

6-29-2017

An Optimal Execution Problem with S-shaped Market Impact Functions

Takashi Kato

Association of Mathematical Finance Laboratory, takashi.kato@mathfi-lab.com

Follow this and additional works at: <https://repository.lsu.edu/cosa>



Part of the [Analysis Commons](#), and the [Other Mathematics Commons](#)

Recommended Citation

Kato, Takashi (2017) "An Optimal Execution Problem with S-shaped Market Impact Functions," *Communications on Stochastic Analysis*: Vol. 11: No. 3, Article 1.

DOI: 10.31390/cosa.11.3.01

Available at: <https://repository.lsu.edu/cosa/vol11/iss3/1>

An Optimal Execution Problem with S-shaped Market Impact Functions

Cover Page Footnote

AN OPTIMAL EXECUTION PROBLEM WITH S-SHAPED MARKET IMPACT FUNCTIONS

TAKASHI KATO

ABSTRACT. In this study, we extend the optimal execution problem with convex market impact function studied in Kato [14] to the case where the market impact function is S-shaped, that is, concave on $[0, \bar{x}_0]$ and convex on $[\bar{x}_0, \infty)$ for some $\bar{x}_0 \geq 0$. We study the corresponding Hamilton–Jacobi–Bellman equation and show that the optimal execution speed under the S-shaped market impact is equal to zero or larger than \bar{x}_0 . Moreover, we provide some examples of the Black–Scholes model. We show that the optimal strategy for a risk-neutral trader with small shares is the time-weighted average price strategy whenever the market impact function is S-shaped.

1. Introduction

Optimal execution problems have been widely investigated in mathematical finance as a type of stochastic control problem. There are various studies of optimal execution, such as [1, 2, 4, 7, 23] and references therein, and Gatheral and Schied [8] survey several dynamic models of optimal execution. To study this type of problem, we cannot ignore market impact (MI), which is a market liquidity problem. Here, we consider a situation where a single trader has many shares of a security and tries to sell (liquidate) it until a time horizon. A large selling order induces a gap between supply and demand, causing a decrease in the security price. This effect is called the MI, and the trader should reduce the liquidation speed to avoid the MI cost. However, reducing the liquidation speed also increases the timing cost, which is caused by the random fluctuation of the security price over time. The trader should optimize the execution strategy by considering the MI cost and the timing cost. Therefore, the MI function, $g(x)$, plays an important role in studying optimal execution problems. Here, $g(x)$ implies the decrease of the security price by selling x shares (or selling rate).

The simplest setting for g is a linear function. For instance, in [2, 4, 23], optimal execution problems are treated mainly with linear MI functions and derive optimal execution strategies. However, there are studies on optimization problems with non-linear g [1, 10, 11, 12, 14, 15]. In particular, we derive a mathematically adequate continuous-time model of an optimal execution problem as a limit of discrete-time optimization problems in [14] when g is strictly convex.

Received 2017-6-29; Communicated by Hui-Hsiung Kuo.

2010 *Mathematics Subject Classification.* Primary 91G80; Secondary 93E20, 49L25.

Key words and phrases. Optimal execution problem, market impact, the Hamilton–Jacobi–Bellman equation, time-weighted average price (TWAP) strategy.

It is still unclear what form of g is natural. Recently, it has been proposed that an S-shaped function is suitable for g . That is, $g(x)$ should be concave on $[0, \bar{x}_0]$ and convex on $[\bar{x}_0, \infty)$ for some $\bar{x}_0 \geq 0$. Many traders intuitively expect that MI functions are S-shaped [15]. Moreover, in [24], we find an empirical prediction for a hump-shaped limit order book, which corresponds to the S-shaped MI function. Therefore, our previous study [14] should be extended to include the S-shaped MI function, g . We have tackled this problem partially in [15], but there are still many mathematical and financial questions at this stage. For instance, when we consider the optimal execution problem with S-shaped g , it is intuitive that the optimal execution speed should not reach the range $(0, \bar{x}_0]$. However, we have not proved this finding mathematically. Moreover, we have not discussed the Hamilton–Jacobi–Bellman (HJB) equations corresponding to our optimization problem sufficiently.

In this paper, we resolve these questions as a continuation of our previous study [15]. We completely generalize our previous results [14] to the case of S-shaped g and study the above questions. Moreover, we find that the optimal execution strategy of a risk-neutral trader in the Black–Scholes market model is the time-weighted average price (TWAP) strategy, that is, to sell at a constant speed. This is the same result as Theorem 5.4(ii) in [14] when g is a quadratic function; however, it is not necessary to assume an explicit form of g . We show that this result is true whenever g is S-shaped. This result generalizes Theorem 5.4(ii) in [14] and provides an analytical solution to the optimal execution problem with an uncertain MI given in Section 5.2 of [12].

The rest of this paper is as follows. In Section 2, we introduce a mathematical model of an optimization problem based on our previous work [15] and review the previous results. In Section 3, we characterize our value function as a viscosity solution to the corresponding HJB equation. We show the uniqueness of the viscosity solutions to the HJB equation under adequate conditions. To investigate the properties of optimal strategies, we introduce a verification theorem and show that the optimal execution strategy does not take the value in $(0, \bar{x}_0]$ in Section 4. In Section 5, we present some examples. In particular, we demonstrate the robustness of the TWAP strategy as an optimal strategy in the Black–Scholes model with general shaped MI functions. We summarize our argument and introduce future tasks in Section 6. Section A gives supplemental arguments to guarantee consistency between our present model and our previous model [15]. All proofs are in Section B.

2. Model Settings

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ be a stochastic basis and let $(B_t)_{0 \leq t \leq T}$ be a one-dimensional Brownian motion ($T > 0$). Set $D = \mathbb{R} \times [0, \infty)^2$ and denote by \mathcal{C} the set of non-decreasing, non-negative, and continuous functions with polynomial growth defined on D . For $u \in \mathcal{C}$, we define a function, $J(\cdot; u) : [0, T] \times D \rightarrow \mathbb{R}$, as

$$J(t, c, x, s; u) = \sup_{(x_r)_{r \in \mathcal{A}_t(x)}} \mathbb{E}[u(C_t, X_t, S_t)], \quad (2.1)$$

where $(C_r)_r$, $(X_r)_r$, and $(S_r)_r$ are stochastic processes given by

$$\begin{aligned} dC_r &= x_r S_r dr, \\ dX_r &= -x_r dr, \\ dS_r &= \hat{b}(S_r) dr + \hat{\sigma}(S_r) dB_r - S_r g(x_r) dr, \end{aligned} \quad (2.2)$$

and $(C_0, X_0, S_0) = (c, x, s)$, and $\mathcal{A}_t(x)$ is the set of non-negative $(\mathcal{F}_r)_r$ -progressively measurable process, $(x_r)_{0 \leq r \leq t}$, satisfying $\int_0^t x_r dr \leq x$ a.s. We call an element of $\mathcal{A}_t(x)$ an admissible strategy. Here, \hat{b} and $\hat{\sigma}$ are defined as

$$\hat{b}(s) = \left(b(\log s) + \frac{1}{2} \sigma(\log s)^2 \right) s, \quad \hat{\sigma}(s) = \sigma(\log s) s, \quad s > 0$$

and $\hat{b}(0) = \hat{\sigma}(0) = 0$, where $b, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are bounded and Lipschitz continuous functions. $g \in C([0, \infty)) \cap C^1((0, \infty))$ is a non-negative function with $g(0) = 0$.

Function J implies the value function of an optimal execution problem with MI function g and is derived as a limit of discrete-time value functions in [14] when g is convex, and in [15] when g is S-shaped, that is, when $h := g'$ satisfies the following conditions.

- [A1] $h(x) \geq 0, x > 0$.
- [A2] $\lim_{x \rightarrow 0} x h(x) = 0$.
- [A3] There is an $\bar{x}_0 \geq 0$ such that h is strictly decreasing on $(0, \bar{x}_0]$ and strictly increasing on $[\bar{x}_0, \infty)$.
- [A4] $h(\infty) = \lim_{x \rightarrow \infty} h(x) = \infty$.

Condition [A3] implies that g is concave on $[0, \bar{x}_0]$ and convex on $[\bar{x}_0, \infty)$. In this paper, we always assume [A1]–[A4].

We briefly introduce the financial implications of our model (see [11, 12, 14] for more details). We assume that there is a single trader who has many shares x_0 of a security whose price is s_0 at the initial time. The trader tries to sell the security in the market until time horizon T , but the selling behavior affects the security price via the effect of MI (denoted as the term $-g(x_r) dr$ in (2.3)). S_r is the security price at time r , and C_r (respectively, X_r) describes the cash amount (respectively, shares of the security) held at time r . The trader's purpose is to maximize the terminal expected utility, $J(T, c_0, x_0, s_0; u) = \mathbb{E}[u(C_T, X_T, S_T)]$, by controlling an execution strategy, $(x_r)_r \in \mathcal{A}_T(x)$. Here, x_r implies the liquidation speed at time r ; in other words, the trader sells $x_r dr$ amount in the infinitesimal time interval $[r, r + dr]$. To solve this problem, we introduce the value function $J(t, c, x, s; u)$ for each t, c, x , and s to apply the dynamic programming method.

Remark 2.1.

