Execution of sequential orders with stochastic characteristics

Marek Musiela and Thaleia Zariphopoulou

May 10, 2022

Abstract

We extend the results of [5] to the execution of sequential orders. Orders arrive at known times but their direction (sell, buy, empty) and volume size are random. We study in detail the various kinds of implementation slippages in both the conditional and uncondtional sense, and also establish various properties of the admissible and optimal execution strategies for general random decay functions. We also consider other probabilistic quantities. Once we move beyond the conditional and unconditional means, the dependence of all random characteristics of all orders emerges. This creates various challenges which we discuss in detail. Finally, we study the different kinds of market impacts and discuss issues related to volume time and and to order aggregation.

1 Introduction

TBA

2 The model of sequential orders

We introduce the model for the execution of sequential orders and the objective to be optimized. Throughout, various concepts, definitions and results are extensions of their single-order counterparts proposed in [3], and further extended and modified in [5]. To ease the presentation, we only highlight the main ideas and steps of these extensions, focusing on the new elements that emerge from the multiplicity of orders and their stochastic characteristics. To our knowledge, optimal execution beyond a single order has only been examined in [2] where the case of two orders was studied.

Let $T_0 = 0 < T_1 < .... < T_n = T$ be a sequence of times known at T_0. Let also $\left(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ be a probability space on which we define a Brownian motion W, and random sequences $(V_0, ..., V_{n-1})$ and $(\varepsilon_0, \varepsilon_1, ..., \varepsilon_{n-1})$ with

$$\varepsilon_i = -1, 0, 1 \quad \text{and} \quad V_i > 0, \quad i = 0, ..., n-1. \tag{1}$$

The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is such that, for $i = 0, ..., n-1$, the random variables ε_i and V_i are $\mathcal{F}_{T_i}-$measurable.

For $i = 0, ..., n-1$, order i arrives at T_i and is represented by the product $\varepsilon_i V_i$ in that, if $\varepsilon_i > 0$ (resp. $\varepsilon_i < 0$), the order is to buy (resp. sell) volume V_i while, if $\varepsilon_i = 0$, the order is an empty one. Each order must be fully executed before the next one arrives. For this, we assume that order i must be executed by some $\hat{T}_{i+1} < T_{i+1}$. Herein, we do not discuss how $\hat{T}_{i+1}$ is chosen; in section 6 we provide some preliminary comments on the effects of the lag $\delta_{i+1} := T_{i+1} - \hat{T}_{i+1}$, $i = 1, ..., n-1$ on market impact. For now, we only assume that, similarly to knowning at T_0 the times $T_0 = 0 < T_1 < .... < T_n = T$, we also know at T_0 the *effective execution times* $\hat{T}_1, ..., \hat{T}_n$, which satisfy $T_0 = 0 < \hat{T}_1 < T_1 < \hat{T}_2 < T_2 < .... < T_{n-1} < \hat{T}_n < T_n = T$.

We further assume that filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is such that it can support a sequence of positive random functions $(G_0, G_1, ..., G_{n-1})$ with G_i being $\mathcal{F}_{T_i}-$ measurable. These functions will play the role of the decay function - introduced in [3] for the single order case - during the execution of their respective order. In analogy to [3], we assume that each G_i, $i = 0, ..., n-1$, is represented as the Fourier transform of a finite positive measure ψ_i,

$$G_i(|t|, \omega) = \int e^{i|z|} d\psi_i(\omega), \quad i = 0, ..., n-1 \text{ and } t \in \mathbb{R}, \tag{2}$$

with ψ_i now being $\mathcal{F}_{T_i}-$ measurable.

The admissibility set of execution strategies for order i, $i = 0, 1, ..., n-1$, is defined as

$$\mathcal{A}_{[T_i,T_{i+1}]} := \left\{X^i : X^i_t,\ t \in [T_i, T_{i+1}] \text{ with } X^i_t \in \mathcal{F}_{T_i},\ \text{monotone,} \right. \tag{3a}$$

$$\left. \text{left continuous with right hand limits, } X^i_{T_i} = 0 \text{ and } X^i_{\hat{T}_{i+1}+} = V_i,\ \hat{T}_{i+1} < T_{i+1} \right\}.$$

We recall that, as in the single order case, the condition $\hat{T}_{i+1} < T_{i+1}$ is also needed in order to properly define the execution strategies within their respective interval, given the above imposed continuity assumptions.

The implementation slippage $IS\left(X^i\right)$ of order i, $i = 0, 1, ..., n-1$, is defined as in the single order case. Similar calculations yield that

$$IS\left(X^i\right) = \int_{\left[T_i, \hat{T}_{i+1}\right]} S_t dX^i_t + \frac{G_i(0)}{2} \sum_{T_i \le t \le \hat{T}_{i+1}} \left(\Delta X^i_t\right)^2 - \varepsilon_i V_i S_{T_i}, \tag{4}$$

where ΔX^i_t denotes possible jumps,

$$\Delta X^i_t := X^i_{t+} - X^i_t,\ t \in \left[T_i, \hat{T}_{i+1}\right].$$

The impacted stock price S_t, $t \in (T_i, T_{i+1}]$, is given (as in [3]) by

$$S_t := S_{T_i} + \int_{[T_i,t)} G_i(t-s)\, dX^i_s + \sigma\left(W_t - W_{T_i}\right), \quad t \in \left(T_i, \hat{T}_{i+1}\right] \tag{5}$$

and

$$S_t = S_{\hat{T}_i+} + \sigma(W_t - W_{\hat{T}_i}), \quad t \in \left(\hat{T}_{i+1}, T_{i+1}\right], \tag{6}$$

where $\sigma > 0$ is a given constant; later on, we discuss the case of stochastic volatility (see section 5.1.1. herein).

The main object of study is the aggregate implementation slippage, defined next.

Definition 1 *The implementation slippage of a sequential execution strategy*

$$X := \left(X^0, X^1, ..., X^{n-1}\right) \quad \textit{with } X^i \in \mathcal{A}_{[T_i, T_{i+1}]},\ i = 0, 1, ..., n-1,$$

is defined as the sum of the individual slippages $IS(X^i)$ (cf. (4)),

$$IS\left(X\right) := \sum_{i=0}^{n-1} IS\left(X^i\right) \tag{7}$$

$$= \sum_{i=0}^{n-1} \left(\frac{1}{2} \int_{[T_i, \hat{T}_{i+1}]} \int_{[T_i, \hat{T}_{i+1}]} G_i\left(|t-s|\right) dX_s^i dX_t^i + \int_{[T_i, \hat{T}_{i+1}]} \sigma\left(W_t - W_{T_i}\right) dX_t^i \right).$$

The objective is to *minimize* the expected aggregate implementation slippage over admissible sequential strategies $X = \left(X^0, X^1, ..., X^{n-1}\right)$ with $X^i \in \mathcal{A}_{[T_i, T_{i+1}]}$, $i = 0, 1, ..., n-1$,

$$E\left(IS\left(X^*\right)\right) = \min_X E\left(IS\left(X\right)\right). \tag{8}$$

When a single order arrives at T_0 to be executed at $\hat{T}_1 < T_1$, the above problem reduces to a deterministic problem which can be solved using Fredholm's alternative, or calculus of variations and optimal transport (see, respectively, [3] and [2]). In this case, the order has no stochastic characteristics, like the first order herein. Still, as we pointed out in detail in [5], there are various open questions about the behavior of the model and how meaningful the above objective is.

When sequential orders arrive at *future times* and, moreover, both *their direction and volume size are random*, the situation becomes considerably more complex. We stress that even the optimal execution of a *single* future order with stochastic characteristics is not a mere extension of its non-stochastic counterpart.

The first step in our analysis is to look at probabilistic properties of the individiual and aggregate slippages. Because of the stochastic characteristics of the orders, we need to study the various quantities in both the conditional and unconditional sense.

2.1 Probabilistic properties of implementation slippages

We start with results on conditional and unconditional properties of the individual slippages $IS(X^i)$, $i = 0, ..., n-1$, and the aggregate $IS\left(X\right)$.

2.2 Implementation slippage of a future single order

Consider a single order, say order i, which will arrive at future time T_i and will be executed by $\hat{T}_{i+1} < T_{i+1}$. For the conditional case, the calculations follow along similar arguments as the ones used in Proposition 3 in [5]. For the unconditional case, the mean follows directly but the variance is more involved.

Proposition 2 *Let $i = 0, 1, ..., n-1$. For each order i, the implementation slippage $IS(X^i)$ (cf. (4)) has the following properties:*

i) $IS(X^i)$ is $\mathcal{F}_{T_{i+1}}$-measurable and its conditional on $\mathcal{F}_{T_i}$ distribution is normal.

ii) Its conditional mean and variance are given by

$$E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) = \frac{1}{2}\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right)dX_s^i dX_t^i \tag{9}$$

and

$$Var\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) = \sigma^2\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}\left(t\wedge s - T_i\right)dX_s^i dX_t^i. \tag{10}$$

The random variables

$$E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) \quad \text{and} \quad U_{i+1} := \int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t - W_{T_i}\right)dX_t^i \tag{11}$$

are uncorrelated.

Proof. The normality of the conditional distribution follows from (4), the fact that T_i and $\hat{T}_{i+1}$ are $\mathcal{F}_{T_0}-$measurable and basic properties of the Brownian motion. To show (9), we use that

$$E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t - W_{T_i}\right)dX_t^i\,|\mathcal{F}_{T_i}\right) = 0.$$

For (10), observe that

$$E\left(\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t - W_{T_i}\right)dX_t^i\right)^2|\mathcal{F}_{T_i}\right)$$

$$= \sigma^2 E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}\left(W_t - W_{T_i}\right)\left(W_s - W_{T_i}\right)dX_t^i dX_s^i\,|\mathcal{F}_{T_i}\right)$$

$$= \sigma^2\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]} E\left(\left(W_t - W_{T_i}\right)\left(W_s - W_{T_i}\right)|\mathcal{F}_{T_i}\right)dX_t^i dX_s^i$$

$$= \sigma^2\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}\left(t\wedge s - T_i\right)dX_t^i dX_s^i.$$

The last statement follows easily. ■

Next, we compute the unconditional mean and variance of $IS\left(X^i\right)$. We also look at its distribution, calculating the characteristic function.

