Market Making Problem

Miao-Chin Yen

1 Problem Formulation

There are two important roles in a market. One is liquidity provider and the other is liquidity taker. Market-makers who provide buy and sell limit orders (LOs) are main liquidity providers. Other market participants are mostly liquidity takers who submit market orders (MOs). In this problem, we consider ourselves as a single market-maker attempting to make profits by submitting buy/sell LOs with appropriate size. For maximizing profits, we need to take order book (OB) dynamics into account and adjust sell/buy LOs dynamically. For simplicity, we consider this problem with finite time horizon T. We then first give some notations as follows:

- $W_t \in R$ is market-maker's trading account value at time t.
- $I_t \in Z$ is market-maker's inventory of shares at time t (assume $I_0 = 0$).
- $S_t \in R^+$ is the OB mid price at time t. We assume S_t follows a stochastic process $dS_t = \sigma \cdot dz_t$ where z_t is a brownian motion.
- $P_t^{(b)} \in R^+, N_t^{(b)} \in Z^+$ are market maker's bid price, bid size at time t.
- $P_t^{(a)} \in R^+, N_t^{(a)} \in Z^+$ are market-maker's ask price, ask size at time t.
- $\delta_t^{(b)} = S_t - P_t^{(b)}$ is bid spread.
- $\delta_t^{(a)} = P_t^{(a)} - S_t$ is ask spread.
- $X_t^{(b)} \in Z_{\geq 0}$ represents bid-shares hit up to time t. We assume $X_t^{(b)}$ are independent Poisson processes with hit-rate intensities $\lambda_t^{(b)}$ and $\lambda_t^{(b)} = f^{(b)}\left(\delta_t^{(b)}\right)$ where $f^{(b)}(\cdot)$ is a decreasing function.
- $X_t^{(a)} \in Z_{\geq 0}$ represents ask-shares lifted up to time t. We assume $X_t^{(a)}$ are independent Poisson processes with lift-rate intensities $\lambda_t^{(a)}$ and $\lambda_t^{(a)} = f^{(a)}\left(\delta_t^{(a)}\right)$ where $f^{(a)}(\cdot)$ is a decreasing function.

Given the above notations,

$$dW_t = P_t^{(a)} \cdot dX_t^{(a)} - P_t^{(b)} \cdot dX_t^{(b)} \quad \text{and} \quad I_t = X_t^{(b)} - X_t^{(a)}$$

Our goal is to maximize $E\left[U\left(W_T + I_T \cdot S_T\right)\right]$ where $U(x) = -e^{-\gamma x}$ is a risk-averse utility function and $\gamma > 0$ is the coefficient of risk-aversion. Note that we assume market-maker can add or remove bids/asks costlessly.

2 Derivation of Solution

We can solve this problem by formulating it into a Markov Decision Process (MDP) as follows:

- State $s_t := (S_t, W_t, I_t) \in \mathcal{S}_t$
- Action $a_t := \left(P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)}\right) \in \mathcal{A}_t$

- Reward

$$r_{t+1} := \begin{cases} 0 & \text{for } 0 < t+1 < T \\ U\left(W_{t+1} + I_{t+1} \cdot S_{t+1}\right) & \text{for } t+1 = T \end{cases}$$

Our goal turns into finding an optimal policy $\pi^* = (\pi_t^*)_{0 \le t < T}$ s.t

$$\pi^* = \arg\max_\pi E\left[r_T\right] \text{ where } \pi_t^*\left((S_t, W_t, I_t)\right) = \left(P_t^{(b)}, N_t^{(b)}, P_t^{(a)}, N_t^{(a)}\right)$$

Note that we set the discount factor in the MDP to be 1 since this is a finite horizon MDP. $N_t^{(b)}$ and $N_t^{(a)}$ can be assumed to be 1 since once an LO is executed, the state changes immediately and we need to take an action. Therefore, we could submit an order with quantity 1 without loss of generality. Hence, we modify the action at time t to be a pair $\left(\delta_t^{(b)}, \delta_t^{(a)}\right)$.

