A Branching Particle Approximation to the Filtering Problem with Counting Process Observations*

Jie Xiong † Yong Zeng ‡

November 28, 2006

Abstract

Recently, the filtering model with counting process observations has been demonstrated as a sensible framework for modeling the micromovement of asset price (or ultra-high frequency data). In this paper, we first construct a branching particle system for such a nonlinear filtering model. Then, we show the weighted empirical measures in the constructed branching system converges to the optimal filters uniformly in time by deriving sharp upper bounds for the mean square error. Furthermore, we prove a central limit type theorem to characterize the convergence rate of such weighted empirical measures. The convergence rate is $n^{1/2}$, which is better than the best rate in the classical nonlinear filtering case where the rate is $n^{(1-\alpha)/2}$ for any $\alpha > 0$.

2000 *Mathematics Subject Classification.* Primary: 60H15; Secondary: 60K35, 35R60, 93E11, 60F05, 91B28.

Key Words: Stochastic partial differential equation, particle filters, Monte Carlo approximation, filtering, counting process, and ultra-high frequency data.

---

*This work was done when the second author visited the first author at the Department of Mathematics, University of Tennessee at Knoxville in Fall 2006. The hospitality of and the financial support from the Mathematics Department are gratefully acknowledged. Xiong's research supported in part by NSA Grant H98230-05-1-0043 and Zeng's by NSF Grant DMS-0604722.

†Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300, USA; Tel: (865) 974-4271, Fax: 865-974-6576, Email: jxiong@math.utk.edu and Website: http://www.math.utk.edu/∼jxiong/; and Department of Mathematics, Hebei Normal University, Shijiazhuang 050016, PRC.

‡Department of Mathematics and Statistics, University of Missouri at Kansas City, Kansas City, MO 64110, USA. Tel: (816) 235 5850. Fax: (816) 235 5517. Email: zeng@mendota.umkc.edu. Website: http://mendota.umkc.edu/.

1 Introduction

Recently much research have been developed for modeling the micromovement of asset price referred as the transaction or trade-by-trade price behavior (See [18], [17], and [36] for recent developments). Engle (Nobel Price Laureate, 2003) [17] calls such data as ultra-high frequency data, because of their ultimate disaggregation nature. The micromovement has two characteristics distinguishing from the continuous-time models in asset pricing, or the price macromovement referred as the equally-spaced daily, or weekly closing price behavior in the econometric literature. First, the micromovement observations occur at varying random time intervals. Second, financial noise (or trading noise or market microstructure noise) in the price are not ignorable anymore as in the continuous-time or macromovement cases due to the high frequency transaction nature.

Zeng [32] proposes a general Filtering Micromovement model for asset price (FM model, as we simply call it), where the sample characteristics of micro- and macro-movements are tied in a consistent manner. Economically, the proposed model has the structure similar to a class of the time series structural models developed in many early market microstructure papers (see [21], a survey paper on this topic, and [22]). Namely, price can be decomposed as a permanent component and a transient component. The permanent component has a long-term impact on price while the transient component has only a short-term impact. In FM model, there is an unobservable intrinsic value process for an asset, which corresponds to the usual price process in the option pricing literature and in the empirical econometric literature of macro-movement. The intrinsic value process is the permanent component and has a long-term impact on price. Prices are observed only at random trading times which are modeled by a conditional Poisson process. Moreover, prices are distorted observations of the intrinsic value process at the trading times and trading (or market microstructure) noise is explicitly modeled. It is the transient component and only has a short-term impact (when a trade happens) on price.

The most prominent feature of FM model is that trade-by-trade prices are viewed as a collection of counting processes of price level and the model is framed as a filtering problem with counting process observations. Then, the unnormalized and normalized filtering equations, which correspond to Duncan-Mortensen-Zakai, and KushnerStratonovich or FujisakiKallianpurKunita equations in classical nonlinear filtering, are derived. These equations characterize the evolution of the integrated likelihoods and the conditional distribution of the intrinsic value process (the signal). The Markov chain approximation method is applied to numerically solve the filtering equations. Then, Bayes estimation via filtering for the intrinsic value process and the related parameters in the model is developed in [32]. Bayesian hypothesis testing or model selection via filtering for this class of models is developed in [24]. Furthermore, a risk minimization hedging strategy for a FM model is considered in [28], and a mean-variance portfolio selection for a FM model is studied in [30].

On the other hand, branching and interacting particle filters as approximation of optimal filters in the classical nonlinear filtering have been studied extensively in the last ten years. In the classical Monte Carlo method, the unnormalized filter is approximated by a weighted particle system, but the variances of weights grow exponentially fast. To achieve the goal of variance reduction, the idea is to divide the time interval into small subintervals and the weight for each particle is updated so that the exponential martingale depends on the signal and the noise in the small interval prior to the time of interest. For the approach of interacting particle filters, interested readers are referred to the comprehensive monograph [13] by Del Moral and related references therein. For the approach of branching particle filters, the updating is via branching in small time steps. Precisely, at each time step, each existing particle will die or give birth to a random number of offspring proportional to the weight. Meanwhile, the distribution of this integer-valued variable is selected to have minimal variance subject to this constraint. In this way, particles that stay on the right tract (representing by heavy weights) are explored more thoroughly while particles with unlikely trajectories/positions (representing by little weights) are not carried forward uselessly. Thus, the variation decreases. We refer interested readers to the papers by Crisan and his coauthors in [6] - [11], especially, Crisan and Xiong [12].

In this paper, we adopt the branching particle filtering approach for the FM model. Suppose that V_t is the likelihood and π_t is the conditional distribution of the FM model. V_t (π_t) is characterized by a unnormalized (normalized) filtering equation. First, we construct a branching particle system for the FM model. Then, we define the weighted empirical measure π_t^n (V_t^n) and the unweighted one $\tilde{\pi}_t^n$ ($\tilde{V}_t^n$) of the constructed branching particle filter (Historically, unweighted empirical measures were first studied and were proven convergent to the optimal filters. However, as indicated by Crisan [6] and recently shown by Crisan and Xiong [12], the weighted empirical measure is superior to the unweighted one in convergence rate in the classical nonlinear filtering case. We believe the same holds in this case and focus on the weighted empirical measure in the paper.). The first aim of this paper is to prove the uniform convergence (in time) of V_t^n and $\tilde{V}_t^n$ to V_t as well as π_t^n or $\tilde{\pi}_t^n$ to π_t when $n \to \infty$. We prove them by deriving sharp upper bounds for the mean square errors. The key estimates are in Lemmas 4.5 and 6.1. Moreover, we study the convergence rate of V_t^n and π_t^n. It turns out that the rate is $n^{1/2}$, which is better than the best rate in the classical nonlinear filtering case where the rate is $n^{(1-\alpha)/2}$ for any $\alpha > 0$ (see [12]). This is because the key estimates in Lemmas 4.5 and 6.1 are better than those in the classical nonlinear filtering case (see [6] and [12]). We prove a central limit type (CLT) theorem to characterize the rate. Similar CLT results shown for unweighted particle filters using the interacting particle systems can be found in [13], [15] and [16] for the classical case. Recent results for central limit theorems in the discrete time framework can be found in [1] and [25].

The rest of this paper goes as follows: Section 2 briefly reviews FM models and related results. Section 3 develops a branching particle system and defines the weighted and unweighted empirical measures. Section 4 proves the convergence of the weighted empirical measure for each time t. Section 5 proves the convergence uniformly in time. Section 6 further derives a central limit type theorem. Section 7 concludes. The case of $\tilde{\pi}_t^n$ and $\tilde{V}_t^n$ are dealt with by final remarks in Sections 4 - 5.

Throughout this paper, we shall use K with a subscript to denote a constant whose value might be different in different proofs.

2 The Filtering Model

This section focuses on presenting a FM model whose filters are approximated by a branching particle system in this paper. There are two equivalent representations of the model which are presented in the following two subsections. The equivalence in the sense that both representations have the same probability distribution is proven in [34]. Section 2.3 reviews the related filtering equations.

2.1 Construction of Price from Intrinsic Value

In Representation I, FM model is predicated on a simple intuition that the price is formed from the intrinsic value process of an asset corrupted by trading (or market microstructure) noise.

Assumption 2.1 X_t, the intrinsic value process of an asset follows a diffusion process:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t,$$

where B_t is a standard Brownian motion and X_t has a unique weak solution.

The generator associated with X is

$$Lf(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 f}{\partial x^2}(x) + \mu(x)\frac{\partial f}{\partial x}(x).$$

Obviously, Assumption 2.1 includes geometric Brownian motion (GBM), the Black-Scholes model. The intrinsic value process can not be observed directly, but can be partially observed through the price process, Y. Due to price discreteness, Y is in a discrete state space given by the multiples of *tick*, the minimum price variation set by trading regulation. Y is a distorted observation of X at some random times.

There are three general steps in constructing Y from X. First, we specify $X(t)$ as in Assumption 2.1. Next, we specify the trading times $t_1, t_2, \ldots, t_i, \ldots$, which are driven by a conditional Poisson process with an conditional intensity function $a(X(t), t)$. Finally, $Y(t_i)$, the price at time t_i, is specified by $Y(t_i) = F(X(t_i))$, where $y = F(x)$ is a random transformation with the transition probability $p(y|x)$, modeling trading noise.

Under this construction, information affects $X(t)$, the intrinsic value of an asset, and has a permanent influence on the price while noise affects $F(x)$, the random transformation, and has only a transitory influence on price. This formulation is similar to the time series structural models applied in many market microstructure papers (see Hasbrouck [21], a survey paper, and Hasbrouck [22]) in that $X(t)$ is the permanent component and $F(x)$ is the transient component.

Examples of $F(x)$(or $p(y|x)$) are given in [32], [33], and [29]. These examples well accommodate the three types of well-documented noise in financial literature: discrete noise, clustering noise, and non-clustering noise. Especially, Spalding et. at. [29] applied a simple FM model with new measures of trading noises and trading cost to further support the important financial findings in [5], [4] and [2] (The results of Christie and Schultz [5] led to regulatory investigations, legal activities, and numerous academic studies. This culminated with the Securities and Exchange Commission imposed a series of market reforms in NASDAQ. Barclay et. al. [2] documented effects of the NASDAQ market reforms.).

2.2 Filtering with Counting Process Observations

Alternatively, we can view the transaction prices in the levels of price due to price discreteness. That is, in Representation II, we view the prices as a collection of counting processes in the following form:

$$\vec{Y}(t) = \begin{pmatrix} N_1(\int_0^t \lambda_1(X(s), s)ds) \\ N_2(\int_0^t \lambda_2(X(s), s)ds) \\ \vdots \\ N_w(\int_0^t \lambda_w(X(s), s)ds) \end{pmatrix}, \tag{1}$$

where $Y_k(t) = N_k(\int_0^t \lambda_k(X(s), s)ds)$ is the counting process recording the cumulative number of trades that have occurred at the kth price level (denoted by y_k) up to time t.

The following four mild assumptions are invoked.

Assumption 2.2 N_k's are unit Poisson processes under the physical measure P.

Assumption 2.3 $X, N_1, N_2, \ldots, N_w$ are independent under P.

Assumption 2.4 The total intensity, $a(x,t)$, is uniformly bounded above; i.e., there exist a constant, K, such that $a(x,t) \le K$ for all $t > 0$ and x.

Assumption 2.5 The intensity at price level k, $\lambda_k(x,t) = a(x,t)p(y_k|x)$, where $a(x,t)$ is the total trading intensity at time t with $x = X(t)$ and $p(y_k|x)$ is the transition probability from x to y_k, the kth price level.

Remark 2.1 Note that $p(y_k|x)$ is the same as $p(y|x)$ in $F(x)$ of Representation I. The structure of λ_k implies that $a(X(t),t)$ specifies when the trade might occur while $p(y_k|x)$ specifies at which price level the trade might occur.