Proposition 3 *The mean and variance of the unconditional distribution of the individual implementation slippage* $IS\left(X^i\right)$ *(cf. ((4)) are given by*

$$E\left(IS\left(X^i\right)\right)=E\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right) \tag{12}$$

$$=\frac{1}{2}E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}G_i\left(|t-s|\right)dX_s^i dX_t^i\right),$$

and

$$Var\left(IS\left(X^i\right)\right)=Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right) \tag{13}$$

$$+E\left(\sigma^2\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}\left(t\wedge s-T_i\right)dX_t^i dX_s^i\right).$$

The characteristic function of $IS\left(X^i\right)$ *is given by*

$$Ee^{itIS\left(X^i\right)}=E\left(E\left(e^{itIS\left(X^i\right)}|\mathcal{F}_{T_i}\right)\right) \tag{14}$$

$$=E\left(e^{itE\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-\frac{1}{2}t^2Var\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)}\right),$$

with $E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)$ *and* $Var\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)$ *as in (9) and (10).*

Proof. *We have*

$$E\left(\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-E\left(IS\left(X^i\right)\right)\right)\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i\right)$$

$$E\left(\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-E\left(IS\left(X^i\right)\right)\right)E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i|\mathcal{F}_{T_i}\right)\right)=0,$$

and (12) follows. To derive (13), observe that

$$Var\left(IS\left(X^i\right)\right)=E\left(IS\left(X^i\right)-E\left(IS\left(X^i\right)\right)\right)^2$$

$$=E\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)+\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i-E\left(IS\left(X^i\right)\right)\right)^2$$

$$=E\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-E\left(IS\left(X^i\right)\right)\right)^2$$

$$+2E\left(\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-E\left(IS\left(X^i\right)\right)\right)\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i\right)$$

$$+E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i\right)^2=E\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)-E\left(IS\left(X^i\right)\right)\right)^2$$

$$+E\left(E\left(\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t-W_{T_i}\right)dX_t^i\right)^2|\mathcal{F}_{T_i}\right)\right)$$

$$= E\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) - E\left(IS\left(X^i\right)\right)\right)^2$$

$$+E\left(\sigma^2 \int_{[T_i,T_{i+1}^-]}\int_{[T_i,T_{i+1}^-]} (t\wedge s - T_i)\, dX_t^i dX_s^i\right)$$

$$= Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right) + E\left(\sigma^2 \int_{[T_i,\hat{T}_{i+1}]}\int_{[T_i,\hat{T}_{i+1}]} (t\wedge s - T_i)\, dX_t^i dX_s^i\right)$$

To derive (14), we use the conditional normality of $IS(X^i)$. ■

The above results show that both the conditional on $\mathcal{F}_{T_i}$ and unconditional probabilistic characteristics of order i depend *solely* on its *own* characteristics ε_i, V_i, and the chosen random decay function G_i. In this sense, they resemble the ones of the single order case in [3]. Moreover, the form of $E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)$ and $EIS\left(X^i\right)$ is not surprising (see [2]) and the $Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right)$ is similar to expression (14) in [5].

Note, however, that the unconditional variance $Var\left(IS(X^i\right)$ is quite different. It consists of two terms, $Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right)$ and $E\left(\sigma^2 \int_{[T_i,\hat{T}_{i+1}]}\int_{[T_i,\hat{T}_{i+1}]} (t\wedge s - T_i)\, dX_t^i dX_s^i\right)$. The latter term resembles **(14)** in [5]. The former term *disappears* in the single order model but not in the sequential orders case.

Furthermore, for a nonempty order we have that $Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right) > 0$, independently of the direction of the trade. Therefore, the arrival of non-empty future orders always *increases the unconditional variance.*

2.3 Aggregate implementation slippage

The aggregate implementation slippage (cf. 7) is calculated at initial time T_0. Its mean follows easily from summing the individual means. Its variance, however, is more complex as the second term in (16) indicates. As we will discuss later on, the "cross-terms" appearing in (16) make the problem rather complex.

Proposition 4 *The mean and variance of the implementation slippage* $IS(X)$ *(cf. (7)) are given, respectively, by*

$$E\left(IS\left(X\right)\right) = \frac{1}{2}\sum_{i=0}^{n-1} E\left(\int_{[T_i,\hat{T}_{i+1}]}\int_{[T_i,\hat{T}_{i+1}]} G_i\left(|t-s|\right) dX_s^i dX_t^i\right) \tag{15}$$

and

$$Var\left(IS\left(X\right)\right) = \sum_{i=0}^{n-1} Var\left(IS\left(X^i\right)\right) + \sum_{i\neq j}\left(E\left(Z_i Z_j\right) + E\left(Z_{i\vee j} U_{i\wedge j+1}\right)\right), \tag{16}$$

where, for $i = 0, ..., n-1$,

$$Z_i := \frac{1}{2}\int_{[T_i,\hat{T}_{i+1}]}\int_{[T_i,\hat{T}_{i+1}]} G_i\left(|t-s|\right) dX_s^i dX_t^i$$

$$-\frac{1}{2}E\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}G_i\left(|t-s|\right)dX_s^i dX_t^i\right)$$

and U_{i+1} (cf. (11)) is given by

$$U_{i+1} := \int_{\left[T_i,\hat{T}_{i+1}\right]}\sigma\left(W_t - W_{T_i}\right)dX_t^i.$$

Proof. Equality (15) follows easily. To show (16), observe that

$$Var\left(IS\left(X\right)\right) = E\left(IS\left(X\right) - E\left(IS\left(X\right)\right)\right)^2$$

$$= E\left(\sum_{i=0}^{n-1}\left(IS\left(X^i\right) - E\left(IS\left(X^i\right)\right)\right)\right)^2 = E\left(\sum_{i=0}^{n-1}\left(Z_i + U_{i+1}\right)\right)^2$$

$$= E\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left(Z_i + U_{i+1}\right)\left(Z_j + U_{j+1}\right)$$

$$= E\sum_{i=j}\left(Z_i + U_{i+1}\right)^2 + \sum_{i\neq j}\left(Z_i + U_{i+1}\right)\left(Z_j + U_{j+1}\right).$$

For $i < j$,

$$E\left(\left(Z_i + U_{i+1}\right)\left(Z_j + U_{j+1}\right)\right)$$

$$= E\left(\left(Z_i + U_{i+1}\right)Z_j\right) = E\left(Z_iZ_j\right) + E\left(U_{i+1}Z_j\right),$$

while, for $i = j$,

$$E\left(Z_i + U_{i+1}\right)^2 = EZ_i^2 + EU_{i+1}^2$$

and, for $j < i$,

$$E\left(\left(Z_i + U_{i+1}\right)\left(Z_j + U_{j+1}\right)\right) = EZ_i\left(Z_j + U_{j+1}\right)$$

$$= E\left(Z_iZ_j\right) + E\left(Z_iU_{j+1}\right).$$

Consequently,

$$Var\left(IS\left(X\right)\right) = \sum_{i=0}^{n-1}Var\left(IS\left(X^i\right)\right) + \sum_{i\neq j}\left(E\left(Z_iZ_j\right) + E\left(Z_{i\vee j}U_{i\wedge j+1}\right)\right).$$

The remaining statements follow easily. ■

We see that $E\left(IS\left(X\right)\right)$ is calculated directly by $E\left(IS\left(X^i\right)\right)$, $i = 0, ..., n-1$, for which knowledge of each individual conditional distribution of $(\varepsilon_i, V_i, G_i)$ on $\mathcal{F}_{T_i}$, $i = 0, ..., n-1$, suffices. However, the variance $Var\left(IS\left(X\right)\right)$ requires knowledge of the *joint distribution* of *all* random variables $(V_0, ..., V_{n-1})$ and $(\varepsilon_0, \varepsilon_1, ..., \varepsilon_{n-1})$, as well as the chosen by the trader $(G_0, G_1, ..., G_{n-1})$ random decay functions.

3 Reformulation of the model

In [5], the authors proposed a new formulation of Gatheral's model by modeling the execution strategies through the cumulative distribution functions of probability measures. These cdfs model the cumulative fraction of the traded volume through time, which appears to be a more intuitive quantity to work with. Furthermore, the fraction (instead of volume itself) is a unitless quantity and this facilitates the analysis of various quantities of interest (see, for example, the discussion in section 3 in [5])

Working with this version of the model also allows us to work with processes that are aligned with the classical cadlag assumption in stochastic calculus. Furthermore, we are able to derive universal properties for both the admissible and the optimal policies for general random decay functions. For the optimal policies, in particular, we are able to extend the results of section 6.2 in [5] where we studied the density of the optimal measure in the interior of the execution horizon, its symmetry, and its symmetric discontinuities at the end points. To ease the presentation, we only recall some key lemmata from [5] and state most of the results in the rest of the paper without detailed proofs.

To this end, for $i = 0, 1, ..., n-1$, we introduce probability measures ν^i on $[T_i, T_{i+1}]$, by setting

$$dX^i = \varepsilon_i V_i d\nu^i.$$

We denote by $\nu_t^i : [T_i, T_{i+1}] \to [0,1]$ the cumulative distribution function of ν^i,

$$\nu_t^i = \nu^i\left([T_i, t)\right),$$

which - as mentioned above - represents the aggregate fraction of volume $\varepsilon_i V_i$ that is being traded over $[T_i, t)$, $T_i \leq t \leq \hat{T}_{i+1}$. We have $\nu_{T_i-}^i = 0$ and $\nu_t^i = 1$, for $t \in \left[\hat{T}_{i+1}, T_{i+1}\right]$. Furthermore, ν_t^i is right continuous, nondecreasing with left limits (cadlag).

The individual implementation slippage $IS(X^i)$, $i = 0, ..., n-1$, is expressed as

$$IS\left(X^i\right) = \frac{1}{2}\varepsilon_i^2 V_i^2 \int_{\left[T_i, \hat{T}_{i+1}\right]} \int_{\left[T_i, \hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^i d\nu_t^i \qquad (17)$$
$$+\varepsilon_i V_i \int_{\left[T_i, \hat{T}_{i+1}\right]} \sigma\left(W_t - W_{T_i}\right) d\nu_t^i.$$

Proposition 5 *For $i = 0, ..., n-1$, the conditional on $\mathcal{F}_{T_i}$ mean and variance of $IS\left(X^i\right)$ are given by*

$$E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) = \frac{1}{2}\varepsilon_i^2 V_i^2 \int_{\left[T_i, \hat{T}_{i+1}\right]} \int_{\left[T_i, \hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^i d\nu_t^i$$

and

$$Var\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right) = \sigma^2 \varepsilon_i^2 V_i^2 \left(\int_{\left[T_i, \hat{T}_{i+1}\right]} \int_{\left[T_i, \hat{T}_{i+1}\right]} t \wedge s d\nu_s^i d\nu_t^i - T_i\right).$$

Respectively, the unconditional mean and variance are given by

$$E\left(IS\left(X^i\right)\right)=\frac{1}{2}E\left(\varepsilon_i^2V_i^2\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}G_i\left(|t-s|\right)d\nu_s^id\nu_t^i\right)$$

and

$$Var\left(IS\left(X^i\right)\right)=Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right)$$
$$+E\left(\sigma^2\varepsilon_i^2V_i^2\left(\int_{\left[T_i,\hat{T}_{i+1}\right]}\int_{\left[T_i,\hat{T}_{i+1}\right]}t\wedge s d\nu_s^id\nu_t^i-T_i\right)\right).$$

The above representations expose the dependence of the means and variances on the random order characteristics V_i and ε_i, and the random decay function G_i that is chosen by the trader. Among others, they show that they depend *only on the frequency* of empty orders and not on whether order i is a buy or sell order.

4 Optimality results for orders with stochastic characteristics

We present the optimal policies and the minimized implementation slippages for two cases. Firstly, we consider only a single order with stochastic characteristics, arriving at a future time. We then study the optimal policies and the minimized aggregate implementation slippage. We also consider the case of random exponential decay functions and provide explicit results. For completeness, we present most of the results using both formulations for the strategies (i.e. with X^i and ν^i, and X and ν).

