Market Microstructure and Algorithmic Trading

University of Oxford

Fayçal Drissi
faycal.drissi@eng.ox.ac.uk
http://www.faycaldrissi.com

lecture notes

January 18, 2024

Optimal order routing of aggressive orders

The decision problem

- A “**marketable**” order is a buy (resp. sell) order at a price higher than the best ask (resp. lower than the best bid).
- Often, operators have to be split a large order over N available venues.

Optimal order routing of aggressive orders

The decision problem

- A "**marketable**" order is a buy (resp. sell) order at a price higher than the best ask (resp. lower than the best bid).
- Often, operators have to be split a large order over N available venues.

The model of the environment

- Let ($Q^\star$ at max price $P^\star$) be a marketable **buy** order.
  Let $Q_n(p)$ be the **visible quantity** that is **available** at **price** p in **trading venue** n.

Optimal order routing of aggressive orders

The decision problem

- A “**marketable**” order is a buy (resp. sell) order at a price higher than the best ask (resp. lower than the best bid).

- Often, operators have to be split a large order over N available venues.

The model of the environment

- Let ($Q^\star$ at max price $P^\star$) be a marketable **buy** order.
  Let $Q_n(p)$ be the **visible quantity** that is **available** at **price** p in **trading venue** n.

The optimisation problem : Choose $(p_1, \cdots, p_n)$ to minimize $\sum_{n=1}^{N} p_n \cdot Q_n(p_n)$ such that $Q^\star = \sum_{n=1}^{N} Q_n(p_n)$ with the constraint that $P^\star \geq p_n$ for all $n \in \{1, \cdots, N\}$.

The solution:

- Lagrangian: $Q_n(p_n) + p_n\, Q_n^{'}(p_n) = \lambda\, Q_n^{'}(p_n)$ for $n \in \{1, \cdots, N\}$.
- Assume the linear form $Q_n(p) = q_n + c_n \cdot p$.

$$\Longrightarrow \; (q_n + c_n \cdot p_n^\star) + p_n^\star\, c_n = \lambda\, c_n$$
$$\Longrightarrow \; p_n^\star = \frac{\lambda}{2} - \frac{q_n}{2\, c_n}\,.$$

- Inject λ in the constraint $Q^\star = \sum_{n=1}^N Q_n(p_n)$:

$$\Longrightarrow \; Q^\star = \sum_{n=1}^N \{q_n + c_n \cdot p_n^\star\} = \sum_{n=1}^N q_n/2 + c\,\lambda/2, \quad c = \sum_n c_n$$

- Finally

$$\boxed{p_n^\star = \frac{Q^\star}{c} - \frac{q_n}{2\, c_n}\left(1 + \frac{c_n}{q_n} \cdot \frac{\overline{q}}{\overline{c}}\right)} \quad \overline{c} = \frac{1}{N}\sum_n c, \;\; \overline{q} = \frac{1}{N}\sum_n q$$

Some issues with our (simple) model:

- There is **latency** in high frequency markets.
- There are **cancellations** in limit order books.
- The shape of the book **hides** the truth sometimes: iceberg orders + hidden orders.
- etc..

The decision problem

- An operator splits a large LO over N venues: there is structurally **uncertainty** on limit orders splitting; waiting on a bad queue generates **opportunity costs**. Formulate an optimal trading problem ?

The decision problem

- An operator splits a large LO over N venues: there is structurally **uncertainty** on limit orders splitting; waiting on a bad queue generates **opportunity costs**. Formulate an optimal trading problem ?

The model:

- The best queue in each venue has size Q_n.
- Each queue is **consumed** according to a Poisson P^n_t with intensity λ_n.

The objective: find LO quantities $(q_1, \cdots, q_N)$ to minimize, **on average**, the time $t^\star$ to execute the quantity $Q^\star = \sum_n q_n$.

Solution:

- After our LOs, venue n has new queue quantity $Q_n + q_n$.

- Queue is consumed in t_n:

$$\int_0^{t^n} dP_t^n = q_n + Q_n \implies \mathbb{E}\left[P_{t^n}^n\right] = t_n \, \lambda_n = q_n + Q_n .$$

- We minimize the maximum of all t^n, so $t^\star = t^n$ for all $n \in \{1, \cdots, N\}$. So:

$$t^\star = t_n = Q^\star / \sum_n \lambda_n + \sum_n Q_n / \sum_n \lambda_n \implies \boxed{q_n^\star = \rho_n \frac{Q^\star}{N} + \left(\rho_n \overline{Q} - Q\right)}$$

$$\text{where} \quad \rho_n = \lambda_n / \overline{\lambda}, \quad \overline{\lambda} = \frac{1}{N} \sum_n \lambda_n, \quad \overline{Q} = \frac{1}{N} \sum_n Q_n \,.$$

Optimal execution: Large operators need optimal execution models. The first approach is in (Bertsimas and Lo 1998).

