Generalised Particle Filters

Kai Li
Department of Mathematics
Imperial College London

Thesis presented for examination for the degree of
Doctor of Philosophy in Mathematics
of Imperial College London

December 2012

Declaration

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Kai Li

Copyright

Acknowledgement

I would like to express my deepest gratitude and thanks to my supervisor Professor Dan Crisan. It is his continuous and boundless support, both academic and other parts of life, that makes my every single progress and achievement possible throughout these four years.

I would like to thank Professor Andrew Stuart for his generous advice on the key algorithm of my thesis, as well as Dr. Salvador Ortiz-Latorre and Dr. Nikolas Kantas for sharing their ideas and codes on the numerical implementations, with whose help this thesis has been greatly improved.

I would like to give my sincere thanks to numerous PhD students, post-doctorial researchers, and lectures at Imperial College for their many interesting and inspiring discussions I enjoyed very much during my PhD studies.

I would also like to acknowledge the Department of Mathematics, Imperial College for the financial support I received during my study and preparation for this thesis.

Finally, but most importantly, I would like to thank my parents for their dedication and many years of unconditional support that provided the foundation for this work; as well as my wife Yao, who has always been standing besides me with her support and encouragement.

Abstract

The ability to analyse, interpret and make inferences about evolving dynamical systems is of great importance in different areas of the world we live in today. Various examples include the control of engineering systems, data assimilation in meteorology, volatility estimation in financial markets, computer vision and vehicle tracking. In general, the dynamical systems are not directly observable, quite often only partial information, which is deteriorated by the presence noise, is available. This naturally leads us to the area of stochastic filtering, which is defined as the estimation of dynamical systems whose trajectory is modelled by a stochastic process called the signal, given the information accumulated from its partial observation.

A massive scientific and computational effort is dedicated to the development of various tools for approximating the solution of the filtering problem. Classical PDE methods can be successful, particularly if the state space has low dimensions (one to three). In higher dimensions (up to ten), a class of numerical methods called particle filters have proved the most successful methods to-date. These methods produce approximations of the posterior distribution of the current state of the signal by using the empirical distribution of a cloud of particles that explore the signal's state space.

In this thesis, we discuss a more general class of numerical methods which involve *generalised* particles, that is, particles that evolve through spaces larger than the signal's state space. Such generalised particles include Gaussian mixtures, wavelets, orthonormal polynomials, and finite elements in addition to the classical particle methods. This thesis contains a rigorous analysis of the approximation of the solution of the filtering problem using Gaussian mixtures. In particular we deduce the L^2-convergence rate and obtain the central limit theorem for the approximating system. Finally, the filtering model associated to the Navier-Stokes equation will be discussed as an example.

Contents

List of Figures

Notations

- $(\Omega, \mathcal{F}, \mathbb{P})$ - probability triple consisting of a sample space Ω, the σ-algebra $\mathcal{F}$ which is the set of all measurable events, an the probability measure $\mathbb{P}$.
- $(\mathcal{F}_t)_{t\geq 0}$ - a filtration, an increasing family of sub-σ-algebras of $\mathcal{F}$; $\mathcal{F}_s \subset \mathcal{F}_t$, $0 \leq s \leq t$.
- $\mathbb{R}^d$ - the d-dimensional Euclidean space.
- $\overline{\mathbb{R}^d}$ - the one-point compactification of $\mathbb{R}^d$ formed by adding a single point at infinity to $\mathbb{R}^d$.
- $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ - the state space of the signal. Normally $\mathbb{S}$ is taken as a complete separable space, and $\mathcal{B}(\mathbb{S})$ is the associated Borel σ-algebra, that is, the σ-algebra generated by the open sets in $\mathbb{S}$.
- $B(\mathbb{S})$ - the space of bounded $\mathcal{B}(\mathbb{S})$-measurable functions from $\mathbb{S}$ to $\mathbb{R}$.
- $\mathcal{P}(\mathbb{S})$ - the family of Borel probability measures on space $\mathbb{S}$.
- $C_b(\mathbb{R}^d)$ - the space of bounded continuous functions on $\mathbb{R}^d$.
- $C_b^m(\mathbb{R}^d)$ - the space of bounded continuous functions on $\mathbb{R}^d$ with bounded derivatives up to order m.
- $C_0^m(\mathbb{R}^d)$ - the space of continuous functions on $\mathbb{R}^d$, vanishing at infinity with continuous partial derivatives up to order m.
- $\|\cdot\|_\infty$ - the supremum norm for $\varphi : \mathbb{R}^d \to \mathbb{R}$: $\|\varphi\|_\infty = \sup_{x\in\mathbb{R}^d} \|\varphi(x)\|$.
- $\|\cdot\|_{m,\infty}$ - the norm such that for φ on $\mathbb{R}^d$, $\|\varphi\|_{m,\infty} = \sum_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^d} |D_\alpha \varphi(x)|$, where $\alpha = (\alpha^1, \ldots, \alpha^d)$ is a multi-index and $D_\alpha = (\partial_1)^{\alpha_1} \cdots (\partial_d)^{\alpha_d}$.

- $\mathcal{M}_F(\mathbb{R}^d)$ - the set of finite measures on $\mathbb{R}^d$.

- $\mathcal{M}_F(\overline{\mathbb{R}^d})$ - the set of finite measures on $\overline{\mathbb{R}^d}$.

- $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,T]$ - the space of càdlàg functions (or right continuous functions with left limits) $f:[0,T]\to\mathcal{M}_F(\mathbb{R}^d)$.

- $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ - the space of càdlàg functions (or right continuous functions with left limits) $f:[0,\infty)\to\mathcal{M}_F(\mathbb{R}^d)$.

Chapter 1

Introduction

1.1 Preamble

The ability to analyse, manipulate, and interpret data is of crucial and increasing importance nowadays. The scope of the related areas is huge and includes satellite positioning, communications, finance and econometrics, etc (see, for example, [32] and [8]). We are, however, rarely able to fully and directly observe many of these phenomena and data which are crucial to our lives; it is often the case that only partial information about the phenomena is available. It is therefore important to know what analysis one can do and what conclusions one can make about the data, or the *signal*, from its partial *observation* which is perturbed by noises.

Stochastic filtering is an area which enables us to deal with and solve this type of problem. Generally speaking, stochastic filtering deals with the estimation of an evolving dynamical system, called the *signal*, using some *partial observations* and a stochastic model. The signal is modelled by a stochastic process, usually denoted by $X = \{X_t,\ t \geq 0\}$, defined on a generic probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where t is the temporal parameter. As mentioned above, the signal process is not available to observe directly; instead, a partial observation is obtained and it is modelled by another continuous process $Y = \{Y_t,\ t \geq 0\}$. Therefore, the aims of the filtering problem is to make inferences about the process X using the information obtained from recording the process Y.

The original signal cannot be observed or measured fully partly because of the presence of noise. Thus we can define the observation process as a function of the signal X and the (measurement) noise, which is modelled by another stochastic process $W = \{W_t,\ t \geq 0\}$, that is:

$$Y_t = f_t(X, W_t), \qquad t \in [0, \infty).$$

The information available from the observation up to time t is defined as the filtration $\mathcal{Y} = \{\mathcal{Y}_t,\ t \geq 0\}$ generated by the observation process Y, that is:

$$\mathcal{Y}_t = \sigma(Y_s,\ s \in [0, t]), \qquad t \geq 0.$$

Given all these settings, many possible inferences can be made about the signal process X. For example:

- The best estimate $\hat{X}_t$ of the value of the signal at time t given the observation up to time t. If we mean the *best estimate*, for example, as the best mean square estimate, it is equivalent to compute $\mathbb{E}[X_t \mid \mathcal{Y}_t]$, the conditional expectation of X_t given $\mathcal{Y}_t$.

- The estimate of the difference $X_t - \hat{X}_t$ given the observation up to time t. In the case the signal is real-valued, one may wish to compute the conditional variance
$$\mathbb{E}[(X_t - \hat{X}_t)^2 \mid \mathcal{Y}_t] = \mathbb{E}[X_t^2 \mid \mathcal{Y}_t] - \mathbb{E}[X_t \mid \mathcal{Y}_t]^2.$$

- The probability that the signal at time t can be found within a certain set A given the observation up to time t, which means computing
$$\mathbb{P}(X_t \in A \mid \mathcal{Y}_t) = \mathbb{E}[\mathbf{1}_{\{X_t \in A\}} \mid \mathcal{Y}_t],$$
the conditional probability of $\{X_t \in A\}$ given $\mathcal{Y}_t$.

One can see that all the above inferences require the computation of one or more quantities of the form $\mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t]$, where φ is a real-valued function defined on the state space of the signal. Furthermore, instead of the fragments of information about X_t obtained from the above quantities, we would like to know all the information about X_t which is contained in $\mathcal{Y}_t$. That means that, we want to compute π_t — the conditional distribution of X_t given $\mathcal{Y}_t$. This π_t is defined as a probability measure-valued random variable measurable with respect to $\mathcal{Y}_t$ so that

$$\pi_t(\varphi) = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t] = \int_{\mathbb{S}} \varphi(x)\pi_t(dx) \tag{1.1}$$

for all statistics φ for which both sides of (1.1) make sense and $\mathbb{S}$ is the state space of the signal. It follows that knowing π_t will enable us, at least theoretically, to compute any inference of X_t given $\mathcal{Y}_t$ if we integrate function φ with respect to π_t. It is well known (see, for example, [64], [65], [66], [71]) that, when X_t and Y_t are both diffusion processes, π_t is the solution of the *Kushner-Stratonovich equation* and its unnormalised version satisfies the *Zakai equation*. Both equations are stochastic partial differential equations, and the corresponding stochastic integral parts are martingales under certain conditions (details in Chapter 2).

1.2 Overview

Since the analytical solutions to both the Zakai and the Kushner-Stratonovich equations are rarely available, numerical algorithms for solving the filtering equations are required. This thesis is therefore devoted to obtaining numerical approximations of the filtering equations in the continuous time setting.

Among the existing numerical methods for solving the filtering problem, *particle filters* (also known as *sequential Monte Carlo method*) are among the most popular ones. Essentially, particle filters are algorithms that approximate π_t with discrete random measures of the form

$$\sum_i a_i(t)\delta_{v_i(t)},$$

where $\delta_{v_j(t)}$ is the Dirac delta measure centred at $v_j(t)$. In other words, we compute the empirical distributions associated with sets of randomly located particles with weights $a_1(t), a_2(t), \ldots$, and positions $v_1(t), v_2(t), \ldots$. As time increases, typically the trajectories of a large number of particles diverge from the signal's trajectory; with only a small number remaining close to the signal. The weights of the diverging particles decrease rapidly, therefore contributing very little to the approximating system, and causing the approximation to converge very slowly to the conditional distribution. In order to tackle this so-called *sample degeneracy* phenomenon, a *correction procedure* is added. At correction times, each particle is replaced by a random number of offspring. Redundant particles are abandoned and only the particles contributing significantly to the system (i.e. with large weights) are carried forward; so that the most probable region of the trajectory of the signal process X will be more thoroughly explored. This correction mechanism is also called branching or resampling. Currently the tree based branching algorithm (TBBA) and multinomial branching algorithm are two approaches for the correction step.

The introduction of particle methods (or sequential Monte Carlo methods) in solving the filtering problem dates back to 1960's by Handschin and Mayne ([40]), and Akashi and Kumamoto ([1]). In the mid 1990's, several particle filtering algorithms were proposed by various people. See, for example, Gordon, Salmond and Ewing ([38]), Gordon, Salmond and Smith ([39]), Kitagawa ([45]), and Carvalho, Del Moral, Monin and Salut ([12]). The first convergence results for particle filters were published by Del Moral ([27], [28]) and independently by Crisan and Lyons ([21]). Various authors made several improvements on the results subsequently, see, for instance, Crisan and Lyons ([22], [23]), Crisan, Del Moral and Lyons ([17]), Crisan and Doucet ([18]), Del Moral and Guionnet ([30]), Del Moral and Miclo ([31]), and Le Gland and Oudjane ([52]). The area of particle filters is still very active, and has huge amount of research outcomes each year.

The above mentioned classic particle filters make use of a mixture of Dirac measures to construct the particle approximation; several attempts have been made to generalise this idea. Kotecha and Djurić ([46]) first introduced the so called Gaussian particle filters, where they used a single Gaussian to approximate the posterior distribution. They shortly improved their initial work and built the approximations by weighted Gaussian mixtures (See [47]). Up to now, most of the existing work has been closely related to the extended (or ensemble) Kalman filter, because of the Gaussian nature of Kalman filter, and this method of Gaussian mixture approximations may provide a way to improve the asymptotic behaviour for the ensemble Kalman filter (see discussions in [51]). The majority of the previous work is in the discrete time framework. Reich ([59]) recently took a Gaussian mixture to generalise the ensemble Kalman filter and designed a new algorithm based on continuous time formulation, where the approximation was constructed using Gaussian mixtures without weights. See Flament et al ([36]), Van der Merwe and Wan ([69]), Carmi et al ([10]), Iglesias, Law and Stuart ([42]), and Lee ([53]) for more related work.

A major application of the Gaussian mixture approximation is the problem of filtering the Navier-Stokes equation. There has been a huge amount of literatures studying various properties of Navier-Stokes equation; however, the study of the filters of Navier-Stokes equation started quite recently. Brett et al ([6]) chose the Navier-Stokes equation as the forward model, then formulated data assimilation as a Bayesian inverse problem and derived the Gaussian approximation filters. Law and Stuart ([50]) used an MCMC algorithm and showed (by numerical simulation) that, by appropriate parameter choices, approximate filters can perform well in reproducing the mean, whereas do not generally do well in reproducing the covariance. More recent studies on filtering the Navier-Stokes equation are carried out by, for example, Iglesias, Law and Stuart ([41]), and Beskos et al ([4]).

However, to the author's knowledge, there has been no existing literature – including all these mentioned in the above paragraphs – containing rigorous mathematical study on the convergence results of such Gaussian mixture methods. Therefore, it is of great interests to fill this gap; and this is part of the aim of this thesis. In addition, this thesis will aim to generalise the current particle filters framework and build a new approximating system, which I call the *generalised particle filters*. In the generalised particle filters system, the '*generalised positions*' may take values on spaces which are (possibly) larger than the state space of the signal; and the '*generalised weights*' do not necessarily satisfy the same evolution equation as the classic weights.

Four possible settings (Gaussian mixtures, wavelets, orthonormal polynomials, and finite elements) are discussed in this thesis. The key ideas of all these four methods are similar. To be specific, we would like to find appropriate ways to construct the approximations of the solution of the Zakai equation, so that the evolution equation satisfied by the approximating measures are 'sufficiently' close the the Zakai equation. Then we are able to show that the approximating measures are 'sufficiently' close to the solution of the Zakai equation and obtain the convergence of the approximating measures to the real measure. The ultimate aim is to integrate within the framework of generalised particle filters a wide variety of numerical methods including the above ones, and develop a common approach to analysing and comparing the existing numerical methods for solving the filtering problem.

In particular, this thesis will concentrate on the approximation using mixtures of Gaussian measures. Convergence and central-limit type results will be obtained for the Gaussian mixture approximating system. The application to the filtering of the solution of the Navier-Stokes equation will also be discussed.

1.3 Contents of the Thesis

In view of the ideas and settings described above, this thesis is devoted to formulating the stochastic filtering framework and solving the filtering equations, both theoretically and numerically. The thesis is distributed as follows:

In Chapter 2, we review the existing results on stochastic filtering theory[1]. The filtering framework is introduced first, with the focus on the problems where the signal X and observation Y are diffusion processes and the state space $\mathbb{S}$ of the signal process is a complete separable metric space (usually we take $\mathbb{S} = \mathbb{R}^d$). After that, we will discuss the change of measure approach to deducing the Zakai equation, which is the linear stochastic partial differential equation satisfied by the unnormalised version of the conditional distribution of X. Deducing the evolution equation satisfied by the normalised conditional distribution π_t — the Kushner-Stratonovich equation, is a simple consequence of the Zakai equation.

Chapter 3 contains an introduction of the *generalised particle filters with Gaussian mixtures*. We will first briefly discuss the classic particle filtering methods. We

[1] For detailed history of the development of stochastic filtering problem, see, for example, Section 1.3 of [3].

will then generalise the classic particle approximation system and suggest a different approach. A natural generalisation is to replace the Dirac measures by Gaussian measures with non-zero covariance matrix. The approximating system using this mixtures of Gaussian measures will be set out, with the aim of obtaining the approximations of the solutions to the Zakai and Kushner-Stratonovich equations. The Tree Based Branching Algorithm (TBBA) and Multinomial branching algorithm are discussed for the branching mechanism.

Chapters 4 and 5 contain the main results of the thesis. Chapter 4 includes the law of large numbers theorem associated to the approximating system. In this chapter, the evolution equations of the approximating systems introduced in the previous chapter are derived. It is shown that, under certain conditions, the unnormalised and normalised versions of the approximations of the conditional distribution converge to the solutions of the Zakai equation and the Kushner-Stratonovich equation, respectively.

To be specific, we have the following result:

Theorem 1.3.1. *We denote by ρ_t^n the approximating measure of ρ_t – the solution of the Zakai equation, and π_t^n the approximating measure of π_t – the solution of the Kushner-Stratonovich equation, where n is the number of the generalised particles. Then under certain conditions (see Chapter 4 for details), for any $T > 0$ and $m \geq 6$, there exist constants c^T and $\tilde{c}^T$ independent of the n, such that for any test function $\varphi \in C_b^{m+2}(\mathbb{R})$*

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} (\rho_t^n(\varphi) - \rho_t(\varphi))^2\right] \leq \frac{c^T}{n}\|\varphi\|^2_{m+2,\infty}; \tag{1.2}$$

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} |\pi_t^n(\varphi) - \pi_t(\varphi)|\right] \leq \frac{\tilde{c}^T}{\sqrt{n}}\|\varphi\|_{m+2,\infty}. \tag{1.3}$$

$\tilde{\mathbb{E}}$ is the expectation under probability $\tilde{\mathbb{P}}$, which is a new probability obtained by Girsanov transformation such that under $\tilde{\mathbb{P}}$ the observation Y is a Brownian motion independent of the signal X. The derivation of $\tilde{\mathbb{P}}$ is discussed in Chpater 2.

Computer simulation results for Beneš filter are given at the end of this chapter.

In Chapter 5 we obtain a central limit type result. The error between the approximations and the true solutions are recalibrated and shown to form a tight sequence and their limit in distribution is obtained. Specifically, we have:

Theorem 1.3.2. *We adopt the same notations as in Theorem 1.3.1, and we define the measure-valued processes U^n and $\bar{U}^n$ as*

$$U_t^n \triangleq \sqrt{n}(\rho_t^n - \rho_t), \qquad \bar{U}_t^n \triangleq \sqrt{n}(\pi_t^n - \pi_t).$$

Then U^n is a tight sequence and converges in distribution to a $D_{\mathcal{M}(\overline{\mathbb{R}})}[0,\infty)$-valued process U satisfying

$$U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)ds + \int_0^t U_s(h\varphi)dY_s + \Lambda_t^{\varphi}.$$

Furthermore, U is pathwise unique (see Chapter 2 for the definitions of A and h, and Chapter 5 for the definition of Λ^{φ}). The process $\bar{U}$ satisfies

$$\bar{U}_t(\varphi) = \frac{1}{\rho_t(\mathbf{1})}\left(U_t(\varphi) - \pi_t(\varphi)U_t(\mathbf{1})\right).$$

Chapter 6 contains the possible areas of future research. In Section 6.1 we introduce the basic ideas of constructing the generalised particle filters with other possible tools, which include wavelet methods, orthomormal polynomial method, and finite element method. In the following section, we discuss an important application of the generalised particle filters, especially the Gaussian measures. Two-dimensional Navier-Stokes equation is considered as an example. The final chapter will be the conclusion and summary of the whole thesis.