For the notation convenience, we denote $ap_k(X_t,t) = \lambda_k(X_t,t)$ through the rest of the paper.

Remark 2.2 Under this representation, $X(t)$ becomes the signal process, which cannot be observed directly, and $\vec{Y}(t)$ becomes the observation process, which is corrupted by noise. Hence, $(X, \vec{Y})$ is framed as a filtering problem with counting process observations.

2.3 Filtering Equations

We can assume that $(X, \vec{Y})$ is in a filtered complete probability space $(\Omega, \hat{\mathcal{G}}, \hat{\mathbb{F}}, P)$ where $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{0 \le t \le \infty}$ is a given filtration. Assumptions 2.2 - 2.4 imply that there is a reference measure Q under which, X and $\vec{Y}$ become independent, X remains the same probability distribution and $Y_1, Y_2, \ldots, Y_n$ become unit Poisson processes. We consider a fixed time period $[0,T]$. Then, the Radon-Nikodym derivative is:

$$M(T) = \frac{dP}{dQ} = \prod_{k=1}^{w} \exp\left\{ \int_0^T \log ap_k(X(s-), s-) dY_k(s) - \int_0^T \Big[ap_k(X(s), s) - 1 \Big] ds \right\}. \tag{2}$$

Let $M(t) = E^Q[M(T)|\mathcal{F}_t]$. Then, $M(t)$ satisfies the following SDE:

$$dM(t) = \sum_{k=1}^{w} (ap_k - 1) M(t-) d(Y_k(t) - t). \tag{3}$$

Let $\mathcal{F}_t^{\vec{Y}} = \sigma\{(\vec{Y}(s)) | 0 \le s \le t\}$ be all the available information up to time t and let π_t be the conditional distribution of $X(t)$ given $\mathcal{F}_t^{\vec{Y}}$. Define

$$\langle V_t, f \rangle = E^Q[f(\theta(t), X(t)) M(t) | \mathcal{F}_t^{\vec{Y}}] \quad \text{and} \quad \langle \pi_t, f \rangle = E^P[f(\theta(t), X(t)) | \mathcal{F}_t^{\vec{Y}}].$$

By Kallianpur-Striebel formula, the optimal filter in the sense of least mean square error can be written as $\langle \pi_t, f \rangle = \langle V_t, f \rangle / \langle V_t, 1 \rangle$. Hence, the equation governing the evolution of $\langle V_t, f \rangle$ is called the *unnormalized filtering equation*, and that of $\langle \pi_t, f \rangle$ is called the *normalized filtering equation.*

The following proposition is a theorem from [32] summarizing both filtering equations.

Proposition 2.1 *Suppose that $(\theta, X, \vec{Y})$ satisfies Assumptions 1 - 5. Then, V_t is the unique measure-valued solution of the following SDE, the unnormalized filtering equation,*

$$\langle V_t, f\rangle = \langle V_0, f\rangle + \int_0^t \langle V_s, Lf\rangle ds + \sum_{k=1}^{w} \int_0^t \langle V_{s-}, (ap_k - 1)f\rangle d(Y_k(s) - s), \tag{4}$$

for $t > 0$ and $f \in D(L)$, the domain of generator L, where $a = a(X(t), t)$, is the trading intensity, and $p_k = p(y_k|x)$ is the transition probability from x to y_k, the kth price level.

π_t is the unique measure-valued solution of the SDE, the normalized filtering equation,

$$\begin{aligned} \langle \pi_t, f\rangle =& \langle \pi_0, f\rangle + \int_0^t \big[\langle \pi_s, Lf\rangle - \langle \pi_s, fa\rangle + \langle \pi_s, f\rangle\langle \pi_s, a\rangle\big] ds \\ &+ \sum_{k=1}^{w} \int_0^t \Big[\frac{\langle \pi_{s-}, fap_k\rangle}{\langle \pi_{s-}, ap_k\rangle} - \langle \pi_{s-}, f\rangle\Big] dY_k(s). \end{aligned} \tag{5}$$

When $a(X(t), t) = a(t)$, the above equation is simplified as:

$$\langle \pi_t, f\rangle = \langle \pi_0, f\rangle + \int_0^t \langle \pi_s, Lf\rangle ds + \sum_{k=1}^{w} \int_0^t \Big[\frac{\langle \pi_{s-}, fap_k\rangle}{\langle \pi_{s-}, ap_k\rangle} - \langle \pi_{s-}, f\rangle\Big] dY_k(s). \tag{6}$$

3 A Branching Particle System

In this section, we describe a branching particle system and define the weighted and unweighted empirical measures to approximates the optimal filter.

Initially, there are n particles of weight $\frac{1}{n}$ each at x_i^n, $i = 1, 2, \cdots, n$, satisfying the following initial condition
(I): As $n \to \infty$,

$$V_0^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i^n} \to \pi_0 \qquad \textit{in } \mathcal{M}_F(\mathbb{R}),$$

where $\mathcal{M}_F(\mathbb{R})$ is the collection of finite measures on $\mathbb{R}$. Let $\delta = \delta_n = n^{-2\alpha}$ for $0 < \alpha \le 1$ as the length between two time steps. Suppose that at time $t = j\delta$, there are m_j^n particles alive. During the time interval $(j\delta, (j+1)\delta)$, the particles move according to the following diffusions: For $i = 1, 2, \cdots, m_j^n$,

$$X_t^i = X_{j\delta}^i + \int_{j\delta}^t \mu(X_s^i) ds + \int_{j\delta}^t \sigma(X_s^i) dB_s^i.$$

where $\{B^i, i = 1, 2, \cdots, n\}$ are independent standard Brownian motions. The weight of particle i at time $j\delta$ is set to be 1 and for t during $[j\delta, (j+1)\delta)$ is

$$M_j^n(X^i, t) = \prod_{k=1}^{w} \exp\left(\int_{j\delta+}^t \log ap_k(X_{s-}^i, s-) dY_k(s) - \int_{j\delta}^t \big[ap_k(X_s^i, s) - 1\big] ds\right). \tag{7}$$

Obviously, $M^n(X^i, t)$ depends on the observation $\vec{Y}$, and is the likelihood ratio of measure P over measure Q defined by Equation (2) given the simulated path of X^i. If there is no trade (or no jump for $\vec{Y}$) in $(j\delta, t)$, then the first integral in the exponent is zero.

Recall at the beginning, $M_j^n(X^i, j\delta) = 1$. At the end right before branching, the weight becomes $M_{j+1}^n(X^i) = M_j^n(X^i, (j+1)\delta-)$.

To achieve variance reduction, at the end of the interval, the ith particle $(i = 1, 2, \cdots, m_j^n)$ branches (independent of others) into a random number ξ_{j+1}^i of offsprings proportional to its weight. Precisely, we let

$$E^Q\left(\xi_{j+1}^i | \mathcal{F}_{(j+1)\delta-}\right) = \tilde{M}_{j+1}^n(X^i)$$

where

$$\tilde{M}_{j+1}^n(X^i) = \frac{M_{j+1}^n(X^i)}{\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_{j+1}^n(X^\ell)}.$$

Let

$$Var^Q\left(\xi_{j+1}^i | \mathcal{F}_{(j+1)\delta-}\right) = \gamma_{j+1}^n(X^i)$$

Following [8], in order to minimize the variance γ_{j+1}^n, we restrict the possible number of ξ_{j+1}^i to the two integers closest to $\tilde{M}_{j+1}^n(X^i)$ and take

$$\xi_{j+1}^i = \begin{cases} [\tilde{M}_{j+1}^n(X^i)] & \text{with probability } 1 - \{\tilde{M}_{j+1}^n(X^i)\} \\ [\tilde{M}_{j+1}^n(X^i)] + 1 & \text{with probability } \{\tilde{M}_{j+1}^n(X^i)\} \end{cases}$$

where $\{x\} = x - [x]$ is the fraction of x. In this case

$$\gamma_{j+1}^n(X^i) = \{\tilde{M}_{j+1}^n(X^i)\}(1 - \{\tilde{M}_{j+1}^n(X^i)\}). \tag{8}$$

Define

$$\pi_t^n = \frac{1}{n}\sum_{i=1}^{m_j^n} M_j^n(X^i, t)\delta_{X^i(t)}, \quad \text{and} \quad V_t^n = \pi_t^n \eta_t^n, \qquad j\delta \le t < (j+1)\delta.$$

where

$$\eta_t^n = \Pi_{j=0}^{k-1} \frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_{j+1}^n(X^\ell), \qquad \text{if } k\delta \le t < (k+1)\delta.$$

Then, $\langle \pi_t^n, f\rangle = \frac{1}{n}\sum_{i=1}^{m_j^n} M_j^n(X^i, t) f\big(X^i(t)\big)$, $\langle V_t^n, f\rangle = \eta_t^n \langle \pi_t^n, f\rangle$ and $\langle V_t^n, 1\rangle = \eta_t^n$, which is the likelihood function. Also note that $\eta_{j\delta}$ is $\mathcal{F}_{j\delta-}$-measurable.

Note that in most of the literature of the branching particle filters for the classical nonlinear filtering (cf. [10]), the optimal filter is approximated by the unweighted empirical measure $\tilde{\pi}_t^n$ with the same branching mechanism. Following this, we define

$$\tilde{\pi}_t^n = \frac{1}{n}\sum_{i=1}^{m_j^n} \delta_{X^i(t)}, \quad \text{and} \quad \tilde{V}_t^n = \tilde{\pi}_t^n \eta_t^n, \qquad j\delta \le t < (j+1)\delta.$$

In the rest of this paper, we prove the convergence of π^n or $\tilde{\pi}^n$ to π as $n \to \infty$ and to study the convergence rate of π^n.

4 Convergence of V_t^n

In this section, we consider the convergence of V_t^n to V_t for fixed t. The main idea is to use a backward SPDE as the dual of the Zakai equation. This idea has been applied for classical nonlinear filtering models in [8] and [12].

We consider the backward SPDE:

$$\begin{cases} d\psi_s = -L\psi_s ds - \sum_{k=1}^{w}(ap_k - 1)\psi_{s+}\hat{d}(Y_s^k - s), & 0 \le s \le t \\ \psi_t = \phi \end{cases} \tag{9}$$

where $\hat{d}$ denotes the backward Itô's integral and ϕ is a bounded function. For the backward Itô's integral, we take the right point in the Riemann sum when defining the stochastic integral backwardly.

Define

$$\hat{Y}_s^k = Y_t^k - Y_{t-s}^k \text{ and } \hat{\psi}_s = \psi_{t-s}.$$

Then $\hat{\psi}_s$ satisfies the following forward SPDE

$$\begin{cases} d\hat{\psi}_s = L\hat{\psi}_s ds + \sum_{k=1}^{w}(ap_k - 1)\hat{\psi}_{s-}d(\hat{Y}_s^k - s), & 0 \le s \le t \\ \hat{\psi}_0 = \phi \end{cases} \tag{10}$$

which is the Zakai-tpye equation in Theorem 1. Similar to [26], we can prove the uniqueness for the solution to (10), implying the uniqueness of (9). In fact, we need the following technical estimates with the technical (BD) condition.

Lemma 4.1 *Suppose Assumptions 1 - 5 hold and the boundedness condition below holds: (BD): $a, p_k, \phi \in C_b^1(\mathbb{R})$. And μ', σ' and $\sup_{0\le t\le T} E(\sigma^4(X_t))$ are bounded. Let $\psi_u'(x) = \frac{d}{dx}\psi_u(x)$. Then, there exists a constant K such that*

$$E^P[\sup_{0\le s\le t} \|\psi_s\| + \sup_{0\le s\le t} \|\psi_s'\|] \le K.$$

The proof of Lemma 4.1 is in Appendix. It is easy to check that the (BD) condition allows X_t to be GBM. Before we state and prove the main theorem of this section, we state several useful lemmas with their proofs also in Appendix. Lemma 4.2 is a convolution result and Lemmas 4.3 and 4.5 are key moment estimates crucial for the convergence results and the central limit theorem type result. Lemma 4.4 gives the SDEs of two quantities needed in Lemmas 4.5 and 6.1 later. The order of the key estimate in Lemma 4.5 is $o(\delta)$, which is better than $o(\delta^{1/2})$, the order in the classical nonlinear filtering case (see [6] and [12]).

