A Branching Particle Approximation to a Filtering Micromovement Model of Asset Price *

Jie Xiong † Yong Zeng ‡

First Draft: November 28, 2006. This Version: May 24, 2010

Abstract

Recently, a filtering model with counting process observations has been demonstrated as a sensible framework for modeling the micromovement of asset price (or financial ultra-high frequency data). In this paper, we study the simulation-based branching particle approximation for such a nonlinear filtering model. We first construct a branching particle system. Then, we show the weighted (unnormalized and normalized) empirical measures in the constructed branching system converges to the optimal (unnormalized and normalized) filters uniformly in time. This is achieved 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: Particle filters, Monte Carlo approximation, filtering, counting process, stochastic partial differential equation, 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. Yong is grateful to Tom Kurtz for helpful discussion, and to Rene Carmona for his passionate introduction of particle filtering in the early stage of this work. We are grateful to an anonymous referee for the constructive comments, improving the quality of the manuscript. Xiong's research is supported in part by NSF Grant DMS-0906907 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 behaviors. Two important relatively early works, [18] and [17], attempt to model the micromovement from irregularly-spaced time series viewpoint. Engle in [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 to 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 models due to the high frequency transaction nature.

From the standpoint of stochastic process, a general Filtering Micromovement model for asset price (FM model, as we simply call it) is proposed in [35]. In the FM model, there is an unobservable intrinsic value process for an asset, which corresponds to the macro-movement in the empirical econometric literature or the continuous-time price process in the option pricing literature. 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. Therefore, the FM model is capable of matching the two stylized features of micromovements as well as many those of macromovement.

The FM model has the structure similar to two classes of models. One class is the time series structural models developed in many early market microstructure papers (see [23], a survey paper on this topic, and a recent one [24]). Namely, price can be decomposed as a permanent component with a long-term impact on price and a transient component with only a short-term impact. In the FM model, the intrinsic value process is the permanent component and trading noise is the transient component. The other class is the recent two-time-scale frameworks incorporating market microstructure noises in the fast growing literature of realized volatility estimators. See [37], [1], [2], and [20]. Especially, Li and Mykland in [32] shows that rounding noise, which is accommodated in the FM model, may severely distort even the two-scale estimators of realized volatility, and the error could be infinite.

The most prominent feature of the 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 likelihoods and the conditional distribution of the intrinsic value process (the signal). The Markov chain approximation method has been developed and utilized to numerically solve the filtering equations in [35].

On the other hand, simulation-based particle filters have been studied extensively as alternative approximations to the optimal filters in the classical nonlinear filtering in the last ten years. Branching and interacting are two main classes of particle filters. To present the motivation of these particle filters, recall that in the classical Monte Carlo method, the unnormalized filter is approximated by a weighted particle system, but the variances of weights grow exponentially fast. These two particle filters are designed to reduce variances but with different updating schemes. 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. The interacting particle filters employ resampling for updating and interested readers are referred to the comprehensive monograph [13] by Del Moral and related references therein. For the 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] - [10], especially, [11].

In this paper, we study the branching particle approximation to the FM model through sophisticated calculation and accurate moment estimation. Suppose that V_t is the unnormalized conditional measure and π_t is the conditional distribution in the FM model. V_t and π_t are characterized by the unnormalized and normalized filtering equations, respectively. First, we construct a branching particle system to approximate the FM model. Then, we define the weighted empirical measures π_t^n and V_t^n of the constructed branching particle system. The first aim of this paper is to prove the uniform convergence (in time) of V_t^n to V_t as well as π_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 5 and 7. Moreover, we characterize the convergence rate of V_t^n and π_t^n by a central limit type theorem (CLT) on the modified Schwarz space. 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 [11]). This is because the key moment estimates in Lemmas 5 and 7 are sharper than those in the classical nonlinear filtering case (see [6] and [11]).

The unweighted empirical measures in the branching particle system can be defined also. Historically, the unweighted empirical measures were first studied and were proven convergent to the optimal filters. However, as indicated in [6] and recently shown in [11], the weighted empirical measures are superior to the unweighted ones 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 this paper. Similar CLT results shown for some 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 [3] and [28].

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 [31] and the optimal trading strategy for mean-variance portfolio selection problem of the FM models derived in [33].

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.

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

2 The Model and Filtering Equations

This section presents the FM model whose filters are approximated by a branching particle system in this paper. Then, we summarize the related filtering equations.

2.1 The Filtering Micromovement Model

In the model, the signal is the latent intrinsic value process of an asset, X_t, with a mild assumption below.

Assumption 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. Let ϕ be the initial distribution of X_0. We further assume the following bounded conditions: ϕ is a bounded function. $\mu(x)$ and $\sigma(x)$ are continuous with bounded first derivatives. Moreover, $\sup_{0\leq t\leq T} E(\sigma^4(X_t))$ is bounded.

The generator associated with X is

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

Assumption 1 is more restricted than the general assumption of a Markov process used in [35]. However, it can be easily checked that Assumption 1 with the related bounded conditions includes geometric Brownian motion (GBM), the Black-Scholes model.

During trading, the intrinsic value process can not be observed directly, but can be partially observed through the trade-by-trade 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. Therefore, we can employ the point process framework as described in the book [5]. Namely, 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, and w is the number of price level which can be chosen according to the range of transaction prices.

The following four mild assumptions are invoked.

Assumption 2 N_k's are unit Poisson processes under the physical measure P.

Assumption 3 $X, N_1, N_2, \ldots, N_w$ are independent under P.

Assumption 4 The total trading intensity at time t, $a(x,t) = \sum_{k=1}^{w} \lambda_k(x,t)$, is uniformly bounded above; i.e., there exist a constant, K, such that $a(x,t) \leq K$ for all $t > 0$ and x.

Assumption 5 The intensity at price level k, $\lambda_k(x,t) = a(x,t)p(y_k|x;t) > 0$, where $p(y_k|x;t)$ is the time-dependent transition probability from x to y_k, the kth price level. Let $p_k(x) = p(y_k|x;t)$ and $a'(x,t) = \frac{d}{dx}a'(x,t)$. Furthermore, $a'(x,t)$ is continuous in x and uniformly bounded for t, and $p'_k(x)$ is continuous and bounded.

