数理解析研究所講究録
第 1933 巻 2015 年 70-88

Filtering Model for Order Book Dynamics

Hideyuki Takada
Department of Information Science
Faculty of Science
Toho University

Abstract

We propose a stochastic filtering model for limit order book in order to visualize the activity of the market, in another words, selling or buying tendencies in the market. Stochastic dynamics of the limit order book is modeled with a queuing system where incoming orders and cancellations of existing orders arrive according to independent Poisson processes. We suppose that the dynamics of the limit order book is driven by some unobservable latent factor processes. Observing the history of the order book dynamics, the distribution of the latent factor can be derived by solving filtering equation. For example, numerical results with Nikkei 225 futures are presented.

Keywords and phrases: Order book; Markov modulated Poisson process; Branching particle filter
2010 Mathematics Subject Classification codes: 91G40, 91G60.

1 Introduction

Stochastic dynamics of the limit order book has currently addressed with the growing availability of ultra high frequency data record. In fact, understanding the limit order book dynamics would provide effective strategy to save the transaction costs for investors and also provide liquidity at efficient price to market makers.

In this paper, stochastic dynamics of the limit order book is modeled with a queuing system where incoming orders and cancellations of existing orders arrive according to independent Poisson processes as proposed by Cont et al. [3]. In their approach, Cont et al. [3] assumed the technical requirement that limit orders arrive in unit size. Subsequent paper of Huang and Kercheval [6] generalized the model to allow various order sizes. These papers mainly address the probabilities of mid price move which are theoretically captured as the first passage time of the birth-death process.

In our approach, we are interested in drivers which may lying behind the observed order flows. We suppose that the dynamics of the limit order book is driven by some unobservable latent factor processes that represent the activity of the market, in another words, selling or buying tendency of the market. It would be appear that a lot of market participants mount or cancel their orders based on the knowledge of historical order flows including current state of the order book. Therefore it would be supported to build the model within the framework of the stochastic filtering; market participants update their belief about distribution of the latent factors by use of observable data. This enable us to estimate the degree of market activities and further the arrival rates of future market orders would be evaluated.

This paper is organized as follows. Section 2 introduces our model and states some terminology to describe the stochastic dynamics of the order book. Most of Section 2 is devoted to the appropriate modeling of the arrival intensities of the limit/market orders as a function of the latent factors. In section 3, we derive the normalized filter via the linearization method. Numerical calculation of the normalized filter is achieved by particle filter and the detail of the algorithm is described in Section 4. Section 5 provides the numerical example for high frequency order book data of Nikkei 225 Futures.

2 Model Setting

2.1 State of the Limit Order Book (LOB)

We formulate the stochastic dynamics of the limit order book in the context of a high-dimensional queuing system as first introduced in Cont et al. [3]. In contrast to this paper, we do not impose the restriction that all incoming and outgoing orders are of size one.

It is assumed that limit orders and market orders can be placed on a fixed price grid $\mathcal{P} = \{1, 2, \cdots, n\}$ representing multiples of a price tick. Here we assume that the tick size of the limit order book is denoted by δ. For example, in case of the Nikkei225 Future, we pick up the potential range of the prices $\{8000, 8010, 8020, \cdots, 17980, 17990, 18000\}$ and label these grid points as $1, 2, \cdots, n$. So the tick size is δ=10 Yen. The upper boundary n is chosen large enough so that it is highly unlikely that orders at prices higher than n will be placed within the time frame of our analysis[1]. State of the order book at time t is described by the continuous time ($\mathbb{R}^n$-dim) stochastic process $Z(t) = (Z_1(t), Z_2(t), \cdots, Z_n(t))$, where the p-th element $Z_p(t)$ denotes the time t order size waiting for the future market order to be matched. If $Z_p(t) > 0$ then $Z_p(t)$ represents the size of limit sell orders and if $Z_p(t) < 0$ then $Z_p(t)$ represents the size of limit buy orders at time t respectively.

The best-ask price at time t is then defined by $P_{sell}(t) = \inf\{p \in \mathcal{P} | Z_p(t) > 0\} \wedge (n+1)$ and similarly the best-bid price is defined by $P_{buy}(t) = \sup\{p \in \mathcal{P} | Z_p(t) < 0\} \vee 0$. Furthermore, we define the number of outstanding sell orders at a distance k (equivalently $k \cdot \delta$ in price) from the best bid price as $Q^{k\delta}_{sell}(t) = Z_{P_{buy}(t)+k\delta}(t) > 0$. Thus the quantity $Q^{\delta}_{sell}(t)$ indicates the number of orders of best ask at time t. Similarly the number of outstanding buy orders at a distance k from the best ask price is defined as $Q^{k\delta}_{buy}(t) = Z_{P_{sell}(t)-k\delta}(t) < 0$.

2.2 Stochastic Dynamics of the Order Book

Uncertainty is modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t \geq 0}$ that describes the information flow over time. We impose two additional technical conditions, often called the usual conditions. The first is that $\mathcal{F}_t$ is right-continuous and the second is that $\mathcal{F}_0$ contains all $\mathbb{P}$ -null sets, meaning that one can always identify a sure event.

We slightly extend the model proposed by Cont et al. [3] by introducing the size of limit/market orders. The stochastic dynamics of the order book is described as follows.[2]

- a limit sell order of j units at price level $p > P_{buy}(t)$ increases the quantity at price p as

$$Z_p(t + \Delta t) = Z_p(t) + j \tag{1}$$

- a limit buy order of j units at price level $p < P_{sell}(t)$ increases the quantity at price p as

$$Z_p(t + \Delta t) = Z_p(t) - j \tag{2}$$

[1] As described in Cont et al. [3], since the model is intended to to be used on the time scale of days, this finite boundary assumption is reasonable.

[2] Suppose that the small time interval Δt represents 20 millisecond, which is the unit time in our model.

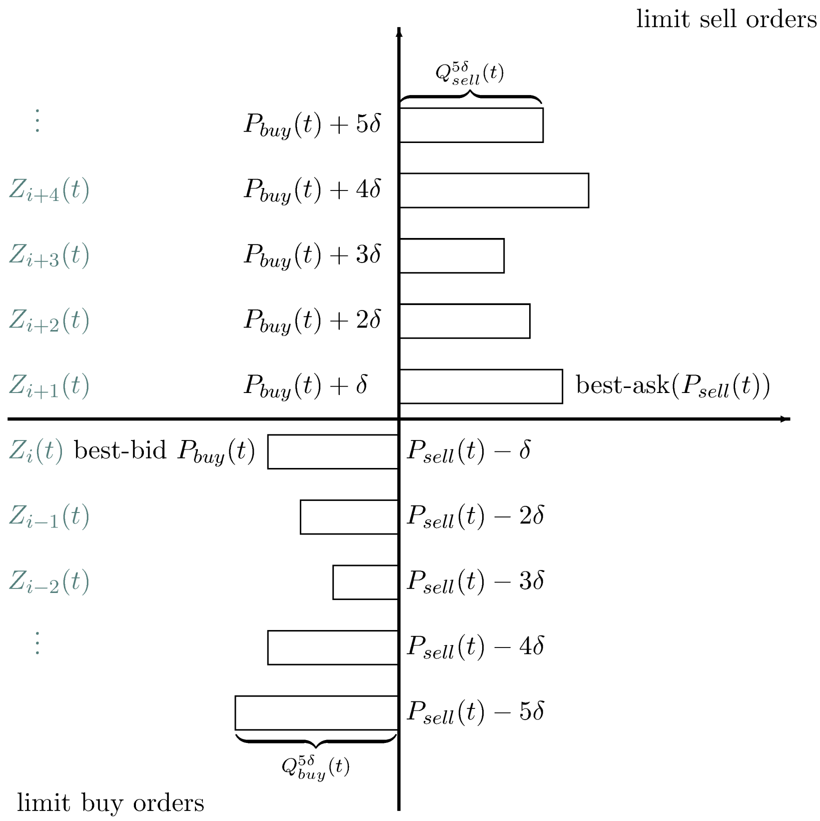

Figure 1: The shape of the Limit Order Book at time t

- a cancellation of an outstanding limit sell order of j units at price level $p > P_{buy}(t)$ decreases the quantity at price p as

$$Z_p(t+\Delta t) = Z_p(t) - j \tag{3}$$

- a cancellation of an outstanding limit buy order of j units at price level $p < P_{sell}(t)$ decreases the quantity at price p as

$$Z_p(t+\Delta t) = Z_p(t) + j \tag{4}$$

- a market buy order of $j < Q^{\delta}_{sell}(t)$ units decreases the quantity at the best ask price

$$Z_{P_{sell}(t)}(t+\Delta t) = Z_{P_{sell}(t)}(t) - j \tag{5}$$

  If $Q^{\delta}_{sell}(t) < j < Q^{2\delta}_{sell}(t)$, $Z_{P_{sell}(t)+\delta}(t+\Delta t) = Z_{P_{sell}(t)+\delta}(t) - (j - Q^{\delta}_{sell}(t))$ and mid price goes up.