Lemma 4.2 *Almost surely, we have*

$$\psi_{(j+1)\delta}(X_{(j+1)\delta}^i)M_{j+1}^n(X^i) - \psi_{j\delta}(X_{j\delta}^i) = \int_{j\delta}^{(j+1)\delta} M_j^n(X^i, s)\psi_s'(X_s^i)\sigma(X_s^i)dB_s^i. \tag{11}$$

For the rest of this paper, we use $\mathbb{E}(X) = E^Q(X)$. Under Q, let $\tilde{Y}_k(t) = Y_k(t) - t$.

Lemma 4.3 $\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right) \le K_1 n.$

Lemma 4.4 *Let*

$$\hat{M}_j^n(t) = \frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_j^n(X^\ell,t), \quad \text{and} \quad \tilde{M}_j^n(X^i,t) = \frac{M_j^n(X^i,t)}{\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_j^n(X^\ell,t)} = \frac{M_j^n(X^i,t)}{\hat{M}_j^n(t)}.$$

Then,

$$d\hat{M}_j^n(t) = \hat{M}_j^n(t-)\sum_{k=1}^{w}\bar{h}_k(t-)d\tilde{Y}_k(t) \tag{12}$$

and

$$\tilde{M}_j^n(X^i) = 1 + \int_{j\delta}^{(j+1)\delta}\tilde{M}_j^n(X^i,s-)\sum_{k=1}^{w}\Big[\frac{ap_k(X_{s-}^i,s-)}{\bar{h}_k(s-)+1} - 1\Big]dY_k(s) \tag{13}$$

where

$$\bar{h}_k(s) = \frac{1}{m_j^n}\sum_{i=1}^{m_j^n}\tilde{M}_j^n(X^i,s)(ap_k(X_s^i,s)-1) \tag{14}$$

Lemma 4.5 *Let $F(x) = \{x\}(1-\{x\})$. Then,*

$$\Big|\mathbb{E}\Big(\gamma_{j+1}^n(X^i)f^2(X_{(j+1)\delta}^i)(\eta_{(j+1)\delta}^n/\eta_{j\delta}^n)^2|\mathcal{F}_{j\delta}\Big) - H(X_{j\delta}^i)\delta\Big| = o(\delta),$$

where $o(\delta)\to 0$ as $\delta \to 0$ and $H(x)$ is nonnegative and given by

$$H(X_{j\delta}^i) = \hat{M}_j^n(j\delta)^2 f^2(X_{j\delta}^i)\sum_{k=1}^{w}F\Big(\frac{ap_k(X_{j\delta}^i,j\delta)}{\bar{h}_k(j\delta)+1}\Big)(\bar{h}_k(j\delta)+1)^2. \tag{15}$$

Theorem 4.1 *Suppose that Assumptions 1 -5 hold with conditions (I) and (BD). Then there exists a constant K_1 such that*

$$\mathbb{E}|\left\langle V_t^n,\phi\right\rangle - \left\langle V_t,\phi\right\rangle|^2 \le K_1 n^{-(1-\alpha)}.$$

Proof: Let $k\delta \le t < (k+1)\delta$. Observe that

$$\begin{aligned}
\langle V_t^n,\phi\rangle - \langle V_0^n,\psi_0\rangle &= \langle V_t^n,\psi_t\rangle - \langle V_{k\delta}^n,\psi_{k\delta}\rangle + \sum_{j=1}^{k}\left(\left\langle V_{j\delta}^n,\psi_{j\delta}\right\rangle - \mathbb{E}\left(\left\langle V_{j\delta}^n,\psi_{j\delta}\right\rangle|\mathcal{F}_{j\delta-}\right)\right) \\
&\quad + \sum_{j=1}^{k}\Big(\mathbb{E}\left(\left\langle V_{j\delta}^n,\psi_{j\delta}\right\rangle|\mathcal{F}_{j\delta-}\right) - \Big\langle V_{(j-1)\delta}^n,\psi_{(j-1)\delta}\Big\rangle\Big) \\
&\equiv I_1^n + I_2^n + I_3^n.
\end{aligned} \tag{16}$$

Then

$$I_1^n = \eta_{k\delta}^n\frac{1}{n}\sum_{i=1}^{m_k^n}\left(M_k^n(X^i,t)\psi_t(X_t^i) - \psi_{k\delta}(X_{k\delta}^i)\right),$$

$$I_2^n = \sum_{j=1}^{k}\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))$$

and

$$
\begin{aligned}
I_3^n &= \sum_{j=1}^{k}\left(\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)\tilde{M}_j^n(X^i)-\eta_{(j-1)\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{(j-1)\delta}(X_{(j-1)\delta}^i)\right)\\
&= \sum_{j=1}^{k}\eta_{(j-1)\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\left(\psi_{j\delta}(X_{j\delta}^i)M_j^n(X^i)-\psi_{(j-1)\delta}(X_{(j-1)\delta}^i)\right).
\end{aligned}
$$

Now, it suffices to estimate the following moments. First, we study I_3 term. By Lemma 4.2 and the independent increments of the Brownian motion, we have

$$
\begin{aligned}
\mathbb{E}((I_3^n)^2) &= \mathbb{E}\left(\sum_{j=0}^{k-1}\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}M_j^n(X^i,s)\psi_s'\sigma(X_s^i)dB_s^i\right)^2\\
&= \sum_{j=0}^{k-1}\mathbb{E}\left(\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}M_j^n(X^i,s)\psi_s'\sigma(X_s^i)dB_s^i\right)^2.
\end{aligned}
$$

Let $\mathcal{F}_t=\mathcal{F}_t^B$ be the natural filtration of $B^i,\ i=1,2,\cdots,m_j^n$ up to t. Since $X^i,\ i=1,2,\cdots,m_j^n$ are conditionally (given $\mathcal{F}_{j\delta}\vee\mathcal{F}_t^{\vec{Y}}$) independent, we can continue with

$$
\begin{aligned}
&\mathbb{E}((I_3^n)^2)\\
=& \sum_{j=0}^{k-1}\mathbb{E}\left(\mathbb{E}\left(\left(\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}M_j^n(X^i,s)\psi_s'\sigma(X_s^i)dB_s^i\right)^2\Big|\mathcal{F}_{j\delta}\vee\mathcal{F}_t^{\vec{Y}}\right)(\eta_{j\delta}^n)^2\right)\\
=& \mathbb{E}\sum_{j=0}^{k-1}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}M_j^n(X^i,s)^2|\psi_s'\sigma(X_s^i)|^2(\eta_{j\delta}^n)^2ds\\
\le& \mathbb{E}\sum_{j=0}^{k-1}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}\mathbb{E}\left(||\psi_s'||_\infty(M_j^n(X^i,s))^2\sigma^2(X_s^i)(\eta_{j\delta}^n)^2|\mathcal{F}_{j\delta}\right)ds\\
=& \mathbb{E}\sum_{j=0}^{k-1}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}\mathbb{E}\left(||\psi_s'||_\infty|\mathcal{F}_{j\delta}\right)\mathbb{E}\left((M_j^n(X^i,s))^2\sigma^2(X_s^i)|\mathcal{F}_{j\delta}\right)(\eta_{j\delta}^n)^2ds
\end{aligned}
$$

where the last equality follows from the independent increments of Y and, given $\mathcal{F}_{j\delta}$, $M_j^n(X^i,s)\sigma(X_s^i)$ is $\mathcal{F}_s^{\vec{Y}}\vee\mathcal{F}_{j\delta,s}^i$-measurable and $||\psi_s'||_\infty$ is $\mathcal{F}_{s,t}^{\vec{Y}}$-measurable. Here $\mathcal{F}_{j\delta,s}^i=\sigma(B_t^i-B_{j\delta}^i: j\delta\le t\le s)$ and $\mathcal{F}_{s,t}^{\vec{Y}}=\sigma(\vec{Y}_u-\vec{Y}_s:\ s\le u\le t)$.

Then,

$$
\mathbb{E}\left((M_j^n(X^i,s))^2\sigma^2(X_s^i)|\mathcal{F}_{j\delta}\right)\le\sqrt{\mathbb{E}\left((M_j^n(X^i,s))^4|\mathcal{F}_{j\delta}\right)}\sqrt{\mathbb{E}\left(\sigma^4(X_s^i)|\mathcal{F}_{j\delta}\right)}.
$$

It is easy to show that $\mathbb{E}\left((M_j^n(X^i,s))^4|\mathcal{F}_{j\delta}\right)\le e^{K\delta}$, and using (BD) condition, $\mathbb{E}\left(|\sigma(X_s^i)|^4|\mathcal{F}_{j\delta}\right)\le K_1$, and by the independent increments of Y and Lemma 4.1, $\mathbb{E}\left(\|\psi_s'\|_\infty^2|\mathcal{F}_{j\delta}\right)=\mathbb{E}\left(\|\psi_s'\|_\infty^2\right)\le K_2$.

Hence, using $\sum_{j=0}^{k-1}\delta \le t \le T$ and applying Lemma 4.3, we obtain

$$\mathbb{E}((I_3^n)^2) \le \mathbb{E}\sum_{j=0}^{k-1}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta} K_3(\eta_{j\delta}^n)^2 ds \le K_4 n^{-2}\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right) \le K_5 n^{-1}.$$

Next, we look at I_2 term. Note that for $j < j'$,

$$\begin{aligned}
& \mathbb{E}\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))\frac{1}{n}\sum_{i=1}^{m_{j'-1}^n}\psi_{j'\delta}(X_{j'\delta}^i)(\xi_{j'}^i - \tilde{M}_{j'}^n(X^i))\eta_{j\delta}^n\eta_{j'\delta}^n \\
= \; & \mathbb{E}\bigg(\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))\frac{1}{n}\sum_{i=1}^{m_{j'-1}^n}\psi_{j'\delta}(X_{j'\delta}^i)\mathbb{E}(\xi_{j'}^i - \tilde{M}_{j'}^n(X^i)|\mathcal{F}_{j'\delta-}\vee\mathcal{F}_t^{\vec{Y}})\eta_{j\delta}^n\eta_{j'\delta}^n\bigg) \\
= \; & 0.
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}((I_2^n)^2) &= \mathbb{E}\left|\sum_{j=1}^{k}\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))\right|^2 \\
&= \sum_{j=1}^{k}\mathbb{E}\left(\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))\right)^2 \\
&= \sum_{j=1}^{k}\mathbb{E}\left(\mathbb{E}\left(\left(\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i))\right)^2\middle|\mathcal{F}_{j\delta-}\right)(\eta_{j\delta}^n)^2\right) \\
&= \mathbb{E}\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)^2\gamma_j^n(X^i)(\eta_{j\delta}^n)^2. \\
&\le \mathbb{E}\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\mathbb{E}\left(||\psi_{j\delta}||_\infty^2\middle|\mathcal{F}_{j\delta-}\right)\gamma_j^n(X^i)(\eta_{j\delta}^n)^2. \\
&\le \mathbb{E}\left(\sup_{0\le s\le T}||\psi_s||_\infty^2\right)\mathbb{E}\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\mathbb{E}\Big[\gamma_j^n(X^i)(\eta_{j\delta}^n)^2\Big|\mathcal{F}_{(j-1)\delta}\Big]
\end{aligned}$$

Applying Lemmas 4.1, 4.5 and 4.3, we have

$$\mathbb{E}((I_2^n)^2) \le K_1\sum_{j=1}^{k}\frac{1}{n^2}E\left(m_{j-1}^n(\eta_{(j-1)\delta}^n)^2\right)\delta \le K_2 n^{-1}.$$

I_1^n can be estimated similar to I_3^n. ∎

Remark 4.1 *For the case of $\tilde{\pi}_t^n$ and $\tilde{V}_t^n$, we have*

$$\begin{aligned}\left\langle \tilde{V}_t^n, \psi_t \right\rangle - \left\langle \tilde{V}_{k\delta}^n, \psi_{k\delta} \right\rangle &= \frac{1}{n} \sum_{i=1}^{m_k^n} \left(\psi_t(X_t^i) M_k^n(X^i, t) - \psi_{k\delta}(X_{k\delta}^i) \right) \\ &\quad + \frac{1}{n} \sum_{i=1}^{m_k^n} \phi(X_t^i) \left(1 - M_k^n(X^i, t)\right).\end{aligned}$$

Namely, the second term is an extra term. It can be proved that its second moment is bounded by $Kn^{-2\alpha}$. Therefore, we have

$$\mathbb{E}| \left\langle \tilde{V}_t^n, \phi \right\rangle - \langle V_t, \phi \rangle |^2 \leq K_1 \left(n^{-1} \vee n^{-2\alpha} \right).$$

5 Convergence of V^n

In this section, we study the convergence of V^n, regarding as a sequence of stochastic processes. More specifically, we prove the convergence uniformly for t in an interval $[0, T]$.