The structure of λ_k implies that $a(X(t),t)$ specifies when the trade might occur while $p(y_k|x;t)$ 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 and denote $ap_k(X_t,t)$ by ap_k at times.

Remark 1 Under this representation, $X(t)$ becomes the signal process, which cannot be observed directly, and $\vec{Y}(t)$ becomes the observation process, corrupted by noise which is modeled by $p(y|x;t)$. Hence, $(X,\vec{Y})$ is framed as a *filtering problem with counting process observations.*

From the standpoint of modeling stochastic volatility, other filtering models for the micromovement are proposed by Frey and Runggaldier in [22] and by Cvitanic, Liptser, and Rozovskii in [12]. However, noise is not incorporated in their models.

Another more heuristic way of modeling is to construct the transaction price Y from the intrinsic value X as below. First, we specify $X(t)$ as in Assumption 1. Then, we specify the trading times $t_1, t_2, \ldots, t_i, \ldots$, which are driven by a conditional Poisson process with a conditional intensity function $a(X(t),t)$. Finally, $Y(t_i)$, the trading price at time t_i, is obtained by a random transformation from the value at that time: $Y(t_i) = F(X(t_i))$, where the random transformation $y = F(x)$ is specified by the transition probability $p(y|x;t)$.

Note that the random transformation models the trading noise as the transition probability does. Examples of $F(x)$(or $p(y|x;t)$) are given in [35] and [36]. These examples well accommodate the three types of well-documented noise in financial literature: discrete noise, clustering noise, and non-clustering noise.

The above heuristic construction has important financial implication: Price is influenced by information and noise. Information affects the intrinsic value of an asset, $X(t)$, and has a permanent influence on the price. Noise, specified by the random transformation $F(x)$, does not affects the intrinsic value and has only a transitory influence on price. The formulation is similar to the time series structural models used in many market microstructure papers (see [23] and [24]). Furthermore, the formulation is closely related the recent two-time-scale frameworks incorporating market microstructure noises in literature of realized volatility estimators. See [37], [1], [2], [20], and especially, [32].

The two approaches of modeling are equivalent in the sense that both representations have the same probability distribution, which is proven in [36]. The structure of λ_k is the key to guarantee the equivalence.

2.2 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 the filtration generated by the pair $(X,\vec{Y})$ and $\hat{\mathcal{G}} = \hat{\mathcal{F}}_\infty$. Assumptions 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 ([26]) 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 $M(t) = E^Q[M(T)|\hat{\mathcal{F}}_t]$. Then, $M(t)$ satisfies the following SDE:

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

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

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

By Kallianpur-Striebel formula ([27]), 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 [35] summarizing both filtering equations.

Proposition 1 *Suppose that $(X,\vec{Y})$ satisfies Assumptions 1 - 5. Then, V_t is the unique measure-valued solution of the following SPDE under Q, 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 $p_k = p(y_k|x;t)$ is the transition probability from x to y_k, the kth price level.

π_t is the unique measure-valued solution of the SPDE under P, 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 normalized and unnormalized weighted empirical measures to approximates the optimal filters.

Recall that ϕ is the initial distribution of X_0. After choosing the initial number of particle n and assigning weight $\frac{1}{n}$ for each particle, we initialize the starting positions of the n particles, each at position x_0^i, $i = 1, 2, \cdots, n$, satisfying the following initial condition.

Assumption 6 *As $n \to \infty$,*

$$V_0^n = \frac{1}{n}\sum_{i=1}^{n}\delta_{x_0^i} \to \phi \qquad in\ \mathcal{M}_F(\mathbb{R}),$$

where $\mathcal{M}_F(\mathbb{R})$, the collection of finite measures on $\mathbb{R}$.

For the distribution V_0^n, the convergence is the same as the weak convergence or the convergence in distribution. However, for the to-be-defined unnormalized empirical conditional measure V_t^n, since it does not sum up to one and it is usually a finite measure only, the convergence is in finite measure.

Let δ as the length between two time steps. After initialization, there are three recursive steps for each time step in the branching particle system. Suppose that at time $t = j\delta$, there are m_j^n particles alive. Set $m_0^n = n$.

First, we simulate the path of each particle independently for the time interval $[j\delta,(j+1)\delta)$ according to the following diffusions satisfying Assumption 1: For $i = 1, 2, \cdots, m_j^n$ and $t \in [j\delta,(j+1)\delta)$,

$$X_t^i = X_{j\delta}^i + \int_{j\delta}^{t}\mu(X_s^i)ds + \int_{j\delta}^{t}\sigma(X_s^i)dB_s^i. \tag{7}$$

where $\{B^i, i = 1, 2, \cdots, n\}$ are independent standard Brownian motions and $X_0^i = x_0^i$ when $j = 0$.

Then, we want to assign a weight to each particle at time $t = (j+1)\delta-$. We first define for particle $i = 1, 2, ..., m_j^n$, its conditional likelihood of measure of $\vec{Y}|_{[j\delta,(j+1)\delta)}$ given that the trajectory of $X|_{[j\delta,(j+1)\delta)}$ equals X^i with the initial conditional likelihood set to be one at time $j\delta$ as

$$M_j^n(X^i,t) = \prod_{k=1}^{w}\exp\left(\int_{j\delta+}^{t}\log ap_k(X_{s-}^i, s-)dY_k(s) - \int_{j\delta}^{t}\left[ap_k(X_s^i,s)-1\right]ds\right). \tag{8}$$

Obviously, the first integral in the exponent is zero unless a trade happens during $(j\delta, t)$. Recall $M_j^n(X^i, j\delta) = 1$ at the beginning and at the time right before branching, let

$$M_{j+1}^n(X^i) = M_j^n(X^i,(j+1)\delta-). \tag{9}$$

In order to keep likely particles, we define the weight proportional to the conditional likelihood for particle i at time $t \in [j\delta,(j+1)\delta)$ as

