MTH 9879 Market Microstructure Models, Spring 2019

Lecture 10: Optimal trading strategies: Transient market impact

Jim Gatheral
Department of Mathematics

Outline of Lecture 10

- The order book model of Obizhaeva and Wang

- The model of Alfonsi, Fruth and Schied

- What choices of decay kernel $G(\cdot)$ preclude price manipulation when instantaneous market impact $f(\cdot)$ is linear?

- Closed-form examples of optimal strategies

- Nonlinear transient market impact

Overview of execution algorithm design

Typically, an execution algorithm has three layers:

- The macrotrader
    - This highest level layer decides how to slice the order: when the algorithm should trade, in what size and for roughly how long.

- The microtrader
    - Given a slice of the order to trade (a child order), this level decides whether to place market or limit orders and at what price level(s).

- The smart order router
    - Given a limit or market order, which venue should this order be sent to?

In this lecture, we are again concerned with the highest level of the algorithm: How to slice the order.

Statement of the problem

- Given a model for the evolution of the stock price, we would like to find an optimal strategy for trading stock, the strategy that minimizes some cost function over all permissible strategies.
    - We will specialize to the case of stock liquidation where the initial position $x_0 = X$ and the final position $x_T = 0$.

- A *static* strategy is one determined in advance of trading.

- A *dynamic* strategy is one that depends on the state of the market during execution of the order, *i.e.* on the stock price.
    - Delta-hedging is an example of a dynamic strategy. VWAP is an example of a static strategy.

- It will turn out, surprisingly, that in many models, a statically optimal strategy is also dynamically optimal.

P&L and cost associated with a trading strategy

Let x_t be a trading strategy. The corresponding P&L (up to time t), denoted by $\Pi_t(x)$, is identified as

$$\Pi_t(x) := x_t(S_t - S_0) + \int_0^t (S_0 - \tilde{S}_\tau)\, dx_\tau.$$

- The first term represents unrealized P&L on transactions (purchases or sales) yet to be executed.

- The second term corresponds to realized P&L on transactions executed up to time t.

Implementation shortfall

- The P&L $\Pi_T(x)$ associated with a trading strategy x terminating at time T is on average negative.

- The quantity $C_T(x) := -\Pi_T(x)$ is often referred to as *implementation shortfall*.
    - Our objective will be to find the strategy x that minimizes the cost $C_T(x)$.

A transient impact model

The price process assumed in [Gatheral][6] is

$$S_t = S_0 + \int_0^t f(v_s)\, G(t - s)\, ds + M_t$$

where $v_t = \dot{x}_t$ is the (signed) trading rate and M_t is a zero mean martingale, a noise term representing the impact of the trading of other market participants.

- $f(v_s)$ is referred to as the *instantaneous market impact function*, which represents the impact of trading at time s, and $G(t - s)$ is a decay factor.
    - Note that $f(v) > 0$ if $v > 0$; whereas $f(v) < 0$ if $v < 0$.

- S_t follows an arithmetic random walk (Brownian motion) with a drift that depends on the accumulated impacts of previous trades.

- The cumulative impact of (others') trading is implicitly in S_0 and the noise term.

- The model is a generalization of processes due to Almgren, Bouchaud, and Obizhaeva and Wang.

- The model corresponds to the "autogressive model" or "propagator model" of Bouchaud et al.

P&L and cost of trading in the transient impact model

As before, at the end of execution period T, the P&L reads

$$\Pi_T(x) = x_T(S_T - S_0) + \int_0^T (S_0 - \tilde{S}_u)\, dx_u,$$

where x_T is the number of shares left unexecuted at time T, and the $\tilde{S}_u$ is the execution price at time u.

- We assume no bid-offer spread or other friction so $\tilde{S}_u = S_u$, the market price.

Hence, in the transient impact (TI) model

$$
\begin{aligned}
\Pi_T(x) = x_T(S_T - S_0) &+ \int_0^T (S_0 - S_u)\, dx_u \\
&= -\int_0^T \int_0^t f(v_u)\, G(t-u)\, du\, dx_t - \int_0^T M_t\, dx_t \quad \text{(since } x_T = 0\text{)} \\
&= -\int_0^T \int_0^t f(v_u)\, G(t-u)\, du\, dx_t - (M_T x_T - M_0 x_0) \\
&\quad + \int_0^T x_t\, dM_t \quad \text{(Integration by parts)} \\
&= -\int_0^T \int_0^t f(v_u)\, G(t-u)\, du\, dx_t + \int_0^T x_t\, dM_t \quad \text{(since } x_T = 0 \text{ and } M_0 = 0\text{).}
\end{aligned}
$$

Thus, the expected cost corresponding to the trading strategy x is given by

$$\mathbb{E}\left[C_T(x)\right] = \mathbb{E}\left[-\Pi_T(x)\right] = \mathbb{E}\left[\int_0^T \int_0^t f(v_u)\, G(t-u)\, du\, dx_t\right].$$

The optimal strategy of a risk neutral trader

For a risk neutral trader whose objective is to minimize the cost of trading, the optimal control problem reads

$$\min_v \mathbb{E}\left[C_T(x)\right] = \min_v \int_0^T \int_0^t f(v_u)\, G(t-u)\, du\, dx_t,$$

where the state variable x_t is driven by $dx_t = v_t dt$ with the constraints $x_0 = X$ and $x_T = 0$.

This is equivalent to the variational problem

$$\min_{v} \mathbb{E}\left[C_T(x)\right] = \min_{v}\left\{\int_0^T v_t\,dt\int_0^t f(v_u)\,G(t-u)\,du\right\},$$

subject to the constraint $\int_0^T v_t dt = X$.

- Recall that the statically optimal strategy is dynamically optimal (Predoiu, Shaikhet and Shreve).

Lagrange multiplier

To derive the Euler-Lagrange equation, consider the Lagrangian

$$\mathcal{L}(v,\lambda) = \int_0^T \int_0^t v_t\,G(t-s)f(v_s)\,ds\,dt$$

$$- \lambda\left(\int_0^T v_t\,dt - X\right)$$

$$= \int_0^T\left[\int_0^t G(t-s)f(v_s)\,ds - \lambda\right]v_t\,dt + \lambda\,X,$$

where λ is a Lagrange multiplier.

Euler-Lagrange equation

Let φ be a perturbation. Consider the first order criterion for the Lagrangian $\mathcal{L}$:

$$0 = \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\mathcal{L}(v_t + \epsilon\,\varphi_t,\lambda)$$

$$= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\int_0^T\left[\int_0^t G(t-s)f(v_s + \epsilon\,\varphi_s)\,ds - \lambda\right](v_t + \epsilon\,\varphi_t)dt$$

$$= \int_0^T\left[\int_0^t G(t-s)f'(v_s)\,\varphi_s\,ds\right]v_t\,dt + \int_0^T\left[\int_0^t G(t-s)f(v_s)\,ds - \lambda\right]\varphi_t\,dt$$

$$= \int_0^T\left[\int_t^T G(s-t)f'(v_t)\,v_s\,ds + \int_0^t G(t-s)f(v_s)\,ds - \lambda\right]\varphi_t\,dt.$$

Since φ_t is arbitrary (because the first order criterion must hold for any perturbation), we must have

(2)

$$\int_0^t f(v_s)\,G(t-s)\,ds + f'(v_t)\int_t^T v_s\,G(s-t)\,ds = \lambda$$

for all $0 \le t \le T$.

