Stochastic Control for Optimal Market-Making

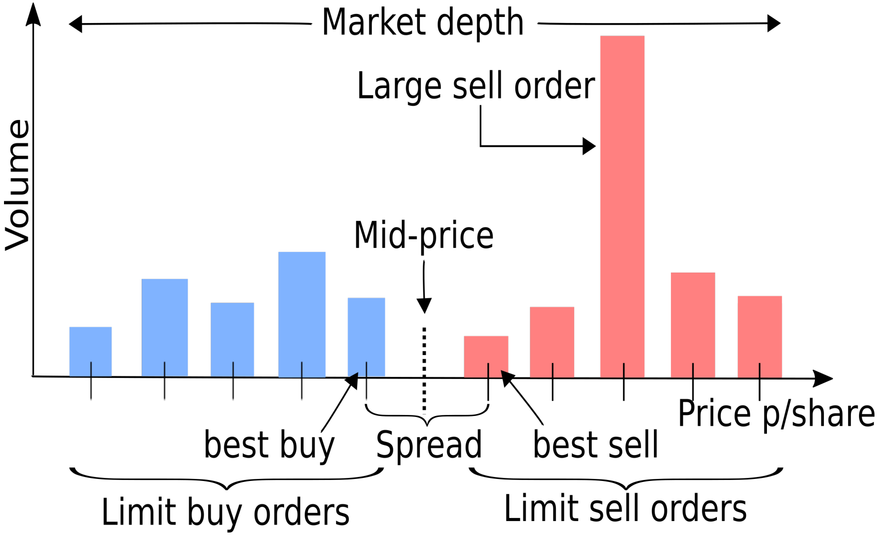
Ashwin Rao

ICME, Stanford University

December 24, 2019

Overview

Trading Order Book (TOB)



Basics of Trading Order Book (TOB)

- Buyers/Sellers express their intent to trade by submitting bids/asks
- These are Limit Orders (LO) with a price P and size N
- Buy LO (P, N) states willingness to buy N shares at a price $\leq P$
- Sell LO (P, N) states willingness to sell N shares at a price $\geq P$
- Trading Order Book aggregates order sizes for each unique price
- So we can represent with two sorted lists of (Price, Size) pairs

$$\text{Bids: } [(P_i^{(b)}, N_i^{(b)}) \mid 1 \leq i \leq m], P_i^{(b)} > P_j^{(b)} \text{ for } i < j$$

$$\text{Asks: } [(P_i^{(a)}, N_i^{(a)}) \mid 1 \leq i \leq n], P_i^{(a)} < P_j^{(a)} \text{ for } i < j$$

- We call $P_1^{(b)}$ as simply *Bid*, $P_1^{(a)}$ as *Ask*, $\frac{P_1^{(a)} + P_1^{(b)}}{2}$ as *Mid*
- We call $P_1^{(a)} - P_1^{(b)}$ as *Spread*, $P_n^{(a)} - P_m^{(b)}$ as *Market Depth*
- A Market Order (MO) states intent to buy/sell N shares at the *best possible price(s)* available on the TOB at the time of MO submission

Trading Order Book (TOB) Activity

- A new Sell LO (P, N) potentially removes best bid prices on the TOB

$$\text{Removal: } [(P_i^{(b)}, \min(N_i^{(b)}, \max(0, N - \sum_{j=1}^{i-1} N_j^{(b)}))) \mid (i : P_i^{(b)} \geq P)]$$

- After this removal, it adds the following to the asks side of the TOB

$$(P, \max(0, N - \sum_{i:P_i^{(b)} \geq P} N_i^{(b)}))$$

- A new Buy MO operates analogously (on the other side of the TOB)
- A Sell Market Order N will remove the best bid prices on the TOB

$$\text{Removal: } [(P_i^{(b)}, \min(N_i^{(b)}, \max(0, N - \sum_{j=1}^{i-1} N_j^{(b)}))) \mid 1 \leq i \leq m]$$

