Theoretical and numerical analysis of an optimal execution problem with uncertain market impact

Kensuke Ishitani
Takashi Kato
THEORETICAL AND NUMERICAL ANALYSIS OF AN OPTIMAL EXECUTION PROBLEM WITH UNCERTAIN MARKET IMPACT

KENSUKE ISHITANI AND TAKASHI KATO*

Abstract. This paper is a continuation of [8], in which we derived a continuous-time value function corresponding to an optimal execution problem with uncertain market impact as the limit of a discrete-time value function. Here, we investigate some properties of the derived value function. In particular, we show that the function is continuous and has the semigroup property, which is strongly related to the Hamilton–Jacobi–Bellman quasi-variational inequality. Moreover, we show that noise in market impact causes risk-neutral assessment to underestimate the impact cost. We also study typical examples under a log-linear/quadratic market impact function with Gamma-distributed noise.

1. Introduction and the Model

In [8], we derive a continuous-time value function corresponding to an optimal execution problem with uncertain market impact (MI) as a limit of a discrete-time value function. In this paper, we study some mathematical properties of the value function, and give an interpretation from the point of view of mathematical finance. First, we recall the continuous-time value function derived in [8]. Denote by \( C \) the set of non-decreasing, non-negative, and continuous functions \( u \) on \( D := \mathbb{R} \times [0, \Phi_0] \times [0, \infty) \), with \( \Phi_0 > 0 \) fixed, such that

\[
    u(w, \varphi, s) \leq C_u(1 + |w|^m + s^m), \quad (w, \varphi, s) \in D
\]

for some constants \( C_u, m_u > 0 \). For \( t \in [0, 1] \), \( (w, \varphi, s) \in D \) and \( u \in C \), define

\[
    V_t(w, \varphi, s; u) = \sup_{(\zeta, \varphi), \in A(\varphi)} \mathbb{E}[u(W_t, \varphi_t, S_t)]
\]

subject to

\[
    dW_t = \zeta_t S_t \, dr, \\
    d\varphi_t = -\zeta_t \, dr, \\
    dX_t = \sigma(X_t)dB_t + b(X_t) \, dr - g(\zeta_t) \, dL_t, \\
    S_t = \exp(X_t)
\]

Received 2015-3-14; Communicated by K. Saitô.
2010 Mathematics Subject Classification. Primary 91G80; Secondary 93E20, 49L20.
Key words and phrases. Optimal execution, market impact, liquidity uncertainty, Lévy process.
* This work was supported by a grant-in-aid from the Zengin Foundation for Studies on Economics and Finance.

343
and \((W_0, \varphi_0, S_0) = (w, \varphi, s)\), where \((B_t)_{0 \leq t \leq 1}\) is a standard one-dimensional Brownian motion defined on a complete probability space \((\Omega, \mathcal{F}, P)\) and \((L_t)_{0 \leq t \leq 1}\) is a one-dimensional non-decreasing Lévy process (subordinator) defined on the same probability space. (Note that \(V_0(w, \varphi, s; u) = u(w, \varphi, s)\).) Assume that \((B_t)_t\) and \((L_t)_t\) are independent. Further assume that \(\sigma, b : \mathbb{R} \to \mathbb{R}\) are Lipschitz continuous bounded functions satisfying
\[
|\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq K|x - y|, \quad |\sigma(x)| + |b(x)| \leq K, \quad x, y \in \mathbb{R}
\] (1.4)
for some \(K > 0\), and \(g : [0, \infty) \to [0, \infty)\) is a function defined by
\[
g(\zeta) = \int_0^\zeta h(\zeta')d\zeta',
\]
where \(h : [0, \infty) \to [0, \infty)\) is a non-decreasing continuous function. \(\mathcal{A}_t(\varphi)\) is the set of \((\mathcal{F}_t)_{0 \leq t \leq 1}\)-adapted and caglad processes (i.e., those that are left-continuous with finite right-limit for arbitrary values of \(r\)) \(\zeta = (\zeta_r)_{0 \leq r \leq t}\) such that \(\zeta_r \geq 0\) for each \(r \in [0, t]\), \(\int_0^r \zeta_r dr \leq \varphi\) almost surely, and
\[
||\zeta||_\infty := \sup_{(r, \omega) \in [0, t] \times \Omega} \zeta_r(\omega) < \infty,
\] (1.5)
where \(\mathcal{F}_r = \sigma\{B_v, L_v; v \leq r\} \cup \{\text{Null sets}\}\). Here, the supremum in (1.5) is taken over all values in \([0, t] \times \Omega\). As noted in [8], we may use the essential supremum in (1.5) in place of the supremum.

We assume that the Lévy measure \(\nu\) of \((L_r)_r\) satisfies
\[
||\nu||_1 + ||\nu||_2 < \infty,
\] (1.6)
where \(||\nu||_p = \left(\int_{(0, \infty)} z^p \nu(dz)\right)^{1/p}\). Note that the Lévy decomposition of \((L_r)_r\) is given by
\[
L_r = \gamma r + \int_0^r \int_{(0, \infty)} zN(dv, dz),
\] (1.7)
where \(\gamma \geq 0\) and \(N(\cdot, \cdot)\) is a Poisson random measure (see, for example, [16, 18]).

Here, we introduce the financial interpretation of these notations. We consider a simple market model in which only two financial assets are traded: cash and a security. Assume that a single trader is to sell (liquidate) the owned shares of the security by time \(t\). Also assume that the price of the cash is always 1 (in other words, the risk-free rate is 0) and that the security price fluctuates due to market noise and in response to the trader’s sales. The function \(u\) in \(C\) is regarded as the trader’s utility function. With this, \(V_t(w, \varphi, s; u)\) is the supremum of the expected utility of the trader with initial cash amount \(w\), initial shares \(\varphi \in [0, \Phi_0]\), and initial security price \(s\). Here, \(\Phi_0 > 0\) denotes an upper bound of \(\varphi\) and can be arbitrarily chosen; \((\zeta_r)_{0 \leq r \leq t}\) denotes the trader’s execution strategy; and \(\zeta_r\) denotes the execution speed at time \(r\). The trader chooses an admissible execution strategy from \(\mathcal{A}_t(\varphi)\) to optimize the expected utility of the triplet \((W_t, \varphi_t, S_t)\), where \(S_t\) describes the security price at time \(t\) and \(X_t\) is its log-price; \(W_t\) denotes the cash amount at time \(t\); and \(\varphi_t\) denotes the shares of the security at time \(t\). The fluctuation of the triplet \((W_t, \varphi_t, S_t)_{0 \leq t \leq t}\) is characterized by the differential
equations in (1.3). \((B_r)_r\) represents the component of the market noise reflected in fluctuation of the security price. The term

\[ g(\zeta_r) dL_r = \gamma g(\zeta_r) dr + g(\zeta_r) \int_{(0,\infty)} z N(dr, dz) \]  

(1.8)

describes the (infinitesimal) MI of the trader’s selling with speed \(\zeta_r\). \(\gamma\) (resp., \(g\)) denotes the magnitude (resp., shape) of the MI. Because \(g\) is non-decreasing and convex, the MI becomes huge when \(\zeta_r\) is large. The last term in the right-hand side of (1.8) indicates the effect of noise in the MI, which is mathematically described by the jump of \((L_r)_r\).

In this paper, we study some properties of the continuous-time value function \(V_t(w, \varphi, s; u)\). We find that the value function is continuous in \((w, \varphi, s) \in D\) and \(t > 0\). In addition, right-continuity at \(t = 0\) depends on the state of \(h(\infty) := \lim_{\zeta \to \infty} h(\zeta)\). In particular, noise in the MI does not affect the continuity of the value function. We also show that the Bellman principle (the semi-group property) holds and perform a comparison with the result in the case of a deterministic MI, which was studied in [10], and show that noise in the MI causes risk-neutral assessment to underestimate the MI cost. This means that a trader who attempts to minimize the expected liquidation cost is not sensitive enough to uncertainty in the MI. Last, we present generalizations of the examples from [10] and investigate the effects of noise in the MI on the optimal strategy of a trader, by numerical experiments. We consider a risk-neutral trader’s execution problem with a log-linear/quadratic MI function with Gamma-distributed noise.

The rest of this paper is organized as follows. In Section 2, we present our results on the properties of the value function. In Section 3, we consider the case where the trader must sell all shares of the security, which is referred to as the “sell-off condition.” We also study the optimization problem under the sell-off condition and show that the results in [10, Sect. 4] also hold in our model. Section 4 compares deterministic MIs with random (stochastic) MIs in a risk-neutral framework. In Section 5, we present some examples based on the proposed model. We conclude this paper in Section 6. All proofs are in Section 7.

2. Properties of Value Functions

Regarding the continuity of the continuous-time value function, we have the following theorem:

**Theorem 2.1.** Let \(u \in \mathcal{C}\).

(i) If \(h(\infty) = \infty\), then \(V_t(w, \varphi, s; u)\) is continuous in \((t, w, \varphi, s) \in [0, 1] \times D\).