Chapter 2

The Classic Filtering Theory

In this chapter, we review the filtering framework and introduce the change of measure methods for deducing the Zakai equation, which describes the evolution of the unnormalised version of the filtering solution. Then the Kushner-Stratonovich equation, which is satisfied by (the normalised version of) the conditional distribution, is also discussed. Both existence and uniqueness conditions are introduced.

2.1 Filtering Framework

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space together with a filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual conditions. On $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a $\mathcal{F}_t$-adapted process $X = \{X_t; t \geq 0\}$ which takes value in a complete separable metric space $\mathbb{S}$ (the state space, usually $\mathbb{S} = \mathbb{R}^d$). Let $\mathcal{S}$ be the associated Borel σ-algebra $\mathcal{B}(\mathbb{S})$. The process X is assumed to have paths which are càdlàg. We call process X the *signal* process. Define

$$\mathcal{X}_t = \sigma(X_s; s \in [0, t]) \vee \mathcal{N},$$

where $\mathcal{N}$ is the collection of all $\mathbb{P}$-null sets on $(\Omega, \mathcal{F})$, then $\mathcal{X}_t$ is the usual augmentation with null sets of the filtration associated with the process X. Then define

$$\mathcal{X} \triangleq \bigvee_{t\in\mathbb{R}^+} \mathcal{X}_t = \sigma\left(\bigcup_{t\in\mathbb{R}^+} \mathcal{X}_t\right).$$

Let $B(\mathbb{S})$ be the space of all bounded $\mathcal{B}(\mathbb{S})$-measurable functions, let $A : B(\mathbb{S}) \to B(\mathbb{S})$ be a possibly unbounded linear operator and denoted by $\mathcal{D}(A)$ the domain of

A which is a subset of $B(\mathbb{S})$. We assume that $\mathbf{1} \in \mathcal{D}(A)$ and $A\mathbf{1} = 0$. This definition implies that if $f \in \mathcal{D}(A)$ then Af is bounded.

Let $\pi_0 \in \mathcal{P}(\mathbb{S})$ and assume that X is a solution of the martingale problem for (A, π_0). In other words, assume the distribution of X_0 is π_0 and the process $M^\varphi = \{M_t^\varphi; t \geq 0\}$ defined as

$$M_t^\varphi = \varphi(X_t) - \varphi(X_0) - \int_0^t (A\varphi)(X_s)ds, \quad t \geq 0 \tag{2.1}$$

is an $\mathcal{F}_t$-adapted martingale for any $\varphi \in \mathcal{D}(A)$. The operator A is called the *infinitesimal generator* of the process X.

Let $h = (h_i)_{i=1}^m : \mathbb{S} \to \mathbb{R}^m$ be a measurable function such that

$$\mathbb{P}\left(\int_0^t \|h(X_s)\|ds < \infty\right) = 1 \tag{2.2}$$

for all $t \geq 0$, where $\|\cdot\|$ is the Euclidean norm, meaning $\|a\| = \sqrt{\sum_{i=1}^d \sum_{j=1}^p a_{ij}^2}$ for a $d \times p$ matrix a. Let W be a standard $\mathcal{F}_t$-adapted m-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ independent of X, and Y be the process which satisfies the following evolution equation

$$Y_t = Y_0 + \int_0^t h(X_s)ds + W_t, \tag{2.3}$$

Condition (2.2) ensures that the Riemann integral in equation (2.3) exists almost surely. This process $Y = \{Y_t; t \geq 0\}$ is called the *observation* process. Let $\{\mathcal{Y}_t, t \geq 0\}$ be the usual augmentation of the filtration associated with the process Y, viz

$$\mathcal{Y}_t = \sigma(Y_s, s \in [0,t]) \vee \mathcal{N},$$

$$\mathcal{Y} = \bigvee_{t \in \mathbb{R}^+} \mathcal{Y}_t = \sigma\left(\bigcup_{t \in \mathbb{R}^+} \mathcal{Y}_t\right).$$

Then note that since by the measurability of h, Y_t is $\mathcal{F}_t$-adapted, it follows that $\mathcal{Y}_t \subset \mathcal{F}_t$.

Now we are in the position to formally define the filtering problem:

Definition 2.1.1. *The filtering problem consists in determining the conditional distribution π_t of the signal X at time t given the information accumulated from observing Y in the interval $[0,t]$; that is, for $\varphi \in B(\mathbb{S})$, computing*

$$\pi_t(\varphi) = \int_{\mathbb{S}} \varphi(x)\pi_t(dx) = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t]. \tag{2.4}$$

From the following theorem (see, for example, in [3]), we see that π_t can be formalised by defining a (probability) measure-valued stochastic process describing the conditional distribution.

Theorem 2.1.2. *Let $\mathbb{S}$ be a complete metric space and $\mathcal{S}$ be the associated Borel σ-algebra. Then there exists a $\mathcal{P}(\mathbb{S})$-valued $\mathcal{Y}_t$-adapted process $\pi = \{\pi_t : t \geq 0\}$ such that for any $f \in \mathcal{B}(\mathbb{S})$*

$$\pi_t(f) = \mathbb{E}[f(X_t) \mid \mathcal{Y}_t] \quad \mathbb{P} - a.s.,$$

where $\mathcal{P}(\mathbb{S})$ is the space of probability measures over $(\mathbb{S}, P(\mathbb{S}))$ and $P(\mathbb{S})$ is the set of all subsets of $\mathbb{S}$. In particular, the identity

$$\pi_t^{\omega}(A) = \mathbb{P}[X_t \in A \mid \mathcal{Y}_t](\omega) \tag{2.5}$$

holds true almost surely for any $A \in \mathcal{B}(\mathbb{S})$.

Moreover if Y satisfies the evolution equation

$$Y_t = Y_0 + \int_0^t h(X_s)ds + W_t, \quad t \geq 0, \tag{2.6}$$

where $W = \{W_t : t \geq 0\}$ is a standard $\mathcal{F}_t$-adapted m-dimensional Brownian motion and $h = (h_i)_{i=1}^m$ is a measurable function such that

$$\mathbb{E}\left[\int_0^t \|h(X_s)\|ds\right] < \infty \tag{2.7}$$

and

$$\mathbb{P}\left(\int_0^t \|\pi_s(h)\|^2 ds < \infty\right) = 1 \tag{2.8}$$

for all $t \geq 0$, then π has a $\mathcal{Y}_t$-adapted progressively measurable modification. Furthermore, if X is càdlàg then π_t can be chosen to have càdlàg paths.

There are two commonly used particular cases of the signal process X_t, which are diffusion process and Markov chain with a finite state space respectively. The following example discusses the first case, which will be used throughout this thesis.

Example 2.1.3 (X is a Diffusion Process)**.** *Let $X = (X^i)_{i=1}^d$ be the solution of a d-dimensional stochastic differential equation driven by a p-dimensional Brownian motion $V = (V^j)_{j=1}^p$:*

$$X_t^i = X_0^i + \int_0^t f^i(X_s)ds + \sum_{j=1}^p \int_0^t \sigma^{ij}(X_s)dV_s^j, \qquad i = 1, \ldots, d \tag{2.9}$$

We assume that both $f = (f^i)_{i=1}^d : \mathbb{R}^d \to \mathbb{R}^d$ *and* $\sigma = (\sigma^{ij})_{i=1,\ldots,d;j=1,\ldots,p} : \mathbb{R}^d \to \mathbb{R}^{d\times p}$ *are globally Lipschitz*[1]*. Under the globally Lipschitz condition,* (2.9) *has a unique solution (Theorem 5.2.9 in [44]). The generator* A *associated with the process* X *is the second-order differential operator*

$$A = \sum_{i=1}^{d} f^i \frac{\partial}{\partial x^i} + \sum_{i=1}^{d}\sum_{j=1}^{d} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j}, \tag{2.10}$$

where $a = (a^{ij})_{i,j=1,\ldots,d} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ *is the matrix-valued function defined as*

$$a^{ij} = \frac{1}{2}\sum_{k=1}^{p} \sigma^{ik}\sigma^{kj} = \frac{1}{2}(\sigma\sigma^\top)^{ij}, \qquad i,j = 1,\ldots,d; \tag{2.11}$$

and $\sigma^\top$ *is the transpose of* σ*.*

Remark 2.1.4. *For the case where* X *is a Markov process with a finite number of states, see section 3.2.2 of [3].*

The proofs of the results in the reminder of this chapter, unless otherwise stated, can be found in [3].

2.2 Theoretical Results

Among the possible ways of deducing the evolution equation for π, there are two commonly used approaches, namely *the change of measure method* and *the innovation process method.* We revisit briefly change of probability measure method. For the second approach, see Section 3.7 of [3].

In the change of measure method, we construct a new probability measure on Ω, under which the process Y is a Brownian motion; and then we represent π in terms of its unnormalised version ρ, which is shown to satisfy a linear evolution equation. An application of Itô's formula gives us the evolution equation satisfied by π.

Firstly, let $Z = \{Z_t, t \geq 0\}$ be the process defined by

$$Z_t = \exp\left(-\sum_{i=1}^{m} \int_0^t h^i(X_s)dW_s^i - \frac{1}{2}\sum_{i=1}^{m}\int_0^t h^i(X_s)^2 ds \right), \quad t \geq 0. \tag{2.12}$$

[1]That is to say, $\exists K > 0$, so that for $\forall x, y \in \mathbb{R}^d$, $\|f(x) - f(y)\| \leq K\|x-y\|$ and $\|\sigma(x) - \sigma(y)\| \leq K\|x-y\|$.

Instead of considering Novikov's condition, which is difficult to verify directly, we consider the following condition

$$\mathbb{E}\left[\int_0^t \|h(X_s)\|^2 ds\right] < \infty, \quad \mathbb{E}\left[\int_0^t Z_s\|h(X_s)\|^2 ds\right] < \infty \tag{2.13}$$

Proposition 2.2.1. *If condition* (2.13) *holds, then the process* $Z = \{Z_t, t \geq 0\}$ *defined in* (2.12) *is an* $\mathcal{F}_t$*-adapted martingale.*

Having these conditions, and notice the fact that $Z_t > 0$ a.s. for fixed $t \geq 0$, we introduce a probability measure $\tilde{\mathbb{P}}^t$ on $\mathcal{F}_t$ by specifying its Radon-Nikodym derivative with respect to $\mathbb{P}$ to be given by Z_t, viz

$$\left.\frac{d\tilde{\mathbb{P}}^t}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t.$$

We define a probability measure $\tilde{\mathbb{P}}$ which is equivalent to $\mathbb{P}$ on $\bigcup_{0\leq t<\infty} \mathcal{F}_t$ and we are able to ignore the superscript t.

Then by Girsanov's Theorem (Theorem 3.5.1 of [44]), the observation process Y is a Brownian motion independent of X under $\tilde{\mathbb{P}}$ provided condition (2.13) is satisfied; and the law of the signal X under $\tilde{\mathbb{P}}$ is the same as its law under $\mathbb{P}$.

Let $\tilde{Z} = \{\tilde{Z}_t, t \geq 0\}$ be the process defined as $\tilde{Z}_t = Z_t^{-1}$ for $t \geq 0$, then under $\tilde{\mathbb{P}}$, $\tilde{Z}_t$ has the following expression:

$$\tilde{Z}_t = \exp\left(\sum_{i=1}^m \int_0^t h^i(X_s)dY_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds\right), \quad t \geq 0. \tag{2.14}$$

It is immediate that $\hat{Z}$ is a local martingale and since $\tilde{\mathbb{E}}[\tilde{Z}_t] = \mathbb{E}[\tilde{Z}_t Z_t] = 1$, $\tilde{Z}_t$ is a genuine $\mathcal{F}_t$-adapted martingale under $\tilde{\mathbb{P}}$ and

$$\left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \tilde{Z}_t \quad t \geq 0.$$

The following proposition holds because of the fact that under $\tilde{\mathbb{P}}$, the $\mathcal{Y}_t$-adapted Brownian motion Y is also a Markov process.

Proposition 2.2.2. *Let* U *be an integrable* $\mathcal{F}_t$*-measurable random variable. Then we have*

$$\tilde{\mathbb{E}}[U \mid \mathcal{Y}_t] = \tilde{\mathbb{E}}[U \mid \mathcal{Y}]. \tag{2.15}$$

Remark 2.2.3. *The importance of this proposition is that it replaces the time-dependent family of σ-algebra $\mathcal{Y}_t$ in the conditional expectation with the fixed σ-algebra $\mathcal{Y}$, enabling us to apply results from Kolmogorov conditional expectation which is applicable only if the conditioning σ-algebras are time-independent.*

Now we are able to introduce Kallianpur-Striebel formula and define the unnormalised conditional distribution process.

Proposition 2.2.4 (Kallianpur-Striebel formula). *Assume that condition* (2.13) *holds. For every $\varphi \in B(\mathbb{S})$, for fixed $t \in [0, \infty)$,*

$$\pi_t(\varphi) = \frac{\tilde{\mathbb{E}}[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}]}{\tilde{\mathbb{E}}[\tilde{Z}_t \mid \mathcal{Y}]} \qquad \tilde{\mathbb{P}}(\mathbb{P}) - a.s.. \tag{2.16}$$

Let $\zeta = \{\zeta_t, t \geq 0\}$ be the process defined by $\zeta_t = \tilde{\mathbb{E}}[Z_t \mid \mathcal{Y}]$, then we can choose a càdlàg version of ζ_t which is a $\mathcal{Y}_t$-martingale. Given such a ζ, we have the following definition:

Definition 2.2.5 (Unnormalised Conditional Distribution). *We define the unnormalised conditional distribution of X to be the measure-valued process $\rho = \{\rho_t, t \geq 0\}$ which is determined by the values of $\rho_t \varphi$ for $\varphi \in B(\mathbb{S})$ which are given for $t \geq 0$ by*

$$\rho_t(\varphi) \triangleq (\pi_t(\varphi)) \cdot \zeta_t.$$

Then from Proposition 2.2.4 and the definition of ζ, for every $t \geq 0$, we have:

$$\rho_t(\varphi) = \tilde{\mathbb{E}}[\tilde{Z}_t \varphi(X_t) \mid \mathcal{Y}] \qquad \tilde{\mathbb{P}}(\mathbb{P}) - a.s. \tag{2.17}$$

and if condition (2.13) holds, we also have:

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})} \qquad \tilde{\mathbb{P}}(\mathbb{P}) - a.s.. \tag{2.18}$$

Now we are in the position to introduce the Zakai Equation, which is satisfied by $\rho_t \varphi$; as well as the Kushner-Stratonovich Equation, which is satisfied by $\pi_t \varphi$. In the following, we further assume that for all $t \geq 0$,

$$\tilde{\mathbb{P}}\left[\int_0^t [\rho_s(\|h\|)]^2 ds < \infty\right] = 1. \tag{2.19}$$

Theorem 2.2.6 (Zakai Equation). *If conditions* (2.13) *and* (2.19) *are satisfied then the process ρ_t satisfies the following evolution equation, called the **Zakai Equation**.*

$$\rho_t(\varphi) = \pi_0(\varphi) + \int_0^t \rho_s(A\varphi)ds + \int_0^t \rho_s(\varphi h^\top)dY_s, \quad \tilde{\mathbb{P}} - a.s. \quad \forall t \geq 0 \tag{2.20}$$

for any $\varphi \in \mathcal{D}(A)$.

To derive the equation satisfied by π, we firstly give the explicit representation of the process $t \to \rho_t(\mathbf{1})$, which is

$$\rho_t(\mathbf{1}) = \exp\left(\int_0^t \pi_s(h^\top)dY_s - \frac{1}{2}\int_0^t \pi_s(h^\top)\pi_s(h)ds\right).$$

With this expression and Kallianpur-Striebel formula, use of Itô's formula leads to the following Theorem:

Theorem 2.2.7 (Kushner-Stratonovich Equation). *If conditions* (2.13) *and* (2.19) *are satisfied then the conditional distribution of the signal X satisfies the following evolution equation called the **Kushner-Stratonovich Equation.***

$$\pi_t(\varphi) = \pi_0(\varphi) + \int_0^t \pi_s(A\varphi)ds + \int_0^t (\pi_s(\varphi h^\top) - \pi_s(h^\top)\pi_s(\varphi))(dY_s - \pi_s(h)ds) \tag{2.21}$$

for any $\varphi \in \mathcal{D}(A)$.

Given the Zakai equation and the Kushner-Stratonovich equation, it is natural to investigate the uniqueness of the solutions. We would like to know under what assumptions on the coefficients of the signal and observation processes the two equations are uniquely characterised by ρ_t and π_t respectively. In fact the question of uniqueness becomes highly important when we approximate ρ and π numerically as most of the analysis of existing numerical algorithms depends on the SPDE characterisation of the two processes.

For the Zakai equation (2.20), we consider the following class of stochastic processes:

Definition 2.2.8. *The set $\mathcal{U}$ is the space of all $\mathcal{Y}_t$-adapted $\mathcal{M}^l(\mathbb{R}^d)$-valued stochastic processes $\mu = \{\mu_t,\ t \geq 0\}$ with càdlàg paths such that for all $t \geq 0$, we have*

$$\tilde{\mathbb{E}}\left[\int_0^t (\mu_s(\psi))^2 ds\right] < \infty,$$

where $\psi : \mathbb{R}^d \to \mathbb{R}$ *is the function* $\psi(x) = 1 + \|x\|$ *for any* $x \in \mathbb{R}^d$*; and* $\mathcal{M}^l(\mathbb{R}^d)$ *is the space of finite measures* μ *over* $\mathcal{B}(\mathbb{R}^d)$ *such that* $\mu(\psi) < \infty$.

Then we have the following theorem on the uniqueness of solution of Zakai equation.

Theorem 2.2.9. *If the functions* f *in* (2.9)*,* a *in* (2.11) *and* h *in* (2.6) *have twice continuously differentiable components and all their derivatives of first- and second-order are bounded, then the Zakai equation* (2.20) *has a unique solution in the class* $\mathcal{U}$*, up to indistinguishability.*

For Kushner-Stratonovich equation (2.21), let $\bar{\mathcal{U}}$ be the class of all $\mathcal{Y}_t$-adapted $\mathcal{M}^l(\mathbb{R}^d)$-valued stochastic processes $\mu = \{\mu_t,\ t \geq 0\}$ with càdlàg paths such that the process $m^\mu \mu$ belongs to the class $\mathcal{U}$, where

$$m_t^\mu = \exp\left(\int_0^t \mu_s(h^\top) dY_s - \frac{1}{2}\int_0^t \mu_s(h^\top)\mu(h) ds\right), \quad t \geq 0.$$

Theorem 2.2.10. *If functions* f *in* (2.9)*,* a *in* (2.11) *and* h *in* (2.6) *have twice continuously differentiable components and their derivatives of first- and second-order are bounded, then the Kushner-Stratonovich equation* (2.21) *has a unique solution in the class* $\bar{\mathcal{U}}$*, up to indistinguishability.*

Chapter 3

Generalised Particle Filters with Gaussian Mixtures

In this chapter, we will firstly review the classic particle filters, and then introduce a generalisation called *generalised particle filters with Gaussian mixtures*. In this case, the positions of the (generalised) particles are in (possibly) larger spaces than the state space of the signal process X.

One of the reason for introducing the generalised particle filters is that the particles involved in the classical particle filter carry information about their positions and their weights. One can interpret the system of particles as a *quantisation* of the posterior distribution π, and of the unnormalised conditional distribution ρ of the signal, respectively. This limited information may be wasteful. Indeed, it may be the case that if we allow more information to be carried by each particle then perhaps we will need a smaller number of particles. Therefore we may be able to reduce the overall computational effort.