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

Then, the total weights of all particles remains m_j^n for $t \in [j\delta,(j+1)\delta)$. For particle i, the weight right before branching (at time $t = (j+1)\delta-$), which depends on *all* $X^1, ..., X^{m_j^n}$, is denoted by

$$\tilde{M}_{j+1}^n(X^i) = \frac{M_{j+1}^n(X^i)}{\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n}M_{j+1}^n(X^\ell)} = m_j^n\left(\frac{M_{j+1}^n(X^i)}{\sum_{\ell=1}^{m_j^n}M_{j+1}^n(X^\ell)}\right). \tag{10}$$

Finally, given the particles at the end of the interval (at time $t = (j+1)\delta$), conditionally independent of all other particles, the ith particle ($i = 1, 2, \cdots, m_j^n$) branches (namely, dies and gives birth) to a random number ξ^i_{j+1} of offsprings, whose conditional expectation is set to be the pre-branching weight, $\tilde{M}^n_{j+1}(X^i)$. Precisely, we let

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

Moreover, let

$$Var^Q\left(\xi^i_{j+1}|\mathcal{F}_{(j+1)\delta-}\right) = \gamma^n_{j+1}(X^i).$$

Following [8], in order to minimize the variance γ^n_{j+1}, we restrict the possible number of ξ^i_{j+1} to the two integers closest to $\tilde{M}^n_{j+1}(X^i)$ and set

$$\xi^i_{j+1} = \begin{cases} [\tilde{M}^n_{j+1}(X^i)] & \text{with probability } 1-\{\tilde{M}^n_{j+1}(X^i)\} \\ [\tilde{M}^n_{j+1}(X^i)]+1 & \text{with probability } \{\tilde{M}^n_{j+1}(X^i)\} \end{cases} \tag{11}$$

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

$$\gamma^n_{j+1}(X^i) = \{\tilde{M}^n_{j+1}(X^i)\}(1-\{\tilde{M}^n_{j+1}(X^i)\}). \tag{12}$$

Variance reduction is achieved through such branching. After branching, at $t = (j+1)\delta$, there are $m^n_{j+1} = \sum_{i=1}^{m^n} \xi^i_j$ particles. Each particle begins from the ending position of its "father" particle with weight $1/m^n_{j+1}$. Then, it goes back to simulate the independent path of each particle. The recursive structure is summarized in the following pseudo code for the branching particle filter:

Step 0, Initializing: `Set` $t \rightarrowtail 0$. `Initialize` n `particles satisfying Assumption 6.`

Step 1, Simulating: `For` $t \in (j\delta, (j+1)\delta)$, `simulate the path of each particle independently according to (7).`

Step 2, Weighting: `At` $t = (j+1)\delta-$, `assign a weight to each particle. This consists two substeps.`

Step 2a: `Compute the conditional likelihood of` $\vec{Y}$ `as (8) with` $t = (j+1)\delta-$ `for each particle.`

Step 2b: `Compute the weight of each particle by (10).`

Step 3, Branching: `At` $t = (j+1)\delta$, `branch each particle conditionally independent of other particles according to its weight as described by (11). New particle evolves from the ending position of its father particle and go back to Step 1. Otherwise stop.`

We summarize the related notations in Table 1 for future reference.

Now, we proceed to define the approximate filters π^n_t and V^n_t.

Definition 1 *Let the empirical conditional distribution of X_t given $\mathcal{F}^{\vec{Y}}_t$ be*

$$\pi^n_t = \frac{1}{m^n_j} \sum_{i=1}^{m^n_j} \tilde{M}^n_j(X^i, t)\delta_{X^i(t)}, \qquad \text{if } j\delta \le t < (j+1)\delta.$$

Table 1: Notations for the Branching Particle System

	$t = j\delta$	$t \in (j\delta, (j+1)\delta)$	$t = (j+1)\delta-$	$t = (j+1)\delta$
# of particle	m_j^n	m_j^n	m_j^n	$\sum_{i=1}^{m_j^n} \xi_j^i = m_{j+1}^n$
position of ith particle	$X_{j\delta}^i$	X_t^i	$X_{(j+1)\delta-}^i$	
weight of ith particle	$\frac{1}{m_j^n}$	$\tilde{M}_j^n(X^i, t)$	$\tilde{M}_{j+1}^n(X^i)$	$\frac{1}{m_{j+1}^n}$

Note: When $t = 0$, $m_0^n = n$, $X_0^i = x_i^n$, and $1/m_0^n = \frac{1}{n}$. At time $(j+1)\delta$, the ith particle at $X_{(j+1)\delta-}^i$ may die or give birth to ξ_j^i particles starting at $X_{(j+1)\delta-}^i$ and the particles are renumbered.

Since the unnormalized filtering equation of V_t is simpler than that of π_t, we first study the convergence of V_t^n to V_t then convert the results to that of π_t^n to π_t. So, we proceed to define V_t^n. From Kallianpur-Striebel formula, $V_t = \pi_t \langle V_t, 1\rangle$ where $\langle V_t, 1\rangle$ is the likelihood ratio of P over Q. We first define the approximate likelihood ratio up to time $j\delta$ as

$$\eta_{j\delta}^n = \Pi_{k=0}^{j-1} \frac{1}{m_k^n} \sum_{\ell=1}^{m_k^n} M_{k+1}^n(X^\ell), .$$

Note that $\eta_{j\delta}$ is $\mathcal{F}_{j\delta-}$-measurable. Since the constructed m_j^n is a martingale with mean n, the likelihood ratio from $j\delta$ to t can be approximated by $\frac{1}{n}\sum_{i=1}^{m_j^n} M_j^n(X^i, t)$. Hence, we have the following definition.

Definition 2 *Let the unnormalized empirical conditional measure of X_t given $\mathcal{F}_t^Y$ be,*

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

Observe that

$$\langle \pi_t^n, f\rangle = \frac{1}{m_j^n} \sum_{i=1}^{m_j^n} \tilde{M}_j^n(X^i, t) f\big(X^i(t)\big), \qquad \langle V_t^n, f\rangle = \frac{1}{n} \eta_{j\delta}^n \sum_{i=1}^{m_j^n} M_j^n(X^i, t) \langle \pi_t^n, f\rangle,$$

. and $\langle \pi_t^n, f\rangle = \langle V_t^n, f\rangle / \langle V_t^n, 1\rangle$.

