top \begin{pmatrix} \widetilde{\sigma_{11}} - 2a & \widetilde{\sigma_{11}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix} \boldsymbol{x} + \Big[\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} + \boldsymbol{s}^\top\Big] \boldsymbol{x} - \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right\} \mathrm{d}\boldsymbol{x}\\
&= \frac{2\pi|(\boldsymbol{\Sigma}^*)^{-1}|^{\frac{1}{2}}}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot \underbrace{\frac{1}{2\pi|(\boldsymbol{\Sigma}^*)^{-1}|^{\frac{1}{2}}} \int_{\mathbb{R}^2} \exp\left\{-\frac{1}{2}\left(\boldsymbol{x} - (\boldsymbol{\Sigma}^*)^{-1}\boldsymbol{\mu}^*\right)^\top \boldsymbol{\Sigma}^* \left(\boldsymbol{x} - (\boldsymbol{\Sigma}^*)^{-1}\boldsymbol{\mu}^*\right)\right\} \mathrm{d}\boldsymbol{x}}_{=1}\\
&\quad \times \exp\left\{\frac{1}{2}(\boldsymbol{\mu}^*)^\top (\boldsymbol{\Sigma}^*)^{-1} \boldsymbol{\mu}^* - \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right\}. \qquad \text{(B.24)}
\end{aligned}
$$

Note that $\mathrm{d}\boldsymbol{x} := \mathrm{d}x_1 \mathrm{d}x_2$. □

Define $\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} := \begin{pmatrix} \widetilde{\sigma_{11}} - 2a & \widetilde{\sigma_{12}} \\ \widetilde{\sigma_{21}} & \widetilde{\sigma_{22}} \end{pmatrix}^{-1} = (\boldsymbol{\Sigma}^*)^{-1}$. Then, Rearranging Eq. (B.24) results in

$$
\mathbb{E}\Big[\exp\Big\{aX^2 + bX + cY\Big\}\Big] = \frac{\sqrt{|\tilde{\boldsymbol{\Sigma}}|}}{\sqrt{|\boldsymbol{\Sigma}|}} \exp\left\{\frac{1}{2}\Big[\pi_{11}b^2 + \pi_{22}c^2 + 2\pi_{12}bc + 2\mu^b b + 2\mu^c c + \mu^a\Big]\right\}, \qquad \text{(B.25)}
$$

where

$$
\begin{aligned}
\mu^a &:= \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \tilde{\boldsymbol{\Sigma}} (\boldsymbol{\Sigma}^*)^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}; && \text{(B.26)}\\
\mu^b &:= (\widetilde{\sigma_{11}}\pi_{11} + \widetilde{\sigma_{12}}\pi_{21})\,\mu_X + (\widetilde{\sigma_{21}}\pi_{11} + \widetilde{\sigma_{22}}\pi_{21})\,\mu_Y; && \text{(B.27)}\\
\mu^c &:= (\widetilde{\sigma_{11}}\pi_{12} + \widetilde{\sigma_{12}}\pi_{22})\,\mu_X + (\widetilde{\sigma_{21}}\pi_{12} + \widetilde{\sigma_{22}}\pi_{22})\,\mu_Y. && \text{(B.28)}
\end{aligned}
$$

(Note that $(\boldsymbol{\Sigma}^*)^{-1}$ is symmetric.) Define

$$
\begin{cases}
\theta_t^i &:= 1 - L_{t+1}^{1i} + \alpha_t \lambda_t \mathrm{e}^{-\rho} L_{t+1}^{2i};\\
\delta_t^i &:= -1 + L_{t+1}^{1i};\\
\phi_t^i &:= \alpha_t \lambda_t \mathrm{e}^{-\rho} L_{t+1}^{2i} - L_{t+1}^{3i}.
\end{cases} \qquad \text{(B.29)}
$$