4.1 Future single order

For $i=0,1,...,n-1$, assume that order i with $\mathcal{F}_{T_i}-$measurable characteristics V_i and ε_i will arrive at time T_i. We also assume that the trader will use a decay function $G_i\in\mathcal{F}_{T_i}$ that satisfies (2) and will apply admissible execution strategies

$$dX_t^i=\varepsilon_iV_id\nu_t^i,$$

in the respective execution interval $\left[T_i,\hat{T}_{i+1}\right]$. The following result follows in analogy to the single order case (see [3], [2] and [5]).

Proposition 6 *For $i=0,1,...n-1$, the minimal expected implementation slippage of order i, with characteristics V_i,ε_i and G_i, is given by*

$$\min_{X^i}E\left(IS\left(X^i\right)\right)=E\left(\min_{X^i}E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right)=\frac{1}{2}E\left(\varepsilon_iV_i\lambda_i\right),$$

where $X^{i,*}$ is the optimal strategy. Its optimality is equivalent to the existence of a unique (a.s.) $\mathcal{F}_{T_i}-$measurable random variable λ_i such that, for each $t \in \left[T_i, \hat{T}_{i+1}\right]$,

$$\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right) dX_s^{i,*} = \varepsilon_i \lambda_i.$$

Proposition 7 *For $i = 0, 1, ...n-1,$, the optimal fraction $\nu_t^{i,*}$ of order i satisfies*

$$\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^{i,*} = \mu_i, \quad ,\text{for each } t \in \left[T_i, \hat{T}_{i+1}\right],$$

where

$$\lambda_i = \varepsilon_i V_i \mu_i.$$

The minimal aggregate expected slippage for order i is given by

$$\min_{X^i} E\left(IS\left(X^i\right)\right) = \frac{1}{2} E\left(\mu_i \varepsilon_i^2 V_i^2\right).$$

The following result provides the variance of the optimal implementation slippage in terms of the optimal measure.

Proposition 8 *Let $\nu^{j,*}$ be the optimal measure for order j, $j = 0, ..., n-1$. Then,*

$$\begin{aligned} Var\left(IS\left(X^{j,*}\right)\right) &= \frac{1}{4} Var\left(\varepsilon_j^2 V_j^2 \mu_j\right) \\ &+ E\left(\sigma_j^2 \varepsilon_j^2 V_j^2 \int_{\left(T_j,\hat{T}_{j+1}\right]} \left(\nu^{j,*}\left(\left[t, \hat{T}_{i+1}\right]\right)\right)^2 dt\right). \end{aligned} \tag{18}$$

The characteristic function of $IS\left(X^{j,*}\right)$, is given by

$$E\left(e^{itIS\left(X^{j,*}\right)}\right) = E\left(e^{it\frac{1}{2}\varepsilon j V_j^2 \mu_j - \frac{1}{2}t^2 \sigma_j^2 \varepsilon_j^2 V_j^2 \int_{\left(T_j,\hat{T}_{j+1}\right]} \left(\nu^{j,*}\left(\left[s,\hat{T}_{j+1}\right]\right)\right)^2 ds}\right).$$

The proof of (18) is based on the following lemma (see (see proof Lemma 8 in [5]), appropriately modified to capture the stochastisity of the orders.

Lemma 9 *Let ν be a probability measure on the interval $[a, b]$, $a < b$, and define $\nu_{a-} = 0$, $\nu_t = \nu\left([a, t]\right)$, $a \leq t \leq b$. Then,*

$$\int_{[a,b]} \int_{[a,b]} (t \wedge s)\, d\nu_s d\nu_t = a + \int_{(a,b]} \left(\nu\left([t, b]\right)\right)^2 dt.$$

We note that the distribution of $IS\left(X^{j,*}\right)$ depends only on the *joint distribution* of the random vector

$$\left(\varepsilon_j^2, V_j^2, \mu_j, \sigma_j^2, \int_{\left(T_j,\hat{T}_{j+1}\right]} \left(\nu^{j,*}\left(\left[s, \hat{T}_{j+1}\right]\right)\right)^2 ds\right).$$

However, it does not depend on previous orders and the decay functions used for their execution.

4.2 Sequential orders

We are now ready to state the main optimality results for problem (8).

Proposition 10 *Let $i = 0, 1, ..., n-1$. Assume a sequence of orders with characteristics V_i, ε_i and G_i, arriving at times T_i, $i = 0, 1, ..., n-1$, and to be executed on $\left[T_i, \hat{T}_i\right]$, $\hat{T}_i < T_{i+1}$, respectively.*

An execution strategy $X^* = \left(X^{0,*}, ..., X^{n-1,*}\right)$, with $X^{i,*} \in \mathcal{A}_{[T_i,T_{i+1}]}$ (cf. (3a)) is optimal if and only if there exist random variables $\lambda_0, \lambda_1, ..., \lambda_{n-1}$, with λ_i being $\mathcal{F}_{T_i}-$measurable, such that, for $i = 0, 1, ..., n-1$,

$$\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right) dX_s^{i,*} = \varepsilon_i \lambda_i, \quad \textit{for each} \quad t \in \left[T_i, \hat{T}_{i+1}\right].$$

The optimal fraction $\nu_t^{i,*}$ satisfies for all $t \in \left[T_i, \hat{T}_{i+1}\right]$ and, $i = 0, 1, ...n-1$,

$$\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^{i,*} = \mu_i \quad \textit{with} \quad \lambda_i = \varepsilon_i V_i \mu_i.$$

Proposition 11 *The minimal aggregate expected implementation slippage is given by*

$$E(IS(X^*)) = \frac{1}{2}\sum_{i=0}^{n-1} E\left(\lambda_i \varepsilon_i V_i\right) \tag{19}$$

$$= \frac{1}{2}\sum_{i=0}^{n-1} E\left(\mu_i \varepsilon_i^2 V_i^2\right) = \frac{1}{2}\sum_{i=0}^{n-1} E\left(\mu_i V_i^2 \mathbf{1}_{\{\varepsilon_i \neq 0\}}\right).$$

Furthermore,

$$Var\left(IS\left(X^*\right)\right) = \sum_{i=0}^{n-1} Var\left(IS\left(X^{i,*}\right)\right) \tag{20}$$

$$+\sum_{i\neq j}\left(E\left(Z_i^* Z_j^*\right) + E\left(Z_{i\vee j}^* U_{i\wedge j+1}^*\right)\right),$$

where

$$Z_i^* := \frac{1}{2}\varepsilon_i^2 V_i^2 \mu_i - \frac{1}{2}E\left(\varepsilon_i^2 V_i^2 \mu_i\right)$$

and

$$U_{i+1}^* := \int_{\left[T_i,T_{i+1}^-\right]} \sigma\left(W_t - W_{T_i}\right) dX_t^{i,*}$$

$$= \varepsilon_i V_i \int_{\left[T_i,T_{i+1}^-\right]} \sigma\left(W_t - W_{T_i}\right) d\nu_t^{i,*} = \varepsilon_i V_i W_{i+1}^*,$$

with

$$W_{i+1}^* := \int_{\left[T_i,T_{i+1}^-\right]} \sigma\left(W_t - W_{T_i}\right) d\nu_t^{i,*}.$$

4.3 Random exponential decay functions

We assume that, for each order $i = 0, 1, ..., n-1$, the random decay functions

$$G_i(|t|) = e^{-\kappa_i |t|}, \quad t \in \mathbb{R}, \quad \kappa_i > 0 \text{ with } \kappa_i \in \mathcal{F}_{T_i}, \tag{21}$$

are used. In analogy to the results in [3], [2] and [5], we deduce that the optimal execution strategy $X^* = (X^{*0}, ..., X^{*,n})$ for a sequence of orders with characteristics $((V_0, \varepsilon_0), (V_1, \varepsilon_1), ..., (V_n, \varepsilon_n))$, is given, for $i = 0, ..., n-1$, by

$$dX_t^{*i} = \varepsilon_i V_i d\nu_t^{i,*},$$

where

$$d\nu_t^{i,*} = \frac{1}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)} \left(\delta_{T_i} + \delta_{\hat{T}_{i+1}} + \kappa_i dt\right).$$

The optimal policy satisfies, for each $t \in \left[T_i, \hat{T}_{i+1}\right]$,

$$\int_{\left[T_i, \hat{T}_{i+1}\right]} e^{-\kappa_i |t-s|} dX_s^{i,*} = \frac{2\varepsilon_i V_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)} = \varepsilon_i \lambda_i = \varepsilon_i V_i \mu_i.$$

The minimal aggregate expected slippage is given by

$$E(IS(X^*)) = \frac{1}{2} \sum_{i=0}^{n-1} E\left(\frac{\varepsilon_i^2 V_i^2}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)}\right). \tag{22}$$

The associated impacted price process satisfies, for $i = 0, 1, ..., n-1$,

- For $T_i < t \leq \hat{T}_{i+1}$,

$$S_t^* = S_{T_i}^* + \int_{[T_i, t)} e^{-\kappa_i |t-s|} dX_s^{*,i} + \sigma (W_t - W_{T_i})$$

$$= S_{T_i}^* + \frac{\varepsilon_i V_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)} + \sigma (W_t - W_{T_i}) = S_{T_i+}^* + \sigma (W_t - W_{T_i}).$$

  The jumps at T_i and $\hat{T}_{i+1}$ are of *equal sizes*, specifically,

$$S_{T_i+}^* - S_{T_i}^* = S_{\hat{T}_{i+1}+}^* - S_{\hat{T}_{i+1}}^* = \frac{\varepsilon_i V_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)}.$$

- For $\hat{T}_{i+1} < t \leq T_{i+1}$,

$$S_t^* = S_{\hat{T}_{i+1}+}^* + \sigma \left(W_t - W_{\hat{T}_{i+1}}\right),$$

  and, hence, it is a Brownian motion starting from $S_{\hat{T}_{i+1}+}^*$ at time $\hat{T}_{i+1}$.

At time T_i, $i = 0, 1, ..., n-1$, the volume $\frac{\varepsilon_i V_i}{2+\kappa_i\left(\hat{T}_{i+1}-T_i\right)}$ is traded, which may be negative (sell order), zero (empty order) or positive (buy order). Up until time t, with $T_i < t < \hat{T}_{i+1}$, the volume

$$\frac{\varepsilon_i V_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)} + \frac{\varepsilon_i V_i \kappa_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)} t$$

is executed while the remainder

$$\frac{\varepsilon_i V_i}{2 + \kappa_i \left(\hat{T}_{i+1} - T_i\right)}$$

is traded all at once at time $\hat{T}_{i+1}$.

Discussion:

i) The impacted price process S_t^*, $t \in [0, T]$, is left continuous with right hand limits, and with jumps at each T_i and $\hat{T}_{i+1}$, $i = 1, ..., n-1$. This is a consequence of the fact that *both* the arrival time of each order and execution horizon are specified in advance ($\mathcal{F}_0-$measurable).

ii) The minimal average expected slippage $E\left(IS\left(X^*\right)\right)$ (cf. (22)) depends on ε_i^2, which takes vales 0 and 1. Thus, $E\left(IS\left(X^*\right)\right)$ does *not* depend on the *direction* of each of the orders.

iii) If order i is an empty one, $\varepsilon_i^2 = 0$, the corresponding term in (22) vanishes. Consequently, the larger the proportion of empty orders, the smaller the optimal aggregate implementation slippage.

iv) The pair (V_i, ε_i) is generated by the market and characterizes an order, which arrive at time T_i, $i = 0, 1, ..., n-1$ and needs to be executed by $\hat{T}_{i+1} < T_{i+1}$. On the other hand, the random decay function G_i is chosen by the trader who uses (V_i, ε_i) and the execution model.

v) While the expected aggregate implementation slippage depends only on the marginal distribution of each order's characteristics, the assessment of the frequency depends on the *joint distribution* of the $\varepsilon_i^2, i = 0, 1, .., n-1$. Ability to forecast the next empty order could lower the overall execution cost.

vi) The description of the triplet dynamics V_i, ε_i and G_i depends on the *joint distribution* of *all* such triplets.