The main idea of this section is to obtain an equation for the process V_t^n and then to derive maximum inequality making use of the martingale theory.

First we consider the equation satisfied by V_t^n. Let $j\delta < t < (j+1)\delta$. By Itô's formula, we have

$$d \langle V_t^n, f \rangle = \langle V_t^n, Lf \rangle dt + \frac{1}{n} \sum_{i=1}^{m_j^n} M_j^n(X^i, t) f' \sigma(X_t^i) dB_t^i \eta_{j\delta}^n + \sum_{k=1}^{w} \langle V_t^n, f(ap_k - 1) \rangle d\tilde{Y}_k(t).$$

The jump at $(j+1)\delta$ is

$$\begin{aligned}& \eta_{(j+1)\delta}^n \frac{1}{n} \sum_{i=1}^{m_j^n} \xi_{j+1}^i \delta_{X_{(j+1)\delta}^i} - \eta_{j\delta}^n \frac{1}{n} \sum_{i=1}^{m_j^n} M_{j+1}^n(X^i)) \delta_{X_{(j+1)\delta}^i} \\ = \; & \eta_{(j+1)\delta}^n \frac{1}{n} \sum_{i=1}^{m_j^n} \left(\xi_{j+1}^i - \tilde{M}_{j+1}^n(X^i) \right) \delta_{X_{(j+1)\delta}^i}.\end{aligned}$$

Therefore,

$$\langle V_t^n, f \rangle = \langle V_0, f \rangle + \int_0^t \langle V_s^n, Lf \rangle ds + \sum_{k=1}^{w} \int_0^t \langle V_s^n, f(ap_k - 1) \rangle d\tilde{Y}_k(s) + N_t^{n,f} + \hat{N}_t^{n,f}, \tag{17}$$

where

$$N_t^{n,f} = \sum_{j=0}^{[t/\delta]} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta) \wedge t} f' \sigma(X_s^i) dB_s^i \eta_{j\delta}^n$$

and

$$\hat{N}_t^{n,f} = \sum_{j=1}^{[t/\delta]} \eta_{j\delta}^n \frac{1}{n} \sum_{i=1}^{m_{j-1}^n} (\xi_j^i - \tilde{M}_j^n(X^i)) f(X_{j\delta}^i).$$

It is easy to see that $N_t^{n,f}$, $\hat{N}_t^{n,f}$ are two uncorrelated martingales with quadratic variational processes

$$\left\langle N^{n,f}\right\rangle_t = \sum_{j=0}^{[t/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)\wedge t} |f'\sigma(X_s^i)|^2 ds (\eta_{j\delta}^n)^2$$

and

$$\begin{aligned} \left\langle \hat{N}^{n,f}\right\rangle_t = \left\langle \hat{N}^{n,f}\right\rangle_{[t/\delta]\delta} &= \sum_{j=1}^{[t/\delta]} \frac{1}{n^2} \mathbb{E}\left(\left(\sum_{i=1}^{m_{j-1}^n} (\xi_j^i - M_j^n(X^i)) f(X_{j\delta}^i)\right)^2 |\mathcal{F}_{j\delta-}\right)(\eta_{j\delta}^n)^2 \\ &= \sum_{j=1}^{[t/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_{j-1}^n} \gamma_j^n(X^i) f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2. \end{aligned} \tag{18}$$

Define the usual distance

$$d(\nu_1,\nu_2) = \sum_{k=1}^{\infty} 2^{-k} \left(|\left\langle \nu_1 - \nu_2, f_k \right\rangle| \wedge 1\right)$$

where $\{f_k\}$ satisfy the following conditions: $f_k \in C_b^2(\mathbb{R})$ with $\|Lf_k\|_\infty \le 1$. Moreover, we define

$$\tilde{d}(\nu_1,\nu_2) = \sum_{k=1}^{\infty} 2^{-k} \left(|\left\langle \nu_1 - \nu_2, f_k \right\rangle| \right)$$

with the same assumptions on $\{f_k\}$. Obviously, $d \le \tilde{d}$, but $\tilde{d}$ may not be a distance.

Theorem 5.1 *Under the assumptions of Theorem 4.1, there exists a constant K_1 such that*

$$\mathbb{E} \sup_{t\le T} d(V_t^n, V_t)^2 \le K_1 n^{-1}.$$

Proof: Note that

$$\mathbb{E}\sup_{t\le T} \tilde{d}(V_t^n, V_t)^2 \le \sum_{k=1}^{\infty} 2^{-k} \left(\mathbb{E}\sup_{t\le T} \left\langle V_t^n - V_t, f_k\right\rangle^2\right) + \mathbb{E}\sup_{t\le T} \left\langle V_t^n - V_t, 1\right\rangle^2 \tag{19}$$

By Equation (17) and Doob's maximum inequality,

$$\begin{aligned} &\mathbb{E}\sup_{t\le T} \left\langle V_t^n - V_t, f\right\rangle^2 \\ \le& K_2 \int_0^T \mathbb{E}\left\langle V_t^n - V_t, Lf\right\rangle^2 dt + K_2 \mathbb{E} \sum_{j=0}^{[T/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)} |f'\sigma(X_s^i)|^2 ds (\eta_{j\delta}^n)^2 \\ &+ K_2 \sum_{k=1}^{w} \int_0^T \mathbb{E}\left\langle V_t^n - V_t, f(ap_k - 1)\right\rangle^2 dt + K_2 \mathbb{E} \sum_{j=1}^{[T/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \gamma_j^n(X^i) f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2. \end{aligned} \tag{20}$$

By Theorem 4.1, the first and third terms are bounded by $K_3 n^{-1}$. By Lemma 4.3,

$$\text{2nd term} \le K_3 \sum_{j=1}^{[T/\delta]} \frac{\delta}{n^2} \mathbb{E}\left(m_j^n (\eta_{j\delta}^n)^2 \right) \le K_4 n^{-1}.$$

By Lemma 4.5, we have

$$\text{4th term} \le K_5 \sum_{j=0}^{[T/\delta]} \frac{\delta}{n^2} \mathbb{E}\left(m_j^n (\eta_{j\delta}^n)^2 \right) \le K_6 n^{-1}.$$

Finally, we consider the last term in (19). Take $f = 1$ in (20), we have

$$\begin{aligned} &\mathbb{E} \sup_{t\le T} \langle V_t^n - V_t, 1 \rangle^2 \\ \le & K_2 \sum_{k=1}^{w} \int_0^T \mathbb{E} \langle V_t^n - V_t, (ap_k - 1) \rangle^2 \, dt + K_2 \mathbb{E} \sum_{j=1}^{[T/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \gamma_j^n (X^i) (\eta_{j\delta}^n)^2. \end{aligned} \tag{21}$$

Again, Theorem 4.1 implies that the first term is bounded by $K_7 n^{-1}$. Clearly, Lemma 4.5 is true with $f = 1$, and a similar argument implies the second term of (21) is bounded by $K_8 n^{-1}$. Putting all the above estimates back to (20), we establish the desired result, since $d \le \tilde{d}$. ■

Remark 5.1 *For the case of $\tilde{V}_t^n$, the jump at $(j+1)\delta$ is*

$$\eta_{(j+1)\delta}^n \frac{1}{n} \sum_{i=1}^{m_j^n} \left(\xi_{j+1}^i - \frac{\eta_{j\delta}^n}{\eta_{(j+1)\delta}^n} \right) \delta_{X_{(j+1)\delta}^i}.$$

Write

$$\xi_{j+1}^i - \frac{\eta_{j\delta}^n}{\eta_{(j+1)\delta}^n} = \left(\xi_{j+1}^i - \tilde{M}_j^n (X^i) \right) + \frac{M_j^n (X^i) - 1}{\frac{1}{m_j^n} \sum_{k=1}^{m_j^n} M_j^n (X^k)}.$$

Then the new $\hat{N}^{n,f}$ can be written as two terms. A careful estimate of the second term leads to the bound $K n^{-2\alpha}$. Thus, we have

$$\mathbb{E} \sup_{t\le T} d(\tilde{V}_t^n, V_t)^2 \le K \left(n^{-2\alpha} \vee n^{-1} \right).$$

Now, we convert the convergence result for π^n (the convergence of $\tilde{\pi}^n$ can be obtained similarly).

Theorem 5.2 *Under the assumptions of Theorem 4.1, there exists a constant K such that*

$$E^P \sup_{0\le t\le T} d(\pi_t^n, \pi_t) \le K n^{-\frac{1}{2}}. \tag{22}$$

Proof: Note that for f bounded by 1, we have

$$\begin{aligned} |\langle \pi_t^n - \pi_t, f \rangle| &= \left| \frac{\langle V_t^n, f \rangle \langle V_t - V_t^n, 1 \rangle + \langle V_t^n, 1 \rangle \langle V_t^n - V_t, f \rangle}{\langle V_t^n, 1 \rangle \langle V_t, 1 \rangle} \right| \\ &\le \frac{|\langle V_t - V_t^n, 1 \rangle|}{\langle V_t, 1 \rangle} + \frac{|\langle V_t^n - V_t, f \rangle|}{\langle V_t, 1 \rangle}. \end{aligned} \tag{23}$$

Thus

$$d(\pi_t^n, \pi_t) \le \frac{1}{\langle V_t, 1\rangle} | \langle V_t - V^n, 1\rangle | + \frac{1}{\langle V_t, 1\rangle} \tilde{d}(V_t^n, V_t).$$

Now,

$$\begin{aligned} E^P \sup_{0\le t\le T} d(\pi_t^n, \pi_t) &= \mathbb{E} \sup_{0\le t\le T} \left\{ \frac{1}{\langle V_t, 1\rangle} | \langle V_t - V^n, 1\rangle | + \frac{1}{\langle V_t, 1\rangle} \tilde{d}(V_t^n, V_t) \right\} M_T \\ &\le \left(\mathbb{E} \sup_{0\le t\le T} | \langle V_t - V^n, 1\rangle |^2 \right)^{\frac{1}{2}} \left(\mathbb{E} \sup_{0\le t\le T} \frac{M_T^2}{\langle V_t, 1\rangle^2} \right)^{\frac{1}{2}} \\ &\quad + \left(\mathbb{E} \sup_{0\le t\le T} \tilde{d}(V_t^n, V_t)^2 \right)^{\frac{1}{2}} \left(\mathbb{E} \sup_{0\le t\le T} \frac{M_T^2}{\langle V_t, 1\rangle^2} \right)^{\frac{1}{2}}. \end{aligned} \tag{24}$$

With Assumption 4 and the SDEs for M_t and $\langle V_t, 1\rangle$, it is straightforward to prove that $\hat{\mathbb{E}} M_T^4 < \infty$ and

$$\mathbb{E} \sup_{0\le t\le T} \langle V_t, 1\rangle^{-4} < \infty.$$

Thus, by Theorem 5.1 and (24), there is a constant K such that (22) holds. ■

6 A Central Limit Type Theorem

In this section, we prove the exact rate of convergence by a central limit type theorem. Let

$$U_t^n = n^{\frac{1}{2}} (V_t^n - V_t).$$

We first prove tightness for $\{U_n\}$ in an appropriate space and then characterize the limit and obtain a central limit type theorem. The exact rate of convergence for the FM model is $n^{\frac{1}{2}}$ which is better than that for the classical filtering model, which is $n^{(1-\alpha)/2}$ for $\alpha > 0$ (see [12]).

