**IMS Collections**
**Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz**
Vol. 4 (2008) 123–136

DOI: 10.1214/074921708000000345

# A Class of Multivariate Micromovement Models of Asset Price and Their Bayesian Model Selection via Filtering*

**Laurie C. Scott**[1] **and Yong Zeng**[1,†]

*University of Missouri–Kansas City*

**Abstract:** A filtering model with counting process observations has been recently developed as a reasonable framework for the micromovement of asset price. In this paper, we first highlight such an extension to multivariate case for modeling multi-stocks and related results on Bayes estimation via filtering. For this rich class of multivariate models, we develop the Bayesian model selection using Bayes factor. Based on the unnormalized, Duncan-Mortensen-Zakai-like filtering equation, we derive a system of SPDE characterizing the evolution of the Bayes factors and prove their uniqueness. Furthermore, applying Kushner's Markov chain approximation method, we propose a numerical scheme to derive recursive algorithms, and we prove the consistency (or robustness) of the recursive algorithms.

## 1. Introduction

Micro-movements of asset price are referred to transactional price behavior in contrast to the macro-movements referring to daily, weekly, or monthly closing price behavior. Such micromovement data are called *ultra high frequency data* by Engle [4]. There are strong connections as well as striking distinctions between the macro- and micro-movements. Zeng [17] proposed a general class of Filtering Micro-movement models (FM models, as we simply call it) that bridge the gap between the macro- and micro- movements caused by noise, and amalgamate the sample characteristics of micro-movement and macro-movement of price in a consistent manner. The main appeal of the proposed model is that prices are viewed as a collection of counting processes, each of which represents a price level. Then, the model is framed as a filtering problem with counting process observations. Alternatively, the price process can be constructed from the intrinsic value process by incorporating the trading noise. Zeng [17] also developed continuous-time Bayes parameter estimation via filtering for the model and Kouritzin and Zeng [9] further developed Bayesian model selection via filtering for the class of models based on Bayes factor.

The multivariate extension of FM model and its Bayes estimation via filtering is studied in Scott and Zeng [14]. In this paper, we study the model selection of this class of multivariate FM model. Model selection, a significant and persistent area of research, evaluates which of competing models best fits transaction price data.

*This work was supported in part by NSF grant DMS-0604722.

†Yong is grateful to Tom Kurtz for his inspiration, insight, encouragement and support during Yong's Ph.D study under Tom and along Yong's academic career. We also thank an anonymous referee.

[1]Department of Mathematics and Statistics, University of Missouri–Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA, e-mail: lcscott@dstsystems.com; zeng@mendota.umkc.edu; url: http://mendota.umkc.edu

*Keywords and phrases:* Bayes factor, filtering, high frequency data, counting process, model selection

Moreover, to the degree that economic theory can be modeled statistically, model selection provides a powerful tool for testing the economic theories related to market microstructure. The Bayesian approach offers a general methodology for hypothesis testing and model selection based on Bayes factor with many merits. See Kass and Raftery [8], a survey of Bayes factor in both methodology and applications. Here, we also adopt the Bayesian approach for the model selection of the class of multivariate FM models.

Kouritzin and Zeng [9] introduce a two-step approach to calculate the Bayes factor. The first step is to derive the system of two stochastic differential equations (SDEs) that govern the evolution of the Bayes factors of Models 1 vs. 2 and that of Models 2 vs. 1. The second step is to apply the Markov chain approximation method to the system of SDEs to develop a recursive algorithm for computing the Bayes factors. Following this approach, we first derive the system of SDEs governing the evolutions of Bayes factors for the multivariate FM models and prove the uniqueness of the solution. Then, we prove a convergence theorem guaranteeing the consistency (or robustness) of the recursive algorithms generated from the Markov chain approximation method.

The rest of the paper goes as follows. Section 2 highlights the multivariate FM models in two equivalent representations. Section 3 reviews the unnormalized, Duncan Mortensen Zakai-like filtering equation, and the normalized, Kushner Stratonovich (or Fujisaki Kallianpur Kunita)-like filtering equations respectively. Section 4 first reviews the methodology of Bayesian model selection and then derives the system of SDEs for the Bayes factors for the multivariate FM models and prove the uniqueness. Section 5 proves the convergence theorem and provides a numerical scheme to calculate the Bayes factors. We conclude in Section 6.

## 2. Multivariate Micromovement Models

This section presents the class of multivariate FM models in two equivalent ways: one as a collection of counting processes of price levels and the other as a construction of price from the intrinsic value of an asset by incorporating noises. These two representations are equivalent in the sense that they have the same probability distribution, which is proven in [14].

### *2.1. Counting Process Observations*

In real world trading, the price of an asset fluctuates as the result of inflowing information and trading noise. In micromovement level, the price does not move continuously as the common asset price models such as diffusion or jump-diffusion processes suggest, but moves level-by-level due to price discreteness (caused by the minimum price variation set by trading regulation). From this viewpoint, we can formulate the prices of an asset as a collection of counting processes as described below.

Now, suppose that $\vec{\theta}$ is a vector of parameters and $\vec{X}$ is a vector Markov process representing the intrinsic values of $m$ assets, which are not observable directly.

Define $\vec{Y}_j(t)$ for the $j$th asset as

$$\vec{Y}_j(t) = \begin{pmatrix} N_{j,1}(\int_0^t \lambda_{j,1}(\vec{\theta}(s), \vec{X}(s), s)ds) \\ N_{j,2}(\int_0^t \lambda_{j,2}(\vec{\theta}(s), \vec{X}(s), s)ds) \\ \vdots \\ N_{j,n_j}(\int_0^t \lambda_{j,n_j}(\vec{\theta}(s), \vec{X}(s), s)ds) \end{pmatrix}, \tag{2.1}$$

Here $Y_{j,k}(t) = N_{j,k}(\int_0^t \lambda_{j,k}(\vec{\theta}(s), \vec{X}(s), s)ds)$ is the observable counting process recording the cumulative number of trades for the $j$th asset that have occurred at the $k$th price level (denoted by $y_{j,k}$) up to time $t$. For notational convenience we refer to the collection of the counting processes of $m$ assets as

$$\mathbf{Y}(t) = (\vec{Y}_1(t), \vec{Y}_2(t), \ldots, \vec{Y}_m(t)). \tag{2.2}$$

We extend the five assumptions of FM model to multivariate FM ones. These assumptions are needed not only in the initial development of the models and of the related filtering equations, but also in showing the consistency of the recursive algorithms to actually computing the Bayes factors.

