

## High-Frequency Algorithmic Trading with Momentum and Order Imbalance

My goal is to establish and solve the stochastic optimal control problem that captures the momentum and order imbalance dynamics of the Limit Order Book (LOB). The solution will yield an optimal trading strategy that will permit statistical arbitrage of the underlying stock, which will then be backtested on historical data.

### Progress Timeline

DATE	THESIS	STA4505
Dec 2014	Complete CTMC calibration	
Dec 2014	Backtest naive strategies based on CTMC	
Jan-May	Study stochastic controls: ECE1639, STA4505	
Jun 5	Establish models	Exam Study
Jun 12	Establish performance criteria	Exam Study
Jun 15	Derive DPP/DPE	EXAM
Jun 19	Derive DPP/DPE	
Jun 26	Derive DPE/Solve PDEs	
Jul 3	Solve PDEs	
Jul 10	Solve PDEs and implement numerical solution	Numerical solution
Jul 31	Backtest solution on historical data	Implement simulations
Aug 15	Dissertation Writeup	Project Writeup

# For Our Readers in the Middle East...

## The Academic Week in Review

### The Dissertation

Recall that we're working on solving the DPE:

$$\max \left\{ \partial_t H + \sup_{\delta \in \mathcal{A}} \mathcal{L}_t^\delta H ; H(t, x - (s + \pi), s, \mathbf{z}, q + 1) - H(\cdot) ; \right. \\ \left. H(t, x + (s - \pi), s, \mathbf{z}, q - 1) - H(\cdot) \right\} = 0 \quad (1)$$

with boundary conditions

$$H(T, x, s, \mathbf{z}, q) = x + q(s - \text{sgn}(q)\pi) - \alpha q^2 \quad (2)$$

$$H(T, x, s, \mathbf{z}, 0) = x \quad (3)$$

and where the infinitesimal generator is given by

$$\begin{aligned} \mathcal{L}_t^\delta H &= \mu^+(\mathbf{z})e^{-\kappa\delta^-} \mathbb{E}[H(t, x + (s + \pi + \delta^-), s, \mathbf{z}, q - 1) - H(\cdot)] \\ &\quad + \mu^-(\mathbf{z})e^{-\kappa\delta^+} \mathbb{E}[H(t, x - (s - \pi - \delta^+), s, \mathbf{z}, q + 1) - H(\cdot)] \\ &\quad + \sum_{k \in P} \sum_{l \in \{-1, 0, 1\}} G_{\mathbf{z}, (k, l)} \mathbb{E}[H(t, x, s + l\eta_{0, \mathbf{z}}, \mathbf{z} + (k, l), q) - H(\cdot)] \end{aligned} \quad (4)$$

We introduced the ansatz  $H(\cdot) = x + q(s - \text{sgn}(q)\pi) + h(t, \mathbf{z}, q)$  with boundary conditions  $h(T, \mathbf{z}, q) = -\alpha q^2$  and  $h(T, \mathbf{z}, 0) = 0$ , and substituted it into infinitesimal generator. So now we proceed...

$$\begin{aligned} \mathcal{L}_t^\delta H &= \mu^+(\mathbf{z})e^{-\kappa\delta^-} [\delta^- + \pi[1 + \text{sgn}(q - 1) + q(\text{sgn}(q) - \text{sgn}(q - 1))] + h(t, \mathbf{z}, q - 1) - h(t, \mathbf{z}, q)] \\ &\quad + \mu^-(\mathbf{z})e^{-\kappa\delta^+} [\delta^+ + \pi[1 - \text{sgn}(q + 1) + q(\text{sgn}(q) - \text{sgn}(q + 1))] + h(t, \mathbf{z}, q + 1) - h(t, \mathbf{z}, q)] \\ &\quad + \sum_{k \in P} \sum_{l \in \{-1, 0, 1\}} G_{\mathbf{z}, (k, l)} [ql\mathbb{E}[\eta_{0, \mathbf{z}}] + h(t, (k, l), q) - h(t, \mathbf{z}, q)] \end{aligned}$$

In the DPE, the first term requires finding the supremum over all  $\delta^\pm$  of the infinitesimal generator. For this we can set the partial derivatives with respect to both  $\delta^+$  and  $\delta^-$  equal to zero to solve for the optimal posting depth. For  $\delta^+$  we get:

$$\begin{aligned} 0 &= \partial_{\delta^+} \left[ e^{-\kappa\delta^+} [\delta^+ + \pi[1 - \text{sgn}(q + 1) + q(\text{sgn}(q) - \text{sgn}(q + 1))] + h(t, \mathbf{z}, q + 1) - h(t, \mathbf{z}, q)] \right] \\ &= -\kappa e^{-\kappa\delta^+} [\delta^+ + \pi[1 - \text{sgn}(q + 1) + q(\text{sgn}(q) - \text{sgn}(q + 1))] + h(t, \mathbf{z}, q + 1) - h(t, \mathbf{z}, q)] + e^{-\kappa\delta^+} \\ &= e^{-\kappa\delta^+} [-\kappa(\delta^+ + \pi[1 - \text{sgn}(q + 1) + q(\text{sgn}(q) - \text{sgn}(q + 1))] + h(t, \mathbf{z}, q + 1) - h(t, \mathbf{z}, q)) + 1] \end{aligned}$$