We then denote the optimal value function as $V^*(t, S_t, W_t, I_t)$ and the optimal value function satisfies a recursive formulation for $0 \le t < t_1 < T$:

$$V^*(t, S_t, W_t, I_t) = \max_{\delta_u^{(b)}, \delta_u^{(a)}: t \le u < T} E\left[-e^{-\gamma \cdot (W_T + I_T \cdot S_T)}\right] = \max_{\delta_u^{(b)}, \delta_u^{(a)}: t \le u < t_1} E\left[V^*(t_1, S_{t_1}, W_{t_1}, I_{t_1})\right]$$

We have the Hamilton-Jacobi-Bellman (HJB) equation if we rewrite the recursion in stochastic differential form:

$$\begin{cases} \max_{\delta_t^{(b)}, \delta_t^{(a)}} E\left[dV^*(t, S_t, W_t, I_t)\right] = 0 \text{ for } t < T \\ \\ V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)} \end{cases} \tag{1}$$

$dV^*(t, S_t, W_t, I_t)$ consists of three elements:

- Pure movement in time: t
- Randomness in OB mid-price: S_t
- Randomness in hitting/lifting the market-maker's bid/ask: $\lambda_t^{(a)}, \lambda_t^{(b)}$

Hence, we can rewrite the HJB as

$$\begin{cases} \max_{\delta_t^{(b)}, \delta_t^{(a)}} \{\dfrac{\partial V^*}{\partial t} \cdot dt + E\left[\sigma \cdot \dfrac{\partial V^*}{\partial S_t} \cdot dz_t + \dfrac{\sigma^2}{2} \cdot \dfrac{\partial^2 V^*}{\partial S_t^2} \cdot (dz_t)^2\right] \\ \qquad + \lambda_t^{(b)} \cdot dt \cdot V^*\left(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1\right) \\ \qquad + \lambda_t^{(a)} \cdot dt \cdot V^*\left(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1\right) \\ \qquad + \left(1 - \lambda_t^{(b)} \cdot dt - \lambda_t^{(a)} \cdot dt\right) \cdot V^*(t, S_t, W_t, I_t) \\ \qquad - V^*(t, S_t, W_t, I_t)\} = 0 \\ \\ V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)} \end{cases} \tag{2}$$

Since $E\left[dz_t\right] = 0$ and $E\left[(dz_t)^2\right] = dt$ (Quadratic Variation), we reorganize the equation and divide it by dt:

$$\begin{cases} \max_{\delta_t^{(b)}, \delta_t^{(a)}} \left\{\dfrac{\partial V^*}{\partial t} + \dfrac{\sigma^2}{2} \dfrac{\partial^2 V^*}{\partial S_t^2} + \lambda_t^{(b)} \cdot \left(V^*\left(t, S_t, W_t - S_t + \delta_t^{(b)}, I_t + 1\right) - V^*(t, S_t, W_t, I_t)\right) \right. \\ \qquad \left. + \lambda_t^{(a)} \cdot \left(V^*\left(t, S_t, W_t + S_t + \delta_t^{(a)}, I_t - 1\right) - V^*(t, S_t, W_t, I_t)\right)\right\} = 0 \\ \\ V^*(T, S_T, W_T, I_T) = -e^{-\gamma \cdot (W_T + I_T \cdot S_T)} \end{cases} \tag{3}$$

Since $\lambda_t^{(b)}$ and $\lambda_t^{(a)}$ are only related to $\delta_t^{(b)}$ and $\delta_t^{(a)}$ respectively, we apply the max only on them:

$$\begin{cases} \frac{\partial V^*}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 V^*}{\partial S_t^2} + \max_{\delta_t^{(b)}}\left\{f^{(b)}\left(\delta_t^{(b)}\right)\cdot\left(V^*\left(t,S_t,W_t-S_t+\delta_t^{(b)},I_t+1\right)-V^*\left(t,S_t,W_t,I_t\right)\right)\right\} \\ + \max_{\delta_t^{(a)}}\left\{f^{(a)}\left(\delta_t^{(a)}\right)\cdot\left(V^*\left(t,S_t,W_t+S_t+\delta_t^{(a)},I_t-1\right)-V^*\left(t,S_t,W_t,I_t\right)\right)\right\} = 0 \\ \\ V^*\left(T,S_T,W_T,I_T\right) = -e^{-\gamma\cdot(W_T+I_T\cdot S_T)} \end{cases} \tag{4}$$