- a market sell order of $j < Q^{\delta}_{buy}(t)$ units decreases the quantity at the best bid price

$$Z_{P_{buy}(t)}(t+\Delta t) = Z_{P_{buy}(t)}(t) + j \tag{6}$$

  If $Q^{\delta}_{buy}(t) < j < Q^{2\delta}_{buy}(t)$, $Z_{P_{buy}(t)+\delta}(t+\Delta t) = Z_{P_{buy}(t)+\delta}(t) - (j - Q^{\delta}_{buy}(t))$ and mid price goes down.

Thus the time evolution of the order book is driven by the incoming flow of limit orders, cancellations and market orders, each of which can be represented as a counting process with mark, where the mark represents the size of orders.

Assumption 2.1. *Filtrations generated by observed limit orders are denoted by* $\mathcal{G}_t = \sigma(Z(s) : 0 \le s \le t)$.

Furthermore we assume that the arrival rate of orders described above are modulated by some latent factors[3].

2.3 Latent Factor

Let $X(t) = (X_{sell}(t), X_{buy}(t))$, representing the activity of the market, be a 2-dimensional unobservable state process (signal process) governed by the next stochastic differential equations

$$\begin{aligned} dX_{sell}(t) &= -c_{sell}X_{sell}(t)dt + \sigma_{sell}dB_1(t), \\ dX_{buy}(t) &= -c_{buy}X_{buy}(t)dt + \sigma_{buy}dB_2(t), \\ dB_1(t)dB_2(t) &= \rho dt, \end{aligned}$$

where $c_{sell}, \sigma_{sell}, c_{buy}, \sigma_{buy} \in \mathbb{R}_+, \rho \in \mathbb{R}$, and $B_1(t)$ and $B_2(t)$ are $\mathbb{P}$-Brownian motion. More precisely, $X_{sell}(t)$ and $X_{buy}(t)$ represent potential demands for selling and buying the security at time t with respectively. Then we can translate the role of $X(t) = (X_{sell}(t), X_{buy}(t))$ as follows.

- Process $X_{sell}(t)$ represents selling tendency, i.e., increasing of $X_{sell}(t)$ would mean overheating of sell orders and cancelation of the sell orders are relatively low.
- Process $X_{buy}(t)$ represents buying tendency, i.e., increasing of $X_{buy}(t)$ would mean overheating of buy orders and then cancelation of the buy orders are relatively low.
- If both increases the market is active while both decreases the market is quiet.

Remark 2.2. *We defined* $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{F}_t^X$ *and then* $X(t)$ *is* $\mathcal{F}_t$*-measurable while is not* $\mathcal{G}_t$*-measurable.*

Thus the filtering problem with point process observations arise. That is, public investors want to estimate $\hat{\lambda}_t = \mathbb{E}_{\mathbb{P}}[\lambda(X(t))|\mathcal{G}_t]$, $\hat{\nu}_t = \mathbb{E}_{\mathbb{P}}[\nu(X(t))|\mathcal{G}_t]$, and $\hat{\mu}_t = \mathbb{E}_{\mathbb{P}}[\mu(X(t))|\mathcal{G}_t]$ for ask side and bid side.

Assumption 2.3. *For given state* $X(t) = (X_{sell}(t), X_{buy}(t))$ *at time* t*, the events such that the arrivals and cancellations of the limit/market orders are modeled with conditionally independent exponentially distributed inter arrival times.*

2.4 Detailed Descriptions of the Dynamics of Order Book

2.4.1 Arrivals of the Limit Sell Orders (ask side)

We first model the stochastic dynamics of $Z_p(t) > 0$, which are the state of the ask side of the order book. We assume that the arrival rate of limit sell order of size j to the price p, where $p > P_{buy}(t)$, at time t is given by

$$\lambda_{ask}(j, p\,|X(t)) = \bar{\lambda}_{ask}(\,p - P_{buy}(t), X_{sell}(t), X_{buy}(t)\,) \times J_{\lambda}^{ask}(\,j \mid X_{sell}(t), X_{buy}(t)\,). \quad (7)$$

[3]The shape of the LOB are renewed at random times which are modeled by a compound Poisson process.

Here the function $\bar{\lambda}_{ask} : \mathbb{Z}_+ \times \mathbb{R}^2 \to \mathbb{R}_+$ represents the arrival rate (arrival intensity) of limit sell order and the function $J^{ask}_{\lambda}(\,\cdot\,|x_1, x_2) : \mathbb{Z}_+ \to \mathbb{R}_+$ represents the conditional probability that the incoming limit sell order has j unit of amounts given that the limit sell order is posted to the price level p at time t. Both the arrival intensity and conditional probability of the order size depend on the latent factor $X(t) = (X_{sell}(t), X_{buy}(t))$.

Functional form of $\bar{\lambda}_{ask}$ According to empirically observed facts presented in Zovko and Farmer [10] and Bouchaud et al. [1], order arrival rates depend on the distance to the bid/ask in such a way that most orders are placed close to the current price. In addition, frequency of incoming limit sell orders would get higher in selling pressure market and then $\bar{\lambda}_{ask}$ should be the increasing function of $X_{sell}(t)$. Next we consider whether $\bar{\lambda}_{ask}$ is an increasing function or a decreasing function of $X_{buy}(t)$. If investors who wish to sell, come to recognize that the buying tendency is high in the market, they may place limit sell orders with the expectation of smooth execution, or they may hold on their attitude with the expectation to sell at a higher price catching the bull market consensus. In this study, we assume that the arrival rate of the limit sell orders $\bar{\lambda}_{ask}$ could be an increasing function of X_{buy}, but its impact(sensitivity) is smaller than that of X_{sell}. This is achieved by the inequality $\beta > \gamma$ in definition of arrival intensity bellow.

$$\bar{\lambda}_{ask}(\ell, x_1, x_2) = \frac{\exp(\beta x_1 + \gamma x_2)}{(\ell/\delta)^{\alpha}}, \quad \alpha, \beta, \gamma > 0. \tag{8}$$

Thus $\bar{\lambda}_{ask}\big(\ell, X(t)\big)$ determines when the next limit sell order of price level at a distance ℓ ticks from the best bid will occur and $J_{\lambda}(\,j \mid X(t))$ determines the order size of this limit sell order.

Functional form of J^{ask}_{λ} In general, limit orders tend to arrive with the minimum unit or at most severalfold, and huge amount of limit orders rarely come with one shot[4]. Therefore, we assume that the conditional probability $J_{\lambda}(j)$ is a monotonically decreasing function with respect to j for fixed $X(t)$ and here we take $J^{ask}_{\lambda}(j) \propto ge^{-g \cdot j}$ as a candidate of the functional form (see Figure 2). More explicitly, we choose

$$J^{ask}_{\lambda}(j) = \frac{ge^{-g \cdot j}}{\sum_{k=1}^{\infty} ge^{-g \cdot k}}, \quad g \geq 0, \tag{9}$$

with some g, which will be defined later as a function of $X_{sell}(t) - X_{buy}(t)$. When the net selling tendency gets large, that is, $X_{sell}(t) - X_{buy}(t)$ gets large, it would be appear that the amount of the limit sell order would increase. On the other hand, when $X_{sell}(t) - X_{buy}(t)$ gets small, the amount of the limit sell order would decrease.

In order to determine whether the function g is increasing function or decreasing function with respect to $X_{sell}(t) - X_{buy}(t)$, let us consult the Figure 2. For example, when $g = 0.12$, the shape of the function $j \mapsto ge^{-g \cdot j}$ is shown by the solid line and $g = 0.03$ case is plotted by the dashed line. When the net selling tendency $X_{sell}(t) - X_{buy}(t)$ is relatively large, the amount of the limit sell orders would increase and then such a situation corresponds to the dashed line in Figure 2 rather than the solid line. Thus we can conclude that it should be better to consider that g is a decreasing function of $X_{sell}(t) - X_{buy}(t)$. Put it all together, we assume that

$$J^{ask}_{\lambda}(j|x_1, x_2) \;=\; \frac{g(x_1 - x_2)\exp(-g(x_1 - x_2) \cdot j)}{\sum_k g(x_1 - x_2)\exp(-g(x_1 - x_2) \cdot k)}, \tag{10}$$

[4]In fact, big investors tend to mount their limit orders of relatively large size but round number of order such as 500 or 1000, and so on. In this paper, we consider such phenomena as an irregular event and do not pay special attention to this. One of the typical technique to take account this phenomena would be a mixture of two distributions as proposed in Kato, Takada and Ogihara, "Empirical analysis of limit order books and power laws in financial markets" also appear in this RIMS Kôkyûroku.