(ii) If \(h(\infty) < \infty\), then \(V_t(w, \varphi, s; u)\) is continuous in \((t, w, \varphi, s) \in [0, 1] \times D\) and \(V_t(w, \varphi, s; u)\) converges to \(Ju(w, \varphi, s)\) uniformly on any compact subset of \(D\) as \(t \downarrow 0\), where \(Ju(w, \varphi, s)\) is given as

\[
\begin{cases}
\sup_{\psi \in [0, \varphi]} u\left(w + \frac{1-e^{-\gamma h(\infty)}}{\gamma h(\infty)} s, \varphi - \psi, s e^{-\gamma h(\infty)\psi}\right) & (\gamma h(\infty) > 0), \\
\sup_{\psi \in [0, \varphi]} u\left(w + \psi s, \varphi - \psi, s\right) & (\gamma h(\infty) = 0).
\end{cases}
\]
Remark 2.2.

(i) The assertions of Theorem 2.1 are also quite similar to the result in [10], which showed that continuities in \( w, \phi, \) and \( s \) of the value function are always guaranteed, but continuity in \( t \) at the origin depends on the state of the function \( h \) at infinity. When \( h(\infty) = \infty \), MI for large sales is sufficiently strong \((g(\zeta) \text{ diverges rapidly with } \zeta \to \infty)\) to prevent the trader from performing instant liquidation: an optimal policy is "no trading" in infinitesimal time, and thus \( V_t \) converges to \( u \) as \( t \downarrow 0 \). When \( h(\infty) < \infty \), the value function is not always continuous at \( t = 0 \) and has the right limit \( Ju(w, \phi, s) \). In this case, MI for large sales is not particularly strong \((g(\zeta) \text{ still diverges, although with low divergence speed})\) and there is room for liquidation within infinitesimal time. The function \( Ju(w, \phi, s) \) corresponds to the utility of liquidation by the trader, who sells part of the shares of a security \( \psi \) by dividing it infinitely within an infinitely short time (sufficiently short that the fluctuation in the price of the security can be ignored) and obtains an amount \( \phi - \psi \); that is,

\[
\zeta^\delta_r = \frac{\psi}{\delta} 1_{[0,\delta]}(r), \quad r \in [0,t] \quad (\delta \downarrow 0).
\]  

(ii) Note that the jump part

\[
g(\zeta_r) \int_{(0,\infty)} z N(dr, dz)
\]  

does not change the result. Also note that if \( \gamma = 0 \) and \( h(\infty) < \infty \), then the effect of MI disappears in \( Ju(w, \phi, s) \). This situation may occur even if \( E[c^0] \geq \epsilon_0 \) (or \( E[L_1] \geq \epsilon_0 \)) for some \( \epsilon_0 > 0 \).

Here, we present the Bellman principle (dynamic programming principle or "semi-group" property). Let us define \( Q_t : \mathcal{C} \rightarrow \mathcal{C} \) by \( Q_t u(w, \phi, s) = V_t(w, \phi, s; u) \). Then we can easily show that \( Q_t \) is well defined as a nonlinear operator. The same proof as that for Theorem 3.2 in [10] gives the following proposition:

**Proposition 2.3.** For each \( r, t \in [0,1] \) with \( t + r \leq 1 \), \((w, \phi, s) \in D \) and \( u \in \mathcal{C} \), it holds that \( Q_{t+r} u(w, \phi, s) = Q_t Q_r u(w, \phi, s) \).

**Remark 2.4.** By using the above proposition, we can formally derive the Hamilton–Jacobi–Bellman (HJB) equation corresponding to our value function on the generalized domain of the utility function \( \tilde{D} = \mathbb{R} \times [0,\infty) \times [0,\infty) \):

\[
\frac{\partial}{\partial t} V_t(w, \phi, s; u) - \sup_{\zeta \geq 0} \mathcal{L}^\zeta V_t(w, \phi, s; u) = 0
\]  

(2.3)
with the same boundary conditions as (3.5) in [10], where

\[
\mathcal{L}^\xi v(t, w, \varphi, s) = \mathcal{L}^\xi v(t, w, \varphi, s) - \mathcal{L}^\xi v(t, w, \varphi, s),
\]

\[
\mathcal{L}^\xi v(t, w, \varphi, s) = \frac{1}{2} \sigma(s)^2 \frac{\partial^2}{\partial s^2} v(t, w, \varphi, s) + b(s) \frac{\partial}{\partial s} v(t, w, \varphi, s)
\]

\[
+ \zeta \left( s \frac{\partial}{\partial w} v(t, w, \varphi, s) - \frac{\partial}{\partial \varphi} v(t, w, \varphi, s) \right) - \gamma g(\zeta) s \frac{\partial}{\partial s} v(t, w, \varphi, s),
\]

\[
\mathcal{L}^\xi v(t, w, \varphi, s) = \int_{(0, \infty)} \left\{ v(w, \varphi, s) - v(w, \varphi, se^{-g(z)}) \right\} v(dz).
\]

(2.3) is a partial integro-differential equation (PIDE). When \( \mathcal{L}^\xi \equiv 0 \), that is, when there is no jump, characterization of our value function as the unique viscosity solution of (2.3) is studied by [10] under some additional technical conditions. Showing these properties in the general case is a more challenging task. Here we introduce some related literature in place of presenting a detailed argument on the solvability of (2.3): in [7], the existence (i.e., characterization of a value function as a viscosity solution) and uniqueness of the solution of the HJB equation corresponding to the optimal investment/consumption problem with durability and local substitution in the Lévy version of the Black–Scholes-type market model is studied. Reference [20] shows existence and uniqueness of a solution to the Hamilton–Jacobi–Bellman quasi-variational inequalities (HJBQVIs) appearing in combined impulse and (regular) stochastic control problems with jump diffusions (existence in this case is also introduced in [15] without detailed technical arguments). In [3], by means of the weak dynamic programming principle, the characterization of a value function of stochastic control problems under Lévy processes with finite Lévy measure, which arises as a discontinuous viscosity solution of the corresponding HJB equation, is studied. The strong comparison principle (which is closely related to the uniqueness of viscosity solutions) for second-order non-linear PIDEs on a bounded domain is studied in [5].

### 3. Sell-Off Condition

In this section, we consider the optimal execution problem under the “sell-off condition” introduced in [10]. A trader has a certain quantity of shares of a security at the initial time, and must liquidate all of them by the time horizon. Then, the space of admissible strategies is reduced to

\[
\mathcal{A}_t^{SO}(\varphi) = \left\{ (\zeta_r)_r \in \mathcal{A}_r(\varphi) : \int_0^t \zeta_r dr = \varphi \right\}.
\]

We define a value function with the sell-off condition by

\[
V_t^{SO}(w, \varphi, s; U) = \sup_{(\zeta_r)_r \in \mathcal{A}_t^{SO}(\varphi)} E[U(W_t)]
\]

for a continuous, non-decreasing and polynomial growth function \( U : \mathbb{R} \rightarrow \mathbb{R} \).

The following theorem is analogous to Theorem 4.1 in [10] (we omit the proof because it is nearly identical):

**Theorem 3.1.** \( V_t^{SO}(w, \varphi, s; U) = V_t(w, \varphi, s; u) \), where \( u(w, \varphi, s) = U(w) \).
By Theorem 3.1, we see that the sell-off condition does not introduce changes in the value of the value function in a continuous-time model.

Analogously to Theorem 4.2 in [10], a similar result to Theorem 3 in [13] holds when \( g(\zeta) \) is linear:

**Theorem 3.2.** Assume \( g(\zeta) = \alpha_0 \zeta \) for \( \alpha_0 > 0 \).

(i) \( V_S^{SO}(w, \varphi; s; U) = \overline{V}_t^\varphi \left( w + \frac{1 - e^{-\gamma_0\varphi}}{\gamma_0} s, e^{-\gamma_0\varphi} s; U \right) \), where

\[
\overline{V}_t^\varphi(w, \tilde{s}; U) = \sup_{(\overline{S}_r, \overline{W}_r) \in \overline{A}_t(\varphi)} E[U(\overline{W}_t)]
\]

s.t. \( \overline{S}_r = e^{-\gamma_0 \varphi \overline{W}_r} \overline{S}_0 + e^{-\gamma_0 \varphi \overline{W}_r} \sigma(\overline{W}_r)dB_r \)

\( \overline{W}_r = \frac{e^{\gamma_0 \varphi \overline{W}_r} - 1}{\gamma_0} \overline{S}_r \),

\( \overline{S}_0 = \tilde{s}, \overline{W}_0 = \tilde{w} \)

and

\[
\overline{A}_t(\varphi) = \left\{ \frac{\varphi - \int_0^r \zeta dv}{0 \leq r \leq t} ; (\zeta_r)_{0 \leq r \leq t} \in A^S_t(\varphi) \right\}.
\]

\[ G_r = \int_0^r \int_{(0, \infty)} (1 - e^{-\alpha_0 \zeta s})N(ds, dz). \]

(ii) If \( U \) is concave and \( \overline{b}(s) \leq 0 \) for \( s \geq 0 \), then

\[ V_S^{SO}(w, \varphi; s; U) = U \left( w + \frac{1 - e^{-\gamma_0 \varphi}}{\gamma_0} s \right). \]  (3.1)

The proof is in Section 7.2. Note that the assertion (ii) is the same as Theorem 3 in [13], and in this case we can also obtain the explicit form of the value function. The right side of (3.1) is equal to \( J_u(w, \varphi, s) \) for \( u(w, \varphi, s) = U(w) \) and the nearly optimal strategy for \( V_S^{SO}(w, \varphi; s; U) = V_t(w, \varphi; s; U) \) is given by (2.1). This implies that when considering a linear MI function, a risk-averse (or risk-neutral) trader’s optimal liquidation strategy with negative risk-adjusted drift is nearly the same as block liquidation (i.e., selling all shares at once) at the initial time.