**Assumption 2.1.** $N_{j,k}$'s are unit Poisson processes under the physical measure $P$.

**Assumption 2.2.** $(\vec{\theta}, \vec{X}), N_{1,1}, N_{1,2}, \ldots, N_{1,n_1}, N_{2,1}, \ldots, N_{m,n_m}$ are independent under measure P.

**Assumption 2.3.** Each of the total intensity processes, $a_j(\vec{\theta}(t), \vec{X}(t), t)$, is uniformly bounded above, namely, there exist a positive constant, $C$, such that $0 < a_j(\vec{\theta}(t), \vec{X}(t), t) \leq C$ for each $j = 1, 2, \ldots, m$ and for all $t > 0$.

**Remark 2.1.** Assumptions 2.1–2.3 imply the existence of a reference measure $Q$ such that under $Q$, $\vec{X}, \vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_m$ are independent and $Y_{1,1}, \ldots, Y_{1,n_1}$, $Y_{2,1}, \ldots$, $Y_{m,n_m}$ are unit Poisson processes. This reference measure is crucial in deriving the filtering equations and proving the convergence theorem.

**Remark 2.2.** Assumption 2.1–2.3 also imply that $Y_{j,k}(t)$ is a conditional Poisson process and $Y_{j,k}(t) - \int_0^t \lambda_{j,k}(\vec{\theta}(s), \vec{X}(s), s)ds$ is a martingale.

**Assumption 2.4.** The intensity is

$$\lambda_{j,k}(\vec{\theta}(t), \vec{X}(t), t) = a_j(\vec{\theta}(t), \vec{X}(t), t) g_{j,k}(y_{j,k} \mid X_j(t)),$$

where $a_j(\vec{\theta}(t), \vec{X}(t), t)$ is the total trading intensity for the $j$th stock and $g_{j,k}(y_{j,k} \mid X_j(t))$ is the transition probability from $X_j(t)$, the intrinsic value of th $j$th asset, to $y_{j,k}$, the $k$th price level for the $j$th asset at time $t$.

**Remark 2.3.** Assumption 2.4 imposes a desirable structure. Namely, the overall rate of trade occurrence of asset $j$ at time $t$ is determined by the total trading intensity $a_j(\vec{\theta}(t), \vec{X}(t), t)$ and the proportion of trading intensity at the $k$th price level of asset $j$ is determined by $g_{j,k}(y_{j,k} \mid X_j(t))$. The intensity structure plays an essential role in the equivalence of the two approaches of modeling.

Finally, we impose a general assumption on the intrinsic value vector process and the parameter vector process.

**Assumption 2.5.** $(\vec{\theta}, \vec{X})$ is the solution of a martingale problem for a generator $\mathbf{A}$ such that

$$M_f(t) = f(\vec{\theta}(t), \vec{X}(t)) - \int_0^t \mathbf{A} f(\vec{\theta}(s), \vec{X}(s)) ds$$

is a $\mathcal{F}_t^{\vec{\theta},\vec{X}}$-martingale, where $\mathcal{F}_t^{\vec{\theta},\vec{X}}$ is the $\sigma$-field generated by $(\vec{\theta}(s), \vec{X}(s))_{0 \le s \le t}$.

Under this representation, $(\vec{\theta}(t), \vec{X}(t))$ becomes a signal process, which cannot be observed directly, but can be partially observed through the collection of counting processes, $\mathbf{Y}(t)$, corrupted by trading noise, which is modeled by $g_{j,k}(y_{j,k} \mid x_j)$. Hence, $(\vec{\theta}, \vec{X}, \mathbf{Y})$ is framed as a *multivariate filtering problem with counting process observations.*

### *2.2. Price Construction from the Value Process*

More intuitively, we can construct the model for the price behavior of a set of assets by explicitly constructing the observed price of each asset from its intrinsic value. For this method, we assume that the underlying $(\vec{\theta}, \vec{X})$ satisfies Assumption 2.5. Further we assume the prices of asset $j$ are the marks associated with a collection of trading times $t_1, t_2, \ldots$ described by a conditional Poisson process with intensity $a_j(\vec{\theta}(t), \vec{X}(t), t)$. Finally, we model the price of asset $j$ at a trading time $t_{j,i}$ as $Z_j(t_i) = F_j(X_j(t_i))$, where $F_j(\cdot)$ is a random transformation taking the value of the asset to the observed price with the transition probability function $g_{j,k}(y_{j,k} \mid x_j)$.

Under this construction, the impact of $F_j(X_j(t_i))$ is transitory at trading times just as the impact of trading noise. Only the underlying stock value, $X_j(t_i)$, is considered to have permanent influence on the price process. The random transformation, $F(x)$, is flexible. Different examples that accommodate the three well-documented important types of noise: discrete, clustering and non-clustering, can be found in [17], [18], and [15].

## 3. Filtering Equations

This section reviews the foundations of statistical inference for the multivariate FM models, namely, the likelihoods and posteriors and their related filtering equations.

### *3.1. Likelihood Functions and Posteriors*

Let $\mathbf{F}_t^{\mathbf{Y}} = \sigma\{(\mathbf{Y}(s)) \mid 0 \le s \le t\}$ be the $\sigma$-algebra generated by the observed sample path of the price $\mathbf{Y}$ up to time $t$. Here $\mathbf{Y}$ is defined by (2.2)

Now, taken together Assumptions 2.1–2.3, we have the existence of a reference measure $Q$ such that $P$ is absolutely continuous with respect to $Q$. We can write $Q$ as $Q = P_{\vec{\theta}(t),\vec{X}(t)} \times Q_{\mathbf{Y}}$, where $P_{\vec{\theta}(t),\vec{X}(t)}$ is the probability measure generated by $\vec{X}(t)$ and $\vec{\theta}(t)$ and $Q_{\mathbf{Y}}$ is the measure under which the counting processes $Y_{1,1}, \ldots, Y_{1,n_1}, Y_{2,1}, \ldots, Y_{m,n_m}$ are independent unit Poisson processes. This in turn implies that $\hat{Y}_{j,k}(t) = Y_{j,k}(t) - t$ is martingale with respect to $Q$ for each $(j,k)$. Another property under this reference measure is the independence of $(\vec{\theta}, \vec{X})$ and $\mathbf{Y}$.