Since  $e^{-\kappa\delta^+} > 0$ , the term inside the square braces must be equal to zero:

$$\begin{aligned} 0 &= -\kappa(\delta^+ + \pi[1 - \text{sgn}(q+1) + q(\text{sgn}(q) - \text{sgn}(q+1))] + h(t, \mathbf{z}, q+1) - h(t, \mathbf{z}, q)) + 1 \\ \delta^{+*} &= \frac{1}{\kappa} - \pi[1 - \text{sgn}(q+1) + q(\text{sgn}(q) - \text{sgn}(q+1))] - h(t, \mathbf{z}, q+1) + h(t, \mathbf{z}, q) \end{aligned}$$

We can further simplify the factor of  $\pi$ .

$$\begin{aligned} 1 - \text{sgn}(q+1) + q(\text{sgn}(q) - \text{sgn}(q+1)) &= 1 - (-\mathbb{1}_{q \leq -2} + \mathbb{1}_{q \geq 0}) + \mathbb{1}_{q=-1} \\ &= 1 + (\mathbb{1}_{q \leq -1} - \mathbb{1}_{q \geq 0}) \\ &= 2 \cdot \mathbb{1}_{q \leq -1} \end{aligned}$$

Thus, we find that the optimal buy limit order posting depth can be written in feedback form as

$$\delta^{+*} = \frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \leq -1} - h(t, \mathbf{z}, q+1) + h(t, \mathbf{z}, q) \quad (5)$$

We can follow similar steps to solve for the optimal sell limit order posting depth

$$\delta^{-*} = \frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \geq 1} - h(t, \mathbf{z}, q-1) + h(t, \mathbf{z}, q) \quad (6)$$

Turning our attention to the stopping regions of the DPE, we can use the ansatz to simplify the expressions:

$$\begin{aligned} H(t, x - (s + \pi), s, \mathbf{z}, q+1) - H(\cdot) &= x - s - \pi + (q+1)(s - \text{sgn}(q+1)\pi) + h(t, \mathbf{z}, q+1) \\ &\quad - [x + q(s - q \text{sgn}(q)\pi) + h(t, \mathbf{z}, q)] \\ &= -\pi[(q+1) \text{sgn}(q+1) - q \text{sgn}(q) + 1] + h(t, \mathbf{z}, q+1) - h(t, \mathbf{z}, q) \\ &= -2\pi \cdot \mathbb{1}_{q \geq 0} + h(t, \mathbf{z}, q+1) - h(t, \mathbf{z}, q) \end{aligned}$$

and

$$\begin{aligned} H(t, x + (s - \pi), s, \mathbf{z}, q-1) - H(\cdot) &= x + s - \pi + (q-1)(s - \text{sgn}(q-1)\pi) + h(t, \mathbf{z}, q-1) \\ &\quad - [x + q(s - q \text{sgn}(q)\pi) + h(t, \mathbf{z}, q)] \\ &= -\pi[(q-1) \text{sgn}(q-1) - q \text{sgn}(q) + 1] + h(t, \mathbf{z}, q-1) - h(t, \mathbf{z}, q) \\ &= -2\pi \cdot \mathbb{1}_{q \leq 0} + h(t, \mathbf{z}, q-1) - h(t, \mathbf{z}, q) \end{aligned}$$

Substituting all these results into the DPE, we find that  $h$  satisfies

$$\begin{aligned} 0 &= \max \left\{ \partial_t h + \mu^+(\mathbf{z}) \frac{1}{\kappa} e^{-\kappa(\frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \geq 1} - h(t, \mathbf{z}, q-1) + h(t, \mathbf{z}, q))} \right. \\ &\quad + \mu^-(\mathbf{z}) \frac{1}{\kappa} e^{-\kappa(\frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \leq -1} - h(t, \mathbf{z}, q+1) + h(t, \mathbf{z}, q))} \\ &\quad + \sum_{k \in P} \sum_{l \in \{-1, 0, 1\}} G_{\mathbf{z}, (k, l)} [q l \mathbb{E}[\eta_{0, \mathbf{z}}] + h(t, (k, l), q) - h(t, \mathbf{z}, q)] ; \\ &\quad - 2\pi \cdot \mathbb{1}_{q \geq 0} + h(t, \mathbf{z}, q+1) - h(t, \mathbf{z}, q) ; \\ &\quad \left. - 2\pi \cdot \mathbb{1}_{q \leq 0} + h(t, \mathbf{z}, q-1) - h(t, \mathbf{z}, q) \right\} \quad (7) \end{aligned}$$

Looking at the simplified feedback form in the stopping region, we see that a buy market order will be executed at time  $\tau_q^+$  whenever

$$h(\tau_q^+, \mathbf{z}, q+1) - h(\tau_q^+, \mathbf{z}, q) = 2\pi \cdot \mathbb{1}_{q \geq 0} \quad (8)$$

In particular, with negative inventory, we will execute a buy market order so long as it does not change our value function; and with zero or positive inventory, only if it increases the value function by the value of the spread. The opposite holds for sell market orders. Together, these indicate a penchant for using market orders to drive inventory levels back toward zero when it has no effect on value, and using them to gain extra value only when the expected gain is equal to the size of the spread. This is reminiscent of what we saw in the exploratory data analysis: if a stock is worth  $S$ , we can purchase it at  $S + \pi$  and immediately be able to sell it at  $S - \pi$ , at a loss of  $2\pi$ ; this was the most significant source of loss in the naive trading market order strategy. Hence we need to expect our value to increase by at least  $2\pi$  when executing market orders for gain.