In what follows, we discuss the general ideas the classic particle filters and the generalised particle filters constructed by Gaussian mixtures. We will, in this chapter, introduce the approximating algorithm of the Gaussian mixture approximation; the corresponding convergence analysis and central limit theorem will be discussed in the next two chapters.

3.1 Introduction to the Classic Particle Filters

Explicit solutions (ρ or π) of the filtering equations are rarely available. There are only a few exceptions where one can obtain π_t explicitly (see Chapter 6 of [3] for explicit formula for the Beneš filter and Kalman-Bucy filter). Several classes of numerical methods were therefore developed to approximate the solution of the filtering problem. These methods include the projection filter and moments methods, the spectral methods, the PDE methods, as well as the particle methods.

The projection filter is an algorithm which is used to provide an approximation of the conditional distribution of the signal. It is based on the differential geometric approach to statistics. To the author's knowledge, no general convergence theorem has been developed for this method.

The spectral approach, introduced in 1997 by Lototsky, Mikulevicius and Rozovskii in [54], for numerically estimating the conditional distribution of the signal is based on Cameron-Martin decomposition of L^2-functionals of a Gaussian process. This approach allows the computations involving the observations and the ones involving the system parameters to be separated.

The partial differential equations (PDE) method makes use of the fact that the density of the unnormalized conditional distribution of the signal is the solution of a partial differential equation (see, for example, Chapter 7 in [3]). Although this is a stochastic partial differential equation, classical partial differential equations methods can still be applied to it. These methods are very successful in low dimensions but cannot be applied to high-dimensional problems. This is because they require the use of a space grid whose size increases exponentially with the dimension of the state space of the signal.

Particle methods are one of the most effective and versatile methods for solving the filtering problem numerically. The main idea is to represent the process π_t (or ρ_t) by an approximating system of (weighted) particles whose positions and weights satisfy certain SDEs which are numerically solvable. Roughly speaking, they are algorithms which approximate the stochastic process π_t with discrete random measures (sum of Dirac measures) of the form

$$\sum_i a_i(t)\delta_{v_i(t)}$$

with stochastic masses (weights) $a_1(t), a_2(t), \ldots$, and corresponding stochastic positions $v_1(t), v_2(t), \ldots$, where $v_i(t) \in \mathbb{S}$. Kallianpur-Striebel formula is the basis of this class of numerical method because we are essentially doing Monte Carlo approximation to the unnormalised conditional distribution $\rho_t\varphi = \tilde{\mathbb{E}}[\tilde{Z}_t\varphi(X_t) \mid \mathcal{Y}]$ (defined in (2.17)). This idea will be explained further as below:

From the Kallianpur-Striebel formula we know that:

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})} \qquad \tilde{\mathbb{P}}(\mathbb{P}) - a.s.,$$

Let $\{v_j^n\}_{j=1}^n$ be n mutually independent stochastic processes which are all independent of the observation Y, and each of them is a solution of the martingale problem (A, π_0); in other words, the evolution of v_j^n is

$$v_j^n(t) = v_j^n(0) + \int_0^t f\left(v_j^n(s)\right) ds + \int_0^t \sigma\left(v_j^n(s)\right) dV_s^{(j)}, \tag{3.1}$$

where the processes $(V^{(j)})_{j=1}^n$ are mutually independent $\mathcal{F}_t$-adapted p-dimensional Brownian motions which are independent of Y and all other random variables in the system.

Also let $\{a_j^n\}_{j=1}^n$ be the following exponential martingale:

$$a_j^n(t) = 1 + \sum_{k=1}^m \int_0^t a_j^n(s)h^k(v_j^n(s))dY_s^k; \tag{3.2}$$

in other words

$$a_j^n(t) = \exp\left(\int_0^t h(v_j^n(s))^\top dY_s - \frac{1}{2}\int_0^t \|h(v_j^n(s))\|^2 ds\right). \tag{3.3}$$

Then (v_j^n, a_j^n, Y), $j = 1, \ldots, n$, are identically distributed and have the same distribution as $(X, \tilde{Z}, Y)$ under $\tilde{\mathbb{P}}$; and furthermore, the pairs $(v_j^n(t), a_j^n(t))$, $j = 1, \ldots, n$ are mutually independent conditional on the σ-algebra $\mathcal{Y}_t$.

We now have the numerical approximations $\rho^n = \{\rho_t^n; t \geq 0\}$ (and $\pi^n = \{\pi_t^n; t \geq 0\}$) of the solutions of the filtering problem ρ (and π). Define the measure-valued process ρ_t^n to be the following weighted sum of Dirac measure:

$$\rho_t^n \triangleq \frac{1}{n} \sum_{j=1}^{n} a_j^n(t) \delta_{v_j^n(t)}; \tag{3.4}$$

and its normalised version

$$\pi_t^n \triangleq \frac{\rho_t^n}{\rho_t^n(\mathbf{1})} = \sum_{j=1}^{n} \bar{a}_j^n(t) \delta_{v_j^n(t)}, \tag{3.5}$$

where the normalised weights $\bar{a}_j^n$ have the form

$$\bar{a}_j^n(t) \triangleq \frac{a_j^n(t)}{\sum_{k=1}^{n} a_k^n(t)}.$$

It can now be seen that this is a Monte Carlo approximation and the independent realisations v_j^n of the signal X can be interpreted as the trajectories of the particles. From Corollary 8.2.1 in [3] we have that ρ^n and π^n converge (both in expectation and almost surely) to ρ and π respectively.

As time increases, the unnormalised weights of the majority of the particles decrease to zero, with only few becoming very large (or equivalently, the normalised weights of the majority of the particles decrease to zero, with only few becoming close to one), this phenomenon is called the *sample degeneracy*. As a consequence, only a small number of particles contribute significantly to the approximations, and therefore a large number of particles are needed in order to obtain the required accuracy; in other words, the convergence of this approximation is very slow. In order to solve this, *particle filters* (or *sequential Monte Carlo Methods*) are employed. To be specific, a resampling (or branching) procedure is used so that it culls particles with small weights and multiplies particles with large weights. The resampling depends both on the weights of the particles and the observation data, and by doing this particles with small weights (and hence their trajectories are far from the signal) are not carried forward and therefore the more likely region where the signal might be can be explored.

3.2 The Approximation with Gaussian Mixtures

The Gaussian mixture approximation is similar to the classic particle filters, except that the approximating measures are constructed using a sum of Gaussian measures, rather than Dirac measures. In this section we will introduce the approximating algorithm involving mixtures of Gaussian measures.

For ease of notations, we assume, hereinafter from this section, that the state space of the signal $\mathbb{S} = \mathbb{R}$. The approximating algorithm discussed in this section, together with the L^2-convergence analysis in Chapter 4 and Central Limit Theorem result in Chapter 5, are all based on this assumption. We should also note that all the results hereinafter can be extended without significant technical difficulties to the multi-dimensional case where $\mathbb{S} = \mathbb{R}^d$.

Firstly, we let $\Delta = \{0 = \delta_0 < \delta_1 < \cdots < \delta_N = T\}$ be an equidistant partition of the interval $[0, T]$ with equal length, with $\delta_i = i\delta,\ i = 1, \ldots, N$; and $N = \frac{T}{\delta}$. The approximating algorithm is then introduced as follows.

Initialisation: At time zero, the particle system consists of n Gaussian measures all with equal weights $1/n$, initial means $v_j^n(0)$, and initial variances $\omega_j^n(0)$, for $j = 1, \ldots, n$; denoted by $\Gamma_{v_j^n(0), \omega_j^n(0)}$. The approximation of π_0 has the form

$$\pi_0^n \triangleq \frac{1}{n} \sum_{j=1}^{n} \Gamma_{v_j^n(0), \omega_j^n(0)}, \tag{3.6}$$

Recursion: During the interval $t \in [i\delta, (i+1)\delta)$, $i = 1, \ldots, N$, the approximation ρ^n of the unnormalised conditional distribution ρ will take the form

$$\pi_t^n \triangleq \sum_{j=1}^{n} \bar{a}_j^n(t) \Gamma_{v_j^n(t), \omega_j^n(t)}, \tag{3.7}$$

where $v_j^n(t)$ denotes the mean and $\omega_j^n(t)$ denotes the variance of the Gaussian measure $\Gamma_{v_j^n(t), \omega_j^n(t)}$, and $a_j^n(t)$ is the (unnormalised) weight of the particle, and

$$\bar{a}_j^n(t) = \frac{a_j^n(t)}{\sum_{k=1}^{n} a_k^n(t)}$$

is the normalised weight. Obviously, each particle is characterised by the triple process $(a_j^n, v_j^n, \omega_j^n)$ which is chosen to evolve as

$$\begin{cases} a_j^n(t) = 1 + \int_{i\delta}^t a_j^n(s) h(v_j^n(s)) dY_s, \\ v_j^n(t) = v_j^n(i\delta) + \int_{i\delta}^t f\left(v_j^n(s)\right) ds + \sqrt{1-\alpha} \int_{i\delta}^t \sigma\left(v_j^n(s)\right) dV_s^{(j)}, \\ \omega_j^n(t) = \alpha\left(\beta + \int_{i\delta}^t \sigma^2\left(v_j^n(s)\right) ds\right), \end{cases} \tag{3.8}$$

where $\{V^{(j)}\}_{j=1}^n$ are mutually independent Brownian motions and independent of Y. The parameter α is a real number in the interval $[0,1]$. For $\alpha = 0$ we recover the classic particle approximation (see, for example, Chapter 9 in [3]); for $\alpha = 1$ the mean of the Gaussian measures evolve deterministically (the stochastic term is eliminated). The parameter β is a positive real number, which we call the *smoothing parameter*, ensures that the approximating measure has smooth density at the branching time.

Branching/Resampling: As in the classic particle filters, we need a branching/resampling mechanism in order to minimise the effect of sample degeneracy. To be specific, at the end of the interval $[i\delta, (i+1)\delta)$, immediately prior to branching, each Gaussian measure is replaced by a random number of offsprings, which are Gaussian measures with mean $X_j^n((i+1)\delta)$ and variance $\alpha\beta$, where the mean X_j^n is a normally distributed random variable, i.e.

$$X_j^n((i+1)\delta) \sim \mathcal{N}\left(v_j^n(i+1)\delta_-, \omega_j^n(i+1)\delta_-\right), \quad j = 1, \ldots, n.$$

We denote by $o_j^{n,(i+1)\delta}$ the number of "offsprings" produced by jth generalised particle. The total number of offsprings is fixed to be n at each branching event.

After branching all the particles are re-indexed from 1 to n and all of the unnormalised weights are re-initialised back to 1; and the particles evolve following (3.8) again. The recursion is repeated N times until we reach the terminal time T, where we obtain the approximation π_T^n of π_T.

There are two commonly adopted branching methods, namely the Tree Based Branching Algorithm (TBBA) and Multinomial Branching, to determine the distribution of $\{o_j^n\}_{j=1}^n$. In what follows we discuss each of them respectively.

3.2.1 Tree Based Branching Algorithm

We set

$$o_j^{n,(i+1)\delta} = \begin{cases} \left[n\bar{a}_j^{n,(i+1)\delta}\right] & \text{with prob.} \quad 1 - \{n\bar{a}_j^{n,(i+1)\delta}\} \\ \\ \left[n\bar{a}_j^{n,(i+1)\delta}\right] + 1 & \text{with prob.} \quad \{n\bar{a}_j^{n,(i+1)\delta}\}; \end{cases} \tag{3.9}$$

where $\bar{a}_j^{n,(i+1)\delta}$ is the value of the Gaussian particle's weight immediately prior to the branching, in other words,

$$\bar{a}_j^{n,(i+1)\delta} = \bar{a}_j^n((i+1)\delta-) = \lim_{t\nearrow(i+1)\delta} \bar{a}_j^n(t).$$

If $\mathcal{F}_{(i+1)\delta-}$ is the σ-algebra of events up to time $(i+1)\delta$, i.e.

$$\mathcal{F}_{(i+1)\delta-} = \sigma(\mathcal{F}_s : s < (i+1)\delta),$$

then we have the following proposition.

Proposition 3.2.1. *The random variables* $\{o_j^n\}_{j=1}^n$ *defined in* (3.9) *have the following properties*

$$\begin{aligned} &\mathbb{E}\left[o_{j'}^{n,(i+1)\delta}|\mathcal{F}_{(i+1)\delta-}\right] = n\bar{a}_{j'}^{n,(i+1)\delta}, \\ &\mathbb{E}\left[\left(o_j^{n,(i+1)\delta} - n\bar{a}_j^{n,(i+1)\delta}\right)^2 |\mathcal{F}_{(i+1)\delta-}\right] = \left\{n\bar{a}_j^{n,(i+1)\delta}\right\}\left(1 - \left\{n\bar{a}_j^{n,(i+1)\delta}\right\}\right). \end{aligned}$$

Remark 3.2.2. *By Exercise 9.1 in [3] we know that the random variables* $o_j^{n,(i+1)\delta}$ *defined* (3.9) *have conditional minimal variance in the set of all integer-valued random variables* ξ *satisfying* $\mathbb{E}[\xi|\mathcal{F}_{(i+1)\delta-}] = n\bar{a}_j^{n,(i+1)\delta}$. *This property is important as it is the variance of* o_j^n *that influences the speed of the corresponding algorithm.*

We wish to control the branching process so that the number of particles in the system remains constant at n; that is, we require that for each i,

$$\sum_{j=1}^n o_j^{n,(i+1)\delta} = n. \tag{3.10}$$

We apply the algorithm introduced in Section 9.2.1 in [3] to ensure (3.10) is satisfied, and by Proposition 9.3 in [3] we know that the distribution of o_j^n satisfies (3.9) and Proposition 3.2.1.

3.2.2 Multinomial Branching Algorithm

Under this algorithm, the offspring distribution is determined by the multinomial distribution

$$O_{(i+1)\delta} = \text{Multinomial}(n, \bar{a}_1^n((i+1)\delta_-), \ldots, \bar{a}_n^n((i+1)\delta_-))$$

defined by

$$\mathbb{P}\left(O_{(i+1)\delta}^{(j)} = o_j^{n,(i+1)\delta}, j = 1, \ldots, n\right) = \frac{n!}{\prod_{j=1}^n o_j^{n,(i+1)\delta}!} \prod_{j=1}^n \left(\bar{a}_j^n((i+i)\delta-)\right)^{o_j^{n,(i+1)\delta}} \tag{3.11}$$

with $\sum_{j=1}^n o_j^{n,(i+1)\delta} = n$.

We then, by properties of multinomial distribution, have the following result.

Proposition 3.2.3. *At branching time $(i+1)\delta$, $\left\{O_{(i+1)\delta}^{(j)} = o_j^{n,(i+1)\delta}\right\}_{j=1}^n$ has a multinomial distribution, then the conditional mean is proportional to the normalised weights of their parents:*

$$\tilde{\mathbb{E}}\left[o_j^{n,(i+1)\delta} \middle| \mathcal{F}_{(i+1)\delta-}\right] = n\bar{a}_{j'}^{n,(i+1)\delta} \tag{3.12}$$

for $1 \leq j \leq n$; and the condition variance and covariance satisfy

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(o_l^{n,(i+1)\delta} - n\bar{a}_l^{n,(i+1)\delta}\right)\left(o_j^{n,(i+1)\delta} - n\bar{a}_j^{n,(i+1)\delta}\right) \middle| \mathcal{F}_{(i+1)\delta-}\right] \\
&= \begin{cases} n\bar{a}_j^{n,(i+1)\delta}\left(1 - \bar{a}_j^{n,(i+1)\delta}\right), & l = j \\ -n\bar{a}_l^{n,(i+1)\delta}\bar{a}_j^{n,(i+1)\delta}, & l \neq j \end{cases}
\end{aligned} \tag{3.13}$$

for $1 \leq l, j \leq n$.

This multinomial sampling algorithm essentially states that, at branching times, we sample n times (with replacement) from the population of Gaussian random variables $X_j^n((i+1)\delta)$ (with means $v_j^n((i+1)\delta_-)$ and variances $\omega_j^n((i+1)\delta_-)$), $j = 1, \ldots, n$ according to the multinomial probability distribution given by the corresponding normalised weights $\bar{a}_j^n((i+1)\delta_-), j = 1, \ldots, n$. Therefore, by definition of multinomial distribution, $o_j^{n,(i+1)\delta}$ is the number of times $X_j^n((i+1)\delta)$ is chosen at time $(i+1)\delta$; that is to say, $o_j^{n,(i+1)\delta}$ is the number of offspring produced by this Gaussian random variable.

Chapter 4

Convergence Analysis

In this chapter we deduce the evolution equation of the approximating measure ρ^n for the generalised particle filters with Gaussian mixtures, and show its convergence to the target measure ρ – the solution of the Zakai equation, as well as the convergence of π^n to π – the solution of the Kushner-Stratonovich equation. The correction mechanism for the generalised particle system involves either the use of the Tree Based Branching Algorithm (TBBA) or the multinomial sampling branching. These will be investigated in Sections 4.2 and 4.3 respectively.

4.1 Evolution Equation for ρ^n

We firstly define the process $\xi^n = \{\xi_t^n; t \geq 0\}$ by

$$\xi_t^n \triangleq \left(\prod_{i=1}^{[t/\delta]} \frac{1}{n} \sum_{j=1}^{n} a_j^{n,i\delta} \right) \left(\frac{1}{n} \sum_{j=1}^{n} a_j^n(t) \right).$$

Then ξ^n is a martingale and by Exercise 9.10 in [3] we know for any $t \geq 0$ and $p \geq 1$, there exist two constants $c_1^{t,p}$ and $c_2^{t,p}$ which depends only on $\max_{k=1,\ldots,m} \|h_k\|_{0,\infty}$, such that

$$\sup_{n\geq 0} \sup_{s\in[0,t]} \tilde{\mathbb{E}} \left[(\xi_s^n)^p \right] \leq c_1^{t,p}, \tag{4.1}$$

and

$$\max_{j=1,\ldots,n} \sup_{n\geq 0} \sup_{s\in[0,t]} \tilde{\mathbb{E}} \left[(\xi_s^n a_j^n(s))^p \right] \leq c_2^{t,p}. \tag{4.2}$$

We use the martingale ξ^n to linearise π^n in order to make it easier to analyse its convergence. Let $\rho^n = \{\rho^n_t; t \geq 0\}$ be the measure-valued process defined by

$$\rho^n_t \triangleq \xi^n_t \pi^n_t = \frac{\xi_{[t/\delta]\delta}}{n} \sum_{j=1}^{n} a^n_j(t) \Gamma_{v^n_j(t), \omega^n_j(t)}. \tag{4.3}$$

We will show the convergence of ρ^n to ρ as the number of particles n increases. In the following, we will use the norm $\|\cdot\|_{m,\infty}$ $(m \geq 0)$ defined as

$$\|\varphi\|_{m,\infty} = \sum_{|\eta| \leq m} \sup_{x \in \mathbb{R}^d} |D_\eta \varphi(x)|,$$

where $\eta = (\eta^1, \ldots, \eta^d)$ is a multi-index and $D_\eta = (\partial_1)^{\eta_1} \cdots (\partial_d)^{\eta_d}$. We also define $D_{(0)}\varphi(x) \triangleq \varphi(x)$, where $\|\varphi\|_{0,\infty} \triangleq \|\varphi\| = \sup_{x \in \mathbb{R}^d} |\varphi(x)|$.

The following proposition describes the evolution equation satisfied by the approximating sequence $\rho^n = \{\rho^n_t; t \geq 0\}$ constructed using the algorithm described in the previous chapter. As discussed in Chapter 3, the approximation algorithm is constructed for the case where the state space of the signal process X is $\mathbb{R}$. We adopt this assumption in this chapter and Chapter 5.