The joint distribution needs to represent the movement that is consistent with the way that a) *market generates orders* and b) the *traders adjust* the random decay functions G_i.

The challenge here is to statistically capture this movement and reconcile it with the way the market prices single orders in succession. This comes down to defining the transition mechanism of order characteristics from one execution period to the next.

vii) Historical data should be used to determine how the volumes $V_0, V_1, ..., V_n$ move with the order flow and the market.

5 Rethinking of optimality criteria

In [5] the authors pointed out a deficiency of Gatheral's model, namely, the expected implementation slippage may become arbitrarily small, or even vanish, while its variance remains finite**.** The elimination of the expected slippage may be, for example, achieved by choosing a suitable decay function. They then proposed two possible ways to remedy this flaw.

The first direction is to still work with the same criterion (minime the expected slippage) but, also, impose constraints on the optimal solution. For example, one may want to control the variance of the optimal slippage. This, in turn, imposes constraints on the model parameters. A major challenge, however, is that very little, if anything, is known for the structure of the optimal solution once we depart from the tractable exponential case. For general decay functions, the authors in [5] derived various new results. Among others, they proved a general robustness result of the optimal solution with respect to the decay function, and also established universal bounds of the variance of the implementaion slippage. For the latter, they worked with the reformulated Gatheral's model (as (17) herein) and characterized in detail the optimal measure.

The second alternative proposed in [5] is to work with an alternative optimization criterion which combines the mean and the variance of the implementation slippage. For the single order case, they showed that this results in a deterministic two-dimensional calculus of variations problem.

When we deal with sequential orders, which also have stochastic characteristics, the problem becomes considerably more complex. A major difficulty stems from the rather convoluted dependence of the variance of the implementaion slippage on the joint distribution of all random variables involved, namely, the volumes $(V_1, V_2, ..., V_n)$, the directions $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ and the (to be chosen) decay functions $(G_1, G_2, ..., G_n)$.

Another difficulty arises from the multiplicity of the orders itself. For example, in order to follow the first alternative proposed by the authors, one needs to decide a priori what constraints should be imposed "order-by-order" on the optimized individual slippages. But these slippages are affected by the sequential choice of the random decay functions, which are themselves influenced by the upcoming realized market impacts and other quantities.

Below, we do not give answers to these questions but, rather, take the step to first generalize the results in [5] to individual future orders with stochastic charcateristics. Furthermore, we only focus on the first alternative which is, neverthless, already challenging enough.

In lack of a better word, we will be using the "model calibration" terminology to describe the effects on the model parameters that the additional constraints on the optimal solution will have. To motivate the reader, we start with a tractable case of sequential deterministic decay functions of exponential type.

5.1 A motivational example

Assume that the trader chooses to work with sequential deterministic decay functions,

$$G_i(|t|) = e^{-\kappa_i |t|}, \quad t \in \mathbb{R}\ , \quad \kappa_i > 0,$$

for order i, $i = 0, 1, ..., n-1$. The related optimal execution strategy $X^{i,*}$ is then given by

$$dX_t^{*i} = \varepsilon_i V_i d\nu_t^{i,*} = \frac{\varepsilon_i V_i}{2 + \kappa_i(\hat{T}_{i+1} - T_i)} \left(\delta_{T_i} + \delta_{\hat{T}_{i+1}} + \kappa_i dt \right).$$

Moreover,

$$\int_{\left[T_i, \hat{T}_{i+1}\right]} e^{-\kappa_i |t-s|} dX_s^{i,*} = \frac{2\varepsilon_i V_i}{2 + \kappa_i(\hat{T}_{i+1} - T_i)} = \varepsilon_i \lambda_i = \varepsilon_i V_i \mu_i,$$

with $\mu_i = \frac{\lambda_i}{V_i}$, for $V_i \neq 0$.

The minimal expected implementation slippage is equal to

$$E\left(IS\left(X^{i,*}\right)\right) = \frac{1}{2} E\left(\lambda_i \varepsilon_i V_i\right) = \frac{1}{2 + \kappa_i(\hat{T}_{i+1} - T_i)} E\left(\varepsilon_i^2 V_i^2\right).$$

We easily deduce that

$$\lim_{\kappa_i \uparrow \infty} E\left(IS\left(X^{i,*}\right)\right) = 0,$$

and, hence, the expected implementation slippage for future order i can be made *arbitrarily small.*

Next, we look at $Var\left(IS\left(X^{i,*}\right)\right)$. Working as in Proposition 6 in [5] we deduce the following result. The calculations are tedious and are omitted to ease the presentation.

Proposition 12 *The unconditional variance of $IS\left(X^{i,*}\right)$ is given by*

$$Var\left(IS\left(X^{i,*}\right)\right) = Var\left(\frac{\varepsilon_i^2 V_i^2}{2 + \kappa_i \Delta_i}\right) + E\left(\sigma^2 \varepsilon_i^2 V_i^2 \frac{\Delta_i + \kappa_i \Delta_i^2 + \kappa_i^2 \frac{1}{3} \Delta_i^3}{\left(2 + \kappa_i \Delta_i\right)^2}\right)$$

with

$$\Delta_i := \hat{T}_{i+1} - T_i. \tag{23}$$

Thus,

$$\lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right) = \frac{\Delta_i}{3} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right)$$

and

$$\lim_{\kappa_i \uparrow 0} Var\left(IS\left(X^{i,*}\right)\right) = \frac{1}{4} Var\left(\varepsilon_i^2 V_i^2\right) + \frac{\Delta_i}{4} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right).$$

From the above we deduce that

$$\lim_{\kappa_i \downarrow 0} Var\left(IS\left(X^{i,*}\right)\right) \leq \lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right)$$

if and only if

$$Var\left(\varepsilon_i^2 V_i^2\right) \leq \frac{\hat{T}_{i+1} - T_i}{3} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right) = \lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right).$$

This holds for sufficiently large σ, provided, however, that $E\left(V_i^4\right) < \infty$.

The above example demonstrates that by taking very large $\kappa_i' s$, one can make the expected implementation cost arbitrarily small. In the limit $\kappa_i \uparrow \infty$, the expected implementation slippage is equal to zero but, at the same time, the $Var\left(IS\left(X^{i,*}\right)\right)$ remains *bounded from below.* Consequently, one can reduce the execution cost to zero but, yet, the standard deviation of $IS\left(X^{i,*}\right)$ remains finite.

Observe also that

$$\lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right) = \frac{\hat{T}_{i+1} - T_i}{3} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right),$$

and

$$\lim_{\kappa_i \downarrow 0} Var\left(IS\left(X^{i,*}\right)\right) = \frac{1}{4} Var\left(\varepsilon_i^2 V_i^2\right) + \frac{\hat{T}_{i+1} - T_i}{4} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right).$$

Therefore,

$$\lim_{\kappa_i \downarrow 0} Var\left(IS\left(X^{i,*}\right)\right) \leq \lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right)$$

if and only if

$$Var\left(\varepsilon_i^2 V_i^2\right) \leq \frac{\hat{T}_{i+1} - T_i}{3} E\left(\sigma^2 \varepsilon_i^2 V_i^2\right) = \lim_{\kappa_i \uparrow \infty} Var\left(IS\left(X^{i,*}\right)\right). \tag{24}$$

The latter holds for sufficiently large σ, provided however that $E\left(V_i^4\right) < \infty$.

5.1.1 Choosing the stock's volatility

Inequality (24) actually raises questions about the *choice of volatility* in the underlying model. To explain this point, let us for now assume that all decay deterministic parameters are taken to be equal to the initial one,

$$\kappa_i = \kappa_0, \; i = 0, 1, ..., n-1.$$

It, then, follows that

$$Var\left(IS\left(X^{i,*}\right)\right) = Var\left(\frac{\varepsilon_i^2 V_i^2}{2 + \kappa_0 \Delta_i}\right) + E\left(\sigma^2 \varepsilon_i^2 V_i^2 \frac{\Delta_i + \kappa_0 \Delta_i^2 + \kappa_0^2 \frac{1}{3} \Delta_i^3}{\left(2 + \kappa_0 \Delta_i\right)^2}\right)$$

$$= \frac{1}{\left(2+\kappa_0\Delta_i\right)^2} Var\left(\varepsilon_i^2 V_i^2\right) + \frac{\Delta_i + \kappa_0\Delta_i^2 + \kappa_0^2\frac{1}{3}\Delta_i^3}{\left(2+\kappa_0\Delta_i\right)^2} E\left(\sigma^2\varepsilon_i^2 V_i^2\right),$$

with Δ_i as in (23).

Observe now that if the volatility parameter is taken as a function of κ_0 which also blows up as $\kappa_0 \to \infty$, i.e. $\sigma = \sigma(\kappa_0)$ with $\lim_{\kappa_0\uparrow\infty} \sigma(\kappa_0) = \infty$, then the $Var\left(IS\left(X^{i,*}\right)\right)$ may converge to infinity with an arbitrary rate.

For example, if we choose $\sigma(\kappa_0) = \sqrt{2+\kappa_0\Delta_i}$, we have

$$Var\left(IS\left(X^{i,*}\right)\right) = \frac{1}{\left(2+\kappa_0\Delta_i\right)^2} Var\left(\varepsilon_i^2 V_i^2\right) + \frac{\Delta_i + \kappa_0\Delta_i^2 + \kappa_0^2\frac{1}{3}\Delta_i^3}{2+\kappa_0\Delta_i} E\left(\varepsilon_i^2 V_i^2\right),$$

and the growth rate is linear.

If, on the other hand, we take $\sigma(\kappa_0) = 2+\kappa_0\Delta_i$, then,

$$Var\left(IS\left(X^{i,*}\right)\right) = \frac{1}{\left(2+\kappa_0\Delta_i\right)^2} Var\left(\varepsilon_i^2 V_i^2\right) + \left(\Delta_i + \kappa_0\Delta_i^2 + \kappa_0^2\frac{1}{3}\Delta_i^3\right) E\left(\varepsilon_i^2 V_i^2\right)$$

and the growth rate is quadratic.

In the general case, when the parameters are random, $\kappa_i \in \mathcal{F}_{T_i}$, $i = 0, 1, ..., n-1$, allowing the volatility σ to depend on κ_i requires *modification* of Gatheral's model to a model with *stochastic volatility.* Replacing σ with $\sigma(\kappa_i)$, or even with an arbitrary random variable $\sigma_i \in \mathcal{F}_{T_i}$, does not pose major difficulties as it can be accommodated by the methodology we have developed. Then, in such a framework we have the order characteristics ε_i and V_i, which are given by the market, and the chosen random variables κ_i and σ_i or, in general G_i and σ_i, which determine the model for order execution.