- This a generalized Fredholm integral equation of the first kind.
- The Lagrange multiplier λ is determined by the constraint $\int_0^T v_t dt = X$.

Obizhaeva and Wang 2005

Recall from [Obizhaeva and Wang][9] that in this model,

(1)

$$S_t = S_0 - \eta \int_0^t v_s \, e^{-\rho \, (t-s)} \, ds + \int_0^t \sigma \, dZ_s$$

with $v_t = -\dot{x}_t$.

- Market impact decays exponentially, *i.e.*, $G(\tau) = e^{-\rho \tau}$

- Instantaneous market impact is linear in the rate of trading, $f(v) = -\eta \, v$.

The expected cost of trading becomes:

$$C = \eta \int_0^T v_t \, dt \int_0^t v_s \, e^{-\rho \, (t-s)} \, ds$$
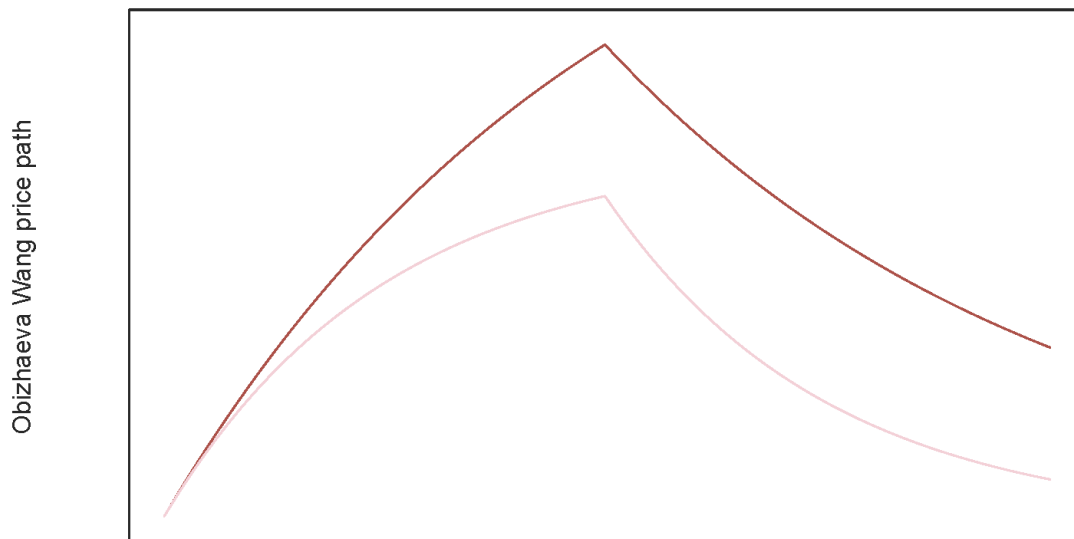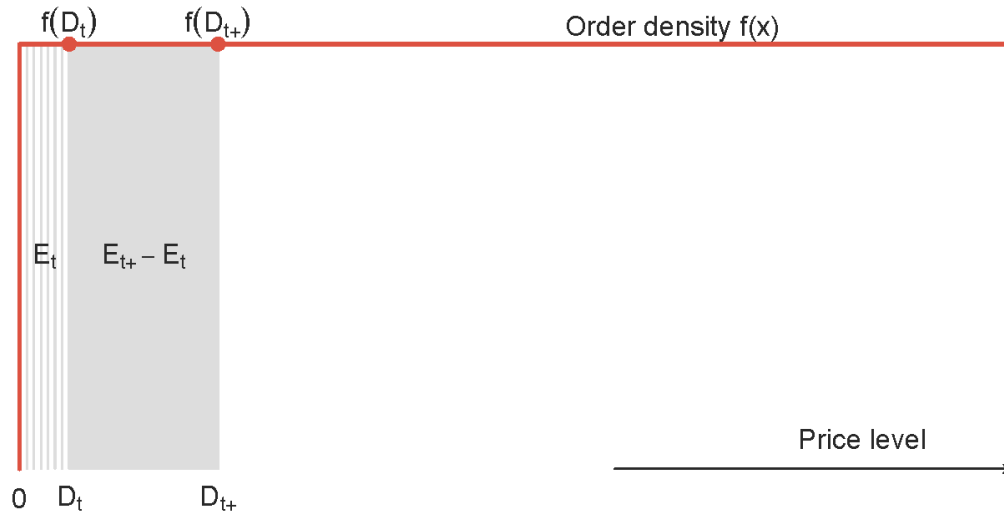
The Obizhaeva and Wang average price path

Figure 1: The Obizhaeva Wang average price path: higher ρ in pink; lower ρ in brown.

Obizhaeva Wang order book process

When a trade of size ξ is placed at time t,

$$E_t \mapsto E_{t+} = E_t + \xi$$
$$D_t = \eta\, E_t \mapsto D_{t+} = \eta\, E_{t+} = \eta\, (E_t + \xi)$$

The Euler-Lagrange equation (2) in this case reads:

(2') $$

- \eta \int_t^T\,v_s\,e^{-\rho\,(s-t)}\,ds - \eta \int_0^t\,v_s\,e^{-\rho\,(t-s)}\,ds = \lambda. $$

(2') may be rewritten as

$$\int_0^T v_s\, e^{-\rho\, |t-s|}\, ds = -\frac{\lambda}{\eta}$$

which is a Fredholm integral equation of the first kind (see [Gatheral, Schied and Slynko][7]).

Now substitute

$$v_s = A\, \delta(s) + A\, \rho + A\, \delta(s - T),$$

where A is a (to be determined) constant, into (2') to obtain

$$A\, e^{-\rho\, t} + A\, \left(1 - e^{-\rho\, t}\right) + A\, \left(1 - e^{-\rho\, (T-t)}\right) + A\, e^{-\rho\, (T-t)} = 2\, A$$

so the Fredholm integral equation is satisfied.

Moreover, from the constraint $\int_0^T v_t dt = X$,

$$\int_0^T v_t \, dt = 2A + A\rho T = X$$

so

$$A = \frac{X}{2 + \rho T}.$$

- The optimal strategy consists of a block trade at time $t = 0$, continuous trading at the rate ρ over the interval $(0, T)$ and another block trade at time $t = T$.

Consider the volume impact process E_t. The initial block-trade causes

$$0 = E_0 \mapsto E_{0+} = A$$

According to the assumptions of the model, the volume impact process reverts exponentially so

$$E_t = E_{0+} \, e^{-\rho t} + A\rho \int_0^t e^{-\rho(t-s)} \, ds = A$$

i.e. the volume impact process is constant when the trading strategy is optimal.