Then, letting

$$
\begin{cases}
a &:= (-\gamma^i) L_{t+1}^{4i};\\
b &:= -\gamma^i \left(\theta_t^i q_t^i + \delta_t^i \overline{Q}_t^i + \mathrm{e}^{-\rho} L_{t+1}^{2i} R_t + L_{t+1}^{3i} \overline{Q}_t^j + L_{t+1}^{5i} + \phi_t^i q_t^j\right);\\
c &:= \gamma^i \left(\overline{Q}_t^i - q_t^i\right),
\end{cases} \qquad \text{(B.30)}
$$

and rearranging Eq. (B.18) results in

$$
\begin{aligned}
&V_t^i(\pi_t^{1*}, \pi_t^{2*})\big[\boldsymbol{s}_t\big]\\
&= \sup_{q_t^i \in \mathbb{R}} -\exp\Big\{-\gamma^i\Big[-A_t^i(q_t^i)^2 + \big[B_t^i \overline{Q}_t^i + C_t^i \overline{Q}_t^j + D_t^i R_t + F_t^i \mathcal{I}_{t-1} + M_t^i + N_t^i q_t^j\big] q_t^i\\
&\quad + W_t^i - P_t \overline{Q}_{T-1}^i \qquad \text{(B.31)}
\end{aligned}
$$

$$
\begin{aligned}
&+\left\{G^{1i}_{t+1}-\frac{1}{2}\gamma^i\delta^i_t\pi^i_t-\frac{1}{2}\gamma^i\pi^{22}_t+\gamma^i\pi^{12}_t\delta^i_t\right\}(\overline{Q}^i_t)^2\\
&+\left\{G^{2i}_{t+1}-\gamma^i\delta^i_tL^{5i}_{t+1}\pi^{11}_t+\gamma^i\pi^{12}_tL^{5i}_{t+1}+\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)a^{\mathcal{I}}_t\delta^i_t-\left(\widetilde{\sigma^{11}_t}\pi^{12}_t+\widetilde{\sigma^{12}_t}\pi^{22}_t\right)a^{\mathcal{I}}_t\right\}\overline{Q}^i_t,\\
&+\left\{\left(1-\mathrm{e}^{-\rho}\right)+\mathrm{e}^{-\rho}H^{1i}_{t+1}-\gamma^i\delta^i_t\mathrm{e}^{-\rho}L^{2i}_{t+1}\pi^{11}_t+\gamma^i\pi^{12}_t\mathrm{e}^{-\rho}L^{2i}_{t+1}\right\}\overline{Q}^i_tR_t\\
&+\left\{\mathrm{e}^{-2\rho}H^{2i}_{t+1}-\frac{1}{2}\gamma^i\mathrm{e}^{-2\rho}\left(L^{2i}_{t+1}\right)^2\pi^{11}_t\right\}R^2_t\\
&+\left\{\mathrm{e}^{-\rho}H^{3i}_{t+1}-\gamma^i\mathrm{e}^{-\rho}L^{2i}_{t+1}L^{5i}_{t+1}\pi^{11}_t+\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)a^{\mathcal{I}}_t\mathrm{e}^{-\rho}L^{2i}_{t+1}\right\}R_t\\
&+\left\{J^{1i}_{t+1}-\gamma^i\delta^i_tL^{3i}_{t+1}\pi^{11}_t+\gamma^i\pi^{12}_tL^{3i}_{t+1}\right\}\overline{Q}^i_t\overline{Q}^j_t\\
&+\left\{\mathrm{e}^{-2\rho}J^{2i}_{t+1}-\gamma^i\mathrm{e}^{-\rho}L^{2i}_{t+1}L^{3i}_{T+1}\pi^{11}_t\right\}R_t\overline{Q}^j_t\\
&+\left\{J^{3i}_{t+1}-\frac{1}{2}\gamma^i\left(L^{3i}_{t+1}\right)^2\pi^{11}_t\right\}(\overline{Q}^j_t)^2\\
&+\left\{J^{4i}_{t+1}-\gamma^iL^{3i}_{t+1}L^{5i}_{t+1}\pi^{11}_t+\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)a^{\mathcal{I}}_tL^{3i}_{t+1}\right\}\overline{Q}^j_t\\
&+\left\{-\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)b^{\mathcal{I}}_t\delta^i_t+\left(\tilde{\sigma}^{11}_t\pi^{12}_t+\tilde{\sigma}^{12}_t\pi^{22}_t\right)b^{\mathcal{I}}_t\right\}\overline{Q}^i_t\mathcal{I}_{t-1}\\
&+\left\{-\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)b^{\mathcal{I}}_t\mathrm{e}^{-\rho}L^{2i}_{t+1}\right\}R_t\mathcal{I}_{t-1}\\
&+\left\{-\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)b^{\mathcal{I}}_tL^{3i}_{t+1}\right\}\overline{Q}^j_t\mathcal{I}_{t-1}\\
&+\left\{-\frac{1}{2\gamma^i}\left\{\left(\widetilde{\sigma^{11}_t}\right)^2\pi^{11}_t+2\widetilde{\sigma^{11}_t}\widetilde{\sigma^{12}_t}\pi^{12}_t+\widetilde{\sigma^{12}_t}\widetilde{\sigma^{22}_t}\pi^{22}_t\right\}\left(b^{\mathcal{I}}_t\right)^2-\frac{\widetilde{\sigma^{11}_t}}{2\gamma^i}\right\}\mathcal{I}^2_{t-1}\\
&+\left\{-\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)b^{\mathcal{I}}_tL^{5i}_{t+1}+\frac{1}{\gamma^i}a^{\mathcal{I}}_tb^{\mathcal{I}}_t\left\{\left(\widetilde{\sigma^{11}_t}\right)^2\pi^{11}_t+2\widetilde{\sigma^{11}_t}\widetilde{\sigma^{12}_t}\pi^{12}_t+\widetilde{\sigma^{12}_t}\widetilde{\sigma^{22}_t}\pi^{22}_t\right\}+\frac{1}{\gamma^i}\tilde{\sigma}^{11}_ta^{\mathcal{I}}_tb^{\mathcal{I}}_t\right\}\mathcal{I}_{t-1}\\
&+Z^i_{t+1}+\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)a^{\mathcal{I}}_tL^{5i}_{t+1}-\frac{1}{2\gamma^i}a^{\mathcal{I}}_t\left\{\left(\widetilde{\sigma^{11}_t}\right)^2\pi^{11}_t+2\widetilde{\sigma^{11}_t}\widetilde{\sigma^{12}_t}\pi^{12}_t+\widetilde{\sigma^{12}_t}\widetilde{\sigma^{22}_t}\pi^{22}_t\right\}-\frac{1}{2\gamma^i}\widetilde{\sigma^{11}_t}\left(a^{\mathcal{I}}_t\right)^2+x^i_t\\
&+X^i_t(q^i_t)^2+\left[Y^{1i}_t\overline{Q}^i_t+Y^{2i}_tR_t+Y^{3i}_t\overline{Q}^j_t+Y^{4i}_t\mathcal{I}_{t-1}+Y^{5i}_{t-1}\right]q^j_t\Big]\Big\},
\end{aligned}
\tag{B.32}
$$