We can then define our joint likelihood function, $L(t)$, in terms of The Radon-Nikodym derivative (see [2], or by Girsanov–Meyer theorem in [12]):

$$
\begin{aligned}
(3.1) \qquad \frac{dP}{dQ}(t) &= \frac{dP_{\vec{\theta}(t),\vec{X}(t)} \times dP_{\mathbf{Y}|\vec{\theta}(t),\vec{X}(t)}}{dP_{\vec{\theta}(t),\vec{X}(t)} \times dQ_{\mathbf{Y}}}(t) = \frac{dP_{\mathbf{Y}|\vec{\theta}(t),\vec{X}(t)}}{dQ_{\mathbf{Y}}}(t) = L(t) \\
&= \prod_{j=1}^{m}\prod_{k=1}^{n} \exp\bigg[\int_0^t \log(\lambda_{j,k}(\vec{\theta}(s-),\vec{X}(s-),s-))dY_{j,k}(s) \\
&\qquad - \int_0^t [\lambda_{j,k}(\vec{\theta}(s-),\vec{X}(s-),s-)-1]ds\bigg].
\end{aligned}
$$

Or, in stochastic differential equation (SDE) form,

$$
(3.2) \qquad L(t) = 1 + \sum_{j=1}^{m}\sum_{k=1}^{n}\int_0^t [\lambda_{j,k}(\vec{\theta}(s-),\vec{X}(s-),s-)-1]L(s-)d(Y_{j,k}(s)-s).
$$

We need two definitions for the filtering equations.

**Definition 3.1.** $\phi(f,t) = E^Q[f(\vec{\theta}(t),\vec{X}(t))L(t) \mid \mathbf{F}_t^{\mathbf{Y}}]$.

**Remark 3.1.** If $(\vec{\theta}(0),\vec{X}(0))$ is fixed, then the likelihood of $\mathbf{Y}$ is $E^Q[L(t) \mid \mathcal{F}^{\mathbf{Y}}] = \phi(1,t)$. If a prior is assumed on $(\vec{\theta}(0),\vec{X}(0))$ as in Bayesian paradigm, then the integrated (or marginal) likelihood of $\mathbf{Y}$ is also $\phi(1,t)$.

**Definition 3.2.** *Let $\pi_t$ be the conditional distribution of $(\vec{\theta}(t),\vec{X}(t))$ given $\mathbf{F}_t^{\mathbf{Y}}$ over $P$ and $\pi(f,t) = E^P[f(\vec{\theta}(t),\vec{X}(t)) \mid \mathbf{F}_t^{\mathbf{Y}}] = \int f(\vec{\theta},\vec{X})\pi_t(d\vec{\theta},d\vec{X})$.*

**Remark 3.2.** If a prior is assumed on $(\theta(0),X(0))$ as in Bayesian inference, then $\pi_t$ becomes the continuous-time posterior.

### *3.2. Filtering Equations*

The Kallianpur–Striebel formula provides the relationship between $\phi(f,t)$ and $\pi(f,t)$: $\pi(f,t) = \phi(f,t)/\phi(1,t)$. Hence, the equation governing the evolution of $\phi(f,t)$ is called the *unnormalized filtering equation*, and that of $\pi(f,t)$ is called the *normalized filtering equation*.

The filtering equation provides an effective way to characterize $\pi_t$. The following theorem proven in [14] presents both filtering equations.

**Theorem 3.1.** Suppose that $\mathbf{Y}$ is the counting process observations as specified in by (2.2) satisfying Assumptions 2.1 to 2.4 and $(\vec{\theta},\vec{X})$ satisfies Assumption 2.5. Then, for every $t > 0$ and every $f$ in the domain of $\mathbf{A}$, $\phi_t$ is the unique measure-valued solution of the unnormalized filtering equation:

$$
\begin{aligned}
(3.3) \qquad \phi(f,t) &= \phi(f,0) + \int_0^t \phi\bigg(\mathbf{A}f - \sum_{j=1}^{m}\sum_{k=1}^{n_j} f(\lambda_{j,k}-1), s\bigg)ds \\
&\quad + \sum_{j=1}^{m}\sum_{k=1}^{n_j}\int_0^t \phi((\lambda_{j,k}-1)f, s-)dY_{j,k}(s)
\end{aligned}
$$

and $\pi_t$ is the unique measure-valued solution of the normalized filtering equation:

$$\begin{aligned}\pi(f,t) = \pi(f,0) + \int_0^t \pi(\mathbf{A}f,s)ds - \sum_{j=1}^{m}\int_0^t (\pi(fa_j,s) - \pi(f,s)\pi(a_j,s))ds \\ + \sum_{j=1}^{m}\sum_{k=1}^{n_j}\int_0^t \left[\frac{\pi(fg_{j,k},s-)}{\pi(g_{j,k},s-)} - \pi(f,s-)\right] dY_{j,k}(s),\end{aligned} \tag{3.4}$$

where $a_j = a_j(\vec{\theta}(t), \vec{X}(t), t)$ is the trading intensity, and $g_{j,k} = g_{j,k}(y_{j,k} \mid x_j)$ is the transition probability from $x_j = X_j(t)$ to $y_{j,k}$, the $k$-th price level for the $j$th stock. In the special case that $a_j(\vec{\theta}(t), \vec{X}(t), t) = a_j(t)$, the filtering equation of $\pi_t$ reduces to

$$\begin{aligned}\pi(f,t) = \pi(f,0) + \int_0^t \pi(\mathbf{A}f,s)ds \\ + \sum_{j=1}^{m}\sum_{k=1}^{n_j}\int_0^t \left[\frac{\pi(fg_{j,k},s-)}{\pi(g_{j,k},s-)} - \pi(f,s-)\right] dY_{j,k}(s).\end{aligned} \tag{3.5}$$

## 4. Bayesian Model Selection via Filtering

With the irregular likelihood functions traditional parameter estimation and model selection techniques can be prohibitively difficult to use. However, as Kass and Rafferty [8] demonstrated critical evaluation is still possible that through Bayesian inference, i.e., while the overall "fit" of a model to the data may be difficult to calculate, it can be determined which of two models is a better "fit." The calculated Bayes factors allow the researcher to estimate the pair-wise relative fit of models for the same data and to select the best of the available models.

Bayes factors are used to select the best of two models using the integrated (or marginal) likelihood functions of each model given the observed data. The Bayes factor for model 1 over model 2 is the ratio of the two integrated likelihood functions. Kass and Rafferty, [8] suggested the following rules for the interpretation of Bayes factor. A Bayes factor with a value between 1 and 3 is considered insignificant, i.e., Model 1 is not clearly "better" than Model 2. A Bayes factor with a value between 3 and 12 indicates Model 1 is a somewhat better fit for the data than Model 2. From a value between 12 and 150, it indicates that Model 1 is an overall better fit than Model 2 can be inferred. A Bayes factor with a value greater than 150 implies that Model 1 is decisively better. Naturally, if the Bayes factor is less than 1 we simply consider the Bayes factor of Model 2 over Model 1 and interpret the result accordingly.