### 4. Effect of Uncertainty in MI in the Risk-neutral Framework

The purpose of this section is to investigate how noise in the MI function affects the trader. Particularly, we focus on the case where the trader is risk-neutral, that is, \( u(w, \varphi, s) = u_{RN}(w, \varphi, s) = w \). Note that such a risk-neutral setting is a typical and standard assumption in the study of the execution problem (see e.g. [1, 4, 9, 11, 12, 14, 19]).

First, we prepare a value function of the execution problem with a deterministic MI function to compare with the case of random MI. Let \( V_t(w, \varphi; s; U) \) be the same as in (1.2) by replacing \( g(\zeta) \) and \( L_t \) with \( \tilde{g}(\zeta) \) and \( t \), that is, the SDE for \((X_r)_r\) is given as

\[
\frac{dX_r}{\sigma(X_r)dB_r + b(X_r)dr - \tilde{g}(\zeta_r)dr},
\]
where
\[ \bar{\gamma} = E[L_1] = \gamma + \int_{(0,\infty)} z\nu(dz). \] (4.1)
The following proposition is proved in Section 7.3:

**Proposition 4.1.** We have
\[ V_t(w, \varphi, s; u_{RN}) \geq \bar{V}_t(w, \varphi, s; u_{RN}). \] (4.2)

This proposition shows that noise in MI is welcome because it decreases the liquidation cost for a risk-neutral trader.

For instance, we consider a situation where the trader estimates the MI function from historical data and tries to minimize the expected liquidation cost. Then, a higher sensitivity of the trader to the volatility risk of MI results in a lower estimate for the expected proceeds of the liquidation. This implies that accommodating the uncertainty in MI makes the trader prone to underestimating the liquidation cost. Thus, as long as the trader’s target is the expected cost, the uncertainty in MI is not an incentive for being conservative with respect to the unpredictable liquidity risk. In Section 5, we present the results of numerical experiments conducted to simulate the above phenomenon.

### 5. Examples

In this section, we show two examples of our model, which are both generalizations of the ones in [10].

Motivated by the Black–Scholes-type market model, we assume that \( b(x) \equiv -\mu \) and \( \sigma(x) \equiv \sigma \) for some constants \( \mu, \sigma \geq 0 \) and assume that \( \bar{\mu} := \mu - \sigma^2/2 \) is positive. We also assume a risk-neutral trader with utility function \( u(w, \varphi, s) = u_{RN}(w) = w \). In this case, if there is no MI, then a risk-neutral trader will fear a decrease in the expected stock price, and thus will liquidate all the shares immediately at the initial time.

We consider MI functions that are log-linear and log-quadratic with respect to liquidation speed, and assume Gamma-distributed noise; that is, \( g(\zeta) = \alpha_0 \zeta^p \) for \( \alpha_0 > 0 \) and \( p = 1, 2 \), and \( L_t \) satisfies
\[ P(L_t - \gamma t \in dx) = \Gamma(\alpha_1 t)(\beta_1) x^{\alpha_1 t - 1} e^{-x/\beta_1} \mathbb{1}_{(0,\infty)}(x) dx, \]
where \( \Gamma(x) \) is the Gamma function. Here, \( \alpha_1, \beta_1 \), and \( \gamma > 0 \) are constants. The corresponding Lévy measure is
\[ \nu(dz) = \frac{\alpha_1}{z} e^{-z/\beta_1} \mathbb{1}_{(0,\infty)}(z) dz. \]
Note that for the discrete-time model studied in [8], we can define the corresponding discrete-time MI function as \( g_k^n(\psi) = c_k^n g_n(\psi) \), where \( g_n(\psi) = n^{p-1} \alpha_p \psi^p \) and \( (c_k^n) \) is a sequence of i.i.d. random variables with distribution
\[ P(c_k^n - \gamma \in dx) = \Gamma(\alpha_1/n, n\beta_1)(dx). \]
In each case, assumptions [A], [B1]–[B3], and [C] of [8] are satisfied.
5.1. Log-Linear Impact & Gamma Distribution. In this subsection, we set
$g(\zeta) = \alpha_0 \zeta$ ($p = 1$). Theorem 3.2 directly implies the following:

**Theorem 5.1.** We have

$$V_t(w, \varphi; s; u_{RN}) = w + \frac{1 - e^{-\gamma \alpha_0 \varphi}}{\gamma \alpha_0} s$$  \hspace{1cm} (5.1)

for each $t \in (0, 1]$ and $(w, \varphi, s) \in D$.

The implication of this result is the same as in [10]: the right side of (5.1) is
equal to $J u(w, \varphi, s)$ and converges to $w + \varphi s$ as $\alpha_0 \downarrow 0$ or $\gamma \downarrow 0$, which is the profit
gained by choosing the execution strategy of block liquidation at $t = 0$. Therefore,
the optimal strategy in this case is to liquidate all shares by dividing infinitely
within an infinitely short time at $t = 0$ (we refer to such a strategy as a nearly
block liquidation at the initial time). Note that the jump part of MI (2.2) does
not influence the value of $V_t(w, \varphi; s; u_{RN})$.

5.2. Log-Quadratic Impact & Gamma Distribution. Next we study the
case of $g(\zeta) = \alpha_0 \zeta^2$ ($p = 2$). In [10], we obtained a partial analytical solution to
the problem: when $\varphi$ is sufficiently small or large, we obtain the explicit form of
optimal strategies. However, the noise in MI complicates the problem, and deriving
the explicit solution is more difficult. Thus, we rely on numerical simulations.
Under the assumption that the trader is risk-neutral, we can assume that an
optimal strategy is deterministic. Here, we introduce the following additional
condition:

$[D] \quad \gamma \geq \alpha_1 \beta_1 / 8$.

In fact, we can replace our optimization problem with the deterministic control
problem

$$f(t, \varphi) = \sup_{(\zeta_r), r} \int_0^t \exp \left( - \int_0^r q(\zeta_r) dv \right) \zeta_r dr$$

for a deterministic process $(\zeta_r), r$ under the above assumption, where

$$q(\zeta) = \bar{\mu} + \hat{g}(\zeta),$$

$$\hat{g}(\zeta) = \gamma \alpha_0 \zeta^2 + \alpha_1 \log(\alpha_0 \beta_1 \zeta^2 + 1).$$

This gives the following theorem:

**Theorem 5.2.** $V_t(w, \varphi; s; u_{RN}) = w + sf(t, \varphi)$ under $[D]$.

This theorem is obtained by a similar proof to Proposition 5.1 in [10] by using
the following Laplace transform of the Gamma distribution:

$$E[e^{-\lambda \zeta}] = \exp \left( -\gamma \lambda - \frac{\alpha_1}{n} \log(n \beta_1 \lambda + 1) \right).$$

From Theorem 5.2 and (2.3), we derive the HJB equation for the function $f$ as

$$\frac{\partial}{\partial t} f + \bar{\mu} f - \sup_{\zeta \geq 0} \left\{ \zeta \left( 1 - \frac{\partial}{\partial \varphi} f \right) - \hat{g}(\zeta) f \right\} = 0$$  \hspace{1cm} (5.2)
with the boundary condition
\[ f(0, \varphi) = f(t, 0) = 0. \] (5.3)
When \( \gamma \geq \alpha_1 / 2 \), the function \( \bar{g} \) becomes convex, so we can apply Theorems 3.3 and 3.6 in [10] to show the following proposition:

**Proposition 5.3.** Assume \( \gamma \geq \alpha_1 / 2 \). Then \( f(t, \varphi) \) is the viscosity solution of (5.2). Moreover, if \( \bar{f} \) is a viscosity solution of (5.2) and (5.3) and has a polynomial growth rate, then \( f = \bar{f} \).

It is difficult to obtain an explicit form of the solution of (5.2) and (5.3). Instead, we solve this problem numerically by considering the deterministic control problem \( f^n_{[nt]}(\varphi) \) in the discrete-time model for a sufficiently large \( n \):

\[
f^n_k(\varphi) = \sup_{(\psi^n_{0:k-1}) \in [0,\varphi]^k} \sum_{l=0}^{k-1} \psi^n_l \exp \left( -\mu \times \frac{l}{n} - \sum_{m=0}^{l} I_m \right),
\]

\[ I_m = n \gamma \alpha_0 (\psi^n_m)^2 + \frac{\alpha_1}{n} \log(n^2 \alpha_0 \beta_1 (\psi^n_m)^2 + 1). \]

Note that the convergence \( \lim_{n \to \infty} f^n_{[nt]}(\varphi) = f(t, \varphi) \) is guaranteed by Theorem 2.3 of [8]. We set each parameter as follows: \( \alpha_0 = 0.01, t = 1, \mu = 0.05, w = 0, s = 1, \) and \( n = 500 \). We examine three patterns for \( \varphi \), \( \varphi = 1, 10, \) and 100.