The impacted process is then calibrated to the order characteristics to generate behavior consistent with the dynamics of the characteristics and with the observed performance of the execution strategies. The modified dynamics of the impacted process are of the form

$$S_t = S_{T_i} + \int_{[T_i,t)} G_i(t-s)\, dX_s^i + \sigma_i\left(W_t - W_{T_i}\right), \quad t \in (T_i, T_{i+1}],$$

where $\sigma_i \in \mathcal{F}_{T_i}$ with $E\left(\sigma_i^4\right) < \infty$.

Under the above assumptions, all previous results hold true when the volatility parameter σ is replaced with σ_i in $[T_i, T_{i+1}]$, $i = 0, 1, ..., n-1$. Specifically,

$$E\left(IS\left(X^i\right)\right) = \frac{1}{2} E\left(\varepsilon_i^2 V_i^2 \int_{\left[T_i,\hat{T}_{i+1}\right]} \int_{\left[T_i,\hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^i d\nu_t^i\right)$$

and

$$Var\left(IS\left(X^i\right)\right) = Var\left(E\left(IS\left(X^i\right)|\mathcal{F}_{T_i}\right)\right)$$
$$+E\left(\sigma_i^2\varepsilon_i^2 V_i^2\left(\int_{\left[T_i,\hat{T}_{i+1}\right]} \int_{\left[T_i,\hat{T}_{i+1}\right]} t\wedge s\, d\nu_s^i d\nu_t^i - T_i\right)\right).$$

5.2 General random decay functions

We consider order i, $i = 0, 1, ..., n-1$, with characteristics V_i, ε_i and decay function G_i. We will be using the self-evident notation $IS\left(X^i, G_i\right)$.

Proposition 13 *Let $i = 0, 1, ..., n-1$ and assume that the random decay functions G_i, H_i satisfy*

$$G_i \leq H_i, \quad a.s..$$

Then,

$$E\left(IS\left(X^i, G_i\right)\right) = E\left(\int_{\left[T_i, \hat{T}_{i+1}\right]} \int_{\left[T_i, \hat{T}_{i+1}\right]} G_i\left(|t-s|\right) dX_s^i dX_t^i\right)$$

$$\leq E\left(\int_{\left[T_i, \hat{T}_{i+1}\right]} \int_{\left[T_i, \hat{T}_{i+1}\right]} H_i\left(|t-s|\right) dX_s^i dX_t^i\right) = E\left(IS\left(X^i, H_i\right)\right).$$

Therefore, as in the single order case in [5], the implementation slippage can be made arbitrarily small by taking a decreasing sequence of decay functions $G_i^{(n)}$,

$$0 < G_i^{(m+1)} \leq G_i^{(m)}, a.s.$$

As mentioned earlier, we then need to consider additional quantities related to the random variables $IS\left(X^i, G_i^{(m)}\right)$ and $\sum_{i=0}^{n-1} IS\left(X^i, G_i^{(m)}\right)$. In particular, we considered the unconditional variance

$$Var\left(\sum_{i=0}^{n-1} IS\left(X^i, G_i^{(m)}\right)\right)$$

and observed that it displays explicit dependence among the various random modeling components, namely, the random sequences $(V_0, ..., V_{n-1})$, $(\varepsilon_0, \varepsilon_1, ..., \varepsilon_{n-1})$, $\left(G_0^{(m)}, G_1^{(m)}, ..., G_{n-1}^{(m)}\right)$ *and* the Brownian motion $(W_t : t \geq 0)$.

5.3 Targeted interplay between $E(IS(X^*))$ and $Var\left(IS\left(X^*\right)\right)$

We present results for $E(IS(X^*))$ and its conncetion with $Var\left(IS\left(X^*\right)\right)$. We first recall the following lemma (see section 4.2 in [5])

Lemma 14 *Let G and H be decay functions such that*

$$0 < G \leq H.$$

Then, for any probability measure ν on an interval $[a, b]$

$$\int_{[a,b]} \int_{[a,b]} G\left(|t-s|\right) d\nu_s d\nu_t \leq \int_{[a,b]} \int_{[a,b]} H\left(|t-s|\right) d\nu_s d\nu_t.$$

Let ν^G and ν^H be the probability measures for which, respectively,

$$\int_{[a,b]} G(|t-s|)\, d\nu_s^G = \mu^G \quad \text{and} \quad \int_{[a,b]} H(|t-s|)\, d\nu_s^H = \mu^H,$$

for all $t \in [a,b]$. Then,

$$\mu^G \leq \mu^H.$$

To calibrate our model, we take a decreasing sequence of decay functions $G_i^{(n)}$ such that $0 < G_i^{(m+1)} \leq G_i^{(m)}$ and assume that, for each i and m,

$$\int_{\left[T_i,\hat{T}_{i+1}\right]} G_i^{(m)}(|t-s|)\, d\nu_s^{i,*,(m)} = \mu_i^{(m)},$$

for a probability measure $\nu_s^{i,*,(m)}$. It, then, follows from the above lemma that $\mu_i^{(m+1)} \leq \mu_i^{(m)}$. Next, assume that, for each i, $\lim_{m\uparrow\infty} G_i^{(m)} = 0$. Then, for any probability measure ν,

$$\lim_{m\uparrow\infty} \mu_i^{(m)} = \lim_{m\rightarrow\infty} \int_{\left[T_i,\hat{T}_{i+1}\right]} G_i^{(m)}(|t-s|)\, d\nu_s^{i,*,(m)}$$

$$= \lim_{m\uparrow\infty} \int_{\left[T_i,\hat{T}_{i+1}\right]} \int_{\left[T_i,\hat{T}_{i+1}\right]} G_i^{(m)}(|t-s|)\, d\nu_s^{i,*,(m)} d\nu_t^{i,*,(m)}$$

$$\leq \lim_{m\uparrow\infty} \int_{\left[T_i,\hat{T}_{i+1}\right]} \int_{\left[T_i,\hat{T}_{i+1}\right]} G_i^{(m)}(|t-s|)\, d\nu_s d\nu_t.$$

Consequently,

$$\lim_{m\uparrow\infty} \mu_i^{(m)} \leq \int_{\left[T_i,\hat{T}_{i+1}\right]} \int_{\left[T_i,\hat{T}_{i+1}\right]} \lim_{m\uparrow\infty} G_i^{(m)}(|t-s|)\, d\nu_s d\nu_t = 0.$$

Next, we recall that, for each i and m,

$$E\left(IS\left(X^{i,*,(m)}\right)\right) = \frac{1}{2} E\left(\varepsilon_i^2 V_i^2 \mu_i^{(m)}\right)$$

and

$$Var\left(IS\left(X^{i,*,(m)}\right)\right) = \frac{1}{4} Var\left(\varepsilon_i^2 V_i^2 \mu_i^{(m)}\right)$$

$$+E\left(\varepsilon_i^2 V_i^2 \left(\sigma_i^{(m)}\right)^2 \int_{\left(T_i,\hat{T}_{i+1}\right]} \left(\nu^{i,*,(m)}\left(\left[t,\hat{T}_{i+1}\right]\right)\right)^2 dt\right).$$

Therefore, in order to calibrate the model, we choose the decay function $G_i^{(m)}$ *together* with the stochastic volatility $\sigma_i^{(m)}$ in a way to obtain the *desired levels* of the mean $E\left(IS\left(X^{i,*,(m)}\right)\right)$ and variance $Var\left(IS\left(X^{i,*,(m)}\right)\right)$. Clearly,

$$E\left(IS\left(X^{i,*,(m+1)}\right)\right) \leq E\left(IS\left(X^{i,*,(m)}\right)\right)$$

and

$$\lim_{m\uparrow\infty} E\left(IS\left(X^{i,*,(m)}\right)\right) = \frac{1}{2}\lim_{m\uparrow\infty} E\left(\varepsilon_i^2 V_i^2 \mu_i^{(m)}\right) = 0.$$

Moreover, observe that

$$\lim_{m\uparrow\infty} Var\left(IS\left(X^{i,*,(m)}\right)\right)$$

$$= \lim_{m\uparrow\infty} E\left(\varepsilon_i^2 V_i^2 \left(\sigma_i^{(m)}\right)^2 \int_{\left(T_i,\hat{T}_{i+1}\right]} \left(\nu^{i,*,(m)}\left(\left[t,\hat{T}_{i+1}\right]\right)\right)^2 dt\right) = \infty,$$

under appropriate conditions on the sequence

$$\left(\sigma_i^{(m)}\right)^2 \int_{\left(T_i,\hat{T}_{i+1}\right]} \left(\nu^{i,*,(m)}\left(\left[t,\hat{T}_{i+1}\right]\right)\right)^2 dt,\ m = 1,2,3...$$

When the mean converges to zero and the variance grows to infinity, the decrease in the expected implementation slippage can be *balanced* against the increase in the standard deviation of $IS\left(X^{i,*,(m)}\right)$.

5.4 Distribution of $IS\left(X^*\right)$

The distribution of the optimized implementation splippage

$$IS\left(X^*\right) = \sum_{i=0}^{n-1} IS\left(X^{i,*}\right),$$

is considerably more involved. Indeed, it depends on the *joint distribution of all orders and on their interaction with the Brownian motion* defining the model dynamics. To see this, recall that

$$IS\left(X^{i,*}\right) = \frac{1}{2}\varepsilon_i^2 V_i^2 \mu_i + \varepsilon_i V_i \sigma_i M_{i+1}$$

where

$$M_{i+1} = \int_{\left[T_i,T_{i+1}^-\right]} \left(W_t - W_{T_i}\right) d\nu_t^{i,*}$$

is $\mathcal{F}_{\hat{T}_{i+1}}$−measurable with $E\left(M_{i+1}\left|\mathcal{F}_{T_i}\right.\right) = 0$. Consequently, for $i < j$, we have

$$E\left(IS\left(X^{i,*}\right) IS\left(X^{j,*}\right)\right) = E\left(IS\left(X^{i,*}\right) E\left(IS\left(X^{j,*}\right)\left|\mathcal{F}_{\hat{T}_j}\right.\right)\right)$$

$$= E\left(IS\left(X^{i,*}\right)\frac{1}{2}\varepsilon_j^2 V_j^2 \mu_j\right) = \frac{1}{4}E\left(\varepsilon_i^2 V_i^2 \mu_i \varepsilon_j^2 V_j^2 \mu_j\right)$$

$$+\frac{1}{2}E\left(\varepsilon_i V_i \sigma_i M_{i+1} \varepsilon_j^2 V_j^2 \mu_j\right).$$

Hence, the covariance displays not only the dependence among the orders but also on the Brownian motion W through the random variable M_{i+1}. In order to analyze further this dependence, we will use the following result; for its proof see Proposition 9 in [5].