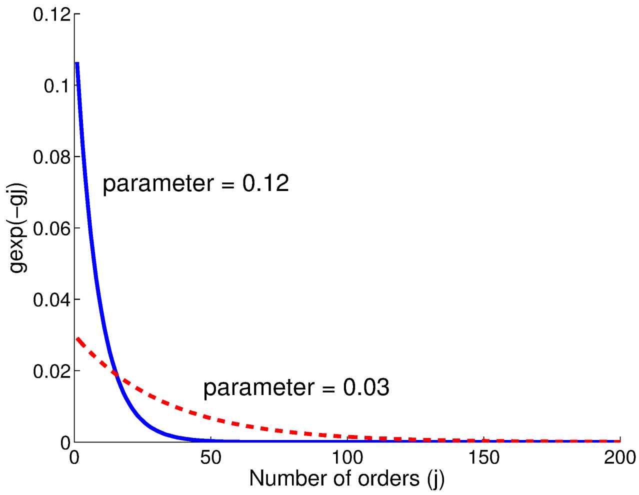

Figure 2: The shape of the function $j \mapsto g\exp(-gj)$ for $g = 0.03$ and 0.12.

where the function $g : \mathbb{R} \to \mathbb{R}_+$ is decreasing with respect to x. In practice, we simply chose $g(x) = g_0 \exp(-g_1 \cdot x), g_0, g_1 \in \mathbb{R}_+$ for tractability.

2.4.2 Cancellation of the Limit Sell Orders

We assume that the limit sell orders of size j at the price $p > P_{buy}(t)$ are canceled with rate

$$\nu_{ask}(j, p \mid X(t)) = \bar{\nu}_{ask}\big(p - P_{buy}(t), X_{sell}(t), X_{buy}(t)\big) \times J^{ask}_{\nu}\big(j \mid X_{sell}(t), X_{buy}(t)\big) \qquad (11)$$

and $\bar{\nu}_{ask}$ is assumed to be

$$\bar{\nu}_{ask}(\ell, x_1, x_2) = \frac{\exp(-\kappa x_1 + \gamma x_2)}{(\ell/\delta)^{\alpha}}, \ \alpha > 0, \kappa > 0. \qquad (12)$$

Functional form of J^{ask}_{ν} Next, we formulate $J^{ask}_{\nu}(j)$, which represents the conditional probability that j unit of the limit sell orders are cancelled given that the cancellation is placed to the price level p.

$$J^{ask}_{\nu}(j) = \frac{he^{-h \cdot j}}{\sum_{k=1}^{K} he^{-h \cdot k}}, \quad h \geq 0. \qquad (13)$$

Here we suppose that the amount of the limit sell orders waiting at the price p just before the cancellation is K units. By the similar argument of the previous subsection, we assume that

$$J^{ask}_{\nu}(j|x_1, x_2) = \frac{h(x_1 - x_2)\exp(-h(x_1 - x_2) \cdot j)}{\sum_{k=1}^{K} h(x_1 - x_2)\exp(-h(x_1 - x_2) \cdot k)} \qquad (14)$$

and conclude h is a increasing function with respect to $X_{sell}(t) - X_{buy}(t)$. In practice, we take $h(x) = h_0 \exp(h_1 \cdot x), h_0, h_1 \in \mathbb{R}_+$ for tractability.

Remark 2.4. *If $g = -h$ then $J^{ask}_{\lambda} = J^{ask}_{\nu}$.*

2.4.3 Execution of the Market Buy Orders

As many practitioners are confronted (facing), one can hardly distinguish the two events; (i) *execution of the market buy orders* and (ii) *cancellation of the limit sell order placed in the best ask price* by observing the time series of order book combined with the executed price tick data. Hence, in our study, we assume the following.

Assumption 2.5. *Decrements in $Q^{\delta}_{sell}(t)$ is the consequence of the execution of the market buy orders.*[5]

Under the above assumption, we assume that the execution of market sell order of j unit occur with rate

$$\mu_{ask}(j \mid X(t)) = \bar{\lambda}_{bid}\big(10, X_{sell}(t), X_{buy}(t)\big) \times J^{bid}_{\lambda}\big(\, j \mid X_{sell}(t), X_{buy}(t)\big), \tag{15}$$

where $\bar{\lambda}_{bid}$ and J^{bid}_{λ} would be described explicitly in the next section formulating the arrival rate of the limit buy order.

2.4.4 Arrivals of the Limit Buy Orders (bid side)

Next we model the stochastic dynamics of $Z_p(t) < 0$, which are the bid side of the order book. We assume that the arrival rate of limit buy order of size j to the price p, where $p < P_{sell}(t)$, at time t is given by

$$\lambda_{bid}(j, p \mid X(t)) \;=\; \bar{\lambda}_{bid}\big(\, P_{buy}(t) - p, X_{sell}(t), X_{buy}(t)\,\big) \times J^{bid}_{\lambda}\big(\, j \mid X_{sell}(t), X_{buy}(t)\,\big), \tag{16}$$

where $\bar{\lambda}_{bid} : \mathbb{Z}_+ \times \mathbb{R}^2 \to \mathbb{R}_+$ represents the arrival rate (arrival intensity) of limit buy order and defined as

$$\bar{\lambda}_{bid}(\ell, x_1, x_2) = \frac{\exp(\gamma x_1 + \beta x_2)}{(\ell/\delta)^{\alpha}}, \quad \alpha, \beta, \gamma > 0. \tag{17}$$

And the function $J^{bid}_{\lambda}(\,\cdot\, |x_1, x_2) : \mathbb{Z}_+ \to \mathbb{R}_+$ represents the conditional probability that the incoming limit buy order has j unit given that the order is posted to the price level p. Thus we assume

$$J^{bid}_{\lambda}(j|x_1, x_2) \;=\; \frac{u(x_2 - x_1)\exp(-u(x_2 - x_1)\cdot j)}{\sum_{k=1}^{\infty} u(x_2 - x_1)\exp(-u(x_2 - x_1)\cdot k)}, \tag{18}$$

with $u(x) = u_0 \exp(-u_1 \cdot x), u_0, u_1 \in \mathbb{R}_+$, decreasing function of x.

Remark 2.6. *If $g = -u$ then $J^{ask}_{\lambda} = J^{bid}_{\lambda}$.*

2.4.5 Cancellation of the Limit Buy Orders

We assume that the limit buy orders of size j at the price $p < P_{sell}(t)$ are cancelled with rate

$$\nu_{bid}(j, p \mid X(t)) = \bar{\nu}_{bid}\big(P_{sell}(t) - p, X_{sell}(t), X_{buy}(t)\big) \times J^{bid}_{\nu}\big(\, j \mid X_{sell}(t), X_{buy}(t)\big), \tag{19}$$

where $\bar{\nu}_{bid} : \mathbb{Z}_+ \times \mathbb{R}^2 \to \mathbb{R}_+$ represents the arrival rate of cancelation and defined as

$$\bar{\nu}_{bid}(\ell, x_1, x_2) = \frac{\exp(\gamma x_1 - \kappa x_2)}{(\ell/\delta)^{\alpha}}, \quad \alpha, \gamma, \kappa > 0. \tag{20}$$

[5] In order to eliminate the simultaneous events in our model formulation.

And the function $J_{\nu}^{bid}(\ \cdot\ |x_1, x_2) : \mathbb{Z}_+ \to \mathbb{R}_+$ represents the conditional probability that the amount of the cancel has j unit given that the cancelation is occurred to the price level p.

$$J_{\nu}^{bid}(\ j \mid x_1, x_2) \ = \ \frac{v(x_2 - x_1)\exp(-v(x_2 - x_1)\cdot j)}{\sum_{k=1}^{K} v(x_2 - x_1)\exp(-v(x_2 - x_1)\cdot k)}, \tag{21}$$

with $v(x) = v_0 \exp(v_1 \cdot x), v_0, v_1 \in \mathbb{R}_+$, increasing function of x. Here, K denotes the amount of the limit sell orders just before cancellation.