### 5.2.1. The case of fixed \( \gamma \)
In this subsection, we set \( \gamma = 1 \) to examine the effects of the shape parameter \( \alpha_1 \) of the noise in MI. Here, we also set \( \beta_1 = 2 \). As seen in the numerical experiment in [10], the forms of optimal strategies vary according to the value of \( \varphi \). Therefore, we summarize our results separately for each \( \varphi \).

Figure 1 shows graphs of the optimal strategy \( (\zeta_t) \) and its corresponding process \( (\varphi_t) \) of the security holdings in the case of \( \varphi = 1 \), that is, the number of initial shares of the security is small. As found in [10], if there is no noise in the MI function (i.e., if \( \alpha_1 = 0 \)), then the optimal strategy is to sell the entire amount at the same speed (note that the roundness at the corner in the left graph of Figure 1 represents the discretization error and is not essential). The same tendency is found in the case of \( \alpha_1 = 1 \), but in this case the execution time is longer than in the case of \( \alpha_1 = 0 \). When we take \( \alpha_1 = 3 \), the situation is completely different. In this case, the optimal strategy is to increase the execution speed as the time horizon approaches.

When the amount of the security holdings is 10, which is larger than in the case of \( \varphi = 1 \), the optimal strategy and the corresponding process of the security holdings are as shown in Figure 2. In this case, a trader’s optimal strategy is to increase the execution speed as the end of the trading time approaches, which is the same as in the case of \( \varphi = 1 \) with \( \alpha_1 = 3 \). Clearly, a larger value of \( \alpha_1 \) corresponds to a higher speed of execution closer to the time horizon. We should add that a trader cannot complete the liquidation when \( \alpha_1 = 3 \). However, as mentioned in Section 3, we can choose a nearly optimal strategy from \( A^{SO}_1(\varphi) \) without changing the value of the expected proceeds of liquidation by combining the execution strategy in Figure 2 (with \( \alpha_1 = 3 \)) and the terminal (nearly) block liquidation. See Section 5.2 of [10] for details.
When the amount of the security holdings is too large, as in the case of $\varphi = 100$, a trader cannot complete the liquidation regardless of the value of $\alpha_1$, as Figure 3 shows. This is similar to the case of $\varphi = 10$ with $\alpha_1 = 3$. The remaining amount of shares of the security at the time horizon is larger for larger noise in MI. Note that the trader can also sell all the shares of the security without decreasing the profit by combining the strategy with the terminal (nearly) block liquidation.

Figure 1. Result for $\varphi = 1$ in the case of fixed $\gamma$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$.

Figure 2. Result for $\varphi = 10$ in the case of fixed $\gamma$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$.

Figure 3. Result for $\varphi = 100$ in the case of fixed $\gamma$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$. 
5.2.2. The case of fixed $\hat{\gamma}$. In the above subsection, we presented a numerical experiment performed to compare the effects of the parameter $\alpha_1$ by fixing $\gamma$. Here, we perform numerical comparison from a different viewpoint.

The results in Section 4 imply that accounting for the uncertainty in MI will cause a risk-neutral trader to be optimistic about the estimation of liquidity risks. To obtain a deeper insight, we investigate the structure of the MI function in more detail. In Theorems 2.1(ii) and 5.1, the important parameter is $\gamma$, which is the infimum of $L_1$ and is smaller than or equal to $E[L_1]$. We can interpret this as a characteristic feature whereby the (nearly) block liquidation eliminates the effect of positive jumps of $(L_t)_t$. However, there is another decomposition of $L_t$ such that

$$L_t = \hat{\gamma}t + \int_0^t \int_{(0,\infty)} z\tilde{N}(dr,dz),$$

where $\hat{\gamma}$ is given by (4.1) and

$$\tilde{N}(dr,dz) = N(dr,dz) - \nu(dz)dr.$$

This representation is essential from the viewpoint of martingale theory. Here, $\tilde{N}(\cdot,\cdot)$ is the compensator of $N(\cdot,\cdot)$ and $\hat{\gamma}$ can be regarded as the “expectation” of the noise in MI. Just for a risk-neutral world (in which a trader is risk-neutral), as studied in Section 4, we can compare our model with the case of deterministic MI functions as in [10] by setting $\hat{\gamma} = 1$. Based on this, we conduct another numerical experiment with a constant value of $\hat{\gamma}$.

Note that in our example

$$\hat{\gamma} = \gamma + \alpha_1\beta_1$$

and

$$\frac{1}{T}\text{Var}\left(\int_0^T \int_{(0,\infty)} z\tilde{N}(dr,dz)\right) = \alpha_1\beta_1^2$$

hold. Here, (5.4) (respectively, (5.5)) corresponds to the mean (respectively, the variance) of the noise in the MI function at unit time. Comparisons in this subsection are performed with the following assumptions: We set the parameters $\beta_1$ and $\gamma$ to satisfy

$$\gamma + \alpha_1\beta_1 = 1, \quad \alpha_1\beta_1^2 = 0.5.$$  

We examine the cases of $\alpha_1 = 0.5$ and 1, and compare them with the case of $\gamma = 1$ and $\alpha_1 = 0$.

Figure 4 shows the case of $\varphi = 1$, where the trader has a small amount of security holdings. Compared with the case in Section 5.2.1, the forms of all optimal strategies are the same; that is, the trader should sell the entire amount at the same speed. The execution times for $\alpha_1 > 0$ are somewhat shorter than for $\alpha_1 = 0$.

Figure 5 corresponds to the case of $\varphi = 10$. The forms of the optimal strategies are similar to the case of $\varphi = 10$, $\alpha_1 = 0, 1$ in Section 5.2.1. Clearly, the speed of execution near the time horizon increases with increasing $\alpha_1$.

The results for $\varphi = 100$ are shown in Figure 6. The forms of the optimal strategies are similar to the case of $\varphi = 100$ in Section 5.2.1. However, in contrast
to the results in the previous subsection, the remaining amount of shares of the
security at the time horizon is smaller for larger $\alpha_1$.

Finally, we investigate the total MI cost introduced in [9] (which is essentially
equivalent to an implementation shortfall (IS) cost [2, 17]):

$$\text{TC}(\varphi) = - \log \frac{V_T(0, \varphi, s)}{\varphi s}.$$  

As noted at the beginning of this section, when the market is fully liquid and there
is no MI, then the total proceeds of liquidating $\varphi$ shares of the security at $t = 0$ are
equal to $\varphi s$. In the presence of MI, however, the optimal total proceeds decrease
to $V_T(0, \varphi, s) = \varphi s \times \exp(-\text{TC}(\varphi))$. Thus, the total MI cost $\text{TC}(\varphi)$ denotes the
loss rate caused by MI in a risk-neutral world.

Figure 7 shows the total MI costs in the cases of $\varphi = 1$ and 10. Here, we omit
the case of $\varphi = 100$ because the amount of shares of the security is too large to
complete the liquidation unless otherwise combining terminal block liquidations
(which may crash the market). In both cases of $\varphi = 1$ and 10, we find that the
total MI cost decreases by increasing $\alpha_1$. Since the expected value $\tilde{\gamma}$ of the noise in
MI is fixed, an increase in $\alpha_1$ implies a decrease in $\gamma$ and $\beta_1$. Risk-neutral traders
seem to be more sensitive to the parameter $\gamma$ than to $\alpha_1$, and thus the trader
can liquidate the security without concern about the volatility of the noise in MI.
Therefore, the total MI cost for $\alpha_1 > 0$ is lower than that for $\alpha_1 = 0$.

6. Concluding Remarks

In this paper, we studied an optimal execution problem with uncertain MI by
using the model derived in [8]. Our main results discussed in Sections 2 and 3 are
almost the same as in [10].

When considering uncertainty in MI, there are two typical barometers of the
“level” of MI: $\gamma$ and $\tilde{\gamma}$. By using the parameter $\gamma$, we can decompose MI into a
deterministic part $\gamma g(\zeta_t) dt$ and a pure jump part $g(\zeta_t) \int_{0,\infty} 2N(dt,dz)$. Then, the
pure jump part can be regarded as the difference from the deterministic MI case
studied in [10]. On the other hand, as mentioned in Sections 4 and 5, the parameter
$\tilde{\gamma}$ is important not only in martingale theory but also in a risk-neutral world.
Studying $\tilde{\gamma}$ also provides some hints about actual trading practices. Regardless of
whether we accommodate uncertainty into MI, it may result in an underestimate
of MI for a risk-neutral trader.