Two seminal papers: (Almgren and Chriss 1999) and (Almgren and Chriss 2001). Considered to be the pioneers in optimal execution.

- A model taking into account both the expected cost of execution and the risk that the price moves.
- Large orders are split into smaller ones, that are executed progressively over a given time window.
- A trader **executing fast** pays high execution costs: bid-ask spread and limited available liquidity at each price in the order book.
- **Slow execution** exposes to possible adverse price fluctuations.
- There is a **trade-off** between execution costs and price risk.

- An agent with a single stock.
  At time $t = 0$, the position (in number of shares) is q_0.

- The **objective**: **unwind the position by time** T. Find the optimal **schedule**.

- The **model**: We assume a regular temporal grid

$$t_0 = 0 < \cdots < t_n = n\,\Delta t < \cdots < t_N = N\,\Delta t = T.$$

  At the start of each $[t_n, t_{n+1}]$, the agent sends an MO of size $\nu_{n+1}\,\Delta t$.

  If $\nu_{n+1} \leq 0$ then the agent sells shares, if $\nu_{n+1} \geq 0$, then the agent buys shares.

- The dynamics for the **inventory** (number of shares in the agent's portfolio)

$$\boxed{q_{n+1} = q_n + \nu_{n+1}\,\Delta t}$$

- The **mid-price** follows a Brownian motion

$$\boxed{S_{n+1} = S_n + \underbrace{\sigma\sqrt{\Delta t}\,\epsilon_{n+1}}_{\text{market risk}} + \underbrace{k\,\nu_{n+1}\,\Delta t}_{\text{linear perm. impact}}}$$

  - ϵ_n are i.i.d $\mathcal{N}(0,1)$ variables
  - $\sigma > 0$ is the arithmetic volatility
  - $k > 0$ scales the magnitude of the **linear permanent impact**. (more on this later ...)

Permanent price impact: large MOs leave a long-term effect on the midprice.

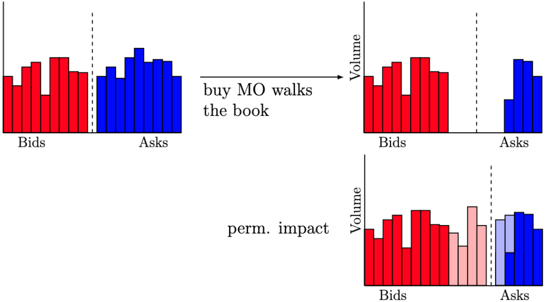


Figure 7: In the first two panels, an MO walks the book so the next midprice exhibits the temporary price impact. Immediately after the MO, market participants replenish the LOB. The difference between the midprice in the last panel and that of the first panel is the permanent impact.

- **Execution costs** (temporary impact):

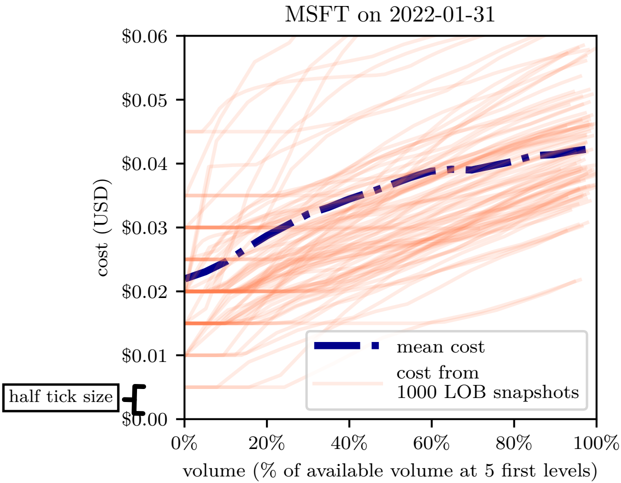
  - We introduce the deterministic **market volume** V_{n+1}, which is the volume traded by other agents throughout $[t_n, t_{n+1}]$.

  - The price $\tilde{S}_{n+1}$ obtained for each share in $[t_n, t_{n+1}]$ depends on the quantity $\nu_{n+1}\,\Delta t$ and on the market volume V_{n+1}. We assume the linear form

    $$\boxed{\tilde{S}_{n+1} = S_n + \eta\,\nu_{n+1}/V_{n+1}}.$$

    $\eta > 0$ so the agent buys (sells) at prices higher (lower) than the mid-price S_n.

  - ν_n/V_n is the **participation rate** and it is very important to estimate η. e.g., the costs are 1 ticks (or bid-ask spread) spread per 5% of participation rate.



Execution costs defined as a function of participation rate for multiple snapshots of the LOB of MSFT quoted on Nasdaq. The total trade volume is approximated by the total available liquidity.

Market volume

The market volume V_{n+1} corresponds to the total volume of the market over a time slice $[t_n, t_{n+1}]$. In practice, it is difficult to consider this deterministic ...

