Optimal execution and absence of price manipulations in limit order book models

Aurélien Alfonsi
Université Paris-Est, CERMICS,
Project team MathFi ENPC-INRIA-UMLV,
Ecole des Ponts, 6-8 avenue Blaise Pascal,
77455 Marne La Vallée, France
alfonsi@cermics.enpc.fr

Alexander Schied
Zentrum Mathematik
TU Munich
Boltzmannstr. 3
85748 Munich, Germany
schied@tum.de

June 22, 2009

Abstract: We continue the analysis of optimal execution strategies in the model for a limit order book with nonlinear price impact and exponential resilience that was considered in our earlier paper with A. Fruth [2]. We now allow for non-homogeneous resilience rates and arbitrary trading dates and consider the extended problem of optimizing jointly over trading dates and sizes. Our main results show that, under general conditions on the shape function of the limit order book, placing the deterministic trade sizes identified in [2] at trading dates that are homogeneously spaced is optimal also within the large class of adaptive strategies with arbitrary trading dates. Perhaps even more importantly, our analysis yields as a corollary that our model does not admit price manipulation strategies in the sense of Huberman and Stanzl [14]. This latter result contrasts the recent findings of Gatheral [13], where, in a related but different model, exponential resilience was found to give rise to price manipulation strategies when price impact is nonlinear.

1 Introduction.

The problem of optimal execution is concerned with the optimal acquisition or liquidation of large asset positions. In doing so, it is usually beneficial to split up the large order into a sequence of partial order, which are then spread over a certain time horizon, so as to reduce the overall price impact and the execution costs. The optimization problem at hand is thus to find a trading strategy that minimizes a cost criterion under the constraint of overall order execution within a given time frame. There are several reasons why studying this problem is interesting.

First, liquidity risk is one of the least understood sources of financial risk, and one of its various aspects is the risk resulting from price impact created by trading large positions. Due to the nonlinear feedback effects on dynamic trading strategies, market impact risk is probably also among the most fascinating aspects of liquidity risk for mathematicians. The optimal execution problem allows studying market impact risk in its purest form. Moreover, the results obtained for this problem can serve as building blocks in a realistic analysis of more complex problems such as the hedging of derivatives in illiquid markets.

Second, the mathematical analysis of optimal execution strategies can help in the ongoing search for viable market impact models. As argued by Huberman and Stanzl [14] and Gatheral [13], any reasonable market impact model should not admit *price manipulation strategies* in the sense that there are no 'round trips' (i.e., trading strategies with zero balance in shares), whose expected trading costs are negative. Since every round trip can be regarded as the execution of a zero-size order, a solution of the optimal execution problem also includes an analysis of price manipulation strategies in the model (at least as limiting case when the order size tends to zero).

In recent years, the problem of optimal execution was considered for various market impact models and cost functions by authors such as Bertsimas and Lo [10], Almgren and Chriss [5, 6], Almgren [4], Obizhaeva and Wang [16], Almgren and Lorenz [7], the second author and Schöneborn [18, 19], our joint papers with A. Fruth [1, 2], and Bayraktar and Ludkovski [9], to mention only a few.

Here, we continue our analysis from [1, 2]. The underlying model is a time-inhomogeneous version of the one in [2] and thus an extension of the limit order book model with linear impact and exponential resilience that was proposed by Obizhaeva and Wang [16]. In our model, the ask part of the limit order book consists of a certain distribution of shares offered in the form of limit orders at prices higher than the current best ask price. When the large trader has not yet been active, the best ask price fluctuates according to the actions of noise traders. A buy market order placed by the large trader consumes a block of shares located immediately to the right of the best ask and thus increases the ask price by a certain amount, the price impact. Since the distribution of limit orders is allowed to be non-uniform, the price impact created by a market order is typically a nonlinear function of the order size. In reaction to the increased bid-ask spread, the best ask price will recover from the impact of the buy order within a certain time span, i.e., it will show a certain resilience. Thus, the price impact of a market order will neither be completely instantaneous nor entirely permanent but will decay exponentially with a time-dependent resilience rate. As in [2], we consider the following two distinct possibilities for modeling the resilience of the limit order book after a large market order: the exponential recovery of the number of limit orders, i.e., of the volume of the limit order book (Model 1), or the exponential recovery of the bid-ask spread (Model 2).

This model is quite close to descriptions of price impact on limit order books found in empirical studies such as Biais *et al.* [11], Potters and Bouchaud [17], Bouchaud *et al.* [12], and Weber and Rosenow [20]. In particular, the existence of a strong resilience effect, which stems from the placement of new limit orders close to the bid-ask spread, seems to be a well established fact, although its quantitative features seem to be the subject of an ongoing discussion. As for the order of the immediate price impact, our assumptions include in particular the cases of power-law impact as proposed in Almgren *et al.* [8] and of logarithmic impact as suggested in Bouchaud *et al.* [12].

While in [2] we considered only strategies whose trades are placed at equidistant times, we now allow the trading times to be stopping times. This problem description is clearly much more natural than prescribing *a priori* the dates at which trading may take place. It is also more realistic than the idealization of trading in continuous time. In addition, the time-inhomogeneous description allows us to account for time-varying liquidity and thus in particular for the well-known U-shape patterns in intraday market parameters; see, e.g., [15].

Optimal execution in this extended framework leads to the problem of optimizing simultaneously over both trading times and sizes. This problem is more complex than the one considered in [2] and requires new arguments. Nevertheless, our main results show that the unique optimum is attained by placing the deterministic trade sizes identified in [2] at trading dates that are homogeneously spaced with respect to the average resilience rate in between trades.

As a corollary, we show that neither of the two variants of our model admits price manipulation strategies in the sense of Huberman and Stanzl [14] and Gatheral [13], provided that the shape function of the limit order book belongs to a certain class of functions (which slightly differs for each variant). In either case, this class of functions contains the power-law family and logarithmic price impact. This corollary is fairly surprising and puzzling in view of recent results by Gatheral [13]. There it was shown that, in a closely related market impact model, exponential resilience leads to the existence of price manipulation strategies as soon as price impact is nonlinear.

As of now, it is not clear to us which of the features in Gatheral's model that are different from ours are responsible for the creation of price manipulation strategies. In fact, one of our motivations for the research reported in this paper was to understand whether the restriction to equidistant time grids in [2] prevented the emergence of price manipulation strategies in our model.

In view of our results on the nonexistence of price manipulation strategies, we must therefore deduce that, in contrast to Gatheral's conclusion, exponential resilience of a limit order book is a viable possibility for describing the decay of market impact, at least from a theoretical perspective.

This paper is organized as follows. In Section 2.1 we introduce the limit order book model with its two variants. The cost optimization problem is explained in Section 2.2. In Section 2.3 we state our main results for the case of a block-shaped limit order book, which corresponds to linear price impact. This special case is much simpler than the case with nonlinear price impact. We therefore give a self-contained description and proof for this case, so that the reader can gain a quick intuition on why our results are true. The proofs for the block-shaped case rely on the results from our earlier paper [1] with A. Fruth and are provided in Section 3.2. As in [2], we refer to the variant of our model with exponential resilience for the volume of the limit order book as Model 1. The main results for this case are stated in Section 2.4. The corresponding proofs are given in Section 3.3. Model 2 refers to the exponential resilience of the bid-ask spread. The results for this variant are stated in Section 2.5, while proofs are given in Section 3.4.

2 Setup and main results

In this section we first introduce the two variants of our model and state the optimization problem. We then state our results for the particularly simple case of a block-shaped limit order book. Subsequently, we formulate our theorems for each model variant individually.

2.1 Model description

The model variants that we consider here are time-inhomogeneous versions of the ones in [2]. The general aim is to model the dynamics of a limit order book that is exposed to repeated market orders by a large trader, whose goal is to purchase a large amount $X_0 > 0$ of shares within a certain time period $[0,T]$. Hence, emphasis is on buy orders, and we concentrate first on the upper part of the limit order book, which consists of shares offered at various ask prices. The lowest ask price at which shares are offered is called the *best ask price*. Symmetric statements hold of course for the liquidation of a given asset position of $X_0 > 0$ shares.

When the large trader is inactive, the dynamics of the limit order book are determined by the actions of noise traders only. We assume that the corresponding *unaffected best ask price* A^0 is an arbitrary martingale on a given filtered probability space $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$ and satisfies $A^0_0 = A_0$.

Above the unaffected best ask price A^0_t, we assume a continuous ask price distribution for available shares in the limit order book: the number of shares offered at price $A^0_t + x$ is given by $f(x)\,dx$ for a bounded and continuous density function $f : \mathbb{R}\backslash\{0\} \longrightarrow (0,\infty)$. We call f the *shape function* of the limit order book. It determines the impact of a market order placed by our large trader. The choice of a constant shape function corresponds to linear market impact and to the block-shaped limit order book model of Obizhaeva and Wang [16].

The *actual best ask price* at time t, i.e., the best ask price after taking the price impact of previous buy orders of the large trader into account, is denoted by A_t. It is clearly above the unaffected best ask price, and the *extra spread* caused by the actions of the large trader is denoted by

$$D^A_t := A_t - A^0_t.$$

A buy market order of $x_t > 0$ shares placed by the large trader at time t will consume all the shares offered at prices between A_t and

$$A_{t+} := A_t + D^A_{t+} - D^A_t = A^0_t + D^A_{t+},$$

where D^A_{t+} is determined by the condition

$$\int_{D^A_t}^{D^A_{t+}} f(x)dx = x_t.$$

Thus, the process D^A captures the impact of market orders on the current best ask price. Clearly, the price impact $D^A_{t+} - D^A_t$ will be a nonlinear function of the order size x_t unless f is constant between D^A_t and D^A_{t+}.

Example 2.1. Consider the shape functions of the form

$$f(x) = \begin{cases} \dfrac{q_+}{(1+\lambda_+ x)^{\alpha_+}} & \text{for } x > 0, \\[2ex] \dfrac{q_-}{(1-\lambda_- x)^{\alpha_-}} & \text{for } x < 0, \end{cases}$$

where $q_\pm$ and $\lambda_\pm$ are positive constants and $\alpha_\pm \in (0,1]$. For $\alpha_+ = 1$, D^A_{t+} will depend logarithmically on the trade size x_t, a functional dependence suggested by Bouchaud *et al.* [12].

For $0 < \alpha_+ < 1$ we get the power law impact found by Almgren *et al.* [8]. The particular case $\alpha_+ = 1/2$ corresponds to the popular square-root dependence; see [13, Section 6.2] for a discussion. The parameters q_-, λ_-, and α_- will of course determine the shape of the limit order book for sell orders.

The quantity

$$E_t^A = \int_0^{D_t^A} f(x)dx, \tag{1}$$

describes the volume impact in the limit order book present at time t. By introducing the antiderivative

$$F(z) = \int_0^z f(x)\,dx \tag{2}$$

of f, the relation (1) can also be expressed as

$$E_t^A = F(D_t^A) \qquad \text{and} \qquad D_t^A = F^{-1}(E_t^A),$$

where we have used our assumption that f is strictly positive to obtain the second identity. For simplicity, the function F is supposed to be unbounded in the sense that

$$\lim_{x\uparrow\infty} F(x) = \infty \quad \text{and} \quad \lim_{x\downarrow-\infty} F(x) = -\infty, \tag{3}$$

i.e., we assume that the limit order book has infinite depth. Relaxing this assumption is possible but would require constraints on the order sizes and thus complicate the problem description.

When the large trader is inactive in between market orders, D^A and E^A revert back to zero on an exponential scale with a deterministic, time-dependent rate $t \mapsto \rho_t$, called *resilience speed*. More precisely, we will consider the following two model variants for the resilience of the market impact:

Model 1: E^A evolves according to

$$dE_t^A = -\rho_t E_t^A\,dt$$

if the large investor is not placing buy orders during the time interval $[t, t+s)$.

Model 2: D^A evolves according to

$$dD_t^A = -\rho_t D_t^A\,dt$$

if the large investor is not placing buy orders during the time interval $[t, t+s)$.

With these features we model the well-established empirical fact that order books exhibit a certain resilience as to the price impact of a large buy market orders. That is, after the initial impact the best ask price reverts back to its previous position; cf. Biais *et al.* [11], Potters and Bouchaud [17], Bouchaud *et al.* [12], and Weber and Rosenow [20] for empirical studies.