Studying the effects of uncertainty in MI in a risk-averse world is also meaning-
ful. As mentioned in Section 3, when the deterministic part of the MI function is
linear, the uncertainty in MI does not significantly influence the trader’s behavior,
even when the trader is risk-averse. In future work, we will investigate the case of
nonlinear MI.

Explicitly introducing trading volume processes is another important general-
ization. In some studies of the optimization problem of volume-weighted average
price (VWAP) slippage, the trading volume processes are introduced as stochastic
processes. For instance, [6] studies a minimization problem of the tracking error of
VWAP execution strategies (see [11] for a definition of VWAP execution strate-
gies). In [6], a cumulative trading volume process is defined as a Gamma process.
Figure 4. Result for $\varphi = 1$ in the case of fixed $\bar{\gamma}$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$.

Figure 5. Result for $\varphi = 10$ in the case of fixed $\bar{\gamma}$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$.

Figure 6. Result for $\varphi = 100$ in the case of fixed $\bar{\gamma}$. Left: The optimal strategy $\zeta_r$. Right: The amount of security holdings $\varphi_r$. 

\[ \text{OPTIMAL EXECUTION WITH UNCERTAIN MARKET IMPACT} \]
Moreover, [11] treats a generalized Almgren–Chriss model such that a temporary 
MI function depends on instantaneous trading volume processes, and shows that 
an optimal execution strategy of a risk-neutral trader is actually the VWAP exe-
cution strategy. Since a trading volume process is unobservable, we can regard it 
as a source of the uncertainty of MI functions. Therefore, studying the case where 
MI functions are affected by trading volumes is within our focus.

Finally, in our settings the MI function is stationary in time, but in the real 
market the characteristics of MI change according to the time zone. Therefore, it 
is meaningful to study the case where the MI function is not time-homogeneous. 
This is another topic for future work.

7. Proofs

We first recall some lemmas from [8].

**Lemma 7.1.** Let $\Gamma_k (k \in \mathbb{N})$ be sets, $u \in \mathcal{C}$, and let $(W^i(k, \gamma), \varphi^i(k, \gamma), S^i(k, \gamma)) \in \mathcal{D} (\gamma \in \Gamma_k, k \in \mathbb{N}, i = 1, 2)$ be random variables. Assume that

$$\lim_{k \to \infty} \sup_{\gamma \in \Gamma_k} E[|W^1(k, \gamma) - W^2(k, \gamma)|^{m_1} + |\varphi^1(k, \gamma) - \varphi^2(k, \gamma)|^{m_2} + |S^1(k, \gamma) - S^2(k, \gamma)|^{m_3}] = 0$$

and

$$\sum_{i=1}^2 \sup_{\gamma \in \Gamma_k} E[|W^i(k, \gamma)|^{m_4} + (S^i(k, \gamma))^{m_4}] < \infty$$

for some $m_1, m_2, m_3 > 0$ and $m_4 > m_u$, where $m_u$ is as appeared in (1.1). Then we have

$$\lim_{k \to \infty} \sup_{\gamma \in \Gamma_k} |E[u(W^1(k, \gamma), \varphi^1(k, \gamma), S^1(k, \gamma)) - E[u(W^2(k, \gamma), \varphi^2(k, \gamma), S^2(k, \gamma))]| = 0.$$
Lemma 7.2. Let $Z(t; r, s) = \exp(Y(t; r, \log s))$ and $\hat{Z}(s) = \sup_{0 \leq r \leq 1} Z(r; 0, s)$. Then, for each $m > 0$, there is a constant $C_{m,K} > 0$ depending only on $K$ and $m$ such that $E[\hat{Z}(s)^m] \leq C_{m,K}s^m$, where $K > 0$ is a constant appearing in (1.4).

Lemma 7.3. Let $(X_r^{k,i})_{r \in [0,1]}$, $i = 1, 2$, $k \in \mathbb{N}$, be $\mathbb{R}$-valued $(\mathcal{F}_r)_r$-progressive processes satisfying

$$X_r^{k,i} = x^{k,i} + \int_0^r b(X_v^{k,i}) dv + \int_0^r \sigma(X_v^{k,i}) dB_v + F_r^{k,i}, \quad r \in [0, 1],$$

with $x^{k,i} \in \mathbb{R}$ for $i = 1, 2$ and $k \in \mathbb{N}$, where $(F_r^{k,i})_r$ are $(\mathcal{F}_r)_r$-adapted processes of bounded variation, and let $\Pi_k \subset [0, 1]$, $k \in \mathbb{N}$, be Borel sets. Moreover, assume that

(i): $x^{k,1} - x^{k,2} \to 0$, $k \to \infty$,

(ii): $\lim_{k \to \infty} \left\{ D_k^0 + \int_0^1 D_k^dr \right\} = 0$, where

$$D_r^k = E \left[ \sup_{v \in \Pi_k(r)} |F_v^{k,1} - F_v^{k,2}| \right], \quad \Pi_k(r) = ([0, r] \cap \Pi_k) \cup \{r\}.$$

Then it holds that

$$E \left[ \sup_{v \in \Pi_k} |X_v^{k,1} - X_v^{k,2}| \right] \to 0, \quad k \to \infty.$$

Lemma 7.4. Let $t \in [0, 1]$, $\varphi \geq 0$, $x \in \mathbb{R}$, $(\zeta_r)_{0 \leq r \leq t}$, $(\zeta'_r)_{0 \leq r \leq t} \in \mathcal{A}_t(\varphi)$ and suppose $(X_r)_{0 \leq r \leq t}$ (resp., $(X'_r)_{0 \leq r \leq t}$) is given by (1.3) with $(\zeta_r)_r$ (resp., $(\zeta'_r)_r$) and $X_0 = x \leq X'_0$. Suppose $\zeta_r \leq \zeta'_r$ for any $r \in [0, t]$ almost surely. Then $X_r \geq X'_r$ for any $r \in [0, t]$ almost surely.

7.1. Proof of Theorem 2.1. Continuity in $(w, \varphi, s)$ can be easily proved in the same manner as in the previous study [10], so we focus on the continuity in $t$ (uniformly on any compact subset of $D$).

First of all, we prove the following lemma:

Lemma 7.5. Assume $h(\infty) = \infty$. Then, for any $t \in [0, 1]$, $\varphi \in [0, \Phi_0]$, and $(\zeta_r)_{0 \leq r \leq t} \in \mathcal{A}_t(\varphi),$

$$E \left[ \int_0^r \exp \left( -\int_0^v g(\zeta_w) dL_w \right) \zeta_v dv \right] \leq \phi(r), \quad r \in [0, t],$$

where $\phi(r)$, $r \in (0, 1]$ is a continuous function depending only on function $h(\zeta)$ and $\Phi_0$, such that $\lim_{t \to 0} \phi(t) = 0$.

Proof of Lemma 7.5. We may assume that $\tilde{\gamma} > 0$. Let $\tau_r = \int_0^r g(\zeta_v) dL_v$ and $\tau_R = \inf\{v \in [0, t] : \pi_v > R\} \land t$ for $r \in (0, t]$ and $R > 0$. Since $(\pi_v)_v$ is nondecreasing and $(\exp(-\pi_v)\zeta_v)_v$ is left-continuous, we have that

$$E \left[ \int_0^k \exp(-\pi_v) \zeta_v dv \right] \leq E \left[ \int_0^\tau \exp(-\pi_v) \zeta_v dv \right] = \frac{1}{\tilde{\gamma}} E \left[ \int_0^{(\tau_R + \epsilon) \land \tau} \zeta_v dL_v \right] + \frac{e^{-R}}{\tilde{\gamma}} E \left[ \int_0^{(\tau_R + \epsilon) \land \tau} \zeta_v dL_v \right]$$

(7.2)
holds for \( r \in (0, \epsilon], \ R > 0 \) and \( \epsilon > 0 \). Using the left-continuity of \((\zeta_\nu)_\nu\), we obtain
\[
\frac{e^{-R}}{\gamma} E \left[ \int_0^r \zeta_\nu dL_\nu \right] \leq \frac{e^{-R}}{\gamma} E \left[ \int_0^r \zeta_\nu dL_\nu \right] = e^{-R} \int_0^r E[\zeta_\nu] d\nu \leq \Phi_0 e^{-R}.
\]
The first term on the right side of (7.2) is rewritten as
\[
\frac{1}{\gamma} E \left[ \int_0^{(\tau_\nu+\epsilon)^{\wedge} \nu} \zeta_\nu dL_\nu \right] = r E \left[ \int_0^{\tau_\nu(0, \tau_\nu+\epsilon)} dL_\nu \right],
\]
where the inverse function of Lemma B.12 in [10], we define \( \gamma \). Summarizing the above arguments, we arrive at
\[
E \left[ \int_0^r \exp(-\pi_\nu) \zeta_\nu d\nu \right] \leq r g^{-1} \left( \frac{R}{\gamma} \right) + \Phi_0 e^{-R}.
\]
Therefore, if we can find a positive function \( R(r) \) that satisfies
\[
R(r) \rightarrow \infty \text{ and } r g^{-1} \left( \frac{R(r)}{\gamma} \right) \rightarrow 0 \text{ as } r \rightarrow 0,
\]
we complete the proof of (7.1). To construct such an \( R(r) \), mimicking the proof of Lemma B.12 in [10], we define
\[
R(r) = \gamma r g(M(r)), \quad M(r) = f^{-1} \left( \frac{1}{r} \right), \quad f(\zeta) = \sqrt{h \left( \frac{\zeta}{2} \right)}, \quad r > 0,
\]
where the inverse function \( f^{-1}(y) \) is defined in the same manner as \( g^{-1}(y) \). We can easily verify (7.4) by the same arguments as in [10].