where $x^i_t := \dfrac{1}{2\gamma^i}\log\dfrac{|(\mathbf{\Sigma}^*)^{-1}|^{\frac{1}{2}}}{|\mathbf{\Sigma}|^{\frac{1}{2}}}$, and

$$
\begin{pmatrix}\widetilde{\sigma^{11}_t} & \widetilde{\sigma^{12}_t}\\ \widetilde{\sigma^{21}_t} & \widetilde{\sigma^{22}_t}\end{pmatrix}:=\begin{pmatrix}\left(\sigma^{\mathcal{I}}_{t+1}\right)^2 & \rho^{\mathcal{I},\epsilon}\sigma^{\mathcal{I}}_{t+1}\sigma^{\epsilon}_{t+1}\\ \rho^{\mathcal{I},\epsilon}\sigma^{\mathcal{I}}_{t+1}\sigma^{\epsilon}_t & (\sigma^{\epsilon}_{t+1})^2\end{pmatrix}^{-1},\quad \begin{pmatrix}\pi^{11}_t & \pi^{12}_t\\ \pi^{21}_t & \pi^{22}_t\end{pmatrix}:=\begin{pmatrix}\widetilde{\sigma^{11}_t}-2\gamma^iL^{4i}_{t+1} & \widetilde{\sigma^{12}_t}\\ \widetilde{\sigma^{21}_t} & \widetilde{\sigma^{22}_t}\end{pmatrix}^{-1},
\tag{B.33}
$$