However, market activity depends on the time of day. On average, it is deterministic and has a characteristic U-shape.

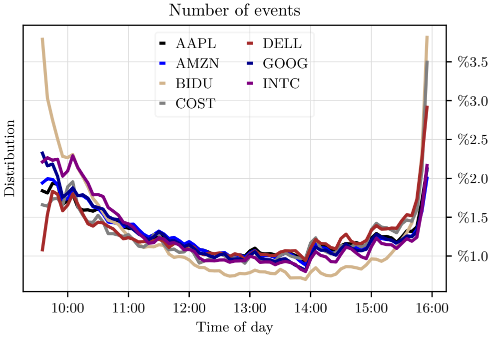


Figure: Distribution of the activity throughout the trading day, measured in portion of LOB events, averaged through trading days between October and December 2022. Source: (Cartea et al. 2023).

- The amount paid (received) for $\nu_{n+1}\,\Delta t$ shares bought (sold) between t_n and t_{n+1} is
$$\nu_{n+1}\,\tilde{S}_{n+1}\,\Delta t = \nu_{n+1}\,(S_n + \eta\,\nu_{n+1}/V_{n+1})\,\Delta t.$$
So the dynamics of the cash account X are
$$\boxed{X_{n+1} = X_n - \nu_{n+1}\,S_n\,\Delta t - \eta\,\frac{\nu_{n+1}^2}{V_{n+1}}\,\Delta t}$$
- The execution costs paid are relative to S_n over $[t_n, t_{n+1}]$, so no price risk within each slice ...

Find a **liquidation strategy** $v = (\nu_1, \ldots, \nu_n)$ maximizing the mean-variance objective function

$$\boxed{\mathbb{E}[X_N] - \frac{\gamma}{2}\mathbb{V}[X_N]}.$$

We focus on deterministic admissible strategies (more on this later ...)

$$(\nu_n)_n \in \mathcal{A}^d = \left\{ (\nu_1, \ldots, \nu_N) \in \mathbb{R}^n, \sum_{n=0}^{N-1} \nu_{n+1}\,\Delta t = -q_0 \right\}$$

To solve the problem, we compute the value of the terminal wealth X_N :

$$
\begin{aligned}
X_N &= X_0 - \sum_{n=0}^{N-1} (q_{n+1} - q_n)\, S_n - \eta \sum_{n=0}^{N-1} \frac{\nu_{n+1}^2}{V_{n+1}} \Delta t \\
&= X_0 - \sum_{n=0}^{N-1} q_{n+1} \left(S_{n+1} - \sigma\sqrt{\Delta t}\, \epsilon_{n+1} - k\, \nu_{n+1} \Delta t \right) + \sum_{n=0}^{N-1} q_n\, S_n - \eta \sum_{n=0}^{N-1} \frac{\nu_{n+1}^2}{V_{n+1}} \Delta t \\
&= X_0 + q_0\, S_0 + \sigma \sqrt{\Delta t} \sum_{n=0}^{N-1} q_{n+1}\, \epsilon_{n+1} + \underbrace{k \sum_{n=0}^{N-1} q_{n+1}\, \nu_{n+1}\, \Delta t}_{\star} - \eta \sum_{n=0}^{N-1} \frac{\nu_{n+1}^2}{V_{n+1}} \Delta t.
\end{aligned}
$$

Observe that:

$$
\begin{aligned}
\star &= k \sum_{n=0}^{N-1} \left(\frac{q_{n+1} + q_n}{2} + \frac{q_{n+1} - q_n}{2} \right) (q_{n+1} - q_n) \\
&= \frac{k}{2} \sum_{n=0}^{N-1} \left(q_{n+1}^2 - q_n^2 \right) + \frac{k}{2} \sum_{n=0}^{N-1} (q_{n+1} - q_n)^2 \\
&= -\frac{k}{2} q_0^2 + \frac{k}{2} \sum_{n=0}^{N-1} \nu_{n+1}^2\, \Delta t^2\,.
\end{aligned}
$$

The terminal wealth is

$$X_N = X_0 + q_0\,S_0 - \frac{k}{2}q_0^2 + \sigma\sqrt{\Delta t}\sum_{n=0}^{N-1} q_{n+1}\,\epsilon_{n+1} + \frac{k}{2}\sum_{n=0}^{N-1}\nu_{n+1}^2\,\Delta t^2 - \eta\sum_{n=0}^{N-1}\frac{\nu_{n+1}^2}{V_{n+1}}\Delta t$$

$$= X_0 + q_0\,S_0 - \frac{k}{2}q_0^2 + \sigma\sqrt{\Delta t}\sum_{n=0}^{N-1} q_{n+1}\,\epsilon_{n+1} - \sum_{n=0}^{N-1}\nu_{n+1}^2\left(\frac{\eta - \frac{k}{2}\,V_{n+1}\,\Delta t}{V_{n+1}}\right)\Delta t\,.$$

To obtain analytic formulae, **assume** a flat volume curve $V_n = V, \forall n$.