6.1 The Modified Schwartz Space and Tightness of $\{U_n\}$

As in Hitsuda and Mitoma [23], we use the modified Schwartz space Φ. Let $\rho(x) = K_1 1_{\{|x|<1\}} \exp\left(-1/(1-|x|^2)\right)$, where K_1 is a constant such that $\int \rho(x)dx = 1$. Let $\psi(x) = \int e^{-|y|} \rho(x-y)dy$. Then for any integer k and $e = \psi^{-1}$, we have $|e^{(k)}(x)| \le K_2(k)(1 + e^{|x|})$. Let $\Phi = \{\phi : \phi\psi \in \mathcal{S}\}$, where $\mathcal{S}$ is the Schwartz space. For $\kappa = 0, 1, 2, \ldots$, define

$$\|\phi\|_\kappa^2 = \sum_{0\le |k|\le\kappa} \int_{\mathbb{R}} (1+|x|^2)^{2\kappa} \left| \frac{\partial^k}{\partial x^k} (\phi(x)\psi(x)) \right|^2 dx$$

the k above is a multi-index $(k_1, \cdots, k_d)$ with $|k| = k_1 + \cdots + k_d$. Let Φ_κ be the completion of Φ with respect to $\|\cdot\|_\kappa$. Then Φ_κ is a Hilbert space with inner product

$$\langle \phi_1, \phi_2\rangle_\kappa = \sum_{0\le |k|\le\kappa} \int_{\mathbb{R}} (1+|x|^2)^{2\kappa} \left(\frac{\partial^k}{\partial x^k} (\phi_1(x)\psi(x)) \right) \left(\frac{\partial^k}{\partial x^k} (\phi_2(x)\psi(x)) \right) dx.$$

Note that $\Phi_\kappa \supset \Phi_{\kappa+1}$ and that Φ_0 is $L^2(\mu_\psi)$, where $\mu_\psi(dx) = \psi^2(x)dx$. For $\hat{\phi} \in \Phi_0$ and $\phi \in \Phi_\kappa$,

$$\langle \hat{\phi}, \phi \rangle \equiv \langle \hat{\phi}, \phi \rangle_0 = \int_{\mathbb{R}} \hat{\phi}(x)\phi(x)\psi^2(x)dx$$

defines a continuous linear functional on Φ_κ with norm

$$\|\hat{\phi}\|_{-\kappa} = \sup_{\phi \in \Phi_\kappa} \frac{|\langle \hat{\phi}, \phi \rangle|}{\|\phi\|_\kappa},$$

and we let $\Phi_{-\kappa}$ denote the completion of Φ_0 with respect to this norm. Then $\Phi_{-\kappa}$ is a representation of the dual of Φ_κ. If $\{\phi_j^\kappa\}$ is a complete, orthonormal system for Φ_κ, then the inner product for $\Phi_{-\kappa}$ can be written as

$$\langle \hat{\phi}_1, \hat{\phi}_2 \rangle_{-\kappa} = \sum_{j=1}^{\infty} \langle \hat{\phi}_1, \phi_j^\kappa \rangle \langle \hat{\phi}_2, \phi_j^\kappa \rangle. \tag{25}$$

By a slight modification of Theorem 7, page 82, of [20], these norms determine a nuclear space, so in particular, for each κ there exists a $\kappa' > \kappa$ such that the embedding $T_\kappa^{\kappa'} : \Phi_{\kappa'} \to \Phi_\kappa$ is a Hilbert-Schmidt operator. The adjoint $T_\kappa^{\kappa' *} : \Phi_{-\kappa} \to \Phi_{-\kappa'}$ is also Hilbert-Schmidt. $\Phi' = \cup_{k=0}^{\infty} \Phi_{-k}$ gives a representation of the dual of Φ (see [20], page 59). Next, we prove tightness for $\{U^n\}$ in $D_{\Phi_{-\kappa}}[0,\infty)$ for an appropriate κ.

By (17) and Zakai equation, we have

$$\begin{aligned}\langle U_t^n, f \rangle &= \langle U_0^n, f \rangle + \int_0^t \langle U_s^n, Lf \rangle \, ds + \sum_{k=1}^{w} \int_0^t \langle U_s^n, f(ap_k - 1) \rangle \, d\tilde{Y}_k(s) \\ &\quad + n^{\frac{1}{2}} N_t^{n,f} + n^{\frac{1}{2}} \hat{N}_t^{n,f},\end{aligned} \tag{26}$$

Theorem 6.1 *Under the assumptions of Theorem 4.1, there exists κ such that $\{U^n\}$ is tight in $D_{\Phi_{-\kappa}}[0,\infty)$.*

Proof: For $u \leq \epsilon$, we have

$$\mathbb{E}\left(\left\langle U_{t+u}^n - U_t^n, f \right\rangle^2 \Big| \mathcal{F}_t \right) \leq \mathbb{E}\left(\sum_{i=1}^{4} \zeta_f^{i,n}(\epsilon) \right)$$

where

$$\zeta_f^{1,n}(\epsilon) = \int_t^{t+\epsilon} \langle U_s^n, Lf \rangle^2 \, ds,$$

$$\zeta_f^{2,n}(\epsilon) = \int_t^{t+\epsilon} \langle U_s^n, \nabla^* f c + hf \rangle^2 \, ds,$$

$$\zeta_f^{3,n}(\epsilon) = n \sum_{t \leq j\delta < t+\epsilon} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)\wedge t} |\nabla^* f \sigma(X_s^i)|^2 ds (\eta_{j\delta}^n)^2$$

and

$$\zeta_f^{4,n}(\epsilon) = n \sum_{t \leq j\delta < t+\epsilon} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \gamma_j^2(X^i) f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2.$$

Similar to the previous section, we can show that

$$\lim_{\epsilon\to 0}\sup_n \mathbb{E}\left(\sum_{i=1}^{4}\zeta_f^{i,n}(\epsilon)\right)=0.$$

Then, using Remark 8.7 (p.138) in Ethier and Kurtz [19], we prove the tightness of $\langle U^n, f\rangle$ in $D_{\mathbb{R}}[0,\infty)$. Finally, following the proof of Theorem 3.1 in [27] and applying Mitoma' theorem, we prove that there exists a κ such that the tightness of U^n holds in $D_{\Phi_{-\kappa}}[0,\infty)$. ■

6.2 Characterization of the Limits

It is easy to show that

$$n^{\frac{1}{2}}N_t^{n,f}\to 0. \tag{27}$$

To characterize the limit of $n^{1/2}\hat{N}^{n,f}$, we need the following two more technical estimates in two lemmas with their proofs in Appendix. The other key estimate in Lemma 6.1 is in the order of $O(\delta^{3/2})$, which is better than $O(\delta)$, the order in the classical nonlinear filtering case (see [12]).

Lemma 6.1

$$\left|\mathbb{E}\Big(\gamma_{j+1}^n(X^i)^2(\eta_{(j+1)\delta}^n/\eta_{j\delta}^n)^4|\mathcal{F}_{j\delta}\Big)\right|\le K\delta^{3/2}.$$

Lemma 6.2

$$\mathbb{E}\big((m_j^n)^2(\eta_{j\delta}^n)^4\big)\le Kn^2.$$

Note that

$$\begin{aligned}
\left\langle n^{1/2}\hat{N}^{n,f}\right\rangle_t &= n\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2\\
&= n\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\mathbb{E}\Big(\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2\Big|\mathcal{F}_{(j-1)\delta}\Big)\\
&\quad +n\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\Big(\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2-\mathbb{E}\Big(\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2\Big|\mathcal{F}_{(j-1)\delta}\Big)\Big)
\end{aligned}$$

By Lemma 4.5, the first term is approximated by

$$\begin{aligned}
& n\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}H(X_{(j-1)\delta}^i)\delta f^2(X_{(j-1)\delta}^i)(\eta_{(j-1)\delta}^n)^2\\
=\ & \sum_{j=1}^{[t/\delta]}\left\langle\tilde{V}_{j\delta}^n, Hf^2(X_{j\delta}^i)\right\rangle\left\langle\tilde{V}_{j\delta}^n,1\right\rangle\delta\\
\to\ & \int_0^t\left\langle V_s,Hf^2\right\rangle\langle V_s,1\rangle\,ds,
\end{aligned}$$

where the approximation means the difference tends to 0 as $n\to\infty$.

The second moment of the second term is bounded by

$$\begin{aligned}
& n^2\sum_{j=1}^{[t/\delta]}\mathbb{E}\Big(\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)f^2(X^i_{j\delta})(\eta^n_{j\delta})^2\Big)^2\\
\le\ & ||f||^4_\infty n^{-2}\sum_{j=1}^{[t/\delta]}\mathbb{E}\Big(m^n_{j-1}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)^2(\eta^n_{j\delta})^4\Big)\\
\le\ & K_1n^{-2}\sum_{j=1}^{[t/\delta]}\delta^{3/2}\mathbb{E}\big(K(m^n_{j-1})^2(\eta^n_{(j-1)\delta})^4\big)\\
\le\ & K_2\delta^{1/2}\to 0
\end{aligned}$$

Lemma 6.1 is applied in the second inequality and Lemma 6.2 in the last inequality. Combining the above results, we obtain

Lemma 6.3
$$n^{\frac12}\hat N^{n,f}_t\Longrightarrow M^f_t$$
which is a martingale uncorrelated to B and $\vec{Y}$ such that

$$\left\langle M^f_t\right\rangle_t=\int_0^t\left\langle V_s,Hf^2\right\rangle\langle V_s,1\rangle\,ds.$$

Further, there exists a space-time white noise $W(dtdx)$ (independent of B and $\vec{Y}$ such that

$$M^f_t=\int_0^t\int_{\mathbb{R}}\sqrt{H(x)V(s,x)\langle V_s,1\rangle}f(x)W(dsdx).$$

Summarizing these, we obtain

Theorem 6.2 *Under the assumptions of Theorem 4.1, $U^n\Rightarrow U$ which is the unique solution to:*

$$\begin{aligned}
\langle U_t,f\rangle &= \langle U_0,f\rangle+\int_0^t\langle U_s,Lf\rangle\,ds+\sum_{k=1}^{w}\int_0^t\langle U_{s-},f(ap_k-1)\rangle\,d\tilde{Y}_k(s)\\
&\quad+\int_0^t\int_{\mathbb{R}}\sqrt{H(x)V(s,x)\langle V_s,1\rangle}f(x)W(dsdx). \qquad (28)
\end{aligned}$$

Proof: By Theorem 6.1, we can take U being a limit point. Without loss of generality, we assume that $U^n\Longrightarrow U$. By Lemma 6.3 and (27), it is easy to show that U satisfies (28). To prove the uniqueness, we take another solution $\tilde U$ of (28) and define $\hat U_t=U_t-\tilde U_t$. Then $\hat U_t$ satisfies the following homogeneous linear equation

$$\left\langle\hat U_t,f\right\rangle=\int_0^t\left\langle\hat U_s,Lf\right\rangle ds+\sum_{k=1}^{w}\int_0^t\left\langle\hat U_s,f(ap_k-1)\right\rangle d\tilde{Y}_k(s).$$

Similar to Lemma 4.2 in [27] we get $\hat U=0$. ■

Finally, we convert the convergence result to that for the optimal filter.