One main result of the paper is the convergence of V_t^n to V_t for fixed t.

Theorem 1 *Suppose that Assumptions 1 -5 hold for the FM model and Assumption 6 holds for the branching particle system constructed in Section 3. Then for a bounded initial distribution function ϕ of X_0, there exists a constant K_1 such that*

$$\mathbb{E}|\langle V_t^n, \phi\rangle - \langle V_t, \phi\rangle|^2 \le K_1 n^{-1} \qquad \text{as } \delta \to 0.$$

Moreover, we improve the above result to the uniform convergence in time and convert the result of V^n to that of π^n. Before we state the two more main results, we define the usual distance for two finite measures ν_1 and ν_2

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

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

Theorem 2 *Under the assumptions of Theorem 1, there exists a constant K_1 such that*

$$\mathbb{E}\sup_{t\le T}d(V_t^n,V_t)^2\le K_1n^{-1},\qquad as\ \ \delta\to 0.$$

Theorem 3 *Under the assumptions of Theorem 1, there exists a constant K such that*

$$E^P\sup_{0\le t\le T}d(\pi_t^n,\pi_t)\le Kn^{-\frac{1}{2}},\qquad as\ \ \delta\to 0. \tag{13}$$

In Section 4, we prove Theorem 1, namely, the convergence of V_t^n. In Section 5, we prove Theorems 2 and 3, namely, the uniform convergence of V^n and π_n. In Section 6, we characterize the exact convergence rates by proving central limit type theorems. In each of the following section, we first present the main ideas and the key lemmas or theorems with all their proofs in the subsequent subsections.

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 the classical nonlinear filtering models in [8] and [11].

We consider the backward SPDE:

$$\left\{\begin{array}{ll} d\psi_s=-L\psi_s ds-\sum_{k=1}^{w}(ap_k-1)\psi_{s+}\hat{d}(Y_s^k-s), & 0\le s\le t\\ \psi_t=\phi \end{array}\right. \tag{14}$$

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

$$\left\{\begin{array}{ll} d\hat{\psi}_s=L\hat{\psi}_s ds+\sum_{k=1}^{w}(ap_k-1)\hat{\psi}_{s-}d(\hat{Y}_s^k-s), & 0\le s\le t\\ \hat{\psi}_0=\phi \end{array}\right. \tag{15}$$

which is the Zakai-tpye equation in Proposition 1. Similar to [29], we can prove the uniqueness for the solution to (15), implying the uniqueness of (14). In fact, we need the following technical estimates in Lemma 1.

Lemma 1 *Suppose Assumptions 1 - 5 hold for the FM model. Let $\psi_u'(x) = \frac{d}{dx}\psi_u(x)$. Then, there exists a constant K such that*

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

Lemma 2 is a convolution result. Lemmas 3 and 5 are key moment estimates crucial for the convergence results and the central limit theorem type result. Lemma 4 gives the SDEs of two quantities needed in Lemmas 5 and 7 later. The order of a key moment estimate in Lemma 5 is $o(\delta)$, which is sharper than $o(\delta^{1/2})$, the order in the classical nonlinear filtering case (see [6] and [11]).

Lemma 2 *Almost surely, we have*

$$\psi_{(j+1)\delta}(X^i_{(j+1)\delta})M^n_{j+1}(X^i) - \psi_{j\delta}(X^i_{j\delta}) = \int_{j\delta}^{(j+1)\delta} M^n_j(X^i,s)\psi_s'(X^i_s)\sigma(X^i_s)dB^i_s. \tag{16}$$

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 3 $\mathbb{E}\left(m^n_j(\eta^n_{j\delta})^2\right) \le K_1 n.$

Lemma 4 *Let*

$$\hat{M}^n_j(t) = \frac{1}{m^n_j}\sum_{\ell=1}^{m^n_j} M^n_j(X^\ell,t), \quad \text{and} \quad \tilde{M}^n_j(X^i,t) = \frac{M^n_j(X^i,t)}{\frac{1}{m^n_j}\sum_{\ell=1}^{m^n_j} M^n_j(X^\ell,t)} = \frac{M^n_j(X^i,t)}{\hat{M}^n_j(t)}.$$

Then,

$$d\hat{M}^n_j(t) = \hat{M}^n_j(t-)\sum_{k=1}^{w}\bar{h}^n_k(t-)d\tilde{Y}_k(t) \tag{17}$$

and

$$\tilde{M}^n_j(X^i) = 1 + \int_{j\delta}^{(j+1)\delta}\tilde{M}^n_j(X^i,s-)\sum_{k=1}^{w}\Big[\frac{ap_k(X^i_{s-},s-)}{\bar{h}^n_k(s-)+1} - 1\Big]dY_k(s) \tag{18}$$

where

$$\bar{h}^n_k(s) = \frac{1}{m^n_j}\sum_{i=1}^{m^n_j}\tilde{M}^n_j(X^i,s)(ap_k(X^i_s,s)-1) \tag{19}$$

Lemma 5 *Let $F(x) = \{x\}(1-\{x\})$. Then, for bounded $f(x)$ with bounded Lf^2,*

$$\Big|\mathbb{E}\Big(\gamma^n_{j+1}(X^i)f^2(X^i_{(j+1)\delta})(\eta^n_{(j+1)\delta}/\eta^n_{j\delta})^2|\mathcal{F}_{j\delta}\Big) - f^2\tilde{H}^{n,\delta}_{j\delta}(X^i_{j\delta})\delta\Big| = o(\delta),$$

where $o(\delta)\to 0$ as $\delta\to 0$ and $\tilde{H}^{n,\delta}_s(x)$ is nonnegative and given by

$$\tilde{H}^{n,\delta}_{j\delta}(X^i_{j\delta}) = \sum_{k=1}^{w}F\Big(\frac{ap_k(X^i_{j\delta},j\delta)}{\bar{h}^n_k(j\delta)+1}\Big)(\bar{h}^n_k(j\delta)+1)^2. \tag{20}$$

With the above lemmas, we are able to derive a sharp upper bound for the mean squared error at fixed time t, implying the convergence of V^n_t to V_t for each time t, namely, Theorem 1.