- (i) The log-price process $Y_r = \log S_r$ satisfies the stochastic differential equation (SDE),

$$dY_r = b(Y_r) dr + \sigma(Y_r) dB_r - g(x_r) dr, \quad (2.3)$$

whenever $S_r > 0, r \geq 0$.

- (ii) In stochastic control theory, $(x_r)_r \in \mathcal{A}_t(x)$ is called a control process and $(C_r, X_r, S_r)_r$ defined in (2.2) is its controlled process. However, the existence and uniqueness of $(C_r, X_r, S_r)_r$ for each $(x_r)_r$ is not obvious in

our case, because $(Y_r)_r$ may diverge due to the term $-g(x_r)dr$. We can overcome this difficulty by regarding $S_r = 0$ after Y_r diverges to $-\infty$ (see Section A for details).

- (iii) In [11, 14, 15], we require an additional assumption such that each admissible strategy is essentially bounded; that is, we consider the optimization problem

$$J^\infty(t, c, x, s; u) = \sup_{(x_r)_r \in \mathcal{A}_t^\infty(x)} \mathbb{E}[u(C_t, X_t, S_t)] \tag{2.4}$$

instead of (2.1), where

$$\mathcal{A}_t^\infty(x) = \left\{ (x_r)_r \in \mathcal{A}_t(x) ; \operatorname{esssup}_{r, \omega} x_r(\omega) < \infty \right\}.$$

This condition arises in the process of taking the limit from the discrete-time model to the continuous-time model; however, it is a mathematical technical condition and is unnatural in relation to finance. We can show that J coincides with J^∞ , and thus we are not overly concerned about this problem (also see Section A).

In [15], we show that $J(\cdot ; u)$ is continuous on $[0, T] \times D$ and $J(r, \cdot ; u) \in \mathcal{C}$ for each $r \geq 0$ and $u \in \mathcal{C}$. Moreover, J satisfies the dynamic programming principle,

$$J(t + r, c, x, s; u) = J(t, c, x, s; J(r, \cdot ; u)),$$

for each $(c, x, s) \in D$, $u \in \mathcal{C}$ and $t, r \geq 0$ with $t + r \leq T$.

By using these results, we characterize J as a viscosity solution of the corresponding HJB equation in the next section. From now on, we fix $u \in \mathcal{C}$ and denote $J(t, c, x, s; u) = J(t, c, x, s)$ for brevity.

3. Main Results I: Viscosity Properties

Our first main result is as follows.

Theorem 3.1.

- (i) *We assume that*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (J(t, c, x, s + \varepsilon) - J(t, c, x, s)) > 0, \quad (t, c, x, s) \in (0, T] \times \tilde{D}, \tag{3.1}$$

where $\tilde{D} = \operatorname{int}D = \mathbb{R} \times (0, \infty)^2$. Then J is a viscosity solution of the following HJB equation on $(0, T] \times \tilde{D}$:

$$\frac{\partial}{\partial t} J - \sup_{y \geq 0} \mathcal{L}^y J = 0, \tag{3.2}$$

where

$$\mathcal{L}^y = (\hat{b}(s) - sg(y)) \frac{\partial}{\partial s} + \frac{1}{2} \hat{\sigma}(s)^2 \frac{\partial^2}{\partial s^2} + y \left(s \frac{\partial}{\partial c} - \frac{\partial}{\partial x} \right).$$

- (ii) *We assume that equation (3.1) holds, \hat{b} and $\hat{\sigma}$ are Lipschitz continuous, and $\liminf_{x \rightarrow \infty} h(x)/x > 0$. Then we see the uniqueness of a viscosity solution of (3.2) in the following sense. If a continuous function $v : [0, T] \times D \rightarrow \mathbb{R}$*

with polynomial growth is a viscosity solution of (3.2) and satisfies the boundary conditions,

$$v(0, c, x, s) = u(c, x, s), \quad (3.3)$$

$$v(t, c, 0, s) = \mathbf{E}[u(c, 0, Z_t(s))], \quad (3.4)$$

$$v(t, c, x, 0) = u(c, x, 0), \quad (3.5)$$

then it holds that $J = v$, where $(Z_r(s))_r$ is a unique solution to the SDE:

$$dZ_r(s) = \hat{b}(Z_r(s))dr + \hat{\sigma}(Z_r(s))dB_r, \quad Z_0(s) = s. \quad (3.6)$$

Remark 3.2.

- (i) The assertions of Theorem 3.1 are the same as those of Theorems 3.3 and 3.6 in [14]. Thus, Theorem 3.1 was already obtained when g is convex, namely, when $\bar{x}_0 = 0$. As mentioned in Remark 3.7 of [14], our HJB equation (3.2) does not satisfy standard assumptions to apply a standard argument to viscosity characterization discussed in, for instance, [5, 6, 18, 21]. We demonstrate Theorem 3.1(i) through a refinement of the proof of Theorem 3.3 in [14]. In the proof of Theorem 3.6 in [14], we do not use the convexity of g mainly, so Theorem 3.1(ii) is obtained in a similar way to the proof of Propositions B.21–B.23 in [14].

- (ii) The following condition is a standard natural condition for a utility function in mathematical finance:

$$[\text{B}] \quad u(c, x, s) = U(c) \text{ for some concave function } U \in C^1(\mathbb{R}).$$

Under [B], the boundary conditions (3.3)–(3.5) are simplified as

$$v(0, c, x, s) = v(t, c, 0, s) = v(t, c, x, 0) = U(c).$$

- (iii) It is not easy to check (3.1) in general. When g is convex, the natural and simple sufficient conditions of (3.1) are introduced in [14] as
 [C1] u satisfies [B]. Moreover, it holds that $U'(c) \geq \delta$, $c \in \mathbb{R}$, for some $\delta > 0$.

[C2] b and σ are differentiable and their derivatives are Lipschitz continuous and uniformly bounded.

In our case, by the same proof as for Proposition 3.5 in [14], we also verify that (3.1) holds under [C1]–[C2].

4. Main Results II: Verification Arguments

Theorem 7.4 in [15] gives us a typical example where an optimal execution strategy takes the value zero or larger than \bar{x}_0 . This result is consistent with financial intuition, such as selling with the speed in the range of the concave part of g (i.e., $(0, \bar{x}_0]$) induces superfluous transaction cost. In this section, we present a verification theorem to demonstrate that the optimal execution speed is in $\{0\} \cup (\bar{x}_0, \infty)$ in general.

First, we introduce notation to state our second main result. Conditions [A3] and [A4] show that there is an inverse function, $h^{-1} : [h(\bar{x}_0), \infty) \rightarrow [\bar{x}_0, \infty)$, of

h. Then we define

$$\mathcal{H}(s, p) = \frac{sp_c - p_x}{sp_s} 1_{\{sp_s > 0\}}, \quad (4.1)$$

$$\Xi(s, p) = h^{-1}(\mathcal{H}(s, p)) 1_\Lambda(s, p) \quad (4.2)$$

for $s \geq 0$ and $p = (p_c, p_x, p_s)' \in \mathbb{R}^3$, where

$$\Lambda = \left\{ (s, p) \in (0, \infty) \times \mathbb{R}^3 \ ; \ p_s > 0, \mathcal{H}(s, p) > h(\bar{x}_0), \right. \\ \left. g(h^{-1}(\mathcal{H}(s, p))) < \mathcal{H}(s, p) h^{-1}(\mathcal{H}(s, p)) \right\},$$

and A' denotes the transpose of A . Moreover, for each continuously differentiable function, $v : [0, T] \times D \rightarrow \mathbb{R}$, we define

$$\bar{b}^v(t, c, x, s) = \begin{pmatrix} s\Xi(s, \mathcal{D}v(T-t, c, x, s)) \\ -\Xi(s, \mathcal{D}v(T-t, c, x, s)) \\ \hat{b}(s) - sg(\Xi(s, \mathcal{D}v(T-t, c, x, s))) \end{pmatrix},$$

where $\mathcal{D} = \left(\frac{\partial}{\partial c}, \frac{\partial}{\partial x}, \frac{\partial}{\partial s} \right)'$.

Now, we present our second main result.

Theorem 4.1. *We assume (3.1) and [B], that $J \in C^{1,1,1,2}((0, T] \times D)$, and that for given $(c_0, x_0, s_0) \in \tilde{D}$, there is a continuous process, $(\bar{C}_t, \bar{X}_t, \bar{S}_t)_t$, which satisfies*

$$d \begin{pmatrix} \bar{C}_t \\ \bar{X}_t \\ \bar{S}_t \end{pmatrix} = \bar{b}^J(t, \bar{C}_t, \bar{X}_t, \bar{S}_t) dt + \begin{pmatrix} 0 \\ 0 \\ \hat{\sigma}(\bar{S}_t) \end{pmatrix} dB_t, \quad t \in [0, \bar{\tau}] \quad (4.3)$$

and $(\bar{C}_0, \bar{X}_0, \bar{S}_0) = (c_0, x_0, s_0)$, where

$$\bar{\tau} = \inf\{t \geq 0 \ ; \ (\bar{C}_t, \bar{X}_t, \bar{S}_t) \in \partial D\} \wedge T.$$

Then there is an optimizer, $(\hat{x}_t)_t$ to $J(T, c_0, x_0, s_0)$, such that $\hat{x}_t \in \{0\} \cup (\bar{x}_0, \infty)$, $t \in [0, T]$ a.s.