### *4.1. Evolution of the Bayes Factors*

In our efforts to develop the Bayes factors for the multivariate FM models, we begin with notational adjustments for simplicity in working with two models. Since, the Bayes factors are used to compare the fit of two models given the observed prices, we will let $c = 1, 2$ through out this section. Then we denote Model-$c$ as $(\vec{\theta}^{(c)}, \vec{X}^{(c)}, \mathbf{Y})$ and its joint likelihood by $L^{(c)}(t)$ given as before by (3.1) and have the following definitions.

**Definition 4.1.** *Let* $\phi_c(f_c, t) = E^{Q^{(c)}}[f_c(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t))L^{(c)}(t) \mid \mathbf{F}_t^{\mathbf{Y}}]$.

Then the integrated likelihood of $\mathbf{Y}$ would be $\phi_c(1,t)$ for Model $c$ with proper priors.

**Definition 4.2.** *Denote* $\pi_t^{(c)}$ *as the conditional distribution of* $(\vec{\theta}^{(c)}, \vec{X}^{(c)})$ *given* $\mathbf{F}_t^{\mathbf{Y}}$. *Let* $\pi_c(f_c,t) = E^P[f_c(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t)) \mid \mathbf{F}_t^{\mathbf{Y}}]$.

Next, we need to define filter ratio processes as

**Definition 4.3.** *Let the filter ratio processes,* $q_1$ *and* $q_2$, *be written as*

$$q_1(f_1,t) = \frac{\phi_1(f_1,t)}{\phi_2(1,t)}, \qquad q_2(f_2,t) = \frac{\phi_2(f_2,t)}{\phi_1(1,t)}.$$

**Remark 4.1.** *the Bayes factors can then be written as* $B_{1,2}(t) = q_1(1,t)$ *and* $B_{2,1}(t) = q_2(1,t)$.

The evolution of the filter ratio processes can be characterized by Theorem 4.1.

**Theorem 4.1.** Suppose Model-$c$ $(c=1,2)$ has generator $\mathbf{A}^{(c)}$ for $(\vec{\theta}^{(c)}, \vec{X}^{(c)})$, trading intensity for the $j$th asset $a_j^{(c)} = a_j^{(c)}(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t), t)$, and transition probability $g_{j,k}^{(c)} = g_{j,k}^{(c)}(y_{j,k} \mid x_j)$ from $x_j$ to $y_{j,k}$ for the random transformation $F^{(c)}$. If Model-$c$ satisfies Assumptions 2.1 to 2.5, then $(q_t^{(1)}, q_t^{(2)})$ are the unique measure-valued pair solution of the system of SDEs

$$\begin{aligned} q_1(f_1,t) &= q_1(f_1,0) \\ &+ \int_0^t \left[ q_1(\mathbf{A}^{(1)} f_1, s) - \sum_{j=1}^m \left( q_1(a_j^{(1)} f_1, s) - \frac{q_1(f_1,s) q_2(a_j^{(2)}, s)}{q_2(1,s)} \right) \right] ds \\ &+ \sum_{j=1}^m \sum_{k=1}^{n_j} \int_0^t \left[ \frac{q_1(a_j^{(1)} f_1 g_{j,k}^{(1)}, s-)}{q_2(a_j^{(2)} g_{j,k}^{(2)}, s-)} q_2(1,s-) - q_1(f_1,s-) \right] dY_{j,k}(s) \end{aligned} \tag{4.1}$$

for all $t > 0$ and $f_1 \in D(\mathbf{A}^{(1)})$, and for all $t > 0$ and $f_2 \in D(\mathbf{A}^{(2)})$

$$\begin{aligned} q_2(f_2,t) &= q_2(f_2,0) \\ &+ \int_0^t \left[ q_2(\mathbf{A}^{(2)} f_2, s) - \sum_{j=1}^m \left( q_2(a_j^{(2)} f_2, s) - \frac{q_2(f_2,s) q_1(a_j^{(1)}, s)}{q_1(1,s)} \right) \right] ds \\ &+ \sum_{j=1}^m \sum_{k=1}^{n_j} \int_0^t \left[ \frac{q_2(a_j^{(2)} f_2 g_{j,k}^{(2)}, s-)}{q_1(a_j^{(1)} g_{j,k}^{(1)}, s-)} q_1(1,s-) - q_2(f_2,s-) \right] dY_{j,k}(s). \end{aligned} \tag{4.2}$$

In the special case that the intensity for the trading times of the $j$th asset depends only on $t$, namely, $a_j^{(c)}(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t), t) = a(t)$, the preceding equations reduce to

$$\begin{aligned} q_1(f_1,t) &= q_1(f_1,0) + \int_0^t q_1(\mathbf{A}^{(1)} f_1, s) ds \\ &+ \sum_{j=1}^m \sum_{k=1}^{n_j} \int_0^t \left[ \frac{q_1(f_1 g_{j,k}^{(1)}, s-)}{q_2(g_{j,k}^{(2)}, s-)} q_2(1,s-) - q_1(f_1,s-) \right] dY_{j,k}(s) \end{aligned} \tag{4.3}$$

and

$$q_2(f_2,t) = q_2(f_2,0) + \int_0^t q_2(\mathbf{A}^{(2)} f_2, s)ds$$
$$+\sum_{j=1}^{m}\sum_{k=1}^{n_j}\int_0^t \left[\frac{q_2(f_2 g_{j,k}^{(2)}, s-)}{q_1(g_{j,k}^{(1)}, s-)} q_1(1,s-) - q_2(f_2,s-)\right] dY_{j,k}(s). \tag{4.4}$$

*Proof.* There are two steps.

*Step 1: Derivation of the evolution equations for $q_1(f_1,t)$ and $q_2(f_2,t)$*

We will show that $q_1(f_1,t)$ satisfies (4.1), and when $a_j^{(c)}(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t), t) = a_j(t)$ (4.1) reduces to (4.3). Then, by symmetry, $q_2(f_2,t)$ satisfies (4.2) and, in the special case, (4.4). Recall that $\phi_c(f_s,t)$ satisfies (3.3). Then applying Ito's formula for semimartingales ([12]) and simplifying gives us

$$\frac{\phi_1(f_1,t)}{\phi_2(1,t)} = \frac{\phi_1(f_1,0)}{\phi_2(1,0)} + \int_0^t \frac{\phi_1(\mathbf{A}^{(1)} f_1, s)}{\phi_2(1,t)} ds$$
$$-\sum_{j=1}^{m}\int_0^t \left[\frac{\phi_1(a_j^{(1)} f_1, s)}{\phi_2(1,s)} - \frac{\phi_1(f_1,s)\phi_2(a_j^{(2)},s)}{\phi_2^2(1,s)}\right] ds$$
$$+\sum_{j=1}^{m}\sum_{k=1}^{n_j}\int_0^t \left[\frac{\phi_1(f_1,s)}{\phi_2(1,s)} - \frac{\phi_1(f_1,s-)}{\phi_2(1,s-)}\right] dY_{j,k}(s). \tag{4.5}$$