Proposition 4.1.1. *The measure-valued process $\rho^n = \{\rho^n_t : t \geq 0\}$ satisfies the following evolution equation:*

$$\rho^n_t(\varphi) = \rho^n_0(\varphi) + \int_0^t \rho^n_s(A\varphi)ds + \int_0^t \rho^n_s(h\varphi)dY_s + M^{n,\varphi}_{[t/\delta]} + B^{n,\varphi}_t \tag{4.4}$$

for any $\varphi \in C^m_b(\mathbb{R})$ and $t \in [0,T]$ with $m \geq 6$.

In (4.4), $M^{n,\varphi} = \{M^{n,\varphi}_i, i > 0 \text{ and } i \in \mathbb{N}\}$ is the discrete process

$$M^{n,\varphi}_{[t/\delta]} = \frac{1}{n} \sum_{i=0}^{[t/\delta]} \xi^n_{i\delta} \sum_{j=1}^{n} \left[o^{n,i\delta}_j \int_{\mathbb{R}} \varphi(x) \frac{e^{-\frac{(x - X^n_j(i\delta))^2}{2\alpha\beta}}}{\sqrt{2\pi\alpha\beta}} dx - n\bar{a}^n_j(i\delta-) \int_{\mathbb{R}} \varphi(x) \frac{e^{-\frac{(x - v^n_j(i\delta-))^2}{2\omega^n_j(i\delta-)}}}{\sqrt{2\pi\omega^n_j(i\delta-)}} dx \right] \tag{4.5}$$

where $X^n_j(i\delta) \sim N(v^n_j(i\delta-), \omega^n_j(i\delta-))$ is a Gaussian random variable.

$B^{n,\varphi}_t$ is the following process:

$$B^{n,\varphi}_t = \frac{1}{n} \sum_{j=1}^{n} \int_0^t \xi^n_{[s/\delta]\delta} a^n_j(s) \left[R^1_{s,j}(\varphi)ds + R^2_{s,j}(\varphi)dY_s + R^3_{s,j}(\varphi)dV^{(j)}_s \right];$$

the processes $R^1_{s,j}(\varphi)$, $R^2_{s,j}(\varphi)$, *and* $R^3_{s,j}(\varphi)$ *are*

$$\begin{aligned} R^1_{s,j}(\varphi) =& \omega^n_j(s)\left[\frac{1}{2}(f\varphi''')(v^n_j(s)) + \frac{\alpha}{4}(\sigma\varphi^{(4)})(v^n_j(s)) + 2\alpha\sigma^2(v^n_j(s))I^{(4)}_{4,j}(\varphi) - I_j(A\varphi)\right] \\ &+ (\omega^n_j(s))^2\left[f(v^n_j(s))I^{(5)}_{4,j}(\varphi) + \frac{\alpha\sigma^2(v^n_j(s))}{2\sqrt{\omega^n_j(s)}}I_{5,j}(\varphi) + \frac{1-\alpha}{2}\sigma^2(v^n_j(s))I^{(6)}_{4,j}(\varphi)\right], \end{aligned} \tag{4.6}$$

$$R^2_{s,j}(\varphi) = \omega^n_j(s)\left[\frac{1}{2}h(v^n_j(s))\varphi''(v^n_j(s)) - I_j(h\varphi)\right] + (\omega^n_j(s))^2 h(v^n_j(s))I^{(4)}_{4,j}(\varphi), \tag{4.7}$$

$$\begin{aligned} R^3_{s,j}(\varphi) =& \sqrt{1-\alpha}\Bigg[\sigma(v^n_j(s))\varphi'(v^n_j(s)) + \frac{1}{2}\omega^n_j(s)\sigma(v^n_j(s))\varphi'''(v^n_j(s)) \\ &+ (\omega^n_j(s))^2\sigma(v^n_j(s))I^{(5)}_{4,j}(\varphi)\Bigg]; \end{aligned} \tag{4.8}$$

and

$$\begin{aligned} I^{(k)}_{4,j}(\varphi) &= \int_{\mathbb{R}} \frac{y^4 e^{\frac{-y^2}{2}}}{\sqrt{2\pi}} \int_0^1 \varphi^{(k)}\left(v^n_j(s) + uy\sqrt{\omega^n_j(s)}\right)\frac{(1-u)^3}{6}du dy, \quad \text{for } k = 4,5,6; \\ I_{5,j}(\varphi) &= \int_{\mathbb{R}} \frac{y^5 e^{\frac{-y^2}{2}}}{\sqrt{2\pi}} \int_0^1 \varphi^{(5)}\left(v^n_j(s) + uy\sqrt{\omega^n_j(s)}\right)\frac{u(1-u)^3}{6}du dy; \\ I_j(\psi) &= \int_{\mathbb{R}} \frac{y^2 e^{\frac{-y^2}{2}}}{\sqrt{2\pi}} \int_0^1 (\psi)''\left(v^n_j(s) + uy\sqrt{\omega^n_j(s)}\right)(1-u)du dy, \quad \text{for } \psi = A\varphi, h\varphi. \end{aligned}$$

Proof. For any $\varphi \in C^m_b(\mathbb{R})$ and $t \in [i\delta, (i+1)\delta)$, we have from (4.3) that

$$\begin{aligned} \rho^n_t(\varphi) &= \frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n a^n_j(t)\int_{\mathbb{R}}\varphi(x)\Gamma_{v^n_j(t),\omega^n_j(t)}(dx) \\ &= \frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n a^n_j(t)\int_{\mathbb{R}}\varphi(x)\frac{1}{\sqrt{2\pi\omega^n_j(t)}}\exp\left(-\frac{(x-v^n_j(t))^2}{2\omega^n_j(t)}\right)dx \\ &= \frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n a^n_j(t)\int_{\mathbb{R}}\varphi\left(v^n_j(t) + y\sqrt{\omega^n_j(t)}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy, \end{aligned} \tag{4.9}$$

with similar formulas for $A\varphi$ and $h\varphi$.

We have the following Taylor expansions

$$\psi\left(v_j^n(t)+y\sqrt{\omega_j^n(t)}\right)=\sum_{k=0}^{2p-1}\frac{y^k}{k!}(\omega_j^n(t))^{\frac{k}{2}}\psi^{(k)}(v_j^n(t))$$
$$+y^{2p}\left(\omega_j^n(t)\right)^p\int_0^1\frac{1}{(2p)!}\psi^{(2p)}\left(v_j^n(t)+uy\sqrt{\omega_j^n(t)}\right)(1-u)^{2p-1}du, \tag{4.10}$$

where ψ can be φ, $A\varphi$, or $h\varphi$.

By applying (4.10) (for $p=2$ and $p=1$) to (4.9) and the similar ones for $A\varphi$ and $h\varphi$, note the fact that for any $k\geq 1$ and $k\in\mathbb{N}$,

$$\int_{\mathbb{R}}y^{2k-1}\frac{1}{\sqrt{2\pi}}\exp(-\frac{y^2}{2})dy=0,\quad \int_{\mathbb{R}}y^{2k}\frac{1}{\sqrt{2\pi}}\exp(-\frac{y^2}{2})dy=\prod_{j=1}^{k}(2j-1),$$

we obtain that

$$\rho_t^n(\varphi)=\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\left[\varphi(v_j^n(t))+\frac{1}{2}\omega_j^n(t)\varphi''(v_j^n(t))\right]+\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\left(\omega_j^n(t)\right)^2 I_{4,j}^{(4)}(\varphi); \tag{4.11}$$

$$\rho_t^n(A\varphi)=\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\left[(A\varphi)\left(v_j^n(t)\right)\right]+\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\omega_j^n(t)I_j(A\varphi); \tag{4.12}$$

$$\rho_t^n(h\varphi)=\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\left[(h\varphi)\left(v_j^n(t)\right)\right]+\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n(t)\omega_j^n(t)I_j(h\varphi). \tag{4.13}$$

Next we apply Itô's formula to equation (4.11), with the particles satisfying equations (3.8). After substituting (4.12) and (4.13), we have for $t\in[i\delta,(i+1)\delta)$

$$\rho_t^n(\varphi)=\rho_{i\delta}^n(\varphi)+\int_{i\delta}^t\rho_s^n(A\varphi)ds+\int_{i\delta}^t\rho_s^n(h\varphi)dY_s$$
$$+\int_{i\delta}^t\frac{1}{n}\sum_{j=1}^n\xi_{i\delta}^n a_j^n(s)\Big[R_{s,j}^1(\varphi)ds+R_{s,j}^2(\varphi)dY_s+R_{s,j}^3(\varphi)dV_s^{(j)}\Big]. \tag{4.14}$$

Let $\mathcal{F}_{i\delta-}=\sigma\left(\mathcal{F}_s,0\leq s<i\delta\right)$ be the σ-algebra of the events up to time $i\delta$ (the time of the i-th-branching) and $\rho_{i\delta-}^n=\lim_{t\nearrow i\delta}\rho_t^n$. For any $t\geq 0$, we have[1]

$$\rho_t^n(\varphi)=\rho_0^n(\varphi)+\sum_{i=1}^{[t/\delta]}(\rho_{i\delta}^n(\varphi)-\rho_{i\delta-}^n(\varphi))+\sum_{i=1}^{[t/\delta]}(\rho_{i\delta-}^n(\varphi)-\rho_{(i-1)\delta}^n(\varphi))+(\rho_t^n(\varphi)-\rho_{[t/\delta]\delta}^n(\varphi)), \tag{4.15}$$

[1]We use the standard convention $\sum_{k=1}^{0}=0$.

At the i-th branching event, each Gaussian measure is replaced by a random number ($o_j^{n,i\delta}$) of offsprings. Each offspring is a Gaussian measure with mean $X_j^n(i\delta)$ and variance $\alpha\beta$, where $X_j^n(i\delta) \sim \mathcal{N}(v_j^n(i\delta-), \omega_j^n(i\delta-))$. The weights of the offspring generalised particles are re-initialised to 1, i.e. $a_j^n(i\delta) = 1$; hence $\bar{a}_j^n(i\delta) = 1/n$. So

$$\pi_{i\delta}^n = \frac{1}{n}\sum_{j=1}^n o_j^{n,k\delta}\Gamma_{X_j^n(k\delta),\alpha\beta}, \qquad \textit{and} \qquad \pi_{i\delta}^n(\varphi) = \frac{1}{n}\sum_{j=1}^n o_j^{n,k\delta}\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-X_j^n(i\delta))^2}{2\alpha\beta}}}{\sqrt{2\pi\alpha\beta}}dx;$$

Before the branching event, we still have generalised particles, thus

$$\pi_{i\delta-}^n(\varphi) = \sum_{j=1}^n \bar{a}_j^n(i\delta-)\int_{\mathbb{R}}\varphi\left(v_j^n(i\delta-) + y\sqrt{\omega_j^n(i\delta-)}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy.$$

We then obtain

$$\begin{aligned}
M_{t/\delta}^{n,\varphi} &\triangleq \sum_{i=0}^{[t/\delta]}\left(\rho_{i\delta}^n(\varphi) - \rho_{i\delta-}^n(\varphi)\right) = \sum_{i=0}^{[t/\delta]}\xi_{i\delta}^n\left(\pi_{i\delta}^n(\varphi) - \pi_{i\delta-}^n(\varphi)\right) \\
&= \frac{1}{n}\sum_{i=0}^{[t/\delta]}\xi_{i\delta}^n\sum_{j=1}^n\left[o_j^{n,i\delta}\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-X_j^n(i\delta))^2}{2\alpha\beta}}}{\sqrt{2\pi\alpha\beta}}dx - n\bar{a}_j^n(i\delta-)\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-v_j^n(i\delta-))^2}{2\omega_j^n(i\delta-)}}}{\sqrt{2\pi\omega_j^n(i\delta-)}}dx\right].
\end{aligned} \tag{4.16}$$

For $t \in [(i-1)\delta, i\delta)$, for $i = 1, 2, \ldots, [t/\delta]$, we have

$$\rho_t^n(\varphi) - \rho_{(i-1)\delta}^n(\varphi) = \int_{(i-1)\delta}^t d\rho_s^n(\varphi),$$

Similarly, let $t \nearrow i\delta-$, we have

$$\rho_{i\delta-}^n(\varphi) - \rho_{(i-1)\delta}^n(\varphi) = \int_{(i-1)\delta}^{i\delta} d\rho_s^n(\varphi).$$

Then by (4.14), we obtain that

$$\begin{aligned}
&\sum_{i=1}^{[t/\delta]}(\rho_{i\delta-}^n(\varphi) - \rho_{(i-1)\delta}^n(\varphi)) + (\rho_t^n(\varphi) - \rho_{i\delta}^n(\varphi)) = \sum_{i=1}^{[t/\delta]}\int_{(i-1)\delta}^{i\delta} d\rho_s^n(\varphi) + \int_{[t/\delta]\delta}^t d\rho_s^n(\varphi) \\
&= \int_0^t d\rho_s^n(\varphi) = \rho_0^n(\varphi) + \int_0^t \rho_s^n(A\varphi)ds + \int_0^t \rho_s^n(h\varphi)dY_s \\
&\qquad + \int_0^t \frac{1}{n}\sum_{j=1}^n \xi_{[s/\delta]\delta}^n a_j^n(s)\left[R_{s,j}^1(\varphi)ds + R_{s,j}^2(\varphi)dY_s + R_{s,j}^3(\varphi)dV_s^{(j)}\right],
\end{aligned} \tag{4.17}$$

Finally, (4.16) and (4.17) imply (4.4), which completes the proof. ■

Corollary 4.1.2. *Under the same assumption as in Proposition 4.1.1, if we further assume that* $\alpha = 0$ *in* (3.8)*, which implies the classic particle filters with mixture of Dirac measures, then the approximating measure* ρ^n *satisfies the following evolution equation:*

$$\rho_t^n(\varphi) = \rho_0^n(\varphi) + \int_0^t \rho_s^n(A\varphi)ds + \int_0^t \rho_s^n(h\varphi)dY_s + M_{[t/\delta]}^{n,\varphi} + B_t^{n,\varphi} \tag{4.18}$$

for any $\varphi \in C_b^1(\mathbb{R})$*, where*

$$M_{[t/\delta]}^{n,\varphi} = \frac{1}{n}\sum_{i=0}^{[t/\delta]} \xi_{i\delta}^n \sum_{j=1}^{n} \varphi(v_j^n(i\delta_-))\left[o_j^{n,i\delta} - n\bar{a}_j^n(i\delta_-)\right];$$

$$B_t^{n,\varphi} = \frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi_{[s/\delta]\delta}^n a_j^n(s)\sigma(v_j^n(s))\varphi'(v_j^n(s))dV_s^{(j)}.$$

4.2 Convergence Results for Generalised Particle Filters using the TBBA

In order to investigate the convergence of the approximating measure ρ^n, we consider the mild form of the Zakai equation. One should note that the proof of the convergence in [3] using the dual, $\psi_s^{t,\varphi}$, of the measure-valued process ρ does not work for our model. $\psi_s^{t,\varphi}$ is measurable with respect to the backward filtration $\mathcal{Y}_s^t = \sigma(Y_t - Y_r,\ r \in [s,t])$, and so is $R_{s,j}^2(\psi_s^{t,\varphi})$; however, the Itô's integral $\int_0^t R_{s,j}^2(\psi_s^{t,\varphi})dY_s$ requires $R_{s,j}^2(\psi_s^{t,\varphi})$ is measurable with respect to the forward filtration $\mathcal{Y}_s = \sigma(Y_r,\ r \in [0,s])$. This leads to an anticipative integration which cannot be tracked in a standard manner. Another approach is therefore required. Markov semigroups was used in [56] to obtain relevant bounds on the error which in turn enables us to discuss the convergence rate. In the following this idea will be discussed in some details.

We introduce first the Zakai equation for time-inhomogeneous test functions. Let $\tilde{\varphi} : [0,\infty) \times \mathbb{S} \to \mathbb{R}$ be a bounded measurable function with continuous bounded derivatives up to order m ($m \geq 6$). Then for any $\tilde{\varphi} \in \mathcal{D}(A)$, the time-inhomogeneous Zakai equation is (see, for example, Chapter 3 in [3])

$$d\rho_t(\tilde{\varphi}) = \rho_t\left(\frac{\partial\tilde{\varphi}}{\partial t} + A\tilde{\varphi}\right)dt + \rho_t(h\tilde{\varphi})dY_t. \tag{4.19}$$

Now fix $s > 0$ and define

$$\tilde{\varphi}(t,x) = P_{s-t}\varphi(x), \qquad t \in [0,s]$$

where $(P_r)_{r\geq 0}$ is the Markov semigroup whose infinitesimal generator is the operator A and φ is the single variable function which does not depend on t. It follows by the properties of semigroup (see, for example, [63]) that

$$\frac{\partial \tilde{\varphi}}{\partial t} = -AP_{s-t}\varphi,$$

therefore (4.19) becomes

$$d\rho_t(P_{s-t}\varphi) = \rho_t(hP_{s-t}\varphi)dY_t,$$

and the integration form is

$$\rho_t(P_{s-t}\varphi) = \rho_0(P_s\varphi) + \int_0^t \rho_r(hP_{s-r}\varphi)dY_r.$$

Similarly for $\rho_t^n(\varphi)$ we rewrite (4.4) for $t \in [0,s]$ and get

$$\rho_t^n(P_{s-t}\varphi) = \rho_0^n(P_s\varphi) + \int_0^t \rho_r^n(hP_{s-r}\varphi)dY_r + M_{[t/\delta]}^{n,P\varphi} + B_t^{n,P\varphi} \tag{4.20}$$

and the error of the approximation has the representation

$$(\rho_t^n - \rho_t)(P_{s-t}\varphi) = (\rho_0^n - \rho_0)(P_s\varphi) + \int_0^t (\rho_r^n - \rho_r)(hP_{s-r}\varphi)dY_r + M_{[t/\delta]}^{n,P\varphi} + B_t^{n,P\varphi}, \tag{4.21}$$

where $M_i^{n,\varphi}$ and $B_t^{n,\varphi}$ are the same as in Proposition 4.1.1, except that φ replaced by $P_{s-r}\varphi$.

In order to prove the convergence of the approximating measures ρ_t^n to the actual measure ρ_t, we need to control all the terms on the right hand side of (4.21). Now we will discuss each of them respectively in the following Lemmas.