We then make a guess for $V^*\left(t,S_t,W_t,I_t\right)$:

$$V^*\left(t,S_t,W_t,I_t\right) = -e^{-\gamma(W_t+\theta(t,S_t,I_t))}$$

and reduce the problem to a PDE in terms of $\theta\left(t,S_t,I_t\right)$.

$$\begin{cases} \frac{\partial V}{\partial t} = -\gamma\cdot V\frac{\partial\theta}{\partial t} \\ \frac{\partial V}{\partial S_t} = -\gamma\cdot V\frac{\partial\theta}{\partial S_t} \\ \frac{\partial^2 V}{\partial {S_t}^2} = \gamma\cdot V\left(\gamma\left(\frac{\partial\theta}{\partial S_t}\right)^2-\frac{\partial^2\theta}{\partial {S_t}^2}\right) \end{cases} \tag{5}$$

We substitute (5) into the PDE of (4) and get:

$$\begin{cases} \frac{\partial\theta}{\partial t} + \frac{\sigma^2}{2}\left(\frac{\partial^2\theta}{\partial S_t^2}-\gamma\left(\frac{\partial\theta}{\partial S_t}\right)^2\right) \\ + \max_{\delta_t^{(b)}}\left\{\frac{f^{(b)}\left(\delta_t^{(b)}\right)}{\gamma}\cdot\left(1-e^{-\gamma\left(\delta_t^{(b)}-S_t+\theta(t,S_t,I_t+1)-\theta(t,S_t,I_t)\right)}\right)\right\} \\ + \max_{\delta_t^{(a)}}\left\{\frac{f^{(a)}\left(\delta_t^{(a)}\right)}{\gamma}\cdot\left(1-e^{-\gamma\left(\delta_t^{(a)}+S_t+\theta(t,S_t,I_t-1)-\theta(t,S_t,I_t)\right)}\right)\right\} = 0 \\ \\ \theta\left(T,S_T,I_T\right) = I_T\cdot S_T \end{cases} \tag{6}$$

We then introduce reservation bid price $Q^{(b)}\left(t,S_t,I_t\right)$ to be the price to buy a share with guarantee of immediate purchase resulting in optimum expected utility being unchanged. Therefore, $Q^{(b)}\left(t,S_t,I_t\right)$ is defined as:

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

Similarly, reservation ask price $Q^{(a)}\left(t,S_t,I_t\right)$ is the price to sell a share with guarantee of immediate sell resulting in optimum expected utility being unchanged. Therefore, $Q^{(a)}\left(t,S_t,I_t\right)$ is defined as:

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

We then express reservation prices in terms of θ:

$$\begin{gathered} -e^{-\gamma\left(W_t-Q^{(b)}(t,S_t,I_t)+\theta(t,S_t,I_t+1)\right)} = -e^{-\gamma(W_t+\theta(t,S_t,I_t))} \\ \Rightarrow Q^{(b)}\left(t,S_t,I_t\right) = \theta\left(t,S_t,I_t+1\right)-\theta\left(t,S_t,I_t\right) \end{gathered} \tag{7}$$

$$\begin{gathered} -e^{-\gamma\left(W_t+Q^{(a)}(t,S_t,I_t)+\theta(t,S_t,I_t-1)\right)} = -e^{-\gamma(W_t+\theta(t,S_t,I_t))} \\ \Rightarrow Q^{(a)}\left(t,S_t,I_t\right) = \theta\left(t,S_t,I_t\right)-\theta\left(t,S_t,I_t-1\right) \end{gathered} \tag{8}$$

We further substitute $Q^{(b)}(t, S_t, I_t)$ and $Q^{(a)}(t, S_t, I_t)$ in the PDE of (6) for θ :

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

To find the optimal bid and ask spread, we use the first order condition:

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

Thus, Eqn. (6) in terms of $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ turns into:

$$
\begin{cases}
\frac{\partial \theta}{\partial t}+\frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2}-\gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right) \\
+\frac{f^{(b)}\left(\delta_t^{(b)^*}\right)}{\gamma}\cdot\left(1-e^{-\gamma\left(\delta_t^{(b)^*}-S_t+\theta(t,S_t,I_t+1)-\theta(t,S_t,I_t)\right)}\right) \\
+\frac{f^{(a)}\left(\delta_t^{(a)^*}\right)}{\gamma}\cdot\left(1-e^{-\gamma\left(\delta_t^{(a)^*}+S_t+\theta(t,S_t,I_t-1)-\theta(t,S_t,I_t)\right)}\right)=0 \\
\theta(T,S_T,I_T)=I_T\cdot S_T
\end{cases}
\tag{9}
$$

We then make some assumptions to derive analytical approximations. First, let

$$
f^{(b)}(\delta)=f^{(a)}(\delta)=A\cdot e^{-k\cdot\delta}
$$

This reduces $\delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ to:

$$
\delta_t^{(b)^*}=S_t-Q^{(b)}(t,S_t,I_t)+\frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) \tag{10}
$$

$$
\delta_t^{(a)^*}=Q^{(a)}(t,S_t,I_t)-S_t+\frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) \tag{11}
$$

We substitute (10) and (11) into (9):

$$
\begin{cases}
\frac{\partial \theta}{\partial t}+\frac{\sigma^2}{2}\left(\frac{\partial^2 \theta}{\partial S_t^2}-\gamma\left(\frac{\partial \theta}{\partial S_t}\right)^2\right)+\frac{A}{k+\gamma}\left(e^{-k\cdot\delta_t^{(b)^*}}+e^{-k\cdot\delta_t^{(a)^*}}\right)=0 \\
\theta(T,S_T,I_T)=I_T\cdot S_T
\end{cases}
\tag{12}
$$

We then make a linear approximation by using Taylor expansion to the first order term for $e^{-k\cdot\delta_t^{(b)^*}}$ and $e^{-k\cdot\delta_t^{(a)^*}}$ in the PDE of (12):

$$\frac{\partial\theta}{\partial t}+\frac{\sigma^2}{2}\left(\frac{\partial^2\theta}{\partial S_t^2}-\gamma\left(\frac{\partial\theta}{\partial S_t}\right)^2\right)+\frac{A}{k+\gamma}\left(1-k\cdot\delta_t^{(b)^*}+1-k\cdot\delta_t^{(a)^*}\right)=0$$

Given (7), (8), (10), and (11), we get

$$\delta_t^{(b)^*}+\delta_t^{(a)^*}=\frac{2}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)+2\theta\left(t,S_t,I_t\right)-\theta\left(t,S_t,I_t+1\right)-\theta\left(t,S_t,I_t-1\right)$$

Thus,

$$\begin{aligned}\frac{\partial\theta}{\partial t}+\frac{\sigma^2}{2}\left(\frac{\partial^2\theta}{\partial S_t^2}-\gamma\left(\frac{\partial\theta}{\partial S_t}\right)^2\right)+\frac{A}{k+\gamma}\Bigg(2-\frac{2k}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)\\ -k\left(2\theta\left(t,S_t,I_t\right)-\theta\left(t,S_t,I_t+1\right)-\theta\left(t,S_t,I_t-1\right)\right)\Bigg)=0\end{aligned}$$

To solve the above PDE, we consider an asymptotic expansion of θ in I_t :

$$\theta\left(t,S_t,I_t\right)=\sum_{n=0}^{\infty}\frac{I_t^n}{n!}\cdot\theta^{(n)}\left(t,S_t\right)$$

We only approximate this expansion to the first 3 terms:

$$\theta\left(t,S_t,I_t\right)\approx\theta^{(0)}\left(t,S_t\right)+I_t\cdot\theta^{(1)}\left(t,S_t\right)+\frac{I_t^2}{2}\cdot\theta^{(2)}\left(t,S_t\right)$$