4.1 Related Proofs for the Convergence of V_t^n

Proof: (for Lemma 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 (14) reduces to

$$d\psi_s = -L\psi_s ds + \sum_{k=1}^{w}(ap_k(X_s) - 1)\psi_s ds. \tag{21}$$

Feynman-Kac Formula ([34]) 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 (21) 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 2) After simplifying notations, it is equivalent to proving:

$$\psi_t(X_t)M_t - \psi_0(X_0) = \int_0^t M_s\psi_s'\sigma(X_s)dB_s. \tag{22}$$

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 [4, page 81].

Lemma 6 *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 6, it is sufficient to show that

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

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

$$\mathbb{E}\left(\psi_r(X_r)M_r\theta_f^{\vec{Y}}(t)\theta_g^B(t)|\mathcal{F}_r^{\vec{Y}} \vee \mathcal{F}_r^B\right) = \Theta_r(X_r)M_r\theta_f^{\vec{Y}}(r)\theta_g^B(r) \tag{23}$$

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

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

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

$$d\Theta_r(X_r) \;=\; -\sqrt{-1}\left(\sum_{k=1}^w (f_k(r)-1)(ap_k(X_r,r)-1)\Theta_r(X_r)\right)dr + \Theta_r'\sigma(X_r)dB_r. \tag{24}$$

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

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

Combining with (23), we get

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

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

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

This finishes the proof of the lemma. ■

Proof: (of Lemma 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 $\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right) \le e^{K^2T}n \le K_1 n$. ■

Proof: (of Lemma 4) By Equation (3), we have

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

Observe that

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

This gives (17). Applying Itô's formula to the last two equations and simplifying, we obtain

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

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

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

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

Proof: (of Lemma 5) Note that $\eta^n_{(j+1)\delta}/\eta^n_{j\delta} = \hat{M}_j^n((j+1)\delta)$ and $\hat{M}_j^n(t)$ follows (17). Then,

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

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

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

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

Equation (12) gives $\gamma_{j+1}^n(X^i) = F(\tilde{M}_{j+1}^n(X^i))$. By telescoping and using (18), we obtain

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

Applying Itô's formula again, we have

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

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

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

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

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

and (the last inequality above is by $\mathbb{E}\big[(1-\tilde{M}^n_j(X^i,s))^2|\mathcal{F}_{j\delta}\big] \le K_1\delta$ and $\mathbb{E}\big(\hat{M}^n_j(s)^4|\mathcal{F}_{j\delta}\big)\le K_2$) as $\delta\to 0$,

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

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

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

$$\begin{aligned}
\langle V_t^n, \phi\rangle - \langle V_0^n, \psi_0\rangle &= \langle V_t^n, \psi_t\rangle - \langle V_{k\delta}^n, \psi_{k\delta}\rangle + \sum_{j=1}^{k} \left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle - \mathbb{E}\left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle | \mathcal{F}_{j\delta-}\right)\right) \\
&\quad + \sum_{j=1}^{k} \left(\mathbb{E}\left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle | \mathcal{F}_{j\delta-}\right) - \left\langle V_{(j-1)\delta}^n, \psi_{(j-1)\delta}\right\rangle\right) \\
&\equiv I_1^n + I_2^n + I_3^n. \qquad (25)
\end{aligned}$$

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 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

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

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 $\mathcal{F}_{j\delta,s}^i = \sigma(B_t^i - B_{j\delta}^i : j\delta \leq t \leq s)$ and $\mathcal{F}_{s,t}^{\vec{Y}} = \sigma(\vec{Y}_u - \vec{Y}_s : \; s \leq u \leq t)$.

Then,

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

It is easy to show that $\mathbb{E}\left((M_j^n(X^i,s))^4|\mathcal{F}_{j\delta}\right) \leq e^{K\delta}$, and using (BD) condition, $\mathbb{E}\left(|\sigma(X_s^i)|^4|\mathcal{F}_{j\delta}\right) \leq K_1$, and by the independent increments of Y and Lemma 1, $\mathbb{E}\left(\|\psi_s'\|_\infty^2|\mathcal{F}_{j\delta}\right) = \mathbb{E}\left(\|\psi_s'\|_\infty^2\right) \leq K_2$.

Hence, using $\sum_{j=0}^{k-1}\delta \leq t \leq T$ and applying Lemma 3, we obtain

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

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

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

Applying Lemmas 1, 5 and 3, we have

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

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

5 Convergence of V^n

Next, we study the convergence of V^n, regarding as a sequence of stochastic processes. 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 a 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

$$\begin{aligned} & \eta^n_{(j+1)\delta}\frac{1}{n}\sum_{i=1}^{m^n_j}\xi^i_{j+1}\delta_{X^i_{(j+1)\delta}} - \eta^n_{j\delta}\frac{1}{n}\sum_{i=1}^{m^n_j} M^n_{j+1}(X^i))\delta_{X^i_{(j+1)\delta}} \\ = \; & \eta^n_{(j+1)\delta}\frac{1}{n}\sum_{i=1}^{m^n_j}\left(\xi^i_{j+1} - \tilde{M}^n_{j+1}(X^i)\right)\delta_{X^i_{(j+1)\delta}}. \end{aligned}$$

Therefore,

$$\langle V^n_t, f\rangle = \langle V_0, f\rangle + \int_0^t \langle V^n_s, Lf\rangle\, ds + \sum_{k=1}^{w}\int_0^t \langle V^n_s, f(ap_k - 1)\rangle\, d\tilde{Y}_k(s) + N^{n,f}_t + \hat{N}^{n,f}_t, \tag{26}$$

where

$$N^{n,f}_t = \sum_{j=0}^{[t/\delta]}\frac{1}{n}\sum_{i=1}^{m^n_j}\int_{j\delta}^{((j+1)\delta)\wedge t} f'\sigma(X^i_s)dB^i_s\eta^n_{j\delta}$$

and

$$\hat{N}^{n,f}_t = \sum_{j=1}^{[t/\delta]}\eta^n_{j\delta}\frac{1}{n}\sum_{i=1}^{m^n_{j-1}}(\xi^i_j - \tilde{M}^n_j(X^i))f(X^i_{j\delta}).$$