Assumption (A). *We assume that the coefficients σ, f, and h are bounded and Lipschitz, σ and f are six times differentiable, and h is twice differentiable.*

Lemma 4.2.1. *If Assumption (A) is satisfied. Then for any $T > 0$ there exists a constant c_1^T independent of n such that for any $p \geq 1$ and $\varphi \in C_b(\mathbb{R})$, we have*

$$\tilde{\mathbb{E}}\left[(\rho_0^n(P_s\varphi) - \rho_0(P_s\varphi))^p\right] \leq \frac{c_1^T}{n^{p/2}}\|\varphi\|^p, \qquad t \in [0,T]$$

Proof. Note that $\rho_0^n(P_s\varphi) - \rho_0(P_s\varphi) = \pi_0^n(P_s\varphi) - \pi_0(P_s\varphi)$, and also note that

$$\xi_j \triangleq \frac{1}{n}\sum_{j=1}^{n}(P_s\varphi(v_j^n(0)) - \pi_0(P_s\varphi) = \pi_0^n(P_s\varphi) - \pi_0(P_s\varphi), \quad j = 1, \ldots, n$$

are independent identically distributed random variables with mean 0, therefore

$$\begin{aligned}
\tilde{\mathbb{E}}\left[(\rho_0^n(P_s\varphi) - \rho_0(P_s\varphi))^{2p'}\right] &= \tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\xi_j\right)^{2p'}\right] \\
&= \frac{1}{n^{2p'}}\tilde{\mathbb{E}}\left[\sum_{j_1,\ldots,j_n}\begin{pmatrix} 2p' \\ j_1, \ldots, j_n\end{pmatrix}[\xi_1^{j_1}\cdots\xi_n^{j_n}]\right] \\
&= \frac{1}{n^{2p'}}\sum_{j_1,\ldots,j_n}\begin{pmatrix} 2p \\ j_1, \ldots, j_n\end{pmatrix}\tilde{\mathbb{E}}[|\xi_1|^{j_1}\cdots|\xi_n|^{j_n}] \\
&\leq \frac{C_{p'}}{n^{p'}}\|\varphi\|^{2p'};
\end{aligned}$$

then the result follows by setting $p = 2p'$ and $c_1^T = C_{p'}$. ■

Lemma 4.2.2. *If Assumption (A) is satisfied. Then for any $T > 0$, there exists a constant c_2^T independent of n such that for any $\varphi \in C_b^6(\mathbb{R})$,*

$$\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi_{[r/\delta]\delta}^n a_j^n(r) R_{r,j}^1(P_{s-r}\varphi)dr\right)^2\right] \leq c_3^T(\alpha\delta)^2\|\varphi\|_{6,\infty}^2.$$

Proof. From the facts that f and σ are bounded, for $\alpha\delta \leq 1$ we have

$$|R_{r,j}^1(P_{s-r}\varphi)| \leq \left(C_1\alpha\delta + C_2(\alpha\delta)^2\right)\|\varphi\|_{6,\infty} \leq C\alpha\delta\|\varphi\|_{6,\infty},$$

Then by Jensen's inequality, Fubini's theorem and (4.2), we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi_{[r/\delta]\delta}^n a_j^n(r) R_{r,j}^1(P_{s-r}\varphi)dr\right)^2\right] \\
\leq&\frac{1}{n}\sum_{j=1}^{n} t\tilde{\mathbb{E}}\left[\left(\int_0^t \xi_{[r/\delta]\delta}^n a_j^n(r) R_{r,j}^1(P_{s-r}\varphi)dr\right)^2\right] \\
\leq&\frac{1}{n}\sum_{j=1}^{n} tC^2(\alpha\delta)^2\|\varphi\|_{6,\infty}^2\int_0^t\tilde{\mathbb{E}}\left[\left(\xi_{[r/\delta]\delta}^n a_j^n(r)\right)^2\right]dr \\
\leq&\frac{1}{n}\sum_{j=1}^{n} tC^2(\alpha\delta)^2\|\varphi\|_{6,\infty}^2 c_2^{t,2}t \\
=&T^2C^2c_2^{T,2}(\alpha\delta)^2\|\varphi\|_{6,\infty}^2.
\end{aligned}$$

The result follows by letting $c_3^T = T^2C^2c_2^{T,2}$. ■

Lemma 4.2.3. *If Assumption (A) is satisfied. Then for any $T > 0$, there exists a constant c_2^T independent of n such that for any $\varphi \in C_b^4(\mathbb{R})$,*

$$\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)R^2_{r,j}(P_{s-r}\varphi)dY_r\right)^2\right] \leq c_4^T(\alpha\delta)^2\|\varphi\|^2_{4,\infty}.$$

Proof. From the facts that f and σ are bounded, we have for $\alpha\delta \leq 1$

$$R^2_{r,j}(P_{s-r}\varphi) \leq \left(C_1\alpha\delta + C_2(\alpha\delta)^2\right)\|\varphi\|_{4,\infty} \leq C\alpha\delta\|\varphi\|_{4,\infty}.$$

Then Burkholder-Davis-Gundy and Jensen's inequalities, Fubini's theorem, and (4.2) yield

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)R^2_{r,j}(P_{s-r}\varphi)dY_r\right)^2\right]\\
\leq&\frac{1}{n^2}C^2(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\tilde{\mathbb{E}}\left[\left(\int_0^t\sum_{j=1}^{n}\xi^n_{[r/\delta]\delta}a^n_j(r)dY_r\right)^2\right]\\
\leq&\frac{1}{n^2}C^2(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\tilde{C}\tilde{\mathbb{E}}\left[\left\langle\int_0^\cdot\sum_{j=1}^{n}\xi^n_{[r/\delta]\delta}a^n_j(r)dY_r\right\rangle_t\right]\\
=&\frac{1}{n^2}C^2(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\tilde{C}\int_0^t\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\xi^n_{[r/\delta]\delta}a^n_j(r)\right)^2\right]dr\\
=&\frac{1}{n^2}C^2\tilde{C}(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\int_0^t\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\tilde{\mathbb{E}}\Big(\xi^n_{[r/\delta]\delta}a^n_j(r)\xi^n_{[r/\delta]\delta}a^n_k(r)\Big)\right]dr\\
\leq&\frac{1}{n^2}C^2\tilde{C}(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\int_0^t\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\sqrt{\tilde{\mathbb{E}}\Big[(\xi^n_{[r/\delta]\delta}a^n_j(r))^2\Big]\tilde{\mathbb{E}}\Big[(\xi^n_{[r/\delta]\delta}a^n_k(r))^2\Big]}\right]dr\\
\leq&\frac{1}{n^2}C^2\tilde{C}(\alpha\delta)^2\|\varphi\|^2_{4,\infty}\int_0^t n^2c_2^{t,2}dr\\
\leq&TC^2\tilde{C}c_2^{t,2}(\alpha\delta)^2\|\varphi\|^2_{4,\infty},
\end{aligned}$$

and the result follows by letting $c_4^T = TC^2\tilde{C}c_2^{t,2}$. ■

Lemma 4.2.4. *If Assumption (A) is satisfied. Then for any $T > 0$, there exists a constant c_2^T independent of n such that for any $\varphi \in C_b^5(\mathbb{R})$,*

$$\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)R^3_{r,j}(P_{s-r}\varphi)dV_r^{(j)}\right)^2\right] \leq \frac{c_5^T}{n}\|\varphi\|^2_{5,\infty}.$$

Proof. By the facts that f and σ are bounded, we have for $\alpha\delta \leq 1$

$$R^3_{r,j}(P_{s-r}\varphi) \leq \left(C_0 + C_1\alpha\delta + C_2(\alpha\delta)^2\right)\|\varphi\|_{5,\infty} \leq (C_0 + C\alpha\delta)\|\varphi\|_{5,\infty}.$$

Then by Burkholder-Davis-Gundy and Jensen's inequalities, Fubini's theorem, and (4.2), and noticing the fact that $\{V^{(j)}\}_{j=1}^n$ are mutually independent Brownian motions, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^n\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)R^3_{r,j}(P_{s-r}\varphi)dV^{(j)}_r\right)^2\right]\\
=&\frac{1}{n^2}\sum_{j=1}^n\tilde{\mathbb{E}}\left[\left(\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)R^3_{r,j}(P_{s-r}\varphi)dV^{(j)}_r\right)^2\right]\\
\leq&\frac{1}{n^2}(C_0+C\alpha\delta)^2\|\varphi\|^2_{5,\infty}\sum_{j=1}^n\tilde{\mathbb{E}}\left[\left(\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)dV^{(j)}_r\right)^2\right]\\
\leq&\frac{1}{n^2}(C_0+C\alpha\delta)^2\|\varphi\|^2_{5,\infty}\sum_{j=1}^n\tilde{C}\tilde{\mathbb{E}}\left[\left\langle\int_0^t \xi^n_{[r/\delta]\delta}a^n_j(r)dV^{(j)}_r\right\rangle_t\right]\\
=&\frac{1}{n^2}\tilde{C}(C_0+C\alpha\delta)^2\|\varphi\|^2_{5,\infty}\sum_{j=1}^n\int_0^t\tilde{\mathbb{E}}\left[(\xi^n_{i\delta}a^n_j(r))^2\right]dr\\
\leq&\frac{1}{n^2}\tilde{C}(C_0+C\alpha\delta)^2\|\varphi\|^2_{5,\infty}nc^{t,2}_2 t\\
\leq&\frac{1}{n}T\tilde{C}(C_0+C\alpha\delta)^2c^{t,2}_2\|\varphi\|^2_{5,\infty},
\end{aligned}$$

and the result follows by letting $c^T_5 = T\tilde{C}(C_0+C\alpha\delta)^2c^{t,2}_2$. ∎

Recalling Proposition 4.1.1 and the semigroup operator P, we can decompose $M^{n,\varphi}$ in the following way

$$\begin{aligned}
M^{n,\varphi}_{[t/\delta]} =&\frac{1}{n}\sum_{i=0}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^n\left[o^{n,i\delta}_j\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-X^n_j(i\delta))^2}{2\alpha\beta}}}{\sqrt{2\pi\alpha\beta}}dx - n\bar{a}^n_j(i\delta-)\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-v^n_j(i\delta-))^2}{2\omega^n_j(i\delta-)}}}{\sqrt{2\pi\omega^n_j(i\delta-)}}dx\right]\\
\triangleq& A^{n,\varphi}_{[t/\delta]} + D^{n,\varphi}_{[t/\delta]} + G^{n,\varphi}_{[t/\delta]},
\end{aligned}$$

where $X^n_j(i\delta) \sim \mathcal{N}\left(v^n_j(i\delta-), \omega^n_j(i\delta-)\right)$ is a Gaussian distributed random variable, and

$$A^{n,\varphi}_{[t/\delta]} = \frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^{n}\left[\left(o^{n,i\delta}_j - n\bar{a}^n_j(i\delta-)\right)\varphi(X^n_j(i\delta))\right]; \tag{4.22}$$

$$D^{n,\varphi}_{[t/\delta]} = \frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^{n}\left[o^{n,i\delta}_j\left(\int_{\mathbb{R}}\varphi\left(X^n_j(i\delta)+y\sqrt{\alpha\beta}\right)\frac{e^{\frac{-y^2}{2}}}{\sqrt{2\pi}}dy - \varphi(X^n_j(i\delta))\right)\right]; \tag{4.23}$$

$$\begin{aligned} G^{n,\varphi}_{[t/\delta]} &= \frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^{n}n\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(i\delta)) - \int_{\mathbb{R}}\varphi\left(v^n_j(i\delta-)+y\sqrt{\omega^n_j(i\delta-)}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy\right] \\ &= \sum_{i=1}^{[t/\delta]}\sum_{j=1}^{n}\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(k\delta))\right)\right]. \end{aligned} \tag{4.24}$$

Then we have the following lemma:

Lemma 4.2.5. *If Assumption (A) is satisfied. Then for any $T > 0$, and for any $\varphi \in C^1_b(\mathbb{R})$, we have the following bounds for $A^{n,\varphi}_{[t/\delta]}$, $D^{n,\varphi}_{[t/\delta]}$ and $B^{n,\varphi}_{[t/\delta]}$:*

$$\begin{aligned} \tilde{\mathbb{E}}\left[|A^{n,\varphi}_{[t/\delta]}|^2\right] &\le \frac{c^T_6}{n\sqrt{\delta}}\|\varphi\|^2_{0,\infty}, \\ \tilde{\mathbb{E}}\left[|D^{n,\varphi}_{[t/\delta]}|^2\right] &\le \tilde{c}^T_6(\alpha\beta)^2\|\varphi\|^2_{2,\infty}, \\ \tilde{\mathbb{E}}\left[|G^{n,\varphi}_{[t/\delta]}|^2\right] &\le \frac{c^T_7\alpha}{n}\|\varphi\|^2_{1,\infty}; \end{aligned} \tag{4.25}$$

where c^T_6, $\tilde{c}^T_6$ and c^T_7 are constants independent of n.

Proof. Without loss of generality, we choose the test function φ to be non-negative (since we can write $\varphi = \varphi^+ - \varphi^-$) and the random variables $\{o^{n,k\delta}_{j'},\ j' = 1,\dots,n\}$

are negatively correlated (see Proposition 9.3 in [3]), it follows that

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\xi_{i\delta}^{n}\left(\left(o_j^{n,i\delta}-n\bar{a}_j^n(i\delta-)\right)\varphi(X_j^n(i\delta))\right)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi_{i\delta}^n)^2}{n^2}\;\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\left(\left(o_j^{n,i\delta}-n\bar{a}_j^n(i\delta-)\right)\varphi(X_j^n(i\delta))\right)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
\le&\frac{(\xi_{i\delta}^n)^2}{n^2}\|\varphi\|_{0,\infty}^2\sum_{j=1}^{n}\tilde{\mathbb{E}}\left[\left(o_j^{n,i\delta}-n\bar{a}_j^n(i\delta-)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi_{i\delta}^n)^2}{n^2}\|\varphi\|_{0,\infty}^2\sum_{j=1}^{n}\left\{n\bar{a}_j^n(i\delta-)\right\}\left\{1-n\bar{a}_j^n(i\delta-)\right\}\\
\le&\frac{(\xi_{i\delta}^n)^2}{n^2}\|\varphi\|_{0,\infty}^2\sum_{j=1}^{n}\left|1-n\bar{a}_j^n(i\delta-)\right|\\
\le&\frac{(\xi_{i\delta}^n)^2}{n^2}\|\varphi\|_{0,\infty}^2 nC_\delta\sqrt{\delta}=C_\delta\sqrt{\delta}\frac{(\xi_{i\delta}^n)^2}{n}\|\varphi\|_{0,\infty}^2;
\end{aligned}
$$

By taking the expectation on both sides, we have

$$
\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\xi_{i\delta}^{n}\left(\left(o_j^{n,i\delta}-n\bar{a}_j^n(i\delta-)\right)\varphi(X_j^n(i\delta))\right)\right)^2\right]\le\frac{C_\delta c_1^{t,2}}{n}\sqrt{\delta}\|\varphi\|_{0,\infty}^2.
$$

Therefore

$$
\tilde{\mathbb{E}}\left[\left(A_{[t/\delta]}^{n,\varphi}\right)^2\right]\le\sum_{i=1}^{[t/\delta]}\frac{C_\delta c_1^{t,2}}{n}\sqrt{\delta}\|\varphi\|_{0,\infty}^2\le\frac{[t/\delta]C_\delta c_1^{t,2}}{n}\sqrt{\delta}\|\varphi\|_{0,\infty}^2\le\frac{tC_\delta c_1^{t,2}}{\sqrt{\delta}n}\|\varphi\|_{0,\infty}^2. \tag{4.26}
$$

For $G_{[t/\delta]}^{n,\varphi}$, first by noting that

$$
\int_{\mathbb{R}}\varphi\left(X_j^n(i\delta)+y\sqrt{\alpha\beta}\right)\frac{e^{\frac{-y^2}{2}}}{\sqrt{2\pi}}dy-\varphi(X_j^n(i\delta))=\quad\frac{\alpha\beta}{2}\varphi''\left(X_j^n(i\delta)\right)+\mathcal{O}\left((\alpha\beta)^2\right); \tag{4.27}
$$

then it is clear that we only need to show

$$
\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi_{i\delta}^n\sum_{j=1}^{n}o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X_j^n(i\delta)\right)\right)\right)^2\right]\le\tilde{c}_6^T(\alpha\beta)^2\|\varphi\|_{0,\infty}^2. \tag{4.28}
$$

Observe that

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X^n_j(i\delta)\right)\right)\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n o_j^{n,i\delta}\varphi''\left(X^n_j(i\delta)\right)\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
\leq&\frac{(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\|\varphi\|^2_{2,\infty}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)+n\bar{a}^n_j(i\delta-)\right]\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
\leq&\frac{2(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\|\varphi\|^2_{2,\infty}\left[\sum_{j=1}^n\tilde{\mathbb{E}}\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)^2\middle|\mathcal{F}_{i\delta-}\right]+\left(\sum_{j=1}^n n\bar{a}^n_j(i\delta-)\right)^2\right].
\end{aligned}
\tag{4.29}
$$

For Tree Based Branching Algorithm (TBBA), by Proposition 3.2.1,

$$
\tilde{\mathbb{E}}\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)^2\middle|\mathcal{F}_{i\delta-}\right]\leq\left\{n\bar{a}^n_j(i\delta-)\right\}\left(1-\left\{n\bar{a}^n_j(i\delta-)\right\}\right)\leq\frac{1}{4};\tag{4.30}
$$

and then by taking expectation on both sides of (4.57), we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X^n_j(i\delta)\right)\right)\right)^2\right]\\
\leq&\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}\frac{n}{4}+\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}n^2\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^2\left(\sum_{j=1}^n\bar{a}^n_j(i\delta-)\right)^2\right]\\
\leq&\frac{\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{2n}+2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2\sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^4\right]\sum_{l=1}^n\sum_{j=1}^n\sqrt{\tilde{\mathbb{E}}\left[\bar{a}^n_l(i\delta-)^2\right]\tilde{\mathbb{E}}\left[\bar{a}^n_j(i\delta-)^2\right]}}\\
\leq&\frac{\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{2n}+2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2\sqrt{c_1^{t,4}e^{c_2t}}\\
\leq&c^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty},
\end{aligned}
\tag{4.31}
$$

where $c^T=\frac{1}{2}+2\sqrt{c_1^{t,4}e^{c_2t}}$.

Therefore, we obtain

$$
\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X^n_j(i\delta)\right)\right)\right)^2\right]\leq[t/\delta]\sum_{i=1}^{[t/\delta]}c^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty}=\tilde{c}_6^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty}
$$

if we let $\tilde{c}_6^T = T^2 c^T/\delta^2$.

As for $G^{n,\varphi}_{[t/\delta]}$, first note that $X^n_j(i\delta) \sim N\left(v^n_j(i\delta), \omega^n_j(i\delta)\right)$ and X^n_js are mutually independent ($j = 1, \dots, n$), then we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n \xi^n_{i\delta}\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(k\delta))\right)\right]\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
=&\sum_{j=1}^n \left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^2 \tilde{\mathbb{E}}\left[\left(\varphi(X^n_j(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(k\delta))\right)\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq&\sum_{j=1}^n \left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^2 4\|\varphi'\|^2_{0,\infty}\tilde{\mathbb{E}}\left[\left(X^n_j(k\delta) - \tilde{\mathbb{E}}\left(X^n_j(k\delta)\right)\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq&4\sum_{j=1}^n \left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^2 \|\varphi\|^2_{1,\infty}\|\sigma\|^2_{0,\infty}\alpha\delta \\
\triangleq&c_\sigma\alpha\delta\|\varphi\|^2_{1,\infty}\sum_{j=1}^n \left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^2.
\end{aligned}
$$

We know from the proof of Lemma 4.4 in [56] that for any $p > 0$,

$$
\tilde{\mathbb{E}}\left[\left(\bar{a}^n_j(t)\right)^p\right] \leq \frac{1}{n^p}\exp(c_p t);
$$

then by taking the expectation on both sides, we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n \xi^n_{i\delta}\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(k\delta))\right)\right]\right)^2\right] \\
\leq&c_\sigma\alpha\delta\|\varphi\|^2_{1,\infty}\sum_{j=1}^n \tilde{\mathbb{E}}\left[\left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^2\right] \leq c_\sigma\alpha\delta\|\varphi\|^2_{1,\infty}\sum_{j=1}^n \sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^4\right]\tilde{\mathbb{E}}\left[\left(\bar{a}^n_j(i\delta-)\right)^4\right]} \\
\leq&c_\sigma\alpha\delta\|\varphi\|^2_{1,\infty}\frac{1}{n^2}\sum_{j=1}^n \sqrt{c_1^{t,4}\exp(c_4 t)} = \frac{c_\sigma\sqrt{c_1^{t,4}\exp(c_4 t)}}{n}\alpha\delta\|\varphi\|^2_{1,\infty}.
\end{aligned}
$$

Finally we have

$$
\tilde{\mathbb{E}}\left[\left(G^{n,\varphi}_{[t/\delta]}\right)^2\right] \leq \sum_{i=1}^{[t/\delta]} \frac{c_\sigma\sqrt{c_1^{t,4}\exp(c_4 t)}}{n}\alpha\delta\|\varphi\|^2_{1,\infty} \leq \frac{tc_\sigma\sqrt{c_1^{t,4}\exp(c_4 t)}}{n}\alpha\|\varphi\|^2_{1,\infty}. \tag{4.32}
$$

The result follows by letting $c_6^T = TC_\delta c_1^{T,2}$ and $c_7^T = Tc_\sigma\sqrt{c_1^{T,4}\exp(c_4 T)}$. ■

The following Theorem, which is a variation of Theorem 4.10 in [56], establishes the convergence of finite *signed* measure valued processes and allows us to use the bounds obtained from the above Lemmas to get the convergence results of ρ_t^n.