Lemma 15 *Let ν be a probability measure on the interval $[a,b], a < b$ and define $\nu_{a-} = 0$, $\nu_t = \nu([a,t])$, $a \le t \le b$. Then,*

$$\int_{[a,b]} (W_t - W_a)\, d\nu_t = \int_{(a,b]} \nu([t,b])\, dW_t$$

and

$$E\left(\int_{[a,b]} (W_t - W_a)\, d\nu_t\right)^2 = \int_{(a,b]} (\nu([t,b]))^2\, dt.$$

Moreover,

$$\frac{\left(\int_{[a,b]} (t-a)\, d\nu_t\right)^2}{b-a} \le \int_{(a,b]} (\nu([t,b]))^2\, dt \le \int_{[a,b]} (t-a)\, d\nu_t, \tag{25}$$

$$\inf_\nu \int_{(a,b]} (\nu([t,b]))^2\, dt = \int_{(a,b]} (\delta_a([t,b]))^2\, dt = 0$$

and

$$\sup_\nu \int_{(a,b]} (\nu([t,b]))^2\, dt = \int_{(a,b]} (\delta_b([t,b]))^2\, dt = b-a.$$

Remark 16 *Note that the above lemma yields*

$$E\left(\int_{[a,b]} (W_t - W_a)\, d\nu_t\right)^2 = E\left(\int_{(a,b]} \nu([t,b])\, dW_t\right)^2 = \int_{(a,b]} (\nu([t,b]))^2\, dt,$$

which can be used to compute the variances $Var\left(IS\left(X^{i,*}\right)\right)$.

It is well known that if $\varphi_X(t) := Ee^{itX}$, the characteristic function of a random variable X, is integrable, the cumulative distribution function F_X of X is absolutely continuous. Hence, it has a density, f_X, given by

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t)\, dt.$$

Next, we recall that the optimal measure $\nu^{i,*}$ satisfies $\nu^{i,*}\left(\left(\hat{T}_{i+1}, T_{i+1}\right]\right) = 0$ and, hence, we have

$$M_{i+1} = \int_{\left[T_i, \hat{T}_{i+1}\right]} (W_t - W_{T_i})\, d\nu_t^{i,*} = \int_{[T_i, T_{i+1}]} (W_t - W_{T_i})\, d\nu_t^{i,*}.$$

Consequently, using the previous lemma we deduce that

$$M_{i+1} = \int_{(T_i, T_{i+1}]} \nu^{i,*}([t, T_{i+1}])\, dW_t.$$

This representation of M_{i+1} is particularly useful in the analysis of the aggregate implementation slippage, as it is demonstrated below.

$$IS\left(X^{*}\right)=\sum_{i=0}^{n-1} IS\left(X^{i,*}\right)=\sum_{i=0}^{n-1}\left(\frac{1}{2}\varepsilon_i^2 V_i^2\mu_i+\varepsilon_i V_i\sigma_i M_{i+1}\right)$$

$$=\frac{1}{2}\sum_{i=0}^{n-1}\varepsilon_i^2 V_i^2\mu_i+\sum_{i=0}^{n-1}\varepsilon_i V_i\sigma_i\int_{(T_i,T_{i+1}]}\nu^{i,*}\left([t,T_{i+1}]\right)dW_t$$

$$=\frac{1}{2}\sum_{i=0}^{n-1}\varepsilon_i^2 V_i^2\mu_i+\sum_{i=0}^{n-1}\varepsilon_i V_i\sigma_i\int_{(0,T_n]}\mathbf{1}_{(T_i,T_{i+1}]}\left(t\right)\nu^{i,*}\left([t,T_{i+1}]\right)dW_t$$

$$=\frac{1}{2}\sum_{i=0}^{n-1}\varepsilon_i^2 V_i^2\mu_i+\int_{(0,T_n]}\left(\sum_{i=0}^{n-1}\varepsilon_i V_i\sigma_i\mathbf{1}_{(T_i,T_{i+1}]}\left(t\right)\nu^{i,*}\left([t,T_{i+1}]\right)\right)dW_t.$$

Using the above we deduce the following result.

Proposition 17 *i) The conditional on*

$$\left(\varepsilon_i,V_i,\mu_i,\sigma_i,\int_{(T_i,T_{i+1}]}\left(\nu^{k,*}\left([t,T_{i+1}]\right)\right)^2 dt;\ i=0,1,...,n-1\right),$$

distribution of $IS\left(X^{*}\right)$ *is normal with mean*

$$A:=\frac{1}{2}\sum_{i=0}^{n-1}\varepsilon_i^2 V_i^2\mu_i \tag{26}$$

and variance

$$B:=\int_{(0,T_n]}\left(\sum_{i=0}^{n-1}\varepsilon_i V_i\sigma_i\mathbf{1}_{(T_i,T_{i+1}]}\left(t\right)\nu^{i,*}\left([t,T_{i+1}]\right)\right)^2 dt$$

$$=\sum_{i=0}^{n-1}\varepsilon_i^2 V_i^2\sigma_i^2\int_{(T_i,T_{i+1}]}\left(\nu^{i,*}\left([t,T_{i+1}]\right)\right)^2 dt. \tag{27}$$

ii) The characteristic function and density of $IS\left(X^{*}\right)$ *are given, respectively, by*

$$Ee^{itIS\left(X^{*}\right)}=Ee^{itA-\frac{1}{2}t^2B}=\varphi\left(t\right) \tag{28}$$

and

$$f_{IS\left(X^{*}\right)}\left(x\right)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{-itx}Ee^{itA-\frac{1}{2}t^2B}dt \tag{29}$$

$$=E\frac{1}{\sqrt{2\pi B}}\exp\left(-\frac{1}{2B}\left(x-A\right)^2\right),$$

provided $\varphi(t)$ is integrable, which holds when

$$E\left(\frac{1}{\sqrt{B}}\right) < \infty. \tag{30}$$

iii) Furthermore,

$$EIS(X^*) = \frac{1}{2}\sum_{k=0}^{n-1} E\left(\varepsilon_k^2 V_k^2 \mu_k\right)$$

and

$$Var(IS(X^*)) = VarA + EB$$
$$= Var\left(\frac{1}{2}\sum_{k=0}^{n-1}\varepsilon_k^2 V_k^2 \mu_k\right) + E\sum_{k=0}^{n-1}\varepsilon_k^2 V_k^2 \sigma_k^2 \int_{(T_k,T_{k+1}]} \left(\nu^{k,*}([t,T_{k+1}])\right)^2 dt.$$

Proof. Observe that

$$E\left(e^{itIS(X^*)}\middle|\varepsilon_k, V_k, \mu_k, \sigma_k, \int_{(T_k,T_{k+1}]} \left(\nu^{k,*}([t,T_{k+1}])\right)^2 dt;\ k = 0,1,...n-1\right)$$
$$= e^{itA - \frac{1}{2}t^2 B}$$

and (28) follows. To prove the integrability of φ, we note that

$$|\varphi(t)| = \left|Ee^{itA-\frac{1}{2}t^2B}\right| \leq Ee^{-\frac{1}{2}t^2B}$$

and, hence,

$$\int_{\mathbb{R}} E\left(e^{-\frac{1}{2}t^2B}\right) dt = E\left(\frac{1}{\sqrt{B}}\int_{\mathbb{R}} e^{-\frac{1}{2}t^2} dt\right) = \sqrt{2\pi}E\left(\frac{1}{\sqrt{B}}\right) < \infty.$$

Next, using that $\varphi'(0) = iEA$ and $\varphi''(0) = -EA^2 - EB$, we deduce that

$$Var(IS(X^*)) = -\varphi''(0) + (\varphi'(0))^2 = VarA + EB.$$

■

6 Types of market impact

The impact the execution of large trades has on market prices has been widely discussed in the literature. Not surprisingly, there are differences in the way this impact is defined and, in turn, measured. Our focus, so far, has been on the analysis of the implementation slippages induced by the execution strategies.

Next, we analyze the *realized*, *permanent* and *temporary* price impacts as defined in [1]. We denote by S_0 market price before order execution begins, by S_{post} market price after the order is completed, and by $\overline{S}$ volume weighted average realized price on the order.

The post trade price S_{post} should capture the permanent effect of the order execution on market prices. The argument used in [1] is that it should be taken "long enough" after the order is completed. Empirical analysis suggests that one half-hour is adequate to achieve this. Consequently, if execution stops at time 1, then $S_{post} = S_{1+\delta}$, where δ is this extra time which in [1] is one half-hour. Based on these prices, and following [1], we define the permanent impact as

$$I = \frac{S_{post} - S_0}{S_0},$$

and the realized impact as

$$J = \frac{\overline{S} - S_0}{S_0}.$$

The difference between the realized and permanent impacts is used to define the temporary price impact, namely,

$$J - I = Temporary + Noise.$$

We stress that all computations in [1] are performed in *volume time* τ_t, which represents the fraction of an average market volume that has been executed up to and including clock time t. Note, however, that the market volume time corresponds to our ν_t, which represents the fraction of the volume, for a given order, executed up to time t. This representation of a strategy is convenient because it does *not* depend on the specific volume to be executed, or traded, in the market. Until now, we did not take into consideration the volume traded in the market. Clearly we cannot execute an order for which the volume exceeds what is traded. From now on, we assume that the volume traded in the market over any interval within the trading period exceeds the volume we want to execute.

6.1 Single order arbitrary strategy

To incorporate the notation of [1] within our framework, we assume from now on that for each order i, $i = 0, 1, ..., n-1$ we use

$$I_i = \frac{S_{T_{i+1}} - S_{T_i}}{S_{T_i}}$$

to *represent the permanent* price impact. Note that we choose the *post* time to be T_{i+1} while the execution of order i is completed at time $\hat{T}_{i+1}$. Hence, the "gap" is given by

$$\delta_{i+1} = T_{i+1} - \hat{T}_{i+1}. \tag{31}$$

Recall that the impacted price process is given by

$$S_t = S_{T_i} + \int_{[T_i,t)} G_i\left(t-s\right) dX_s^i + \sigma_i \left(W_t - W_{T_i}\right), \quad t \in (T_i, T_{i+1}]$$

$$= S_{T_i} + \varepsilon_i V_i \int_{[T_i,t)} G_i\left(t-s\right) d\nu_s^i + \sigma_i\left(W_t - W_{T_i}\right), \quad t \in (T_i, T_{i+1}], \tag{32}$$

where σ_i is, in general, a random variable measurable with respect to $\mathcal{F}_{T_i}$ and such that $E\left(\sigma_i^4\right) < \infty$.

The order is executed over the interval $\left[T_i, \hat{T}_{i+1}\right]$ with $\hat{T}_{i+1} < T_{i+1}$. The process S is left continuous with right hand limits on the interval $\left[T_i, \hat{T}_{i+1}\right]$ and continuous on $\left(\hat{T}_{i+1}, T_{i+1}\right]$. Therefore,

$$S_{T_{i+1}} - S_{T_i} = \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i\left(T_{i+1}-s\right) d\nu_s^i + \sigma_i\left(W_{T_{i+1}} - W_{T_i}\right).$$

Consequently, the permanent price impact takes the form

$$I_i = \frac{\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i\left(T_{i+1}-s\right) d\nu_s^i + \sigma_i\left(W_{T_{i+1}} - W_{T_i}\right)}{S_{T_i}},$$

using the notation of (32).