The model of Alfonsi, Fruth and Schied

[Alfonsi, Fruth and Schied][2] consider the following (AS) model of the order book:

- There is a continuous (in general nonlinear) density of orders $f(x)$ above some martingale ask price A_t. The cumulative density of orders up to price level x is given by
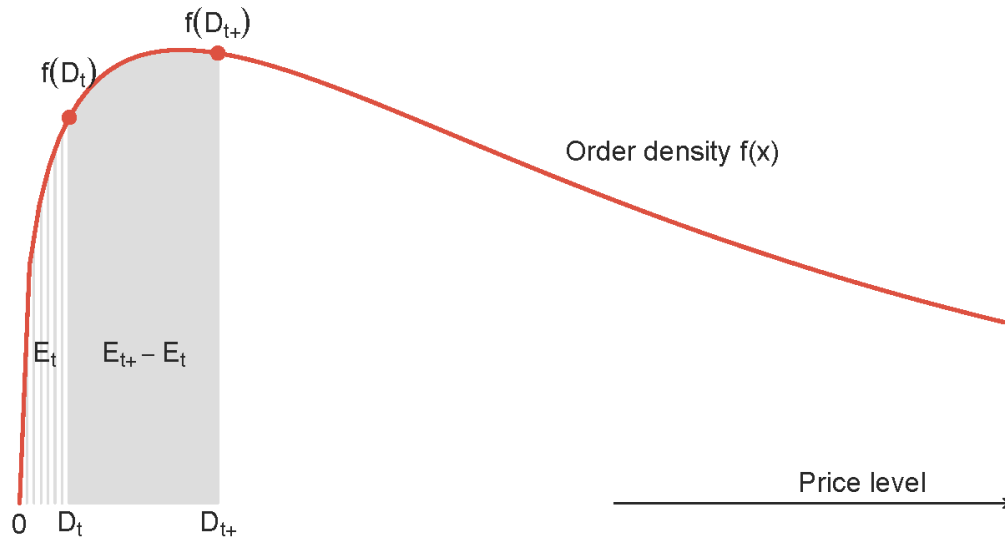
$$F(x) := \int_0^x f(y) \, dy$$

Note

The f here is not the f as in the impact function.

- Executions eat into the order book (*i.e.* executions are with market orders).

- A purchase of ξ shares at time t causes the ask price to increase from $A_t + D_t$ to $A_t + D_{t+}$ with

$$\xi = \int_{D_t}^{D_{t+}} f(x) \, dx = F(D_{t+}) - F(D_t)$$

Schematic of the model

When a trade of size ξ is placed at time t,

$$E_t = F(D_t) \mapsto E_{t+} = E_t + \xi = F(D_{t+})$$
$$D_t = F^{-1}(E_t) \mapsto D_{t+} = F^{-1}(E_{t+}) = F^{-1}(E_t + \xi)$$

Optimal liquidation strategy in the AS model

The cost of trade execution in the AS model is given by:

(3)

$$C = \int_0^T F^{-1}(E_t)\, dx_t^c + \sum_{t \leq T} [H(E_{t+}) - H(E_t)]$$

where x_t^c denotes the continuous part of x_t and

$$E_t = \int_0^t e^{-\rho\,(t-s)}\, dx_s := \int_0^t e^{-\rho\,(t-s)}\, dx_s^c + \sum_{0 \leq s \leq t} e^{-\rho(t-s)} \Delta x_s$$

is the volume impact process. Recall that $\Delta x_s = x_s - x_{s-}$ denotes the jump size of x at time s, should there be a jump.

$$H(x) = \int_0^x F^{-1}(y)\, dy$$

gives the cost of executing an instantaneous block trade of size x.

One can show that the optimal trading strategy is given by $v_t = \xi_0\, \delta(t) + \xi_0\, \rho + \xi_T \delta(T-t)$, where ξ_T satisfies

$$F^{-1}(\xi_0 + \xi_T) = F^{-1}(\xi_0) + F^{-1\,\prime}(\xi_0)\, \xi_0 \qquad \Longleftrightarrow \qquad \xi_T = -\xi_0 + F\left(F^{-1}(\xi_0) + F^{-1\,\prime}(\xi_0)\, \xi_0\right)$$

and ξ_0 is determined by

$$\int_0^T v_t\, dt = \rho \xi_0 T + F\left(F^{-1}(\xi_0) + F^{-1\,\prime}(\xi_0)\, \xi_0\right) = X.$$

Example

With $F^{-1}(x) = \sqrt{x}$,

$$\begin{aligned}
\sqrt{\xi_0 + \xi_T} &= F^{-1}(\xi_0 + \xi_T) \\
&= F^{-1}(\xi_0) + F^{-1\,\prime}(\xi_0)\, \xi_0 \\
&= \sqrt{\xi_0} + \frac{1}{2}\,\sqrt{\xi_0}
\end{aligned}$$

which has the solution $\xi_T = \frac{5}{4}\,\xi_0$. Morevoer, $\xi_0 = \frac{X}{\rho T + \frac{9}{4}}$.

Optimality of the bucket-shaped strategy

- [Predoiu, Shaikhet and Shreve][8] have shown that the bucket-shaped strategy is optimal under more general conditions than exponential resiliency.
  - Specifically, if resiliency is a function of E_t (or equivalently D_t) only, the optimal strategy has a block trades at inception and completion and continuous trading at a constant rate in-between.

Optimality and price manipulation

- For all of the models considered so far, there was an optimal strategy.

- The optimal strategy always involved trades of the same sign. So no sells in a buy program, no buys in a sell program.

- Recall that [Gatheral][4]</sup> showed that we can write down models for which price manipulation is possible.
  - For example, square root price impact with exponential decay admits price manipulation.

- In such cases, a round-trip trade can generate cash on average.
  - You would want to repeat such a trade over and over.
  - There would be no optimal strategy.

Linear transient market impact

Recall Gatheral's transient impact model assumed in [Gatheral][4]

$$S_t = S_0 + \int_0^t f(v_s)\, G(t-s)\, ds + \text{ noise/zero mean martingale}$$

- In [Gatheral, Schied and Slynko][5], this model is on the one hand extended to explicitly include discrete optimal strategies and on the other hand restricted to the case of linear market impact, $f(v) = \eta\, v$.

When the admissible strategy x is used, the price S_t is given by

(5) $$S_t = S_0 + \eta \int_{\{s$

and the expected cost of liquidation is given by

(6)

$$C(x) := \frac{\eta}{2} \int_0^T \int_0^T G(|t - s|) \, dx_s \, dx_t.$$

Condition for no price manipulation

Definition (Huberman and Stanzl)

> A *round trip* is an admissible strategy with $x_0 = x_T$. A *price manipulation strategy* is a round trip with strictly negative expected costs.

Theorem (Bochner)

> $C(x) \geq 0$ for all admissible strategies x if and only if $G(| \cdot |)$ can be represented as the Fourier transform of a positive finite Borel measure μ on $\mathbb{R}$, i.e.,
> $$G(|\tau|) = \int e^{i z \tau} \mu(dz).$$

First order condition

For linear market impact, the Euler-Lagrange equation (2) reduces to

$$\int_0^T G(|t - s|) \, dx_s = \lambda \qquad \text{for all } t \in [0, T].$$

Moreover, we have

Theorem

> Suppose that G is positive definite. Then x^* minimizes $C(\cdot)$ if and only if there is a constant λ such that x^* solves the generalized Fredholm integral equation
> (7)
> $$\int_0^T G(|t - s|) \, dx_s^* = \lambda \qquad \text{for all } t \in [0, T].$$
> In this case, $C(x^*) = \frac{1}{2} \lambda X$. In particular, λ must be nonzero as soon as G is strictly positive definite and $X \neq 0$.