Theorem 4.2.6. *Let $\mu^n = \{\mu_t^n : t \geq 0\}$ be a signed measure-valued process such that for any $\varphi \in C_b^m(\mathbb{R}^d)$, $m \geq 6$, any fixed $\alpha \geq 1$ and fixed $s > t$, we have*

$$\mu_t^n(\varphi) = \mu_0^n(a_s(\varphi)) + \sum_{l=1}^{\alpha} R_{t,l}^{n,\varphi} + \sum_{k=1}^{\beta} \int_0^t \mu_r^n(a_{s,r}^k(\varphi))dW_r^k, \tag{4.33}$$

where $W = (W^k)_{k=1}^{\beta}$ is an β-dimensional Brownian motion, and $a_s,\ a_{s,r}^k : C_b^m(\mathbb{R}^d) \to C_b^m(\mathbb{R}^d)$ are bounded linear operators with bounds c and C_k ($k = 1, \dots, \beta$) respectively, i.e., $\|a_s(\varphi)\|_{m,\infty} \leq c\|\varphi\|_{m,\infty}$ and $\|a_{s,r}^k(\varphi)\|_{m,\infty} \leq C_k\|\varphi\|_{m,\infty}$. If for any $T > 0$ there exist constants $\gamma_0, \gamma_1, \dots, \gamma_\alpha$ such that for $t \in [0, T]$, $p \geq 2$ and $q_l > 0$ ($l = 0, 1, \dots, \alpha$),

$$\tilde{\mathbb{E}}\left[|\mu_0^n(a_s(\varphi))|^p\right] \leq \frac{\gamma_0}{n^{q_0}}\|\varphi\|_{m,\infty}^p, \qquad \tilde{\mathbb{E}}\left[|R_{t,l}^{n,\varphi}|^p\right] \leq \frac{\gamma_l}{n^{q_l}}\|\varphi\|_{m,\infty}^p,\ l = 1, \dots, \alpha. \tag{4.34}$$

Then for any $t \in [0, T]$, we have

$$\tilde{\mathbb{E}}\left[|\mu_t^n(\varphi)|^p\right] \leq \frac{c_t}{n^q}\|\varphi\|_{m,\infty}^p, \tag{4.35}$$

where c_t is a constant independent of n and $q = \min(q_0, q_1, \dots, q_\alpha)$.

Proof. See Appendix B.3. ■

Applying the bounds in Lemmas 4.2.1 to 4.2.5, one obtains the rate of convergence of the approximation in terms of the three parameters n, δ and α. In what follows we will assume, without loss of generality that $\delta < 1$. The following theorem can then be viewed as a direct corollary of Theorem 4.2.6.

Proposition 4.2.7. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists a constant c_0^T independent of n, δ or α such that for any $\varphi \in C_b^m(\mathbb{R})$, we have for $t \in [0, T]$*

$$\tilde{\mathbb{E}}\left[(\rho_t^n(\varphi) - \rho_t(\varphi))^2\right] \leq c_0^T\|\varphi\|_{m,\infty}^2 c(n, \delta, \alpha, \beta),$$

where

$$c(n, \delta, \alpha, \beta) = \max\left\{\frac{1}{n},\ (\alpha\delta)^2,\ \frac{1}{n\sqrt{\delta}},\ (\alpha\beta)^2,\ \frac{\alpha}{n}\right\}.$$

In what follows, we will discuss $c(n, \delta, \alpha, \beta)$ to obtain the L^2-convergence rate of the approximation process ρ_t^n.

When $\alpha = 0$ in (3.8), the component Gaussian measures have null covariance matrices, in other words they are Dirac measures. In this case ρ_n is nothing other than the classic particle filter (see, for example, [3]). In this case several terms in $c(n, \delta, \alpha)$ coming from the covariance term disappear. The rate of convergence $c(n, \delta, 0)$ becomes:

$$c(n, \delta, 0) = \max \left\{ \frac{1}{n}, \ \frac{1}{n\sqrt{\delta}} \right\}.$$

Obviously the fastest rate is obtained when δ is a fixed constant independent of n. The L^2-convergence rate will be in this case of order $1/n$, which coincides with the results in [3].

When $\alpha \in (0, 1]$, the rate of convergence can deteriorate. First of all let us observe that we still need to choose δ to be a fixed constant independent of n. Then the convergence depends on the simpler coefficient $c(n, \alpha)$ given by

$$c(n, \alpha, \beta) = \max \left\{ \frac{1}{n}, \ \alpha^2, \ (\alpha\beta)^2, \ \frac{\alpha}{n} \right\}$$

In this case we need to choose $\alpha = \frac{1}{\sqrt{n}}$ (or of order $1/\sqrt{n}$) and β to be a fixed constant independent of n to ensure the optimal rate of convergence, which equals to $1/n$. This discussion therefore leads to the following convergence theorem:

Theorem 4.2.8. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists constant c_1^T independent of n, such that for any $\varphi \in C_b^m(\mathbb{R})$, $t \in [0, T]$ and $\alpha \propto \frac{1}{\sqrt{n}}$ (defined in* (3.8)*), we have*

$$\tilde{\mathbb{E}} \left[(\rho_t^n(\varphi) - \rho_t(\varphi))^2 \right] \leq \frac{c_1^T}{n} \|\varphi\|_{m,\infty}^2. \tag{4.37}$$

For the normalised approximating measure π^n, we have the following result.

Theorem 4.2.9. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists a constant c_7^T independent of n such that for $\alpha \propto \frac{1}{\sqrt{n}}$ and $\varphi \in C_b^m(\mathbb{R})$, we have*

$$\tilde{\mathbb{E}} \left[|\pi_t^n(\varphi) - \pi_t(\varphi)| \right] \leq \frac{c_7^T}{\sqrt{n}} \|\varphi\|_{m,\infty}, \qquad t \in [0, T]. \tag{4.38}$$

Proof. Observe that $\rho_t^n(\varphi) = \rho_t^n(\mathbf{1})\pi_t^n(\varphi)$, we then have

$$\begin{aligned}&\pi_t^n(\varphi) - \pi_t(\varphi)\\ =&\left(\rho_t^n(\varphi) - \rho_t(\varphi)\right)\left(\rho_t(\mathbf{1})\right)^{-1} - \pi_t^n(\varphi)\left(\rho_t^n(\mathbf{1}) - \rho_t(\mathbf{1})\right)\left(\rho_t(\mathbf{1})\right)^{-1}.\end{aligned}$$

Use the fact that $m_t \triangleq \sqrt{\tilde{\mathbb{E}}\left[(\rho_t(\mathbf{1}))^{-2}\right]} < \infty$ (see Exercise 9.16 of [3] for details), and by Cauchy-Schwartz inequality:

$$\begin{aligned}&\tilde{\mathbb{E}}\left[|\pi_t^n(\varphi) - \pi_t(\varphi)|\right]\\ \leq& m_t\left(\sqrt{\tilde{\mathbb{E}}\left[(\rho_t^n(\varphi) - \rho_t(\varphi))^2\right]} + \|\varphi\|_{0,\infty}\sqrt{\tilde{\mathbb{E}}\left[(\rho_t^n(\mathbf{1}) - \rho_t(\mathbf{1}))^2\right]}\right),\end{aligned} \tag{4.39}$$

and the result follows by applying Theorem 4.2.8 to the two expectations of the right hand side of (4.39). ■

The above results applied to the convergence of $\rho_t^n(\varphi)$ to $\rho_t(\varphi)$ and $\pi_t^n(\varphi)$ to $\pi_t(\varphi)$ for a fixed test function φ. We now discuss the convergence of ρ_t^n to ρ_t and π_t^n to π_t. Let $\mathcal{M} = \{\varphi_i,\ i \geq 0\} \in C_b^6(\mathbb{R}^d)$ be a countable convergence determining set[2] such that $\|\varphi_i\|_{6,\infty} \leq 1$ for any $i > 0$ and $d_{\mathcal{M}}$ be the metric on $\mathcal{M}_F(\mathbb{R}^d)$

$$d_{\mathcal{M}} : \mathcal{M}_F(\mathbb{R}^d) \times \mathcal{M}_F(\mathbb{R}^d) \to [0,\infty);$$

$$d_{\mathcal{M}}(\mu,\nu) = \sum_{i=0}^{\infty} \frac{|\mu\varphi_i - \nu\varphi_i|}{2^i}.$$

Theorem 4.2.8 and Theorem 4.2.9 imply the following corollary:

Corollary 4.2.10. *If Assumption (A) is satisfied. Then for any $T \geq 0$, there exist two constants c_8^T and c_9^T independent of n, such that*

$$\sup_{t\in[0,T]} \tilde{\mathbb{E}}[d_{\mathcal{M}}(\rho_t^n, \rho_t)] \leq \frac{2\sqrt{c_8^T}}{\sqrt{n}} \qquad \sup_{t\in[0,T]} \tilde{\mathbb{E}}[d_{\mathcal{M}}(\pi_t^n, \pi_t)] \leq \frac{2c_9^T}{\sqrt{n}} \tag{4.40}$$

This corollary means that ρ_t^n converges to ρ_t in expectation and π_t^n converges to π_t in expectation. A stronger convergence result for ρ_t^n and π_t^n will be proved in the following two theorems, from which we can see that their convergence are uniform in time t.

[2]See Theorem 2.18 in [3] for the existence of such set.

Proposition 4.2.11. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists a constant c_{10}^T independent of n such that for any $\varphi \in C_b^{m+2}(\mathbb{R})$,*

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}(\rho_t^n(\varphi)-\rho_t(\varphi))^2\right] \leq \frac{c_{10}^T}{n}\|\varphi\|_{m+2,\infty}^2. \tag{4.41}$$

Proof. By Proposition 4.1.1 and the fact that $\rho_t(\varphi)$ satisfies Zakai equation, we have

$$\begin{aligned}\rho_t^n(\varphi) = {} & (\pi_0^n(\varphi)-\pi_0(\varphi)) \\ & + \int_0^t (\rho_s^n(A\varphi)-\rho_s(A\varphi))ds + \int_0^t (\rho_s^n(h\varphi)-\rho_s(h\varphi))dY_s + M_{[t/\delta]}^{n,\varphi} \\ & + \frac{1}{n}\sum_{j=1}^n\sum_{i=0}^\infty \int_{i\delta\wedge t}^{(i+1)\delta\wedge t} \xi_{i\delta}^n a_j^n(s)\Big[R_{s,j}^1(\varphi)ds + R_{s,j}^2(\varphi)dY_s + R_{s,j}^3(\varphi)dV_s^{(j)}\Big].\end{aligned} \tag{4.42}$$

By Lemmas 4.2.2 – 4.2.4, we know that,

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\frac{1}{n}\sum_{j=1}^n\sum_{i=0}^\infty\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi_{i\delta}^n a_j^n(r)R_{s,j}^1(\varphi)ds\right)^2\right] \leq c_3^T(\alpha\delta)^2\|\varphi\|_{6,\infty}^2;$$

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\frac{1}{n}\sum_{j=1}^n\sum_{i=0}^\infty\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi_{i\delta}^n a_j^n(r)R_{s,j}^2(\varphi)dY_s\right)^2\right] \leq c_4^T(\alpha\delta)^2\|\varphi\|_{4,\infty}^2;$$

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\frac{1}{n}\sum_{j=1}^n\sum_{i=0}^\infty\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi_{i\delta}^n a_j^n(r)R_{s,j}^3(\varphi)dV_s^{(j)}\right)^2\right] \leq \frac{c_5^T}{n}\|\varphi\|_{5,\infty}^2.$$

By Doob's maximal inequality and Lemma 4.2.5

$$\tilde{\mathbb{E}}\left[\max_{i=1,\dots,[T/\delta]}(M_i^{n,\varphi})^2\right] \leq 4\tilde{\mathbb{E}}\left[\left(M_{[T/\delta]}^{n,\varphi}\right)^2\right] \leq \frac{4c_6^{[T/\delta]}}{n}\|\varphi\|_{1,\infty}^2;$$

Now we only need to bound the first three terms on the right-hand side of (4.42). For the first term, using the mutual independence of the initial locations of the particles $v_j^n(0)$,

$$\tilde{\mathbb{E}}\left[(\,\pi_0^n(\varphi)-\pi_0(\varphi))^2\right] = \frac{1}{n}\left(\pi_0(\varphi^2)-\pi_0(\varphi)^2\right) \leq \frac{1}{n}\|\varphi\|_{2,\infty}^2 \leq \frac{1}{n^q}\|\varphi\|_{2,\infty}^2.$$

For the second term, by Jensen's inequality and Fubini's Theorem, together with Theorem 4.2.8, we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\int_0^t (\rho_s^n(A\varphi)-\rho_s(A\varphi))ds\right)^2\right] \le \tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} t\int_0^t (\rho_s^n(A\varphi)-\rho_s(A\varphi))^2 ds\right]\\
=&\tilde{\mathbb{E}}\left[T\int_0^T (\rho_s^n(A\varphi)-\rho_s(A\varphi))^2 ds\right] = T\int_0^T \tilde{\mathbb{E}}\left[(\rho_s^n(A\varphi)-\rho_s(A\varphi))^2\right] ds\\
\le&\frac{c_0^T T^2}{n}\|A\varphi\|_{m,\infty}^2 = \frac{\tilde{c}_0^T T^2}{n}\|\varphi\|_{m+2,\infty}^2.
\end{aligned}
$$

For the third term, similarly, by Burkholder-Davis-Gundy inequality and Fubini's Theorem, together with Theorem 4.2.8, we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\int_0^t (\rho_s^n(h\varphi)-\rho_s(h\varphi))dY_s\right)^2\right] \le \tilde{C}\tilde{\mathbb{E}}\left[\int_0^T (\rho_s^n(h\varphi)-\rho_s(h\varphi))^2 ds\right]\\
\le&\tilde{C}\int_0^T \tilde{\mathbb{E}}\left[(\rho_s^n(h\varphi)-\rho_s(h\varphi))^2\right] ds \le \frac{\tilde{C}c_0^T T\|h\|_{0,\infty}^2}{n}\|\varphi\|_{m,\infty}^2.
\end{aligned}
$$

The above obtained bounds together imply (4.41). ■

Proposition 4.2.12. *If Assumption (A) is satisfied. Then for any $T \ge 0$, $m \ge 6$, there exists a constant c_{11}^T independent of n such that for and $\varphi \in C_b^{m+2}(\mathbb{R})$,*

$$
\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} |\pi_t^n(\varphi)-\pi_t(\varphi)|\right] \le \frac{c_{11}^T}{\sqrt{n}}\|\varphi\|_{m+2,\infty}. \tag{4.43}
$$

Proof. As in the proof of Theorem 4.2.9,

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} |\pi_t^n(\varphi)-\pi_t(\varphi)|\right]\\
\le&\hat{m}_T\left(\sqrt{\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}(\rho_t^n(\varphi)-\rho_t(\varphi))^2\right]} + \|\varphi\|_{0,\infty}\sqrt{\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}(\rho_t^n(\mathbf{1})-\rho_t(\mathbf{1}))^2\right]}\right),
\end{aligned}
$$

where

$$
m_t \triangleq \sqrt{\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}(\rho_t(\mathbf{1}))^{-2}\right]} < \infty
$$

and the result follows from Theorem 4.2.11. ■

Let $\bar{\mathcal{M}} = \{\varphi_i,\ i \geq 0\} \in C_b^8(\mathbb{R}^d)$ be a countable convergence determining set such that $\|\varphi_i\|_{8,\infty} \leq 1$ for any $i > 0$ and with the same $d_{\bar{\mathcal{M}}}$ be the metric on $\mathcal{M}_F(\mathbb{R}^d)$. Then the following corollary follows immediately from Theorem 4.2.11 and Theorem 4.2.12

Corollary 4.2.13. *If Assumption (A) is satisfied. Then for any $T \geq 0$, there exist two constants c_{12}^T and c_{13}^T, such that*

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} d_{\bar{\mathcal{M}}}(\rho_t^n, \rho_t)\right] \leq \frac{2\sqrt{c_{12}^T}}{\sqrt{n}} \qquad \tilde{\mathbb{E}}\left[\sup_{t\in[0,T]} d_{\bar{\mathcal{M}}}(\pi_t^n, \pi_t)\right] \leq \frac{2c_{13}^T}{\sqrt{n}} \tag{4.44}$$

Remark 4.2.14. *The fact that the optimal value for α decreases with n is not surprising. As the number of particles increases, the quantisation of the posterior distribution becomes finer and finer. Therefore,* asymptotically*, the position and the weight of the particle provide sufficient information to obtain a good approximation.*

Remark 4.2.15. *Since the approximations ρ_t^n and π_t^n have smooth densities with respect to the Lebesgue measure, it makes it possible to study various properties the density of ρ_t and that from its approximation ρ_t^n (for example, the position of their maximum value, the decay in time, the properties of their derivatives, etc). This would be possible under the classic particle filtering framework, where the approximations are linear combinations of Dirac measures, only if a smoothing procedure is applied first (see [24]).*

In this section, the convergence results and L^2-error are obtained under probability $\tilde{\mathbb{P}}$; however, it is more natural to investigate these results under the original probability $\mathbb{P}$. The following proposition states the L^2-convergence result under $\mathbb{P}$.

Proposition 4.2.16. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists constant c^T independent of n, such that for any $\varphi \in C_b^m(\mathbb{R})$, $t \in [0,T]$ and $\alpha \propto \frac{1}{\sqrt{n}}$ (defined in* (3.8)*), we have*

$$\mathbb{E}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|\right] \leq \frac{c^T}{\sqrt{n}}\|\varphi\|_{m,\infty}. \tag{4.45}$$

Proof. Recalling the derivation of the new probability $\tilde{\mathbb{P}}$, we know that

$$\tilde{Z}_t = \exp\left(\int_0^t h(X_s)dY_s^i - \frac{1}{2}\int_0^t h^2(X_s)ds\right) \quad (t \geq 0) \tag{4.46}$$

is an $\mathcal{F}_t$-adapted martingale under $\tilde{\mathbb{P}}$ and

$$\left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \tilde{Z}_t \quad t \geq 0.$$

Therefore

$$\begin{aligned}\mathbb{E}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|\right] =& \tilde{\mathbb{E}}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|\tilde{Z}_t\right] \\ \leq& \sqrt{\tilde{\mathbb{E}}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|^2\right]\tilde{\mathbb{E}}\left[(\tilde{Z}_t)^2\right]} \\ \leq& \sqrt{\frac{c_1^T c_{\tilde{z}}}{n}}\|\varphi\|_{m,\infty}. \end{aligned} \tag{4.47}$$

The result follows by letting $c^T = \sqrt{c_1^T c_{\tilde{z}}}$. ■

Remark 4.2.17. *Under the Tree Based Branching Algorithm (TBBA), L^p-convergence result for ρ^n cannot be generally obtained. This is because, in general, pth moment of $M^{n,\varphi}_{[t/\delta]}$ can not be obtained and controlled under $\tilde{\mathbb{P}}$. As a result, one can only obtain L^1-convergence result for ρ^n under the original probability $\mathbb{P}$.*

4.3 Convergence Results using the Multinomial Branching Algorithm

In this section we show the convergence result for the case where the resampling is conducted by using Multinomial branching algorithm. The following theorem states the main convergence result.