Note that V^* can depend on S_t only through I_t. Thus,

$$\begin{gathered}V^*\left(t,S_t,W_t,0\right)=-e^{-\gamma\left(W_t+\theta^{(0)}(t,S_t)\right)}\text{ is independent of }S_t\Rightarrow\theta^{(0)}\left(t,S_t\right)\text{ is independent of }S_t\\ \Rightarrow\theta^{(0)}\left(t,S_t\right)=\theta^{(0)}\left(t\right)\end{gathered}$$

Therefore, the approximated expansion for $\theta\left(t,S_t,I_t\right)$ is

$$\theta\left(t,S_t,I_t\right)=\theta^{(0)}(t)+I_t\cdot\theta^{(1)}\left(t,S_t\right)+\frac{I_t^2}{2}\cdot\theta^{(2)}\left(t,S_t\right)$$

We substitute this approximation for $\theta\left(t,S_t,I_t\right)$ and get

$$\begin{cases}\frac{\partial\theta^{(0)}}{\partial t}+I_t\frac{\partial\theta^{(1)}}{\partial t}+\frac{I_t^2}{2}\frac{\partial\theta^{(2)}}{\partial t}+\frac{\sigma^2}{2}\left(I_t\frac{\partial^2\theta^{(1)}}{\partial S_t^2}+\frac{I_t^2}{2}\frac{\partial^2\theta^{(2)}}{\partial S_t^2}\right)\\ -\frac{\gamma\sigma^2}{2}\left(I_t\frac{\partial\theta^{(1)}}{\partial S_t}+\frac{I_t^2}{2}\frac{\partial\theta^{(2)}}{\partial S_t}\right)^2+\frac{A}{k+\gamma}\left(2-\frac{2k}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)+k\cdot\theta^{(2)}\right)=0\\ \theta^{(0)}(T)+I_T\cdot\theta^{(1)}\left(T,S_T\right)+\frac{I_T^2}{2}\cdot\theta^{(2)}\left(T,S_T\right)=I_T\cdot S_T\end{cases}\tag{13}$$

We separately collect terms involving specific powers of I_t and yield three PDEs:

- I_t :

$$\begin{cases}\frac{\partial\theta^{(1)}}{\partial t}+\frac{\sigma^2}{2}\cdot\frac{\partial^2\theta^{(1)}}{\partial S_t^2}=0\\ \theta^{(1)}\left(T,S_T\right)=S_T\end{cases}$$

$$\Rightarrow\theta^{(1)}\left(t,S_t\right)=S_t$$

- I_t^2 :

$$\begin{cases} \frac{\partial \theta^{(2)}}{\partial t} + \frac{\sigma^2}{2} \cdot \frac{\partial^2 \theta^{(2)}}{\partial S_t^2} - \gamma\sigma^2 \cdot \left(\frac{\partial \theta^{(1)}}{\partial S_t}\right)^2 = \frac{\partial \theta^{(2)}}{\partial t} + \frac{\sigma^2}{2} \cdot \frac{\partial^2 \theta^{(2)}}{\partial S_t^2} - \gamma\sigma^2 \cdot 1 = 0 \\ \theta^{(2)}\left(T, S_T\right) = 0 \end{cases}$$

$$\Rightarrow \theta^{(2)}\left(t, S_t\right) = -\gamma\sigma^2(T-t)$$

- Constant term (without I_t and I_t^2) :

$$\begin{cases} \frac{\partial \theta^{(0)}}{\partial t} + \frac{A}{k+\gamma}\left(2 - \frac{2k}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + k \cdot \theta^{(2)}\right) = \frac{\partial \theta^{(0)}}{\partial t} + \frac{A}{k+\gamma}\left(2 - \frac{2k}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - k \cdot \gamma\sigma^2(T-t)\right) = 0 \\ \theta^{(0)}(T) = 0 \end{cases}$$

$$\Rightarrow \theta^{(0)}(t) = \frac{A}{k+\gamma}\left(\left(2 - \frac{2k}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)\right)(T-t) - \frac{k\gamma\sigma^2}{2}(T-t)^2\right)$$