We now make the following observations:

1. In order to compare the model proposed in [1] with (32) we must first reconcile the notation in both models. Recall that in (32) the process S represents price, say, in dollars. However, the strategy X^i and volume V_i represent number of shares. Consequently, we have *unit inconsistency.* Fortunately, there is a simple remedy to this issue, which also helps to reconcile both models. Indeed, note that the relative performance of the strategies, and in particular of the optimal strategy for a given function G_i, does not change when we multiply G_i by an $\mathcal{F}_{T_i}-$measurable quantity C_i. It then turns out that to compare model [1] with (32) one simply needs to take $C_i = S_{T_i}$. Then, the new decay function G_i^{new} becomes

$$G_i^{new} = S_{T_i} G_i. \tag{33}$$

2. Another difficulty we encounter is how to interpret the stock's volatility. In (32) , σ represents the so called normal volatility of the Bachelier model. In [1], however, volatility represents the log-normal volatility of the Black and Scholes model. In order to compare both models, we just need to convert the normal volatility to the log-normal one. We do this by setting

$$\sigma_i^{new} = S_{T_i} \sigma_i, \tag{34}$$

where now σ_i represents the log-normal volatility. This change in notation brings model (32) closer to the framework of the model in [1].

3. There is still one remaining difference related to how time is treated in [1]. Model (32) runs in *clock time* while the model in [1] runs in *volume time*. We will focus on this difference later on. For now, we introduce new parametrization to our model, which still runs in clock time. Specifically, we take, for $t \in (T_i, T_{i+1}]$,

$$\begin{aligned} S_t &= S_{T_i} + \int_{[T_i,t)} G_i^{new}(t-s)\,dX_s^i + \sigma_i^{new}(W_t - W_{T_i}) \\ &= S_{T_i} + \varepsilon_i V_i \int_{[T_i,t)} G_i^{new}(t-s)\,d\nu_s^i + \sigma_i^{new}(W_t - W_{T_i}), \end{aligned} \tag{35}$$

where G_i^{new} and σ_i^{new} are given by (33) and (34), respectively.

Using "*new*" notation, the permanent price impact takes the form

$$\begin{aligned} I_i &= \frac{\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i^{new}(T_{i+1}-s)\,d\nu_s^i + \sigma_i^{new}\left(W_{T_{i+1}} - W_{T_i}\right)}{S_{T_i}} \\ &= \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\,d\nu_s^i + \sigma_i\left(W_{T_{i+1}} - W_{T_i}\right). \end{aligned} \tag{36}$$

In order to compare notation and results in [1] with our framework, we use the "*new*" parameters.

Proposition 18 *The conditional on $\mathcal{F}_{T_i}$ distribution of the permanent impact I_i is normal with mean*

$$E\left(I_i \middle| \mathcal{F}_{T_i}\right) = \frac{\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i^{new}(T_{i+1}-s)\,d\nu_s^i}{S_{T_i}} = \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\,d\nu_s^i$$

and variance

$$Var\left(I_i \middle| \mathcal{F}_{T_i}\right) = \frac{\left(\sigma_i^{new}\right)^2}{S_{T_i}^2}(T_{i+1}-T_i) = \sigma_i^2 (T_{i+1}-T_i),$$

where σ_i denotes log-normal volatility.

The realized impact J defined in [1] corresponds in (35) to

$$J_i = \frac{\int_{[T_i,T_{i+1}]} S_t d\nu_t^i - \varepsilon_i S_{T_i}}{S_{T_i}},$$

where $\int_{[T_i,T_{i+1}]} S_t d\nu_t^i$ represents the realized price. The latter is, in fact, the price weighted by the volume of order i, since

$$\frac{\int_{[T_i,T_{i+1}]} S_t dX_t^i}{X^i([T_i,T_{i+1}])} = \int_{[T_i,T_{i+1}]} S_t d\nu_t^i.$$

Using (**??**) we deduce that

$$\int_{[T_i,T_{i+1}]} S_t dX_t^i = S_{T_i}\varepsilon_i V_i + \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i^{new}(t-s)\, dX_s^i dX_t^i + \int_{[T_i,T_{i+1}]} \sigma_i^{new} W_t dX_t^i$$

$$= S_{T_i}\varepsilon_i V_i + \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} S_{T_i} G_i(t-s)\, dX_s^i dX_t^i + \int_{[T_i,T_{i+1}]} S_{T_i}\sigma_i W_t dX_t^i$$

$$= S_{T_i}\left(\varepsilon_i V_i + \varepsilon_i^2 V_i^2 \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i + \varepsilon_i V_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i\right).$$

Consequently,

$$\int_{[T_i,T_{i+1}]} S_t d\nu_t^i = S_{T_i}\left(\varepsilon_i + \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i + \varepsilon_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i\right)$$

and, hence,

$$J_i = \varepsilon_i + \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i + \varepsilon_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i - \varepsilon_i$$

$$= \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i + \varepsilon_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i. \qquad (37)$$

Proposition 19 *The conditional on $\mathcal{F}_{T_i}$ distribution of the realized impact J_i is normal with mean*

$$E\left(J_i|\,\mathcal{F}_{T_i}\right) = \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]}\int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i$$

and variance

$$Var\left(J_i|\,\mathcal{F}_{T_i}\right) = \varepsilon_i^2\sigma_i^2 \int_{[T_i,T_{i+1}]}\int_{[T_i,T_{i+1}]} (t\wedge s - T_i)\, d\nu_s^i d\nu_t^i$$

$$= \varepsilon_i^2\sigma_i^2 \int_{(T_i,T_{i+1}]} \left(\nu^i([t,T_{i+1}])\right)^2 dt.$$

We easily deduce that

$$IS\left(X^i\right) = \frac{1}{2}\int_{[T_i,T_{i+1}]}\int_{[T_i,T_{i+1}]} G_i^{new}(|t-s|)\, dX_s^i dX_t^i + \int_{[T_i,T_{i+1}]} \sigma_i^{new} W_t dX_t^i$$

$$= S_{T_i}\left(\frac{1}{2}\varepsilon_i^2 V_i^2 \int_{[T_i,T_{i+1}]}\int_{[T_i,T_{i+1}]} G_i(|t-s|)\, d\nu_s^i d\nu_t^i + \varepsilon_i V_i \sigma_i \int_{[T_i,T_{i+1}]} W_t d\nu_t^i\right).$$

Consequently, relative to the value of volume traded (i.e. relative to $S_{T_i}V_i$) the implementation slippage, defined by

$$R_i := \frac{IS\left(X^i\right)}{S_{T_i} V_i},$$

is given by

$$R_i = \frac{1}{2}\varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]} \int_{[T_i,T_{i+1}]} G_i\left(|t-s|\right) d\nu_s^i d\nu_t^i + \varepsilon_i \sigma_i \int_{[T_i,T_{i+1}]} W_t d\nu_t^i. \quad (38)$$

The conditional on $\mathcal{F}_{T_i}$ distribution of R_i is normal with mean

$$E\left(R_i|\,\mathcal{F}_{T_i}\right) = \frac{1}{2}\varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]} \int_{[T_i,T_{i+1}]} G_i\left(|t-s|\right) d\nu_s^i d\nu_t^i$$

and variance

$$Var\left(R_i|\,\mathcal{F}_{T_i}\right) = \varepsilon_i^2 \sigma_i^2 \int_{(T_i,T_{i+1}]} \left(\nu^i\left([t, T_{i+1}]\right)\right)^2 dt.$$

Concepts of permanent, realized and temporary price impacts defined in [1] are derived within models referring to the *rate of trading*. Such models, however, have *continuous* trajectories. In our framework the price process *may have jumps and the jumps are priced using space and not time infinitesimal arguments*.

When the execution strategy is infinitesimal in time we refer to the rate of trading. However, when the execution strategy involves a jump, which can be associated with a block trade, that jump is priced (in a fixed time) through infinitesimal arguments in space. Here, the order book dynamics are used to develop the price concept. It turns out that the block trades are implemented at the *average* price of S, calculated before and after the jump they generate (this was first observed in [5] and can be easily extended herein).

The unconditional distributions of I_i, J_i and R_i can be easily derived using arguments developed so far. In particular, the density $f_{R_i}(x)$ of R_i is given by

$$f_{R_i}(x) = E\frac{1}{\sqrt{2\pi Var\left(R_i|\,\mathcal{F}_{T_i}\right)}} \exp\left(-\frac{1}{2Var\left(R_i|\,\mathcal{F}_{T_i}\right)}\left(x - E\left(R_i|\,\mathcal{F}_{T_i}\right)\right)^2\right),$$

provided $\varepsilon_i^2 = 1$ and

$$E\frac{1}{\sqrt{Var\left(R_i|\,\mathcal{F}_{T_i}\right)}} < \infty.$$

The densities for I_i and J_i are derived in the same way. It is, also, straightforward to show that

$$R_i - J_i = \frac{1}{2}G_i(0)\,\varepsilon_i^2 V_i \sum_{T_i \le t \le T_{i+1}} \left(\Delta\nu_t^i\right)^2$$

and, hence, $R_i - J_i \in \mathcal{F}_{T_i}$. Moreover,

$$J_i - I_i = \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]} \int_{[T_i,t[} G_i(t-s)\, d\nu_s^i d\nu_t^i + \varepsilon_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i$$

$$-\left(\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i + \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right)\right)$$

$$= \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]} \int_{[T_i,t[} G_i(t-s)\, d\nu_s^i d\nu_t^i - \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i$$

$$+\varepsilon_i \int_{[T_i,T_{i+1}]} \sigma_i W_t d\nu_t^i - \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right).$$

Therefore, in accordance with the definition given in [1], the quantity

$$K_i = \varepsilon_i^2 V_i \int_{[T_i,T_{i+1}]} \int_{[T_i,t)} G_i(t-s)\, d\nu_s^i d\nu_t^i - \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i$$

represents the *temporary impact*. Clearly, we also have

$$R_i - J_i + I_i = \frac{1}{2} G_i(0)\, \varepsilon_i^2 V_i \sum_{T_i \le t \le T_{i+1}} \left(\Delta \nu_t^i\right)^2$$

$$+\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i + \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right).$$

Introducing

$$L_i := \frac{1}{2} G_i(0)\, \varepsilon_i^2 V_i \sum_{T_i \le t \le T_{i+1}} \left(\Delta \nu_t^i\right)^2 + \varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i$$

we obtain the decomposition

$$R_i - J_i + I_i = L_i + \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right). \tag{39}$$

The term $\varepsilon_i V_i \int_{[T_i,T_{i+1}]} G_i(T_{i+1}-s)\, d\nu_s^i$ represents the conditional mean of the permanent impact, in the terminology of [1].

On the other hand, the quantity $\frac{1}{2} G_i(0)\, \varepsilon_i^2 V_i \sum_{T_i \le t \le T_{i+1}} \left(\Delta \nu_t^i\right)^2$ represents the impact of jumps which, however, are *excluded* by the model used in [1].

6.2 Single order optimal strategy and the power law

The optimal fraction $\nu^{i,*}$ corresponding to the optimal strategy $X^{i,*}$ of the future order i satisfies for $i = 0, 1, ...n-1$,

$$\int_{[T_i,\hat{T}_{i+1}]} G_i^{new}(|t-s|)\, d\nu_s^{i,*} = \mu_i^{new}, \text{ for each } t \in \left[T_i, \hat{T}_{i+1}\right] \tag{40}$$

with $\mu_i^{new} = S_{T_i}\mu_i$, where for all $t \in \left[T_i, \hat{T}_{i+1}\right]$ and $i = 0, 1, ..., n-1$,

$$\int_{\left[T_i, \hat{T}_{i+1}\right]} G_i\left(|t-s|\right) d\nu_s^{i,*} = \mu_i, \text{ for each } t \in \left[T_i, \hat{T}_{i+1}\right].$$

Working as in the single order case in [3], we deduce the following result.