Up to now, we have only described the effect of buy orders on the upper half of the limit order book. Since the overall goal of the larger trader is to buy $X_0 > 0$ shares up to time T, a restriction to buy orders would seem to be reasonable. However, we do not wish to exclude the *a priori* possibility that it could be beneficial to also *sell* some shares and to buy them back at a later point in time. To this end, we also model the impact of sell market orders on the lower part of the limit order book, which consists of a certain number of bids for shares at each price below the *best bid price*. As for ask prices, we will distinguish between an unaffected best bid price B_t^0 and the actual best bid price B_t, for which the price impact of previous sell orders of the large trader is taken into account. All we assume on the dynamics of B^0 is

$$B_t^0 \leq A_t^0 \qquad \text{at all times } t.$$

The distribution of bids below B_t^0 is modeled by the restriction of the shape function f to the domain $(-\infty, 0)$. In analogy to the ask part, we introduce the the *extra spread* in the bid price distribution,

$$D_t^B := B_t - B_t^0,$$

which will be nonpositive. A sell market order of $x_t < 0$ shares placed at time t will consume all the shares offered at prices between B_t and

$$B_{t+} := B_t + D_{t+}^B - D_t^B = B_t^0 + D_{t+}^B,$$

where D_{t+}^B is determined by the condition

$$x_t = \int_{D_t^B}^{D_{t+}^B} f(x)dx = F(D_{t+}^B) - F(D_t^B) = E_{t+}^B - E_t^B,$$

for $E_s^B := F(D_s^B)$. Note that F is defined via (2) also for negative arguments. If the large trader is inactive during the time interval $[t, t+s[$, then the processes D^B and E^B behave just as their counterparts D^A and E^A, i.e.,

$$\begin{aligned} dE_t^B &= -\rho_t E_t^B \, dt \qquad \text{in Model 1,} \\ dD_t^B &= -\rho_t D_t^B \, dt \qquad \text{in Model 2.} \end{aligned} \tag{4}$$

2.2 The cost optimization problem

When placing a single buy market order of size $x_t > 0$ at time t, the large trader will purchase $f(x)\,dx$ shares at price $A_t^0 + x$, with x ranging from D_t^A to D_{t+}^A. Hence, the total cost of the buy market order amounts to

$$\pi_t(x_t) := \int_{D_t^A}^{D_{t+}^A} (A_t^0 + x) f(x)\, dx = A_t^0 x_t + \int_{D_t^A}^{D_{t+}^A} x f(x)\, dx.$$

Similarly, for a sell market order $x_t < 0$, we have

$$\pi_t(x_t) := B_t^0 x_t + \int_{D_t^B}^{D_{t+}^B} x f(x)\, dx.$$

We assume that the large trader needs to buy a total of $X_0 > 0$ shares until time T and that trading can occur at $N+1$ trades within the time interval $[0, T]$. An *admissible sequence of trading times* will be a sequence $\mathcal{T} = (\tau_0, \dots, \tau_N)$ of stopping times such that $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_N = T$. For such an admissible sequence of trading times, $\mathcal{T}$, we define a $\mathcal{T}$*-admissible trading strategy* as a sequence $\boldsymbol{\xi} = (\xi_0, \xi_1, \dots, \xi_N)$ of random variables such that

- $\sum_{n=0}^{N} \xi_n = X_0$,
- each ξ_n is measurable with respect to $\mathcal{F}_{\tau_n}$,
- each ξ_n is bounded from below.

The quantity ξ_n corresponds to the size of the market order placed at time t_n. Note that we do not *a priori* require ξ_n to be positive, i.e., we also allow for intermediate sell orders, but we assume that there is some lower bound on sell orders. Finally, an *admissible strategy* is a pair $(\mathcal{T}, \boldsymbol{\xi})$ consisting of an admissible sequence of trading times $\mathcal{T}$ and a $\mathcal{T}$-admissible trading strategy $\boldsymbol{\xi}$.

The *average cost* $\mathcal{C}(\boldsymbol{\xi}, \mathcal{T})$ of an admissible strategy $(\boldsymbol{\xi}, \mathcal{T})$ is defined as the expected value of the total costs incurred by the consecutive market orders:

$$\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}) = \mathbb{E}\Big[\sum_{n=0}^{N} \pi_{\tau_n}(\xi_n) \Big]. \tag{5}$$

In our earlier paper with A. Fruth, [2], we considered the case of a constant resilience ρ and a fixed, equidistant time spacing $\mathcal{T}_{eq} = \{iT/N \,|\, i = 0, \dots, N\}$. In this setting, we determined trading strategies that minimize the cost $\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}_{eq})$ among all $\mathcal{T}_{eq}$-admissible trading strategies $\boldsymbol{\xi}$. Our goal in this paper consists in *simultaneously* minimizing over trade times and sizes, i.e., to minimize the cost $\mathcal{C}(\boldsymbol{\xi}, \mathcal{T})$ among all admissible strategies $(\boldsymbol{\xi}, \mathcal{T})$.

In our present setting of an inhomogeneous resilience function ρ_t it is natural to replace the equidistant time spacing by the *homogeneous time spacing*

$$\mathcal{T}^* = (t_0^*, \dots, t_N^*)$$

defined via

$$\int_{t_{i-1}^*}^{t_i^*} \rho_s \, ds = \frac{1}{N} \int_0^T \rho_s \, ds, \qquad i = 1, \dots, N.$$

We also define

$$a^* := e^{-\frac{1}{N} \int_0^T \rho_u du}. \tag{6}$$

Our main result states that, under certain technical assumptions, $\mathcal{T}^*$ is in fact the *unique optimal* time grid for portfolio liquidation with $N+1$ trades in $[0, T]$. In addition, the unique optimal $\mathcal{T}^*$-admissible strategies in Model 1 and 2 are given by the corresponding trading strategies in [2].

As a corollary to our main results, we are able to show that our models do not admit *price manipulation strategies* in the following sense, introduced by Huberman and Stanzl [14] (see also Gatheral [13]).

Definition 2.2. A *round trip* is an admissible strategy $(\overline{\boldsymbol{\xi}}, \overline{\mathcal{T}})$ such that $\sum_{i=0}^{N} \overline{\xi}_i = 0$. A *price manipulation strategy* is a round trip $(\overline{\boldsymbol{\xi}}, \overline{\mathcal{T}})$ such that $\mathcal{C}(\overline{\boldsymbol{\xi}}, \overline{\mathcal{T}}) < 0$.

Our result on the non-existence of profitable price manipulation strategies strongly contrasts Gatheral's conclusion [13] that "the widely-assumed exponential decay of market impact is compatible only with linear market impact."

2.3 Main results for the block-shaped limit order book

We first discuss our problem in the particularly easy case of a block-shaped limit order book in which $f(x) = q$. In that case Models 1 and 2 coincide. It follows from the results in [1] that for every admissible sequence of trading times $\mathcal{T} = (\tau_0, \dots, \tau_N)$ there is a $\mathcal{T}$-admissible trading strategy that minimizes the cost $\mathcal{C}(\cdot, \mathcal{T})$ among all $\mathcal{T}$-admissible trading strategies. This strategy can even be computed explicitly; see [1, Theorem 3.1]. In the following theorem we consider the problem of optimizing jointly over trading times and sizes.

Theorem 2.3. *In a block-shaped limit order book, there is a unique optimal strategy* $(\boldsymbol{\xi}^*, \mathcal{T}^*)$ *consisting of homogeneous time spacing* $\mathcal{T}^*$ *and the deterministic trading strategy* $\boldsymbol{\xi}^*$ *defined by*

$$\xi_0^* = \xi_N^* = \frac{X_0}{2 + (N-1)(1-a^*)} \qquad \text{and} \qquad \xi_1^* = \cdots = \xi_{N-1}^* = \xi_0^*(1-a^*), \tag{7}$$

where a^* *is as in* (6).

While the preceding theorem is a special case of or main results, Theorem 2.7 and Theorem 2.13, it admits a particularly easy proof based on the results in [1]. This proof is given in Section 3.2.

Corollary 2.4. *In a block-shaped limit order book, any nontrivial round trip has a strictly positive average cost. In particular, there are no profitable price manipulation strategies.*

Figure 1 gives an illustration of the situation when $\rho(t) = a + b\cos(t/(2\pi))$, $0 \le t \le 1$. For $a > b > 0$ the resilience is greater near the opening and the closure of the stock exchange. We plot here the relative gain, i.e., quotient of the respective expected costs, for the optimal strategies corresponding to the optimal time grid $\mathcal{T}^*$ and the equidistant time grid $\mathcal{T}_{eq}$. More precisely, we plot the quotient of the respective cost functions defined in equation (18) below.

2.4 Main results for Model 1

In this section we state our main results for Model 1. They hold under the following assumption, which includes the important cases of an immediate price impact following a power law or a logarithmic price impact; see Example 2.1.

Assumption 2.5. In Model 1, we assume in addition to (3) that the shape function f is nondecreasing on $\mathbb{R}_-$ and nonincreasing on $\mathbb{R}_+$.

We start by looking at optimal trading strategies when an admissible sequence of trading times $\mathcal{T} = (\tau_0, \dots, \tau_N)$ is fixed. If $\boldsymbol{\xi}$ is a $\mathcal{T}$-admissible trading strategy and it happens that $\tau_i = \tau_{i+1}$, then the corresponding trades, ξ_i and ξ_{i+1}, are executed simultaneously. We therefore say that two $\mathcal{T}$-admissible trading strategies $\boldsymbol{\xi}$ and $\overline{\boldsymbol{\xi}}$ are *equivalent* if $\xi_i + \xi_{i+1} = \overline{\xi}_i + \overline{\xi}_{i+1}$ $\mathbb{P}$-a.s. on $\{\tau_i = \tau_{i+1}\}$.

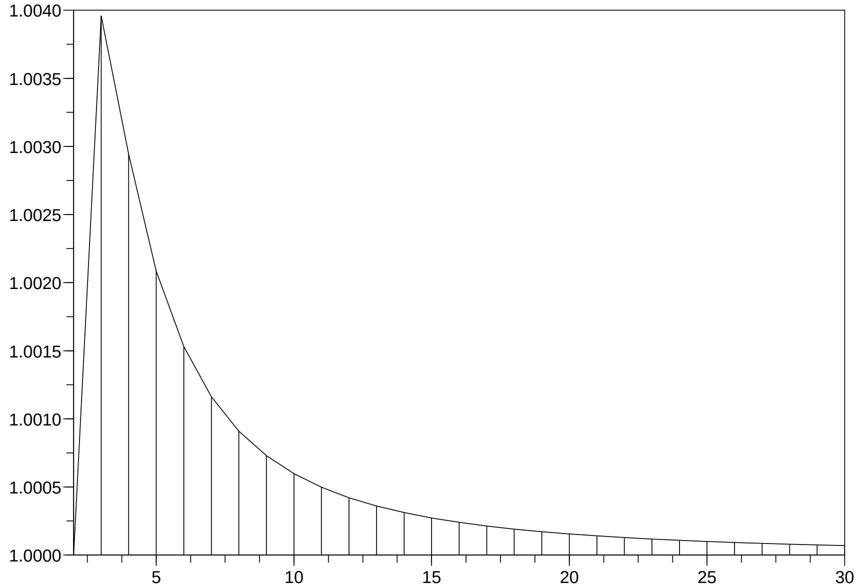

Figure 1: Relative gain between the extra liquidity cost of the optimal strategy on the optimal grid $\mathcal{T}^*$ and the optimal strategy on the equidistant grid $\mathcal{T}_{eq}$ as a function of N, when $T = 1$ and $\rho(t) = 2 + \cos(t/2\pi)$.