- A Buy Market Order N will remove the best ask prices on the TOB

$$\text{Removal: } [(P_i^{(a)}, \min(N_i^{(a)}, \max(0, N - \sum_{j=1}^{i-1} N_j^{(a)}))) \mid 1 \leq i \leq n]$$

TOB Dynamics and Market-Making

- Modeling TOB Dynamics involves predicting arrival of MOs and LOs
- Market-makers are liquidity providers (providers of Buy and Sell LOs)
- Other market participants are typically liquidity takers (MOs)
- But there are also other market participants that trade with LOs
- Complex interplay between market-makers & other mkt participants
- Hence, TOB Dynamics tend to be quite complex
- We view the TOB from the perspective of a single market-maker who aims to gain with Buy/Sell LOs of appropriate width/size
- By anticipating TOB Dynamics & dynamically adjusting Buy/Sell LOs
- Goal is to maximize *Utility of Gains* at the end of a suitable horizon
- If Buy/Sell LOs are too narrow, more frequent but small gains
- If Buy/Sell LOs are too wide, less frequent but large gains
- Market-maker also needs to manage potential unfavorable inventory (long or short) buildup and consequent unfavorable liquidation

Notation for Optimal Market-Making Problem

- We simplify the setting for ease of exposition
- Assume finite time steps indexed by $t = 0, 1, \ldots, T$
- Denote $W_t \in \mathbb{R}$ as Market-maker's trading PnL at time t
- Denote $I_t \in \mathbb{Z}$ as Market-maker's inventory of shares at time t ($I_0 = 0$)
- $S_t \in \mathbb{R}^+$ is the TOB Mid Price at time t (assume stochastic process)
- $P_t^{(b)} \in \mathbb{R}^+, N_t^{(b)} \in \mathbb{Z}^+$ are market maker's Bid Price, Bid Size at time t
- $P_t^{(a)} \in \mathbb{R}^+, N_t^{(a)} \in \mathbb{Z}^+$ are market-maker's Ask Price, Ask Size at time t
- Assume market-maker can add or remove bids/asks costlessly
- Denote $\delta_t^{(b)} = S_t - P_t^{(b)}$ as Bid Spread, $\delta_t^{(a)} = P_t^{(a)} - S_t$ as Ask Spread
- Random var $X_t^{(b)} \in \mathbb{Z}_{\geq 0}$ denotes bid-shares "hit" *up to* time t
- Random var $X_t^{(a)} \in \mathbb{Z}_{\geq 0}$ denotes ask-shares "lifted" *up to* time t

$$W_{t+1} = W_t + P_t^{(a)} \cdot (X_{t+1}^{(a)} - X_t^{(a)}) - P_t^{(b)} \cdot (X_{t+1}^{(b)} - X_t^{(b)}) \ , \ I_t = X_t^{(b)} - X_t^{(a)}$$

- Goal to maximize $\mathbb{E}[U(W_T + I_T \cdot S_T)]$ for appropriate concave $U(\cdot)$

Markov Decision Process (MDP) Formulation

- Order of MDP activity in each time step $0 \leq t \leq T-1$:
  - Observe *State* $:= (t, S_t, W_t, I_t)$
  - Perform *Action* $:= (P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)})$
  - Experience TOB Dynamics resulting in:
    - random bid-shares hit $= X_{t+1}^{(b)} - X_t^{(b)}$ and ask-shares lifted $= X_{t+1}^{(a)} - X_t^{(a)}$
    - update of W_t to W_{t+1}, update of I_t to I_{t+1}
    - stochastic evolution of S_t to S_{t+1}
  - Receive next-step $(t+1)$ *Reward* R_{t+1}

$$R_{t+1} := \begin{cases} 0 & \text{for } 1 \leq t+1 \leq T-1 \\ U(W_{t+1} + I_{t+1} \cdot S_{t+1}) & \text{for } t+1 = T \end{cases}$$

- Goal is to find an *Optimal Policy* π^*:

$$\pi^*(t, S_t, W_t, I_t) = (P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)}) \text{ that maximizes } \mathbb{E}[\sum_{t=1}^{T} R_t]$$