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

$$\left\langle N^{n,f}\right\rangle_t = \sum_{j=0}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_j}\int_{j\delta}^{((j+1)\delta)\wedge t} |f'\sigma(X^i_s)|^2 ds(\eta^n_{j\delta})^2$$

and

$$\begin{aligned} \left\langle \hat{N}^{n,f}\right\rangle_t = \left\langle \hat{N}^{n,f}\right\rangle_{[t/\delta]\delta} &= \sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\mathbb{E}\left(\left(\sum_{i=1}^{m^n_{j-1}}(\xi^i_j - M^n_j(X^i))f(X^i_{j\delta})\right)^2 |\mathcal{F}_{j\delta-}\right)(\eta^n_{j\delta})^2 \\ &= \sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)f^2(X^i_{j\delta})(\eta^n_{j\delta})^2. \end{aligned} \tag{27}$$

With the distance defined in the end of Section 3, we are able to derive a sharp upper bound of the same order for the uniform (in t) mean squared error, implying the uniform convergence of V^n to V for t in $[0,T]$. Namely, we can prove Theorem 2. Then, we convert the uniform convergence result of V^n to that of π^n, that is, we prove Theorem 3.

5.1 Related Proofs for the Convergence of V^n

Proof: (of Theorem 2) We first 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. Note that

$$\mathbb{E}\sup_{t\le T}\tilde{d}(V_t^n, V_t)^2 \le \sum_{k=1}^{\infty} 2^{-k}\left(\mathbb{E}\sup_{t\le T}\langle V_t^n - V_t, f_k\rangle^2\right) + \mathbb{E}\sup_{t\le T}\langle V_t^n - V_t, 1\rangle^2 \tag{28}$$

By Equation (26) and Doob's maximum inequality,

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\langle V_t^n - V_t, f\rangle^2 \\
\le& K_2\int_0^T \mathbb{E}\langle V_t^n - V_t, Lf\rangle^2 dt + K_2\mathbb{E}\sum_{j=0}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{((j+1)\delta)} |f'\sigma(X_s^i)|^2 ds (\eta_{j\delta}^n)^2 \\
&+ K_2\sum_{k=1}^{w}\int_0^T \mathbb{E}\langle V_t^n - V_t, f(ap_k - 1)\rangle^2 dt + K_2\mathbb{E}\sum_{j=1}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2.
\end{aligned} \tag{29}$$

By Theorem 1, the first and third terms are bounded by $K_3 n^{-1}$. By Lemma 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 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 (28). Take $f = 1$ in (29), we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\langle V_t^n - V_t, 1\rangle^2 \\
\le& K_2\sum_{k=1}^{w}\int_0^T \mathbb{E}\langle V_t^n - V_t, (ap_k - 1)\rangle^2 dt + K_2\mathbb{E}\sum_{j=1}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\gamma_j^n(X^i)(\eta_{j\delta}^n)^2.
\end{aligned} \tag{30}$$

Again, Theorem 1 implies that the first term is bounded by $K_7 n^{-1}$. Clearly, Lemma 5 is true with $f = 1$, and a similar argument implies the second term of (30) is bounded by $K_8 n^{-1}$. Putting all the above estimates back to (29), we establish the desired result, since $d \le \tilde{d}$. ■

Proof: (of Theorem 3) Note that for f bounded by 1, we have

$$\begin{aligned}
|\langle \pi_t^n - \pi_t, f\rangle| &= \left|\frac{\langle V_t^n, f\rangle\langle V_t - V_t^n, 1\rangle + \langle V_t^n, 1\rangle\langle V_t^n - V_t, f\rangle}{\langle V_t^n, 1\rangle\langle V_t, 1\rangle}\right| \\
&\le \frac{|\langle V_t - V_t^n, 1\rangle|}{\langle V_t, 1\rangle} + \frac{|\langle V_t^n - V_t, f\rangle|}{\langle V_t, 1\rangle}.
\end{aligned} \tag{31}$$

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,

$$
\begin{aligned}
E^P \sup_{0\le t\le T} d(\pi_t^n,\pi_t) &= \mathbb{E}\sup_{0\le t\le T}\left\{\frac{1}{\langle V_t,1\rangle}|\,\langle V_t-V^n,1\rangle\,| + \frac{1}{\langle V_t,1\rangle}\tilde{d}(V_t^n,V_t)\right\}M_T \\
&\le \left(\mathbb{E}\sup_{0\le t\le T}|\,\langle V_t-V^n,1\rangle\,|^2\right)^{\frac12}\left(\mathbb{E}\sup_{0\le t\le T}\frac{M_T^2}{\langle V_t,1\rangle^2}\right)^{\frac12} \\
&\quad + \left(\mathbb{E}\sup_{0\le t\le T}\tilde{d}(V_t^n,V_t)^2\right)^{\frac12}\left(\mathbb{E}\sup_{0\le t\le T}\frac{M_T^2}{\langle V_t,1\rangle^2}\right)^{\frac12}. \qquad (32)
\end{aligned}
$$

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 2 and (32), there is a constant K such that (13) holds. ■

6 A Central Limit Type Theorem

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

$$U_t^n = n^{\frac12}(V_t^n - V_t) \quad \text{and} \quad \zeta_t^n = n^{\frac12}(\pi_t^n-\pi_t)$$

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 $n^{\frac12}$ which is better than that for the classical filtering model, which is $n^{(1-\alpha)/2}$ for $\alpha>0$ (see [11]). Then, we convert the results for ζ^n.

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

It turns out that the modified Schwarz space Φ is an appropriate space as it was used in [25]. We first briefly describe the modified Schwartz space.

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

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

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

$$\langle \hat{\phi}_1, \hat{\phi}_2 \rangle_{-\kappa} = \sum_{j=1}^{\infty} \langle \hat{\phi}_1, \phi_j^\kappa \rangle \langle \hat{\phi}_2, \phi_j^\kappa \rangle. \tag{33}$$

By a slight modification of Theorem 7, page 82, of [21], 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 [21], page 59).