Proposition 2.6. *Suppose that an admissible sequence of trading times $\mathcal{T}$ is given and that Assumption 2.5 holds. Then there exists a $\mathcal{T}$-admissible trading strategy $\boldsymbol{\xi}^{(1),\mathcal{T}}$, unique up to equivalence, that minimizes the cost $\mathcal{C}(\cdot,\mathcal{T})$ among all $\mathcal{T}$-admissible trading strategies. Moreover, it consists only of nontrivial buy orders, i.e., $\xi_i^{(1),\mathcal{T}} > 0$ $\mathbb{P}$-a.s. for all i up to equivalence.*

As we will see in the proof of Proposition 2.6, the optimal trading strategy $\boldsymbol{\xi}^{(1),\mathcal{T}}$ can be implicitly characterized via a certain nonlinear equation. Our main result for Model 1 states, however, that things become much easier when optimizing simultaneously over trading times and sizes:

Theorem 2.7. *Under Assumption 2.5, there is a unique optimal strategy $(\boldsymbol{\xi}^{(1)},\mathcal{T}^*)$ consisting of homogeneous time spacing $\mathcal{T}^*$ and the deterministic trading strategy $\boldsymbol{\xi}^{(1)}$ that is defined as follows. The initial market order $\xi_0^{(1)}$ is the unique solution of the equation*

$$F^{-1}\big(X_0 - N\xi_0^{(1)}(1-a^*)\big) = \frac{F^{-1}(\xi_0^{(1)}) - a^* F^{-1}(a^*\xi_0^{(1)})}{1-a^*}, \tag{8}$$

the intermediate orders are given by

$$\xi_1^{(1)} = \cdots = \xi_{N-1}^{(1)} = \xi_0^{(1)}(1-a^*), \tag{9}$$

and the final order is determined by

$$\xi_N^{(1)} = X_0 - \xi_0^{(1)} - (N-1)\xi_0^{(1)}(1-a^*).$$

Moreover, $\boldsymbol{\xi}^{(1)}$ consists only of nontrivial buy orders, i.e., $\xi_n^{(1)} > 0$ for all n.

In the limit $X_0 \downarrow 0$, the preceding result yields the nonexistence of price manipulation strategies in the sense of Definition 2.2:

Corollary 2.8. *Under Assumption 2.5, any nontrivial round trip has a strictly positive average cost. In particular, there are no price manipulation strategies.*

Remark 2.9. The preceding corollary shows that, in our Model 1, exponential resilience of price impact is well compatible with nonlinear impact governed by a shape function that satisfies Assumption 2.5. This fact is in stark contrast to Gatheral's observation that, in a related but different continuous-time model, exponential decay of price impact gives rise to price manipulation as soon as price impact is nonlinear. We thus deduce that, at least from a theoretical perspective, exponential resilience of a limit order book is a viable possibility for describing the decay of market impact.

Note that the positivity of the optimal trading strategies in Proposition 2.6 and Theorem 2.7 can be regarded as an additional regularity and viability result for our model. Indeed, it is possible that a market impact model does not admit price manipulation strategies in the strict sense of Definition 2.2, while an optimal strategy for a buy program can contain nontrivial sell orders or the other way round. Such trading strategies can be regarded as *price manipulation strategies in a weak sense.* We refer to our forthcoming paper [3] with A. Slynko for a systematic study and corresponding examples.

Remark 2.10. Let us briefly discuss the asymptotic behavior of the optimal strategy when the number N of trades tends to infinity. Since for any N we have $\xi_0^{(1)} \in (0, X_0)$, we can extract a subsequence that converges to some $\xi_0^{(1),\infty} \geq 0$. One therefore checks that the right-hand side of (8) tends to

$$h_1^\infty(\xi_0^{(1),\infty}) := F^{-1}(\xi_0^{(1),\infty}) + \frac{\xi_0^{(1),\infty}}{f(F^{-1}(\xi_0^{(1),\infty}))}.$$

Since $N(1-a^*) \to \int_0^T \rho_s\,ds$, the left-hand side of (8) converges as well, and so $\xi_0^{(1),\infty}$ must be a solution of the equation

$$F^{-1}\Big(X_0 - y\int_0^T \rho_s\,ds\Big) = h_1^\infty(y).$$

Note that, under our assumptions, h_1^∞ is strictly increasing. Hence, the preceding equation has a unique solution, which consequently must be the limit of $\xi_0^{(1)}$ as $N \uparrow \infty$. It follows moreover that $N\xi_1^{(1)} \to \xi_0^{(1),\infty} \int_0^T \rho_s\,ds$ and that

$$\xi_N^{(1)} \longrightarrow X_0 - \xi_0^{(1),\infty} - \xi_0^{(1)} \int_0^T \rho_s\,ds =: \xi_T^{(1),\infty}.$$

Thus, the optimal strategy, described in "resilience time" $r(t) := \int_0^t \rho_s\,ds$, consists of an initial block trade of size $\xi_0^{(1),\infty}$, continuous buying at constant rate $\xi_0^{(1),\infty}$ during $(0,T)$, and a final block trade of size $\xi_T^{(1),\infty}$. Transforming back to standard time leaves the initial and final block trades unaffected, and continuous buying in $(0,T)$ now occurs at the time-dependent rate $\rho_t\xi_0^{(1),\infty}$.

2.5 Main results for Model 2

In this section we state our main results for Model 2. This case is analytically more complicated than Model 1, because the quantity that decays exponentially is no longer a linear function of the order size. We therefore need a stronger assumption:

Assumption 2.11. In Model 2, we assume in addition to (3) that the shape function f is twice differentiable on $\mathbb{R}\backslash\{0\}$, nondecreasing on $\mathbb{R}_-$ and nonincreasing on $\mathbb{R}_+$. We moreover assume:

$$x \mapsto xf'(x)/f(x) \text{ is nondecreasing on } \mathbb{R}_-, \text{ nonincreasing on } \mathbb{R}_+, \text{ and } (-1,0]\text{-valued,} \tag{10}$$

$$1 + x\frac{f'(x)}{f(x)} + 2x^2\left(\frac{f'(x)}{f(x)}\right)^2 - x^2\frac{f''(x)}{f(x)} \geq 0 \qquad \text{for all } x \geq 0. \tag{11}$$

We will see in Example 2.15 below that this assumption is satisfied for the power law shape functions from Example 2.1.

We start by looking at optimal trading strategies when an admissible sequence of trading times $\mathcal{T} = (\tau_0, \dots, \tau_N)$ is fixed. As in Section 2.4, we say that two $\mathcal{T}$-admissible trading strategies $\boldsymbol{\xi}$ and $\overline{\boldsymbol{\xi}}$ are equivalent if $\xi_i + \xi_{i+1} = \overline{\xi}_i + \overline{\xi}_{i+1}$ $\mathbb{P}$-a.s. on $\{\tau_i = \tau_{i+1}\}$.

Proposition 2.12. *Suppose that an admissible sequence of trading times $\mathcal{T}$ is given and that Assumption 2.11 holds. Then there exists a $\mathcal{T}$-admissible trading strategy $\boldsymbol{\xi}^{(2),\mathcal{T}}$, unique up to equivalence, that minimizes the cost $\mathcal{C}(\cdot,\mathcal{T})$. Moreover, it consists only of nontrivial buy orders, i.e., $\xi_i^{(2),\mathcal{T}} > 0$ $\mathbb{P}$-a.s. for all i up to equivalence.*

As in Proposition 2.6, computing the optimal trading strategy $\boldsymbol{\xi}^{(2),\mathcal{T}}$ for an arbitrary sequence $\mathcal{T}$ can be quite complicated. But, again, the structure becomes much easier when optimizing also over the sequence of trading times, $\mathcal{T}$. To state the corresponding result, let us recall from (6) the definition of a^* and let us introduce the function

$$h_{2,a^*}(x) := x\frac{f(x/a^*)/a^* - a^*f(x)}{f(x/a^*) - a^*f(x)}.$$

We will see in Lemma 3.7 (a) below that $h_{2,a^*}(x)$ is indeed well-defined for all $x \in \mathbb{R}$ as soon as Assumption 2.11 is satisfied.

Theorem 2.13. *Suppose that that the shape function f satisfies Assumption 2.11. Then there is a unique optimal strategy $(\boldsymbol{\xi}^{(2)},\mathcal{T}^*)$, consisting of homogeneous time spacing $\mathcal{T}^*$ and the deterministic trading strategy $\boldsymbol{\xi}^{(2)}$ that is defined as follows. The initial market order $\xi_0^{(2)}$ is the unique solution of the equation*

$$F^{-1}\Big(X_0 - N\big[\xi_0^{(2)} - F\big(a^*F^{-1}(\xi_0^{(2)})\big)\big]\Big) = h_{2,a^*}\big(F^{-1}(\xi_0^{(2)})\big), \tag{12}$$

the intermediate orders are given by

$$\xi_1^{(2)} = \cdots = \xi_{N-1}^{(2)} = \xi_0^{(2)} - F\big(a^*F^{-1}(\xi_0^{(2)})\big), \tag{13}$$

and the final order is determined by

$$\xi_N^{(2)} = X_0 - N\xi_0^{(2)} + (N-1)F\big(a^*F^{-1}(\xi_0^{(2)})\big).$$

Moreover, the optimal strategy consists only of nontrivial buy orders, i.e., $\xi_n^{(2)} > 0$ for all n.

As for Model 1, the limit $X_0 \downarrow 0$ yields the following result on the nonexistence of price manipulation strategies in our model. We refer to Remark 2.9 for a discussion of this fact.

Corollary 2.14. *Under Assumption 2.11, any nontrivial round trip has a strictly positive average cost. In particular, there are no profitable price manipulation strategies in Model 2.*

We conclude this section by showing that the power law shape functions from Example 2.1 are admissible also for Model 2.

Example 2.15 (Power law)**.** Let us show that the power law shape functions from Example 2.1 satisfy our Assumption 2.11. For checking (10) and (11) we concentrate on the branch of f on the positive part of the real line. So let us suppose that

$$f(x) = \frac{q}{(1+\lambda x)^\alpha}, \qquad \text{for } x > 0,$$

with $\alpha \in [0,1]$, $q, \lambda > 0$. We have $xf'(x)/f(x) = -\frac{\alpha\lambda x}{1+\lambda x} \in (-1, 0]$ which is nonincreasing on $\mathbb{R}_+$. Moreover, for $x \geq 0$ we have

$$1 + x\frac{f'(x)}{f(x)} + 2x^2\left(\frac{f'(x)}{f(x)}\right)^2 - x^2\frac{f''(x)}{f(x)} = \frac{1 + (2-\alpha)\lambda x + (1 - 2\alpha + \alpha^2)(\lambda x)^2}{(1+\lambda x)^2} \geq 0.$$

Remark 2.16. As in Remark 2.10, we can study the asymptotic behavior of the optimal strategy as the number N of trades tends to infinity. First, one checks that h_{2,a^*} converges to

$$h_2^\infty(x) := x\Big(1 + \frac{f(x)}{f(x) + xf'(x)}\Big),$$

and that $N(y - F(a^*F^{-1}(y)))$ tends to $F^{-1}(y)f(F^{-1}(y))\int_0^T \rho_s\,ds$. Now, suppose that the equation

$$F^{-1}\Big(X_0 - F^{-1}(y)f(F^{-1}(y))\int_0^T \rho_s\,ds\Big) = h_2^\infty(F^{-1}(y))$$

has a unique solution in $(0, X_0)$, which we will call $\xi_0^{(2),\infty}$. We then see as in Remark 2.10 that $\xi_0^{(2),\infty}$ is the limit of $\xi_0^{(2)}$. Next, $N\xi_1^{(2)}$ converges to $F^{-1}(\xi_0^{(2),\infty})f(F^{-1}(\xi_0^{(2),\infty}))\int_0^T \rho_s\,ds$ and $\xi_N^{(2)}$ to

$$\xi_T^{(2),\infty} := X_0 - \xi_0^{(2),\infty} - F^{-1}(\xi_0^{(2),\infty})f(F^{-1}(\xi_0^{(2),\infty}))\int_0^T \rho_s\,ds.$$

This yields a description of the continuous-time limit in "resilience time" $r(t) := \int_0^t \rho_s\,ds$. Using a time change as in Remark 2.10, we obtain that the optimal strategy consists of an initial block order of $\xi_0^{(2),\infty}$ shares at time 0, continuous buying at rate $\rho_t F^{-1}(\xi_0^{(2),\infty})f(F^{-1}(\xi_0^{(2),\infty}))$ during $(0, T)$, and a final block order of $\xi_T^{(2),\infty}$ shares at time T.

3 Proofs

3.1 Reduction to a deterministic problem

First, it will be convenient to work with the quantities

$$\alpha_k := \int_{\tau_{k-1}}^{\tau_k} \rho_s ds, \qquad k = 1, \dots, N, \tag{14}$$

instead of the τ_k themselves. The condition $0 = \tau_0 \le \tau_1 \le \cdots \le \tau_N = T$ is clearly equivalent to $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_N)$ belonging to

$$\mathcal{A} := \Big\{ \boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}_+^N \,\Big|\, \sum_{k=1}^N \alpha_k = \int_0^T \rho_s \, ds \Big\}.$$

As explained in [2, Section A], the martingale assumption on $(A_t^0, t \ge 0)$, allows to reduce the optimization problem to a deterministic one. This remains true in our extended setting in which trading times are allowed to be stopping times. Following [2], we introduce a new pair of discrete-time processes D and E that react on both sell and buy orders according to the following dynamics.