When executing a large amount of the security, it is important to decrease the execution speed to reduce the execution cost. However, Theorem 4.1 tells us that when the MI function is S-shaped (especially, concave on $[0, \bar{x}_0]$), it is undesirable to decrease the execution speed beyond the threshold, \bar{x}_0 . An optimal execution strategy in this case is to sell with the execution speed greater than \bar{x}_0 or to stop selling.

We can apply Theorem 4.1 if we verify the smoothness of the value function J . Even if we find a classical (sub)solution of (3.2), which does not necessarily satisfy the boundary conditions (3.3)–(3.5), we can construct an optimal strategy to $J(T, c_0, x_0, s_0)$. We introduce the following verification theorem.

Theorem 4.2. *Let $(c_0, x_0, s_0) \in \tilde{D}$ and let $v \in C([0, T] \times D) \cap C^{1,1,1,2}((0, T] \times \tilde{D})$ be a function that satisfies the following conditions.*

(i) *There are $K, m > 0$ such that*

$$|v(t, c, x, s)| \leq K(1 + c^m + x^m + s^m), \quad t \in [0, T], \quad (c, x, s) \in D.$$

(ii) *$v(0, c, x, s) \geq u(c, x, s)$ holds for each $(c, x, s) \in D$.*

(iii) *$\frac{\partial}{\partial t} v - \sup_{y \geq 0} \mathcal{L}^y v \geq 0$ on $(0, T] \times \tilde{D}$.*

- (iv) *There is an $(\hat{x}_t)_t \in \mathcal{A}_T(x_0)$ such that $E[u(\hat{C}_T, \hat{X}_T, \hat{S}_T)] \geq v(T, c_0, x_0, s_0)$, where $(\hat{C}_t, \hat{X}_t, \hat{S}_t)_t$ is given by (2.2) with $(\hat{C}_0, \hat{X}_0, \hat{S}_0) = (c_0, x_0, s_0)$.*

Then we have $J(T, c_0, x_0, s_0) = E[u(\hat{C}_T, \hat{X}_T, \hat{S}_T)] = v(T, c_0, x_0, s_0)$ and $(\hat{x}_t)_t$ is its optimizer.

In the next section, we introduce some examples in which we derive optimal execution strategies by using Theorem 4.2.

5. Examples

Similar to Section 5 in [14], we introduce some examples where the security price process is given as the Black–Scholes model. We assume that $b(\cdot) \equiv \mu$ and $\sigma(\cdot) \equiv \sigma$ are constants and the utility function is set as $u_{\text{RN}}(c, x, s) = c$; that is, the trader is risk-neutral.

By using the same argument as the proof of Proposition 5.2 in [14], we have the following theorem.

Theorem 5.1. *We have $J(t, c, x, s) = c + sW(t, x)$, where*

$$\begin{aligned} W(t, x) &= \sup_{(x_r)_r \in \mathcal{A}_t^{\text{stat}}(x)} \int_0^t \exp\left(-\tilde{\mu}r - \int_0^r g(x_v)dv\right) x_r dr, \\ \mathcal{A}_t^{\text{stat}}(x) &= \{(x_r)_r \in \mathcal{A}_t(x) ; (x_r)_r \text{ is deterministic}\}, \\ \tilde{\mu} &= -\mu - \frac{1}{2}\sigma^2. \end{aligned}$$

Theorems 3.1 and 5.1 lead us to

Theorem 5.2. *W is a viscosity solution to the partial differential equation*

$$\frac{\partial}{\partial t}W + \tilde{\mu}W + \inf_{y \geq 0} \left\{ Wg(y) - \left(1 - \frac{\partial}{\partial x}W\right)y \right\} = 0 \quad (5.1)$$

with the boundary condition

$$W(t, 0) = W(0, x) = 0. \quad (5.2)$$

Moreover, if $\liminf_{x \rightarrow \infty} h(x)/x > 0$, then a viscosity solution to (5.1)–(5.2) is unique in the following sense. If a continuous function w with polynomial growth is a viscosity solution to (5.1)–(5.2), then $W = w$.

Until the end of this section, we assume $\tilde{\mu} > 0$ to focus on the case where the expected security price decreases over time.

5.1. Mixed Power MI Function. Here, we consider the case where g is given by

$$g(x) = \beta x^{\tilde{\pi}} \quad (0 \leq x \leq \bar{x}_0), \quad \alpha x^\pi + \gamma \quad (x > \bar{x}_0) \quad (5.3)$$

for some $\bar{x}_0 \geq 0$, $\alpha > 0$ and $0 < \tilde{\pi} < 1 < \pi$. Because g is continuously differentiable, β and γ must satisfy

$$\beta = \frac{\pi}{\tilde{\pi}} \alpha \bar{x}_0^{\pi - \tilde{\pi}}, \quad \gamma = \left(\frac{\pi}{\tilde{\pi}} - 1\right) \alpha \bar{x}_0^\pi.$$

Figure 1 shows the form of g when $\tilde{\pi} = 0.5$ and $\pi = 2$.

The next result is a pure extension of Theorem 5.4 in [14].

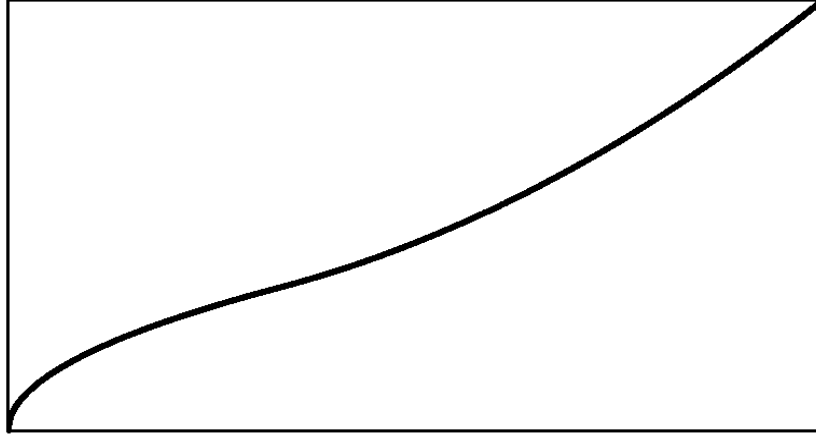


FIGURE 1. Form of the MI function $g(x)$ defined as (5.3) with $\tilde{\pi} = 0.5$ and $\pi = 2$. The horizontal axis corresponds to x . The vertical axis corresponds to $g(x)$.

Theorem 5.3. *Set*

$$\begin{aligned} x^{*,1} &= \frac{1}{\delta_\pi} \mathbb{B} \left(1 - \exp \left(-\frac{\pi}{\pi-1} (\tilde{\mu} + \gamma) T \right); \frac{1}{\pi} + 1, 2 \right), \\ x^{*,2} &= \nu_\pi T, \end{aligned}$$

where

$$\delta_\pi = \alpha^{1/\pi} \pi \left(\frac{\tilde{\mu} + \gamma}{\pi - 1} \right)^{\frac{\pi-1}{\pi}}, \quad \nu_\pi = \left(\frac{\tilde{\mu} + \gamma}{(\pi - 1)\alpha} \right)^{1/\pi}$$

and

$$\mathbb{B}(z; a, b) = \int_0^z \frac{dx}{x^{a-1}(1-x)^{b-1}}$$

is the incomplete Beta function.

(i) If $x_0 \geq x^{*,1}$, then we have

$$J(T, c_0, x_0, s_0) = c_0 + \frac{s_0}{\delta_\pi} \left(1 - \exp \left(-\frac{\pi}{\pi-1} (\tilde{\mu} + \gamma) T \right) \right)^{\frac{\pi-1}{\pi}},$$

and its optimizer is given by

$$\hat{x}_t = \nu_\pi \left(1 - \exp \left(-\frac{\pi}{\pi-1} (\tilde{\mu} + \gamma) (T - t) \right) \right)^{-1/\pi}.$$

(ii) If $x_0 \leq x^{*,2}$, then we have

$$J(T, c_0, x_0, s_0) = c_0 + s_0 \cdot \frac{1 - e^{-\delta_\pi x_0}}{\delta_\pi},$$

and its optimizer is given by

$$\hat{x}_t = \nu_\pi 1_{[0, x_0/\nu_\pi]}(t).$$

Similarly to Theorem 5.4 in [14], the form of the optimal strategy changes drastically according to the initial shares x_0 , and we do not have an analytical solution when $x^{*,2} < x_0 < x^{*,1}$. Moreover, when $x_0 \leq x^{*,2}$, the optimal strategy is the TWAP strategy, that is to sell with constant speed ν_π . The TWAP strategy is the optimal strategy for the Almgren–Chriss model, which is a standard model of optimal execution, for the risk-neutral trader [2, 7, 16, 17]. Theorem 5.3(ii) is also obtained as a corollary of the result of the next subsection. In addition,

$$\nu_\pi > \left(\frac{\gamma}{(\pi - 1)\alpha} \right)^{1/\pi} = \left(\frac{\pi - \tilde{\pi}}{\tilde{\pi}(\pi - 1)} \right)^{1/\pi} \quad \bar{x}_0 > \bar{x}_0;$$

hence we can verify that $\hat{x}_t \in \{0\} \cup (\bar{x}_0, \infty)$ in both cases of Theorem 5.3(i)(ii). This is consistent with Theorem 4.1.