Next, either we neglect the term in Δt^2 (define $\tilde{\eta} = \eta$) or we assume $\eta >> \frac{k}{2}\,V\,\Delta t$ (define $\tilde{\eta} = \eta - \frac{k}{2}\,V\,\Delta t$)

The controls are deterministic, so X_N is normally distributed with mean

$$\boxed{\mathbb{E}\left[X_N\right] = X_0 + q_0\,S_0 - \frac{k}{2}q_0^2 - \tilde{\eta}\sum_{n=0}^{N-1}\frac{\nu_{n+1}^2}{V}\Delta t}$$

and variance $$\boxed{\mathbb{V}\left[X_N\right] = \sigma^2\,\Delta t\sum_{n=0}^{N-1} q_{n+1}^2}$$

The problem **reduces** to minimising the following functional over $\mathcal{A}^d$

$$\tilde\eta \sum_{n=0}^{N-1} \frac{\nu_{n+1}^2}{V} \Delta t + \frac{\gamma}{2}\sigma^2 \Delta t \sum_{n=0}^{N-1} q_{n+1}^2 \,,$$

which is equivalent to minimising J over $\mathcal{C}_d = \{q = (q_0, \dots, q_N)\,, q_0 = q_0, q_N = 0\}$

$$\boxed{J : q \in \mathbb{R}^{N+1} \mapsto \tilde\eta \sum_{n=0}^{N-1} \frac{(q_{n+1} - q_n)^2}{V\,\Delta t} + \frac{\gamma}{2}\sigma^2 \Delta t \sum_{n=0}^{N-1} q_{n+1}^2}$$

The Legendre-Fenchel transform of $g : x \mapsto \tilde\eta \frac{x^2}{V\,\Delta t}$ is easily found (FOC) to be

$$g^\star : p \mapsto \sup_x p\,x - \tilde\eta \frac{x^2}{V\,\Delta t} = \frac{V\,\Delta t}{4\,\tilde\eta} p^2.$$

The **optimal trading curve** $q^\star$ is characterized by the Hamiltonian system

$$\begin{cases} p_{n+1} &= p_n + \gamma\,\sigma^2\,\Delta t\, q_{n+1}^\star\,, \quad 0 \le n < N-1 \\ q_{n+1}^\star &= q_n^\star + \frac{V}{2\,\tilde\eta}\,\Delta t\, p_n\,, \qquad 0 \le n < N \end{cases}$$

The **optimal inventory** $q^\star$ to hold is the solution of the second-order recursive equation

$$q^\star_{n+2} - \left(2 + \frac{\gamma\,\sigma^2\,V}{2\,\tilde{\eta}}\Delta t^2\right) q^\star_{n+1} + q^\star_n = 0\,,$$

with boundary conditions

$$q^\star_0 = q_0 \quad \text{and} \quad q^\star_N = 0.$$

Solving the equation gives

$$\boxed{q^\star_n = q_0\,\frac{\sinh\left(\alpha\,(T - t_n)\right)}{\sinh(\alpha T)}}$$

where α solves

$$2\,\cosh\left(\alpha\,\Delta t\right) = \frac{\gamma\,\sigma^2\,V}{2\,\tilde{\eta}}\Delta t^2\,.$$

- Consider an asset with volatility $1\$ \cdot \text{day}^{-1/2}$ (approx. 32% annualized vol) with $S_0 = 100$.

- Assume the market trades $V = 4,000,000$ shares per day, and assume $\eta = 0.1\$ \cdot \text{share}^{-1}$.

- The initial inventory to liquidate is $q_0 = 200,000$ corresponding to 5% participation rate.

Effect of model parameters: η, σ, γ, and V ?

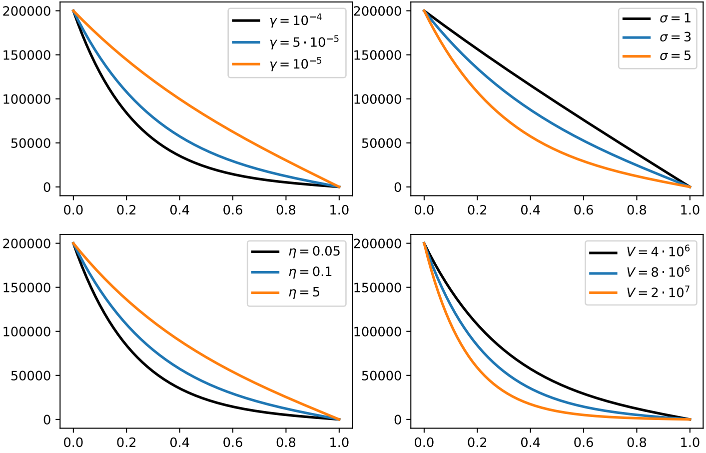


Figure: Optimal trading curves.