- Note: Discount Factor when aggregating Rewards in the MDP is 1

Avellaneda-Stoikov Continuous Time Formulation

- We go over the landmark paper by Avellaneda and Stoikov in 2006
- They derive a simple, clean and intuitive analytical solution
- We adapt our discrete-time notation to their continuous-time setting
- $X_t^{(b)}, X_t^{(a)}$ are *Poisson processes* with *arrival-rate* means $\lambda_t^{(b)}, \lambda_t^{(a)}$

$$dX_t^{(b)} \sim Poisson(\lambda_t^{(b)} \cdot dt) \; , \; dX_t^{(a)} \sim Poisson(\lambda_t^{(a)} \cdot dt)$$

$$\lambda_t^{(b)} = f^{(b)}(\delta_t^{(b)}) \; , \; \lambda_t^{(a)} = f^{(a)}(\delta_t^{(a)}) \text{ for decreasing functions } f^{(b)}, f^{(a)}$$

$$dW_t = P_t^{(a)} \cdot dX_t^{(a)} - P_t^{(b)} \cdot dX_t^{(b)} \; , \; I_t = X_t^{(b)} - X_t^{(a)} \text{ (note: } I_0 = 0)$$

- Since infinitesimal Poisson random variables $dX_t^{(b)}$ (shares hit in time dt) and $dX_t^{(a)}$ (shares lifted in time dt) are Bernoulli (shares hit/lifted in time dt are 0 or 1), $N_t^{(b)}$ and $N_t^{(a)}$ can be assumed to be 1
- This simplifies the *Action* at time t to be just the pair: $(\delta_t^{(b)}, \delta_t^{(a)})$
- TOB Mid Price Dynamics: $dS_t = \sigma \cdot dz_t$ (scaled brownian motion)
- Utility function $U(x) = -e^{-\gamma x}$ where γ is coefficient of risk-aversion

Hamilton-Jacobi-Bellman (HJB) Equation

- We denote the Optimal Value function as $V^*(t, S_t, W_t, I_t)$

$$V^*(t, S_t, W_t, I_t) = \max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[-e^{-\gamma\cdot(W_T + I_t \cdot S_T)}]$$

- $V^*(t, S_t, W_t, I_t)$ satisfies a recursive formulation for $0 \leq t < t_1 < T$:

$$V^*(t, S_t, W_t, I_t) = \max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[V^*(t_1, S_{t_1}, W_{t_1}, I_{t_1})]$$

- Rewriting in stochastic differential form, we have the HJB Equation

$$\max_{\delta_t^{(b)}, \delta_t^{(a)}} \mathbb{E}[dV^*(t, S_t, W_t, I_t)] = 0 \text{ for } t < T$$

$$V^*(T, S_T, W_T, I_T) = -e^{-\gamma\cdot(W_T + I_T \cdot S_T)}$$

Converting HJB to a Partial Differential Equation

- Change to $V^*(t, S_t, W_t, I_t)$ is comprised of 3 components:
  - Due to pure movement in time t
  - Due to randomness in LOB Mid-Price S_t
  - Due to randomness in hitting/lifting the Bid/Ask
- With this, we can expand $dV^*(t, S_t, W_t, I_t)$ and rewrite HJB as:

$$\max_{\delta_t^{(b)}, \delta_t^{(a)}} \{ \frac{\partial V^*}{\partial t} dt + \mathbb{E}[\sigma \frac{\partial V^*}{\partial S_t} dz_t + \frac{\sigma^2}{2} \frac{\partial^2 V^*}{\partial S_t^2} (dz_t)^2]$$

$$+ \lambda_t^{(b)} \cdot dt \cdot V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1)$$

$$+ \lambda_t^{(a)} \cdot dt \cdot V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1)$$

$$+ (1 - \lambda_t^{(b)} \cdot dt - \lambda_t^{(a)} \cdot dt) \cdot V^*(t, S_t, W_t, I_t)$$

$$- V^*(t, S_t, W_t, I_t) \} = 0$$

Converting HJB to a Partial Differential Equation

We can simplify this equation with a few observations:

- $\mathbb{E}[dz_t] = 0$
- $\mathbb{E}[(dz_t)^2] = dt$
- Organize the terms involving $\lambda_t^{(b)}$ and $\lambda_t^{(a)}$ better with some algebra
- Divide throughout by dt

$$
\begin{aligned}
\max_{\delta_t^{(b)},\delta_t^{(a)}} \{ &\frac{\partial V^*}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 V^*}{\partial S_t^2} \\
&+ \lambda_t^{(b)} \cdot (V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1) - V^*(t, S_t, W_t, I_t)) \\
&+ \lambda_t^{(a)} \cdot (V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1) - V^*(t, S_t, W_t, I_t))\} = 0
\end{aligned}
$$

Converting HJB to a Partial Differential Equation

Next, note that $\lambda_t^{(b)} = f^{(b)}(\delta_t^{(b)})$ and $\lambda_t^{(a)} = f^{(a)}(\delta_t^{(a)})$, and apply the max only on the relevant terms

$$\frac{\partial V^*}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 V^*}{\partial S_t^2}$$
$$+ \max_{\delta_t^{(b)}}\{f^{(b)}(\delta_t^{(b)}) \cdot (V^*(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1) - V^*(t, S_t, W_t, I_t))\}$$
$$+ \max_{\delta_t^{(a)}}\{f^{(a)}(\delta_t^{(a)}) \cdot (V^*(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1) - V^*(t, S_t, W_t, I_t))\} = 0$$

This combines with the boundary condition:

$$V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)}$$

Converting HJB to a Partial Differential Equation

- We make an "educated guess" for the structure of $V^*(t, S_t, W_t, I_t)$:
$$V^*(t, S_t, W_t, I_t) = -e^{-\gamma(W_t + \theta(t, S_t, I_t))} \tag{1}$$
and reduce the problem to a PDE in terms of $\theta(t, S_t, I_t)$
- Substituting this into the above PDE for $V^*(t, S_t, W_t, I_t)$ gives:
$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right)$$
$$+ \max_{\delta_t^{(b)}} \left\{ \frac{f^{(b)}(\delta_t^{(b)})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(b)} - S_t + \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t))}\right) \right\}$$
$$+ \max_{\delta_t^{(a)}} \left\{ \frac{f^{(a)}(\delta_t^{(a)})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(a)} + S_t + \theta(t, S_t, I_t - 1) - \theta(t, S_t, I_t))}\right) \right\} = 0$$
- The boundary condition is:
$$\theta(T, S_T, I_T) = I_T \cdot S_T$$

Indifference Bid/Ask Price

- It turns out that $\theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t)$ and $\theta(t, S_t, I_t) - \theta(t, S_t, I_t - 1)$ are equal to financially meaningful quantities known as *Indifference Bid and Ask Prices*
- Indifference Bid Price $Q^{(b)}(t, S_t, I_t)$ is defined as:

$$V^*(t, S_t, W_t - Q^{(b)}(t, S_t, I_t), I_t + 1) = V^*(t, S_t, W_t, I_t) \tag{2}$$

- $Q^{(b)}(t, S_t, I_t)$ is the price to buy a share with *guarantee of immediate purchase* that results in Optimum Expected Utility being unchanged
- Likewise, Indifference Ask Price $Q^{(a)}(t, S_t, I_t)$ is defined as:

$$V^*(t, S_t, W_t + Q^{(a)}(t, S_t, I_t), I_t - 1) = V^*(t, S_t, W_t, I_t) \tag{3}$$

- $Q^{(a)}(t, S_t, I_t)$ is the price to sell a share with *guarantee of immediate sale* that results in Optimum Expected Utility being unchanged
- We abbreviate $Q^{(b)}(t, S_t, I_t)$ as $Q_t^{(b)}$ and $Q^{(a)}(t, S_t, I_t)$ as $Q_t^{(a)}$