Transaction-triggered price manipulation

Definition (Alfonsi, Schied, Slynko (2009))

> A market impact model admits *transaction-triggered price manipulation* if the expected costs of a sell (buy) program can be decreased by intermediate buy (sell) trades.

- As discussed in [Alfonsi, Schied and Slynko][3], transaction-triggered price manipulation can be regarded as an additional model irregularity that should be excluded.
- Transaction-triggered price manipulation can exist in models that do not admit standard price manipulation in the sense of Huberman and Stanzl definition.

Remark

It follows that in admissible models of market impact, the optimal strategy consists only of orders of the same sign. In particular, this strategy must be of bounded variation, in which case the integration by parts trick of Predoiu, Shaikhet and Shreve from Lecture 9 applies.

As we saw in Lecture 9, the practical consequence of this is that we should always set up our model in such a way that the risk term (if any) does not depend on the stock price.

Condition for no transaction-triggered price manipulation

Theorem
Suppose that the decay kernel $G(\cdot)$ is convex, satisfies $\int_0^1 G(t)\, dt < \infty$ and that the set of admissible strategies is nonempty. Then there exists a unique admissible optimal strategy x^*. Moreover, x_t^* is a monotone function of t, and so there is no transaction-triggered price manipulation.

Remark
If G is not convex in a neighborhood of zero, there is transaction-triggered price manipulation.

An instructive example

We solve a discretized version of the Fredholm equation (with 512 time points) for two similar decay kernels.

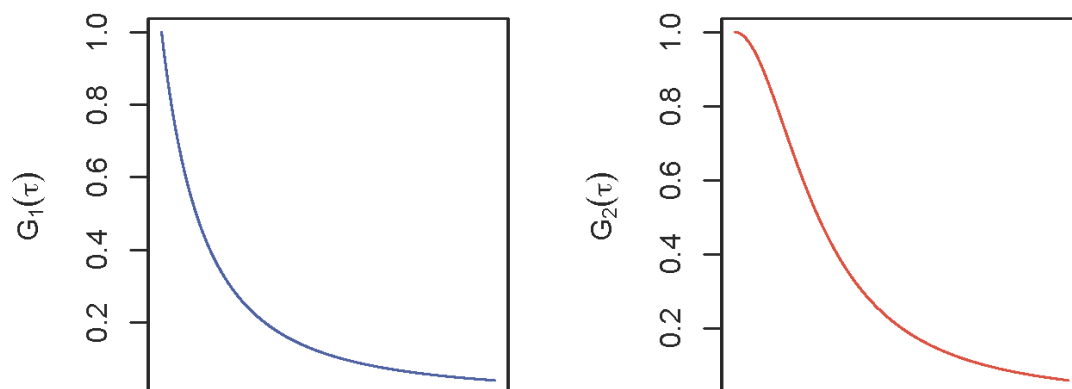$$G_1(\tau) = \frac{1}{(1+t)^2}; \qquad G_2(\tau) = \frac{1}{1+t^2}$$

Figure 2: $G_1(\cdot)$ is convex, but $G_2(\cdot)$ is concave near $\tau = 0$ so there should be a unique optimal strategy with $G_1(\cdot)$ as a choice of kernel but there should be transaction-triggered price manipulation with $G_2(\cdot)$ as the choice of decay kernel.

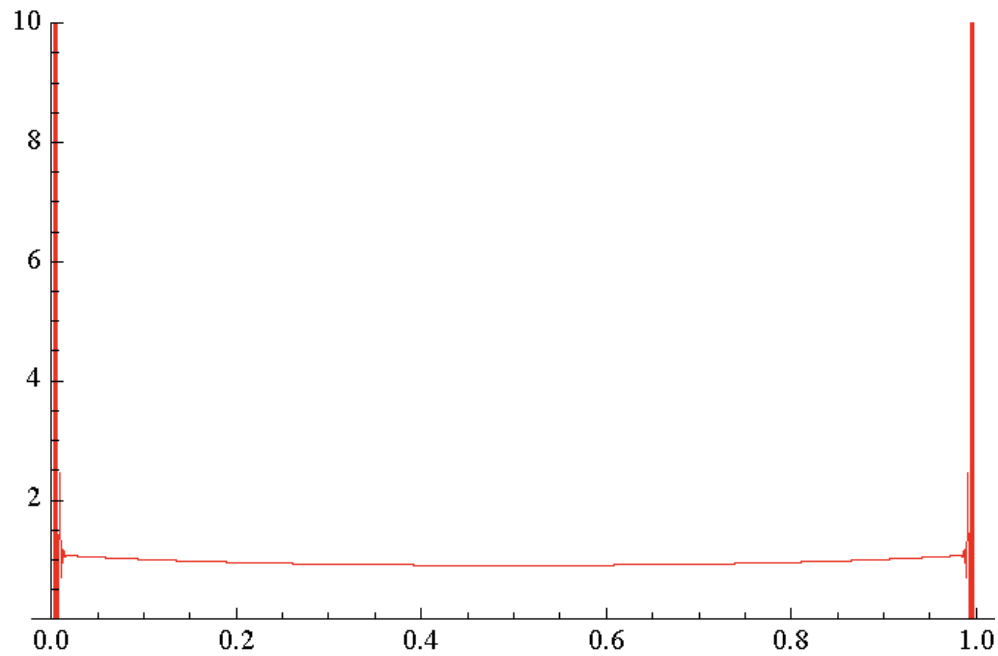## Numerical solution of the Fredholm equation with $G_1(\tau) = \frac{1}{(1+\tau)^2}$

Figure 2: With G_1 as the decay kernel, we observe block trades at $t = 0$ and $t = 1$ with continuous (nonconstant) trading in $(0, 1)$.

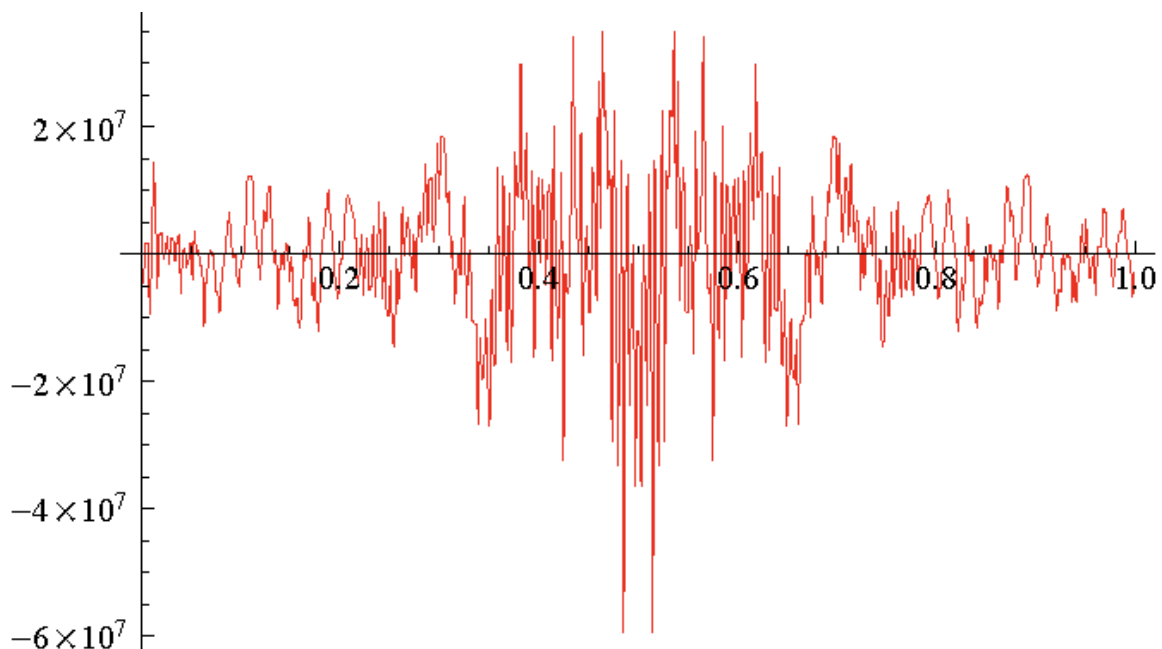## Numerical solution of the Fredholm equation with $G_2(\tau) = \frac{1}{1+\tau^2}$