We then derive formulas for $Q^{(b)}\left(t, S_t, I_t\right), Q^{(a)}\left(t, S_t, I_t\right), \delta_t^{(b)^*}$ and $\delta_t^{(a)^*}$ as follows:

$$\begin{aligned} Q^{(b)}\left(t, S_t, I_t\right) &= \theta^{(1)}\left(t, S_t\right) + \left(2I_t+1\right) \cdot \theta^{(2)}\left(t, S_t\right) = S_t - \left(2I_t+1\right)\frac{\gamma\sigma^2(T-t)}{2} \\ Q^{(a)}\left(t, S_t, I_t\right) &= \theta^{(1)}\left(t, S_t\right) + \left(2I_t-1\right) \cdot \theta^{(2)}\left(t, S_t\right) = S_t - \left(2I_t-1\right)\frac{\gamma\sigma^2(T-t)}{2} \end{aligned}$$

$$\begin{aligned} \delta_t^{(b)^*} &= \frac{\left(2I_t+1\right)\gamma\sigma^2(T-t)}{2} + \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) \\ \delta_t^{(a)^*} &= \frac{\left(1-2I_t\right)\gamma\sigma^2(T-t)}{2} + \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) \end{aligned}$$

3 Numerical Simulations

We compare the performance of the optimal policy (derived above) against a policy (call it "naive policy") which is always symmetric around the OB mid price. Naive policy is with a constant bid-ask spread equal to the average bid-ask spread of the optimal policy for the optimal market-making problem. We first generate a large number of simulation traces. Each simulation trace consists of $\frac{T}{\Delta t}$ time steps. The process for performing a time step in each simulation trace for the optimal policy is as follows:

- At each time t, we observe the state $\left(S_t, W_t, I_t\right)$ and calculate the optimal action $\left(\delta_t^{(b)^*}, \delta_t^{(a)^*}\right)$.
- With probability $A \cdot e^{-k \cdot \delta_t^{(a)^*}} \cdot \Delta t$, I_t is decremented by 1 and W_t (the trading PnL) is increased by $P_t^{(a)^*}$.
- With probability $A \cdot e^{-k \cdot \delta_t^{(b)^*}} \cdot \Delta t$, I_t is incremented by 1 and W_t (the trading PnL) is decreased by $P_t^{(b)^*}$.
- S_t is incremented or decremented randomly (each with probability 0.5) by $\sigma \cdot \sqrt{\Delta t}$.
- These updates to $\left(S_t, W_t, I_t\right)$ give us the state for the next time $t + \Delta t$.

We run 10,000 simulation traces and calculate the average bid-ask spread across all time steps across all simulation traces for the naive policy. We then run the same large number of simulations for the naive policy. We use the following parameters in our simulations which are provided by Avallaneda-Stoikov:

$$S_0 = 100, T = 1, \Delta t = 0.005, \sigma = 2, I_0 = 0, k = 1.5, A = 140, W_0 = 0$$

Note that we could buy a share of stock even when we do not have enough money in our trading account. Also, we can short sell a stock in our simulations.

We first set γ to be 0.1 and run a simulation trace for the optimal policy to compare the evolution of OB mid price, optimal ask price, and optimal bid price. We further let optimal mid price be the average of optimal ask price and optimal bid price. We can view the optimal mid price as inventory-risk-adjusted mid price.

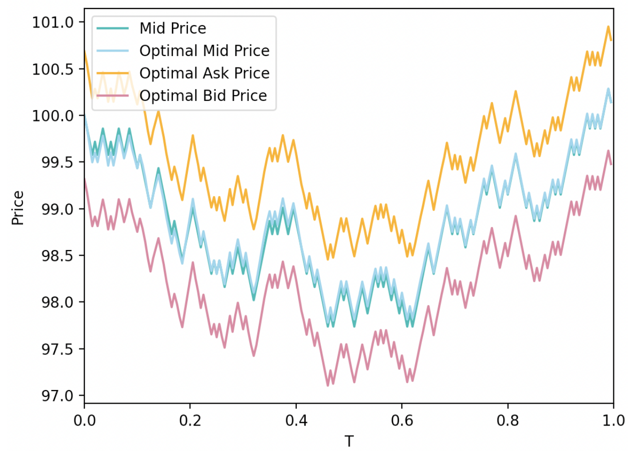

Figure 1: OB Mid Price vs. Optimal Bid Price vs. Optimal Mid Price vs. Optimal Ask Price