Next, we prove tightness for $\{U^n\}$ in $D_{\Phi_{-\kappa}}[0, \infty)$ for an appropriate κ.

By (26) and (4), we have

$$\begin{aligned} \langle U_t^n, f \rangle &= \langle U_0^n, f \rangle + \int_0^t \langle U_s^n, Lf \rangle \, ds + \sum_{k=1}^{w} \int_0^t \langle U_s^n, f(ap_k - 1) \rangle \, d\tilde{Y}_k(s) \\ &\quad + n^{\frac{1}{2}} N_t^{n,f} + n^{\frac{1}{2}} \hat{N}_t^{n,f}, \end{aligned} \tag{34}$$

Using the above expression with suitable moment estimates, we are able to prove the tightness of $\{U^n\}$.

Theorem 4 *Under the assumptions of Theorem 1, there exists κ such that $\{U^n\}$ is tight in $D_{\Phi_{-\kappa}}[0, \infty)$.*

The tightness of $\{U^n\}$ implies that there exists a subsequence of $\{U^n\}$ converging to a U. Without loss of generality, we can just assume U^n converges weakly to U in the above modified Schwarz Space and we denote it by $U^n \Rightarrow U$.

6.2 Characterization of the Limits

From (34), in order to characterize the limit U, it suffices to characterize the limits of $n^{\frac{1}{2}} N_t^{n,f}$ and $n^{\frac{1}{2}} \hat{N}_t^{n,f}$. For the first one, it is easy to show that

$$n^{\frac{1}{2}} N_t^{n,f} \to 0. \tag{35}$$

The quadratic variation of the second one can be separated into two terms:

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

By Lemma 5, the first term is approximated by

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

where the approximation means the difference tends to 0 as $n\to\infty$ and $H_s(x)$ is given by

$$H_s(X_s)=\sum_{k=1}^{w}F\big(\frac{ap_k(X_s,s)}{\bar{h}_k(s)+1}\big)(\bar{h}_k(s)+1)^2,$$

with $h_k(s)=M(s)(ap_k(X_s,s)-1)$. Note that $h_k^n(s)\to h_k(s)$ in finite measure where $h_k^n(s)$ is given in Equation (19).

To characterize the second term, we need the following two more lemmas with the needed technical estimates. Lemma 7 provides the other key estimate, whose order is of $O(\delta^{3/2})$. This order is better than $O(\delta)$, the order in the classical nonlinear filtering case (see [11]), leading to a better convergence rate in the case of the FM model.

Lemma 7

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

Lemma 8

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

Thus, the second moment of the second term is bounded by

$$
\begin{aligned}
& n^2 \sum_{j=1}^{[t/\delta]} \mathbb{E}\Big(\frac{1}{n^2} \sum_{i=1}^{m_{j-1}^n} \gamma_j^n(X^i) f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2\Big)^2 \\
\leq \quad & ||f||_\infty^4 n^{-2} \sum_{j=1}^{[t/\delta]} \mathbb{E}\Big(m_{j-1}^n \sum_{i=1}^{m_{j-1}^n} \gamma_j^n(X^i)^2 (\eta_{j\delta}^n)^4\Big) \\
\leq \quad & K_1 n^{-2} \sum_{j=1}^{[t/\delta]} \delta^{3/2} \mathbb{E}\big(K(m_{j-1}^n)^2 (\eta_{(j-1)\delta}^n)^4\big) \\
\leq \quad & K_2 \delta^{1/2} \to 0
\end{aligned}
$$

Lemma 7 is applied in the second inequality and Lemma 8 in the last inequality. Combining the above results, we obtain:

Lemma 9

$$n^{\frac{1}{2}} \hat{N}_t^{n,f} \Longrightarrow M_t^f$$

which is a martingale uncorrelated to B and $\vec{Y}$ such that

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

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

$$M_t^f = \int_0^t \int_{\mathbb{R}} \sqrt{H_s(x) V_s(x) \langle V_s, 1\rangle} f(x) W(dsdx).$$

Summarizing these, we obtain the characterization of U.

Theorem 5 *Under the assumptions of Theorem 1, $U^n \Rightarrow U$ which is the unique solution to:*

$$
\begin{aligned}
\langle U_t, f\rangle \quad = \quad & \langle U_0, f\rangle + \int_0^t \langle U_s, Lf\rangle \, ds + \sum_{k=1}^{w} \int_0^t \langle U_{s-}, f(ap_k - 1)\rangle \, d\tilde{Y}_k(s) \\
& + \int_0^t \int_{\mathbb{R}} \sqrt{H_s(x) V_s(x) \langle V_s, 1\rangle} f(x) W(dsdx).
\end{aligned}
\tag{36}
$$

Finally, we would like to convert the characterization of U^n to that of ζ^n. The tightness of ζ^n is immediate by Kallianpur-Striebel formula and Theorem 5. Then, we obtain the characterization of ζ in the following theorem.

Theorem 6 *Under the assumptions of Theorem 1, $n^{\frac{1}{2}}(\pi_t^n - \pi_t) \Rightarrow \zeta_t$ which is the unique solution to:*

$$
\begin{aligned}
d\langle \zeta_t, f\rangle = & \langle \zeta_t, Lf - (a-w)f - f\langle \pi_t, a-w\rangle + (a-w)\langle \pi_t, f\rangle\rangle \, dt \\
& + \sum_{k=1}^{w} \Big[\frac{\langle \zeta_{t-}, f a p_k\rangle}{\langle \pi_{t-}, a p_k\rangle} - \frac{\langle \zeta_{t-}, a p_k\rangle \langle \pi_{t-}, f a p_k\rangle}{\langle \pi_{t-}, a p_k\rangle^2} - \langle \zeta_{t-}, f\rangle \Big] dY_k(t) \\
& + \int_{\mathbb{R}} \frac{f(x) - \langle \pi_t, f\rangle}{\langle V_t, 1\rangle} \sqrt{H_t(x) V_t(x) \langle V_t, 1\rangle} W(dxdt).
\end{aligned}
\tag{37}
$$