5.2. TWAP Strategies for Small Amount Execution. Next, we consider the case where the amount x_0 of initial shares is small. Here, we do not restrict the form of g without [A1]–[A4].

Before stating the result, we prepare the following proposition.

Proposition 5.4. *Set $G_h(x) = xh(x) - g(x)$. Then there is a unique $\nu_h \in (\bar{x}_0, \infty)$ such that $G_h(\nu_h) = \tilde{\mu}$.*

Theorem 5.5. *If $x_0 \leq \nu_h T$, then we have*

$$J(T, c_0, x_0, s_0) = c_0 + s_0 \cdot \frac{1 - e^{-h(\nu_h)x_0}}{h(\nu_h)}, \quad (5.4)$$

and its optimizer is given by

$$\hat{x}_t = \nu_h 1_{[0, x_0/\nu_h]}(t). \quad (5.5)$$

This theorem implies the robustness of the optimality of the TWAP strategy for general shaped MI functions. When x_0 is small, the optimal execution strategy is to sell the security at speed $\nu_h (> \bar{x}_0)$ until the time when the remaining shares become zero.

Remark 5.6.

(i) As mentioned in Theorems 4.2 and 5.1 in [14], when $g(x) = \alpha x$ ($\alpha > 0$) is given as a linear function, we have

$$J(T, c_0, x_0, s_0) = c_0 + s_0 \cdot \frac{1 - e^{-\alpha x_0}}{\alpha}, \quad (5.6)$$

and the corresponding nearly optimal execution strategy is a quasi-block liquidation with the initial time; that is, $\hat{x}_t^\delta = (x_0/\delta) 1_{[0, \delta]}(t)$ with $\delta \rightarrow 0$. This strategy formally corresponds to (5.5) taking the limit $\nu_h \rightarrow \infty$. Note that $h(x) \equiv \alpha$; hence (5.4) coincides with (5.6).

(ii) Let us consider an extreme case where

$$g(x) = \hat{g}(x - \bar{x}_0)1_{[\bar{x}_0, \infty)}(x) \quad (5.7)$$

for some increasing convex function $\hat{g} \in C^1([0, \infty); [0, \infty))$ with $\hat{g}(0) = \hat{g}'(0) = 0$ and $\hat{g}'(\infty) = \infty$. The form of $g(x)$ is shown in Figure 2 for $\hat{g}(x)$ set as x^3 . In this case, we can completely avoid the MI cost by selling at a speed lower than or equal to \bar{x}_0 . Therefore, the optimal execution strategy seems to be $\tilde{x}_t = \bar{x}_0 1_{[0, x_0/\bar{x}_0]}(t)$ at a glance. Following the strategy, $(\tilde{x}_t)_t$, we get the expected proceeds

$$\tilde{\mathcal{C}} := s_0 \int_0^{x_0/\bar{x}_0} e^{-\tilde{\mu}t} \bar{x}_0 dt = s_0 \iota(\tilde{\mu}/\bar{x}_0; x_0),$$

where $\iota(y; x) = (1 - e^{-xy})/y$. However, Theorem 5.5 implies that this strategy is not optimal; the optimal execution speed, ν_h , is strictly greater than \bar{x}_0 . We compare the expected proceeds

$$\hat{\mathcal{C}} := J(T, 0, x_0, s_0) = s_0 \iota(h(\nu_h); x_0)$$

obtained by the optimal strategy, $(\hat{x}_t)_t$, with $\tilde{\mathcal{C}}$ obtained by $(\tilde{x}_t)_t$. Let us denote $\hat{G}(x) = x\hat{g}'(x) - \hat{g}(x)$. Then we see that $\hat{G}(\nu_h - \bar{x}_0) > \hat{G}(0) = 0$, and thus

$$\tilde{\mu} = G_h(\nu_h) = \hat{G}(\nu_h - \bar{x}_0) + \bar{x}_0 h(\nu_h) > \bar{x}_0 h(\nu_h).$$

This implies that $h(\nu_h) < \tilde{\mu}/\bar{x}_0$. Because $\iota(\cdot; x_0)$ is decreasing, we have $\hat{\mathcal{C}} > \tilde{\mathcal{C}}$. This is because selling at a lower speed increases the execution time and the timing cost. The trader should sell with the optimal speed, ν_h , and accept the MI cost.

5.3. Generalization of a Previous Result in Ishitani and Kato [12]. As an application of Theorem 5.5, we provide an analytical solution to an optimal execution problem with uncertain MI studied in Section 5.2 of [12]. We consider the optimization problem

$$\sup_{(x_t)_t \in \mathcal{A}_T(x_0)} \mathbb{E} \left[\int_0^T S_t x_t dt \right], \quad (5.8)$$

where $(S_t)_t$ is given by the SDE:

$$dS_t = S_t(-\tilde{\mu}dt + \sigma dB_t - g(x_t)dL_t), \quad S_0 = s_0.$$

Here, $(L_t)_t$ is the Lévy process, which is independent of $(B_t)_t$ and whose distribution is given by the Gamma distribution

$$P(L_t - \gamma t \in dz) = \frac{1}{\Gamma(\alpha_1 t) \beta_1^{\alpha_1 t}} z^{\alpha_1 t - 1} e^{-z/\beta_1} 1_{(0, \infty)}(z) dz,$$

where $\alpha_1, \beta_1, \gamma > 0$ satisfy $\alpha_1 \beta_1 \leq 8\gamma$ and $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ is the Gamma function. Moreover, we assume that $g(x) = \alpha_0 x^2$ is given as a quadratic function with $\alpha_0 \geq 0$.

In Section 5.2 of [12], we do not find the explicit form of the optimal strategy to (5.8), even when x_0 is small. However, numerical experiments suggest that the

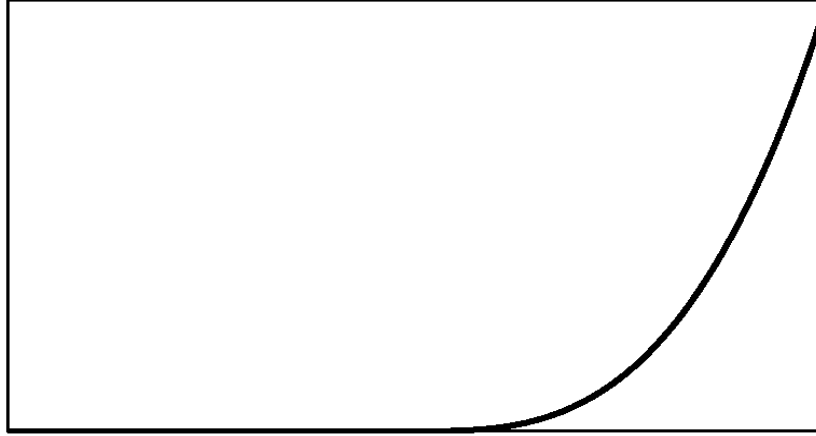


FIGURE 2. Form of the MI function, $g(x)$, defined as (5.7) with $\hat{g}(x) = x^3$. The horizontal axis corresponds to x . The vertical axis corresponds to $g(x)$.

optimal strategy with small x_0 is the TWAP strategy. Here, we prove mathematically that this conjecture is true.

Theorem 5.7. *Let \hat{v} be the solution to*

$$\gamma\alpha_0\hat{v}^2 + \alpha_1 \left\{ 2 \left(1 - \frac{1}{1 + \alpha_0\beta_1\hat{v}^2} \right) - \log(\alpha_0\beta_1\hat{v}^2 + 1) \right\} = \tilde{\mu}.$$

If $x_0 \leq \hat{v}T$, then the optimal strategy for (5.8) is given by the TWAP strategy

$$\hat{x}_t = \hat{v}1_{[0, x_0/\hat{v}]}(t). \quad (5.9)$$

6. Concluding Remarks

In this paper, we studied the optimal execution problem with S-shaped MI functions as a continuation of [15]. We showed that our value function is characterized as a viscosity solution of the corresponding HJB equation. This is an extended result of that in [14]. Moreover, we provided the verification theorem to show that the optimal execution speed is not in the range $(0, \bar{x}_0]$. This implies that the trader should not blindly decrease the execution speed to reduce the MI cost.

In the Black–Scholes market model, we found that an optimal execution strategy is the TWAP strategy when the number of shares of the security held is small. A concrete form is not required for the MI function, g , so this result is robust and suggests the optimality of TWAP strategy in practice.

The volume-weighted average price (VWAP) strategy is widely used in trading practice rather than the TWAP strategy [19]. Gatheral and Schied [8] pointed out that we should regard the time parameter, t , not as physical time but as volume time. Volume time implies a stochastic clock, which is measured by a market trading volume process [3, 9, 20, 25]. If we consider the model on a volume time line, we may find the optimality of the VWAP strategy in a similar way to Theorem 5.5. However, we should not ignore the randomness of the market trading volume. One of our future tasks is to construct a model of optimal execution with S-shaped MI functions on a volume time line.