and

$$
\begin{cases}
A^i_t := \left(1-\alpha^t\right)\lambda_t-G^{1i}_{t+1}+\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1}+\alpha^2_t\lambda^2_t\mathrm{e}^{-2\rho}H^{2i}_{t+1}+\frac{1}{2}\gamma^i\left(\theta^i_t\right)^2\pi^{11}_t+\frac{1}{2}\gamma^i\pi^{22}_t+\gamma^i\pi^{12}_t\theta^i_t;\\
B^i_t := -\alpha^t\lambda_t-2G^{1i}_{t+1}+\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1}-\gamma^i\theta^i_t\delta^i_t\pi^{11}_t+\gamma^i\pi^{22}_t+\gamma^i\pi^{12}_t\theta^i_t-\gamma^i\pi^{12}_t\delta^i_t;\\
C^i_t := -J^{1i}_{t+1}+\alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1}-\gamma^i\theta^i_tL^{3i}_{t+1}\pi^{11}_t-\gamma^i\pi^{12}_tL^{3i}_{t+1};\\
D^i_t := -(1-\mathrm{e}^{-\rho})-\mathrm{e}^{-\rho}H^{1i}_{t+1}+2\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{2i}_{t+1}-\gamma^i\theta^i_t\mathrm{e}^{-\rho}L^{2i}_{t+1}\pi-\gamma^i\pi^{12}_t\mathrm{e}^{-\rho}L^{2i}_{t+1};\\
F^i_t := -\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)b^{\mathcal{I}}_t\theta^i_t-\left(\tilde{\sigma}^{11}_t\pi^{12}_t+\tilde{\sigma}^{12}_t\pi^{22}_t\right)b^{\mathcal{I}}_t;\\
M^i_t := -G^{2i}_{t+1}+\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{3i}_{t+1}-\theta^i_tL^{5i}_{t+1}\pi^{11}_t-\gamma^i\pi^{12}_tL^{5i}_{t+1}+a^{\mathcal{I}}_t\theta^i_t\left(\widetilde{\sigma^{11}_t}\pi^{11}_t+\widetilde{\sigma^{12}_t}\pi^{21}_t\right)\\
\qquad +a^{\mathcal{I}}_t\left(\tilde{\sigma}^{11}_t\pi^{12}_t+\tilde{\sigma}^{12}_t\pi^{22}_t\right);\\
N^i_t := -\left(1-\alpha^t\right)-\alpha_t\lambda_t\mathrm{e}^{-\rho}H^{1i}_{t+1}+2\alpha^2_t\lambda^2_t\mathrm{e}^{-2\rho}H^{2i}_{t+1}+J^{1i}_{t+1}-\alpha_t\lambda_t\mathrm{e}^{-\rho}J^{2i}_{t+1}-\gamma^i\theta^i_t\phi^i_t\pi^{11}_t-\gamma^i\pi^{12}_t\phi^i_t,
\end{cases}
\tag{B.34}
$$

and

$$\begin{cases} X_t^i & := \alpha_t^2 \lambda_t^2 \mathrm{e}^{-\rho} H_{t+1}^{2i} - \alpha_t \lambda_t \mathrm{e}^{-\rho} J_{t+1}^{2i} + J_{t+1}^{3i} - \frac{1}{2}\gamma^i \left(\phi_t^i\right)^2 \pi_t^{11}; \\ Y_t^{1i} & := -\alpha^t \lambda_t + \alpha_t \lambda_t \mathrm{e}^{-\rho} H_{t+1}^{1i} - J_{t+1}^{1i} - \gamma^i \delta_t^i \phi_t^i \pi_t^{11} + \gamma^i \pi_t^{12} \phi_t^i; \\ Y_t^{2i} & := 2\alpha_t \lambda_t \mathrm{e}^{-2\rho} H_{t+1}^{2i} - \mathrm{e}^{-\rho} J_{t+1}^{2i} - \gamma^i \mathrm{e}^{-\rho} L_{t+1}^{2i} \phi_t^i \pi_t^{11}; \\ Y_t^{3i} & := \alpha_t \lambda_t \mathrm{e}^{-\rho} J_{t+1}^{2i} - 2J_{t+1}^{3i} - \gamma^i L_{t+1}^{3i} \phi_t^i \pi_t^{11}; \\ Y_t^{4i} & := -\left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) b_t^{\mathcal{I}} \phi_t^i; \\ Y_t^{5i} & := \alpha_t \lambda_t \mathrm{e}^{-\rho} H_{t+1}^{3i} - J_{t+1}^{4i} - \gamma^i L_{t+1}^{5i} \phi_t^i \pi_t^{11} + \left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) a_t^{\mathcal{I}} \phi_t^i. \end{cases}$$