- We have $E_0 = D_0 = 0$ and
$$E_n = F(D_n) \qquad \text{and} \qquad D_n = F^{-1}(E_n).$$
- For $n = 0, \ldots, N$, regardless of the sign of ξ_n,
$$E_{n+} = E_n + \xi_n \qquad \text{and} \qquad D_{n+} = F^{-1}\left(\xi_n + F\left(D_n\right)\right).$$
- For $k = 0, \ldots, N-1$,
$$\begin{aligned} E_{k+1} &= e^{-\alpha_{k+1}} E_{k+} = e^{-\alpha_{k+1}} (E_k + \xi_k) \qquad \text{in Model 1,} \\ D_{k+1} &= e^{-\alpha_{k+1}} D_{k+} = e^{-\alpha_{k+1}} F^{-1}\left(\xi_k + F\left(D_k\right)\right) \qquad \text{in Model 2.} \end{aligned} \tag{15}$$

Next, observe that, in each Model $i = 1, 2$, the simplified extra spread process D evolves deterministically once the values of $\boldsymbol{\alpha}(\omega) = (\alpha_1(\omega), \ldots, \alpha_N(\omega))$ and $\boldsymbol{\xi}(\omega) = (\xi_0(\omega), \ldots, \xi_N(\omega))$ are given. Hence, there exists a deterministic function $C^{(i)} : \mathbb{R}^{N+1} \times \mathcal{A} \to \mathbb{R}$ such that

$$\sum_{n=0}^N \int_{D_n}^{D_{n+}} x f(x) \, dx = C^{(i)}(\boldsymbol{\xi}, \boldsymbol{\alpha}). \tag{16}$$

The relation between the functions $C^{(i)}$ and our original cost function $\mathcal{C}$ now is as follows: if $(\boldsymbol{\xi}, \mathcal{T})$ is an admissible strategy and $\boldsymbol{\alpha}$ is as in (14), then

$$\mathcal{C}(\boldsymbol{\xi}, \mathcal{T}) \ge A_0 X_0 + \mathbb{E}\big[\, C^{(i)}(\boldsymbol{\xi}, \boldsymbol{\alpha}) \,\big] \qquad \text{with equality if } \xi_i \ge 0 \ \mathbb{P}\text{-a.s. for all } i. \tag{17}$$

This follows from the same arguments as in [2, Section A].

We will show in the respective Sections 3.3 and 3.4 that, under our assumptions, the functions $C^{(i)}$, $i = 1, 2$, have unique minima within the set $\Xi \times \mathcal{A}$, where

$$\Xi := \Big\{ \boldsymbol{x} = (x_0, \ldots, x_N) \in \mathbb{R}^{N+1} \,\big|\, \sum_{n=0}^N x_n = X_0 \Big\}.$$

It then follows from (17) that the corresponding minimizers also minimize the original cost functional (5), provided that the resulting trading strategies consist only of buy orders. When working with deterministic trading strategies in Ξ rather than with random variables, we will mainly use Roman letters like $\boldsymbol{x}$ instead of Greek letters such as $\boldsymbol{\xi}$. We conclude this section with the following easy lemma.

Lemma 3.1. *For $X_0 > 0$, there is no $\boldsymbol{x} \in \Xi$ such that $E_{n+} = E_n + x_n \le 0$ (or, equivalently, $D_{n+} \le 0$) for all $n = 0, \dots, N$.*

Proof. Since the effect of resilience is to drive the extra spread back to zero, we have $E_{n+} \ge x_0 + \cdots + x_n$ up to and including the first n at which $x_0 + \cdots + x_n > 0$. □

3.2 Proofs for a block-shaped limit order book

In this section, we give quick and direct proofs for our results in case of a block-shaped limit order book with $f(x) = q$. In this setting, Models 1 and 2 coincide. As explained in [1], the cost function in (16) is an increasing affine function of

$$C(\boldsymbol{x}, \boldsymbol{\alpha}) = \frac{1}{2}\langle \boldsymbol{x}, M(\boldsymbol{\alpha})\boldsymbol{x}\rangle, \qquad \boldsymbol{x} \in \Xi,\ \alpha \in \mathcal{A}, \tag{18}$$

where $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product and $M(\boldsymbol{\alpha})$ is the positive definite symmetric matrix with entries

$$M(\boldsymbol{\alpha})_{n,m} = \exp\Big(-\int_{\tau_{n\wedge m}}^{\tau_{n\vee m}} \rho_u\,du\Big) = \exp\Big(-\Big|\sum_{i=1}^{n}\alpha_i - \sum_{j=1}^{m}\alpha_j\Big|\Big), \qquad 0 \le n, m \le N.$$

Proof of Theorem 2.3. For $\boldsymbol{\alpha}$ belonging to

$$\mathcal{A}^* := \big\{\boldsymbol{\alpha} \in \mathcal{A} \,|\, \alpha_i > 0,\ i = 1, \dots, N\big\},$$

the inverse $M(\boldsymbol{\alpha})^{-1}$ of the matrix $M(\boldsymbol{\alpha})$ can be computed explicitly, and the unique optimal trading strategy for fixed $\boldsymbol{\alpha}$ is

$$\boldsymbol{x}^*(\boldsymbol{\alpha}) = \frac{X_0}{\langle \mathbf{1}, M(\boldsymbol{\alpha})^{-1}\mathbf{1}\rangle} M(\boldsymbol{\alpha})^{-1}\mathbf{1},$$

and this strategy has only strictly positive components; see [1, Theorem 3.1]. From [1, Eq. (17)], it follows that

$$\begin{aligned}
\min_{\boldsymbol{x}\in\Xi} C(\boldsymbol{x}, \boldsymbol{\alpha}) &= C(\boldsymbol{x}^*(\boldsymbol{\alpha}), \boldsymbol{\alpha}) = \frac{X_0^2}{2\langle \mathbf{1}, M(\boldsymbol{\alpha})^{-1}\mathbf{1}\rangle} \\
&= \frac{X_0^2}{2}\left(\frac{2}{1+e^{-\alpha_1}} + \sum_{n=2}^{N}\frac{1-e^{-\alpha_n}}{1+e^{-\alpha_n}}\right)^{-1} \\
&= \frac{X_0^2}{2}\left(\sum_{n=1}^{N}\frac{2}{1+e^{-\alpha_n}} - (N-1)\right)^{-1}.
\end{aligned} \tag{19}$$

Minimizing $\min_{\boldsymbol{x}\in\Xi} C(\boldsymbol{x}, \boldsymbol{\alpha})$ over $\boldsymbol{\alpha} \in \mathcal{A}^*$ is thus equivalent to maximizing $\sum_{n=1}^{N}\frac{2}{1+e^{-\alpha_n}}$. The function $x \mapsto \frac{2}{1+e^{-x}}$ is strictly concave in $x > 0$. Hence,

$$\sum_{n=1}^{N}\frac{2}{1+e^{-\alpha_n}} \le \frac{2N}{1+e^{-\frac{1}{N}\sum_{n=1}^{N}\alpha_n}} = \frac{2N}{1+e^{-\frac{1}{N}\int_0^T \rho_u du}},$$

with equality if and only if $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$, where $\boldsymbol{\alpha}^*$ corresponds to to homogeneous time spacing $\mathcal{T}^*$, i.e.,

$$\alpha_i^* = \frac{1}{N}\int_0^T \rho_s\,ds, \qquad i = 1, \dots, N. \tag{20}$$

Next, $C(\boldsymbol{x}, \boldsymbol{\alpha})$ is clearly jointly continuous in $\boldsymbol{x} \in \Xi$ and $\boldsymbol{\alpha} \in \mathcal{A}$, so $\inf_{\boldsymbol{x}\in\Xi} C(\boldsymbol{x}, \boldsymbol{\alpha})$ is upper semicontinuous in $\boldsymbol{\alpha}$. One thus sees that the minimum cannot be attained at the boundary of $\mathcal{A}$. Finally, the formula (7) for the optimal trading strategy with homogeneous time spacing can be found in [1, Remark 3.2] or in [2, Corollary 6.1]. □

Proof of Corollary 2.4. Suppose that $(\boldsymbol{x}, \boldsymbol{\alpha})$ is a round trip such that $C(\boldsymbol{x}, \boldsymbol{\alpha}) \le 0$. We can assume w.l.o.g. that $\boldsymbol{\alpha} \in \mathcal{A}^*$, for otherwise we can simply merge those trades that occur at the same time into a single trade. Then

$$C(\boldsymbol{x}, \boldsymbol{\alpha}) = \lim_{\varepsilon \downarrow 0} C(\boldsymbol{x} + \varepsilon\mathbf{1}, \boldsymbol{\alpha}).$$

But $C(\boldsymbol{x}+\varepsilon\mathbf{1}, \boldsymbol{\alpha}) > 0$ for each $\varepsilon > 0$, due to (19). Hence we must have $C(\boldsymbol{x}, \boldsymbol{\alpha}) = 0$. According to [1, Theorem 3.3], the matrix $M(\boldsymbol{\alpha})$ is positive definite for $\alpha \in \mathcal{A}^*$, and so there can be at most one minimizer of $C(\cdot, \boldsymbol{\alpha})$ in the class of round trips. Since we clearly have $C(\mathbf{0}, \boldsymbol{\alpha}) = 0$, we must conclude that $\boldsymbol{x} = \mathbf{0}$. The result now follows from (17). □

3.3 Proofs for Model 1

We need a few lemmas before we can prove our main results for Model 1.

Lemma 3.2. *Under Assumption 2.5, the following conclusions hold.*

(a) *For each $a \in (0, 1)$, the function*

$$\begin{aligned} h_{1,a} : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\longmapsto F^{-1}(y) - aF^{-1}(ay) \end{aligned} \tag{21}$$

is strictly increasing.

(b) *For all $a, b \in (0, 1)$ and $\nu > 0$, we have the inequalities*

$$h_{1,a}^{-1}\big(\nu(1-a)\big) > b \cdot h_{1,b}^{-1}\big(\nu(1-b)\big) \tag{22}$$

and

$$b \cdot h_{1,b}^{-1}\big(\nu(1-b)\big) < F(\nu). \tag{23}$$

(c) *The function*

$$\begin{aligned} H_1 : (0,\infty) \times (0,1) &\longrightarrow \mathbb{R}^2 \\ (y, a) &\longmapsto \Big(\frac{F^{-1}(y) - aF^{-1}(ay)}{1-a}, ay\frac{F^{-1}(y) - F^{-1}(ay)}{1-a}\Big) \end{aligned}$$

is one-to-one.

Proof: Part (a) of the assertion follows from [2, Remark 4.3].

For the proof of part (b), let $y := h_{1,a}^{-1}\big(\nu(1-a)\big)$. Then $y > 0$ since $h_{1,a}(0) = 0$. Note also that F^{-1} is convex on $\mathbb{R}_+$. Let $\widehat{f}$ be its derivative. Then,

$$\begin{aligned}\nu &= \frac{F^{-1}(y) - aF^{-1}(ay)}{1-a} = F^{-1}(ay) + \frac{F^{-1}(y) - F^{-1}(ay)}{1-a}\\ &= F^{-1}(ay) + \frac{1}{1-a}\int_{ay}^{y} \widehat{f}(x)\,dx < F^{-1}(y) + y\widehat{f}(y) =: g(y).\end{aligned}$$

Clearly, g is a strictly increasing function on $\mathbb{R}_+$, and so we have $y > g^{-1}(\nu)$.

Next, let $z := b \cdot h_{1,b}^{-1}\big(\nu(1-b)\big)$. Then,

$$\begin{aligned}\nu &= \frac{F^{-1}(z/b) - bF^{-1}(z)}{1-b} = F^{-1}(z) + \frac{F^{-1}(z/b) - F^{-1}(z)}{1-b}\\ &= F^{-1}(z) + \frac{1}{1-b}\int_{z}^{z/b} \widehat{f}(x)\,dx \geq F^{-1}(z) + z\widehat{f}(z) = g(z),\end{aligned}$$

since $\widehat{f}(z) = 1/f(F^{-1}(z))$ is nondecreasing for $z > 0$. Thus, $z \leq g^{-1}(\nu) < h_{1,a}^{-1}\big(\nu(1-a)\big)$, and (22) follows. For (23) it now suffices to note that $g(z) > F^{-1}(z)$.