2.4.6 Execution of the Market Sell Orders

For essentially the same reason as Assumption 2.5, we assume the following.

Assumption 2.7. *Increments in $Q_{buy}^{\delta}(t)$ is the consequence of the execution of the market sell orders.*

Under the above assumption, we assume that the execution of market sell order of j unit occur with rate

$$\mu_{bid}(j\ |X(t)) = \bar{\lambda}_{ask}\big(\delta, X_{sell}(t), X_{buy}(t)\big) \times J_{\lambda}^{ask}\big(\ j \mid X_{sell}(t), X_{buy}(t)\big). \tag{22}$$

Remark 2.8. *If $u = -v$ then $J_{\lambda}^{bid} = J_{\nu}^{bid}$, and If $h = -v$ then $J_{\nu}^{ask} = J_{\nu}^{bid}$.*

Above discussions are summarized as bellow.

$$\begin{aligned}
\lambda_{ask}(j,p|x_1,x_2) &= \frac{\exp(\beta x_1 + \gamma x_2)}{((p - P_{buy})/\delta)^{\alpha}} \times \frac{g(x_1 - x_2)e^{-g(x_1-x_2)\cdot j}}{\sum_{k=1}^{\infty} g(x_1 - x_2)e^{-g(x_1-x_2)\cdot k}}, \\
g(x) &= g_0 \exp(-g_1 \cdot x), \\
\nu_{ask}(j,p|x_1,x_2) &= \frac{\exp(-\kappa x_1 + \gamma x_2)}{((p - P_{buy})/\delta)^{\alpha}} \times \frac{h(x_1 - x_2)e^{-h(x_1-x_2)\cdot j}}{\sum_{k=1}^{\infty} h(x_1 - x_2)e^{-h(x_1-x_2)\cdot k}}, \\
h(x) &= h_0 \exp(h_1 \cdot x), \\
\mu_{ask}(j,p|x_1,x_2) &= \exp(\gamma x_1 + \beta x_2) \quad \times \frac{u(x_2 - x_1)e^{-u(x_2-x_1)\cdot j}}{\sum_{k=1}^{K} u(x_2 - x_1)e^{-u(x_2-x_1)\cdot k}}, \\
u(x) &= u_0 \exp(-u_1 \cdot x),
\end{aligned}$$

$$\begin{aligned}
\lambda_{bid}(j,p|x_1,x_2) &= \frac{\exp(\gamma x_1 + \beta x_2)}{((P_{buy} - p)/\delta)^{\alpha}} \times \frac{u(x_2 - x_1)e^{-u(x_2-x_1)\cdot j}}{\sum_{k=1}^{\infty} u(x_2 - x_1)e^{-u(x_2-x_1)\cdot k}}, \\
\nu_{bid}(j,p|x_1,x_2) &= \frac{\exp(\gamma x_1 - \kappa x_2)}{((P_{sell} - p)/\delta)^{\alpha}} \times \frac{v(x_2 - x_1)e^{-v(x_2-x_1)\cdot j}}{\sum_{k=1}^{\infty} v(x_2 - x_1)e^{-v(x_2-x_1)\cdot k}}, \\
v(x) &= v_0 \exp(v_1 \cdot x), \\
\mu_{bid}(j,p|x_1,x_2) &= \exp(\beta x_1 + \gamma x_2) \quad \times \frac{g(x_1 - x_2)e^{-g(x_1-x_2)\cdot j}}{\sum_{k=1}^{K} g(x_1 - x_2)e^{-g(x_1-x_2)\cdot k}}.
\end{aligned}$$

3 Non-linear filtering

Filtrations generated by observed limit orders are denoted by $\mathcal{G}_t = \sigma(Z(s) : 0 \le s \le t)$, which is available to market participants. Gloval history is denoted by $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{F}_t^X$ where $\mathcal{F}_t^X = \sigma(X(s) : 0 \le s \le t)$. We suppose that $X(t)$ is not directly observable for market participants and that the available information is specified by the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ generated by the history of $Z(t)$. Thus the filtering problem with counting process observations arise. The next aim is to estimate

$$\hat{\lambda}_t = \mathbb{E}_{\mathbb{P}}[\lambda(X(t))|\mathcal{G}_t], \;\; \hat{\nu}_t = \mathbb{E}_{\mathbb{P}}[\nu(X(t))|\mathcal{G}_t], \text{ and } \;\; \hat{\mu}_t = \mathbb{E}_{\mathbb{P}}[\mu(X(t))|\mathcal{G}_t] \tag{23}$$

for both ask side and bid side in a recursive form.

Definition 3.1. *We define the observation processes (counting processes) as follows.*

- $Y_\lambda^{ask}(j;t)$ *counts the number of times that the size j limit sell order arrived during the time interval* $[0,t]$
- $Y_\lambda^{bid}(j;t)$ *counts the number of times that the size j limit buy order arrived during the time interval* $[0,t]$
- $Y_\nu^{ask}(j;t)$ *counts the number of times that the size j limit sell order canceled during the time interval* $[0,t]$
- $Y_\nu^{bid}(j;t)$ *counts the number of times that the size j limit buy order canceled during the time interval* $[0,t]$
- $Y_\mu^{ask}(j;t)$ *counts the number of times that the size j market sell order executed during the time interval* $[0,t]$
- $Y_\mu^{bid}(j;t)$ *counts the number of times that the size j market buy order executed during the time interval* $[0,t]$

Assumption 3.2. *For given state $X(t)$, the limit/market orders form conditionally independent exponentially distributed inter arrival times. More precisely, expressions of the counting processes for limit sell/buy order, cancellation of limit sell/buy order and market sell/buy order are given as follows.*

$$\begin{aligned}
Y_\lambda^{ask}(j;t) &= N^j\Big(\int_0^t \lambda_{ask}(j,\ell(s-),X(s-))ds\Big),\\
Y_\lambda^{bid}(j;t) &= N^j\Big(\int_0^t \lambda_{bid}(j,\ell(s-),X(s-))ds\Big),\\
Y_\nu^{ask}(j;t) &= N^j\Big(\int_0^t \nu_{ask}(j,\ell(s-),X(s-))ds\Big),\\
Y_\nu^{bid}(j;t) &= N^j\Big(\int_0^t \nu_{bid}(j,\ell(s-),X(s-))ds\Big),\\
Y_\mu^{ask}(j;t) &= N^j\Big(\int_0^t \mu_{ask}(j,X(s-))ds\Big),\\
Y_\mu^{bid}(j;t) &= N^j\Big(\int_0^t \mu_{bid}(j,X(s-))ds\Big),
\end{aligned}$$

where $N^j(t)$ are mutually independent $\mathbb{P}$-Standard Poisson process which are independent of $X(t)$. Assume that $\lambda_{ask}, \lambda_{bid}, \nu_{ask}, \nu_{bid}, \mu_{ask}, \mu_{bid}$ are measurable functions.

Assumption 3.3. *(i) The total intensity $\Lambda(j;\ell,x) = \lambda_{ask}(j;\ell,x) + \lambda_{bid}(j;\ell,x) + \nu_{ask}(j;\ell,x) + \nu_{bid}(j;\ell,x) + \mu_{ask}(j;x) + \mu_{bid}(j;x)$ is uniformly bounded from below and above.*
(ii) Partial derivative of total intensity $\Lambda(j;\ell,x)$ with respect to x, and partial derivatives of $J_\lambda and J_\nu$ with respect to x are bounded and continuous in x.