Indifference Bid/Ask Price in the PDE for θ

- Express $V^*(t, S_t, W_t - Q_t^{(b)}, I_t + 1) = V^*(t, S_t, W_t, I_t)$ in terms of θ:

$$-e^{-\gamma(W_t - Q_t^{(b)} + \theta(t, S_t, I_t + 1))} = -e^{-\gamma(W_t + \theta(t, S_t, I_t))}$$

$$\Rightarrow Q_t^{(b)} = \theta(t, S_t, I_t + 1) - \theta(t, S_t, I_t) \tag{4}$$

- Likewise for $Q_t^{(a)}$, we get:

$$Q_t^{(a)} = \theta(t, S_t, I_t) - \theta(t, S_t, I_t - 1) \tag{5}$$

- Using equations (4) and (5), bring $Q_t^{(b)}$ and $Q_t^{(a)}$ in the PDE for θ

$$\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right) + \max_{\delta_t^{(b)}} g(\delta_t^{(b)}) + \max_{\delta_t^{(a)}} h(\delta_t^{(b)}) = 0$$

$$\text{where } g(\delta_t^{(b)}) = \frac{f^{(b)}(\delta_t^{(b)})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(b)} - S_t + Q_t^{(b)})})$$

$$\text{and } h(\delta_t^{(a)}) = \frac{f^{(a)}(\delta_t^{(a)})}{\gamma} \cdot (1 - e^{-\gamma(\delta_t^{(a)} + S_t - Q_t^{(a)})})$$

Optimal Bid Spread and Optimal Ask Spread

- To maximize $g(\delta_t^{(b)})$, differentiate g with respect to $\delta_t^{(b)}$ and set to 0

$$e^{-\gamma(\delta_t^{(b)*} - S_t + Q_t^{(b)})} \cdot (\gamma \cdot f^{(b)}(\delta_t^{(b)*}) - \frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)*})) + \frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)*}) = 0$$

$$\Rightarrow \delta_t^{(b)*} = S_t - Q_t^{(b)} + \frac{1}{\gamma} \cdot \ln{(1 - \gamma \cdot \frac{f^{(b)}(\delta_t^{(b)*})}{\frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}(\delta_t^{(b)*})})} \tag{6}$$

- To maximize $g(\delta_t^{(a)})$, differentiate g with respect to $\delta_t^{(a)}$ and set to 0

$$e^{-\gamma(\delta_t^{(a)*} + S_t - Q_t^{(a)})} \cdot (\gamma \cdot f^{(a)}(\delta_t^{(a)*}) - \frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)*})) + \frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)*}) = 0$$

$$\Rightarrow \delta_t^{(a)*} = Q_t^{(a)} - S_t + \frac{1}{\gamma} \cdot \ln{(1 - \gamma \cdot \frac{f^{(a)}(\delta_t^{(a)*})}{\frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}(\delta_t^{(a)*})})} \tag{7}$$

- (6) and (7) are implicit equations for $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$ respectively

Solving for θ and for Optimal Bid/Ask Spreads

- Let us write the PDE in terms of the Optimal Bid and Ask Spreads

$$\begin{aligned}
&\frac{\partial \theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2} - \gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right) \\
&+ \frac{f^{(b)}(\delta_t^{(b)*})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(b)*} - S_t + \theta(t,S_t,I_t+1) - \theta(t,S_t,I_t))}\right) \\
&+ \frac{f^{(a)}(\delta_t^{(a)*})}{\gamma} \cdot \left(1 - e^{-\gamma(\delta_t^{(a)*} + S_t + \theta(t,S_t,I_t-1) - \theta(t,S_t,I_t))}\right) = 0
\end{aligned} \tag{8}$$

with boundary condition $\theta(T, S_T, I_T) = I_T \cdot S_T$

- First we solve PDE (8) for θ in terms of $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$
- In general, this would be a numerical PDE solution
- Using (4) and (5), we have $Q_t^{(b)}$ and $Q_t^{(a)}$ in terms of $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$
- Substitute above-obtained $Q_t^{(b)}$ and $Q_t^{(a)}$ in equations (6) and (7)
- Solve implicit equations for $\delta_t^{(b)*}$ and $\delta_t^{(a)*}$ (in general, numerically)