To transform this equation to the desired form, we make two observations. First,

$$\frac{\phi_1(f_1,s)\phi_2(a_j^{(2)},s)}{\phi_2^2(1,s)} = \frac{\frac{\phi_1(f_1,s)}{\phi_2(1,s)}\frac{\phi_2(a_j^{(2)},s)}{\phi_1(1,s)}}{\frac{\phi_2(1,s)}{\phi_1(1,s)}} = \frac{q_1(f_1,s)q_2(a_j^{(2)},s)}{q_2(1,s)}.$$

Next, again using (3.3) and assuming that the $k$th trade of the $j$th stock occurs at time $s$

$$\frac{\phi_1(f_1,s)}{\phi_2(1,s)} = \frac{\phi_1(f_1,s-) + \phi_1((a_j^{(1)} g_{j,k}^{(1)} - 1)f_1, s-)}{\phi_2(1,s-) + \phi_2(a_j^{(2)} g_{j,k}^{(2)} - 1, s-)}$$
$$= \frac{\phi_1(a_j^{(1)} g_{j,k}^{(1)} f_1, s-)}{\phi_2(a_j^{(2)} g_{j,k}^{(2)}, s-)} = \frac{q_1(a_j^{(1)} f_1 g_{j,k}^{(1)}, s-)}{q_2(a_j^{(2)} g_{j,k}^{(2)}, s-)} q_2(1,s-). \tag{4.6}$$

Substituting these results into the definition of the filter ratio process gives us (4.1). If $a_j^{(c)}(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t), t) = a_j(t)$ for $c = 1,2$ then

$$\frac{q_1(f_1,s)q_2(a_j,s)}{q_2(1,s)} = \frac{q_1(f_1,s)a_j q_2(1,s)}{q_2(1,s)} = q_1(a_j f_1, s)$$

and

$$\frac{q_1(a_j f_1 g_{j,k}^{(1)}, s-)}{q_2(a_j g_{j,k}^{(2)}, s-)} q_2(1,s-) = \frac{q_1(f_1 g_{j,k}^{(1)}, s-)}{q_2(g_{j,k}^{(2)}, s-)} q_2(1,s-)$$

giving us (4.3).

*Step 2: Uniqueness of $q_1(f_1,t)$ and $q_2(f_2,t)$*

It remains only to show the uniqueness of the filtering ratio processes, which follows closely from the perturbation arguments of SPDE such as given in Kouritzin and Zeng [9]. Towards this end, let $T^{(1)}, T^{(2)}$ be the semi-groups with weak generators $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}$. Define $\{\tau_i\}_{i=1}^{\infty}$ as the jump times associated with the $\mathbf{Y}$ where $\tau_0 = 0$. Here $\tau_i$ represents a trading time for one of the $j = 1, \dots m$ assets. For $c = 1, 2$ and $l = 3 - c$, let $(q_1, q_2)$ be finite measure process satisfying

$$q_c(f_c, t) = q_c(f_c, \tau_i) + \int_{\tau_i}^{t} q_c(\mathbf{A}^{(c)} f_c, s) ds - \sum_{j=1}^{m} \int_{\tau_i}^{t} \left[ q_c(a_j^{(c)} f_c, s) - \frac{q_c(f_c, s) q_l(a_j^{(l)}, s)}{q_l(1, s)} \right] ds \tag{4.7}$$

for all $t \in [\tau_i, \tau_i + 1)$ and $f_c \in D(\mathbf{A}^{(c)})$. Hence, by Assumption 2.3 and using (4.7) for $t \in [\tau_i, \tau_i + 1)$, we have

$$\exp(-C(t - \tau_i)) q_c(1, \tau_i) \leq q_c(1, t) \leq \exp(C(t - \tau_i)) q_c(1, \tau_i). \tag{4.8}$$

Now, we apply a standard technique in SPDE and define a convolution form $\chi_c$ for $c = 1, 2$ and $l = 3 - c$ by

$$\chi_c(t, u, f_c) = q_c(T_{t-u}^{(c)} f_c, u) + \sum_{j=1}^{m} \int_{u}^{t} \left[ \frac{q_c(T_{t-s}^{(c)} f_c, s) q_l(a_j^{(l)}, s)}{q_l(1, s)} - q_c(a_j^{(c)} T_{t-s}^{(c)} f_c, s) \right] ds$$

for all $u \leq t \in [\tau_i, \tau_i + 1)$ and $f_c \in D(\mathbf{A}^{(c)})$. Then, taking the fact that for all $s \geq 0$, $T_s^{(c)} f_c \in D(\mathbf{A}^{(c)})$ and applying Leibniz's rule give us

$$\frac{d}{du} \chi_c(t, u, f_c) = \frac{d}{du} q_c(T_{t-u}^{(c)} f_c, u) - \sum_{j=1}^{m} \left[ \frac{q_c(T_{t-u}^{(c)} f_c, u) q_l(a_j^{(l)}, u)}{q_l(1, u)} - q_c(a_j^{(c)} T_{t-u}^{(c)} f_c, u) \right].$$

Observe that

$$\frac{d}{du} q_c(T_{t-u}^{(c)} f_c, u) = -q_c(\mathbf{A}^{(c)} T_{t-u}^{(c)} f_c, u) + q_c(\mathbf{A}^{(c)} T_{t-u}^{(c)} f_c, u) + \sum_{j=1}^{m} \left[ \frac{q_c(T_{t-u}^{(c)} f_c, u) q_l(a_j^{(l)}, u)}{q_l(1, u)} - q_c(a_j^{(c)} T_{t-u}^{(c)} f_c, u) \right]. \tag{4.9}$$

Therefore, $\frac{d}{du} \chi_c(t, u, f_c) = 0$ for $u \in [\tau_i, t]$. This implies $\chi_c(t, t, f_c) = \chi_c(t, \tau_i, f_c)$, which produces

$$q_c(f_c, t) = q_c(T_{t-\tau_i}^{(c)} f_c, \tau_i) + \sum_{j=1}^{m} \int_{\tau_i}^{t} \left[ \frac{q_c(T_{t-s}^{(c)} f_c, s) q_l(a_j^{(l)}, s)}{q_l(1, s)} - q_c(a_j^{(c)} T_{t-s}^{(c)} f_c, s) \right] ds. \tag{4.10}$$