To prove part (c), let $a_1, a_2 \in (0,1)$ and $y_1, y_2 > 0$ and assume that $H_1(a_1, y_1) = H_1(a_2, y_2)$. Since

$$\frac{F^{-1}(y) - aF^{-1}(ay)}{1-a} = F^{-1}(y) + a\frac{F^{-1}(y) - F^{-1}(ay)}{1-a},$$

we get

$$\begin{aligned}F^{-1}(y_1) + a_1\frac{F^{-1}(y_1) - F^{-1}(a_1y_1)}{1-a_1} &= F^{-1}(y_2) + a_2\frac{F^{-1}(y_2) - F^{-1}(a_2y_2)}{1-a_2},\\ a_1y_1\frac{F^{-1}(y_1) - F^{-1}(a_1y_1)}{1-a_1} &= a_2y_2\frac{F^{-1}(y_2) - F^{-1}(a_2y_2)}{1-a_2}.\end{aligned}$$

Assume that $y_1 \neq y_2$, say, $y_1 > y_2 > 0$. Multiplying the first identity by y_1 and subtracting the second identity yields

$$y_1\frac{F^{-1}(y_1) - F^{-1}(y_2)}{y_1 - y_2} = \frac{a_2}{1-a_2}\big[F^{-1}(y_2) - F^{-1}(a_2y_2)\big].$$

Since $(F^{-1})'(y) = \widehat{f}(y)$ is nondecreasing for $y > 0$, we obtain that

$$y_1\frac{F^{-1}(y_1) - F^{-1}(y_2)}{y_1 - y_2} \geq y_1\widehat{f}(y_2) \quad \text{and} \quad \frac{a_2}{1-a_2}\big[F^{-1}(y_2) - F^{-1}(a_2y_2)\big] \leq a_2y_2\widehat{f}(y_2),$$

which contradicts the previous equation since $y_1 > y_2 \geq a_2y_2$. Therefore we must have $y_1 = y_2$.

It is therefore sufficient to show that

$$\widetilde{h}(a) := \frac{a}{1-a}\big[F^{-1}(y) - F^{-1}(ay)\big], \qquad a \in]0,1[,$$

is one-to-one for any $y > 0$. Its derivative is equal to

$$\widetilde{h}'(a) = \frac{1}{(1-a)^2}\big[F^{-1}(y) - F^{-1}(ay)\big] - \frac{ay}{1-a}\widehat{f}(ay). \tag{24}$$

Using again that $\widehat{f}(y)$ is nondecreasing for $y > 0$, we get

$$F^{-1}(y) - F^{-1}(ay) > (1-a)y\widehat{f}(ay)$$

and in turn $\widetilde{h}'(a) > 0$. □

Let us introduce the functions

$$\tilde{F}(x) := \int_0^x zf(z)\,dz \qquad \text{and} \qquad G = \tilde{F} \circ F^{-1}.$$

With this notation, the simplified cost function (16) in Model 1 can be represented as

$$C^{(1)}(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_{n=0}^{N} \Big[G(E_n + x_n) - G(E_n)\Big], \qquad \boldsymbol{x} \in \Xi,\ \boldsymbol{\alpha} \in \mathcal{A}, \tag{25}$$

where

$$E_0 = 0 \qquad \text{and} \qquad E_n = \sum_{i=0}^{n-1} x_i e^{-\sum_{k=i+1}^{n} \alpha_k}, \qquad 1 \le n \le N.$$

Lemma 3.3. *For $i = 0, \ldots, N-1$, we have the following recursive formula,*

$$\frac{\partial C^{(1)}}{\partial x_i} = F^{-1}(E_i + x_i) - e^{-\alpha_{i+1}} F^{-1}(E_{i+1}) + e^{-\alpha_{i+1}} \frac{\partial C^{(1)}}{\partial x_{i+1}}. \tag{26}$$

Moreover, for $i = 1, \ldots, N$,

$$\frac{\partial C^{(1)}}{\partial \alpha_i} = E_i \sum_{n=i}^{N} \big[F^{-1}(E_n + x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=i+1}^{n} \alpha_k}. \tag{27}$$

Proof: To prove (26), we first need to calculate $\partial E_n / \partial x_i$. We obtain:

$$\frac{\partial E_n}{\partial x_i} = 0 \qquad \text{if } i \ge n, \text{ and} \qquad \frac{\partial E_n}{\partial x_i} = e^{-\sum_{k=i+1}^{n} \alpha_k} \qquad \text{if } i < n.$$

Using the fact that $G' = F^{-1}$, we therefore get

$$\begin{aligned}
\frac{\partial C^{(1)}}{\partial x_i} &= F^{-1}(E_i + x_i) + \sum_{n=i+1}^{N} \big[F^{-1}(E_n + x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=i+1}^{n} \alpha_k} \\
&= F^{-1}(E_i + x_i) - e^{-\alpha_{i+1}} F^{-1}(E_{i+1}) \\
&\quad + e^{-\alpha_{i+1}} \left(F^{-1}(E_{i+1} + x_{i+1}) + \sum_{n=i+2}^{N} \big[F^{-1}(E_n + x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=i+2}^{n} \alpha_k} \right),
\end{aligned}$$

which yields (26).

For the proof of (27), we have first to compute $\partial E_n/\partial\alpha_i$. We obtain:

$$\frac{\partial E_n}{\partial \alpha_i}=0 \qquad \text{if } i>n, \text{ and} \qquad \frac{\partial E_n}{\partial \alpha_i}=-\sum_{m=0}^{i-1} x_m e^{-\sum_{k=m+1}^n \alpha_k} \qquad \text{for } i\le n.$$

From here, we get

$$\begin{aligned}\frac{\partial C^{(1)}}{\partial \alpha_i} &= -\sum_{n=i}^{N}\big[F^{-1}(E_n+x_n)-F^{-1}(E_n)\big]\sum_{m=0}^{i-1} x_m e^{-\sum_{k=m+1}^n \alpha_k}\\ &= E_i\sum_{n=i}^{N}\big[F^{-1}(E_n+x_n)-F^{-1}(E_n)\big]e^{-\sum_{k=i+1}^n \alpha_k},\end{aligned}$$

which is (27). □

Remark 3.4. A consequence of this lemma is that homogeneous time spacing $\boldsymbol{\alpha}^*$ and the optimal strategy $\boldsymbol{\xi}^{(1)}$ given in [2] yield a critical point for the minimization in $(\boldsymbol{x},\boldsymbol{\alpha})$. Indeed, we have then $E_i=a^*\xi_0^{(1)}$ for any i, and therefore $\frac{\partial C^{(2)}}{\partial\alpha_i}$ does not depend on i.

Lemma 3.5. *For each $\boldsymbol{\alpha}\in\mathcal{A}$ the function $C^{(1)}(\cdot,\boldsymbol{\alpha})$ has a minimizer $\boldsymbol{x}^*(\boldsymbol{\alpha})\in\Xi$, which is unique up to equivalence.*

Proof: First note that we may assume without loss of generality that $\boldsymbol{\alpha}\in\mathcal{A}^*=\{\boldsymbol{\alpha}\in\mathcal{A}\,|\,\alpha_i>0,\ i=1,\dots,N\}$. Indeed, if $\alpha_i=0$ we can merge the trades x_{i-1} and x_i into a single one and reduce N to $N-1$.

We next extend the arguments in [2, Lemma B.1] to prove the existence of a unique minimizer of $C^{(1)}(\cdot,\boldsymbol{\alpha})$ in Ξ.

Using the convention $\sum_{k=n+1}^n \alpha_k:=0$, we obtain by rearranging the sum in (25) that

$$\begin{aligned}C^{(1)}(\boldsymbol{x},\boldsymbol{\alpha}) &= G\Big(\sum_{i=0}^{N} x_i e^{-\sum_{k=i+1}^N \alpha_k}\Big)-G(0)\\ &\quad+\sum_{n=0}^{N-1}\Big[G\Big(\sum_{i=0}^{n} x_i e^{-\sum_{k=i+1}^n \alpha_k}\Big)-G\Big(e^{-\alpha_{n+1}}\sum_{i=0}^{n} x_i e^{-\sum_{k=i+1}^n \alpha_k}\Big)\Big].\end{aligned}$$

Let us define the linear map $T:\mathbb{R}^{N+1}\to\mathbb{R}^{N+1}$ via

$$(T\boldsymbol{x})_n=\sum_{i=0}^{n} x_i e^{-\sum_{k=i+1}^n \alpha_k}, \qquad n=0,\dots,N.$$

We can thus write

$$C^{(1)}(\boldsymbol{x},\boldsymbol{\alpha})=G\big((T\boldsymbol{x})_N\big)-G(0)+\sum_{n=0}^{N-1}\Big[G\big((T\boldsymbol{x})_n\big)-G\big(e^{-\alpha_{n+1}}(T\boldsymbol{x})_n\big)\Big]. \tag{28}$$

Note first that G is strictly convex since $G' = F^{-1}$ is strictly increasing. Second, for $a \in (0,1)$, the function $x \to G(x) - G(ax)$ is also strictly convex, because its derivative is equal to the strictly increasing function $h_{1,a}$ in Lemma 3.2 (a). And third, T is one-to-one. Hence, $C^{(1)}(\cdot, \boldsymbol{\alpha})$ is strictly convex in is first argument, and there can be at most one minimizer.

To show the existence of a minimizer, note that $G' = F^{-1}$ is increasing with $F^{-1}(0) = 0$, and hence $G(y) - G(ay) \geq (1-a)|y| \cdot |F^{-1}(ay)|$. Therefore, (28) yields

$$\begin{aligned} C^{(1)}(\boldsymbol{x}, \boldsymbol{\alpha}) \;\; &\geq \;\; G\big((T\boldsymbol{x})_N\big) - G(0) \\ &\quad + \sum_{n=0}^{N-1} (1 - e^{-\alpha_{n+1}}) \cdot \big|F^{-1}\big(e^{-\alpha_{n+1}} (T\boldsymbol{x})_n\big)\big| \cdot \Big|(T\boldsymbol{x})_n\Big|. \end{aligned}$$

Hence,

$$C^{(1)}(\boldsymbol{x}, \boldsymbol{\alpha}) \geq \Lambda\big(|T\boldsymbol{x}|_\infty\big) - G(0),$$

where $|\cdot|_\infty$ is the ℓ^∞-norm on $\mathbb{R}^{N+1}$ and Λ is the function

$$\Lambda(y) := G(y) \wedge G(-y) \wedge \min_{n=0,\dots,N-1} \Big\{ |y| \cdot (1 - a_{n+1}) \Big(\big|F^{-1}(a_{n+1} \cdot y)\big| \wedge \big|F^{-1}(-a_{n+1} \cdot y)\big| \Big) \Big\},$$

where $a_{n+1} := e^{-\alpha_{n+1}}$. Since F is unbounded, both $G(y)$ and $|F^{-1}(y)|$ tend to $+\infty$ for $|y| \to \infty$, and the fact that T is one-to-one implies that $\Lambda\big(|T\boldsymbol{x}|_\infty\big) \to +\infty$ for $|\boldsymbol{x}| \to \infty$. Note also that by assumption $\alpha_n > 0$ for each n. Hence, $C^{(1)}(\cdot, \boldsymbol{\alpha})$ must attain its minimum on Ξ. □

We are now in a position to prove the main results for Model 1.

Proof of Proposition 2.6: The result will follow via (17) if we can show that the minimizer in Lemma 3.5 consists only of strictly positive components. Here, we may assume without loss of generality that the admissible sequence of trading times is strictly increasing, or equivalently that $\boldsymbol{\alpha} \in \mathcal{A}^*$, for otherwise we can simply merge two trades occurring at the same time into a single trade.