Execution costs

Setting temporary market impact parameter η in practice ((Almgren and Chriss 2001)):

Traders suppose that the additional cost incurred per share when trading a given volume is **proportional** to the **participation rate**.

For each percent of participation rate, a cost corresponding to some % of the **the bid-ask spread** is incurred.

Objective functions

- There is a-priori no reason to believe agents have a mean-variance risk profile.
- In practice, the choice of **utility function** is difficult. In the simple case of mean-variance, choosing the **risk aversion parameter** γ is complex.
- If the agent is a Brokerage firm or a cash trader executing for clients, which value of γ to use ?
  - The client must, somehow, indicate the value of γ.
  - One interpretation: the risk aversion parameter encodes the **urgency** of trading.
  - Or the broker needs to tailor γ to the size of the trade / participation rate ($\gamma = \overline{\gamma}/|q_0\, S_0|$)

Extensions

The framework is flexible:

- The criterion can be challenged: PoV, VWAP, TWAP, TC, etc.
- The variance term has a strong influence.
  More interesting risk measures: CVaR, etc ..
- In practice: traders use **participation constraints** to their trading flow.

Some problems / extensions:

- We don't deal with orderbook dynamics $\implies \eta$ and k are stochastic.
- The market impact model is far from being realistic. The market price is partially resilient (it does not vanish instantaneously).

- An agent holds an initial position q_0 at time $t = 0$ that they wish to unwind over a time window $[0, T]$.

- The trader's **inventory** over $[0, T]$ is modelled by the process $(q_t)_{t\in[0,T]}$ with dynamics

$$\boxed{dq_t = \nu_t \, dt}$$

- At time t, ν_t is the **trading velocity**, or **trading speed**, or **instantaneous trading volume**.

- The set of **admissible** strategies (controls) is[1]

$$\boxed{\mathcal{A} = \left\{ (\nu_t)_{t\in[0,T]} \in \mathbb{H}^0 \left(\mathbb{R}, (\mathcal{F}_t)_t\right), \; \int_0^T \nu_t \, dt = -q_0, \; \int_0^T |\nu_t| \, dt \in L^\infty(\Omega) \right\}}$$

[1] $L^\infty(\Omega)$ is the set of bounded processes. $\mathbb{H}^0\left(\mathbb{R}, (\mathcal{F}_t)_t\right)$ is the set of real-valued progressively measurable processes.

- The **mid-price** of the asset has a **linear permanent impact** component and is modelled by the (controlled) process $(S_t)_{t\in[0,T]}$ with dynamics

$$\boxed{dS_t = k\,\nu_t\,dt + \sigma dW_t}$$

- The **market volume** process is $(V_t)_{t\in[0,T]}$. We assume it is deterministic, continuous, positive, and bounded.

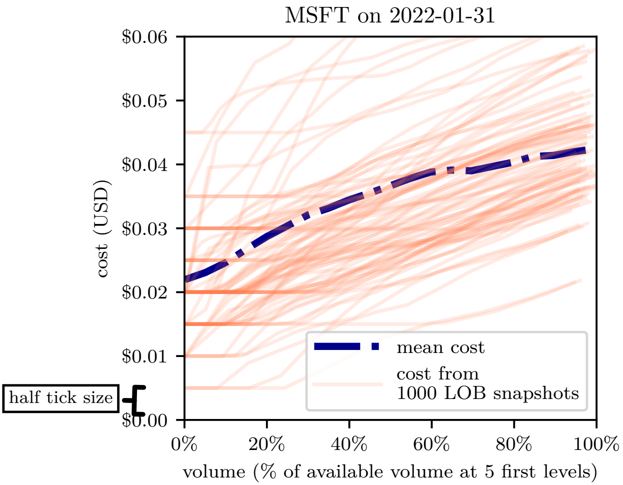
- **Execution costs**: the price obtained for each asset at time t is

$$\tilde{S}_t = S_t + g\left(\nu_t/V_t\right) = S_t + V_t\,L\left(\nu_t/V_t\right).$$

We assume $L : \mathbb{R} \mapsto \mathbb{R}$ verifies

- *No fixed cost*: $L(0) = 0$.
- *There is always a cost to trade when selling and buying*: L is strictly convex, increasing in $\mathbb{R}^+$ and decreasing in $\mathbb{R}^-$
- *L grows faster than a linear cost*: L is asymptotically super-linear, i.e., $\lim_{|x|\to\infty} \frac{L(x)}{|x|} = \infty$.

In some examples, L is a strictly convex power function: $L(x) = \eta\,|x|^{1+\rho}$ where $\rho > 0$.
Or $L(x) = \kappa\,|x| + \eta\,|x|^{1+\rho}$ where $\kappa|x|$ is the bid-ask component of the execution costs. The initial AC framework considers $L(x) = \eta\,x^2$.