Now suppose that $(r_1, r_2)$ is a second process satisfying (4.7) such that

$$(q_1(\cdot, \tau_i), q_2(\cdot, \tau_i)) = (r_1(\cdot, \tau_i), r_2(\cdot, \tau_i)),$$

i.e., that the processes agree at the trading times. Then, using (4.10) for both pairs we have that for all $t \in [\tau_i, \tau_i + 1)$ and $f_c \in D(\mathbf{A}^{(c)})$, $c = 1, 2$

$$\begin{aligned}
&|r_1(f_1, t) - q_1(f_1, t)| + |r_2(f_2, t) - q_2(f_2, t)| \\
&\quad \leq \sum_{j=1}^{m} \Big[ \int_{\tau_i}^{t} \Big| \frac{r_1(T^{(1)}_{t-s} f_1, s) r_2(a^{(2)}_j, s)}{r_2(1, s)} - \frac{q_1(T^{(1)}_{t-s} f_1, s) q_2(a^{(2)}_j, s)}{q_2(1, s)} \Big| ds \\
&\qquad + \int_{\tau_i}^{t} \Big| \frac{r_2(T^{(2)}_{t-s} f_2, s) r_1(a^{(1)}_j, s)}{r_1(1, s)} - \frac{q_2(T^{(2)}_{t-s} f_2, s) q_2(a^{(1)}_j, s)}{q_1(1, s)} \Big| ds \\
&\qquad + \int_{\tau_i}^{t} |r_1(a^{(1)}_j T^{(1)}_{t-s} f_1, s) - q_1(a^{(1)}_j T^{(1)}_{t-s} f_1, s)| ds \\
&\qquad + \int_{\tau_i}^{t} |r_2(a^{(2)}_j T^{(2)}_{t-s} f_2, s) - q_2(a^{(2)}_j T^{(2)}_{t-s} f_2, s)| ds \Big].
\end{aligned}$$

However, together Assumption 2.3 and equation (4.8) imply

$$\begin{aligned}
(4.11) \qquad &\Big| \frac{r_1(T^{(1)}_{t-s} f_1, s) r_2(a^{(2)}_j, s)}{r_2(1, s)} - \frac{q_1(T^{(1)}_{t-s} f_1, s) q_2(a^{(2)}_j, s)}{q_2(1, s)} \Big| \\
&\quad + \Big| \frac{r_2(T^{(2)}_{t-s} f_2, s) r_1(a^{(1)}_j, s)}{r_1(1, s)} - \frac{q_2(T^{(2)}_{t-s} f_2, s) q_1(a^{(1)}_j, s)}{q_1(1, s)} \Big| \\
&\quad + |r_1(a^{(1)}_j T^{(1)}_{t-s} f_1, s) - q_1(a^{(1)}_j T^{(1)}_{t-s} f_1, s)| \\
&\quad + |r_2(a^{(2)}_j T^{(2)}_{t-s} f_2, s) - q_2(a^{(2)}_j T^{(2)}_{t-s} f_2, s)| \\
&\qquad \leq 2C \sup_{f_1 \in \bar{C}(E), \|f_1\|_\infty \leq 1} |r_1(f_1, t) - q_1(f_1, t)| \\
&\qquad\quad + 2C \sup_{f_2 \in \bar{C}(E), \|f_2\|_\infty \leq 1} |r_2(f_2, t) - q_2(f_2, t)|
\end{aligned}$$

where $s \in [\tau_i, \tau_i + 1)$, $f_c \in D(\mathbf{A}^{(c)})$ with $\|f\| \leq 1$. Now, using (4.8) and the compact containment condition, there exist increasing compact sets $K^{(c)}_n$ for $c = 1, 2$ such that $r_c(K^{(c)}_n, t) \wedge q_c(K^{(c)}_n, t) \geq 1 - \frac{1}{n}$ for all $t \in [\tau_i, \tau_i + 1)$. Then Assumption 2.3, (4.11) and (4.11), and Stone–Weierstrass imply

$$\begin{aligned}
&\sup_{f_1 \in \bar{C}(E), \|f_1\|_\infty \leq 1} \Big| \int_{K^{(1)}_n} f_1 [dr^{(1)}_t - dq^{(1)}_t] \Big| + \sup_{f_2 \in \bar{C}(E), \|f_2\|_\infty \leq 1} \Big| \int_{K^{(2)}_n} f_2 [dr^{(2)}_t - dq^{(2)}_t] \Big| \\
&\quad \leq \frac{2}{n} + \sup_{f_1 \in D(\mathbf{A}^{(1)}), \|f_1\|_\infty \leq 1} |r_1(f_1, t) - q_1(f_1, t)| \\
&\qquad + \sup_{f_2 \in D(\mathbf{A}^{(2)}), \|f_2\|_\infty \leq 1} |r_2(f_2, t) - q_2(f_2, t)| \\
&\quad \leq \frac{2}{n} + \frac{8C}{n}(t - \tau_i) + 4C \int_{\tau_i}^{t} \Big[ \sup_{f_1 \in \bar{C}(E), \|f_1\|_\infty \leq 1} \Big| \int_{K^{(1)}_n} f_1 [dr^{(1)}_t - dq^{(1)}_t] \Big| \\
&\qquad\qquad + \sup_{f_2 \in \bar{C}(E), \|f_2\|_\infty \leq 1} \Big| \int_{K^{(2)}_n} f_2 [dr^{(2)}_t - dq^{(2)}_t] \Big| \Big] ds.
\end{aligned}$$

Finally, we apply Gronwell's inequality ([5]) and let $n \to \infty$ and obtain

$$\sup_{f_1 \in \bar{C}(E), \|f_1\|_\infty \leq 1} \Big| \int_E f_1 [dr^{(1)}_t - dq^{(1)}_t] \Big| + \sup_{f_2 \in \bar{C}(E), \|f_2\|_\infty \leq 1} \Big| \int_E f_2 [dr^{(2)}_t - dq^{(2)}_t] \Big| = 0.$$

Hence, we have the uniqueness for $t \in [\tau_0, \tau_1)$ and the updating equation implies the same holds for $t = \tau_1$. By induction, we have the uniqueness on $[0, \infty)$, and the theorem follows. $\square$

## 5. A Convergence Theorem and a Numerical Scheme

Theorem 4.1 characterizes the evolution of the Bayes factors. To compute the Bayes factors, one constructs an algorithm to approximate $q_k(f_k, t)$, where $q_1(1, t) = B_{12}(t)$. The algorithm, based on the evolution of SDEs, is naturally recursive, handling a datum at a time. Thus, the algorithm makes real-time updates and can handle large data sets.