Furthermore, to apply Theorem 4.1, we require the value function, J , to be smooth, whereas it is difficult to show smoothness in general. Moreover, the solvability of SDE (4.3) is not clear. Further study is needed.

Appendix A. Supplemental Arguments

We present the following propositions, which link the results in our previous study [15] with the present model.

Proposition A.1. *Let $t > 0$ and let $(c, x, s) \in D$. For each $(x_r)_{r \leq t} \in \mathcal{A}_t(x)$, there is a unique process, $(C_r, X_r, S_r)_{r \leq t}$, that satisfies (2.2) and $(C_0, X_0, S_0) = (c, x, s)$.*

The comparison theorem for solutions of SDEs (see Proposition 5.2.18 in [13] for instance) tells us that

$$0 \leq S_r \leq Z_r(s) \text{ a.s.}, \quad (\text{A.1})$$

where $(Z_r(s))_r$ is defined in (3.6). Moreover, Lemma B.1 in [14] tells us that

$$\mathbb{E}[\sup_{0 \leq r \leq t} Z_r(s)^m] < \infty \quad (\text{A.2})$$

for each t, s and $m > 0$. Based on (A.1)–(A.2), we see that our value function, $J(t, c, x, s)$, is well-defined and finite.

Proposition A.2. $J(t, c, x, s) = J^\infty(t, c, x, s)$.

In [15], we show some properties of $J^\infty(t, c, x, s)$. Proposition A.2 implies that these results also hold for $J(t, c, x, s)$.

Appendix B. Proofs

Proof of Proposition 5.4. First, [A3] implies that

$$G_h(\bar{x}_0) = \bar{x}_0 h(\bar{x}_0) - \int_0^{\bar{x}_0} h(x) dx \leq \bar{x}_0 h(\bar{x}_0) - \bar{x}_0 h(\bar{x}_0) = 0. \quad (\text{B.1})$$

[A3] also tells us that

$$G_h(x) - G_h(y) \geq (h(x) - h(y))y > 0 \quad (\text{B.2})$$

for each $x > y > \bar{x}_0$. Hence, G_h is strictly increasing on (\bar{x}_0, ∞) . Moreover, letting $x \rightarrow \infty$ in (B.2), we see that

$$\lim_{x \rightarrow \infty} G_h(x) = \infty, \quad (\text{B.3})$$

owing to condition [A4]. Because G_h is continuous on $[\bar{x}_0, \infty)$ and $\tilde{\mu}$ is positive, (B.1)–(B.3) immediately give the assertion. \square

To show Proposition A.1, we prepare a lemma.

Lemma B.1. *Let $(\varphi_t)_t$ be an $(\mathcal{F}_t)_t$ -progressively measurable process such that $|\varphi_t| \leq K$ for some positive constant K . Then there is a $C_{K,T} > 0$ that depends only on K and T , such that*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \exp \left(\int_0^t \varphi_r dB_r \right) \right] \leq C_{K,T}.$$

Proof. Put

$$N_t = \exp \left(\int_0^t \varphi_r dB_r - \frac{1}{2} \int_0^t \varphi_r^2 dr \right) - 1.$$

It's formula immediately implies that $(N_t)_t$ is a continuous local martingale starting at 0 and $d\langle N \rangle_t = (N_t + 1)^2 \varphi_t^2 dt$.

Take any $R > 0$ and define $\tau_R = \inf\{t \geq 0 ; \langle N \rangle_t \geq R\} \wedge T$ and $m_t^R = \mathbb{E}[\langle N \rangle_{t \wedge \tau_R}] (\leq R < \infty)$. Then we observe

$$0 \leq m_t^R \leq 2K^2 \mathbb{E} \left[\int_0^{t \wedge \tau_R} (N_r^2 + 1) dr \right] \leq 2K^2 T + 2K^2 \int_0^t m_r^R dr.$$

We apply the Gronwall inequality to obtain

$$m_t^R \leq 2K^2 T + 4K^2 T^2 e^{2K^2 T} =: C'_{K,T}.$$

The Chebyshev inequality implies that $\tau_R \nearrow T$, $R \rightarrow \infty$ a.s., and hence $\mathbb{E}[\langle N \rangle_T] = \lim_{R \rightarrow \infty} m_t^R \leq C'_{K,T}$ by the monotone convergence theorem. Now we arrive at

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} \exp \left(\int_0^t \varphi_r dB_r \right) \right] &\leq e^{K^2 T/2} (2 \mathbb{E}[\langle N \rangle_T]^{1/2} + 1) \\ &\leq e^{K^2 T/2} \left(2\sqrt{C'_{K,T}} + 1 \right). \end{aligned}$$

□

Proof of Proposition A.1. It suffices to show the existence and uniqueness of process $(S_r)_{r \leq t}$ for each given $(x_r)_r \in \mathcal{A}_t(x)$ and $s > 0$.

Step 1. For each $n \in \mathbb{N}$, define

$$\tau_n = \inf \left\{ r \geq 0 ; \int_0^r g(x_v) dv \geq n \right\} \wedge t$$

and put $x_r^n = x_r 1_{[0, \tau_n]}(r)$. Then we can show that there is a unique solution $(Y_r^n)_r$ to the following SDE by the standard argument:

$$dY_r^n = b(Y_r^n) dr + \sigma(Y_r^n) dB_r - g(x_r^n) dr, \quad Y_0^n = \log s.$$

It's formula implies that the process $S_r^n := \exp(Y_r^n)$ satisfies

$$dS_r^n = \hat{b}(S_r^n) dr + \hat{\sigma}(S_r^n) dB_r - S_r^n g(x_r^n) dr, \quad S_0^n = s.$$

We see that $\tau_n \leq \tau_m$ and $S_r^n = S_r^m$, $r \in [0, \tau_n]$ a.s. for each $n < m$. Therefore, we can define $S_r^\infty = \lim_{n \rightarrow \infty} S_r^n$ for each $r \in [0, \tau) \cap [0, t]$ a.s., where $\tau = \lim_{n \rightarrow \infty} \tau_n$.

Next, we show that $\lim_{r \rightarrow \tau} S_r^\infty = 0$ a.s. on $\{\tau \leq t\}$. For each $\delta > 0$, we see that

$$0 \leq S_{\tau-\delta}^\infty = \lim_{n \rightarrow \infty} S_{\tau-\delta}^n \leq s D_t G_\delta \quad \text{on } \{\tau \leq t\},$$

where

$$\begin{aligned} D_t &= \liminf_{n \rightarrow \infty} \sup_{0 \leq r \leq t} \exp \left(\int_0^r b(Y_v^n) dv + \int_0^r \sigma(Y_v^n) dB_v \right), \\ G_\delta &= \exp \left(- \int_0^{\tau - \delta} g(x_r) dr \right). \end{aligned}$$

Because b and σ are bounded, Lemma B.1 implies that $E[D_t] < \infty$, hence $D_t < \infty$ a.s. Moreover, based on the definition of τ , it holds that $G_\delta \mathbf{1}_{\{\tau \leq t\}} \rightarrow 0$, $\delta \rightarrow 0$ a.s. Thus, we have $\lim_{\delta \rightarrow 0} S_{\tau - \delta}^\infty = 0$ a.s. on $\{\tau \leq t\}$.

Therefore, we can define $S_r := S_{r \wedge \tau}^\infty$ as a continuous process on $[0, t]$, and it holds that

$$\begin{aligned} & s + \int_0^r \hat{\sigma}(S_v) dB_v + \int_0^r (\hat{b}(S_v) - S_v g(x_v)) dv \\ &= s + \int_0^{r \wedge \tau} \hat{\sigma}(S_v^\infty) dB_v + \int_0^{r \wedge \tau} (\hat{b}(S_v^\infty) - S_v^\infty g(x_v)) dv \\ &= \lim_{n \rightarrow \infty} S_{r \wedge \tau_n}^n = S_r, \quad r \leq t. \end{aligned}$$

Thus, $(S_r)_r$ satisfies (2.2).

Step 2. Next, we show the uniqueness of the solution to (2.2). Assume that $(\tilde{S}_r)_r$ satisfies (2.2) and $\tilde{S}_0 = s$. We see that $Y_{r \wedge \tau_n} = \log S_{r \wedge \tau_n}$ and $\tilde{Y}_{r \wedge \tau_n} = \log \tilde{S}_{r \wedge \tau_n}$ satisfy (2.3). Because b and σ are Lipschitz continuous, we have $E[\sup_{0 \leq r \leq t} |Y_{r \wedge \tau_n} - \tilde{Y}_{r \wedge \tau_n}|^2] = 0$. This implies that $S_{r \wedge \tau_n} = \tilde{S}_{r \wedge \tau_n}$, $r \leq t$ a.s. Then we have $S_{r \wedge \tau_n} = \tilde{S}_{r \wedge \tau_n}$, $r \leq t$, a.s. Letting $n \rightarrow \infty$, we arrive at $S_{r \wedge \tau} = \tilde{S}_{r \wedge \tau}$, $r \leq t$ a.s. Based on (2.2), $S_r = \tilde{S}_r = 0$ for each r larger than τ a.s. on $\{\tau \leq t\}$, so we conclude that $(S_r)_r$ is equal to $(\tilde{S}_r)_r$ a.s. \square

Proof of Proposition A.2. Because $J(t, c, x, s) \geq J^\infty(t, c, x, s)$ is clear, we may prove the opposite inequality.