If $\boldsymbol{x}$ is the minimizer of $C^{(1)}(\cdot, \boldsymbol{\alpha})$ on Ξ, then there must be a Lagrange multiplier ν such that $\boldsymbol{x}$ is a critical point of $\boldsymbol{y} \mapsto C^{(1)}(\boldsymbol{y}, \boldsymbol{\alpha}) - \nu \sum_{i=0}^N y_i$. Hence, (26) yields that

$$\nu(1 - a_{i+1}) = F^{-1}(E_i + x_i) - a_{i+1} F^{-1}(E_{i+1}) = h_{1,a_{i+1}}(E_i + x_i), \qquad i = 0, \dots, N-1, \quad (29)$$

where $a_{i+1} = e^{-\alpha_{i+1}}$ and $h_{1,a}$ is as in (21). For the final trade, we have

$$\nu = F^{-1}(E_N + x_N). \quad (30)$$

Since $h_{1,a}(0) = 0 = F^{-1}(0)$ and both $h_{1,a}$ and F^{-1} are strictly increasing, we conclude that $E_0 + x_0, \dots, E_N + x_N$ have all the same sign as ν. Thus, $\nu > 0$ by Lemma 3.1. Next, (29) implies that $E_i + x_i = h^{-1}_{1,a_{i+1}}(\nu(1 - a_{i+1}))$ and hence $E_{i+1} = a_{i+1} h^{-1}_{1,a_{i+1}}(\nu(1 - a_{i+1}))$. Using (29) once again yields

$$x_0 = h^{-1}_{1,a_1}(\nu(1 - a_1)) \quad \text{and} \quad x_i = h^{-1}_{1,a_{i+1}}(\nu(1 - a_{i+1})) - a_i h^{-1}_{1,a_i}(\nu(1 - a_i)), \qquad i = 1, \dots, N-1.$$

The inequality (22) thus gives $x_i > 0$ for $i = 0, \dots, N-1$. As for the final trade, (30) gives $x_N = F(\nu) - a_N h^{-1}_{1,a_N}(\nu(1 - a_N))$, which is strictly positive by (23). □

Proof of Theorem 2.7: We will show that $(\boldsymbol{\xi}^{(1)}, \boldsymbol{\alpha}^*)$, defined via (8), (9), and (20), is the unique minimizer of $C^{(1)}$ on $\Xi \times \mathcal{A}$. The result will then follow from (17). The first step is to show the existence of a minimizer. To this end, note that Proposition 2.6 allows us to restrict the minimization of $C^{(1)}$ to $\Xi_+ \times \mathcal{A}$, where $\Xi_+ = \{\boldsymbol{x} \in \Xi \,|\, x_i \geq 0,\ i = 0,\dots,N\}$. The set $\Xi_+ \times \mathcal{A}$ is in fact the product of two compact simplices, and so the continuity of $C^{(1)}$ yields the existence of a global minimizer, which lies in $\Xi_+ \times \mathcal{A}$.

We next argue that any minimizer must belong to the relative interior of $\Xi_+ \times \mathcal{A}$. To this end, suppose that $\boldsymbol{x} \in \Xi_+$ and $\boldsymbol{\alpha} \in \mathcal{A}$ are given and such that $\alpha_i = 0$ for some i. We then define $\overline{\boldsymbol{\alpha}} := (\alpha_0, \dots, \alpha_{i-1}, \alpha_{i+1}/2, \alpha_{i+1}/2, \alpha_{i+2}, \dots, \alpha_N)$ and $\overline{\boldsymbol{x}} := (x_0, \dots, x_{i-2}, x_{i-1} + x_i, 0, x_{i+1}, \dots, x_N)$ and observe that $C^{(1)}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\alpha}}) = C^{(1)}(\boldsymbol{x}, \boldsymbol{\alpha})$. But Proposition 2.6 implies that $\overline{\boldsymbol{x}}$ cannot be optimal for $\overline{\boldsymbol{\alpha}}$ since $\overline{x}_i = 0$. In particular, $(\boldsymbol{x}, \boldsymbol{\alpha})$ cannot be optimal. Thus, the $\boldsymbol{\alpha}$-component of any minimizer must lie in the relative interior of $\mathcal{A}$. Finally, for $\boldsymbol{\alpha}$ in the relative interior of $\mathcal{A}$, Proposition 2.6 states that $\overline{\boldsymbol{x}}^*(\boldsymbol{\alpha})$ belongs to the relative interior of Ξ_+.

Now suppose that $(\boldsymbol{x}, \boldsymbol{\alpha})$ is a minimizer of $C^{(1)}$. Due to the preceding step, there must be Lagrange multipliers ν and λ such that $(\boldsymbol{x}, \boldsymbol{\alpha})$ is a critical point of $(\boldsymbol{y}, \boldsymbol{\beta}) \mapsto C^{(1)}(\boldsymbol{y}, \boldsymbol{\beta}) - \nu \sum_{i=0}^N y_i - \lambda \sum_{j=1}^N \beta_j$. The identity (26) thus again yields

$$\nu(1 - e^{-\alpha_{i+1}}) = F^{-1}(E_i + x_i) - e^{-\alpha_{i+1}} F^{-1}(E_{i+1}), \qquad i = 0, \dots, N-1, \tag{31}$$

and

$$\nu = F^{-1}(E_N + x_N). \tag{32}$$

Using the same argument as in the proof of Proposition 2.6, we have $\nu > 0$. Note that this can also be obtained by writing

$$X_0 = \sum_{i=0}^N x_i = F(\nu) + \sum_{i=1}^N (1 - a_i) h_{1,a_i}^{-1}(\nu(1 - a_i)).$$

Indeed, the right-hand side is strictly increasing in ν (F and the functions h_{1,a_i}^{-1} are strictly increasing) and vanishes for $\nu = 0$, so $\nu > 0$.

Next, (27) gives

$$\lambda = E_j \sum_{n=j}^N \big[F^{-1}(E_n + x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=j+1}^n \alpha_k}, \qquad j = 1, \dots, N. \tag{33}$$

We now rewrite the sum in (33) as follows:

$$\begin{aligned}
&\sum_{n=j}^N \big[F^{-1}(E_n + x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=j+1}^n \alpha_k} \\
&\quad = \; -F^{-1}(E_j) \\
&\qquad + \big[F^{-1}(E_j + x_j) - F^{-1}(E_{j+1}) e^{-\alpha_{j+1}}\big] + \dots \\
&\qquad + \big[F^{-1}(E_{N-1} + x_{N-1}) - F^{-1}(E_N) e^{-\alpha_N}\big] e^{-\sum_{k=j+1}^{N-1} \alpha_k} \\
&\qquad + F^{-1}(E_N + x_N) e^{-\sum_{k=j+1}^N \alpha_k}.
\end{aligned}$$

Plugging in (31) and (32), simplifications occur and we get

$$\sum_{n=j}^{N}\left[F^{-1}(E_n+x_n)-F^{-1}(E_n)\right]e^{-\sum_{k=j+1}^{n}\alpha_k}=\nu-F^{-1}(E_j).$$

Plugging this back into (33) yields $\lambda=(\nu-F^{-1}(E_j))E_j$ for $j=1,\dots,N$. Solving this equation together with (31) for ν and λ implies that necessarily

$$\begin{aligned}\nu &= \frac{F^{-1}(E_{i-1}+x_{i-1})-e^{-\alpha_i}F^{-1}(e^{-\alpha_i}(E_{i-1}+x_{i-1}))}{1-e^{-\alpha_i}},\\ \lambda &= e^{-\alpha_i}(E_{i-1}+x_{i-1})\frac{F^{-1}(E_{i-1}+x_{i-1})-F^{-1}(e^{-\alpha_i}(E_{i-1}+x_{i-1}))}{1-e^{-\alpha_i}},\end{aligned}$$

for $i=1,\dots,N$. Lemma 3.2 (c) thus implies that

$$\alpha_1=\cdots=\alpha_N \qquad \text{and} \qquad x_0=E_1+x_1=\cdots=E_{N-1}+x_{N-1}.$$

This gives $\boldsymbol{\alpha}=\boldsymbol{\alpha}^*$. Moreover, (9) holds since $x_i=(1-a^*)x_0$ for $i=1,\dots,N-1$. We also get $E_i=a^*x_0$ for $i=1,\dots,N$. Note next that $x_N=X_0-x_0-(N-1)(1-a^*)x_0$ and therefore $E_N+x_N=X_0-N(1-a^*)x_0$. Equation (8) now follows from the fact that

$$F^{-1}(X_0-N(1-a^*))=F^{-1}(E_N+x_N)=\frac{\partial C^{(1)}}{\partial x_N}(\boldsymbol{x},\boldsymbol{\alpha})=\nu=\frac{F^{-1}(x_0)-a^*F^{-1}(a^*x_0)}{1-a^*}.$$

This concludes the proof of the theorem. □

Proof of Corollary 2.8: Suppose that $(\boldsymbol{x},\boldsymbol{\alpha})$ is a round trip such that $C^{(1)}(\boldsymbol{x},\boldsymbol{\alpha})\le 0$. We can assume w.l.o.g. that $\boldsymbol{\alpha}\in\mathcal{A}^*$, for otherwise we can simply merge those trades that occur at the same time into a single trade. Then

$$C^{(1)}(\boldsymbol{x},\boldsymbol{\alpha})=\lim_{\varepsilon\downarrow 0}C^{(1)}(\boldsymbol{x}+\varepsilon\mathbf{1},\boldsymbol{\alpha}).$$

But $C^{(1)}(\boldsymbol{x}+\varepsilon\mathbf{1},\boldsymbol{\alpha})>0$ for each $\varepsilon>0$, due to our previous results. Hence we must have $C^{(1)}(\boldsymbol{x},\boldsymbol{\alpha})=0$. The strict convexity of $C^{(1)}(\cdot,\boldsymbol{\alpha})$, established in the proof of Lemma 3.5, implies that there can be at most one minimizer of $C^{(1)}(\cdot,\boldsymbol{\alpha})$ in the class of round trips. Since we clearly have $C^{(1)}(\mathbf{0},\boldsymbol{\alpha})=0$, we must conclude that $\boldsymbol{x}=\mathbf{0}$. □

3.4 Proofs for Model 2

We first need two technical lemmas. The first one states some properties for functions satisfying Assumption 2.11, while the second one is the analogue of Lemma 3.2 and prepares for the uniqueness of a critical point.

Lemma 3.6. *Under Assumption 2.11, the following conclusions hold.*

(a) $x\mapsto xf(x)$ *is increasing on* $\mathbb{R}$ *(or, equivalently,* $\tilde{F}$ *is convex).*

(b) *For all* $a\in(0,1)$, $x\mapsto af(ax)/f(x)$ *is nondecreasing on* $\mathbb{R}_+$ *and nonincreasing on* $\mathbb{R}_-$ *and takes values in* $(0,1)$.

(c) *For all* $x > 0$,

$$(0,1) \ni a \longmapsto \frac{1 - a^2 f(ax)/f(x)}{1 - af(ax)/f(x)}$$

is increasing.

(d) *For all* $x > 0$,

$$(0,1) \ni a \longmapsto a^{-1} \frac{1 - a^2 f(x)/f(x/a)}{1 - af(x)/f(x/a)}$$

is decreasing.

Proof: (a) The derivative is positive since $xf'(x)/f(x) > -1$ by Assumption 2.11.

(b) Since $x \mapsto xf(x)$ is increasing, $af(ax)/f(x) = [axf(ax)]/[xf(x)] \in (0,1)$. The derivative of $x \mapsto af(ax)/f(x)$ is equal to $[a^2 f'(ax)f(x) - af(ax)f'(x)]/f(x)^2$. It is nonnegative on $\mathbb{R}_+$ and nonpositive on $\mathbb{R}_-$ if and only if

$$\frac{af'(ax)}{f(ax)} \geq \frac{f'(x)}{f(x)} \quad \text{for } x \geq 0, \text{ and} \quad \frac{af'(ax)}{f(ax)} \leq \frac{f'(x)}{f(x)} \quad \text{for } x \leq 0.$$

These conditions hold as a direct consequence of (10).