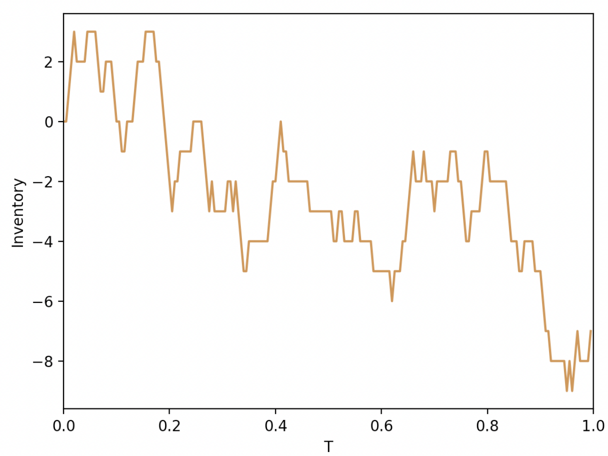

Figure 2: Inventory Trace

In fig. 1, we notice that $P_t^{(a)^*}$ and $P_t^{(b)^*}$ move following the pattern of optimal mid price. We further compare with fig. 2 and find that if the optimal mid price is less than S_t, we incline to sell than buy a share of stock. The reason is that S_t is higher than the price we expected and we follow the key rule - buy low & sell high.

We then study the PnL and inventory metrics. In table 1, we observe that naive strategy gives higher PnL. However, the standard deviation of inventory and PnL for naive strategy are extremely higher than that for optimal strategy. We also show comparisons between the two policies for final PnL, final inventory, total hits and total lifts. The variance of these metrics for naive strategy are all higher. Thus, we conclude that optimal strategy is better than naive strategy because of the stability.

Strategy	Average Spread	Final PnL Mean	Final PnL Std	Final Inventory Mean	Final Inventory Std
Optimal Strategy	1.49	67.69	295.49	-0.03	2.95
Naive Strategy	1.49	74.56	837.17	-0.06	8.37

Table 1: $\gamma = 0.1$

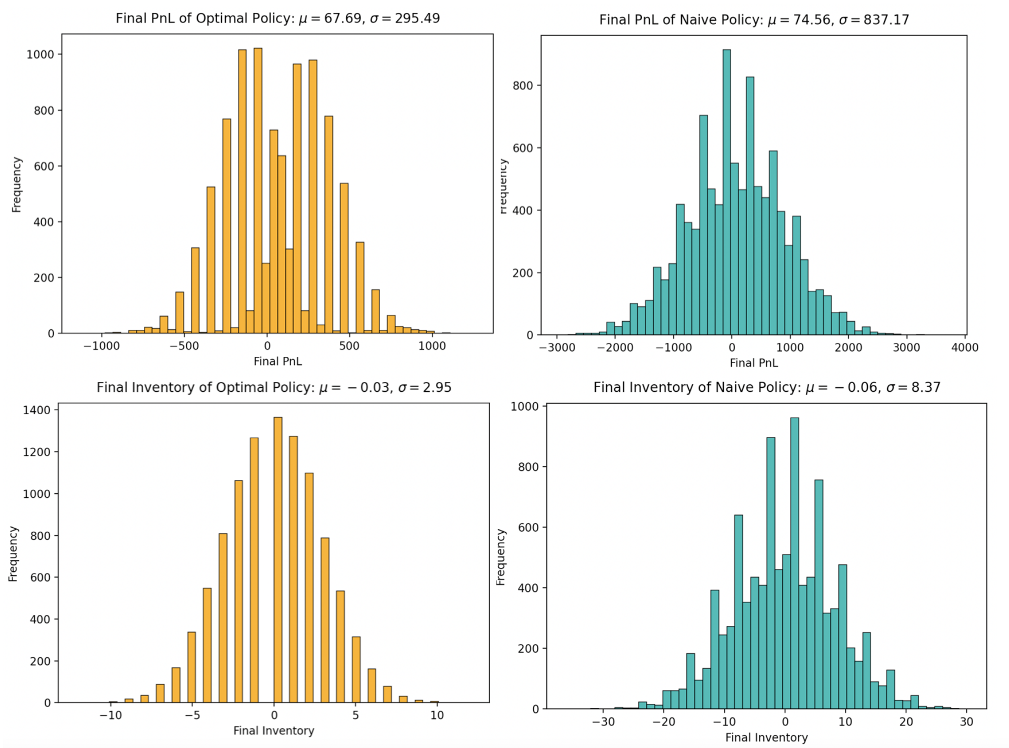

Figure 3: Final PnL & Final Inventory ($\gamma = 0.1$)