One basic requirement for the recursive algorithm is consistency: The approximate $q_k$, computed by the recursive algorithm, must converge to the true one. The following theorem summarizes the related convergence results and provides the theoretical foundation for consistency. Furthermore, the theorem furnishes a recipe for constructing consistent recursive algorithms.

Let $c = 1, 2$ throughout this section. First denote $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x})$ as an approximation of $(\vec{\theta}^{(c)}, \vec{X}^{(c)})$. Further, denote $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x}) \Rightarrow (\vec{\theta}^{(c)}, \vec{X}^{(c)})$ as the weak convergence in the Skorohod topology as $(\vec{\epsilon}, \vec{\epsilon}_x) \to 0$. Define for each $j = 1, 2, \ldots, m$

$$\vec{Y}^{\varepsilon,c}_j(t) = \begin{pmatrix} N^{(c)}_{j,1}(\int_0^t \lambda^{(c)}_{j,1}(\vec{\theta}^{(c)}_{\vec{\epsilon}}(s), \vec{X}^{(c)}_{\vec{\epsilon}_x}(s), s)ds) \\ N^{(c)}_{j,2}(\int_0^t \lambda^{(c)}_{j,2}(\vec{\theta}^{(c)}_{\vec{\epsilon}}(s), \vec{X}^{(c)}_{\vec{\epsilon}_x}(s), s)ds) \\ \vdots \\ N^{(c)}_{j,n_j}(\int_0^t \lambda^{(c)}_{j,n_j}(\vec{\theta}_{\vec{\epsilon}^{(c)}}(s), \vec{X}^{(c)}_{\vec{\epsilon}_x}(s), s)ds) \end{pmatrix}, \tag{5.1}$$

where $\varepsilon = \max(|\vec{\epsilon_x}|, |\vec{\epsilon}|)$ and $|\vec{\epsilon}|$ is Euclidean norm of a vector. Analogously to the continuous case, we define the collection of the counting processes for the approximate model $c$ as

$$\mathbf{Y}^{(c)}_\varepsilon(t) = (\vec{Y}^{\varepsilon,c}_1(t), \vec{Y}^{\varepsilon,c}_2(t), \ldots, \vec{Y}^{\varepsilon,c}_m(t)). \tag{5.2}$$

Let $\mathbf{F}^{\mathbf{Y}^{(c)}_\varepsilon}_t = \sigma(\mathbf{Y}^{(c)}_\varepsilon(s), 0 \le s \le t)$.

Now, let $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x}, \mathbf{Y}^{(c)}_\varepsilon)$ be defined on $(\Omega^{(c)}_\varepsilon, \mathbf{F}^{(c)}_\varepsilon, P^{(c)}_\varepsilon)$ with Assumptions 2.1 to 2.5. Assumptions 2.1 to 2.3 imply the existence of a reference measure $Q^{(c)}_\varepsilon$ having similar properties. The corresponding Radon–Nikodym derivative is $dP_\varepsilon/dQ_\varepsilon$ which is

$$\begin{aligned} L_\varepsilon(t) = \prod_{j=1}^{m}\prod_{k=1}^{n_j} \exp\bigg[\int_0^t &\log(\lambda_{j,k}(\vec{\theta}_{\vec{\epsilon}}(s-), \vec{X}_{\vec{\epsilon}_x}(s-), s-))dY_{j,k}(s) \\ &- \int_0^t [\lambda_{j,k}(\vec{\theta}_{\vec{\epsilon}}(s-), \vec{X}_{\vec{\epsilon}_x}(s-), s-) - 1]ds\bigg]. \end{aligned} \tag{5.3}$$

Given this reference measure and the joint likelihood function we can similarly define for $c = 1, 2$ and $l = 3 - c$

$$\phi_{\varepsilon,c}(f_c, t) = E^{Q^{(c)}_\varepsilon}[f_c(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x})L^{(c)}_\varepsilon(t) \mid \mathbf{F}^{\mathbf{Y}^{(c)}_\varepsilon}_t],$$

$$\pi_{\varepsilon,c}(f_c, t) = E^{P^{(c)}_\varepsilon}[f_c(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x}) \mid \mathbf{F}^{\mathbf{Y}^{(c)}_\varepsilon}_t],$$

and

$$q_{\varepsilon,c}(f_c,t) = \frac{\phi_{\varepsilon,c}(f_c,t)}{\phi_{\varepsilon,l}(1,t)}.$$

We have the following theorem.

**Theorem 5.1.** Let $c = 1, 2$. Suppose that Assumptions 2.1 to 2.5 hold for the model $(\vec{\theta}^{(c)}, \vec{X}^{(c)}, \mathbf{Y}^{(c)})$ and for the approximate model $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x}, \mathbf{Y}^{(c)}_{\varepsilon})$. Suppose $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x}) \Rightarrow (\vec{\theta}^{(c)}, \vec{X}^{(c)})$ as $(\vec{\epsilon}, \vec{\epsilon}_x) \to 0$. Then, as $\varepsilon \to 0$, for bounded and continuous $f_1$ and $f_2$,
(i) $\mathbf{Y}_\varepsilon \Rightarrow \mathbf{Y}$, as $\varepsilon \to 0$
(ii) $\phi_{\varepsilon,c}(f_c,t) \Rightarrow \phi_c(f_c,t)$
(iii) $\pi_{\varepsilon,c}(f_c,t) \Rightarrow \pi_c(f_c,t)$
(iv) $q_{\varepsilon,1}(f_1,t) \Rightarrow q_1(f_1,t)$ and $q_{\varepsilon,2}(f_2,t) \Rightarrow q_2(f_2,t)$ simultaneously.

**Remark 5.1.** Part (i) implies the convergence of the observation in approximate model to that in the true one. We note that Part (ii) implies the consistency of the approximate (integrated) likelihood, while Part (iii) shows the consistency of approximate posterior. Lastly, Part (iv) implies the consistency of the approximate Bayes factors. Taken as a whole, Theorem 5.1 shows that there are discrete, computationally feasible, and consistent versions of the likelihoods, posterior, and Bayes factors for the class of models.