(c) For a fixed $x \geq 0$, we set $\psi(a) = af(ax)/f(x)$, which takes values in $(0,1)$. We need to show that

$$\frac{d}{da} \frac{1 - a\psi(a)}{1 - \psi(a)} = \frac{(1-a)\psi'(a) - \psi(a)(1 - \psi(a))}{(1 - \psi(a))^2} > 0.$$

This condition holds if and only if

$$\frac{\psi'(a)}{\psi(a)} > \frac{1 - \psi(a)}{1 - a}.$$

It is thus sufficient to show that ψ'/ψ is nonincreasing, since then we would have

$$1 - \psi(a) < \int_a^1 \frac{\psi'(u)}{\psi(u)}\, du \leq (1-a)\frac{\psi'(a)}{\psi(a)}.$$

This leads to requiring $\psi\psi'' - (\psi')^2 \leq 0$, which in turn leads to the following condition:

$$1 + (ax)^2 \left(\frac{f'(ax)}{f(ax)}\right)^2 - (ax)^2 \frac{f''(ax)}{f(ax)} \geq 0 \qquad \text{for } a \in (0,1).$$

The latter condition is ensured by Assumption (11), since $xf'(x)/f(x) \in (-1, 0]$ and thus

$$\left(\frac{xf'(x)}{f(x)}\right)^2 + \frac{xf'(x)}{f(x)} < 0.$$

(d) We fix $x > 0$ and let $\tilde{\psi}(a) := af(x)/f(x/a)$. We need to show that

$$\frac{d}{da} a^{-1} \frac{1 - a\tilde{\psi}(a)}{1 - \tilde{\psi}(a)} = \frac{\tilde{\psi}(a) - 1 + a\tilde{\psi}'(a)(1-a)}{a^2(1 - \tilde{\psi}(a))^2} < 0.$$

This condition holds if and only if

$$a\tilde{\psi}'(a) < \frac{1-\tilde{\psi}(a)}{1-a}.$$

Hence it is enough to show that $a \mapsto a\tilde{\psi}'(a)$ is nondecreasing, because then we would have

$$1-\tilde{\psi}(a) > \int_a^1 u\tilde{\psi}'(u)\,du \geq (1-a)a\tilde{\psi}'(a).$$

Some calculations lead to

$$\frac{d}{da}a\tilde{\psi}'(a) = \frac{1}{f(x/a)}\left(1+\frac{x}{a}\frac{f'(x/a)}{f(x/a)}+2\left(\frac{x}{a}\frac{f'(x/a)}{f(x/a)}\right)^2-\left(\frac{x}{a}\right)^2\frac{f''(x/a)}{f(x/a)}\right),$$

which is nonnegative by Assumption (11). □

Lemma 3.7. *Under Assumption 2.11, the following conclusions hold.*

(a) *For each $a \in (0,1)$, the function*

$$\begin{aligned} h_{2,a} : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x\frac{f(x/a)/a - af(x)}{f(x/a)-af(x)} \end{aligned} \tag{34}$$

is well-defined and strictly increasing.

(b) *For all $a, b \in (0,1)$ and $\nu > 0$, we have the inequalities*

$$h_{2,a}^{-1}(\nu)/a > h_{2,b}^{-1}(\nu)$$

and

$$h_{2,b}^{-1}(\nu) < \nu.$$

(c) *The function*

$$\begin{aligned} H_2 : (0,\infty)\times(0,1) &\longrightarrow \mathbb{R}^2 \\ (x,a) &\longmapsto \left(x\frac{f(x/a)/a-af(x)}{f(x/a)-af(x)}, -x^2f(x)\frac{f(x/a)(1/a-1)}{f(x/a)-af(x)}\right) \end{aligned}$$

is one-to-one.

Proof: (a) First let us observe that the denominator of $h_{2,a}$ is positive, since $x \mapsto xf(x)$ is increasing by Lemma 3.6 (a). We have

$$h_{2,a}(x) = x\left(1+\frac{a^{-1}-1}{1-af(x)/f(x/a)}\right). \tag{35}$$

Again by Lemma 3.6, the fraction is positive and, as a function of x, nondecreasing on $\mathbb{R}_+$ and nonincreasing $\mathbb{R}_-$, which gives the result.

(b) It is clear from (35) that $h_{2,a}(x) > x$ for $x > 0$ and therefore $h_{2,a}^{-1}(x) < x$. Let us now consider $a, b \in (0,1)$, $\nu > 0$ and set $x' = h_{2,a}^{-1}(\nu)/a$, $x = h_{2,b}^{-1}(\nu)$. Then both x and x' are positive, and we need to show that $x' > x$. It follows that

$$\nu = x' \frac{f(x') - a^2 f(ax')}{f(x') - af(ax')} = x \frac{f(x/b)/b - bf(x)}{f(x/b) - bf(x)}.$$

Let us suppose by a way of contradiction that $x' \leq x$. Then, using Lemma 3.6 (b) and the fact that $u \in [0,1) \mapsto (1-au)/(1-u)$ is increasing, we get:

$$\frac{1 - a^2 f(ax)/f(x)}{1 - af(ax)/f(x)} \geq \frac{1 - a^2 f(ax')/f(x')}{1 - af(ax')/f(x')} \geq b^{-1} \frac{1 - b^2 f(x)/f(x/b)}{1 - bf(x)/f(x/b)}.$$

Again by Lemma 3.6, the left-hand-side is increasing w.r.t a and the right-hand side is decreasing w.r.t. b. Moreover, both have the same limit,

$$\frac{2 + xf'(x)/f(x)}{1 + xf'(x)/f(x)},$$

when $a \uparrow 1$ and $b \uparrow 1$, which leads to a contradiction.

(c) Let $(a_1, y_1), (a_2, y_2) \in (0,1) \times (0, \infty)$ be such that $H_2(a_1, y_1) = H_2(a_2, y_2)$. By (35), we then have

$$\begin{cases} y_1(1+\alpha_1) = y_2(1+\alpha_2) \\ y_1^2 f(y_1)\alpha_1 = y_2^2 f(y_2)\alpha_2, \end{cases} \qquad \text{where } \alpha_i := \frac{(a_i^{-1} - 1)f(y_i/a_i)}{f(y_i/a_i) - a_i f(y_i)} \text{ for } i = 1,2. \tag{36}$$

Let us assume for example that $\alpha_2 \leq \alpha_1$ and set $\eta = \alpha_2/\alpha_1 \in (0,1]$. Eliminating y_1 in (36) yields

$$\phi(\eta) := \left(\frac{1+\eta\alpha_1}{1+\alpha_1}\right)^2 f\left(y_2 \frac{1+\eta\alpha_1}{1+\alpha_1}\right) - \eta f(y_2) = 0.$$

Note that $\alpha_i \in (0,1)$ by Lemma 3.6 (b) and hence $2\alpha_1 < 1 + \alpha_1$. By using in addition that f is nonincreasing on $\mathbb{R}_+$ and that $x \mapsto xf(x)$ is increasing by Lemma 3.6, we obtain

$$\begin{aligned} \phi'(\eta) &= \frac{2\alpha_1}{1+\alpha_1} \frac{1+\eta\alpha_1}{1+\alpha_1} f\left(y_2 \frac{1+\eta\alpha_1}{1+\alpha_1}\right) + \frac{\alpha_1 y_2}{1+\alpha_1} \left(\frac{1+\eta\alpha_1}{1+\alpha_1}\right)^2 f'\left(y_2 \frac{1+\eta\alpha_1}{1+\alpha_1}\right) - f(y_2) \\ &< \frac{1+\eta\alpha_1}{1+\alpha_1} f\left(y_2 \frac{1+\eta\alpha_1}{1+\alpha_1}\right) - f(y_2) \leq 0. \end{aligned}$$

Thus, $\eta = 1$ is the only zero of $\phi(\eta)$. We may thus conclude that $\alpha_1 = \alpha_2$ and in turn that $y_1 = y_2$. Finally, the equality $\alpha_1 = \alpha_2$ leads to $a_1 = a_2$ due to Lemma 3.6 (d), since

$$1 + \alpha_i = a_i^{-1} \frac{1 - a_i^2 f(y_i)/f(y_i/a_i)}{1 - a_i f(y_i)/f(y_i/a_i)}.$$

□

In Model 2, we need to minimize the following the cost functional:

$$C^{(2)}(x_0, \ldots, x_n, \boldsymbol{\alpha}) = \sum_{n=0}^{N} G(F(D_n) + x_n) - G(F(D_n)), \tag{37}$$

where $D_0 = 0$ and $D_n = e^{-\alpha_n} F^{-1}(x_{n-1} + F(D_{n-1}))$ for $1 \le n \le N$. By $\widehat{f}(x) = 1/f(F^{-1}(x))$ we denote again the derivative of F^{-1}.

Lemma 3.8. *We have the following recursive formula for* $i = 0, \dots, N-1$,

$$\frac{\partial C^{(2)}}{\partial x_i} = F^{-1}\big(F(D_i) + x_i\big) + e^{-\alpha_{i+1}} f(D_{i+1}) \widehat{f}\big(x_i + F(D_i)\big) \Big[\frac{\partial C^{(2)}}{\partial x_{i+1}} - D_{i+1}\Big]. \tag{38}$$

Moreover, for $j = 1, \dots, N$,

$$\frac{\partial C^{(2)}}{\partial \alpha_j} = -D_j f(D_j) \left(\frac{\partial C^{(2)}}{\partial x_j} - D_j \right) \tag{39}$$

Proof: We have for $n \in \{1, \dots, N\}$,

$$\begin{array}{llllll}
D_n & & & & & \\
\| & & & & & \\
e^{-\alpha_n} F^{-1}(x_{n-1}+ & F(D_{n-1})) & & & & \\
 & \| & & & & \\
 & \ddots & & & & \\
 & e^{-\alpha_{i+2}} F^{-1}(x_{i+1}+ & F(D_{i+1})) & & & \\
 & & \| & & & \\
 & & e^{-\alpha_{i+1}} F^{-1}(x_i+ & F(D_i)) & & \\
 & & & \| & & \\
 & & & \ddots & & \\
 & & & e^{-\alpha_1} F^{-1}(x_0). & &
\end{array}$$

Using this scheme, we obtain the following recursive relations between the derivatives of D_n with respect to x_i.

$$\begin{aligned}
\frac{\partial D_n}{\partial x_i} &= 0 \qquad \text{for } i \ge n, \\
\frac{\partial D_n}{\partial x_{n-1}} &= e^{-\alpha_n} \widehat{f}(x_{n-1} + F(D_{n-1})), \\
\frac{\partial D_n}{\partial x_i} &= e^{-\alpha_{i+1}} f(D_{i+1}) \widehat{f}(x_i + F(D_i)) \frac{\partial D_n}{\partial x_{i+1}} \qquad \text{for } 1 \le i \le n-1.
\end{aligned}$$

Thus, by (37),

$$\begin{aligned}
\frac{\partial C^{(2)}}{\partial x_i} &= F^{-1}(x_i + F(D_i)) + \sum_{n=i+1}^{N} f(D_n) \big[F^{-1}(x_n + F(D_n)) - D_n\big] \frac{\partial D_n}{\partial x_i} \qquad (40) \\
&= F^{-1}(x_i + F(D_i)) + e^{-\alpha_{i+1}} f(D_{i+1}) \widehat{f}\big(x_i + F(D_i)\big) \big[F^{-1}(x_{i+1} + F(D_{i+1})) - D_{i+1}\big] \\
&\quad + e^{-\alpha_{i+1}} f(D_{i+1}) \widehat{f}\big(x_i + F(D_i)\big) \sum_{n=i+2}^{N} f(D_n) \big[F^{-1}(x_n + F(D_n)) - D_n\big] \frac{\partial D_n}{\partial x_{i+1}}.
\end{aligned}$$

By (40), the sum in the preceding line satisfies

$$\sum_{n=i+2}^{N} f(D_n)\big[F^{-1}(x_n + F(D_n)) - D_n\big]\frac{\partial D_n}{\partial x_{i+1}} = \frac{\partial C^{(2)}}{\partial x_{i+1}} - F^{-1}(x_{i+1} + F(D_{i+1})).$$

Hence,

$$\frac{\partial C^{(2)}}{\partial x_i} = F^{-1}(x_i + F(D_i)) + e^{-\alpha_{i+1}} f(D_{i+1})\widehat{f}(x_i + F(D_i))\Big[\frac{\partial C^{(2)}}{\partial x_{i+1}} - D_{i+1}\Big],$$

which is our formula (38).