Fix any $(x_r)_r \in \mathcal{A}_t(x)$ and denote by $(C_r, X_r, S_r)_{r \leq t}$ its controlled process. Take any $K > 0$ and set $x_r^K = x_r \wedge K$. Then $(x_r^K)_r \in \mathcal{A}_t^\infty(x)$ holds. Let (C_r^K, X_r^K, S_r^K) be the controlled process of $(x_r^K)_r$. Then we have $X_t^K \geq X_t$. Moreover, Proposition 5.2.18 in [13] implies that $S_r^K \geq S_r$, $r \leq t$ a.s. Therefore, it holds that

$$C_t^K = c + \int_0^t x_r^K S_r^K dr \geq c + \int_0^t x_r S_r dr \quad a.s.,$$

and the monotone convergence theorem tells us that $\liminf_{K \rightarrow \infty} C_t^K \geq C_t$ a.s. Because $u \in \mathcal{C}$, we have $u(C_t, X_t, S_t) \leq \liminf_{K \rightarrow \infty} u(C_t^K, X_t^K, S_t^K)$. Then we apply Fatou's lemma to see that

$$E[u(C_t, X_t, S_t)] \leq \liminf_{K \rightarrow \infty} E[u(C_t^K, X_t^K, S_t^K)] \leq J^\infty(t, c, x, s).$$

Because $(x_r)_r \in \mathcal{A}_t(x)$ is arbitrary, we complete the proof. \square

To prove Theorem 3.1, we define $F : D \times \mathbb{R}^3 \times \mathcal{S} \rightarrow \mathbb{R} \cup \{-\infty\}$ by

$$\begin{aligned} F(z, p, \Sigma) &= -\frac{1}{2}\hat{\sigma}(s)^2\Sigma_{ss} - \hat{b}(s)p_s + H(s, p), \\ H(s, p) &= \inf_{y \geq 0} f(y; s, p), \\ f(y; s, p) &= sp_s g(y) - (sp_c - p_x)y, \end{aligned}$$

where $\mathcal{S} \subset \mathbb{R}^3 \otimes \mathbb{R}^3$ is the set of three-dimensional real symmetric matrices, and we denote

$$z = \begin{pmatrix} c \\ x \\ s \end{pmatrix}, \quad p = \begin{pmatrix} p_c \\ p_x \\ p_s \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{cc} & \Sigma_{cx} & \Sigma_{cs} \\ \Sigma_{xc} & \Sigma_{xx} & \Sigma_{xs} \\ \Sigma_{sc} & \Sigma_{sx} & \Sigma_{ss} \end{pmatrix}.$$

Note that (3.2) is equivalent to

$$\frac{\partial}{\partial t} J + F(z, \mathcal{D}J, \mathcal{D}^2 J) = 0. \quad (\text{B.4})$$

Moreover, put

$$\begin{aligned} \mathcal{U} &= \left\{ (z, p, \Sigma) \in \tilde{D} \times \mathbb{R}^3 \times \mathcal{S} ; F(z, p, \Sigma) > -\infty \right\}, \\ \mathcal{R} &= \tilde{D} \times (\mathbb{R}^2 \times (0, \infty)) \times \mathcal{S}. \end{aligned}$$

Note that $p_s \geq 0$ holds for each $(z, p, \Sigma) \in \mathcal{U}$. Moreover, we have $\mathcal{R} \subset \mathcal{U}$.

Lemma B.2. *For each $(z, p, \Sigma) \in \mathcal{R}$, we have*

$$H(s, p) = f(h^{-1}(\mathcal{H}(s, p) \vee h(x_0)); s, p) \wedge 0 = f(\Xi(s, p); s, p), \quad (\text{B.5})$$

where $\mathcal{H}(s, p)$ and $\Xi(s, p)$ are given by (4.1)–(4.2). In particular, F is continuous on \mathcal{R} .

Proof. First, we note that $H(s, p) = sp_s \bar{H}(\mathcal{H}(s, p))$, where

$$\bar{H}(\bar{y}) = \inf_{y \geq 0} \bar{f}(y; \bar{y}), \quad \bar{f}(y; \bar{y}) = g(y) - \bar{y}y.$$

We see that

$$\bar{H}(\bar{y}) = \bar{f}(h^{-1}(\bar{y} \vee h(\bar{x}_0)); \bar{y}) \wedge 0. \quad (\text{B.6})$$

Indeed, if $\bar{y} \leq h(\bar{x}_0)$, we observe

$$\frac{\partial}{\partial y} \bar{f}(y; \bar{y}) = h(y) - \bar{y} \geq h(y) - h(\bar{x}_0) \geq 0, \quad y \geq 0$$

by [A3]. Thus we get $\bar{H}(\bar{y}) = \bar{f}(0; \bar{y}) = 0$. Moreover, [A3] also implies

$$\bar{f}(h^{-1}(h(\bar{x}_0)); \bar{y}) = \bar{f}(\bar{x}_0; \bar{y}) = \int_0^{\bar{x}_0} h(y') dy' - \bar{y}\bar{x}_0 \geq (h(\bar{x}_0) - \bar{y})\bar{x}_0 \geq 0;$$

hence, (B.6) holds. In contrast, if $\bar{y} > h(\bar{x}_0)$, we see that $\bar{f}(\cdot; \bar{y})$ attains the minimum at $h^{-1}(\bar{y})$ or 0. When $\bar{f}(h^{-1}(\bar{y}); \bar{y}) < 0$, it holds that $\bar{H}(\bar{y}) = \bar{f}(h^{-1}(\bar{y}); \bar{y})$. When $\bar{f}(h^{-1}(\bar{y}); \bar{y}) \geq 0$, it holds that $\bar{H}(\bar{y}) = 0$. In both cases, we see that (B.6) actually holds. (B.6) implies the first equality of (B.5). The second equality of (B.5) is obtained by a straightforward calculation using [A3]. The last assertion is obtained by the continuity of \hat{b} , $\hat{\sigma}$, h , \bar{H} , and $\mathcal{H}(s, p)$. \square

The following proposition is obtained by a standard argument (see [6, 18, 21] for details).

Proposition B.3. *J is the viscosity supersolution of (3.2).*

Proposition B.4. *Assume (3.1). Then J is the viscosity subsolution of (3.2).*

Proof. Fix each $(t, z) \in (0, T] \times \tilde{D}$. Let $v \in C^{1,2}((0, T] \times \tilde{D})$ be a test function, such that $J - v$ attains the local maximum, 0, at (t, z) . Then we can find an $r > 0$ such that

$$J(t', z') < v(t', z') \quad (\text{B.7})$$

holds for each $(t', z') \in \bar{B}_r((t, z)) \setminus \{(t, z)\}$, where

$$\bar{B}_r((t, z)) = \{(t', z') \in (0, T] \times \tilde{D} ; |t' - t|^2 + |z' - z|^2 \leq r^2\}.$$

For each $L > 0$, define

$$\begin{aligned} F^L(z, p, \Sigma) &= -\frac{1}{2}\hat{\sigma}(s)^2\Sigma_{ss} - \hat{b}(s)p_s + \inf_{0 \leq y \leq L} f(y; s, p), \\ J^L(t, c, x, s) &= \sup_{(x_r)_r \in \mathcal{A}_t^L(x)} \mathbb{E}[u(C_t, X_t, S_t)], \\ \mathcal{A}_t^L(x) &= \{(x_r)_r \in \mathcal{A}_t(x) ; |x_r| \leq L\}. \end{aligned}$$

Here, $(C_r, X_r, S_r)_r$ is given as (2.2) and $(C_0, X_0, S_0) = (c, x, s)$. Note that $J^L(t, c, x, s) \nearrow J(t, c, x, s)$, $L \rightarrow \infty$. By the same argument as Proposition B.18 in [14], we see that J^L is a viscosity solution of

$$\frac{\partial}{\partial t} J^L + F^L(z, \mathcal{D}J^L, \mathcal{D}^2 J^L) = 0. \quad (\text{B.8})$$

Because $J^L - v$ is continuous, $J^L - v$ attains a maximum on the set $\bar{B}_r((t, z))$; namely, there is a $(t_L, z_L) \in \bar{B}_r((t, z))$ such that $\max_{\bar{B}_r((t, z))} (J^L - v) = J^L(t_L, z_L) - v(t_L, z_L)$. We show that

$$(t_L, z_L) \longrightarrow (t, z), \quad L \rightarrow \infty. \quad (\text{B.9})$$

Because $\{(t_L, z_L)\}_L$ is a bounded sequence, we see that for each increasing sequence, $(L_n)_n \subset (0, \infty)$, there is a subsequence, $(L_{n_k})_k$, such that $(t_{L_{n_k}}, z_{L_{n_k}})$ converges to a point, $(t^*, z^*) \in \bar{B}_r(t, z)$. Dini's theorem implies that $J^L \rightarrow J$, $L \rightarrow \infty$ is uniform convergence on any compact set; thus, we see that

$$J(t^*, z^*) - v(t^*, z^*) = \lim_{k \rightarrow \infty} (J^{L_{n_k}}(t_{L_{n_k}}, z_{L_{n_k}}) - v((t_{L_{n_k}}, z_{L_{n_k}}))) = 0.$$

Combining this with (B.7), we conclude that (t^*, z^*) must coincide with (t, z) . Therefore, (B.9) is true.