Then, the best response of large trader $i \in \{1,2\}$ to the other large trader aat time t, denoted by $BR^i(q_t^j)$, becomes

$$BR^i(q_t^j) = \frac{1}{2A_t^i}\left(B_t^i \overline{Q}_t^i + C_t^i \overline{Q}_t^j + D_t^i R_t + F_t^i \mathcal{I}_{t-1} + M_t^i + N_t^i q_t^j\right). \tag{B.35}$$

Thus, at the Markov perfect equilibrium, we have

$$\begin{cases} q_t^{i*} = \dfrac{1}{2A_t^i}\left(B_t^i \overline{Q}_t^i + C_t^i \overline{Q}_t^j + D_t^i R_t + F_t^i \mathcal{I}_{t-1} + M_t^i + N_t^i q_t^{j*}\right); \\ q_t^{j*} = \dfrac{1}{2A_t^j}\left(B_t^j \overline{Q}_t^j + C_t^j \overline{Q}_t^i + D_t^j R_t + F_t^j \mathcal{I}_{t-1} + M_t^j + N_t^j q_t^{i*}\right). \end{cases} \tag{B.36}$$

Solving the above simultaneous equations results in

$$\begin{aligned} q_t^{i*} &:= B_t^{i*} \overline{Q}_t^i + C_t^{i*} \overline{Q}_t^j + D_t^{i*} R_t + F_t^{i*} \mathcal{I}_{t-1} + M_t^{i*} \\ &\quad (=: a_t^i + b_t^i \overline{Q}_t^i + c_t^i \overline{Q}_t^j + d_t^i R_t + e_t^i \mathcal{I}_{t-1}), \end{aligned} \tag{B.37}$$

where

$$\begin{cases} \zeta_t^i := 2A_t^i - \dfrac{N_t^i N_t^j}{2A_t^j}; \\ B_t^{i*} := \dfrac{1}{\zeta_t^i}\left(B_t^i + \dfrac{N_t^i C_t^j}{2A_t^j}\right); \quad C_t^{i*} := \dfrac{1}{\zeta_t^i}\left(C_t^i + \dfrac{N_t^i B_t^j}{2A_t^j}\right); \\ D_t^{i*} := \dfrac{1}{\zeta_t^i}\left(D_t^i + \dfrac{N_t^i D_t^j}{2A_t^j}\right); \quad F_t^{i*} := \dfrac{1}{\zeta_t^i}\left(F_t^i + \dfrac{N_t^i F_t^j}{2A_t^j}\right); \\ M_t^{i*} := \dfrac{1}{\zeta_t^i}\left(M_t^i + \dfrac{N_t^i M_t^j}{2A_t^j}\right). \end{cases} \tag{B.38}$$

for each $i \in \{1,2\}$. q_t^{1*} and q_t^{2*} are equilibrium execution volumes at the Markov perfect equilibrium for time t. Then the value function for large trader $i \in \{1,2\}$ at the Markov Perfect equilibrium $(\pi^{1*}, \pi^{2*}) \in \Pi_M^1 \times \Pi_M^2$ becomes

$$\begin{aligned} & V_t^i(\pi_t^{1*}, \pi_t^{2*})\left[W_t^1, W_t^2, P_t, \overline{Q}_t^1, \overline{Q}_t^2, R_t, \mathcal{I}_{t-1}\right] \\ &= -\exp\Big\{-\gamma\Big[W_t^i - P_t \overline{Q}_t^i + G_t^{1i}\left(\overline{Q}_t^i\right)^2 + G_t^{2i} \overline{Q}_t^i + H_t^{1i} \overline{Q}_t R_t \\ &\quad + H_t^{2i} R_t^2 + H_t^{3i} R_t + J_t^{1i} \overline{Q}_t^i \overline{Q}_t^j + J_t^{2i} \overline{Q}_t^j R_t + J_t^{3i}\left(\overline{Q}_t^j\right)^2 + J_t^{4i} \overline{Q}_t^j \\ &\quad + L_t^{1i} \overline{Q}_t^i \mathcal{I}_{t-1} + L_t^{2i} R_t \mathcal{I}_{t-1} + L_t^{3i} \overline{Q}_t^j \mathcal{I}_{t-1} + L_t^{4i} \mathcal{I}_{t-1}^2 + L_t^{5i} \mathcal{I}_{t-1} + Z_t^i\Big]\Big\}, \end{aligned} \tag{B.39}$$