When we are in the continuation region, the equality above is replaced with a  $\leq$ . Noting the feedback form of our optimal buy limit order depth given in equation 5, we thereby obtain a lower bound on  $\delta^+$  given by

$$\begin{aligned} \delta^{+*} &= \frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \leq -1} - h(t, \mathbf{z}, q+1) + h(t, \mathbf{z}, q) \\ &\geq \frac{1}{\kappa} - 2\pi \cdot \mathbb{1}_{q \leq -1} - 2\pi \cdot \mathbb{1}_{q \geq 0} \\ &= \frac{1}{\kappa} - 2\pi \end{aligned}$$

Thus, if we impose that  $\delta^\pm \geq 0$ , we should also have that  $\frac{1}{\kappa} \geq 2\pi$ .

## Case 2: Max Terminal Wealth with Risk Aversion

Our second possible performance criteria was given by:

$$2. \quad \text{Profit with risk aversion: } H^{\tau, \delta}(\cdot) = \mathbb{E} \left[ W_T^{\tau, \delta} - \gamma \cdot \mathbb{1}_{W_T^{\tau, \delta} < 0} \cdot W_T^{\tau, \delta} \right]$$

In this case, the DPE (1) is unchanged, but our boundary conditions must now reflect the risk aversion. Using as before the wealth notation  $W(x, t, s, \mathbf{z}, q) = x + q(s - \text{sgn}(q)\pi) - \alpha q^2$ , our new boundary conditions are:

$$H(T, x, s, \mathbf{z}, q) = W(T, x, s, \mathbf{z}, q) (1 - \gamma \cdot \mathbb{1}_{W(T, x, s, \mathbf{z}, q) < 0}) \quad (9)$$

$$H(T, x, s, \mathbf{z}, 0) = x (1 - \gamma \cdot \mathbb{1}_{x < 0}) \quad (10)$$

Our ansatz now too needs to reflect this risk aversion.

### Case 3: Max Terminal Wealth with Running Inventory Penalty

3. Profit with running inventory penalty:  $H^{\tau,\delta}(\cdot) = \mathbb{E} \left[ W_T^{\tau,\delta} - \varphi \int_t^T (Q_u^{\tau,\delta})^2 du \right]$

In this case, our boundary conditions are unchanged, but a  $-\varphi q^2$  term does percolate down to the DPE. Hence, our value function  $H$  is now the solution to

$$\max \left\{ \partial_t H - \varphi q^2 + \sup_{\delta \in \mathcal{A}} \mathcal{L}_t^\delta H ; H(t, x - (s + \pi), s, \mathbf{z}, q + 1) - H(\cdot) ; \right. \\ \left. H(t, x + (s - \pi), s, \mathbf{z}, q - 1) - H(\cdot) \right\} = 0 \quad (11)$$

It can easily be verified that the analysis otherwise proceeds unchanged using the same ansatz as in the first case, and we produce the same optimal posting depths and MO execution criteria. In turn, we find that  $h$  satisfies the quasi-variational inequality

$$0 = \max \left\{ \partial_t h - \varphi q^2 + \mu^+(\mathbf{z}) \frac{1}{\kappa} e^{-\kappa(\frac{1}{\kappa} - 2\pi \cdot \mathbf{1}_{q \geq 1} - h(t, \mathbf{z}, q - 1) + h(t, \mathbf{z}, q))} \right. \\ + \mu^-(\mathbf{z}) \frac{1}{\kappa} e^{-\kappa(\frac{1}{\kappa} - 2\pi \cdot \mathbf{1}_{q \leq -1} - h(t, \mathbf{z}, q + 1) + h(t, \mathbf{z}, q))} \\ + \sum_{k \in P} \sum_{l \in \{-1, 0, 1\}} G_{\mathbf{z}, (k, l)} [ql \mathbb{E}[\eta_{0, \mathbf{z}}] + h(t, (k, l), q) - h(t, \mathbf{z}, q)] ; \\ - 2\pi \cdot \mathbf{1}_{q \geq 0} + h(t, \mathbf{z}, q + 1) - h(t, \mathbf{z}, q) ; \\ \left. - 2\pi \cdot \mathbf{1}_{q \leq 0} + h(t, \mathbf{z}, q - 1) - h(t, \mathbf{z}, q) \right\} \quad (12)$$

### Looking Ahead