Figure 3: With G_2 as the decay kernel, we see numerical evidence that the optimal strategy does not exist.

Now we give some examples of the optimal strategy with choices of kernel that preclude transaction-triggered price manipulation.

Example I: Linear market impact with exponential decay

$G(\tau) = e^{-\rho\,\tau}$ and the optimal strategy $u(s)$ solves

$$\int_0^T u(s)e^{-\rho\,|t-s|}\,ds = \text{const.}$$

We already derived the solution which is
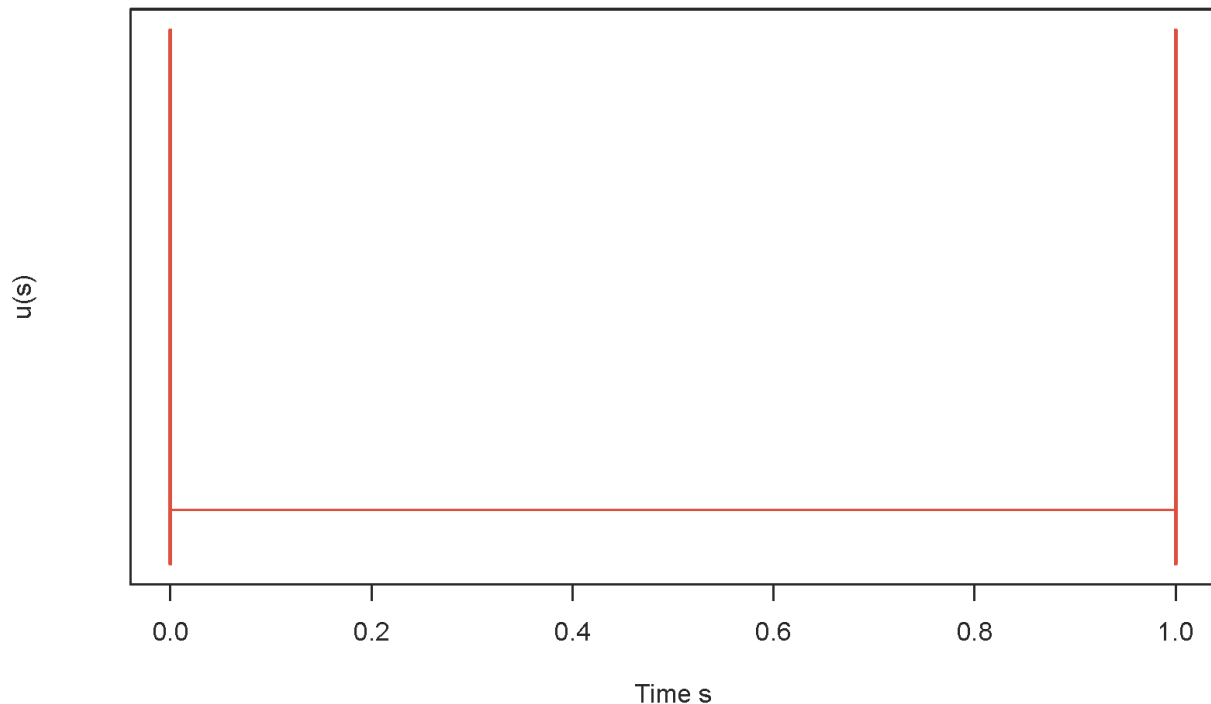
$$u(s) = A\,\{\delta(t) + \rho + \delta(T-t)\}$$

The normalizing factor A is given by

$$\int_0^T u(t)\,dt = X = A\,(2 + \rho\,T)$$

The optimal strategy consists of block trades at $t = 0$ and $t = T$ and continuous trading at the constant rate ρ between these two times.

Schematic of optimal strategy

The optimal strategy with $\rho = 0.1$ and $T = 1$

Example II: Linear market impact with power-law decay

$G(\tau) = \tau^{-\gamma}$ and the optimal strategy $u(s)$ solves

$$\int_0^T \frac{u(s)}{|t-s|^{\gamma}}\, ds = \text{const.}$$

The solution is
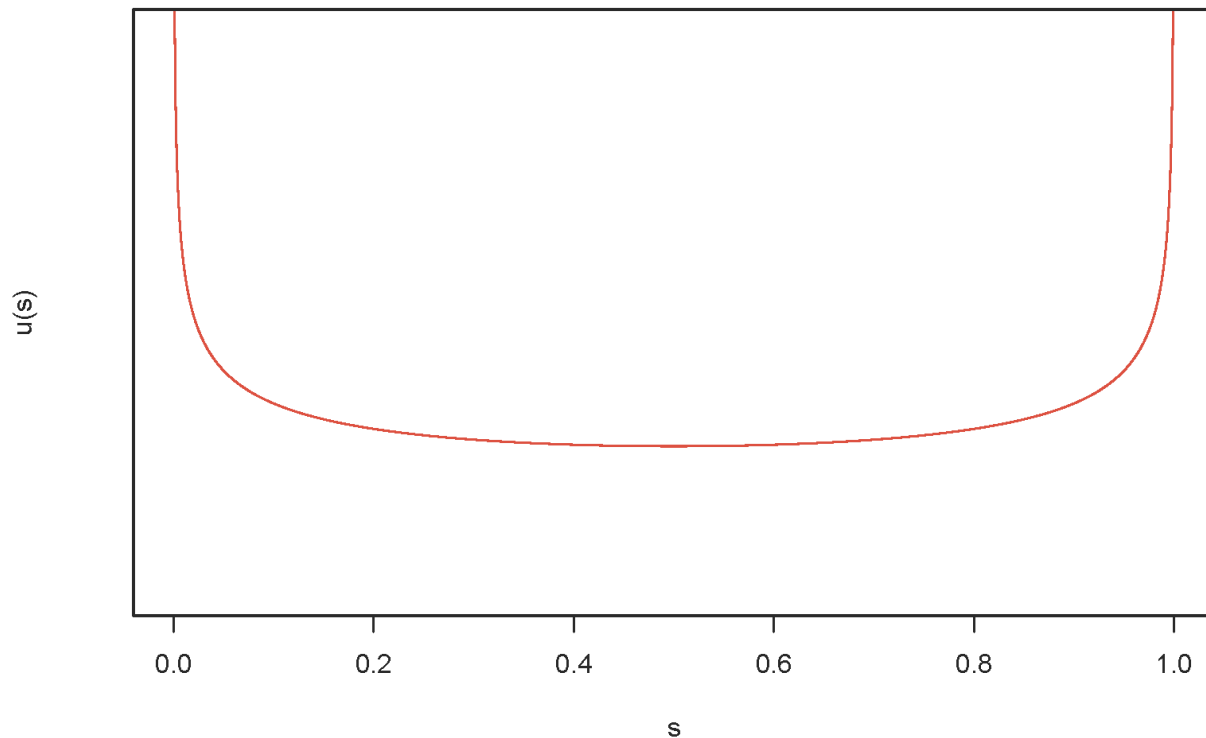
$$u(s) = \frac{A}{[s\,(T-s)]^{(1-\gamma)/2}}.$$