Theorem 4.3.1. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists constant c_1^T independent of n, such that for any $\varphi \in C_b^m(\mathbb{R})$, $t \in [0,T]$ and $\alpha \propto \frac{1}{\sqrt{n}}$ (defined in (3.8)), we have*

$$\tilde{\mathbb{E}}\left[(\rho_t^n(\varphi) - \rho_t(\varphi))^2\right] \leq \frac{c_1^T}{n}\|\varphi\|^2_{m,\infty}. \tag{4.48}$$

In order to prove Theorem 4.3.1, we note that after replacing the TBBA by Multinomial branching, all the analysis in Section 4.2 is automatically valid, except for the analysis on $M_i^{n,\varphi}$. Therefore, it suffices to re-investigate the L^2-bound for $M_i^{n,\varphi}$ only. In other words, we only need to modify the proof of Lemma 4.2.5 as follows.

Lemma 4.3.2. *If Assumption (A) is satisfied. Then for any $T > 0$, given a test function $\varphi \in C_b^1(\mathbb{R})$, $M^{n,\varphi}_{[t/\delta]}$ can be decomposed as*

$$M^{n,\varphi}_{[t/\delta]} = A^{n,\varphi}_{[t/\delta]} + D^{n,\varphi}_{[t/\delta]} + G^{n,\varphi}_{[t/\delta]}; \tag{4.49}$$

where $A^{n,\varphi}_{[t/\delta]}$ and $G^{n,\varphi}_{[t/\delta]}$ are defined the same as the Section 4.2; and we further have

$$\tilde{\mathbb{E}}\left[|A^{n,\varphi}_{[t/\delta]}|^2\right] \leq \frac{c_6^T}{n\delta}\|\varphi\|^2_{0,\infty}, \quad \tilde{\mathbb{E}}\left[|D^{n,\varphi}_{[t/\delta]}|^2\right] \leq \tilde{c}_6^T(\alpha\beta)^2\|\varphi\|^2_{0,\infty}, \quad \tilde{\mathbb{E}}\left[|G^{n,\varphi}_i|^2\right] \leq \frac{c_7^T\alpha}{n}\|\varphi\|^2_{1,\infty}; \tag{4.50}$$

where c_6^T, $\tilde{c}_6^T$ and c_7^T are constants independent of n.

Proof. Recalling proposition 4.1.1 and the semigroup operator P, we can decompose $M^{n,\varphi}$ in the following way

$$\begin{aligned}
M^{n,\varphi}_{[t/\delta]} =& \frac{1}{n}\sum_{i=0}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^{n}\left[o_j^{n,i\delta}\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-X_j^n(i\delta))^2}{2\alpha\beta}}}{\sqrt{2\pi\alpha\beta}}dx - n\bar{a}_j^n(i\delta-)\int_{\mathbb{R}}\varphi(x)\frac{e^{-\frac{(x-v_j^n(i\delta-))^2}{2\omega_j^n(i\delta-)}}}{\sqrt{2\pi\omega_j^n(i\delta-)}}dx\right]\\
\triangleq& A^{n,\varphi}_{[t/\delta]} + D^{n,\varphi}_{[t/\delta]} + G^{n,\varphi}_{[t/\delta]},
\end{aligned}$$

where $X_j^n(i\delta) \sim \mathcal{N}\left(v_j^n(i\delta-), \omega_j^n(i\delta-)\right)$ is a Gaussian distributed random variable. For the first term, by Proposition 3.2.3

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\xi^n_{i\delta}\left(\left(o_j^{n,i\delta} - n\bar{a}_j^n(i\delta-)\right)\varphi(X_j^n(i\delta))\right)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi^n_{i\delta})^2}{n^2}\,\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\left(\left(o_j^{n,i\delta} - n\bar{a}_j^n(i\delta-)\right)\varphi(X_j^n(i\delta))\right)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi^n_{i\delta})^2}{n^2}\|\varphi\|^2_{0,\infty}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\left(o_j^{n,i\delta} - n\bar{a}_j^n(i\delta-)\right)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
\leq&\frac{(\xi^n_{i\delta})^2}{n^2}\|\varphi\|^2_{0,\infty}\tilde{\mathbb{E}}\left[\sum_{j=1}^{n}\left(o_j^{n,i\delta} - n\bar{a}_j^n(i\delta-)\right)^2\middle|\mathcal{Y}_{i\delta-}\vee\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi^n_{i\delta})^2}{n^2}\|\varphi\|^2_{0,\infty}\left[\sum_{j=1}^{n}n\bar{a}_j^n(i\delta-)(1-\bar{a}_j^n(i\delta-))\right].
\end{aligned}\tag{4.51}$$

We know from the proof of Lemma 4.4 in [56] that for any $p > 0$,

$$\tilde{\mathbb{E}}\left[\left(\bar{a}_j^n(t)\right)^p\right] \leq \frac{1}{n^p}\exp(c_pt);$$

then by Cauchy-Schwarz inequality, we know

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^2\bar{a}^n_j(i\delta-)(1-\bar{a}^n_j(i\delta-))\right]\\
\leq&\sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^4\right]\tilde{\mathbb{E}}\left[\left(\bar{a}^n_j(i\delta-)(1-\bar{a}^n_j(i\delta-))\right)^2\right]}\leq\sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^4\right]\tilde{\mathbb{E}}\left[\left(\bar{a}^n_j(i\delta-)\right)^2\right]}\\
\leq&\frac{1}{n}\sqrt{c_1^{T,4}\exp(c_2T)},
\end{aligned}\tag{4.52}$$

then we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{j=1}^n\xi^n_{i\delta}\left(\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)\varphi(X^n_j(i\delta))\right)\right)^2\right]\\
\leq&\frac{\|\varphi\|^2_{0,\infty}}{n^2}n\sum_{j=1}^n\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^2\bar{a}^n_j(i\delta-)(1-\bar{a}^n_j(i\delta-))\right]\\
\leq&\frac{\|\varphi\|^2_{0,\infty}}{n^2}n\sum_{j=1}^n\frac{1}{n}\sqrt{c_1^{T,4}\exp(c_2T)}=\frac{\|\varphi\|^2_{0,\infty}\sqrt{c_1^{T,4}\exp(c_2T)}}{n}.
\end{aligned}\tag{4.53}$$

Therefore

$$\tilde{\mathbb{E}}\left[\left(A^{n,\varphi}_{[t/\delta]}\right)^2\right]\leq\sum_{i=1}^{[t/\delta]}\frac{\sqrt{c_1^{T,4}\exp(c_2T)}}{n}\|\varphi\|^2_{0,\infty}\leq\frac{t\sqrt{c_1^{T,4}\exp(c_2T)}}{\delta n}\|\varphi\|^2_{0,\infty}.\tag{4.54}$$

For $G^{n,\varphi}_{[t/\delta]}$, first by noting that

$$\int_{\mathbb{R}}\varphi\left(X^n_j(i\delta)+y\sqrt{\alpha\beta}\right)\frac{e^{\frac{-y^2}{2}}}{\sqrt{2\pi}}dy-\varphi(X^n_j(i\delta))=\quad\frac{\alpha\beta}{2}\varphi''\left(X^n_j(i\delta)\right)+\mathcal{O}\left((\alpha\beta)^2\right);\tag{4.55}$$

then it is clear that we only need to show

$$\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X^n_j(i\delta)\right)\right)\right)^2\right]\leq\tilde{c}_6^T(\alpha\beta)^2\|\varphi\|^2_{0,\infty}.\tag{4.56}$$

Observe that

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X_j^n(i\delta)\right)\right)\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
=&\frac{(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n o_j^{n,i\delta}\varphi''\left(X_j^n(i\delta)\right)\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
\le&\frac{(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\|\varphi\|^2_{2,\infty}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)+n\bar{a}^n_j(i\delta-)\right]\right)^2\middle|\mathcal{F}_{i\delta-}\right]\\
\le&\frac{2(\xi^n_{i\delta})^2(\alpha\beta)^2}{n^2}\|\varphi\|^2_{2,\infty}\left[\sum_{j=1}^n\tilde{\mathbb{E}}\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)^2\middle|\mathcal{F}_{i\delta-}\right]+\left(\sum_{j=1}^n n\bar{a}^n_j(i\delta-)\right)^2\right].
\end{aligned}
\tag{4.57}
$$

For Multinomial Branching Algorithm, by Proposition 3.2.3,

$$
\tilde{\mathbb{E}}\left[\left(o_j^{n,i\delta}-n\bar{a}^n_j(i\delta-)\right)^2\middle|\mathcal{F}_{i\delta-}\right]=n\bar{a}^n_j(i\delta-)(1-\bar{a}^n_j(i\delta-)),\tag{4.58}
$$

and then by taking expectation on both sides of (4.57), we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X_j^n(i\delta)\right)\right)\right)^2\right]\\
\le&\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}\sum_{j=1}^n\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^2\bar{a}^n_j(i\delta-)(1-\bar{a}^n_j(i\delta-))\right]\\
&+\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}n^2\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^2\left(\sum_{j=1}^n\bar{a}^n_j(i\delta-)\right)^2\right]\\
\le&\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}\sum_{j=1}^n\sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^4\right]\tilde{\mathbb{E}}\left[\bar{a}^n_j(i\delta-)^2(1-\bar{a}^n_j(i\delta-))^2\right]}+2(\alpha\beta)^2\|\varphi\|^2_{2,\infty}c_1^{t,4}e^{c_2t}\\
\le&\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}\sum_{j=1}^n\frac{1}{n}\sqrt{c_1^{t,4}e^{c_2t}}+2(\alpha\beta)^2\|\varphi\|^2_{2,\infty}c_1^{t,4}e^{c_2t}\\
=&\frac{2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2}{n^2}\sqrt{c_1^{t,4}e^{c_2t}}+2\|\varphi\|^2_{2,\infty}(\alpha\beta)^2c_1^{t,4}e^{c_2t}=c^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty},
\end{aligned}
\tag{4.59}
$$

where $c^T=2\sqrt{c_1^{t,4}e^{c_2t}}+2c_1^{t,4}e^{c_2t}$.

Therefore, we obtain

$$
\tilde{\mathbb{E}}\left[\left(\frac{1}{n}\sum_{i=1}^{[t/\delta]}\xi^n_{i\delta}\sum_{j=1}^n o_j^{n,i\delta}\left((\alpha\beta)\varphi''\left(X_j^n(i\delta)\right)\right)\right)^2\right]\le[t/\delta]\sum_{i=1}^{[t/\delta]}c^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty}=\tilde{c}_6^T(\alpha\beta)^2\|\varphi\|^2_{2,\infty}
$$

if we let $\tilde{c}_6^T = T^2 c^T/\delta^2$.

As for $G^{n,\varphi}_{[t/\delta]}$, first note that $X_j^n(i\delta) \sim N\left(v_j^n(i\delta), \omega_j^n(i\delta)\right)$ and X_j^ns are mutually independent ($j = 1, \ldots, n$), then we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n \xi_{i\delta}^n \bar{a}_j^n(i\delta-)\left[\varphi(X_j^n(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X_j^n(k\delta))\right)\right]\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq& \sum_{j=1}^n \left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2 \tilde{\mathbb{E}}\left[\left(\varphi(X_j^n(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X_j^n(k\delta))\right)\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq& \sum_{j=1}^n \left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2 \tilde{\mathbb{E}}\left[\left(\varphi(X_j^n(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X_j^n(k\delta))\right)\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq& \sum_{j=1}^n \left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2 2\|\varphi'\|_{0,\infty}^2 \tilde{\mathbb{E}}\left[\left(X_j^n(k\delta) - \tilde{\mathbb{E}}\left(X_j^n(k\delta)\right)\right)^2 \middle| \mathcal{Y}_{i\delta-}\right] \\
\leq& 2\sum_{j=1}^n \left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2 \|\varphi\|_{1,\infty}^2 \|\sigma\|_{0,\infty}^2 \alpha\delta \triangleq c_\sigma \alpha\delta \|\varphi\|_{1,\infty}^2 \sum_{j=1}^n \left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2. \qquad (4.60)
\end{aligned}$$

then by taking the expectation on both sides, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n \xi_{i\delta}^n \bar{a}_j^n(i\delta-)\left[\varphi(X_j^n(k\delta)) - \tilde{\mathbb{E}}\left(\varphi(X_j^n(k\delta))\right)\right]\right)^2\right] \\
\leq& c_\sigma \alpha\delta \|\varphi\|_{1,\infty}^2 \sum_{j=1}^n \tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n \bar{a}_j^n(i\delta-)\right)^2\right] \leq c_\sigma \alpha\delta \|\varphi\|_{1,\infty}^2 \sum_{j=1}^n \sqrt{\tilde{\mathbb{E}}\left[(\xi_{i\delta}^n)^4\right] \tilde{\mathbb{E}}\left[\left(\bar{a}_j^n(i\delta-)\right)^4\right]} \\
\leq& c_\sigma \alpha\delta \|\varphi\|_{1,\infty}^2 \frac{1}{n^2} \sum_{j=1}^n \sqrt{c_1^{t,4} \exp(c_4 t)} = \frac{c_\sigma \sqrt{c_1^{t,4} \exp(c_4 t)}}{n} \alpha\delta \|\varphi\|_{1,\infty}^2. \qquad (4.61)
\end{aligned}$$

Finally we have

$$\tilde{\mathbb{E}}\left[\left(G^{n,\varphi}_{[t/\delta]}\right)^2\right] \leq \sum_{i=1}^{[t/\delta]} \frac{c_\sigma \sqrt{c_1^{t,4} \exp(c_4 t)}}{n} \alpha\delta \|\varphi\|_{1,\infty}^2 \leq \frac{t c_\sigma \sqrt{c_1^{t,4} \exp(c_4 t)}}{n} \alpha \|\varphi\|_{1,\infty}^2. \qquad (4.62)$$

The result follows by letting $c_6^T = t\sqrt{c_1^{T,4} \exp(c_2 T)}$ and $c_7^T = T c_\sigma \sqrt{c_1^{T,4} \exp(c_4 T)}$. ■

We also have the following convergence result for the approximation π_t^n of the normalised conditional distribution:

Theorem 4.3.3. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists a constant c_7^T independent of n such that for $\alpha \propto \frac{1}{\sqrt{n}}$ and $\varphi \in C_b^m(\mathbb{R})$, we have*

$$\tilde{\mathbb{E}}\left[|\pi_t^n(\varphi) - \pi_t(\varphi)|\right] \leq \frac{c_7^T}{\sqrt{n}}\|\varphi\|_{m,\infty}, \qquad t \in [0,T]. \tag{4.63}$$

Remark 4.3.4. *From the proof of the above theorem, it can be seen that all the other results obtained in Section 4.2 still hold under Multinomial branching algorithm.*

Under the multinomial branching algorithm, one can show that L^p-convergence result for ρ^n and π^n can be obtained for any $p \geq 2$, namely we have the following theorem. The proof of the theorem is similar to that of Chapter 4 of [56].

Theorem 4.3.5. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists constants c_1^T and c_2^T independent of n, such that for any $\varphi \in C_b^m(\mathbb{R})$, $t \in [0,T]$ and $\alpha \propto \frac{1}{\sqrt{n}}$ (defined in* (3.8)*), we have*

$$\tilde{\mathbb{E}}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|^p\right] \leq \frac{c_1^T}{n^{p/2}}\|\varphi\|_{m,\infty}^{p/2}. \tag{4.64}$$

$$\tilde{\mathbb{E}}\left[|\pi_t^n(\varphi) - \pi_t(\varphi)|^p\right] \leq \frac{c_2^T}{n^{p/2}}\|\varphi\|_{m,\infty}^{p/2}. \tag{4.65}$$

The following proposition shows the convergence result under the original probability $\mathbb{P}$.

Proposition 4.3.6. *If Assumption (A) is satisfied. Then for any $T \geq 0$, $m \geq 6$, there exists constant c^T independent of n, such that for any $\varphi \in C_b^m(\mathbb{R})$, $t \in [0,T]$, $q \geq 1$ and $\alpha \propto \frac{1}{\sqrt{n}}$ (defined in* (3.8)*), we have*

$$\mathbb{E}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|^q\right] \leq \frac{c^T}{n^{q/2}}\|\varphi\|_{m,\infty}^{q/2}. \tag{4.66}$$

Proof. Recalling the derivation of the new probability $\tilde{\mathbb{P}}$, we know that

$$\tilde{Z}_t = \exp\left(\int_0^t h(X_s)dY_s^i - \frac{1}{2}\int_0^t h^2(X_s)ds\right) \quad (t \geq 0) \tag{4.67}$$

is an $\mathcal{F}_t$-adapted martingale under $\tilde{\mathbb{P}}$ and

$$\left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \tilde{Z}_t \quad t \geq 0.$$

Therefore by Cauchy-Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}\left[|\rho_t^n(\varphi)-\rho_t(\varphi)|^q\right] =&\tilde{\mathbb{E}}\left[|\rho_t^n(\varphi)-\rho_t(\varphi)|^q\tilde{Z}_t\right]\\ \leq&\sqrt{\tilde{\mathbb{E}}\left[|\rho_t^n(\varphi)-\rho_t(\varphi)|^{2q}\right]\tilde{\mathbb{E}}\left[(\tilde{Z}_t)^2\right]}\\ \leq&\sqrt{\frac{c_1^Tc_{\tilde{z}}}{n^q}\|\varphi\|_{m,\infty}^q}. \end{aligned}\tag{4.68}$$

The result follows by letting $c^T=\sqrt{c_1^Tc_{\tilde{z}}}$. ■

4.4 A Numerical Example

In this section, we present some numerical experiments to test the performance of the approximations with mixture of Gaussian measures. The model chosen in this case is the Beneš filter. This is a stochastic filtering problem with a nonlinear dynamics for the signal process and a linear dynamics the observation process, and this problem has an analytical finite dimensional solution. The main reason for choosing this model is that it has a sufficient nonlinear behaviour to make it interesting, and more importantly, still has a closed form for its solution.

4.4.1 The Model and its Exact Solution

We assume that both the signal and the observation are one-dimensional. The dynamics of the signal X is

$$X_t = X_0 + \int_0^t f(X_s)ds + \sigma V_t, \tag{4.69}$$

where $f(x)=\mu\sigma\tanh(\mu x/\sigma)$. From Exercise 6.1 in [3] we know that f satisfies the Beneš condition. We further assume the observation Y satisfies

$$Y_t = \int_0^t h(X_s)ds + W_t, \tag{4.70}$$

where W is a standard Brownian motion independent of V, and $h(x)=h_1x+h_2$. We also assume that $X_0,\mu,h_1,h_2\in\mathbb{R}$ and $\sigma>0$.