The following proposition can be proved by the same proof as Theorem 3.1(ii) in [10] in combination with Lemma 7.5 and Proposition 2.3.

**Proposition 7.6.** Assume \( h(\infty) = \infty \). Then for any compact set \( E \subset D \),
\[
\lim_{\ell \uparrow D} \sup_{(w, \phi, s) \in E} |V_\ell(w, \phi, s; u) - u(w, \phi, s)| = 0.
\]

Next we consider the case where \( h(\infty) < \infty \). Hereinafter, for each \((w, \phi, s) \in D \) and \((\zeta_\nu)_\nu \in A_\nu(\phi)\), we denote by \( \hat{\Xi}(w, \phi, s; \zeta_\nu)_\nu \) the ordered triplet of processes \((W_r, \phi_r, S_r)_{0 \leq r \leq \ell} \) given by the differential equations in (1.3).
Proposition 7.7. Assume $h(\infty) < \infty$. Then for any compact set $E \subset D$ we have
\[
\limsup_{t \downarrow 0} \sup_{(w, \varphi, s) \in E} (J_u(w, \varphi, s) - V_t(w, \varphi, s; u)) \leq 0.
\]

Proof. Take any $t \in (0, 1)$, $(w, \varphi, s) \in E$, and $\psi \in [0, \varphi]$. Set $(\zeta_r)_{r \in \mathbb{R}}$ by
\[
\zeta_r = \psi \left( 0 \leq r \leq t \right),
\]
and let $(W_r, \varphi_r, S_r)_{0 \leq r \leq t} = \Xi_t(w, \varphi, s; (\zeta_r)_r)$ and $X_r = \log S_r$. A standard argument leads us to
\[
\mathbb{E} \left[ \sup_{r \in [0, t]} \exp(X_r) - s \exp\left(-g(\psi/t)L_r\right) \right] \leq C_K s \sqrt{t},
\]
\[
\mathbb{E} \left[ W_t - w - \psi s \int_0^1 \exp\left(-g(\psi/t)L_{tv}\right) dv \right] \leq C_K \psi s \sqrt{t}
\]
for some $C_K > 0$. Thus, using Lemma 7.1, we get
\[
\sup_{(w, \varphi, s) \in E, \psi \in [0, \varphi]} \left\{ I_1((\zeta_r)_r) - V_t(w, \varphi, s; u) \right\} \leq 0, \quad t \downarrow 0, \quad (7.5)
\]
where
\[
I_1((\zeta_r)_r) = \mathbb{E} [u(w + \psi s) \int_0^1 \exp(-g(\psi/t)L_{tv}) dv, \varphi - \psi, s \exp(-g(\psi/t)L_t)].
\]

Next we will show
\[
\sup_{(w, \varphi, s) \in E, \psi \in [0, \varphi]} \left| I_1((\zeta_r)_r) - I_2((\zeta_r)_r) \right| \rightarrow 0, \quad t \downarrow 0, \quad (7.6)
\]
where
\[
I_2((\zeta_r)_r) = \mathbb{E} \left[ u\left(w + \psi s \int_0^1 \exp(-g(\psi/t)\gamma tv) dv, \varphi - \psi, s \exp(-g(\psi/t)\gamma t) \right) \right].
\]

Theorem 9.43.20 in [18] implies
\[
\lim_{t \downarrow 0} \frac{L_t}{t} = \gamma \quad \text{a.s.} \quad (7.7)
\]
Hence, we obtain
\[
\sup_{(w, \varphi, s) \in E, \psi \in [0, \varphi]} \mathbb{E} \left[ \left| \exp(-g(\psi/t)\gamma t) - \exp(-g(\psi/t)L_t) \right| \right]
\]
\[
\leq \mathbb{E} \left[ 1 - \exp\left( tg(\varphi^*/t) \left\{ \gamma - \frac{L_t}{t} \right\} \right) \right] \rightarrow 0, \quad t \downarrow 0,
\]
where we denote $\varphi^* := \sup_{(w, \varphi, s) \in E} \varphi$. Similarly, we obtain
\[
\lim_{t \downarrow 0} \sup_{(w, \varphi, s) \in E, \psi \in [0, \varphi]} \mathbb{E} \left[ \psi s \int_0^1 \left\{ \exp(-g(\psi/t)\gamma tv) - \exp(-g(\psi/t)L_{tv}) \right\} dv \right] = 0.
\]
Thus we get (7.6) by using Lemma 7.1.
We now complete the proof of Proposition 7.7. By the monotonicity of $u(w, \varphi, s)$ (especially in $w$ and $s$) and the inequality $(0 \leq tg(\psi/t) \leq \psi h(\infty))$, we see that

$$I_2((\zeta_r)_r) \geq u(w + F(\psi)s, \varphi - \psi, se^{-\gamma h(\infty)\psi}),$$

where

$$F(\psi) = \int_0^\psi e^{-\gamma h(\infty)\psi dp} = \psi \int_0^1 \exp(-\gamma h(\infty)\psi dv).$$

Therefore,

$$\sup_{(w, \varphi, s) \in E} \left( J u(w, \varphi, s) - V_1(w, \varphi, s; u) \right) \leq \sup_{(w, \varphi, s) \in E} \left( I_2((\zeta_r)_r) - E[u(W_t, \varphi_t, S_t)] \right).$$

(7.8)

Now our assertion is shown immediately from (7.5), (7.6), and (7.8). \hfill \Box

**Proposition 7.8.** Assume $h(\infty) < \infty$. Then for any compact set $E \subset D$,

$$\limsup_{t \downarrow 0} \sup_{(w, \varphi, s) \in E} (V_t(w, \varphi, s; u) - J u(w, \varphi, s)) \leq 0.$$

**Proof.** Take any $t \in (0, 1)$, $(w, \varphi, s) \in E$, and $(\zeta_r)_r \in A_r(\varphi)$. Denote $(W_r, \varphi_r, S_r)_{0 \leq t \leq t} = \Xi_t(w, \varphi, s; (\zeta_r)_r), \ X_r = \log S_r$.

Since $g$ is convex, the Jensen inequality implies

$$\int_0^r g(\zeta_r) dL_v \geq \gamma \int_0^r g(\zeta_r) dv \geq \gamma r g \left( \frac{1}{r} \int_0^r \zeta_r dv \right) = \gamma \int_0^r h(\zeta_r) d\zeta, \quad r \in [0, t],$$

where $\eta_r = \int_0^r \zeta_r dv$. Then we have

$$u \left( w + s \int_0^t \zeta_r \exp \left( - \int_0^r g(\zeta_r) dL_v \right) dr, \varphi - \eta_t, se^{-\int_0^r g(\zeta_r) dL_v} \right) \leq u \left( w + s \int_0^t \zeta_r \exp \left( - \gamma \int_0^r h(\zeta_r) d\zeta \right) dr, \varphi - \eta_t, se^{-\gamma \int_0^t h(\zeta_r) d\zeta} \right).$$

(7.9)

As in the proof of Proposition 7.7, we get

$$E \left[ \sup_{r \in [0, t]} \left| \exp(X_r) - s \exp \left( - \int_0^r g(\zeta_r) dL_v \right) \right| \right] \leq C_K s \sqrt{t},$$

(7.10)

$$E \left[ \left| W_t - w - s \int_0^t \zeta_r \exp \left( - \int_0^r g(\zeta_r) dL_v \right) dr \right| \right] \leq C_K \Phi_0 s \sqrt{t}$$

(7.11)

for some $C_K > 0$. Then we can apply Lemma 7.1 with (7.10) and (7.11) to obtain

$$\sup_{(w, \varphi, s) \in E} \left| E \left[ u \left( w + s \int_0^t \zeta_r \exp \left( - \int_0^r g(\zeta_r) dL_v \right) dr, \varphi - \eta_t, se^{-\int_0^r g(\zeta_r) dL_v} \right) \right] \right| - E[u(W_t, \varphi_t, S_t)] \rightarrow 0, \ \text{as} \ t \downarrow 0.$$  