We introduce an equivalent measure $\mathbb{Q}$ under which processes $Y_\lambda^{ask}(j;t)$, $Y_\lambda^{bid}(j;t)$, $Y_\nu^{ask}(j;t)$, $Y_\nu^{bid}(j;t)$, $Y_\mu^{ask}(j;t)$ and $Y_\mu^{bid}(j;t)$ are all Standard Poisson. Let us define $L(t)$ as

$$\begin{aligned}
L(t) = 1 \; + \; & \sum_j \int_0^t [\lambda_{ask}(j;\ell(s-),X(s-)) - 1]L(s-)(dY_\lambda^{ask}(j;s) - ds) \\
& + \sum_i \int_0^t [\lambda_{bid}(j;\ell(s-),X(s-)) - 1]L(s-)(dY_\lambda^{bid}(j;s) - ds) \\
& + \sum_i \int_0^t [\nu_{ask}(j;\ell(s-),X(s-)) - 1]L(s-)(dY_\nu^{ask}(j;s) - ds) \\
& + \sum_i \int_0^t [\nu_{bid}(j;\ell(s-),X(s-)) - 1]L(s-)(dY_\nu^{bid}(j;s) - ds) \\
& + \sum_i \int_0^t [\mu_{ask}(j;X(s-)) - 1]L(s-)(dY_\mu^{ask}(j;s) - ds) \\
& + \sum_i \int_0^t [\mu_{bid}(j;X(s-)) - 1]L(s-)(dY_\mu^{bid}(j;s) - ds). \qquad (24)
\end{aligned}$$

The solution of the above SDE can be represented by

$$\begin{aligned}
L(t) = & \prod_j \exp\left(\int_0^t \log\lambda_{ask}(j;\ell(s-),X(s-))dY_\lambda^{ask}(j;s) - \int_0^t [\lambda_{ask}(j;\ell(s-),X(s-)) - 1]ds\right) \\
\times & \prod_j \exp\left(\int_0^t \log\lambda_{bid}(j;\ell(s-),X(s-))dY_\lambda^{bid}(j;s) - \int_0^t [\lambda_{bid}(j;\ell(s-),X(s-)) - 1]ds\right) \\
\times & \prod_j \exp\left(\int_0^t \log\nu_{ask}(j;\ell(s-),X(s-))dY_\nu^{ask}(j;s) - \int_0^t [\nu_{ask}(j;\ell(s-),X(s-)) - 1]ds\right) \\
\times & \prod_j \exp\left(\int_0^t \log\nu_{bid}(j;\ell(s-),X(s-))dY_\nu^{bid}(j;s) - \int_0^t [\nu_{bid}(j;\ell(s-),X(s-)) - 1]ds\right) \\
\times & \prod_j \exp\left(\int_0^t \log\mu_{ask}(j;X(s-))dY_\mu^{ask}(j;s) - \int_0^t [\mu_{ask}(j;X(s-)) - 1]ds\right) \\
\times & \prod_j \exp\left(\int_0^t \log\mu_{bid}(j;X(s-))dY_\mu^{bid}(j;s) - \int_0^t [\mu_{bid}(j;X(s-)) - 1]ds\right).
\end{aligned}$$

Here we define the new probability measure $\mathbb{Q}$ as

$$L(t) = \frac{d\mathbb{P}}{d\mathbb{Q}}\Big|_t$$

then one sees that under $\mathbb{Q}$, the counting processes $Y_\lambda^{ask}(j;t)$, $Y_\lambda^{bid}(j;t)$, $Y_\nu^{ask}(j;t)$, $Y_\nu^{bid}(j;t)$, $Y_\mu^{ask}(j;t)$ and $Y_\mu^{bid}(j;t)$ are Standard Poisson and furthermore we can choose $\mathbb{Q}$ so as to these six processes are independent of $X(t)$.

3.1 Un-normalized filter

Definition 3.4. *For an arbitrary bonded function f, let $\rho_t(f)$ be the conditional expectation of $f(X(t))L(t)$ given $\mathcal{F}_t^Y$ as follows.*

$$\rho_t(f) = \mathbb{E}_{\mathbb{Q}}\big[f(X(t))L(t)|\mathcal{F}_t^Y\big]$$

We derive filtering equation via reference measure approach as in Zeng [11]. Applying integration by parts, we have

$$f(X(t))L(t) = f(X(0))L(0) + \int_0^t L(s-)df(X(s)) + \int_0^t f(X(s-))dL(s) + [f(X), L]_t,$$

where $[\cdot,\cdot]_t$ denote the quadratic covariation process. Since we assumed that $X(t)$ and $Y(t)$ are independent under $\mathbb{P}$, which implies $f(X(t))$ and $L(t)$ have no simultaneous jumps with probability 1, it then follows $[f(X), L]_t = 0$. Let $\mathcal{M}_t^f$ be a $(\mathbb{P}, \mathbb{F})$-martingale specified with

$$\mathcal{M}_t^f := f(X(t)) - f(X(0)) - \int_0^t \mathcal{A}f(X(s))ds. \tag{25}$$

Then from (24)[6] and (25), we have

$$\begin{aligned} f(X(t))L(t) = f(X(0))L(0) &+ \int_0^t L(s-)d\mathcal{M}_s^f + \sum_j \int_0^t f(X(s-))\Big(\Lambda(j; X(s-)) - 1\Big)L(s-)dY(j;s) \\ &+ \int_0^t L(s-)\Big\{\mathcal{A}f(X(s)) - \sum_j f(X(s))\Big(\Lambda(j; X(s-)) - 1\Big)\Big\}ds. \end{aligned}$$

By taking conditional expectations with respect to the reference measure $\mathbb{Q}$ given the filtration $\mathcal{F}_t^Y$ on both sides,

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\Big[f(X(t))L(t)\Big|\mathcal{F}_t^Y\Big] &= f(X(0))L(0) + \sum_j \mathbb{E}_{\mathbb{Q}}\Big[\int_0^t f(X(s-))\Big(\Lambda(j; X(s-)) - 1\Big)L(s-)dY(j;s)\Big|\mathcal{F}_t^Y\Big] \\ &+ \mathbb{E}_{\mathbb{Q}}\Big[\int_0^t L(s-)\Big\{\mathcal{A}f(X(s)) - \sum_j f(X(s))\Big(\Lambda(j; X(s-)) - 1\Big)\Big\}ds\Big|\mathcal{F}_t^Y\Big]. \end{aligned}$$

With the same arguments developed in the Appendix A in Zeng [11], we can exchange the order of integration and expectation of the last term as shown bellow.

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\Big[f(X(t))L(t)\Big|\mathcal{F}_t^Y\Big] &= f(X(0))L(0) + \sum_j \int_0^t \mathbb{E}_{\mathbb{Q}}\Big[f(X(s-))\Big(\Lambda(j; X(s-)) - 1\Big)L(s-)\Big|\mathcal{F}_t^Y\Big]dY(j;s) \\ &+ \int_0^t \mathbb{E}_{\mathbb{Q}}\Big[L(s-)\Big\{\mathcal{A}f(X(s)) - \sum_j f(X(s))\Big(\Lambda(j; X(s-)) - 1\Big)\Big\}\Big|\mathcal{F}_t^Y\Big]ds. \end{aligned}$$

[6] In a SDE form,

$$\begin{aligned} dL(t) = &\sum_j [\lambda_{ask}(j; \ell(t-), X(t-)) - 1]L(t-)(dY_\lambda^{ask}(j;t) - dt) \\ &+ \sum_i [\lambda_{bid}(j; \ell(t-), X(t-)) - 1]L(t-)(dY_\lambda^{bid}(j;t) - dt) \\ &+ \sum_i [\nu_{ask}(j; \ell(t-), X(t-)) - 1]L(t-)(dY_\nu^{ask}(j;t) - dt) \\ &+ \sum_i [\nu_{bid}(j; \ell(t-), X(t-)) - 1]L(t-)(dY_\nu^{bid}(j;t) - dt) \\ &+ \sum_i [\mu_{ask}(j; X(t-)) - 1]L(t-)(dY_\mu^{ask}(j;t) - dt) \\ &+ \sum_i [\mu_{bid}(j; X(t-)) - 1]L(t-)(dY_\mu^{bid}(j;t) - dt). \end{aligned}$$

Then we can conclude

$$\rho_t(f) = \rho_0(f) + \sum_j \int_0^t \rho_{s-}\big(f(\cdot)(\Lambda(j;\cdot)-1)\big)dY(j;s) + \int_0^t \rho_s\Big(\mathcal{A}f(\cdot) - \sum_j f(\cdot)(\Lambda(j;\cdot)-1)\Big)ds. \tag{26}$$

In the full notation, (26) is

$$\begin{aligned}\rho_t(f) = \rho_0(f) + \int_0^t \rho_s(\mathcal{A}f)ds &+ \sum_j \int_0^t \rho_{s-}((\lambda_{ask}(j)-1)f)(dY_\lambda^{ask}(j;s)-ds)\\ &+ \sum_j \int_0^t \rho_{s-}((\lambda_{bid}(j)-1)f)(dY_\lambda^{bid}(j;s)-ds)\\ &+ \sum_j \int_0^t \rho_{s-}((\nu_{ask}(j)-1)f)(dY_\nu^{ask}(j;s)-ds)\\ &+ \sum_j \int_0^t \rho_{s-}((\nu_{bid}(j)-1)f)(dY_\nu^{bid}(j;s)-ds)\\ &+ \sum_j \int_0^t \rho_{s-}((\mu_{ask}(j)-1)f)(dY_\mu^{ask}(j;s)-ds)\\ &+ \sum_j \int_0^t \rho_{s-}((\mu_{bid}(j)-1)f)(dY_\mu^{bid}(j;s)-ds),\end{aligned}$$

where $\lambda_{ask}(j) = \lambda_{ask}(j;p,x)$ and so on.