Next, we define $\tilde{v} \in C^{1,2}((0, T] \times \tilde{D})$ by

$$\tilde{v}(t', z') = v(t', z') + J^L(t_L, z_L) - v(t_L, z_L).$$

Then we see that $J^L - \tilde{v}$ attains a local maximum, 0, at (t_L, z_L) . Moreover, because J^L is a viscosity solution to (B.8), it holds that

$$\begin{aligned} & \frac{\partial}{\partial t} \tilde{v}(t_L, z_L) + F^L(z_L, \mathcal{D}\tilde{v}(t_L, z_L), \mathcal{D}^2\tilde{v}(t_L, z_L)) \\ &= \frac{\partial}{\partial t} v(t_L, z_L) + F^L(z_L, \mathcal{D}v(t_L, z_L), \mathcal{D}^2v(t_L, z_L)) \leq 0. \end{aligned} \quad (\text{B.10})$$

Note that $(z_L, \mathcal{D}v(t_L, z_L), \mathcal{D}^2v(t_L, z_L)) \in \mathcal{R}$ holds for large enough L . Indeed, (3.1) implies that $(\partial/\partial s)v(t, z) > 0$, and the convergence $(t_L, z_L) \rightarrow (t, z)$ and the continuity of $\mathcal{D}v$ lead us to $(\partial/\partial s)v(t_L, z_L) > 0$ for large enough L . Moreover, using Lemma B.2 and Dini's theorem again, we see that F^L converges to F as $L \rightarrow \infty$ uniformly on any compact set in \mathcal{R} . Therefore, taking $L \rightarrow \infty$ in (B.10), we arrive at

$$\frac{\partial}{\partial t} v(t, z) + F(z, \mathcal{D}v(t, z), \mathcal{D}^2v(t, z)) \leq 0.$$

This completes the proof. \square

Proof of Theorem 3.1. Assertion (i) is a consequence of Propositions B.3–B.4. Assertion (ii) is obtained by the same arguments as the proofs of Propositions B.21–B.23 in [14]. We note that Proposition B.22 of [14] requires the convexity of g only on $[x_1, \infty)$ for large enough x_1 . \square

We prepare the following lemma to show Theorem 4.1.

Lemma B.5. *Assume [B] and that $J \in C^{1,1,1,2}((0, T) \times D)$. It holds that*

$$\frac{\partial}{\partial c} J(t, c, 0, s) = U'(c), \quad (\text{B.11})$$

$$\frac{\partial}{\partial x} J(t, c, 0, s) \geq sU'(c), \quad (\text{B.12})$$

$$\frac{\partial}{\partial s} J(t, c, 0, s) = \frac{\partial}{\partial x} J(t, c, x, 0) = 0, \quad (\text{B.13})$$

$$\frac{\partial}{\partial t} J(t, c, 0, s) = \frac{\partial}{\partial t} J(t, c, x, 0) = 0 \quad (\text{B.14})$$

for each $t > 0$ and $(c, x, s) \in D$.

Proof. (B.11), (B.13) and (B.14) are obtained from $J(t, c, 0, s) = J(t, c, x, 0) = U(c)$.

To show (B.12), for each fixed $t > 0$ and each $x \in (0, t^2)$, set $x_r = \sqrt{x}1_{[0, \sqrt{x}]}(r)$ and let $(C_r, X_r, S_r)_r$ be the controlled process associated with $(x_r)_r \in \mathcal{A}_t(x)$. Then we see that

$$\begin{aligned} & \frac{1}{x} (J(t, c, x, s) - J(t, c, 0, s)) \\ & \geq \frac{1}{x} \mathbb{E}[U(C_t) - U(c)] = \mathbb{E} \left[\int_0^1 U'(c + kx A_x) dk A_x \right], \end{aligned} \quad (\text{B.15})$$

where $A_x = \frac{1}{\sqrt{x}} \int_0^{\sqrt{x}} S_r dr$. By the Doob inequality, (3.18) in [11], and (A.1)–(A.2), we have $\mathbb{E}[|A_x - s|] \rightarrow 0$, $x \rightarrow 0$. In particular, A_x converges to s in probability.

Moreover, by (A.1)–(A.2) and the concavity of U , we have

$$\mathbb{E}[\sup_{0 \leq x \leq t^2} A_x^2] < \infty, \quad \mathbb{E}[\sup_{0 \leq k \leq 1, 0 \leq x \leq t^2} (U'(c + kxA_x))^2] \leq (U'(c))^2.$$

Therefore, we can apply the dominated convergence theorem to get

$$\mathbb{E} \left[\int_0^1 U'(c + kxA_x) dk A_x \right] \longrightarrow sU'(c), \quad x \rightarrow 0. \quad (\text{B.16})$$

(B.15)–(B.16) lead us to (B.12). \square

Proof of Theorem 4.1. Define

$$\hat{x}_t = \Xi(\bar{S}_t, \mathcal{D}J(T-t, \bar{C}_t, \bar{X}_t, \bar{S}_t))1_{[0, \bar{\tau})}(t). \quad (\text{B.17})$$

Because $\bar{X}_{T \wedge \bar{\tau}} \geq 0$, it holds that $(\hat{x}_t)_t \in \mathcal{A}_T(x_0)$. Then Proposition A.1 implies that there is a controlled process $(\hat{C}_t, \hat{X}_t, \hat{S}_t)_t$ associated with $(\hat{x}_t)_t$. We see that $\hat{C}_t = \bar{C}_{t \wedge \bar{\tau}}$, $\hat{X}_t = \bar{X}_{t \wedge \bar{\tau}}$ and $\hat{S}_{t \wedge \bar{\tau}} = \bar{S}_{t \wedge \bar{\tau}}$.

Put $\hat{\tau}_R = \inf\{t \geq 0; \hat{S}_t \geq R\} \wedge (T-1/R)_+$ for each $R > 0$. Note that (A.1)–(A.2) and the Chebyshev inequality imply that $\hat{\tau}_R \nearrow T$, $R \rightarrow \infty$. Ito's formula gives us

$$\begin{aligned} & \mathbb{E}[J(T - \hat{\tau}_R, \hat{C}_{\hat{\tau}_R}, \hat{X}_{\hat{\tau}_R}, \hat{S}_{\hat{\tau}_R})] - J(T, c_0, x_0, s_0) \\ &= \mathbb{E} \left[\int_0^{\hat{\tau}_R} \left(-\frac{\partial}{\partial t} J + \mathcal{L}^{\hat{x}_t} J \right) (T-t, \hat{C}_t, \hat{X}_t, \hat{S}_t) dt \right]. \end{aligned} \quad (\text{B.18})$$

Based on Theorem 3.1(i), Lemma B.2, and the smoothness of J , we have

$$\begin{aligned} & \left(-\frac{\partial}{\partial t} J + \mathcal{L}^{\hat{x}_t} J \right) (T-t, \hat{C}_t, \hat{X}_t, \hat{S}_t) \\ &= \left(-\frac{\partial}{\partial t} J + \sup_{y \geq 0} \mathcal{L}^y J \right) (T-t, \bar{C}_t, \bar{X}_t, \bar{S}_t) = 0, \quad t < \bar{\tau}. \end{aligned} \quad (\text{B.19})$$

On $\{t \geq \bar{\tau}\}$, either $\hat{X}_t = 0$ or $\hat{S}_t = 0$ holds. If $\hat{X}_t = 0$, we have

$$\begin{aligned} & \left(-\frac{\partial}{\partial t} J + \sup_{y \geq 0} \mathcal{L}^y J \right) (T-t, \hat{C}_t, 0, \hat{S}_t) \\ &= \sup_{y \geq 0} \left\{ \left(\hat{S}_t U'(\hat{C}_t) - \frac{\partial}{\partial x} J(T-t, \hat{C}_t, 0, \hat{S}_t) \right) y \right\} \\ &= 0 = \left(-\frac{\partial}{\partial t} J + \mathcal{L}^{\hat{x}_t} J \right) (T-t, \hat{C}_t, 0, \hat{S}_t), \quad t \geq \bar{\tau} \end{aligned} \quad (\text{B.20})$$

by Lemma B.5. If $\hat{S}_t = 0$, Lemma B.5 also implies that

$$\begin{aligned} & \left(-\frac{\partial}{\partial t} J + \sup_{y \geq 0} \mathcal{L}^y J \right) (T-t, \hat{C}_t, \hat{X}_t, 0) \\ &= 0 = \left(-\frac{\partial}{\partial t} J + \mathcal{L}^{\hat{x}_t} J \right) (T-t, \hat{C}_t, \hat{X}_t, 0), \quad t \geq \bar{\tau}. \end{aligned} \quad (\text{B.21})$$

Combining (B.18)–(B.21), we arrive at $E[J(T - \hat{\tau}_R, \hat{C}_{\hat{\tau}_R}, \hat{X}_{\hat{\tau}_R}, \hat{S}_{\hat{\tau}_R})] = J(T, c_0, x_0, s_0)$. Letting $R \rightarrow \infty$, we have $E[u(\hat{C}_T, \hat{X}_T, \hat{S}_T)] = J(T, c_0, x_0, s_0)$ due to the dominated convergence theorem. Therefore, $(\hat{x}_t)_t$ is the optimizer to $J(T, c_0, x_0, s_0)$. Based on (4.2) and (B.17), we see that $\hat{x}_t = 0$ or $\hat{x}_t > \bar{x}_0$. \square

Proof of Theorem 4.2. Fix any $(x_t)_t \in \mathcal{A}_T^\infty(x_0)$ and $n \in \mathbb{N}$, and let $x_t^n = (1 - 1/n)x_t$. Denote by $(C_t, X_t, S_t)_t$ (respectively, $(C_t^n, X_t^n, S_t^n)_t$) the controlled process associated with $(x_t)_t$ (respectively, $(x_t^n)_t$). Note that (A.1)–(A.2) imply $E[\sup_{0 \leq t \leq T} (S_t^n)^m] < \infty$ for each $m > 0$, and that $(C_t^n, X_t^n, S_t^n) \in \tilde{D}$, $t \in [0, T]$ holds a.s.