When $a(X_t,t)=a(t)$, depending only on time t, Equation (37) is simplified as below:

$$\begin{aligned} d\left\langle \zeta_t, f\right\rangle =& \left\langle \zeta_t, Lf\right\rangle dt + \int_{\mathbb{R}} \frac{f(x)-\left\langle \pi_t, f\right\rangle}{\left\langle V_t, 1\right\rangle} \sqrt{H_t(x)V_t(x)\left\langle V_t, 1\right\rangle} W(dxdt) \\ &+ \sum_{k=1}^{w} \Big[\frac{\left\langle \zeta_{t-}, fap_k\right\rangle}{\left\langle \pi_{t-}, ap_k\right\rangle} - \frac{\left\langle \zeta_{t-}, ap_k\right\rangle \left\langle \pi_{t-}, fap_k\right\rangle}{\left\langle \pi_{t-}, ap_k\right\rangle^2} - \left\langle \zeta_{t-}, f\right\rangle \Big] dY_k(t). \end{aligned} \tag{38}$$

Observe that $\langle \zeta_t, 1\rangle = 0$ for all t, because $\langle \pi_t^n, 1\rangle = \langle \pi_t, 1\rangle = 1$. This is a necessary condition for ζ_t, which is satisfied in (37) and (38).

6.3 Related Proofs of the Central Limit Type Theorem

Proof: (of Lemma 7) 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

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

Applying Itô's formula, we have

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

For the first term, we have

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

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

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

Proof: (of Lemma 8) 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

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

■

Proof: (of Theorem 5) By Lemma 9 and (35), it is easy to show that U satisfies (36). To prove the uniqueness, we take another solution $\tilde{U}$ of (36) and define $\hat{U}_t = U_t - \tilde{U}_t$. Then $\hat{U}_t$ satisfies the following homogeneous linear equation

$$\Big\langle \hat{U}_t, f\Big\rangle = \int_0^t \Big\langle \hat{U}_s, Lf\Big\rangle ds + \sum_{k=1}^{w}\int_0^t \Big\langle \hat{U}_s, f(ap_k-1)\Big\rangle d\tilde{Y}_k(s).$$

Similar to Lemma 4.2 in [30] we get $\hat{U}=0$. ■

Proof: (of Theorem 6) From Equation (31), we can see that

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

which converges to

$$\zeta_t \equiv \langle V_t, 1\rangle^{-1}\Big(U_t - \langle V_t, 1\rangle^{-1}\langle U_t, 1\rangle V_t\Big).$$

Let $\eta_t = \langle V_t, 1\rangle^{-1} U_t$. 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.

$$\begin{aligned} d\langle \eta_t, f\rangle =& \big(\langle \eta_t, Lf-(a-w)f\rangle + \langle \eta_t, f\rangle\langle \pi_t, a-w\rangle\big)dt \\ &+ \sum_{k=1}^{w}\Big[\frac{\langle \eta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \langle \eta_{t-}, f\rangle\Big]dY_k(t) \\ &+ \int_{\mathbb{R}} \frac{f(x)}{\langle V_t, 1\rangle}\sqrt{H_t(x)V_t(x)\langle V_t, 1\rangle}W(dxdt). \end{aligned} \tag{39}$$

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

$$\begin{aligned} d\langle \eta_t, f\rangle = \langle \eta_t, Lf\rangle\, dt + \sum_{k=1}^{w} \Big[\frac{\langle \eta_{t-}, fap_k\rangle}{\langle \pi_{t-}, ap_k\rangle} - \langle \eta_{t-}, f\rangle \Big] dY_k(t) \\ + \int_{\mathbb{R}} \frac{f(x)}{\langle V_t, 1\rangle} \sqrt{H_t(x) V_t(x) \langle V_t, 1\rangle}\, W(dxdt). \end{aligned} \tag{40}$$

Observe that $\zeta_t = \eta_t - \langle \eta_t, 1\rangle \pi_t$. Applying Itô's formula again, we get Equation (37) for ζ. When $a(X_t, t) = a(t)$, the simplified (40) and (6) gives (38). The uniqueness comes from the similar argument of Theorem 5. ■

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.

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. 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.

References

[1] Y. Aït-Sahalia, P.A. Mykland and L. Zhang (2005), How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise. *Review of Financial Studies*, 18, 351-416.

[2] F. M. Bandi and J. R. Russell (2006), Separating microstructure noise from volatility. *Journal of Financial Economics*, **79**, 655-692.

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

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

[5] P. Brémaud (1981). *Point Processes adn Queues: Martingale Dynamics*, Springer-Verlag, New York.

[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. *Applied Mathematics and Optimization* **54**, 293–314.

[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 and J. Xiong (2006), A central limit type theorem for particle filter. *Comm. Stoch. Analysis* **1**, 103-122.

[12] J. Cvitanic, R. Liptser, and B. Rozovskii (2006), A filtering approach to tracking volatility from prices observed at random times. *Annals of Applied Probability*, 16, 1633-1652.

[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 and 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] J. Fan, and Y. Wang, (2007). Multi-scale jump and volatility analysis for high-Frequency financial data. *Journal of American Statistical Association*, **102**, 1349-1362.

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

[22] R. Frey and W.J. Runggaldier (2001), A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal of Theoretical and Applied Finance* 4, 199-210.

[23] 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.

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

[25] 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.

[26] J. Jacod and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes*, Springer-Verlag, 2nd edition, New York.

[27] G. Kallianpur (1980). *Stochastic Filtering Theory*, Springer-Verlag, New York.

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

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

[30] 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.

[31] K. Lee and Y. Zeng (2010). Risk minimization for a filtering micromovement model of asset price. *Applied Mathematical Finance*, **17**, 177 - 199.

[32] Y. Li and P. A. Mykland (2007). Are volatility estimators robust with respect to modeling assumptions? *Bernoulli*, **13**, 601 - 622.

[33] J. Xiong and Y. Zeng (2008) Mean-variance portfolio selection for A Filtering Point Process Model of Asset Price. Working paper. University of Tennessee.

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

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

[36] 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.

[37] 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.