The **optimisation** problem: find an **optimal strategy** $(\nu_t)_{t\in[0,T]} \in \mathcal{A}$ to liquidate the position.

- The agent adopts the classical CARA (Constant Absolute Risk Aversion) utility function

$$\sup_{\nu} \mathbb{E}\left[-\exp\left(-\gamma X_T\right)\right]$$

where $\gamma > 0$ is the risk aversion coefficient.

- We first restrict admissible controls ν to **deterministic** ones.

$$\boxed{\mathcal{A}^{\text{det}} = \{\nu \in \mathcal{A}, \forall t \in [0,T], \nu_t is \mathcal{F}_0 - \text{ measurable}\}}$$

The solution

For a strategy $\nu \in \mathcal{A}^{\text{det}}$, the terminal cash is

$$\begin{aligned} X_T =& X_0 - \int_0^T \nu_t\, S_t\, dt - \int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\, dt \\ =& X_0 + q_0\, S_0 + \int_0^T k\, \nu_t\, q_t\, dt + \sigma \int_0^T q_t\, dW_t - \int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\, dt \\ =& X_0 + q_0\, S_0 - \frac{k}{2} q_0^2 + \sigma \int_0^T q_t\, dW_t - \int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\, dt. \end{aligned}$$

$\nu \in \mathcal{A}^{\text{det}}$ so the terminal cash is normally distributed with mean and variance

$$\begin{cases} \mathbb{E}\left[X_T\right] &= \underbrace{X_0 + q_0\, S_0}_{\text{MtM}} - \underbrace{\frac{k}{2} q_0^2}_{\text{perm. impact.}} - \underbrace{\int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\, dt}_{\text{execution costs}} \\ \mathbb{V}\left[X_T\right] &= \sigma^2 \int_0^T q_t^2\, dt \end{cases}$$

The **permanent impact** term does not depend on the strategy ν: the permanent impact costs cannot be avoided.

Strategy proposed in (Bertsimas and Lo 1998) (first paper on optimal execution): maximise only $\mathbb{E}[X_T]$, i.e., minimise execution costs.

Exercise: Show that because L is assumed to be convex and super-linear, then execution costs are minimal when ν_t is proportional to V_t.

The variance $\mathbb{V}[X_T] = \sigma^2 \int_0^T q_t^2 \, dt$ is increasing in σ and in inventory. It is minimised when the agent trades quickly.

Use the Laplace transform of a Gaussian variable to write

$$\mathbb{E}\left[-\exp\left(-\gamma X_T\right)\right] = -\exp\left(-\gamma\,\mathbb{E}\left[X_T\right] + \frac{1}{2}\gamma^2\,\mathbb{V}\left[X_T\right]\right)$$

$$= -\exp\left(-\gamma\left(X_0 + q_0 S_0 - \frac{k}{2}q_0^2\right)\right)\exp\left(\gamma\left(\int_0^T V_t\,L\left(\frac{\nu_t}{V_t}\right)dt + \frac{\gamma}{2}\sigma^2\int_0^T q_t^2\,dt\right)\right)$$

The problem reduces to minimising $\int_0^T V_t\,L\left(\frac{\nu_t}{V_t}\right)dt + \frac{\gamma}{2}\sigma^2\int_0^T q_t^2\,dt$.

Equivalently, we minimise J over the set of absolutely cont. func. $W^{1,1}$ with constraints $q(0) = q_0$ and $q(T) = 0$.

$$\boxed{J(q) = \int_0^T\left(V_t\,L\left(\frac{q'(t)}{V_t}\right) + \frac{\gamma}{2}\sigma^2\int_0^T q_t^2\right)dt}$$

(Guéant 2016) shows that there exists a unique minimiser $q^\star$ for general impact functions L (out of scope).

We characterise this minimiser $q^\star$ with the Hamiltonian system

$$\begin{cases} p'(t) & = \gamma\,\sigma^2\, q^\star(t) \\ q^{\star\prime}(t) & = V_t\, H'(p(t)) \\ q^\star(0) & = q_0 \\ q^\star(T) & = 0, \end{cases} \implies p''(t) = \gamma\,\sigma^2\, V_t\, H'(p(t))$$

where H is the Legendre-Fenchel transform of L (L is strictly convex, so H is $\mathcal{C}^1$). Moreover, V is continuous so $q^\star$ is also $\mathcal{C}^1$.