Take any $R > 0$ and set $\tau_R = \inf\{t \geq 0 ; S_t^n \geq R\} \wedge (T - 1/R)_+$. By a standard argument using Ito's formula, we arrive at

$$\begin{aligned} & E[v(T - \tau_R, C_{\tau_R}^n, X_{\tau_R}^n, S_{\tau_R}^n)] - v(T, c_0, x_0, s_0) \\ & \leq E \left[\int_0^{\tau_R} \left(-\frac{\partial}{\partial t} v + \sup_{y \geq 0} \mathcal{L}^y v \right) (T - t, C_t^n, X_t^n, S_t^n) dt \right]. \end{aligned}$$

Combining this with assumption (iii), we have

$$E[v(T - \tau_R, C_{\tau_R}^n, X_{\tau_R}^n, S_{\tau_R}^n)] \leq v(T, c_0, x_0, s_0).$$

Here, based on assumptions (i)–(ii) and the Chebyshev inequality, we see that $\tau_R \nearrow T$, $R \rightarrow \infty$ a.s. and

$$E[u(C_T^n, X_T^n, S_T^n)] \leq v(T, c_0, x_0, s_0). \quad (\text{B.22})$$

Because $(x_t)_t$ is essentially bounded, by using Theorem 2.5.9 in [18], we obtain $E[\sup_{0 \leq t \leq T} |\log S_t^n - \log S_t|^4] \rightarrow 0$, $n \rightarrow \infty$. Then we have $E[\sup_{0 \leq t \leq T} |S_t^n - S_t|^2] \rightarrow 0$, and thus $E[|C_T^n - C_T|] \rightarrow 0$. Moreover, we see that $E[|X_T^n - X_T|] \rightarrow 0$. Therefore, letting $n \rightarrow \infty$ in (B.22) and applying Lemma B.2 in [14], we arrive at

$$E[u(C_T, X_T, S_T)] \leq v(T, c_0, x_0, s_0).$$

Because $(x_t)_t \in \mathcal{A}_T^\infty(x_0)$ is arbitrary, we deduce that

$$J(T, c_0, x_0, s_0) = J^\infty(T, c_0, x_0, s_0) \leq v(T, c_0, x_0, s_0).$$

This and assumption (iv) lead to the conclusion that

$$J(T, c_0, x_0, s_0) = E[u(\hat{C}_T, \hat{X}_T, \hat{S}_T)] = v(T, c_0, x_0, s_0). \quad \square$$

Theorems 5.3 and 5.5 are obtained by a straightforward calculation using Theorem 4.2.

Proof of Theorem 5.7. Theorem 5.2 in [12] tells us that (5.8) is equivalent with the optimization problem

$$c_0 + \sup_{(x_t)_t \in \mathcal{A}_T(x_0)} \int_0^T \hat{S}_t x_t dt,$$

where

$$\begin{aligned} d\hat{S}_t &= -\hat{S}_t(\tilde{\mu} + \hat{g}(x_t))dt, \quad \hat{S}_0 = s_0, \\ \hat{g}(x) &= \gamma\alpha_0x^2 + \alpha_1 \log(\alpha_0\beta_1x^2 + 1). \end{aligned}$$

Moreover, we see that

$$\hat{g}''(x) \geq \frac{\alpha_0\alpha_1\beta_1(\alpha_0\beta_1x^2 - 3)^2}{4(\alpha_0\beta_1x^2 + 1)^2},$$

hence, $\hat{g}(x)$ is strictly convex (see Corollary 1.3.10 in [22] for instance). Therefore, we can apply Theorem 5.5 to complete the proof. The optimal execution speed $\hat{\nu}$ satisfies $G_{\hat{g}'}(\hat{\nu}) = \tilde{\mu}$, where

$$\begin{aligned} G_{\hat{g}'}(x) &= x\hat{g}'(x) - \hat{g}(x) \\ &= \gamma\alpha_0x^2 + \alpha_1 \left\{ 2 \left(1 - \frac{1}{1 + \alpha_0\beta_1x^2} \right) - \log(\alpha_0\beta_1x^2 + 1) \right\}. \end{aligned}$$

□

References

1. Alfonsi, A., Fruth, A., and Schied, A.: Optimal execution strategies in limit order books with general shape functions, *Quant. Finance* **10** (2010) 143–157.
2. Almgren, R. F. and Chriss, N.: Optimal execution of portfolio transactions, *J. Risk*, **18** (2000) 57–62.
3. Ané, T. and Geman, H.: Order flow, transaction clock, and normality of asset returns, *J. Finance*, **55** (2000), no. 5, 2259–2284.
4. Bertsimas, D. and Lo, A. W.: Optimal control of execution costs, *J. Financ. Mark.*, **1** (1998) 1–50.
5. Da Lio, F., Ley, O.: Convex Hamilton–Jacobi equations under superlinear growth conditions on data. *Appl. Math. Optim.* **63** (2011), no. 3, 309–339.
6. Fleming, W. H. and Soner, H. M.: *Controlled Markov Processes and Viscosity Solutions*, Springer, New York, 1992.
7. Gatheral, J. and Schied, A.: Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework, *Int. J. Theor. Appl. Finance* **14** (2011), no. 3, 353–368.
8. Gatheral, J. and Schied, A.: Dynamical models of market impact and algorithms for order execution, *Handbook on Systemic Risk*, Eds. Fouque, J. P. and Langsam, J., Cambridge University Press, Cambridge, 2013, 579–602.
9. Geman, H.: Stochastic clock and financial markets, *Aspects of Mathematical Finance*, Eds. Yor, M., Springer, New York, 2008, 37–52.
10. Guéant, O.: Permanent market impact can be nonlinear, *Preprint* (2014).
11. Ishitani, K. and Kato, T.: Mathematical formulation of an optimal execution problem with uncertain market impact, *Commun. Stoch. Anal.* **9** (2015), no. 3, 113–129.
12. Ishitani, K. and Kato, T.: Theoretical and numerical analysis of an optimal execution problem with uncertain market impact, *Commun. Stoch. Anal.* **9** (2015), no. 3, 343–366.
13. Karatzas, I. and Shreve, S. E.: *Brownian Motion and Stochastic Calculus*, 2nd. edition, Springer, New York, 1991.
14. Kato, T.: An optimal execution problem with market impact, *Finance Stoch.* **18** (2014), no. 3, 695–732.
15. Kato, T.: Non-linearity of market impact functions: empirical and simulation-based studies on convex/concave market impact functions and derivation of an optimal execution model, *Trans. Jpn. Soc. Ind. Appl. Math.* **24** (2014), no. 3, 203–237 (in Japanese).
16. Kato, T.: VWAP execution as an optimal strategy, *JSIAM Lett.* **7** (2015) 33–36.
17. Kato, T.: An optimal execution problem in the volume-dependent Almgren–Chriss model, *Preprint* (2017).

18. Krylov, N. V.: *Controlled Diffusion Processes*, Springer, Berlin, 1980.
19. Madhavan, A.: VWAP Strategies, *Trading* **1** (2002) 32–39.
20. Mazur, S.: Modeling market impact and timing risk in volume time, *Algorithmic Finance* **2** (2013), no. 2, 113–126.
21. Nagai, H.: *Stochastic Differential Equations*, Kyoritsu Shuppan, Tokyo, 1999.
22. Niculescu, C. and Persson, L. -E.: *Convex Functions and Their Applications*, Springer, New York, 2006.
23. Obizhaeva, A. and Wang, J.: Optimal trading strategy and supply/demand dynamics, *J. Financ. Mark.* **16** (2013), no. 1, 1–32.
24. Roşu, I.: A dynamic model of the limit order book, *Rev. Financ. Stud.* **22** (2009), no. 11, 4601–4641.
25. Veraat, E. D. and Winkel, M.: Time change, *Encyclopedia of Quantitative Finance, Eds. Cont, R.*, Wiley, Chichester, 2010, 1812–1816.

TAKASHI KATO: ASSOCIATION OF MATHEMATICAL FINANCE LABORATORY (AMFiL), 2–10,
KOJIMACHI, CHIYODA, TOKYO 102-0083, JAPAN
E-mail address: `takashi.kato@mathfi-lab.com`