where

$$
\begin{aligned}
G_t^{1i} &:= G_{t+1}^{1i} - \frac{1}{2}\gamma^i \delta_t^i \pi_t^i - \frac{1}{2}\gamma^i \pi_t^{22} + \gamma^i \pi_t^{12} \delta_t^i + X_t^i \left(C_t^{j*}\right)^2 + Y_t^{1i} C_t^{j*} + \frac{\left(B_t^{i**}\right)^2}{4A_t^i};\\
G_t^{2i} &:= G_{t+1}^{2i} - \gamma^i \delta_t^i L_{t+1}^{5i} \pi_t^{11} + \gamma^i \pi_t^{12} L_{t+1}^{5i} + \left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) a_t^{\mathcal{I}} \delta_t^i - \left(\widetilde{\sigma_t^{11}} \pi_t^{12} + \widetilde{\sigma_t^{12}} \pi_t^{22}\right) a_t^{\mathcal{I}}\\
&\quad + 2X_t^i C_t^{j*} M_t^{j*} + Y_t^{5i} C_t^{j*} + Y_t^{1i} M_t^{j*} + \frac{B_t^{i**} M_t^{i**}}{2A_t^i};\\
H_t^{1i} &:= \left(1 - \mathrm{e}^{-\rho}\right) + \mathrm{e}^{-\rho} H_{t+1}^{1i} - \gamma^i \delta_t^i \mathrm{e}^{-\rho} L_{t+1}^{2i} \pi_t^{11} + \gamma^i \pi_t^{12} \mathrm{e}^{-\rho} L_{t+1}^{2i} + 2X_t^i C_t^{j*} D_t^{j*}\\
&\quad + Y_t^{2i} C_t^{j*} + Y_t^{1i} D_t^{j*} + \frac{B_t^{i**} D_t^{i**}}{2A_t^i};\\
H_t^{2i} &:= \mathrm{e}^{-2\rho} H_{t+1}^{2i} - \frac{1}{2}\gamma^i \mathrm{e}^{-2\rho} \left(L_{t+1}^{2i}\right)^2 \pi_t^{11} + X_t^i \left(D_t^{j*}\right)^2 + Y_t^{2i} D_t^{j*} + \frac{\left(D_t^{i**}\right)^2}{4A_t^i};\\
H_t^{3i} &:= \mathrm{e}^{-\rho} H_{t+1}^{3i} - \gamma^i \mathrm{e}^{-\rho} L_{t+1}^{2i} L_{t+1}^{5i} \pi_t^{11} + \left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) a_t^{\mathcal{I}} \mathrm{e}^{-\rho} L_{t+1}^{2i}\\
&\quad + 2X_t^i D_t^{j*} M_t^{j*} + Y_t^{5i} D_t^{j*} + Y_t^{2i} M_t^{j*} + \frac{D_t^{i**} M_t^{i**}}{2A_t^i};\\
J_t^{1i} &:= J_{t+1}^{1i} - \gamma^i \delta_t^i L_{t+1}^{3i} \pi_t^{11} + \gamma^i \pi_t^{12} L_{t+1}^{3i} + 2X_t^i B_t^{j*} C_t^{j*} + Y_t^{1i} B_t^{j*} + Y_t^{3i} C_t^{j*} + \frac{B_t^{i**} C_t^{i**}}{2A_t^i};\\
J_t^{2i} &:= \mathrm{e}^{-2\rho} J_{t+1}^{2i} - \gamma^i \mathrm{e}^{-\rho} L_{t+1}^{2i} L_{T+1}^{3i} \pi_t^{11} + 2X_t^i B_t^{j*} D_t^{j*} + Y_t^{2i} B_t^{j*} + Y_t^{3i} D_t^{j*} + \frac{C_t^{i**} D_t^{i**}}{2A_t^i};\\
J_t^{3i} &:= J_{t+1}^{3i} - \frac{1}{2}\gamma^i \left(L_{t+1}^{3i}\right)^2 \pi_t^{11} + X_t^i \left(B_t^{j*}\right)^2 + Y_t^{3i} B_t^{j*} + \frac{\left(C_t^{i**}\right)^2}{4A_t^i};\\
J_t^{4i} &:= J_{t+1}^{4i} - \gamma^i L_{t+1}^{3i} L_{t+1}^{5i} \pi_t^{11} + \left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) a_t^{\mathcal{I}} L_{t+1}^{3i}\\