The normalizing factor A is given by

$$\int_0^T u(t)\, dt = X = A\,\sqrt{\pi}\,\left(\frac{T}{2}\right)^{\gamma}\,\frac{\Gamma\left(\frac{1+\gamma}{2}\right)}{\Gamma\left(1+\frac{\gamma}{2}\right)}$$

The optimal strategy is absolutely continuous with no block trades. However, it is singular at $t = 0$ and $t = T$.

Schematic of optimal strategy

The red line is a plot of the optimal strategy with $T = 1$ and $\gamma = 1/2$.

Example III: Linear market impact with linear decay

$G(\tau) = (1 - \rho\,\tau)^+$ and the optimal strategy $u(s)$ solves

$$\int_0^T u(s)\,(1 - \rho\,|t - s|)^+\,ds = \text{const.}$$

Let $N := \lfloor \rho\,T \rfloor$, the largest integer less than or equal to $\rho\,T$. Then

$$u(s) = A \sum_{i=0}^N \left(1 - \frac{i}{N+1}\right)\left\{\delta\left(s - \frac{i}{\rho}\right) + \delta\left(T - s - \frac{i}{\rho}\right)\right\}$$

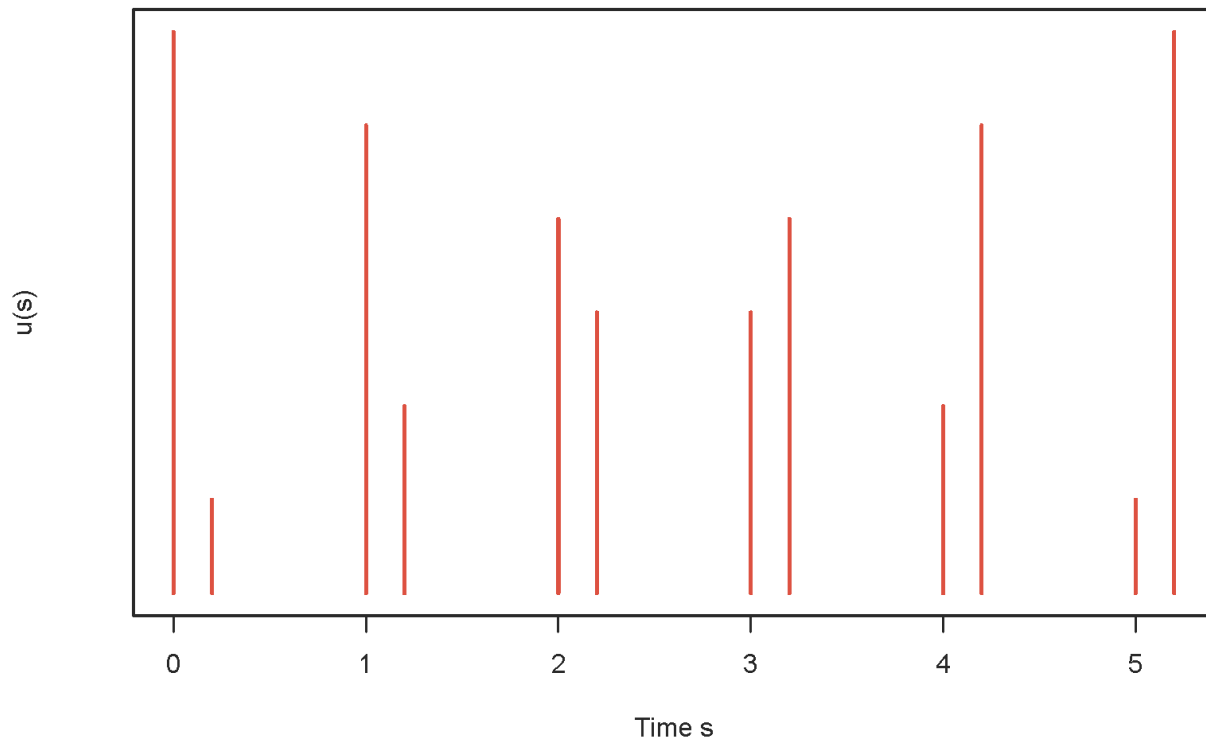The normalizing factor A is given by

$$\int_0^T u(t)\,dt = X = A \sum_{i=0}^N 2\left(1 - \frac{i}{N+1}\right) = A\,(2 + N)$$

The optimal strategy consists only of block trades with no trading between blocks.

Schematic of optimal strategy

Positions and relative sizes of the block trades in the optimal strategy with $\rho = 1$ and $T = 5.2$ (so $N = \lfloor \rho\,T \rfloor = 5$).

When is the optimal strategy bucket-shaped?

- For which nonincreasing convex decay kernels G does the optimal execution strategy have singular components only at $t = 0$ and $t = T$ with absolutely continuous trading on the interval $(0, T)$?
  - This question is answered in Corollary 1 of [Alfonsi and Schied][1].

Corollary

> Suppose that G is completely monotone with $G''(0+) := \lim_{t \downarrow 0} G''(t) < \infty$. Then the capacitary measure $\mu^\star$ has two Dirac components at $t = 0$ and $t = T$ and has a continuous Lebesgue density on $(0, T)$.

Completely monotone functions

Definition

> G is completely monotone if G belongs to $C^\infty((0, \infty))$ and $(-1)^n G^{(n)}$ is nonnegative in $(0, \infty)$ for $n \in \mathbb{N}$.
>

Theorem (Hausdorff–Bernstein–Widder)

> G is completely monotone if and only if it is the Laplace transform of a Borel probability measure λ on $\mathbb{R}+$:

$$G(t) = \int e^{-\rho\, t}\, \lambda(d\rho), \quad t \geq 0.$$

Examples of completely monotone functions

- Examples of completely monotone functions are

\begin{eqnarray} G_1(\tau)&=&\frac{1}{(1+\tau)^\gamma}\ G2(\tau)&=&\sum{i=1}^d\,\lambda_i\,e^{-\rhoi\,\tau} \text{ with } \sum{i=1}^d\,\lambda_i=1 \text{ and } \,\lambda_i,\,\rho_i\geq0\end{eqnarray}

- Roughly speaking then, if the decay kernel looks like a sum of exponential decay terms, the optimal strategy under linear transient impact consists of a block at the beginning, a block at the end and continuous trading in-between.