*Proof.* Parts (i), (ii) and (iii) are proven in [14]. The proof relies on several theorems: *Kurtz and Protter's theorem* on the convergence of stochastic integral (see Theorem 2.2 of [11]), two *theorems on the convergence of conditional expectations* (see [6] and [10]) and the *continuous mapping theorem* (see [5]). Part (iv) follows directly from Part (ii) since, again by the *continuous mapping theorem* $q_{\varepsilon,c}(f_c,t)$ is consistent if $\phi_{\varepsilon,c}(f_c,t)$ and $\phi_{\varepsilon,l}(1,t)$ are for $c = 1, 2$, $l = 3 - c$. □

### *5.1. Overview of the Recursive Algorithm*

The advantage of having a consistent discrete approximation of the model and Bayes factors is that it makes the model estimation and evaluation computationally feasible. To implement the Bayesian model selection via filtering as described here, we can construct recursive algorithms to calculate the approximate Bayes factors $q_{\varepsilon,c}(f_c,t)$, $c = 1, 2$. For simplicity of this discussion we will assume that $a^{(c)}_j(\vec{\theta}^{(c)}(t), \vec{X}^{(c)}(t), t) = a_j(t)$ for the $j$th asset.

The first step is to construct for $c = 1, 2$, $(\vec{\theta}^{(c)}_{\vec{\epsilon}}, \vec{X}^{(c)}_{\vec{\epsilon}_x})$ as a Markov chain approximation to $(\vec{\theta}^{(c)}, \vec{X}^{(c)})$ with generator $\mathbf{A}_\varepsilon$ and obtain $g_{j,k}(y^{(c)}_{j,\varepsilon} \mid X^{(c)}_{j,\varepsilon}(t))$ as an approximation to $g_{j,k}(y_j \mid X_j(t))$. We will restrict our space to the lattice points corresponding to the assumed prior distribution. In the second step, we obtain the approximate Bayes factors from (4.3) and (4.4) broken down into the propagation equation:

$$q^{(c)}_\varepsilon(f_c, t_{i+1}-) = q^{(c)}_\varepsilon(f_c, 0) + \int_{t_i}^{t_{i+1}} q^{(c)}_\varepsilon(\mathbf{A}_\varepsilon f_c, s)ds, \tag{5.4}$$

and the updating equation:

$$q^{(c)}_\varepsilon(f_c, t_{i+1}) = \frac{q^{(c)}_\varepsilon(f_c g^{(c)}_{j,k}, t_{i+1}-)}{q^{(3-c)}_\varepsilon(f_c g^{(c)}_{j,k}, t_{i+1}-)} q^{(3-c)}_\varepsilon(1, t_{i+1}-). \tag{5.5}$$

In the final step, we convert (5.4) and (5.5) into recursive algorithms setting $f_c$ as a lattice-point indicator with two sub-steps: (a) represents $q_\varepsilon(\cdot, t)$ as a finite array with the components being $q_\varepsilon(f, t)$ and (b) approximates the time integral in (5.4) with an Euler scheme.

## 6. Conclusions and Future Works

In this paper, we investigate the model selection problems for a general class of multivariate FM models of asset price and develop Bayesian model selection via filtering in two steps. We first derive the evolution system of SPDEs for the Bayes factors and prove its uniqueness. Then we prove a limit theorem which provide a recipe to develop consistent recursive algorithms for computing the Bayes factors.

The Bayesian model selection via filtering is computationally intensive and even so in the multivariate FM models. To improve efficiency, we will extend the recent developments in particle filtering to the filtering problem with counting process observations. See [16] for the recent development in this direction for univariate case.

With efficient algorithms for implementation, the developed Bayesian model selection via filtering offers a powerful tool to test related market microstructure theories, represented by the micromovement models. For examples, we may test whether NASDAQ has less trading noise after a market reform as argued in [1], test whether information affects trading intensity as argued by [3] and tested by [4], test whether inventory position of a market maker has an impact on price suggested in [7], test whether there is relationship between transaction times and limit order arrival times as in [13], and test whether there is a structure break in transaction periods as in [20].

The multivariate FM models can be further generalized to the filtering models with marked point process observations and likewise the related Bayesian inference via filtering. See [19] for further development.

## References


[1] Barclay, M., Christie, W., Harris, J., Kandel, E. and Schultz, P. H. (1999). The effects of market reform on the trading costs and depths of nasdaq stocks. *Journal of Finance* **54 (1)** 1–34.

[2] Bremaud, P. (1981). *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

[3] Easley, D. and O'Hara, M. (1992). Time and the process of security price adjustment. *Journal of Finance* **47** 577–605.

[4] Engle, R. (2000). The econometrics of ultra-high-frequency data. *Econometrica* **68** 1–22.

[5] Ethier, S. and Kurtz, T. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

[6] Goggin, E. (1994). Convergence in distribution of conditional expectations. *Annals of Probability* **22** 1097–1114.

[7] Hasbrouck, J. (1988). Trades, quotes, inventories and information. *Journal of Financial Economics* **42** 229–252.

[8] Kass, R. E. and Raftery, A. E. (1995). Bayes factors and model uncertainty. *Journal of the American Statistical Association* **90** 773–795.

[9] Kouritzin, M. and Zeng, Y. (2005). Bayesian model selection via filtering for a class of micro-movement models of asset price. *International Journal of Theoretical and Applied Finance* **8** 97–121.

[10] Kouritzin, M. and Zeng, Y. (2005). Weak convergence for a type of conditional expectation: Application to the inference for a class of asset price models. *Nonlinear Analysis: Theory, Methods and Applications* **60** 231–239.

[11] Kurtz, T. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* **19** 1035–1070.

[12] Protter, P. (2003). *Stochastic Integration and Differential Equations.* Springer-Verlag, New York, 2nd ed.

[13] Russell, J. (1999). Econometric modeling of multivariate irregularly-spaced high-frequency data. Working Paper, University of Chicago.

[14] Scott, L. C. and Zeng, Y. (2006). Bayes estimation for a class of multivariate filtering micromovement models of asset price. Working Paper, University of Missouri at Kansas City.

[15] Spalding, R., Tsui, K. W. and Zeng, Y. (2005). A micro-movement model with bayes estimation via filtering: Applications to measuring trading noises and trading cost. *Nonlinear Analysis: Theory, Methods and Applications* **64** 295–309.

[16] Xiong, J. and Zeng, Y. (2006). A branching particle approximation to the filtering problem with counting process observations. Working Paper, University of Tennessee at Knoxville.

[17] Zeng, Y. (2003). A partially-observed model for micro-movement of asset prices with bayes estimation via filtering. *Mathematical Finance* **13** 411–444.

[18] Zeng, Y. (2004). Estimating stochastic volatility via filtering for the micro-movement of asset prices. *IEEE Transactions on Automatic Control* **49** 338–348.

[19] Zeng, Y. (2006). Statistical analysis of the filtering models with marked point process observations: Applications to ultra-high frequency data. Working Paper, University of Missouri at Kansas City.

[20] Zhang, M. Y., Russell, J. R. and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics* **104** 179–207.