As to (39), we use again the recursive scheme at the beginning of this proof to obtain formulas for the derivatives of D_n with respect to α_j:

$$\begin{aligned}
\frac{\partial D_n}{\partial \alpha_j} &= 0 \qquad \text{for } j > n,\\
\frac{\partial D_n}{\partial \alpha_n} &= -D_n,\\
\frac{\partial D_n}{\partial \alpha_j} &= -D_j f(D_j)\frac{\partial D_n}{\partial x_j} \qquad \text{for } 1 \le j \le n-1.
\end{aligned}$$

We therefore obtain from (37):

$$\begin{aligned}
\frac{\partial C^{(2)}}{\partial \alpha_i} &= \sum_{n=i}^{N} f(D_n)\big[F^{-1}(x_n + F(D_n)) - D_n\big]\frac{\partial D_n}{\partial \alpha_i}\\
&= -D_i f(D_i)[F^{-1}(x_i + F(D_i)) - D_i]\\
&\quad - \sum_{n=i+1}^{N} f(D_n) D_i f(D_i)\big[F^{-1}(x_n + F(D_n)) - D_n\big]\frac{\partial D_n}{\partial x_i}\\
&= -D_i f(D_i)\left(F^{-1}(x_i + F(D_i)) - D_i + \frac{\partial C^{(2)}}{\partial x_i} - F^{-1}(x_i + F(D_i))\right)\\
&= -D_i f(D_i)\left(\frac{\partial C^{(2)}}{\partial x_i} - D_i\right). \quad \square
\end{aligned}$$

Remark 3.9. A consequence of this lemma is that the optimal strategy given by [2] on the homogeneous time spacing grid $\mathcal{T}^*$ is a critical point for the minimization in $(\boldsymbol{x}, \boldsymbol{\alpha})$. Indeed, we have then $D_i = a^* F^{-1}(\xi_0^{(2)})$ for any i, and therefore $\frac{\partial C^{(2)}}{\partial \alpha_i}$ does not depend on i.

Lemma 3.10. *Assume that* $\boldsymbol{\alpha} \in \mathcal{A}^*$. *Then,* $C^{(2)}(\boldsymbol{x}, \boldsymbol{\alpha}) \to \infty$ *as* $|\boldsymbol{x}| \to \infty$.

Proof: Equation (37) yields

$$\begin{aligned} &C^{(2)}(\boldsymbol{x},\boldsymbol{\alpha})\\ &= \sum_{n=0}^{N} \tilde F\big(F^{-1}(F(D_n)+x_n)\big) - \tilde F(D_n))\\ &= \sum_{n=0}^{N-1} \Big[\tilde F\big(F^{-1}(F(D_n)+x_n)\big) - \tilde F\big(e^{-\alpha_{n+1}}F^{-1}(F(D_n)+x_n)\big)\Big] + \tilde F\big(F^{-1}(F(D_N)+x_N\big). \end{aligned}$$

Since f is nondecreasing on $\mathbb{R}_-$ and nonincreasing on $\mathbb{R}_+$, we have for $x\in\mathbb{R},\ a\in[0,1)$,

$$\tilde F(x)-\tilde F(ax) = \int_{ax}^{x} xf(x)dx \geq \frac{1}{2}f(x)x^2(1-a^2) =: H(x).$$

Defining $a=\max_{i=1,\dots,N} e^{-\alpha_i}<1$, we thus get

$$C^{(2)}(\boldsymbol{x},\boldsymbol{\alpha}) \geq \frac{1}{2}(1-a^2)H(|T_2(\boldsymbol{x})|_\infty),$$

where

$$T_2(\boldsymbol{x}) = \big(x_0, x_1+F^{-1}(D_1),\dots,x_N+F^{-1}(D_N)\big).$$

From (10), $x\mapsto xf(x)$ is increasing and therefore $H(x)\to+\infty$ as $|x|\to+\infty$. It is therefore sufficient to have $T_2(\boldsymbol{x})\to+\infty$ for $|\boldsymbol{x}|\to+\infty$. To this end, let $(\boldsymbol{x}^k)$ be a sequence such that the sequence $(T_2(\boldsymbol{x}^k))$ is bounded. We will show that $(\boldsymbol{x}^k)$ then must also be bounded. It is clear that the first coordinate x_0^k is bounded. Therefore, $F^{-1}(D_1^k)$ is also bounded, which in turn implies that the second coordinate of $(T_2(\boldsymbol{x}^k))$ is bounded. We then get that (x_1^k) is bounded. An easy induction on coordinates thus gives the desired result. □

We are now in position to prove the main results for Model 2.

Proof of Proposition 2.12: We can assume without loss of generality that $\boldsymbol{\alpha}\in\mathcal{A}^*$, for otherwise we can simply merge two trades occurring at the same time into a single trade. If $\boldsymbol{x}$ is the minimizer of $C^{(2)}(\cdot,\boldsymbol{\alpha})$ on Ξ, then there must be a Lagrange multiplier ν such that $\boldsymbol{x}$ is a critical point of $\boldsymbol{y}\mapsto C^{(1)}(\boldsymbol{y},\boldsymbol{\alpha})-\nu\sum_{i=0}^N y_i$. Hence, (38) yields that

$$\nu = h_{2,a_{i+1}}(D_{i+1}), \qquad i=0,\dots,N-1,$$

where $a_{i+1}=e^{-\alpha_{i+1}}$ and $h_{2,a}$ is defined as in (34). Since $D_{i+1}=a_{i+1}F^{-1}(x_i+F(D_i))$, we get with Lemma 3.7 that

$$x_0 = F(h_{2,a_1}^{-1}(\nu)/a_1),\ x_i = F(h_{2,a_{i+1}}^{-1}(\nu)/a_{i+1}) - F(h_{2,a_i}^{-1}(\nu)), \qquad i=1,\dots,N-1.$$

For the last trade, we also get that $\nu=F^{-1}(x_N+F(D_N))$ and $x_N=F(\nu)-F(h_{2,a_N}^{-1}(\nu))$. Therefore, summing all the trades, we get:

$$X_0 = F(\nu) + \sum_{i=1}^{N}\big[F(h_{2,a_i}^{-1}(\nu)/a_i) - F(h_{2,a_i}^{-1}(\nu))\big]. \tag{41}$$

Now let us observe that F is increasing on $\mathbb{R}$, and for any $a \in (0,1)$, $y \mapsto F(y/a) - F(y)$ is increasing (its derivative is positive by Lemma 3.6 (a)). Besides, F and $h_{2,a}^{-1}$ are increasing for any $a \in (0,1)$ and therefore ν is uniquely determined by the above equation. We have moreover $\nu > 0$ because the left-hand side vanishes when ν is equal to 0. This proves that there a unique critical point, which then is necessarily the global minimum of $C^{(2)}$ by Lemma 3.10.

Finally, $x_i > 0$ for $i = 0, \dots, N$, due to Lemma 3.6 and the fact that F is increasing. □

Proof of Theorem 2.13: We get the existence of a minimizer $(\boldsymbol{\xi}^{(2)}, \boldsymbol{\alpha}^*)$ and the fact that it belongs to $\Xi_+ \times \mathcal{A}^*$ exactly as in the proof of Theorem 2.7.

Now suppose that $(\boldsymbol{x}, \boldsymbol{\alpha})$ is a minimizer of $C^{(2)}$. Due to the preceding step, there must be Lagrange multipliers $\nu, \lambda \in \mathbb{R}$ such that $(\boldsymbol{x}, \boldsymbol{\alpha})$ is a critical point of $(\boldsymbol{y}, \boldsymbol{\beta}) \mapsto C^{(2)}(\boldsymbol{y}, \boldsymbol{\beta}) - \nu \sum_{i=0}^N y_i - \lambda \sum_{j=1}^N \beta_j$.

From (38), we easily obtain that for $i = 1, \dots, N$,

$$\nu = \frac{e^{-\alpha_i} f(D_i)}{f(e^{\alpha_i} D_i)} [\nu - D_i] + e^{\alpha_i} D_i$$

and $\nu = F^{-1}(x_N + F(D_N))$ for the last trade. We then deduce from (39) that

$$\begin{aligned} \nu &= D_i \frac{e^{\alpha_i} f(e^{\alpha_i} D_i) - e^{-\alpha_i} f(D_i)}{f(e^{\alpha_i} D_i) - e^{-\alpha_i} f(D_i)} \\ \lambda &= -D_i^2 f(D_i) \frac{(e^{\alpha_i} - 1) f(e^{\alpha_i} D_i)}{f(e^{\alpha_i} D_i) - e^{-\alpha_i} f(D_i)}, \end{aligned}$$

i.e., $(\nu, \lambda) = H_2(D_i, a_i)$ with $a_i = e^{-\alpha_i}$. As in the proof of Proposition 2.12 we get (41), which ensures $\nu > 0$ and in turn $D_i > 0$ for $i = 1, \dots, N$. Due to Lemma 3.7, H_2 is one-to-one on $(0, \infty) \times (0, 1)$, and therefore $\alpha_1 = \dots = \alpha_N$ and $D_1 = \dots = D_N$. Then, $D_1 = a^* F^{-1}(x_0)$. Since $D_{i+1} = a^* F^{-1}(x_i + F(D_i))$, we get $x_i = x_0 - F(D_i) = x_0 - F(a^* F^{-1}(x_0))$, and therefore $x_N = X_0 - N x_0 + (N-1) F(a^* F^{-1}(x_0))$. Combining this with $\nu = F^{-1}(x_N + F(D_N))$, we get

$$F^{-1}\big(X_0 - N[x_0 - F(a^* F^{-1}(x_0))]\big) = h_{2,a^*}(F^{-1}(x_0)).$$

We refer to [2, Lemma C.3] for the existence, uniqueness, and positivity of the solution x_0 of this equation. It follows that there is a unique critical point of $C^{(2)}$ on $\Xi_+ \times \mathcal{A}^*$, which is necessarily the global minimum. □

Proof of Corollary 2.14: As in Model 2, we can show that a round trip such that $C^{(2)}(\boldsymbol{x}, \boldsymbol{\alpha}) \leq 0$ necessarily satisfies $C^{(2)}(\boldsymbol{x}, \boldsymbol{\alpha}) = 0$. Moreover for $\boldsymbol{\alpha} \in \mathcal{A}^*$, we see looking at the proof of Proposition 2.12 that $(0, \dots, 0)$ is the only critical point when $X_0 = 0$ since we necessarily have $\nu = 0$ by (41). Therefore, it is also the unique minimum of $C^{(2)}$ by Lemma 3.10. □

Acknowledgement. This research was started at Cornell University. The authors would like to thank the School of Operations Research and Information Engineering for its support.

References

[1] Alfonsi, A., Fruth, A., Schied, A. *Constrained portfolio liquidation in a limit order book model.* Banach Center Publ. 83, 9-25 (2008).

[2] Alfonsi, A., Fruth, A., Schied, A. *Optimal execution strategies in limit order books with general shape functions.* To appear in Quantitative Finance (2009).

[3] Alfonsi, A., Schied, A., Slynko, A. *Resilience of limit order books, price manipulations, and optimal execution strategies.* In preparation.

[4] Almgren, R. *Optimal execution with nonlinear impact functions and trading-enhanced risk.* Applied Mathematical Finance 10 , 1-18 (2003).

[5] Almgren, R., Chriss, N. *Value under liquidation.* Risk, Dec. 1999.

[6] Almgren, R., Chriss, N. *Optimal execution of portfolio transactions.* J. Risk 3, 5-39 (2000).

[7] Almgren, R., Lorenz, J. *Adaptive arrival price.* In: Algorithmic Trading III: Precision, Control, Execution, Brian R. Bruce, editor, Institutional Investor Journals (2007).

[8] Almgren, R., Thum, C. Hauptmann, E., Li, E. *Equity market impact.* Risk, July (2005).

[9] Bayraktar, M., Ludkovski, M. *Optimal trade execution in illiquid markets.* Preprint, `http://arxiv.org/abs/0902.2516`.

[10] Bertsimas, D., Lo, A. *Optimal control of execution costs.* Journal of Financial Markets, 1, 1-50 (1998).

[11] Biais, B., Hillion, P., Spatt, C. *An empirical analysis of the limit order book and order flow in Paris Bourse.* Journal of Finance 50, 1655-1689 (1995).

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

[13] Gatheral, J. *No-Dynamic-Arbitrage and Market Impact.* Preprint on `http://papers.ssrn.com`.

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

[15] Mönch, B. *Strategic trading in illiquid markets.* Springer, 2004.

[16] Obizhaeva, A., Wang, J. *Optimal Trading Strategy and Supply/Demand Dynamics*, Preprint (2005).

[17] Potters, M., Bouchaud, J.-P. *More statistical properties of order books and price impact.* Physica A 324, No. 1-2, 133-140 (2003).

[18] Schied, A., Schöneborn, T. *Optimal basket liquidation with finite time horizon for CARA investors.* Preprint, Cornell University (2008).

[19] Schied, A., Schöneborn, T. *Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets.* Finance Stoch. **13**, 181-204 (2009).

[20] Weber, P., Rosenow, B. *Order book approach to price impact.* Quantitative Finance 5, no. 4, 357-364 (2005).