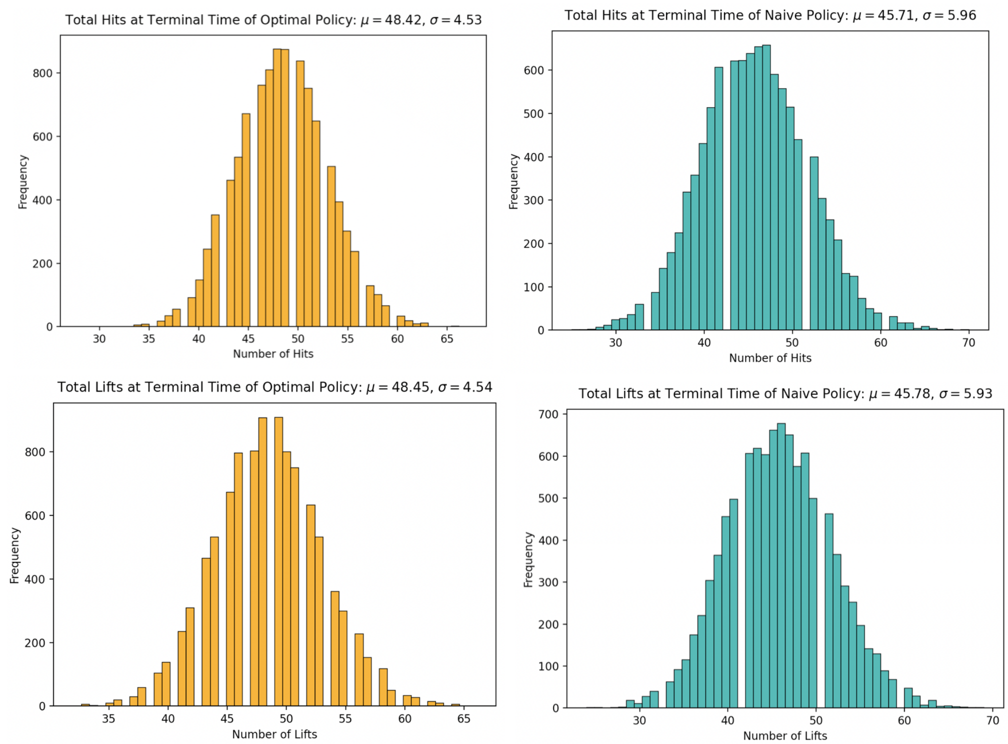

Figure 4: Total Hits & Total Lifts ($\gamma = 0.1$)

We then test for different γ. γ is the degree of risk aversion. Higher γ means the investor is more risk-averse and would not want to take risk. Therefore, the standard deviation of final PnL and final inventory gets lower when the investor is more risk-averse.

Strategy	Average Spread	Final PnL Mean	Final PnL Std	Final Inventory Mean	Final Inventory Std
Optimal Strategy	1.35	56.94	527.01	0.11	5.27
Naive Strategy	1.35	71.17	865.39	-0.03	8.65

Table 2: $\gamma = 0.01$

Strategy	Average Spread	Final PnL Mean	Final PnL Std	Final Inventory Mean	Final Inventory Std
Optimal Strategy	3.03	33.99	162.62	-0.03	1.63
Naive Strategy	3.03	49.53	512.62	-0.06	5.13

Table 3: $\gamma = 1$

References

[1] A. Rao, T. Jelvis. *Foundations of Reinforcement Learning with Applications in Finance*

[2] M. Avellaneda, S. Stoikov *High-frequency trading in a limit order book* In Quantitative Finance, Vol. 8, No. 3, April 2008, 217–224.