Proposition 20 *For $i = 1, ..., n-1$, the random variables I_i, J_i and R_i satisfy*

$$E\left(I_i \middle| \mathcal{F}_{T_i}\right) = \varepsilon_i V_i \mu_i,$$

$$E\left(J_i \middle| \mathcal{F}_{T_i}\right) = \frac{1}{2}\varepsilon_i^2 V_i \left(\mu_i - 2G_i(0)\left(\Delta\nu_{T_i}^i\right)^2\right)$$

and

$$E\left(R_i \middle| \mathcal{F}_{T_i}\right) = \frac{1}{2}\varepsilon_i^2 V_i \mu_i.$$

Furthermore,

$$E\left(J_i \middle| \mathcal{F}_{T_i}\right) < E\left(R_i \middle| \mathcal{F}_{T_i}\right) \quad \text{and} \quad Var\left(J_i \middle| \mathcal{F}_{T_i}\right) = Var\left(R_i \middle| \mathcal{F}_{T_i}\right).$$

Extensive empirical analysis (see ([7] and references within) validates the so-called *power law* which states that the average relative price change is well described by the formula

$$Y\sigma\left(\frac{Q}{V}\right)^{\delta},$$

where σ is the daily volatility of the asset, Q is the volume of a metaorder, and V is the daily traded volume. The numerical constant Y is of order unity. The daily volatility σ and the daily volume V are measured contemporaneously to the trade. As indicated in [7], the power law holds for the levels of $\frac{Q}{V}$ ranging from a few 10^{-4} to a few %. Depending on the markets and on the contracts' types, δ varies between 0.5 and 0.7. It then follows from (36) that the average relative price change is given by

$$E\left(I_i \middle| \mathcal{F}_{T_i}\right) = \varepsilon_i V_i \int_{[T_i, T_{i+1}]} G_i\left(T_{i+1} - s\right) d\nu_s^i$$

for an admissible strategy ν^i, and by

$$E\left(I_i^* \middle| \mathcal{F}_{T_i}\right) = \varepsilon_i V_i \mu_i$$

for the optimal $\nu^{i,*}$. It is important to note that the average has been calculated under the *conditional* on $\mathcal{F}_{T_i}$ distribution. Of course, the average relative price changes, calculated under the unconditional distribution, follow trivially from the above expressions.

Applied at the optimal strategy $\nu^{i,*}$, the power law states that, under the conditional on $\mathcal{F}_{T_i}$ distribution, we have

$$\varepsilon_i V_i \mu_i = \varepsilon_i \lambda_i = \varepsilon_i Y_i \sigma_{D_i}\left(\frac{V_i}{V_{D_i}}\right)^{\delta},$$

where σ_{D_i} is the asset volatility and V_{D_i} the volume traded on the day D_i, both measured contemporaneously to the trade. Consequently, σ_{D_i} and V_{D_i} are *not measurable* with respect to $\mathcal{F}_{T_i}$. However, both ε_i and λ_i are and, hence, the above equality *cannot hold.* This, however, *poses a problem* to the extended model (32) we introduced, in which the order i characteristics, the decay function and the volatility, must be measurable with respect to $\mathcal{F}_{T_i}$.

One way to address this inconsistency is to replace the power law by its *estimate based on the data* which are *not contemporaneous* to the trade but, rather, use the daily volatilities and volumes collected on the days preceding the trade. Then, the power law yields

$$V_i\mu_i = \lambda_i = Y_i V_i^{\delta} E\left(\frac{\sigma_{D_i}}{V_{D_i}^{\delta}} \,|\mathcal{F}_{T_i}\right),$$

where the random variable $E\left(\frac{\sigma_{D_i}}{V_{D_i}^{\delta}} \,|\mathcal{F}_{T_i}\right)$ represents an estimate of the quantity $\frac{\sigma_{D_i}}{V_{D_i}^{\delta}}$. Indeed, because

$$\begin{aligned} I_i^* &= \varepsilon_i V_i \mu_i + \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right) \\ &= \varepsilon_i Y_i \left(V_i\right)^{\delta} E\left(\frac{\sigma_{D_i}}{V_{D_i}^{\delta}} \,|\mathcal{F}_{T_i}\right) + \varepsilon_i Y_i \sigma_{D_i} \left(\frac{V_i}{V_{D_i}}\right)^{\delta} \\ &\quad - \varepsilon_i Y_i V_i^{\delta} E\left(\frac{\sigma_{D_i}}{V_{D_i}^{\delta}} \,|\mathcal{F}_{T_i}\right) + \sigma_i \left(W_{T_{i+1}} - W_{T_i}\right), \end{aligned}$$

we obtain the following, conditionally on $\mathcal{F}_{T_i}$, power law

$$E\left(I_i^* \,|\mathcal{F}_{T_i}\right) = \varepsilon_i Y_i V_i^{\delta} E\left(\frac{\sigma_{D_i}}{V_{D_i}^{\delta}} \,|\mathcal{F}_{T_i}\right).$$

When the daily volume V_{D_i} and the daily volatility σ_{D_i} are estimated using the daily volatilities and volumes collected on the days preceding the trade, they are not estimated contemporaneously to the trade (which we assume from now on). Then, the power law becomes

$$E\left(I_i^* \,|\mathcal{F}_{T_i}\right) = \varepsilon_i Y_i \sigma_{D_i} \left(\frac{V_i}{V_{D_i}}\right)^{\delta} = \varepsilon_i V_i \mu_i = \varepsilon_i \lambda_i. \tag{41}$$

Calibration to the average permanent impact corresponds to the choice of the decay function for which (41) holds. When the decay function is a simple exponential, calibration comes down to the specification of κ for which (41) holds. Recall that, in this case, $\mu_i = \frac{\lambda_i}{V_i} = \frac{2}{2+\kappa_i(T_{i+1}-T_i)}$ and, hence, we must have

$$\frac{2}{2+\kappa_i\left(T_{i+1}-T_i\right)} = \frac{1}{V_i} Y_i \sigma_{D_i} \left(\frac{V_i}{V_{D_i}}\right)^{\delta}.$$

This, in turn, gives

$$\kappa_i = \frac{2}{(T_{i+1} - T_i)} \left(\frac{V_i^{1-\delta} V_{D_i}^{\delta}}{Y_i \sigma_{D_i}} - 1 \right).$$

Note, however, that calibration is only possible when $\kappa_i > 0$. Hence, we must have

$$V_i^{1-\delta} V_{D_i}^{\delta} > Y_i \sigma_{D_i},$$

which, in turn, imposes *constraints* on the volumes of the orders and the stock's volatility.

7 Market impacts of sequential orders

When several orders are executed during a day, one may want to consider variations of the *aggregate impact* they generate in the market. Our model may give some insights as to how impacts of the individual orders are aggregated. Recall that the slippage of a sequential execution strategy is defined as the sum of slippages of individual orders. We adopt the same logic here and aggregate different impacts of individual orders by taking their sums. The idea then is to measure at the model level the impact that the dependence structure of the order flow will have on the distributions of the aggregate impacts.

From the previous analysis, we already know that the cost of execution may *not* be the *only* criterion one wants to consider when optimizing execution of a single or of sequential orders. *Distributions of impacts may help to define an alternative optimality criterion.* For example, small average impact may be associated with its relatively large variance. On the other hand, small variance may be associated with relatively large average impacts.

Rational choices for a single order have been discussed by the authors in [5]. However, when we aggregate impacts during a day, the situation is conisderably less clear. This is because a lot will depend on the nature of the order flow.

Indeed, when we calculate daily permanent impact by summing up permanent impacts of the individual orders, we will notice a difference among the days during which most of the orders are of the same type, say, to buy or to sell, and the days when we get balanced flow of buy and sell orders. This is because on a "balanced day", average permanent impacts may compensate each other to generate a relatively small daily average impact.

This logic is very natural and is similar to the strategy of order aggregation, currently being studied in [6]. The idea therein is to move away from sequential execution of orders and adopt an alternative strategy where one aims for the longest, in calendar time, execution of each order and, hence, needs to manage orders at a portfolio level. There are many ways one could manage such a portfolio. We propose one such approach in [6].

8 Volume time

The main focus of this paper is to provide a framework in which one can analyze and measure the cost of sequential execution of several orders within a fixed period of time, for example, during a day. However, the market activity changes substantially between different periods of the day. Ten-day average intra-day volume and volatility profiles, on 15-minute intervals, are presented in [1].

Volumes tend to decrease from the market opening. They stabilize mid-day and tend to increase towards the end trading. Daily volume profile looks like a symmetric convex function of time during the day. This is a good news for our model because densities of optimal strategies are symmetric and convex, and hence are consistent with the liquidity profile the market provides. However, when we need to execute sequentially several orders during a single day, *symmetric strategies will face non-symmetric liquidity profiles.* This should be taken into account in the execution algorithms.

As demonstrated in [1], the volatility falls initially and then stays roughly at the same level until the close of trading. We have given some attention so far to the impact the volatility parameter σ_i may have on the probabilistic properties of our model outputs. On the other hand, the daily volatility appears in the so called power law describing the average permanent impacts. Daily volumes also appear in the power law. Intra-day average volumes were used in [1] to define the concept of *volume time*, which is an example of business time studied in details in [4].

We now introduce volume time into our model. To this aim, we define increasing cadlag processes A^i_t on the time interval $[T_i, T_{i+1}]$ with $A^i_{\hat{T}_i} = 0$ and $A^i_{\hat{T}_{i+1}}$. The amount A^i_t represents the *average market executed volume* over the interval $[T_i, t]$ and

$$\tau^i_t := \frac{A^i_t}{A_{\hat{T}_{i+1}}}$$

stands for the fraction of $A_{\hat{T}_{i+1}}$ executed over the interval $[T_i, t]$. We assume that the trading stops at clock time T_{i+1}.

References

[1] R. Almgren, C. Thum, E. Hauptmann, and H. Li, Direct estimation of equity market impact. Preprint.

[2] G. Bordigoni, A. Figalli, A. Ledford, P. Ustinov, Strategic execution Strategies, Preprint, November 2020.

[3] J. Gatheral, A. Schied, A. Slynko, Transient linear price impact and Fredholm integral equations, Mathematical Finance: An International Journal of Mathematics, *Statistics and Financial Economics*, 22 (3), 445-474, 2012.

[4] A. Kyle, A. Obizhaeva, *Econometrica*, Marker microstructure invariance: empirical hypothesis, Vol. 84, No.4 (July 2016), pp. 1345-1404.

[5] M. Musiela and T. Zariphopoulou, On Gatheral's model: revision, reformulation and extensions, Preprint, March 2022.

[6] M. Musiela and T. Zariphopoulou, On the aggregation of orders in optimal execution, work in progress, 2022.

[7] B. Toth, Y. Lemperiere, C. Deremble, J. de Lataillade, J. Kockelkoren, J. -P. Bouchaud, Anomalous price impact and the critical nature of liquidity in financial markets, June 29, 2021 arXiv: 1105.1694v.