When **costs are quadratic** and V_t is constant: $L(x) = \eta\, x^2$ so $H(p) = \frac{p^2}{4\,\eta}$ and

$$\begin{cases} p'(t) &= \gamma\,\sigma^2\, q^\star(t) \\ {q^\star}'(t) &= V_t\,\frac{p(t)}{2\,\eta} \\ q^\star(0) &= q_0 \\ q^\star(T) &= 0 \end{cases} \implies {q^\star}''(t) = V\,\frac{\gamma\,\sigma^2}{2\,\eta}\, q^\star(t)\,,$$

The **solution** is

$$\boxed{q^\star(t) = q_0 \frac{\sinh\left((T-t)\sqrt{\frac{\gamma\, V\,\sigma^2}{2\,\eta}}\right)}{\sinh\left(T\sqrt{V\,\frac{\gamma\,\sigma^2}{2\,\eta}}\right)} \implies \nu^\star(t) = -q_0\sqrt{\frac{\gamma\, V\,\sigma^2}{2\,\eta}}\,\frac{\cosh\left((T-t)\sqrt{\frac{\gamma\, V\,\sigma^2}{2\,\eta}}\right)}{\sinh\left(T\sqrt{V\,\frac{\gamma\,\sigma^2}{2\,\eta}}\right)}}$$

Recall the discrete-time counterpart:

$$\begin{cases} q^\star_n &= q_0\,\frac{\sinh(\alpha(T-t_n))}{\sinh(\alpha T)} \\ 2\,\cosh(\alpha\,\Delta t) &= \frac{\gamma\,\sigma^2\, V}{2\,\tilde{\eta}}\Delta t^2 \end{cases} [\Delta t \to 0]\text{Taylor exp.}\longrightarrow \begin{cases} q^\star_n &= q_0\,\frac{\sinh(\alpha(T-t_n))}{\sinh(\alpha T)} \\ \alpha &= \sqrt{\frac{\gamma\,\sigma^2\, V}{2\,\tilde{\eta}}} \end{cases}$$

Role of parameters

- **liquidity parameters** η and V: they are **scaling** factors; the larger η, the more the agent pays costs, the lower V, the more the agent pays costs, so the liquidation process is slower:

$$\frac{dq^\star}{d\frac{\eta}{V}}/q_0 \geq 0.$$

- **volatility** σ: importance of price risk in the performance; the larger σ, the faster the liquidation to reduce exposure to risk:

$$\frac{dq^\star}{d\sigma}/q_0 \leq 0.$$

- **risk aversion** γ: balances the tradeoff between costs and price risk; the larger γ, the more sensitive the agent to price risk. Thus, high γ means fast execution:

$$\frac{dq^\star}{d\gamma}/q_0 \leq 0.$$

  Note that no risk aversion leads to TWAP:

$$\lim_{\gamma\to 0} q^\star(t) = q_0\left(1 - \frac{t}{T}\right)$$

Stochastic and deterministic strategies

Theorem: $\sup_{\nu\in\mathcal{A}} \mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right] = \sup_{\nu\in\mathcal{A}^{\text{det}}} \mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right].$

proof: We need to prove $\sup_{\nu\in\mathcal{A}} \mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right] \leq \sup_{\nu\in\mathcal{A}^{\text{det}}} \mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right].$
Let $\nu\in\mathcal{A}$ and write

$$
\begin{aligned}
\mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right] &= -\exp\left(-\gamma\,\left(X_0 + q_0\, S_0 - \frac{k}{2} q_0^2\right)\right) \\
&\quad \mathbb{E}\left[\exp\left(\gamma\,\int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\,dt\right)\exp\left(-\gamma\,\sigma\,\int_0^T q_t\,dW_t\right)\right] \\
&= -\exp\left(-\gamma\,\left(X_0 + q_0\, S_0 - \frac{k}{2} q_0^2\right)\right) \\
&\quad \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\gamma\,\int_0^T V_t\, L\left(\frac{\nu_t}{V_t}\right)\,dt\right)\exp\left(\frac{\gamma^2}{2}\,\sigma^2\,\int_0^T q_t^2\,dt\right)\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[-\exp\left(-\gamma\,\left(X_0 + q_0\, S_0 - \frac{k}{2} q_0^2 - J\left(q\right)\right)\right)\right] \leq \sup_{\nu\in\mathcal{A}^{\text{det}}} \mathbb{E}\left[-\exp\left(-\gamma\, X_T\right)\right]
\end{aligned}
$$

where $\mathbb{Q}$ is defined by the Radon-Nikodym derivative

$$
\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\gamma\,\sigma\,\int_0^T q_t\,dW_t - \frac{\gamma^2}{2}\,\sigma^2\,\int_0^T q_t^2\,dt\right)
$$

Take the supremum on the left-hand side to get the result.

Deterministic strategies are computed at the start of the liquidation, and stays the same for whichever price path.

In practice, execution algorithms are in **two layers**: the first is strategic and defines the optimal trading curve. The second is **tactical** and tracks the optimal trading curve with different type of orders, different trading venues, etc.