Theorem 6.3 *Under the assumptions of Theorem 4.1, $n^{\frac{1}{2}}(\pi_t^n - \pi_t) \Rightarrow \zeta_t$ which is the unique solution to:*

$$\begin{aligned} d\langle \zeta_t, f\rangle =& \langle \zeta_t, Lf - (a-w)f - f\langle \pi_t, a-w\rangle + (a-w)\langle \pi_t, f\rangle\rangle\, dt \\ &+ \sum_{k=1}^{w}\Big[\frac{\langle \zeta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \frac{\langle \zeta_{t-}, ap_k\rangle\langle \pi_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle^2} - \langle \zeta_{t-}, f\rangle\Big] dY_k(t) \\ &+ \int_{\mathbb{R}} \frac{f(x) - \langle \pi_t, f\rangle}{\langle V_t, 1\rangle}\sqrt{H(x)V(t,x)\langle V_t, 1\rangle}\, W(dxdt). \end{aligned} \tag{29}$$

When $a(X_t, t) = a(t)$, depending only on time t, Equation (29) is simplified as below:

$$\begin{aligned} d\langle \zeta_t, f\rangle =& \langle \zeta_t, Lf\rangle\, dt + \int_{\mathbb{R}} \frac{f(x) - \langle \pi_t, f\rangle}{\langle V_t, 1\rangle}\sqrt{H(x)V(t,x)\langle V_t, 1\rangle}\, W(dxdt) \\ &+ \sum_{k=1}^{w}\Big[\frac{\langle \zeta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \frac{\langle \zeta_{t-}, ap_k\rangle\langle \pi_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle^2} - \langle \zeta_{t-}, f\rangle\Big] dY_k(t). \end{aligned} \tag{30}$$

Proof: From Equation (23), we can see that

$$n^{\frac{1}{2}}(\pi_t^n - \pi_t) = (V_t, 1)^{-1}U_t^n - (V_t^n 1 V_t 1)^{-1}(U_t^n 1)V_t^n$$

which converges to

$$\zeta_t \equiv (V_t 1)^{-1}\left(U_t - (V_t 1)^{-1}(U_t 1)V_t\right).$$

Let $\eta_t = (V_t 1)^{-1}U_t$. By Itô's formula for $\langle \eta_t, f\rangle = \langle U_t, f\rangle / \langle V_t, 1\rangle$, we have the following equation for η_t.

$$\begin{aligned} d\langle \eta_t, f\rangle =& \big(\langle \eta_t, Lf - (a-w)f\rangle + \langle \eta_t, f\rangle\langle \pi_t, a-w\rangle\big) dt \\ &+ \sum_{k=1}^{w}\Big[\frac{\langle \eta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \langle \eta_{t-}, f\rangle\Big] dY_k(t) \\ &+ \int_{\mathbb{R}} \frac{f(x)}{\langle V_t, 1\rangle}\sqrt{H(x)V(t,x)\langle V_t, 1\rangle}\, W(dxdt). \end{aligned} \tag{31}$$

When $a(X_t, t) = a(t)$, the above equation is simplified as:

$$\begin{aligned} d\langle \eta_t, f\rangle =& \langle \eta_t, Lf\rangle\, dt + \sum_{k=1}^{w}\Big[\frac{\langle \eta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \langle \eta_{t-}, f\rangle\Big] dY_k(t) \\ &+ \int_{\mathbb{R}} \frac{f(x)}{\langle V_t, 1\rangle}\sqrt{H(x)V(t,x)\langle V_t, 1\rangle}\, W(dxdt). \end{aligned} \tag{32}$$

Observe that $\zeta_t = \eta_t - (\eta_t 1)\pi_t$. Applying Itô's formula again, we get Equation (29) for ζ. When $a(X_t, t) = a(t)$, the simplified (32) and (6) gives (30). The uniqueness comes from the similar argument of Theorem 6.2. ■

Easy to check that $\langle \zeta_t, 1\rangle = 0$ for all t in (29) and (30), which is a necessary condition for ζ_t.

7 Conclusions

In this paper, we study the branching particle filters to a FM model, which well fit the stylized facts of ultra-high frequency data in financial markets. We construct a branching particle system and its weighted empirical measure. Then, we prove the uniform convergence of the branching particle filters to the optimal filters. Moreover, we study the convergence rate by proving a central limit type theorem. We find out the rate is $n^{1/2}$, which is better than the best rate in the classical nonlinear filtering case.

The branching particle filter developed in this paper can be directly applied to calculate the MSE estimate of X_t, which is important in asset pricing. Moreover, the filter developed can be used to estimate locally risk-minimizing hedging strategy for FM models derived in [28] and the optimal trading strategy for mean-variance portfolio selection problem of the FM models derived in [30].

Future works include studying the large deviation principle of V^n and π^n as the classical nonlinear filtering case in [14], and studying the branching approximation in a more general framework such as X_t becomes a stochastic volatility model (even with jumps) or a general Markov process. See [35] for a very general FM model with statistical analysis. The branching particle filters developed in this paper only estimates X_t. It is intriguing to study branching particle filters for both (X_t, θ_t), where θ_t is the parameter (allowing time-dependent) in a FM model. These topics are currently under investigation by the authors.

A Appendix: Related Proofs

Proof: (for Lemma 4.1) Let $N(t)$ be the counting process for the jumps in $\vec{Y}(t)$. Let $\tau_1, \tau_2, \cdots, \tau_{N(t)}$ be the jump times of $N(t)$ such that $t \geq \tau_1 > \tau_2 \cdots > \tau_{N(t)} > 0$. For $s \in [t, \tau_1)$, there is no jump and (9) reduces to

$$d\psi_s = -L\psi_s ds + \sum_{k=1}^{w}(ap_k(X_s) - 1)\psi_s ds. \tag{33}$$

Feynman-Kac Formula ([31]) and the boundedness of $a(x,t)$ and (BD) condition implies

$$\sup_{\tau_1 < s \leq t} \|\psi_s\|_\infty \leq e^{wC_1(t-\tau_1)}\|\phi\|_\infty$$

After a jump happens at τ_1, $\psi_{\tau_1-} = ap_k\psi_{\tau_1+}$. Hence, $\sup_{\{\tau_1 < s \leq t\}\cup\{\tau_1-\}} \|\psi_s\| \leq K_2\|\phi\|_\infty$ $e^{wK_2(t-\tau_1)}$. By induction, we have

$$\sup_{0 \leq s \leq t} \|\psi_s\| \leq K_2^{N(t)}\|\phi\|_\infty e^{wK_2 t}.$$

Taking expectation, the result for the part of ψ follows.

To obtain the result for ψ', we differentiate Equation (33) with respect to x and obtain

$$d\psi'_s = -L_1\psi'_s ds + \left[\mu' + \sum_{k=1}^{w}(ap_k(X_s) - 1)\right]\psi_s ds + \sum_{k=1}^{w} a' p'_k \psi_s ds$$

where

$$L_1 f(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 f}{\partial x^2}(x) + (\sigma(x)\sigma'(x) + \mu(x))\frac{\partial f}{\partial x}(x).$$

Then, we repeat the steps for ψ to obtain the desired result for ψ'. ■

Proof: (for Lemma 4.2) After simplifying notations, it is equivalent to proving:

$$\psi_t(X_t)M_t - \psi_0(X_0) = \int_0^t M_s \psi'_s \sigma(X_s) dB_s. \tag{34}$$

Let f_k, $k = 1, 2, ..., w$ and g be bounded functions on $[0, t]$,

$$\theta_f^{\vec{Y}}(r) = \prod_{k=1}^{w} \exp\left\{\sqrt{-1}\int_0^r \log f_k(s-) dY_k(s) - \int_0^r (f_k(s) - 1) ds\right\}$$

and

$$\theta_g^B(r) = \exp\left(\sqrt{-1}\int_0^r g_s dB_s + \frac{1}{2}\int_0^r g_s^2 ds\right).$$

First, we need a lemma, whose proof is identical to that of Lemma 4.1.4 in [3, page 81].

Lemma A.1 *If $\xi \in L^2(\Omega, \mathcal{F}_t^{B,\vec{Y}}, \hat{P})$ and for bounded f_k, $k = 1, 2, ..., w$ and g on $[0, t]$,*

$$\mathbb{E}\left(\xi \theta_f^{\vec{Y}}(t) \theta_g^B(t)\right) = 0,$$

then $\xi = 0$ a.s.

By Lemma A.1, it is sufficient to show that

$$\mathbb{E}\left(\left(\psi_t(X_t)M_t - \psi_0(X_0)\right)\theta_f^{\vec{Y}}(t)\theta_g^B(t)\right) = \mathbb{E}\left(\int_0^t M_s \nabla^* \psi_s \tilde{c}(X_s) dB_s \theta_f^{\vec{Y}}(t)\theta_g^B(t)\right).$$

First we observe that for $r \geq 0$,

$$\mathbb{E}\left(\psi_r(X_r)M_r\theta_f^{\vec{Y}}(t)\theta_g^B(t)|\mathcal{F}_r^{\vec{Y}} \vee \mathcal{F}_r^B\right) = \Theta_r(X_r)M_r\theta_f^{\vec{Y}}(r)\theta_g^B(r) \tag{35}$$

where

$$\Theta_r = \mathbb{E}\left(\psi_r\tilde{\theta}_f(r)|\mathcal{F}_r^{\vec{Y}} \vee \mathcal{F}_r^B\right)$$

with

$$\tilde{\theta}_f(r) = \theta_f^{\vec{Y}}(t)/\theta_f^{\vec{Y}}(r) = \prod_{k=1}^{w} \exp\left(\sqrt{-1}\int_r^t \log f_k(s-) dY_k(s) - \int_r^t (f_k(s) - 1) ds\right).$$

Since ψ_r and $\tilde{\theta}_f(r)$ are measurable with respect to the σ-field $\mathcal{F}_{r,t}^{\vec{Y}} = \sigma(\vec{Y}_s - \vec{Y}_r : r \leq s \leq t)$, which is independent of $\mathcal{F}_r^{\vec{Y}} \vee \mathcal{F}_r^B$, we get that

$$\Theta_r = \hat{\mathbb{E}}\left(\psi_r\tilde{\theta}_f(r)\right).$$

Applying backward Itô's formula, we have

$$\hat{d}\tilde{\theta}_f(r) = \sqrt{-1}\tilde{\theta}_f(r+)\sum_{k=1}^{w}(f_k(r+) - 1)\hat{d}\tilde{Y}_k(r).$$

where $\tilde{Y}(r) = Y(r) - r$. Again applying backward Itô's formula, we get

$$\begin{aligned}\hat{d}(\psi_r\tilde{\theta}_f(r)) = \quad & [-L\psi_r - \sqrt{-1}\psi_r \textstyle\sum_{k=1}^w (f_k(r)-1)(ap_k(X_r,r)-1)]\tilde{\theta}_f(r)dr \\ & + \textstyle\sum_{k=1}^w [\sqrt{-1}(f_k(r+)-1) - (ap_k(X_{r+},r+)-1)]\psi_{r+}\tilde{\theta}_f(r+)\hat{d}\tilde{Y}_k(r) \\ & - \textstyle\sum_{k=1}^w \sqrt{-1}(f_k(r+)-1)(ap_k(X_{r+},r+)-1)\psi_{r+}\tilde{\theta}_f(r+)\hat{d}\tilde{Y}_k(r).\end{aligned}$$

Thus

$$d\Theta_r = \left(-L\Theta_r(X_r) - \sqrt{-1}\sum_{k=1}^w (f_k(r)-1)(ap_k(X_r,r)-1)\Theta_r(X_r)\right)dr.$$