3.2 Normalized filter

Definition 3.5. *For an arbitrary bonded function f, let $\pi_t(f)$ be the conditional expectation of $f(X(t))$ given $\mathcal{F}_t^Y$.*

$$\pi_t(f) = \mathbb{E}_{\mathbb{P}}\big[f(X(t))|\mathcal{F}_t^Y\big]$$

By the Bayes formula, one sees that

$$\mathbb{E}_{\mathbb{P}}[f(X(t))|\mathcal{F}_t^Y] = \frac{\mathbb{E}_{\mathbb{Q}}[f(X(t))L(t)|\mathcal{F}_t^Y]}{\mathbb{E}_{\mathbb{Q}}[L(t)|\mathcal{F}_t^Y]},$$

implying that $\pi_t(f)$ can be obtained by normalizing $\rho_t(f)$ with $\rho_t(1)$. Our next aim is to derive the equation governing the evolution of $\pi_t(f)$.

Theorem 3.6 (Kushner-Stratonovich equation). *Let $\mathcal{A}$ be the infinitesimal generator of the* *process $X(t)$. Then $\pi_t(f)$ is the unique solution of the following SDE.*

$$
\begin{aligned}
\pi_t(f) &= \pi_0(f) + \int_0^t \pi_s(\mathcal{A}f)ds \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\lambda_{ask}(j)f\big)}{\pi_{s-}\big(\lambda_{ask}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\lambda^{ask}(j;s) - \pi_s(\lambda_{ask}(j))ds \Big) \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\lambda_{bid}(j)f\big)}{\pi_{s-}\big(\lambda_{bid}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\lambda^{bid}(j;s) - \pi_s(\lambda_{bid}(j))ds \Big) \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\nu_{ask}(j)f\big)}{\pi_{s-}\big(\nu_{ask}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\nu^{ask}(j;s) - \pi_s(\nu_{ask}(j))ds \Big) \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\nu_{bid}(j)f\big)}{\pi_{s-}\big(\nu_{bid}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\nu^{bid}(j;s) - \pi_s(\nu_{bid}(j))ds \Big) \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\mu_{ask}(j)f\big)}{\pi_{s-}\big(\mu_{ask}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\mu^{ask}(j;s) - \pi_s(\mu_{ask}(j))ds \Big) \\
&+ \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\mu_{bid}(j)f\big)}{\pi_{s-}\big(\mu_{bid}(j)\big)} - \pi_{s-}(f) \right] \Big(dY_\mu^{bid}(j;s) - \pi_s(\mu_{bid}(j))ds \Big) \qquad (27)
\end{aligned}
$$

Proof. The proof of the theorem follows in the same way as Zeng [11], applying Ito's formula to $f(X,Y) = \frac{X}{Y}$ with $X = \rho_t(f)$ and $Y = \rho_t(1)$. □

In the short notation, (27) is expressed as

$$
\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(\mathcal{A}f)ds + \sum_j \int_0^t \left[\frac{\pi_{s-}\big(\Lambda(j)f\big)}{\pi_{s-}\big(\Lambda(j)\big)} - \pi_{s-}(f) \right] \Big(dY(j;s) - \pi_s(\Lambda(j))ds \Big).
$$

And this equation can be split into following two equations corresponding to whether the new limit/market order arrived or not at time t.

Between orders: $t \in [\tau_n, \tau_{n+1})$

$$
\pi_t(f) = \pi_{\tau_n}(f) + \int_{\tau_n}^t \pi_s(\mathcal{A}f)ds - \sum_j \int_{\tau_n}^t \Big[\pi_{s-}\big(\Lambda(j)f\big) + \pi_s(\Lambda(j))\pi_{s-}(f) \Big] ds.
$$

At order: τ_n

$$
\pi_{\tau_n}(f) - \pi_{\tau_n-}(f) = \sum_j \left[\frac{\pi_{\tau_n-}\big(\Lambda(j)f\big)}{\pi_{\tau_n-}\big(\Lambda(j)\big)} - \pi_{\tau_n-}(f) \right].
$$

3.3 The linearization method

Further we need to examine the behavior of the filter π_t for $t \in [\tau_n, \tau_{n+1})$, which is governed by

$$
\pi_t(f) = \pi_{\tau_n}(f) + \int_{\tau_n}^t \pi_s(\mathcal{A}f)ds - \sum_j \int_{\tau_n}^t \Big[\pi_{s-}\big(\Lambda(j)f\big) + \pi_s(\Lambda(j))\pi_{s-}(f) \Big] ds. \qquad (28)
$$

Because of the cross term $\pi_s(\Lambda(j))\pi_{s-}(f)$, (28) is a nonlinear equation generally difficult to handle. Therefore we introduce a linearized equation and then finally obtain the solution of (28). The following discussions are based on Kliemann et al. [8] and Ceci and Gerardi [2].

Lemma 3.7. *For $n \in \mathbb{N}$, let ρ_t^n be a solution of*

$$\rho_t^n(f) = \pi_{\tau_n}(f) + \int_{\tau_n}^t \rho_s^n(\mathcal{A}f)ds - \sum_j \int_{\tau_n}^t \rho_{s-}^n(\Lambda(j)f)ds. \tag{29}$$

Then

$$\pi_t(f) = \frac{\rho_t^n(f)}{\rho_t^n(1)}$$

solves equation (28) for $t \in [\tau_n, \tau_{n+1})$.

Proof. The proof follows in the same way as Ceci and Gerardi [2]. □

Proposition 3.8.

$$\Psi_t(x)(f) = \mathbb{E}\Big[f(X(t))\exp\Big(-\int_0^t \Lambda(X(s))ds\Big)\Big]$$

A solution of (29) is given by

$$\rho_t^n(f) = \int_{\mathbb{R}^2} \Psi_t(x)(f)\pi_{\tau_n}(dx)$$

Proof. The proof follows in the same way as Ceci and Gerardi [2]. □

Thus we obtain for $t \in [T_{n-1}, T_n]$,

$$\pi_t(f) = \frac{\int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}}_{(T_{n-1},x)}\left[f(X_{t-T_{n-1}})\exp\left(-\int_0^{t-T_{n-1}} \Lambda(j;X_s)ds\right)\right]\pi_{T_{n-1}}(dx)}{\int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}}_{(T_{n-1},x)}\left[\exp\left(-\int_0^{t-T_{n-1}} \Lambda(j;X_s)ds\right)\right]\pi_{T_{n-1}}(dx)}, \tag{30}$$

where $\pi_{T_{n-1}}(dx)$ is given below in (31). At the limit/market order renewed time T_n, we have

$$\pi_{T_n}(f) = \frac{\int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}}_{(T_{n-1},x)}\left[f\left(X_{T_n-T_{n-1}}\right)\Lambda(j;X_{T_n-T_{n-1}})\exp\left(-\int_0^{T_n-T_{n-1}} \Lambda(j;X_s)ds\right)\right]\pi_{T_{n-1}}(dx)}{\int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}}_{(T_{n-1},x)}\left[\Lambda(j;X_{T_n-T_{n-1}})\exp\left(-\int_0^{T_n-T_{n-1}} \Lambda(j;X_s)ds\right)\right]\pi_{T_{n-1}}(dx)}. \tag{31}$$

4 How to solve the filtering equations

It is necessary to approximate the expected values appeared in equations (30) and (31) so as to numerically compute the filter $\pi_t(f)$. Particle filter is a method to approximate the conditional distribution $\mathbb{P}(X_t \in \bullet | \mathcal{G}_t)$ with some suitable discrete random measures of the form

$$\mathbb{P}(X_t \in \bullet\ | \mathcal{G}_t) \approx \sum_p \eta_t^p \delta_{x_t^p}(\bullet)$$

with some sample points $\{x_t^p\}$ and their consistent stochastic "weights" $\{\eta_t^p\}$. Here, $\delta_x(\bullet)$ is the Dirac measure. In this section, we summarize the numerical algorithm for computing filters $\pi_t(f)$ in case of $f(x) = x$ via a branching particle system as in Frey and Runggaldier [5] and Del Moral and Miclo [4].