Examples of functions that are not completely monotone

- Examples of these include

\begin{eqnarray} G_3(\tau)&=&\frac{1}{1+\tau^2}\ G_4(\tau)&=&(1-\rho\,\tau)^+\end{eqnarray}

- $G_3(\cdot)$ is not monotone because $G''(\tau)$ changes sign.
- $G_4(\cdot)$ is not monotone because it is not smooth (it has a discontinuous first derivative).

Price-manipulation with nonlinear market impact

Proposition ([Gatheral, Schied and Slynko: Exponential resilience][8])

> Assuming the price process
> $$S_t = S_0 + \int_0^t f(v_s)\, G(t-s)\, ds + \text{ noise},$$
> consider a model with a general nonlinear instantaneous market impact function $f(\cdot)$ and a nonincreasing decay kernel $G(t)$ with $G(0) := \lim_{t \downarrow 0} G(t) < \infty$. Then, such a model admits price manipulation.

- The proof follows closely the proof for the special case of exponential decay of market impact that we presented in Lecture 8.

Nonlinear transient market impact

- We know that the reaction of market price to quantity is in general nonlinear.
    - Concave for small quantity and convex for large quantity.

- We also know that the market price reverts after completion of a meta-order (using VWAP say).

A nonlinear transient market impact model

Once again, the price process assumed in [Gatheral][4] is

(8)

$$S_t = S_0 + \int_0^t f(v_s)\, G(t-s)\, ds + \text{ noise}$$

- The instantaneous impact of a trade at time s is given by $f(v_s)$, some function of the rate of trading.
- A proportion $G(t-s)$ of this initial impact is still felt at time $t > s$.

Optimal strategies in the transient impact model

- In [Curato et al.][5], we show by numerical computation that in general the transient impact model admits transaction-triggered price manipulation.

- This undesirable property of the model can be mitigated in at least two reasonable ways that reflect actual properties of the market:
    - Including a bid-offer spread
    - Making the instantaneous market impact function h convex for large trading rates.

- We have not been able to find necessary or sufficient conditions to make the model well-behaved.
    - It follows that in general, there is no optimal strategy in this model.
    - However, in practice we can always restrict the set of admissible strategies.

The square-root model

Consider the following special case of [(8)](#) with $f(v) = \frac{3}{4}\sigma\sqrt{v/V}$ and $G(\tau) = 1/\sqrt{\tau}$:

(9)

$$S_t = S_0 + \frac{3}{4}\sigma \int_0^t \sqrt{\frac{v_s}{V}} \frac{ds}{\sqrt{t-s}} + \text{noise}$$

which we will call the *square-root process*.

It is easy to verify that under the square-root process, the expected cost of a VWAP execution is given by the square-root formula for market impact:

(10)

$$\frac{C}{X} = \sigma\sqrt{\frac{X}{V}}$$

- Of course, that doesn't mean that the square-root process is the true underlying process!

Costing an interval VWAP strategy with square-root decay

Suppose we partition the interval $\{0, T\}$ into m subintervals $\{0, t_1, t_2, \ldots, t_{m-1}, t_m = T\}$ and suppose further that our rate of trading $v_t = \dot{x}_t$ is constant over each subinterval.

The contribution to the cost from trading in the first subinterval is simply

$$\frac{x_1}{t_1} f(v_1) \int_0^{t_1} dt \int_0^t \frac{1}{\sqrt{t-s}} ds$$
$$= 2\frac{x_1}{t_1} f(v_1) \int_0^{t_1} dt \sqrt{t}$$
$$= \frac{4}{3} x_1 f(v_1) \sqrt{t_1}$$

where x_1 is the number of shares liquidated in the first interval.

Matrix elements

In general, the ith self-interaction term is given by

$$C_{ii} = \frac{x_i}{t_i - t_{i-1}} f(v_i) \int_{t_{i-1}}^{t_i} dt \int_{t_{i-1}}^t \frac{1}{\sqrt{t-s}} ds$$
$$= \frac{4}{3} x_i f(v_i) \sqrt{t_i - t_{i-1}}.$$

The contribution to the cost from trading in the jth subinterval related to trading in the ith subinterval with $\$i$

$$C_{ji} = \frac{x_j}{t_j - t_{j-1}} f(v_i) \int_{t_{j-1}}^{t_j} dt \int_{t_{i-1}}^{t_i} \frac{1}{\sqrt{t-s}} \, ds$$

$$= \frac{4}{3} x_j f(v_i) \times$$

$$\frac{\left(t_j - t_{i-1}\right)^{3/2} - \left(t_j - t_i\right)^{3/2} - \left(t_{j-1} - t_{i-1}\right)^{3/2} + \left(t_{j-1} - t_i\right)^{3/2}}{t_j - t_{j-1}}$$

where x_j is the number of shares liquidated in the jth sub-interval.

Cost of liquidation

The cost of liquidation is then given by
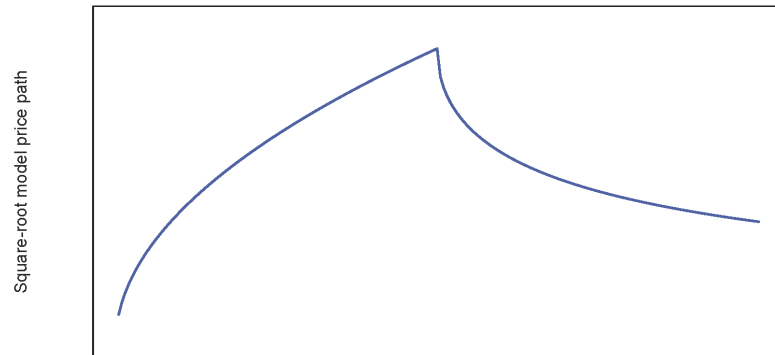
(11)

$$C = \sum_{j=1}^m C_{jj} + \sum_{i<j}^m C_{ji}$$

- For fixed m, we can then minimize C with respect to (for example) the x_i and the positions of the t_i.

The optimal strategy under the square-root process

- It turns out that because $f(\cdot)$ is concave, an optimal strategy does not exist in the square-root model.
  - It is possible to drive the expected cost of trading to zero by increasing the number of slices and decreasing the duration of each slice.
  - To be more realistic, $f(v)$ must be convex for large v and in this case, an optimal strategy does exist that involves trading in bursts, usually more than two.