&\quad + 2X_t^i B_t^{j*} M_t^{j*} + Y_t^{5i} B_t^{j*} + Y_t^{3i} M_t^{j*} + \frac{C_t^{i**} M_t^{i**}}{2A_t^i};\\
L_t^{1i} &:= -\left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) b_t^{\mathcal{I}} \delta_t^i + \left(\tilde{\sigma}_t^{11} \pi_t^{12} + \tilde{\sigma}_t^{12} \pi_t^{22}\right) b_t^{\mathcal{I}} + 2X_t^i C_t^{j*} F_t^{j*} + Y_t^{4i} C_t^{j*} + Y_t^{1i} F_t^{j*}\\
&\quad + \frac{B_t^{i**} F_t^{i**}}{2A_t^i};\\
L_t^{i2} &:= -\left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) b_t^{\mathcal{I}} \mathrm{e}^{-\rho} L_{t+1}^{2i} + 2X_t^i D_t^{j*} F_t^{j*} + Y_t^{4i} D_t^{j*} + Y_t^{2i} F_t^{j*} + \frac{D_t^{i**} F_t^{i**}}{2A_t^i};\\
L_t^{3i} &:= -\left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) b_t^{\mathcal{I}} L_{t+1}^{3i} + 2X_t^i B_t^{j*} F_t^{j*} + Y_t^{4i} B_t^{j*} + Y_t^{3i} F_t^{j*} + \frac{C_t^{i**} F_t^{i**}}{2A_t^i};\\
L_t^{4i} &:= -\frac{1}{2\gamma^i} \left\{ \left(\widetilde{\sigma_t^{11}}\right)^2 \pi_t^{11} + 2\widetilde{\sigma_t^{11}} \widetilde{\sigma_t^{12}} \pi_t^{12} + \widetilde{\sigma_t^{12}} \widetilde{\sigma_t^{22}} \pi_t^{22} \right\} \left(b_t^{\mathcal{I}}\right)^2 - \frac{\widetilde{\sigma_t^{11}}}{2\gamma^i} + X_t^i \left(F_t^{j*}\right)^2 + Y_t^{4i} F_t^{j*} + \frac{\left(F_t^{i**}\right)^2}{4A_t^i};\\
L_t^{5i} &:= -\left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) b_t^{\mathcal{I}} L_{t+1}^{5i} + \frac{1}{\gamma^i} a_t^{\mathcal{I}} b_t^{\mathcal{I}} \left\{ \left(\widetilde{\sigma_t^{11}}\right)^2 \pi_t^{11} + 2\widetilde{\sigma_t^{11}} \widetilde{\sigma_t^{12}} \pi_t^{12} + \widetilde{\sigma_t^{12}} \widetilde{\sigma_t^{22}} \pi_t^{22} \right\} + \frac{1}{\gamma^i} \tilde{\sigma}_t^{11} a_t^{\mathcal{I}} b_t^{\mathcal{I}}\\
&\quad + 2X_t^i F_t^{j*} M_t^{j*} + Y_t^{5i} F_t^{j*} + Y_t^{4i} M_t^{j*} + \frac{F_t^{i**} M_t^{i**}}{2A_t^i};\\
Z_t^i &:= +Z_{t+1}^i + \left(\widetilde{\sigma_t^{11}} \pi_t^{11} + \widetilde{\sigma_t^{12}} \pi_t^{21}\right) a_t^{\mathcal{I}} L_{t+1}^{5i} - \frac{1}{2\gamma^i} a_t^{\mathcal{I}} \left\{ \left(\widetilde{\sigma_t^{11}}\right)^2 \pi_t^{11} + 2\widetilde{\sigma_t^{11}} \widetilde{\sigma_t^{12}} \pi_t^{12} + \widetilde{\sigma_t^{12}} \widetilde{\sigma_t^{22}} \pi_t^{22} \right\}\\
&\quad - \frac{1}{2\gamma^i} \widetilde{\sigma_t^{11}} \left(a_t^{\mathcal{I}}\right)^2 + x_t^i + X_t^i \left(M_t^{j*}\right)^2 + \frac{\left(M_t^{i**}\right)^2}{4A_t^i}. \qquad \text{(B.40)}
\end{aligned}
$$