4.1 Particle filter

The particle system is constructed over time steps $\{t_k = k\Delta\}_{k\in\mathbb{N}}$ with the sequence of occupation measures $\{\tilde{\pi}_{t_k}\}_{k=1,2,\cdots}$ approximating the conditional distributions $\pi_{X_{t_k}|\mathcal{G}_{t_k}} := \mathbb{P}(X_{t_k} \in \bullet \mid \mathcal{G}_{t_k})$ for each time step. As the filtering equations are represented in recursive form, the occupation measure $\tilde{\pi}_{t_k}$ is computed from $\tilde{\pi}_{t_{k-1}}$ and similar procedures are repeated for subsequent time steps. Let $\boldsymbol{x}_k = (x_k^1, x_k^2, \cdots, x_k^p, \cdots, x_k^{n_k})$ denote the set of n_k particles at time t_k living in the state space of X.

Algorithm 4.1 (Branching Particle System). *In order to derive the discrete filter distribution $\{\tilde{\pi}_{t_k}\}_{k=0,1,\ldots}$, we have to follow the several steps:*

Step 0. Initialization

Set the number n_0 of initial particles, a discretized time step size Δ.

Step 1. Initial discrete distribution

The initial discrete distribution $\tilde{\pi}_0$ is given by the occupation measure of n_0 particles of mass $1/n_0$, that is, $\tilde{\pi}_0 = \frac{1}{n_0}\sum_{i=1}^{n_0} \delta_{x^i(0)}$. Here $\boldsymbol{x}_0 = (x^1(0), x^2(0), \cdots, x^{n_0}(0))$ represents independent draws from the initial distribution $\pi_0 := \mathbb{P}(X_0 \in \bullet)$.

Step 2. Prediction stage

Given the particles $\boldsymbol{x}_k = (x_k^1, x_k^2, \cdots, x_k^{n_k})$ at time t_k, generate n_k independent trajectories $(\mathbf{X}_s)_{0\le s\le\Delta} = \left((X_s^1)_{0\le s\le\Delta}, (X_s^2)_{0\le s\le\Delta}, \cdots, (X_s^{n_k})_{0\le s\le\Delta}\right)$ such that, for each $p \in \{1, \cdots, n_k\}$, X_s^p starts from x_k^p at time 0 and then follows the SDE of X_t.

Step 3. Updating stage

Compute $\tilde{w}_{k+1}^p$ for each X_s^p obtained in the last prediction stage as

$$\tilde{w}_{k+1}^p = \exp\left(-\int_0^\Delta \sum_j \Lambda(j; X_s^p)ds\right).$$

Whether some event occurred during $(t_k, t_{k+1}]$ or not, we redefine w_{k+1}^p as follows:

$$w_{k+1}^p := \begin{cases} \tilde{w}_{k+1}^p & \text{if there is no LOB renewals during } (t_k, t_{k+1}], \\ \lambda_{ask}(j;\ \cdot\ , X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an limit sell order of } j \text{ units renewed at } t_{k+1}, \\ \nu_{ask}(j;\ \cdot\ , X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an limit sell order of } j \text{ units canceled at } t_{k+1}, \\ \mu_{ask}(j; X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an market buy order of } j \text{ units executed at } t_{k+1}, \\ \lambda_{bid}(j;\ \cdot\ , X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an limit buy order of } j \text{ units renewed at } t_{k+1}, \\ \nu_{bid}(j;\ \cdot\ , X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an limit buy order of } j \text{ units canceled at } t_{k+1}, \\ \mu_{bid}(j; X_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an market buy order of } j \text{ units executed at } t_{k+1}, \end{cases}$$

Now we compute $v_{k+1}^p = \frac{n_k w_{k+1}^p}{\sum_{q=1}^{n_k} w_{k+1}^q}$ for every $p = 1, \cdots, n_k$ and obtain $\{o_{k+1}^p\}_{p=1,\cdots,n_k}$

$$o_{k+1}^p = \begin{cases} [v_{k+1}^p] & \text{with probability } 1 + [v_{k+1}^p] - v_{k+1}^p \\ [v_{k+1}^p] + 1 & \text{with probability } v_{k+1}^p - [v_{k+1}^p] \end{cases}$$

where $[v]$ *stands for the integer part of* v.
Denote by $n_{k+1} = \sum_{p=1}^{n_k} o_{k+1}^p$ *the total number of particles at time* t_{k+1}. *Then each particle* x_k^p *at time* t_k *independently generates* o_{k+1}^p *offsprings of* $(X_s^p)_{0 \le s \le \Delta}$ *starting at* x_k^p *for every* $p = 1, \cdots, n_k$ *and denote by* $\boldsymbol{x}_{k+1} = (x_{k+1}^1, x_{k+1}^2, \cdots, x_{k+1}^{n_{k+1}})$ *all the realized random samples of* $\mathbf{X}_\Delta$. *Thus one can achieve the approximated discrete distribution at* t_{k+1} *as*

$$\tilde{\pi}_{t_{k+1}} = \frac{1}{n_{k+1}} \sum_{p=1}^{n_{k+1}} \delta_{x_{k+1}^p}.$$

Step 4.

Proceed from k *to* $k+1$ *and go to* **Step 2** *until some time horizon.*

In Step 3, the updating stage, each particle is replaced by the particles of which the number is randomly given by o^p. This procedure is worked in a consistent manner; particles with small weights w^p have almost zero offspring while those with large weights are replaced by several offspring. We mention that most of the calculation time with our algorithm is caused by sampling of the random number o^p in the updating stage.

5 Numerical example

In this section we illustrate the numerical results based on the tick data of Nikkei 225 futures as of June 1, 2012. In order to pursue the particle filter algorithm, we take $n_0 = 1000$ and $\Delta = 20$ milliseconds. Figure 3 shows the historical data of the mid price as of 2012/6/1.

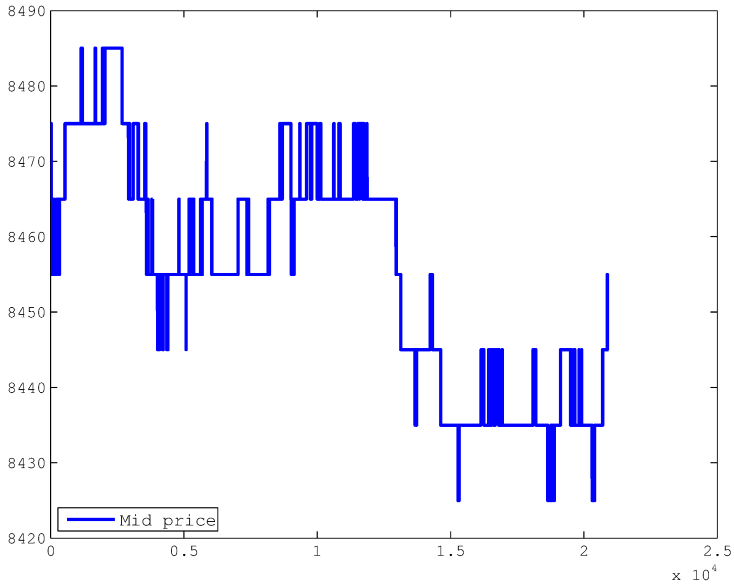

Figure 3: Transitions of the Mid price as of 2012/6/1

Figure 4 and 5 illustrates the transitions of the percentile points of the occupation measure $\tilde{\pi}_t$ of $X_{sell}(t)$ and $X_{buy}(t)$ respectively, calculated every 20 milliseconds from 9:00 am, opening time, to 3:10 pm, closing time in Osaka stock market.