Building Intuition

- Define *Indifference Mid Price* $Q_t^{(m)} = \frac{Q_t^{(b)} + Q_t^{(a)}}{2}$
- To develop intuition for Indifference Prices, consider a simple case where the market-maker doesn't supply any bids or asks. Then,

$$V^*(t, S_t, W_t, I_t) = \mathbb{E}[-e^{-\gamma(W_t + I_t \cdot S_T)}]$$

- Combining this with the diffusion $dS_t = \sigma \cdot dz_t$, we get:

$$V^*(t, S_t, W_t, I_t) = -e^{-\gamma(W_t + I_t \cdot S_t - \frac{\gamma \cdot I_t^2 \cdot \sigma^2 (T-t)}{2})}$$

- Combining this with equations (2) and (3), we get:

$$Q_t^{(b)} = S_t + (1 - 2I_t)\frac{\gamma\sigma^2(T-t)}{2}$$

$$Q_t^{(a)} = S_t + (-1 - 2I_t)\frac{\gamma\sigma^2(T-t)}{2}$$

$$Q_t^{(m)} = S_t - I_t\gamma\sigma^2(T-t)$$

$$Q_t^{(a)} - Q_t^{(b)} = \gamma\sigma^2(T-t)$$

Building Intuition

- Think of $Q_t^{(m)}$ as *inventory-risk-adjusted* mid-price (adjustment to S_t)
- If market-maker is long inventory ($I_t > 0$), $Q_t^{(m)} < S_t$ indicating inclination to sell than buy, and if market-maker is short inventory, $Q_t^{(m)} > S_t$ indicating inclination to buy than sell
- Armed with this intuition, we come back to optimal market-making, observing from eqns (6) and (7): ${P_t^{(b)}}^* < Q_t^{(b)} < Q_t^{(m)} < Q_t^{(a)} < {P_t^{(a)}}^*$
- Think of $[{P_t^{(b)}}^*, {P_t^{(a)}}^*]$ as "centered" at $Q_t^{(m)}$ (rather than at S_t), i.e., $[{P_t^{(b)}}^*, {P_t^{(a)}}^*]$ will (together) move up/down in tandem with $Q_t^{(m)}$ moving up/down (as a function of inventory position I_t)

$$Q_t^{(m)} - {P_t^{(b)}}^* = \frac{Q_t^{(a)} - Q_t^{(b)}}{2} + \frac{1}{\gamma} \cdot \ln\left(1 - \gamma \cdot \frac{f^{(b)}({\delta_t^{(b)}}^*)}{\frac{\partial f^{(b)}}{\partial \delta_t^{(b)}}({\delta_t^{(b)}}^*)}\right) \tag{9}$$

$${P_t^{(a)}}^* - Q_t^{(m)} = \frac{Q_t^{(a)} - Q_t^{(b)}}{2} + \frac{1}{\gamma} \cdot \ln\left(1 - \gamma \cdot \frac{f^{(a)}({\delta_t^{(a)}}^*)}{\frac{\partial f^{(a)}}{\partial \delta_t^{(a)}}({\delta_t^{(a)}}^*)}\right) \tag{10}$$

Real-world Market-Making and Reinforcement Learning

- Arbitrary Price Dynamics $f_t(\cdot)$ and Temporary Price Impact $g_t(\cdot)$
- Non-stationarity/non-linear dynamics/impact require (Numerical) DP
- Frictions: Discrete Prices/Sizes, Constraints on Prices/Sizes, Fees
- Large State space incorporate various external factors in the State
- We could also represent the entire TOB within the State
- So then we'd have to develop a simulator capturing all of the above
- Simulator is a *Data-learnt Sampling Model* of TOB Dynamics
- In practice, we'd need to also capture *Cross-Asset Market Impact*
- Using this simulator and neural-networks func approx, we can do RL
- References: Nevmyvaka, Feng, Kearns; 2006 and Vyetrenko, Xu; 2019
- Exciting area for Future Research as well as Engineering Design