Price path in the square-root model

Figure 2: The expected path of the market price during and after execution of a VWAP order in the square-root model.
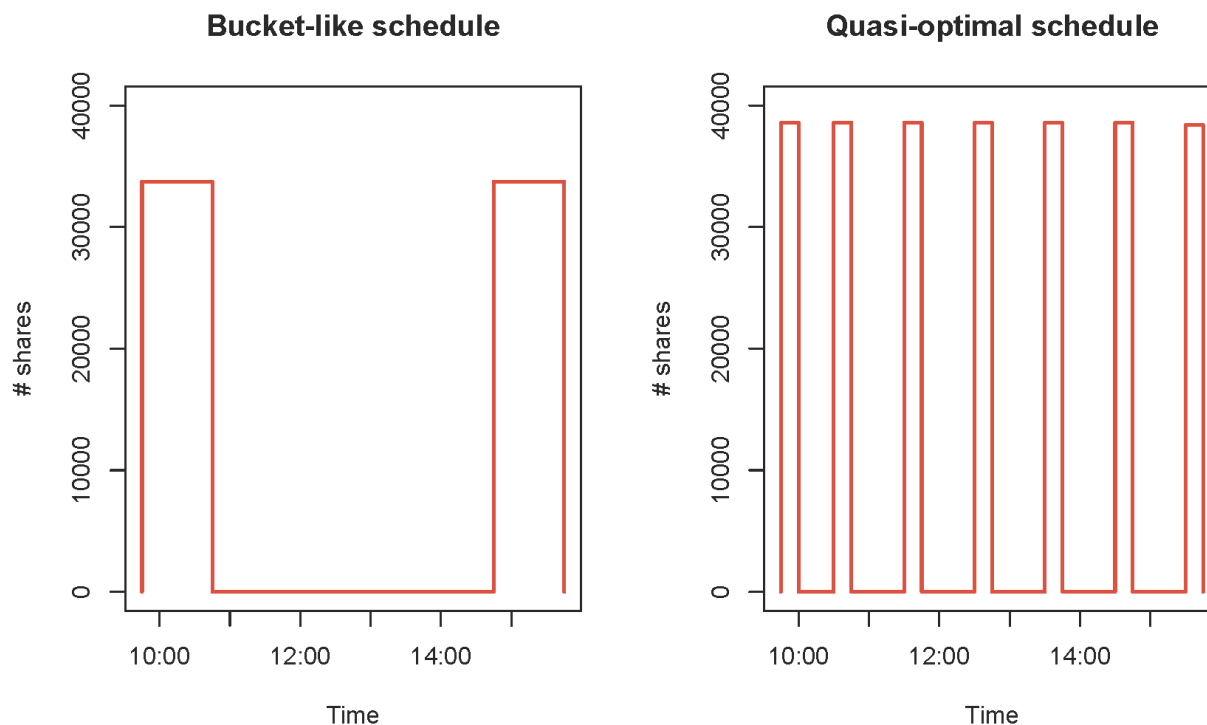

- The optimal strategy does not exist in this model.

Potential cost savings from optimal scheduling

- To estimate potential savings from optimal scheduling, assume that the square-root process (9) is correct and consider a one-day order to sell 540,000 shares of IBM.
    - Daily volatility is assumed to be 2% and daily volume to be 6 million shares.
    - We consider liquidation starting at 09:45 and ending at 15:45 with child orders lasting 15 minutes.

- Because we are not confident in the square-root model for high volume fractions (indeed we don't believe in square-root impact for high volume fractions), we constrain volume fraction to be no greater than 25%.

- We compare the costs of VWAP, a two-slice bucket-like strategy and a quasi-optimal strategy that consists of seven roughly equal slices.
    - The quasi-optimal strategy consists of bursts of trading separated by periods of non-trading.

Stock trading schedules

Bucket-like schedule

Quasi-optimal schedule

Comparison of results

In the square-root model, the cost of a VWAP execution is given exactly by the square-root formula:

$$\sigma\sqrt{\frac{Q}{V}} = 0.02 \times \sqrt{\frac{540}{6000}} = 0.02 \times 0.3 = 60 \text{ bp.}$$

Strategy	Cost	Saving
VWAP	60.0 bp	
Bucket-like	49.6 bp	17%
Quasi-optimal	40.8 bp	32%

Are these saving achievable?

- If market impact of meta orders is all temporary, and ours is the only metaorder trading, then these savings should be achievable.

- If there is a permanent component (roughly 2/3 according to various empirical papers), the saving will be somewhat less.

- Alternatively, if our metaorder is trading together with other metaorders in the same direction, the saving will be somewhat less.

Adaptive trading

- So far, we implicitly assumed that the market impact of a child order depends only on order flow.

- In reality, it can depend on many other variables such as the shape of the order book.

- The expected cost of the quasi-optimal strategy under square-root dynamics is sensitive only to the number of slices (and of course the maximum participation rate).
    - The expected cost of the strategy is rather insensitive to the exact spacing of the slices.

- We can thus send child orders whenever we expect less market impact.
    - For example, trade when others want to trade with us.
    - Trade when volatility is lower.
    - Trade when there is a lot of market volume.

Summary

- The optimal trading strategy depends on the model.
    - For Almgren-Chriss style models, if the price of risk is zero, the minimal cost strategy is VWAP.
    - In Alfonsi-Schied style models with resiliency that depends only on the current spread, the minimal cost strategy is to trade a block at inception, a block at completion and at a constant rate in between.
    - More generally, if market impact is transient, the optimal strategy involves bursts of trading; VWAP is never optimal.

- We can state sufficient conditions on the price process (8) for it to make sense:
    - There is no transaction-triggered price manipulation if the kernel $G(\cdot)$ is convex. It would seem reasonable to also require complete monotonicity.
    - If market impact is nonlinear, the kernel $G(\cdot)$ must be singular at $t = 0$.
    - A nonlinear instantaneous market impact function $h(\cdot)$ must be convex for sufficiently high trading rates for an optimal strategy to exist.

- Assuming square-root dynamics consistent with the square-root market impact formula, we showed that large cost savings are possible by optimizing the scheduling strategy.

References

1. **^** Capacitary measures for completely monotone kernels via singular control, Alexander Schied and Aurélien Alfonsi, *SIAM Journal on Control and Optimization* **51**(2) 1758–1780 (2013).
2. **^** Aurélien Alfonsi, Antje Fruth and Alexander Schied, Optimal execution strategies in limit order books with general shape functions, *Quantitative Finance* **10**(2) 143–157 (2010).
3. **^** Aurélien Alfonsi, Alexander Schied and Alla Slynko, Order book resilience, price manipulation, and the positive portfolio problem, *SIAM Journal on Financial Mathematics* **3**(1) 511–533 (2012).
4. **^** Michael Benzaquen and Jean-Philippe Bouchaud, A fractional reaction-diffusion description of supply and demand, *The European Physical Journal B* **91**(2) 23 (2018).
5. **^** Gianbiagio Curato, Jim Gatheral and Fabrizio Lillo, Optimal execution with nonlinear transient market impact, *Quantitative Finance* **17**(1) 41-54 (2017).
6. **^** Jim Gatheral, No-dynamic-arbitrage and market impact, *Quantitative Finance* **10**(7) 749–759 (2010).
7. **^** Jim Gatheral, Alexander Schied and Alla Slynko, Transient linear price impact and Fredholm integral equations, *Mathematical Finance* **22**, 445–474 (2012).
8. **^** Jim Gatheral, Alexander Schied and Alla Slynko, Exponential resilience and decay of market impact, *Econophysics of Order-Driven Markets*, Springer, 225–236 (2011).
9. **^** Anna Obizhaeva and Jiang Wang, Optimal trading strategy and supply/demand dynamics, *Journal of Financial Markets* **16**(1) 1–32 (2013).
10. **^** Silviu Predoiu, Gennady Shaikhet and Steven Shreve, Optimal execution in a general one-sided limit-order book, *SIAM Journal on Financial Mathematics* (2012).