(7.12)
We can also see that
\[
\sup_{r \in [0,t]} \left| \exp \left( -\gamma \int_0^r h(\zeta/r) d\zeta \right) - e^{-\gamma h(\infty)\eta_r} \right| \leq 2\gamma \varepsilon_t, \tag{7.13}
\]
\[
\left| E \left[ \int_0^t \zeta_r \left\{ \exp \left( -\gamma \int_0^r h(\zeta/r) d\zeta \right) - e^{-\gamma h(\infty)\eta_r} \right\} dr \right] \right| \leq 2\gamma \Phi_0 \varepsilon_t, \tag{7.14}
\]
where \(\varepsilon_t = \int_0^{\Phi_0} (h(\infty) - h(\zeta/t)) d\zeta (\rightarrow 0, \ t \downarrow 0)\). Applying Lemma 7.1 again with (7.13) and (7.14), we have that
\[
\sup_{(w, \varphi, s) \in E} \left\{ E \left[ u \left( w + s \int_0^t \zeta_r \exp \left( -\gamma \int_0^r h(\zeta/r) d\zeta \right) dr, \varphi - \eta_t, se^{-\gamma \int_0^r h(\zeta/t) d\zeta} \right) \right] - E \left[ u \left( w + s \int_0^t \zeta_r e^{-\gamma h(\infty)\eta_r} dr, \varphi - \eta_t, se^{-\gamma h(\infty)\eta_r} \right) \right] \rightarrow 0, \ as \ t \downarrow 0. \tag{7.15}
\]
Moreover, from the definition of \(Ju(w, \varphi, s)\), we see that
\[
\sup_{(w, \varphi, s) \in E} \left\{ E \left[ u \left( w + s \int_0^t \zeta_r e^{-\gamma h(\infty)\eta_r} dr, \varphi - \eta_t, se^{-\gamma h(\infty)\eta_r} \right) \right] - Ju(w, \varphi, s) \right\} \]
\[
= \sup_{(w, \varphi, s) \in E} \left\{ E \left[ u \left( w + sF(\eta_t), \varphi - \eta_t, se^{-\gamma h(\infty)\eta_t} \right) \right] - Ju(w, \varphi, s) \right\} \leq 0. \tag{7.16}
\]
Combining (7.9), (7.12), (7.15), and (7.16), we obtain our assertion. \(\square\)

Finally, we consider the continuity with respect to \(t \in (0,1]\).

**Proposition 7.9.** Let \(E \subset D\) be a compact set. Then we have the following:
(i) \(\lim_{r \downarrow t} \sup_{(w, \varphi, s) \in E} |V_r(w, \varphi, s; u) - V_t(w, \varphi, s; u)| = 0, \ t \in (0,1]\),
(ii) \(\lim_{r \downarrow t} \sup_{(w, \varphi, s) \in E} |V_r(w, \varphi, s; u) - V_t(w, \varphi, s; u)| = 0, \ t \in (0,1]\).

**Proof.** All we have to do is to show that
\[
JV_t(w, \varphi, s; u) \leq V_t(w, \varphi, s; u), \quad (w, \varphi, s) \in D, \ t \in (0,1) \tag{7.17}
\]
under \(h(\infty) < \infty\), because all the other assertions are obtained in the same way as in the proof of Proposition B.17 in [10] combined with Proposition 2.3 and (7.17).

Take any \(t \in (0,1)\), \((w, \varphi, s) \in D\), \(\psi \in [0, \varphi]\), and \((\zeta_r)_{0 \leq r \leq t} \in A_t(\varphi - \psi)\). Define \((W_R, \varphi, s, \zeta_r)_{0 \leq r \leq t} = \Xi_t(w + F(\psi)s, \varphi - \psi, se^{-\gamma h(\infty)\psi}; (\zeta_r)_{t})\) and \(X_r = \log S_r\). For any \(\delta \in (0, t)\), we define \(\left(\tilde{\zeta}_r\right)_{0 \leq r \leq t} \in A_t(\varphi)\) by \(\tilde{\zeta}_r = (\psi/\delta)1_{[0,\delta]}(L_{r-} + \zeta_r)\). Note that the admissibility of \((\tilde{\zeta}_r)_{t}\) comes from \(L_r \geq \gamma r\). Furthermore, we denote \((\tilde{W}_r, \tilde{\varphi}, \tilde{S}_r)_{0 \leq r \leq t} = \Xi_t(w, \varphi, s; (\tilde{\zeta}_r))\) and \(\tilde{X}_r = \log \tilde{S}_r\).

From the definition, we have that
\[
X_r = \log s + \int_0^r \sigma(X_v) dB_v + \int_0^r b(X_v) dv + F_r^{(1)}(\delta), \quad \tilde{X}_r = \log s + \int_0^r \sigma(\tilde{X}_v) dB_v + \int_0^r b(\tilde{X}_v) dv + F_r^{(2)}(\delta), \quad \text{for } r \in [0, t],
\]
Lemma 7.10. \( \frac{\tilde{u}(\delta)}{\delta} \longrightarrow 1, \quad \delta \downarrow 0 \) a.s.

By the above lemma and (7.7), we have

\[ \frac{L_{\tilde{u}(\delta)}}{\delta} \longrightarrow \gamma, \quad \delta \downarrow 0 \text{ a.s.} \quad (7.22) \]

Then the dominated convergence theorem implies that the first term of the right side of (7.20) also converges to 0 as \( \delta \downarrow 0 \). Now we arrive at

\[ D_t^{(\delta)} + \int_0^t D_r^{(\delta)} dr \longrightarrow 0, \quad \delta \downarrow 0, \]

which immediately implies (7.18) together with Lemma 7.3.
A standard argument with (7.18) gives

\[
E \left[ \sup_{r \in [\delta, t]} |\exp(\tilde{X}_r) - \exp(X_r)|^{1/2} \right] 
\leq (2sC_{1,K})^{1/2} E \left[ \sup_{r \in [\delta, t]} |\tilde{X}_r - X_r|^{1/2} \right] \to 0, \quad \delta \downarrow 0. \quad (7.23)
\]

On the other hand, we see that

\[
E[|W_t - \tilde{W}_t|^{1/2}] \leq J_1 + J_2 + J_3,
\]

where

\[
J_1 = E \left[ \frac{\psi}{\delta} \int_0^\delta \exp(\tilde{X}_r)dr - s \int_0^{\gamma \psi} e^{\gamma h(\infty)\psi dr} \right]^{1/2},
\]

\[
J_2 = E \left[ \int_0^t \zeta_r \exp(\tilde{X}_r) - \exp(X_r)|dr \right]^{1/2},
\]

\[
J_3 = E \left[ \int_0^\delta \zeta_r \exp(\tilde{X}_r) - \exp(X_r)|dr \right]^{1/2}.
\]

Easily we get

\[
J_2 \leq \sqrt{\varphi - \psi} E \left[ \sup_{r \in [\delta, t]} |e^{\tilde{X}_r} - e^{X_r}|^{1/2} \right] \to 0, \quad \delta \downarrow 0,
\]

\[
J_3 \leq (\delta \|\zeta\|_\infty)^{1/2} E \left[ \sup_{r \in [0, \delta]} \{e^{\tilde{X}_r} + e^{X_r}\}^{1/2} \right] \to 0, \quad \delta \downarrow 0
\]

by virtue of (7.23) and Lemma 7.2. As for \( J_1 \), a similar calculation to (7.19) gives

\[
J_1 \leq \sqrt{sC_{1,K}\psi} E \left[ \left( \frac{1}{\delta} - \frac{\bar{u}(\delta)}{\delta} \right)^{1/2} \right] 
+ \sqrt{\psi} E \left[ \frac{1}{\delta} \int_0^\delta \exp(\tilde{X}_r) - s \exp \left( -\frac{\gamma h(\infty)\psi r}{\delta} \right) |dr \right]^{1/2}
\]

\[
\leq \sqrt{sC_{1,K}\psi} E \left[ \left( \frac{1}{\delta} - \frac{\bar{u}(\delta)}{\delta} \right)^{1/2} \right] + \sqrt{s(1 + C_{1,K})\psi} \left\{ A_{1/2}^2 + A_{2/2} \right\}, \quad (7.24)
\]

where

\[
A_1 = \frac{1}{\delta} E \left[ \int_0^\delta \left\{ \int_0^r \sigma(\tilde{X}_v)dB_v + \int_0^r b(\tilde{X}_v)dv + \int_0^r g(\zeta_v)dL_v \right\} dr \right],
\]

\[
A_2 = \frac{1}{\delta} E \left[ \int_0^\delta \int_0^r (g(\tilde{\zeta}_v) - g(\zeta_v))dL_v - \frac{\gamma h(\infty)\psi r}{\delta} |dr \right].
\]

Straightforward calculations lead us to

\[
A_1 \leq \frac{2K}{3} \sqrt{\delta} + \frac{K + \hat{g}(\|\zeta\|_\infty)}{2} \delta. \quad (7.25)
\]
Moreover, by Lemma 7.10 and (7.22), we see that

\[
A_2 \leq \gamma \int_0^\psi (h(\infty) - h(\zeta'/\delta))d\zeta' + \frac{\psi h(\infty)}{\delta} E \left[ \frac{1}{\delta} \int_0^\psi \left( \gamma r - L_{r \wedge \hat{u}(\delta)} \right) dr \right] \\
\leq \gamma \int_0^\psi (h(\infty) - h(\zeta'/\delta))d\zeta' + \psi h(\infty) E \left[ \frac{1}{\delta} \int_0^\psi \left\{ \frac{L_r}{\gamma} - \gamma \right\} dr \right] \\
+ \psi h(\infty) E \left[ \left(1 - \frac{\hat{u}(\delta)}{\delta}\right) \left\{ \gamma \left(1 - \frac{\hat{u}(\delta)}{\delta}\right) + \left(\gamma - \frac{L_{\hat{u}(\delta)}}{\delta}\right) \right\} \right] \rightarrow 0, \ \delta \downarrow 0. \\
(7.26)
\]