By Itô's formula, we have

$$d\Theta_r(X_r) \quad = \quad -\sqrt{-1}\left(\sum_{k=1}^w (f_k(r)-1)(ap_k(X_r,r)-1)\Theta_r(X_r)\right)dr + \Theta_r'\sigma(X_r)dB_r. \tag{36}$$

Note that

$$dM_r = \sum_{k=1}^w [ap_k(X_r,r)-1]M_r d\tilde{Y}_r,$$

$$d\theta_f^{\vec{Y}}(r) = \sqrt{-1}\theta_f^{\vec{Y}}(r-)\sum_{k=1}^2 (f_k(r-)-1)d\tilde{Y}_r$$

and

$$d\theta_g^B(r) = \sqrt{-1}\theta_g^B(r)g_r dB_r.$$

Apply Itô's formula to the four equations above, we get

$$d(\Theta_r(X_r)M_r\theta_f^{\vec{Y}}(r)\theta_g^B(r)) = \sqrt{-1}\Theta_r'\sigma(X_r)g_r M_r\theta_f^{\vec{Y}}(r)\theta_g^B(r)dr + d(mart.)$$

Combining with (35), we get

$$\begin{aligned} & \mathbb{E}\left((\psi_t(X_t)M_t - \psi_0(X_0))\,\theta_f^{\vec{Y}}(t)\theta_g^B(t)\right) \\ = \quad & \mathbb{E}\left(\Theta_\delta(X_t)M_t\theta_f^{\vec{Y}}(\delta)\theta_g^B(\delta) - \Theta_0(X_0)\theta_f^{\vec{Y}}(0)\theta_g^B(0)\right) \\ = \quad & \sqrt{-1}\int_0^t \mathbb{E}\left(M_r\theta_f^{\vec{Y}}(r)\theta_g^B(r)\Theta_r'\sigma(X_r)g_r\right)dr.\end{aligned}$$

On the other hand,

$$\mathbb{E}\left(\int_0^r M_s\nabla^*\psi_s\tilde{c}(X_s)dB_s\theta_f^{\vec{Y}}(t)\theta_g^B(t)|\mathcal{F}_t^{\vec{Y}}\vee\mathcal{F}_r^B\right) = \int_0^r M_s\nabla^*\psi_s\tilde{c}(X_s)dB_s\theta_f^{\vec{Y}}(t)\theta_g^B(r).$$

Note that ψ is independent of $\mathcal{F}_r$, we can apply integration by part regarding ψ as nonrandom. Thus,

$$\int_0^r M_s\nabla^*\psi_s\tilde{c}(X_s)dB_s\theta_g^B(r)\int_0^r\cdots dB_s + \sqrt{-1}\int_0^r M_s\psi_s'\sigma(X_s)g_s\theta_g^B(s)ds.$$

This implies that

$$
\begin{aligned}
& \mathbb{E}\left(\int_0^t M_s\psi_s'\sigma(X_s)dB_s\theta_f^{\vec{Y}}(t)\theta_g^B(t)\right) \\
= \; & \mathbb{E}\left(\sqrt{-1}\int_0^t M_s\psi_s'\sigma(X_s)g_s\theta_g^B(s)ds\theta_f^{\vec{Y}}(t)\right) \\
= \; & \mathbb{E}\bigg(\sqrt{-1}\int_0^t M_s\mathbb{E}\left(\psi_s'(X_s)\tilde{\theta}_f(s)|\mathcal{F}_s^{\vec{Y}}\vee\mathcal{F}_s^B\right)\sigma(X_s)g_s\theta_g^B(s)\theta_f^{\vec{Y}}(s)ds\bigg) \\
= \; & \mathbb{E}\left(\sqrt{-1}\int_0^t M_s\Theta_s'(X_s)\sigma(X_s)g_s\theta_g^B(s)\theta_f^{\vec{Y}}(s)ds\right).
\end{aligned}
$$

This finishes the proof of the lemma. ■

Proof: (of Lemma 4.3) Note that

$$
\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right)=\mathbb{E}\mathbb{E}\left(\left(m_j^n(\eta_{j\delta}^n)^2\right)\Big|\mathcal{F}_{j\delta-}\right)=\mathbb{E}\left(m_{j-1}^n(\eta_{j\delta}^n)^2\right)
$$

$$
=\mathbb{E}\left(m_{j-1}^n(\eta_{(j-1)\delta}^n)^2\mathbb{E}\left(\left(\eta_{j\delta}^n/\eta_{(j-1)\delta}^n\right)^2|\mathcal{F}_{(j-1)\delta}\right)\right)\le e^{K^2\delta}\mathbb{E}\left(m_{j-1}^n(\eta_{(j-1)\delta}^n)^2\right)
$$

where the last inequality follows from

$$
\mathbb{E}\left(\left(\frac{1}{m_{j-1}^n}\sum_{k=1}^{m_{j-1}^n}M_j^n(X^k)\right)^2|\mathcal{F}_{(j-1)\delta}\right)\le\frac{1}{m_{j-1}^n}\sum_{k=1}^{m_{j-1}^n}\mathbb{E}\left(M_j^n(X^k)^2|\mathcal{F}_{(j-1)\delta}\right)\le e^{K^2\delta}.
$$

By induction, we have $\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right)\le e^{K^2T}n\le K_1n$. ■

Proof: (of Lemma 4.4) By Equation (3), we have

$$
dM_j^n(X^i,s)=M_j^n(X^i,s-)\sum_{k=1}^w(ap_k(X_{s-}^i,s-)-1)d\tilde{Y}_k(s).
$$

Observe that

$$
d\left(\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n}M_j^n(X^\ell,s)\right)=\left(\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n}M_j^n(X^\ell,s-)\right)\sum_{k=1}^w\bar{h}_k(s-)d\tilde{Y}_k(s)
$$

This gives (12). Applying Itô's formula to the last two equations and simplifying, we obtain

$$
d\tilde{M}_j^n(X^i,s)=-\tilde{M}_j^n(X^i,s)\sum_{k=1}^w(ap_k(X_s^i,s)-1-\bar{h}_k(s))ds+\Delta\tilde{M}_j^n(X^i,s)
$$

Note that $\sum_{k=1}^w(ap_k(X_s^i,s)-1-\bar{h}_k(s))=0$. To make the last term predictable, we observe

$$
\Delta\tilde{M}_j^n(X^i,s)=\tilde{M}_j^n(X^i,s)-\tilde{M}_j^n(X^i,s-)=\sum_{k=1}^w\tilde{M}_j^n(X^i,s-)\big(\frac{ap_k(X_{s-}^i,s-)}{\bar{h}_k(s-)+1}-1\big)dY_k(s).
$$

The conclusion then follows by substituting the above observation into the equation of $\tilde{M}^n_j(X^i,s)$ and by taking integral from $j\delta$ to $(j+1)\delta$. ∎

Proof: (of Lemma 4.5) Note that $\eta^n_{(j+1)\delta}/\eta^n_{j\delta} = \hat{M}^n_j((j+1)\delta)$ and $\hat{M}^n_j(t)$ follows (12). Then,

$$d\hat{M}^n_j(t)^2 = -2\hat{M}^n_j(t)^2(a-w)dt + \hat{M}^n_j(t-)^2\sum_{k=1}^{w}(\bar{h}^2_k(t-)+2\bar{h}_k(t-))dY_k(t).$$

Easy to find that $df^2(X^i_t) = Lf^2(X^i_t)dt + 2ff'\sigma(X^i_t)dB_t$ and by Itô formula, we obtain

$$d\big(\hat{M}^n_j(t)^2 f^2(X^i_t)\big) = \hat{M}^n_j(t)^2(Lf^2 - 2f^2(a-w))dt$$

$$+\hat{M}^n_j(t-)^2 f^2\sum_{k=1}^{w}(\bar{h}^2_k(t-)+2\bar{h}_k(t-))dY_k(t) + 2\hat{M}^n_j(t)^2 ff'\sigma(X^i_t)dB_t.$$

Equation (8) gives $\gamma^n_{j+1}(X^i) = F(\tilde{M}^n_{j+1}(X^i))$. By telescoping and using (13), we obtain

$$\begin{aligned}\gamma^n_{j+1}(X^i) &= \sum_{j\delta<s\le(j+1)\delta}\big[F(\tilde{M}^n_j(X^i,s)) - F(\tilde{M}^n_j(X^i,s-))\big]\\ &= \sum_{k=1}^{w}\int_{j\delta}^{(j+1)\delta}\Big[F\big(\tilde{M}^n_j(X^i,s-)\frac{ap_k(X^i_{s-},s-)}{\bar{h}_k(s-)+1}\big) - F(\tilde{M}^n_j(X^i,s-))\Big]dY_k(s)\end{aligned}$$

Applying Itô's formula again, we have

$$\begin{aligned}&\gamma^n_{j+1}(X^i)f^2(X^i_{(j+1)\delta})(\eta^n_{(j+1)\delta}/\eta^n_{j\delta})^2\\ =&\int_{j\delta}^{(j+1)\delta} F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2\big(Lf^2 + f^2\sum_{k=1}^{2}\bar{h}^2_k(t)\big)dt\\ &+\int_{j\delta}^{(j+1)\delta}\hat{M}^n_j(t)^2 f^2\sum_{k=1}^{2}\Big[F\big(\tilde{M}^n_j(X^i,t)\frac{ap_k(X^i_t,t)}{\bar{h}_k(t)+1}\big) - F(\tilde{M}^n_j(X^i,t))\Big](\bar{h}_k(t)+1)^2dt\\ &+\int_{j\delta}^{(j+1)\delta} F(\tilde{M}^n_j(X^i,t-))\hat{M}^n_j(t-)^2 f^2\sum_{k=1}^{w}\big(\bar{h}^2_k(t-)+2\bar{h}_k(t-)\big)d\tilde{Y}_k(t)\\ &+\int_{j\delta}^{(j+1)\delta} F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2 ff'\sigma(X^i_t)dB_t\\ &+\int_{j\delta}^{(j+1)\delta}\hat{M}^n_j(t-)^2 f^2\sum_{k=1}^{w}\Big[F\big(\tilde{M}^n_j(X^i,t-)\frac{ap_k(X^i_{t-},t-)}{\bar{h}_k(t-)+1}\big) - F(\tilde{M}^n_j(X^i,t-))\Big](\bar{h}_k(t-)+1)^2d\tilde{Y}_k(t)\end{aligned}$$

Taking conditional expectation and noting that the last three terms are zero, we have

$$\begin{aligned}&\mathbb{E}\Big(\gamma^n_{j+1}(X^i)f^2(X^i_{(j+1)\delta})(\eta^n_{(j+1)\delta}/\eta^n_{j\delta})^2|\mathcal{F}_{j\delta}\Big)\\ =&\int_{j\delta}^{(j+1)\delta}\mathbb{E}\Big[F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2\big(Lf^2 + f^2\sum_{k=1}^{2}\bar{h}^2_k(t)\big)\Big|\mathcal{F}_{j\delta}\Big]dt\end{aligned}$$

$$+\int_{j\delta}^{(j+1)\delta}\mathbb{E}\Big[\hat{M}_j^n(t)^2 f^2\sum_{k=1}^{w}\Big[F\big(\tilde{M}_j^n(X^i,t)\frac{ap_k(X_t^i,t)}{\bar{h}_k(t)+1}\big)-F(\tilde{M}_j^n(X^i,t))\Big](\bar{h}_k(t)+1)^2\Big|\mathcal{F}_{j\delta}\Big]dt$$

$$\approx\left(\hat{M}_j^n(j\delta)^2 f^2(X_{j\delta}^i)\sum_{k=1}^{w}F\big(\frac{ap_k(X_{j\delta}^i,j\delta)}{\bar{h}_k(j\delta)+1}\big)(\bar{h}_k(j\delta)+1)^2\right)\delta+o(\delta)$$