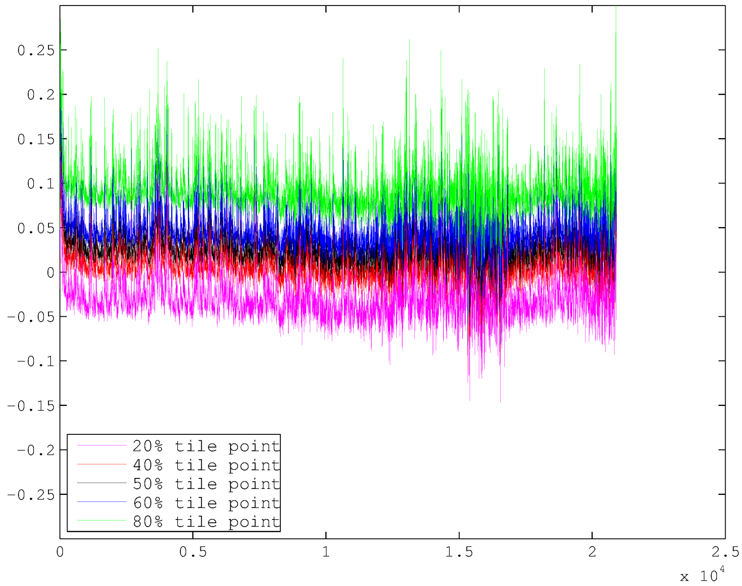

Figure 4: Transitions of the filter $\tilde{\pi}_t$ of $X_{sell}(t)$ calculated for intraday data as of 2012/6/1

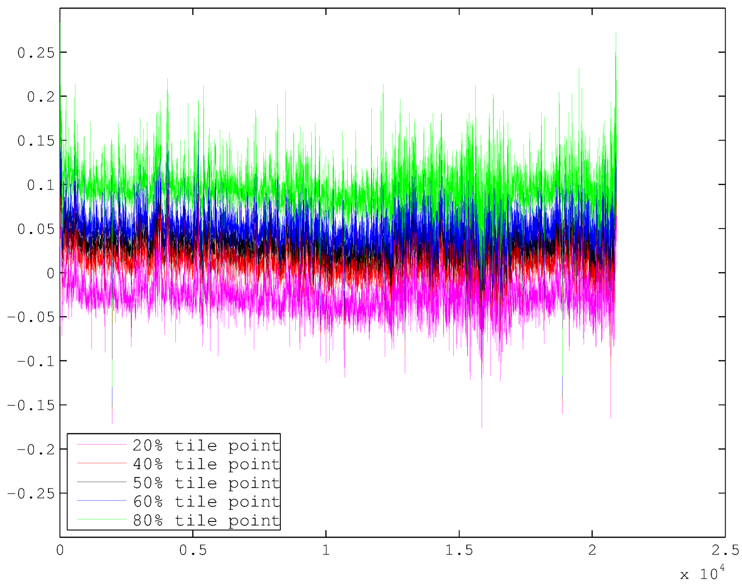

Figure 5: Transitions of the filter $\tilde{\pi}_t$ of $X_{buy}(t)$ calculated for intraday data as of 2012/6/1

Estimated distribution of $X(t) = (X_{sell}(t), X_{buy}(t))$ seems to fluctuate heavily but in average we can see the location of the distribution. In order to look into the detail, we focus on the first 20 minutes to see how transitions of $X(t) = (X_{sell}(t), X_{buy}(t))$ are related to the mid price movements. Figure 6 and 7 illustrates the transitions of the percentile points of the occupation measure $\tilde{\pi}_t$ of $X_{sell}(t)$ and $X_{buy}(t)$ respectively from 9:00 am to 9:20 am. Roughly, when mid price move to downward, X_{sell} tends to jump up and X_{buy} tends to jump down as expected.

However, these are not always true because $X(t) = (X_{sell}(t), X_{buy}(t))$ is sensitive to overall dynamics of the order book.

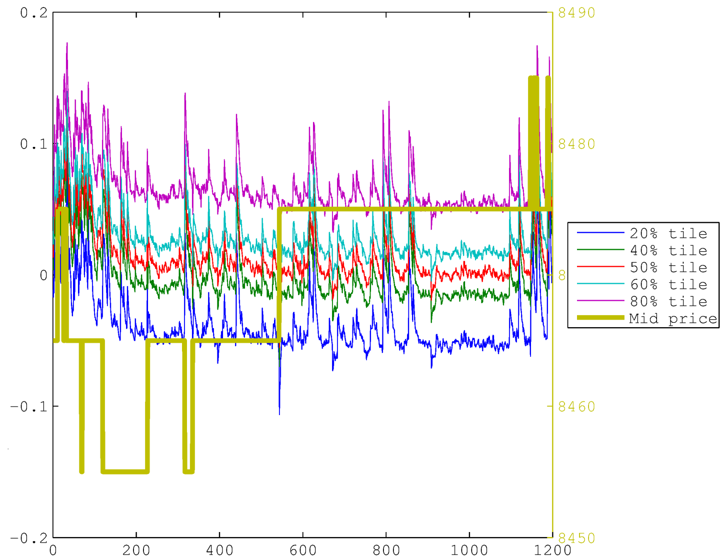

Figure 6: Transitions of $X_{sell}(t)$ from 9:00 am to 9:20 on 2012/6/1

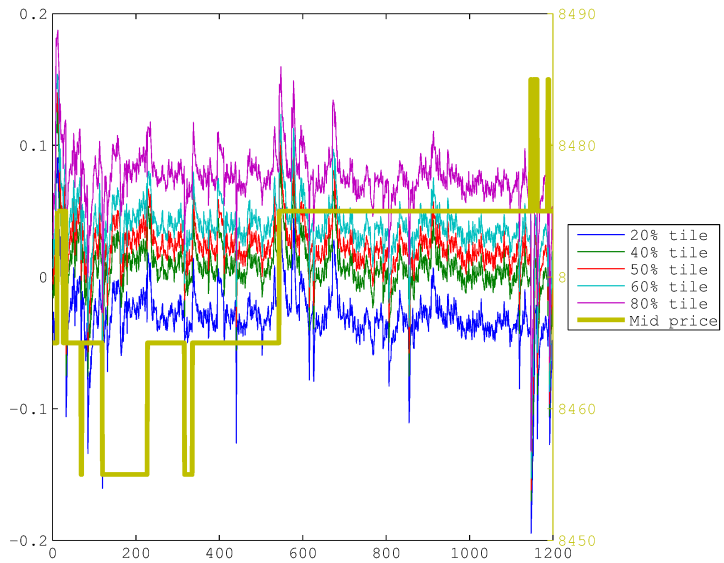

Figure 7: Transitions of $X_{buy}(t)$ from 9:00 am to 9:20 on 2012/6/1

Acknowledgements The author is grateful to participants of the FMA2014 for valuable comments.

References

[1] Bouchaud, J.-P., M., Mezard. and M. Potters, 2002, "Statistical properties of stock books: Empirical results and models", *Quantitative Finance*, **2**, pp.251-256.

[2] Ceci, C. and A. Gerardi, 2006, "A Model for High Frequency Data under Partial Information: A Filtering Approach", *International Journal of Theoretical and Applied Finance*, **9**(4), 555-576.

[3] Cont, R., Stoikov, S. and Talreja, R, 2010, "A Stochastic Model for Order Book Dynamics" *Operations Research*, **58**, No. 3, 549-563.

[4] Del Moral, P. and L. Miclo, 2000, "Branching and Interacting Particle Systems. Approximations of Feynman-Kac Formulae with Applications to Non-linear Filtering", *Seminaire de probabilities (Strasbourg)*, tome 34, 1-145.

[5] Frey, R. and W. Runggaldier, 2010, "Pricing Credit Derivatives under Incomplete Information: a Nonlinear-Filtering Approach", *Finance and Stochastics*, **14**(4), 495-526.

[6] Huang, He and A. H. Kercheval, 2012, "A Generalized Birth-Death Stochastic Model for High-Frequency Order Book Dynamics", *Quantitative Finance*, **12** (4), 547-557.

[7] Frey, R. and T. Schmit, 2012, "Pricing and Hedging of Credit Derivatives via the Innovations Approach to Nonlinear Filtering", *Finance and Stochastics*, **16**(1), 105-133.

[8] Kliemann, W. H., G. Kock, and F. Marchetti, 1990, "On the Unnormalized Solution of the Filtering Problem with Counting Process Observations", *IEEE Transactions and Information Theory*, **36**(6), 1415 - 1425.

[9] Xiong, J. and Y. Zeng, 2011, "A Branching Particle Approximation to a Filtering Micromovement Model of Asset Price", *Statistical Inference for Stochastic Processes*, **14**, 111-140.

[10] Zovko, I. and J. D. Farmer, 2002, "The power of patience; A behavioral regularity in limit order placement", *Quantitative Finance*, **2**, pp.387-392.

[11] Zeng, Y., 2003, "A Partially-Observed Model for Micro-movement of Asset Prices with Bayes Estimation via Filtering", *Mathematical Finance*, **13**, 411-444.

Department of Information Science
Faculty of Science
Toho University, Miyama 2-2-1, Funabashi, Chiba 274-8510, Japan
E-mail address: `hideyuki.takada@is.sci.toho-u.ac.jp`

東邦大学・理学部　高田 英行