Combining Lemma 7.2, Lemma 7.4, (7.24), (7.25), and (7.26), we get \( J_1 \rightarrow 0 \) as \( \delta \downarrow 0 \), hence we arrive at \( \lim_{\delta \downarrow 0} E[|W_t - \tilde{W}_t|^{1/2}] = 0 \). Therefore, by Lemma 7.1 we obtain

\[
E[u(W_t, \varphi_t, \exp(X_t))] - V_t(w, \varphi, s; u) \\
\leq \lim_{\delta \downarrow 0} \left| E[u(W_t, \varphi_t, \exp(X_t))] - E[u(\tilde{X}_t, \tilde{\varphi}_t, \exp(\tilde{X}_t))] \right| = 0.
\]

Since \( (\zeta_t)_{0 \leq t \leq \tau} \in A_t(\varphi - \psi) \) is arbitrary, we get

\[
V_t(w + F(\psi)s, \varphi - \psi, se^{-\gamma h(\infty)\psi}; u) \leq V_t(w, \varphi, s; u).
\]

for an arbitrary \( \psi \in [0, \varphi] \). Now we complete the proof of (7.17).

\[\square\]

**Proof of Lemma 7.10.** We may assume \( \gamma > 0 \). Fix any \( \varepsilon \in (0,1) \) and set \( \varepsilon' = \gamma \varepsilon/(1 - \varepsilon) \). By (7.7), we see that for almost all \( \omega \), there exists a \( \delta_0 = \delta_0(\omega) > 0 \) such that \( L_\delta/\delta < \gamma + \varepsilon' \) for each \( \delta \in (0, \delta_0) \). Let \( \delta_1 = \delta_1(\omega) = (1 + \varepsilon'/\gamma)^{-1}\delta_0 \) and take any \( \delta \in (0, \delta_1) \). Moreover, let \( \delta' = (1 + \varepsilon'/\gamma)^{-1}\delta \). Then we see that \( \delta' < \delta_0 \) and thus \( L_{\delta'} < (\gamma + \varepsilon')\delta' = \gamma \delta' \). By this inequality and the definition of \( \hat{u}(\delta) \), we get \( 1 \geq \hat{u}(\delta)/\delta \geq \delta'/\delta = 1 - \varepsilon \), which implies the assertion.

\[\square\]

**7.2. Proof of Theorem 3.2.** We can confirm assertion (i) by applying Itô’s formula to \( \overline{S}_t \) and \( \overline{W}_t \). By a similar argument to that in Section 7.9 in [10], we obtain

\[
E[U(\overline{W}_t)] \leq U(\overline{w}) + \int_0^t E \left[ \frac{1 - e^{-\gamma a\tilde{\tau}_r}}{\gamma a_0} b(\overline{S}_t e^{\gamma a\tilde{\tau}_r}) \right. \\
\left. - \int_{(0, \infty)} \frac{e^{\gamma a\tilde{\tau}_r} - 1}{\gamma a_0} \overline{S}_r (1 - e^{-\alpha a \zeta_r z}) \nu(dz) \right] dr
\]

for any \((\tilde{\tau}_r)_r \in \overline{A}_t(\varphi)\) by virtue of the Jensen inequality. Since \( \tilde{b} \) is non-positive, the function \( U \) is non-decreasing, and the terms

\[
1 - e^{-\gamma a \tilde{\tau}_r}, \ e^{\gamma a \tilde{\tau}_r} - 1, \ 1 - e^{-\alpha a \zeta_r z}
\]

are all non-negative, we see that \( E[U(\overline{W}_t)] \leq U(\overline{w}) \) for any \((\tilde{\tau}_r)_r \in \overline{A}_t(\varphi)\), which implies \( \overline{V}_t^\phi(\overline{w}, \overline{s}) \leq U(\overline{w}) \). The opposite inequality \( \overline{V}_t^\phi(\overline{w}, \overline{s}) \geq U(\overline{w}) \) is obtained similarly to the result in Section 7.9 in [10]. This completes the proof.
7.3. Proof of Proposition 4.1. The following proposition immediately leads us to (4.2).

Proposition 7.11. $\tilde{V}_k^n(w, \varphi, s; u_{RN}) \geq \tilde{V}_k^n(w, \varphi, s; u_{RN})$, where $\tilde{V}_k^n$ is defined as in [8] and $\tilde{V}_k^n$ is obtained from $V_k^n$ by replacing $c_k^n$ with $\tilde{c}_k^n$.

Proof. We use the notation of [8]. Take any $(\psi_l^n) \in A_k^n(\varphi)$ and let $(W_k^n, \varphi_l^n, S_l^n) = \Xi_k^n(w, \varphi, s; (\psi_l^n)\_l)$ be the triplet for $\tilde{V}_k^n(w, \varphi, s; u_{RN})$. Since $c_k^n$ is independent of $F_l^n$, the Jensen inequality implies

$$E[\tilde{W}_k^n] = w + \sum_{l=0}^{k-1} E[\psi_l^n S_l^n \exp(-E[c_l^n | F_l^n] g_n(\psi_l^n))].$$

$$\leq w + \sum_{l=0}^{k-1} E[\psi_l^n S_l^n \exp(-c_l^n g_n(\psi_l^n)) | F_l^n]] \leq V_k^n(w, \varphi, s; u_{RN}).$$

Since $(\psi_l^n)\_l$ is arbitrary, we obtain the assertion. \qed

Acknowledgment. The authors are grateful to Prof. Tai-Ho Wang (Baruch College, The City University of New York) for helpful comments and discussions on the subject matter. In addition, the authors thank the reviewer for various comments and constructive suggestions to improve the quality of the paper.

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. and Chriss, N.: Optimal execution of portfolio transactions, J. Risk, 3, (2000) 5–39.
3. Bouchard, B. and Touzi, N.: Weak dynamic programming principle for viscosity solutions, SIAM Journal on Control and Optimization, 49(3), (2011) 948–962.
4. Cheng, X. and Wang, T.-H.: Optimal execution with uncertain order fills in Almgren–Chriss framework, SSRN preprint, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2632012, (2015).
5. Ciomaga, A.: On the strong maximum principle for second order nonlinear parabolic integro-differential equations, Advances in Differential Equations, 17, (2012) 635–671.
6. Frei, C. and Westray, N.: Optimal execution of a VWAP order: a stochastic control approach, Mathematical Finance, DOI: 10.1111/mafi.12048, (2013).
7. Holden, N.: Portfolio optimization in a jump-diffusion market with durability and local substitution: A penalty approximation of a singular control problem, Master Thesis, Faculty of Mathematics and Natural Sciences, University of Oslo (2010).
8. Ishitani, K. and Kato, T.: Mathematical formulation of an optimal execution problem with uncertain market impact, Communications on Stochastic Analysis, 9(1), (2015) 113–129.
9. Kato, T.: Optimal execution with a geometric Ornstein-Uhlenbeck price process, arXiv preprint, http://arxiv.org/pdf/1107.1787, (2011).
10. Kato, T.: An optimal execution problem with market impact, Finance and Stochastics, 18(3), (2014) 695–732.
11. Kato, T.: VWAP execution as an optimal strategy, JSIAM Letters, 7, (2015) 33–36.
12. Konishi, H. and Makimoto, N.: Optimal slice of a block trade, Journal of Risk, 3(4), (2001) 33–51.
13. Lions, P.-L. and Lasry, J.-M.: Large investor trading impacts on volatility, Paris-Princeton Lectures on Mathematical Finance 2004, Lecture Notes in Mathematics 1919, Springer, Berlin, (2007) 173–190.
14. Makimoto, N. and Sugihara, Y.: Optimal execution of multiasset block orders under stochastic liquidity, IMES Discussion Paper Series: http://www.imes.boj.or.jp/research/papers/english/10-E-25.pdf, (2010).
15. Øksendal, B. and Sulem, A.: Applied Stochastic Control of Jump Diffusions, Universitext, Springer-Verlag, Berlin, 2005.
16. Papapantoleon, A.: An introduction to Lévy processes with applications in finance, Lecture Notes, TU Vienna, 2008.
17. Perold, A.F.: The implementation shortfall: Paper versus reality, Journal of Portfolio Management, 14, (1988) 4–9.
18. Sato, K.: Lévy processes and infinitely divisible distributions, Cambridge Studies in Advanced Mathematics, 68, Cambridge University Press, Cambridge, 1999.
19. Schied, A. and Zhang, T.: A hot potato game under transient price impact and some effects of a transaction tax, Preprint (2014).
20. Seydel, R.C.: Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions, Stochastic Processes and their Applications, 119(10), (2009) 3719–3748.
21. Subramanian, A. and Jarrow, R.: The liquidity discount, Math. Finance, 11, (2001) 447–474.