The last approximation comes from $\tilde{M}_j^n(X^i,j\delta)=1$, the boundedness of f, Lf^2, $\sum_{k=1}^w\bar{h}_k^2$, and the following two observations:

$$\sup_{j\delta\le s\le(j+1)\delta}\mathbb{E}\Big[\hat{M}_j^n(s)^2F\big(\tilde{M}_j^n(X^i,s)\big)\Big|\mathcal{F}_{j\delta}\Big]$$

$$\le\sup_{j\delta\le s\le(j+1)\delta}\sqrt{\mathbb{E}\big(\hat{M}_j^n(s)^4|\mathcal{F}_{j\delta}\big)}\sqrt{\mathbb{E}\Big[\big(1-\tilde{M}_j^n(X^i,s)\big)^2|\mathcal{F}_{j\delta}\Big]}\le K\sqrt{\delta}.$$

and (the last inequality above is by $\mathbb{E}\big[(1-\tilde{M}_j^n(X^i,s))^2|\mathcal{F}_{j\delta}\big]\le K_1\delta$ and $\mathbb{E}\big(\hat{M}_j^n(s)^4|\mathcal{F}_{j\delta}\big)\le K_2$)

$$\sup_{j\delta\le s\le(j+1)\delta}\Big|\mathbb{E}\Big[\hat{M}_j^n(s)^2F\big(\tilde{M}_j^n(X^i,s-)\frac{ap_k(X_{s-}^i,s-)}{\bar{h}_k(s-)+1}\big)|\mathcal{F}_{j\delta}\Big]-\hat{M}_j^n(j\delta)^2F\big(\frac{ap_k(X_{j\delta}^i,j\delta)}{\bar{h}_k(j\delta)+1}\big)\Big|\to 0.$$

The last observation can be proven similarly as the previous one. ■

Proof: (of Lemma 6.1) Note that

$$d\hat{M}_j^n(t)^4=-4\hat{M}_j^n(t)^4(a-w)dt+\hat{M}_j^n(t-)^4\sum_{k=1}^{w}\big[(\bar{h}_k^2(t-)+1)^4-1\big]dY_k(t),$$

and by telescoping, we obtain

$$\begin{aligned}\gamma_{j+1}^n(X^i)^2&=\sum_{j\delta<s\le(j+1)\delta}\big[F^2(\tilde{M}_j^n(X^i,s))-F^2(\tilde{M}_j^n(X^i,s-))\big]\\&=\sum_{k=1}^{w}\int_{j\delta}^{(j+1)\delta}\Big[F^2\big(\tilde{M}_j^n(X^i,s-)\frac{ap_k(X_{s-}^i,s-)}{\bar{h}_k(s-)+1}\big)-F^2(\tilde{M}_j^n(X^i,s-))\Big]dY_k(s)\end{aligned}$$

Applying Itô's formula, we have

$$\begin{aligned}&\gamma_{j+1}^n(X^i)^2(\eta_{(j+1)\delta}^n/\eta_{j\delta}^n)^4\\=&-8\int_{j\delta}^{(j+1)\delta}F(\tilde{M}_j^n(X^i,t))\hat{M}_j^n(t)^4(a-w)dt\\&+2\int_{j\delta}^{(j+1)\delta}F(\tilde{M}_j^n(X^i,t))\hat{M}_j^n(t)^4\sum_{k=1}^{w}\big[(\bar{h}_k^2(t-)+1)^4-1\big]dY_k(t)\\&+4\int_{j\delta}^{(j+1)\delta}\hat{M}_j^n(t)^3\sum_{k=1}^{w}\Big[F^2\big(\tilde{M}_j^n(X^i,t-)\frac{ap_k(X_{t-}^i,t-)}{\bar{h}_k(t-)+1}\big)-F^2(\tilde{M}_j^n(X^i,t-))\Big]dY_k(t)\\&+\int_{j\delta}^{(j+1)\delta}\hat{M}_j^n(t)^4\sum_{k=1}^{w}\Big[F^2\big(\tilde{M}_j^n(X^i,t-)\frac{ap_k(X_{t-}^i,t-)}{\bar{h}_k(t-)+1}\big)-F^2(\tilde{M}_j^n(X^i,t-))\Big]\\&\qquad\big[(\bar{h}_k^2(t-)+1)^4-1\big]dY_k(t)\end{aligned}$$

For the first term, we have

$$\mathbb{E}\Big(\int_{j\delta}^{(j+1)\delta} F(\tilde{M}_j^n(X^i,t))\hat{M}_j^n(t)^4(a-w)dt\Big|\mathcal{F}_{j\delta}\Big)$$

$$\leq \int_{j\delta}^{(j+1)\delta} \mathbb{E}\Big(\big|\tilde{M}_j^n(X^i,t)-1\big|\hat{M}_j^n(t)^4\Big|\mathcal{F}_{j\delta}\Big)dt$$

$$\leq \int_{j\delta}^{(j+1)\delta} \sqrt{\mathbb{E}\big((\tilde{M}_j^n(X^i,t)-1)^2|\mathcal{F}_{j\delta}\big)}\sqrt{\mathbb{E}(\hat{M}_j^n(t)^4|\mathcal{F}_{j\delta})}dt \leq K\delta^{3/2}.$$

Other terms can be estimated similarly with the same order of $\delta^{3/2}$. ■

Proof: (of Lemma 6.2) We can estimate $\mathbb{E}\big((m_j^n)^2(\eta_{j\delta}^n)^4\big)$ recursively as follows

$$\mathbb{E}\big((m_j^n)^2(\eta_{j\delta}^n)^4\big) = \mathbb{E}\Big(\mathbb{E}\Big((m_j^n)^2(\eta_{j\delta}^n)^4\big|\mathcal{F}_{j\delta-}\Big)\Big) = \mathbb{E}\Big((\eta_{j\delta}^n)^4\mathbb{E}\Big(\big(\sum_{i=1}^{m_{j-1}^n}\xi_i^j\big)^2\Big|\mathcal{F}_{j\delta-}\Big)\Big)$$

$$\leq \mathbb{E}\Big((\eta_{j\delta}^n)^4 m_{j-1}^n \sum_{i=1}^{m_{j-1}^n}\mathbb{E}\big((\xi_i^j)^2\big|\mathcal{F}_{j\delta-}\big)\Big) \leq \mathbb{E}\big((\eta_{j\delta}^n)^4(m_{j-1}^n)^2(1+K\delta)\big)$$

$$\leq \mathbb{E}\Big((\eta_{(j-1)\delta}^n)^4(m_{j-1}^n)^2(1+K\delta)\mathbb{E}\big(\hat{M}_j^n(j\delta)^4\big|\mathcal{F}_{(j-1)\delta}\big)\Big) \leq (1+K\delta)e^{K_1\delta}\mathbb{E}\big((m_{j-1}^n)^2(\eta_{(j-1)\delta}^n)^4\big)$$

Thus, by induction, we have

$$\mathbb{E}\big((m_j^n)^2(\eta_{j\delta}^n)^4\big) \leq (1+K\delta)^j e^{K_1 j\delta}n^2 \leq K_3 n^2.$$

■

References

[1] N. Chopin (2004), Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.* 32, no. 6, 2385–2411.

[2] 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.

[3] A. Bensoussan (1992), *Stochastic control of partially observable systems*, Cambridge University Press.

[4] Christie, W., Harris, J. H. and Schultz, P. H., (1994). Why did NASDAQ market makers stop avoiding odd-eighth quotes?, *Journal of Finance*, **49**, 1841–1860.

[5] Christie, W. and Schultz, P. H., (1994). Why do NASDAQ market makers avoid odd-eighth quotes?, *Journal of Finance*, **49**, 1813–1840.

[6] D. Crisan (2003). Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation Ann. Probab. **31**, 693–718.

[7] D. Crisan (2006). Particle approximations for a class of stochastic partial differential equations. to appear in *The Applied Mathematics and Optimization Journal*.

[8] D. Crisan, J. Gaines, T. Lyons (1998). Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.* **58**, 1568–1590 (electronic).

[9] D. Crisan, T. Lyons (1999). A particle approximation of the solution of the Kushner-Stratonovitch equation. *Probab. Theory Related Fields* **115**, 549–578.

[10] D. Crisan, P. Del Moral and T. Lyons (1999). Interacting particle systems approximations of the Kushner-Stratonovitch equation. *Adv. in Appl. Probab.* **31**, 819–838.

[11] D. Crisan, A. Doucet (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50**, pp 736–746.

[12] D. Crisan and J. Xiong (2006), A central limit type theorem for particle filter. Working Paper, University of Tennessee.

[13] P. Del Moral (2004) *Feynman-Kac formulae. Genealogical and interacting particle systems with applications.* Probability and its Applications, Springer-Verlag, New York.

[14] P. Del Moral, A. Guionnet (1998), Large Deviations for Interacting Particle Systems. Applications to Non Linear Filtering Problems, *Stochastic Processes and their Applications*, **78**, 69-95.

[15] P. Del Moral, A. Guionnet (1999), Central limit theorem for nonlinear filtering and interacting particle systems, *Ann. Appl. Probab.* 9, no. 2, 275–297.

[16] P. Del Moral, L. Miclo (2000), *Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering.* Séminaire de Probabilités, XXXIV, 1–145, Lecture Notes in Math., 1729, Springer, Berlin.

[17] R. Engle (2000) The Econometrics of Ultra-High-Frequency Data. *Econometrica*. **68**, 1-22.

[18] R. Engle and J. Russell (1998) Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*. **66**, 1127-1162.

[19] S.N. Ethier and T.G. Kurtz (1986). *Markov processes : Characterization and convergence.* Wiley, New York.

[20] I. M. Gel'fand and N. Ya. Vilenkin (1964). *Generalized functions. Vol. 4: Applications of harmonic analysis.* Academic Press, New York - London.

[21] J. Hasbrouck (1996). Modeling Market Microstructure Time Series, in *Handbook of Statistics* editted by G.S. Maddala and C.R. Rao, North-Holland, 647-692.

[22] J. Hasbrouck (2002). Stalking the "efficient price" in market microstructure specifications: an overview. *Journal of Finanical Markets*, **5**, 329-339.

[23] M. Hitsuda and I. Mitoma (1986). Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. *J. Multivariate Anal.* 19, 311–328.

[24] M. Kouritzin and Y. 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.

[25] H. R. Kunsch (2005), Recursive Monte Carlo filters: algorithms and theoretical analysis, *Ann. Statist.* 33, no. 5, 1983–2021.

[26] T. Kurtz and J. Xiong (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications,* **83**, 103-126.

[27] T. Kurtz and J. Xiong (2004). A stochastic evolution equation arising from the fluctuation of a class of interacting particle systems. *Communication Mathematical Sciences* **2**, 325-358.

[28] K. Lee and Y. Zeng (2006). Risk minimization for a filtering micromovement model of asset price. Working paper. University of Missouri at Kansas City.

[29] Spalding, R., Tsui, K. W. and Zeng, Y., (2006). A micro-movement model with Bayes estimation via filtering: Applications to measuring trading noises and trading cost, *Nonlinear Analysis*, **64**, 295-309.

[30] J. Xiong and Y. Zeng (2006) Mean-variance portfolio selection for A Filtering Point Process Model of Asset Price. Working paper. University of Tennessee.

[31] J. Yong and X. Zhou (1999). *Stochastic control*, Springer, New York.

[32] Y. Zeng (2003). A partialy observed model for micromovement of asset prices with Bayes estimation via filtering, *Mathematical Finance*, **13** 411-444.

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

[34] Y. Zeng (2005). Bayesian inference via filtering for a class of counting processes: Application to the micromovement of asset price, *Statistical Inference for Stochastic Processes*, **8**, 331 - 354.

[35] Y. Zeng (2006). Statistical analysis of the filtering models with markted point process observations: Application to ultra-high frequency data. Working paper. University of Missouri at Kansas City.

[36] L. Zhang and P. A. Mykland and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high frequency data. *JASA.* **100**, 1394-1411.