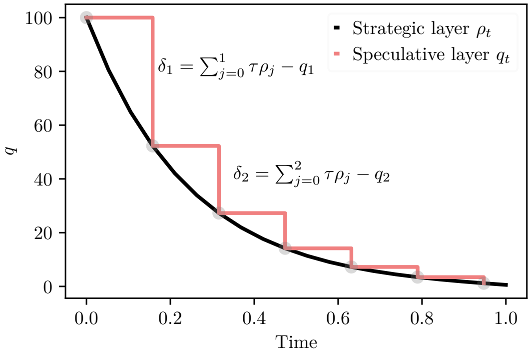


Figure: Execution algorithms in two layers.

Discussion based on the work of (Gatheral 2010).

- Assume the permanent impact is $\kappa(\nu)$ and for simplicity, assume zero execution costs. The dynamics are

$$\begin{cases} dq_t &= \nu_t \, dt \\ dS_t &= \sigma \, dW_t + \kappa(\nu_t) \, dt \\ dX_t &= -\nu_t \, S_t \, dt . \end{cases}$$

- **Definition**: there is dynamic arbitrage if $\exists t_1 < t_2$ and a **roundtrip** strategy ν such that

$$\begin{cases} \int_{t_1}^{t_2} \nu_t \, dt &= 0 \\ \mathbb{E}\left[X_{t_2} \middle| \mathcal{F}_{t_1} \right] &> X_{t_1} . \end{cases}$$

Theorem: $\kappa(\cdot)$ linear is the only possible choice to guarantee absence of dynamic arbitrage.

proof:

- consider the roundtrip strategy $\nu_t = \begin{cases} a & \text{if } t \in \left[t_1, \frac{a\,t_1 + b\,t_2}{a+b}\right] \\ -b & \text{if } t \in \left[\frac{a\,t_1 + b\,t_2}{a+b}, t_2\right] \end{cases}.$

  Then we can show that

$$\mathbb{E}\left[X_{t_2} \middle| \mathcal{F}_{t_1}\right] = X_{t_1} + \frac{a\,b}{2\,(a+b)^2}\,(t_2 - t_1)^2\,(b\,\kappa\,(a) + a\,\kappa\,(-b))\,.$$

$$\begin{aligned}
\text{No dyn. arb.} \implies & \mathbb{E}\left[X_{t_2} \middle| \mathcal{F}_{t_1}\right] \le X_{t_1}\,, \quad \forall a, b \\
\implies & \begin{cases} b\,\kappa\,(a) + a\,\kappa\,(-b) \le 0 \\ a\,\kappa\,(-b) + b\,\kappa\,(a) \ge 0 \quad \text{(replace } (a, b) \text{ by } (-b, -a)\text{)} \end{cases} \\
\implies & b\,\kappa\,(a) = -a\,\kappa\,(-b)\,, \quad \forall a, b \\
\implies & \begin{cases} \kappa\,(a) = -\kappa\,(-a)\,,\ \forall a & \text{for } b = a \\ \kappa\,(a) = -a\,\text{sign}\,(a)\ \kappa\,(-\text{sign}\,(a)) = a\,\kappa\,(1) & \text{for } b = \text{sign}\,(a) \ne 0 \end{cases}
\end{aligned}$$

proof:

- We need to show $\kappa(0) = 0$.

Choose

$$\nu_t = \begin{cases} \kappa\,(0) & \text{if } t \in \left[t_1, t_1 + \frac{t_2 - t_1}{3}\right] \\ 0 & \text{if } t \in \left[t_1 + \frac{t_2 - t_1}{3}, t_1 + 2\frac{t_2 - t_1}{3}\right] \\ -\kappa\,(0) & \text{if } t \in \left[t_1 + 2\frac{t_2 - t_1}{3}, t_2\right] \end{cases}$$

and obtain

$$\mathbb{E}\left[X_{t_2} \middle| \mathcal{F}_{t_1}\right] = X_{t_1} + \kappa\,(0)^2\,\frac{(t_2 - t_1)^2}{9}.$$

No dynamic arbitrage implies $\kappa\,(0) = 0$.

- Conversely, if $\kappa\,(\nu) = k\,\nu$ with $k \geq 0$ then there is no dynamic arbitrage.

Almgren, Robert and Neil Chriss (1999). “Value under liquidation”. In: **Risk** 12.12, pp. 61–63.

— (2001). “Optimal execution of portfolio transactions”. In: **Journal of Risk** 3, pp. 5–40.

Bertsimas, Dimitris and Andrew W Lo (1998). “Optimal control of execution costs”. In: **Journal of financial markets** 1.1, pp. 1–50.

Cartea, Álvaro, Fayçal Drissi, and Pierre Osselin (2023). “Bandits for algorithmic trading with signals”. In: **Available at SSRN 4484004**.

Gatheral, Jim (2010). “No-dynamic-arbitrage and market impact”. In: **Quantitative finance** 10.7, pp. 749–759.

Guéant, Olivier (Mar. 2016). **The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making**. ISBN: 9780429153778. DOI: 10.1201/b21350.