□

C In the case with target close orders

In this subsection, we consider an execution model with a closing price. The time framework $t \in \{ 1, \dots, T, T+1 \}$ is the same in the model mentioned above. However, we add an assumption that a large trader can execute his/her remaining execution volume at time $T+1$, $\overline{Q}^i_{T+1}$, with closing price P_{T+1}. We further assume that the trading at time $T+1$ impose the large trader to pay the additive cost $\chi_{T+1} \in \mathbb{R}$ per unit of the remaining volume. As stated in the last section, we have the following theorem.

Theorem C.1 (Equilibrium Execution Strategy and the Value Function in the Case with Target Close Orders)**.** There exists a Markov perfect equilibrium at which the following properties hold for each large trader $i \in \{ 1, 2 \}$:

1. The execution volume at the Markov perfect equilibrium for the large trader $i \in \{ 1, 2 \}$ at time $t \in \{ 1, \dots, T \}$, denoted as $q^{i*'}_t$, becomes an *affine function* of the Markovian environment at time $t-1$, $\mathcal{I}_{t-1}$, the remaining execution volume of each large trader, $\overline{Q}^i_t$ and $\overline{Q}^j_t$ $(i \neq j, i, j \in \{ 1, 2 \})$, and the cumulative residual effect, R_t, that is,

$$q^{i*'}_t = a^{i*}_t + b^{i*}_t \overline{Q}^i_t + c^{i*}_t \overline{Q}^j_t + d^{i*}_t R_t + e^{i*}_t \mathcal{I}_{t-1}, \quad t = 1, \dots, T. \tag{C.1}$$

2. The value function $V^i_t(\pi^1, \pi^2)\big[\boldsymbol{s}_t\big]$ at time $t \in \{ 1, \dots, T, T+1 \}$ for each large trader $i \in \{ 1, 2 \}$ is represented as a functional form as follows:

$$\begin{aligned} &V^i_t(\pi^1, \pi^2)\big[W^1_t, W^2_t, P_t, \overline{Q}^1_t, \overline{Q}^2_t, R_t, \mathcal{I}_{t-1}\big] \\ &= -\exp\Big\{ -\gamma\Big[W^i_t - P^\top_t \overline{Q}^i_t + G^{1i*}_t \overline{Q}_t + G^{2i*}_t \left(\overline{Q}^i_t\right)^2 + H^{1i*}_t \overline{Q}_t R_t + H^{2i*}_t R^2_t + H^{3i*}_t R_t \\ &\quad + I^{1i*}_t \overline{Q}^i_t \overline{Q}^j_t + I^{2i*}_t \overline{Q}^j_t R_t + I^{3i}_T \left(\overline{Q}^j_t\right)^2 + I^{4i*}_t \overline{Q}^j_t \\ &\quad + J^{1i*}_t \overline{Q}^i_t \mathcal{I}_{t-1} + J^{2i*}_t R_t \mathcal{I}_{t-1} + J^{3i*}_t \overline{Q}^j_t \mathcal{I}_{t-1} + J^{4i*}_t \mathcal{I}^2_{t-1} + J^{5i*}_t \mathcal{I}_{t-1} + Z^{i*}_t\Big]\Big\}, \end{aligned} \tag{C.2}$$

where $G^{1i*}_t, G^{2i*}_t, H^{1i*}_t, H^{2i*}_t, H^{3i*}_t, I^{1i*}_t, I^{2*i}_t, I^{3i*}_t, I^{4i*}_t, J^{1i*}_t, J^{2i*}_t, J^{3i*}_t, J^{4i*}_t, J^{5i*}_t, Z^{i*}_t$ for $t \in \{ 1, \dots, T, T+1 \}$ are deterministic functions of time t which are dependent on the problem parameters, and can be computed backwardly in time t from maturity T.

Proof. Omitted. □