Then from [25] we know that the conditional law of X_t given $\mathcal{Y}_t \triangleq \sigma(Y_s, 0 \le s \le t)$ has the exact expression of a weight mixture of two Gaussian distributions. In other words, the conditional distribution π_t of X_t is

$$\pi_t = w^+\mathcal{N}(A^+/(2B_t), 1/(2B_t)) + w^-\mathcal{N}(A^-/(2B_t), 1/(2B_t)),$$

where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2, and

$$\begin{aligned}
w^\pm &\triangleq \exp\left((A_t^\pm)^2/(4B_t)\right) / \left[\exp\left((A_t^+)^2/(4B_t)\right) + \exp\left((A_t^-)^2/(4B_t)\right)\right] \\
A_t^\pm &\triangleq \pm\frac{\mu}{\sigma} + h_1\Psi_t + \frac{h_1 X_0 + h_2}{\sigma\sinh(h_1\sigma t)} - \frac{h_2}{\sigma}\coth(h_1\sigma t) \\
B_t &\triangleq \frac{h_1}{2\sigma}\coth(h_1\sigma t) \\
\Psi_t &\triangleq \int_0^t \frac{\sinh(h_1\sigma s)}{\sinh(h_1\sigma t)} dY_s.
\end{aligned}$$

We can, however, only observe Y at a finite partition $\Pi^{m,T} = \{0 = t_0 < t_1 < \cdots < t_{m-1} = T\}$ of $[0, T]$ in practice; thus we approximate the integral in the definition of Ψ_t by

$$\Psi_t \approx \sum_{k=0}^{i-1} \frac{\sinh(h_1\sigma t_{k+1})}{\sinh(h_1\sigma t_i)}(Y_{t_{k+1}} - Y_{t_k}), \quad \text{for } t_i \in \Pi^{n,T}.$$

4.4.2 Numerical Simulation Results

We set values for the parameters μ, σ, h_1, h_2, X_0 and T as follows:

$$\mu = 0.3,\ h_1 = 0.8,\ h_2 = 0.0,\ \sigma = 1.0,\ X_0 = 0.0,\ T = 10.0;$$

and then we compute one realisation of X_t and Y_t respectively using the Euler scheme with an equidistant partition $\Pi^{m,T} = \{t_i = \frac{i}{m}T\}_{i=0,\ldots,m}$ with $m = 10^6$. This realisation is then fixed and will act as the given observation path. After that, all the simulations will be done assuming that we are given the previously obtained Y_t computed from that realisation of X_t. With this previously simulated discrete path of Y, we can then approximate Ψ_t and consequently compute the values of $A_t^\pm$, B_t and $w_t^\pm$; so that we can compute the conditional law of $\varphi(X_t)$ given $\mathcal{Y}_t$. At the branching time, we use the Tree Based Branching Algorithm. We will show the convergence of the Gaussian mixture approximation and the classic particle filters in terms of the number of time steps in the partition and the number of particles respectively.

We note that for the test function $\varphi(x) = x$, the Gaussian mixture approximation gives

$$\pi_t^n(x) = \sum_{j=1}^n \bar{a}_j^n(t) \int_{\mathbb{R}} \left(v_j^n(t) + y\sqrt{\omega_j^n(t)} \right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2} \right) dy = \sum_{j=1}^n \bar{a}_j^n(t) v_j^n(t).$$

This is almost the same result as the classic particle filters, except that the evolution equations satisfied by $v_j^n(t)$s are slightly different in two cases (see equations (3.1) and (3.8) for details). It is therefore more meaningful to estimate the normalised conditional distribution $\pi_T(\varphi)$ for $\varphi(x) = x^2$ and $\varphi(x) = x^3$, that is, the second and third moments of the system at time T given the observation Y up to time T. To be specific, we estimate $\pi_T(\varphi)$ by $\pi_T^n(\varphi)$ with the number of particles (of Gaussian generalised particles) $n = 40000$ and we choose various values for the number of time steps m in the partition. We compute $\pi_T^n(\varphi)$ using classic particles and mixture of Gaussian measures respectively. Instead of the absolute error $|\pi_T^n(\varphi) - \pi_T(\varphi)|$, we consider the relative error

$$\frac{|\pi_T^n(\varphi) - \pi_T(\varphi)|}{|\pi_T(\varphi)|}.$$

The convergence of both methods as the number of discretisation time steps m increases can be seen from the following Figure 4.1, and for large number of time steps the Gaussian mixture approximation performs slightly better than the classic particle filters.

In the following we fix the number of time steps $m = 110$ and vary the number of particles n in the approximating system.

From Figure 4.2 and Figure 4.3 we can see the convergence of both approximations with the increase of the number of (generalised) particles. It can be seen (from the right hand side of Figures 4.2 and 4.3) that for small number of (generalised) particles, Gaussian mixture approximation performs much better than the classic particle filters. This is because by using the Gaussian mixture approximation, each (generalised) particle carries more information about the signal (from its variance) than the classic particle does. Therefore a smaller number of particles is required in order to obtain the same level of accuracy.

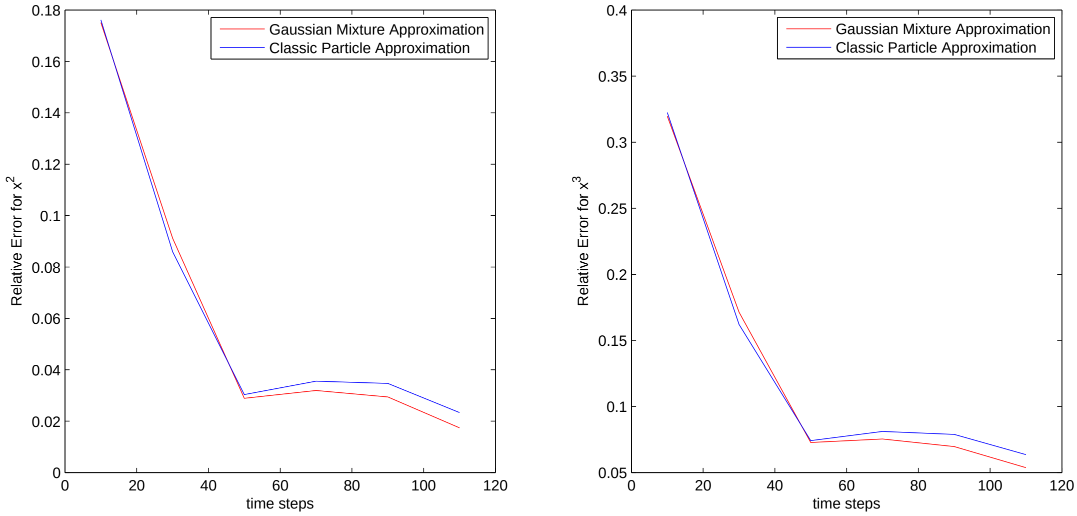

Figure 4.1: Relative Errors with time steps for $\varphi(x) = x^2$ (left) & $\varphi(x) = x^3$ (right)

As the number of (generalised) particles increases, we can see (from the left hand side of Figures 4.2 and 4.3) that the Gaussian mixture approximation converges faster than the classic particle filters; and we are able to obtain a good approximation for both methods with 10^4 particles. There is no significant improvement if we increase the number of (generalised) particles further more in both approximating systems.

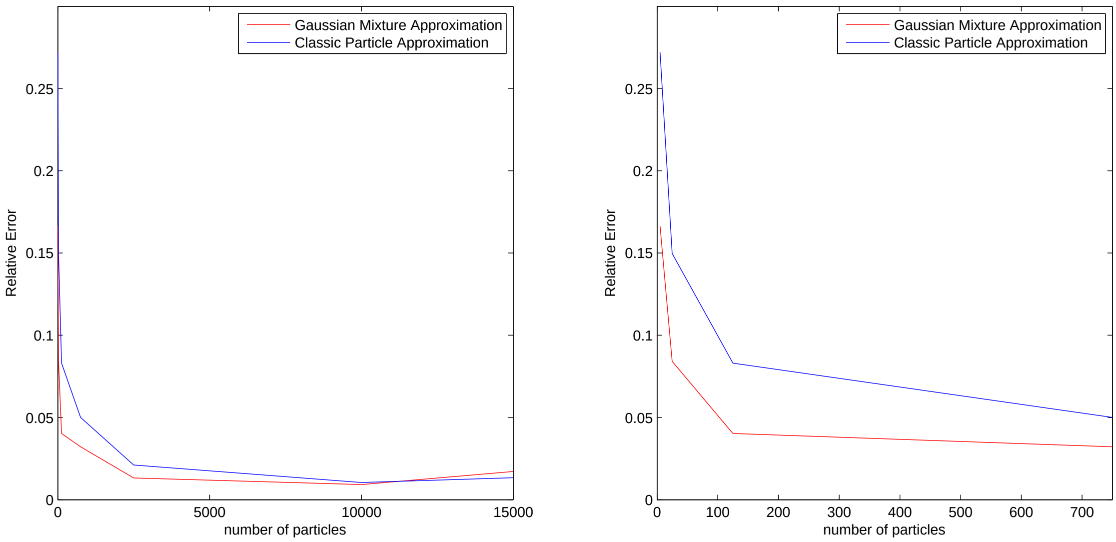

Figure 4.2: Relative Errors with different number of particles for $\varphi(x) = x^2$

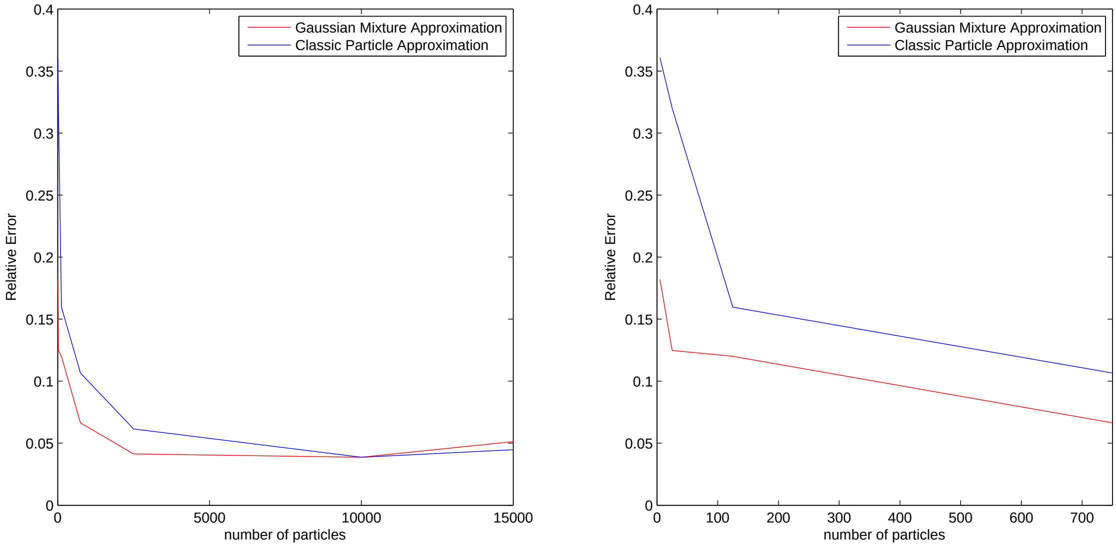

Figure 4.3: Relative Errors with different number of particles for $\varphi(x) = x^3$

Chapter 5

Central Limit Theorem

In this chapter we will obtain a central limit type result for the unnormalised (and normalised) conditional distributions ρ (and π) and their approximating measures ρ^n (and π^n). In other words, we aim to show that

$$\sqrt{n}\,(\rho^n - \rho) \quad \text{and} \quad \sqrt{n}\,(\pi^n - \pi) \tag{5.1}$$

converge in distribution as the number of generalised particles n increases. We proceed in a standard manner: we prove first a tightness result; and then deduce the convergence in distribution. In this chapter, we will prove the central limit theorem under Multinomial branching algorithm. We will also include a further discussion on the preference of multinomial branching over Tree Based Branching Algorithm (TBBA) at the end of this chapter.

5.1 Tightness

First we recall the definitions of relative compactness and tightness. It is possible to obtain the tightness and convergence in distribution results by endowing $\mathcal{M}_F(\mathbb{R})$ with the weak topology. In this topology a sequence of finite measures $\{\mu^n\}_{n\in\mathbb{N}} \subset \mathcal{M}_F(\mathbb{R})$ converges to $\mu \in \mathcal{M}_F(\mathbb{R})$ if and only if for a set $\mathcal{S}(\varphi)$ of test functions, $\mu^n(\varphi)$ converges to $\mu(\varphi)$ for all $\varphi \in \mathcal{S}(\varphi)$. $\mathcal{S}(\varphi)$ can be taken to be $C_b^m(\mathbb{R})$ for any $m \geq 1$.

Definition 5.1.1. *For (X, d) a metric space and $\prod$, a family of probability measures on $(X, \mathcal{B}(X))$, we say*

- $\prod$ *is relatively compact if every sequence of elements of* $\prod$ *contains a convergent subsequence.*

- $\prod$ *is tight if for every* $\varepsilon > 0$*, there exists a compact set* $K \subseteq X$ *such that* $\mathbb{P}(K) \geq 1 - \varepsilon$*, for every* $\mathbb{P} \in \prod$*.*

The following theorem (see, for example, Theorem 2.4.7 in [44]) states the relation between relative compactness and tightness.

Theorem 5.1.2 (Prohorov's Theorem)**.** *Let* Π *be a family of probability measures on a complete, separable metric space* (X, d)*. This family is relatively compact if and only if it is tight.*

Define $U = \{U_t^n : t \geq 0\}$ to be the measure-valued process

$$U_t^n \triangleq \sqrt{n}(\rho_t^n - \rho_t), \tag{5.2}$$

and we aim to show that $\{U^n\}$ converges in distribution to a process U, which is uniquely identified as the solution of a certain martingale problem. This implies that for any continuous and bounded test function,

$$\lim_{n \to \infty} \sqrt{n}(\rho_t^n(\varphi) - \rho_t(\varphi)) = U_t(\varphi); \tag{5.3}$$

hence the error of the approximations $\rho_t^n(\varphi)$ of $\rho_t(\varphi)$ is roughly $U_t(\varphi)/\sqrt{n}$.

Recalling Proposition 4.1.1, we deduced that

$$U_t^n(\varphi) = U_0^n(\varphi) + \int_0^t U_s^n(A\varphi)ds + \int_0^t U_t^n(h\varphi)dY_s + \sqrt{n}M_{[t/\delta]}^{n,\varphi} + \sqrt{n}B_t^{n,\varphi}, \tag{5.4}$$

where

$$\sqrt{n}B_t^{n,\varphi} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \int_0^t \xi_{[s/\delta]\delta}^n a_j^n(s) \Big[R_{s,j}^1(\varphi)ds + R_{s,j}^2(\varphi)dY_s + R_{s,j}^3(\varphi)dV_s^{(j)} \Big]. \tag{5.5}$$

Before proceeding further discussion on U^n, we define the metric on $\mathcal{M}_F(\mathbb{R})$ which generates the weak topology. Let $\varphi_0 = 1$ and $\{\varphi_i\}_{i \geq 0}$ be a sequence of functions which are dense in the space of continuous functions with compact support on $\mathbb{R}$. Then the metric $d_{\mathcal{M}}$ is defined as

$$d_{\mathcal{M}} : \mathcal{M}_F(\mathbb{R}) \times \mathcal{M}_F(\mathbb{R}) \to [0, \infty), \qquad d_{\mathcal{M}}(\mu, \nu) = \sum_{i=0}^{\infty} \frac{\mu(\varphi_i) - \nu(\varphi_i)}{2^i \|\varphi_i\|_{0,\infty}};$$

and $d_{\mathcal{M}}$ generates the weak topology on $\mathcal{M}_F(\mathbb{R})$ in the sense that μ^n converges weakly to μ if and only if $\lim_{n\to\infty} d_{\mathcal{M}}(\mu^n, \mu) = 0$ as $\{\varphi_i\}_{i\geq 0}$ is a convergence determining set of functions over $\mathcal{M}_F(\mathbb{R})$.

However, the space $\left(D_{\mathcal{M}_F(\mathbb{R})}[0,\infty), d_{\mathcal{M}}\right)$ is separable but not complete under this metric because its underlying space $(\mathcal{M}_F(\mathbb{R}), d_{\mathcal{M}})$ is separable but not complete. This inconvenience makes us unable to make use of Prohorov's Theorem. In order to tackle this problem, we consider the one-point compactification of $\mathbb{R}$

$$\overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{\infty\},$$

Then we embed the space $D_{\mathcal{M}_F(\mathbb{R})}[0,\infty)$ into the complete and separable space $D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$ by defining a map such that

$$\mu \in \mathcal{M}_F(\mathbb{R}) \to \overline{\mu} \in \mathcal{M}_F(\overline{\mathbb{R}}) \quad \text{and} \quad \overline{\mu}(A) = \mu(A \cap \mathbb{R}),\ \forall A \in \overline{\mathbb{R}}.$$

The family $\{U_t^n\}$ can then be viewed as a stochastic process with sample paths in the complete and separable space $D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$, or as a random variable with values in the space $\mathcal{P}(D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty))$ – the space of probability measures over $D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$.

We are now ready to show that the family of processes $\{U^n\}$ is tight on $[0,T]$ for all $T > 0$. In other words, let $\{\tilde{\mathbb{P}}_n\} \subset \mathcal{P}\left(D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,T]\right)$ be the family of associated probability distributions of U^n; in other words, $\tilde{\mathbb{P}}_n(B) = \tilde{\mathbb{P}}_n(U^n \in B)$ for all $B \in \mathcal{B}(D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,T])$. We aim to show that $\{\tilde{\mathbb{P}}_n\}$ is relatively compact and hence, by Prohorov's Theorem, tight. To be specific, we will make use of the following theorem (Theorem 2.1 in [62]):

Theorem 5.1.3. *A family of probabilities $\{\tilde{\mathbb{P}}_n\}_n \subset \mathcal{P}\left(D_{\mathcal{M}_F(\overline{\mathbb{R}^d})}[0,T]\right)$ is tight, if there exits a dense sequence $\{\tilde{f}_k\}_{k\geq 0}$ in $C_b(\overline{\mathbb{R}^d})$ such that for each $k \in \mathbb{N}$, $\{\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n\}_n \subset \mathcal{P}\left(D_{\overline{\mathbb{R}}}[0,T]\right)$ is a tight sequence of probabilities; where $\pi_{\tilde{f}_k} : \mathcal{M}_F(\overline{\mathbb{R}^d}) \to \overline{\mathbb{R}}$ is defined by $\pi_{\tilde{f}_k}(\mu) = \mu(\tilde{f}_k)$ for $\mu \in \mathcal{M}_F(\overline{\mathbb{R}^d})$.*

In the remaining of this chapter, because of the definition of the distance $d_{\mathcal{M}}$, we choose $(\tilde{f}_k)_{k\geq 0}$ to be defined as follows: $\tilde{f}_0 \equiv 1$, and $\tilde{f}_k$ $(k \geq 1)$ is chosen so that $\tilde{f}_k\big|_{\mathbb{R}}$ is a dense sequence in $\mathcal{C}_b^6(\mathbb{R})$, the space of six times differentiable continuous functions on $\mathbb{R}$, vanishing at infinity with continuous partial derivatives up to and including the sixth order. According to Theorem 5.1.3, it suffices to prove the tightness result for $\{\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n\}_n$. We will make use of the following criteria, which can be found in [35], to show that $\{\pi_{\tilde{f}_k}U^n\}_n = \{U^n(\tilde{f}_k)\}_n$ is tight, and then the tightness of $\{\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n\}$ follows since Theorem 5.1.3.

Theorem 5.1.4 (Kurtz's criteria of relative compactness)**.** *Let (E,d) be a separable and complete metric space and let $\{X^n\}_{n\in\mathbb{N}}$ be a sequence of processes with sample paths in $D_E[0,\infty)$. Suppose that for every $\eta > 0$ and rational t, there exists a compact set $\Gamma_{\eta,t}$ such that*

$$\sup_n \mathbb{P}(X_t^n \notin \Gamma_{\eta,t}) \leq \eta. \tag{5.6}$$

Then $\{X^n\}_{n\in\mathbb{N}}$ is relatively compact if and only if the following conditions hold:

- *For each $T' > 0$, there exists $\beta > 0$ and a family $\{\gamma^n(\Delta) : 0 < \Delta < 1\}$ of non-negative random variables*

$$\tilde{\mathbb{E}}\left[\left(1 \wedge d(X_{t+u}^n, X_t^n)\right)^\beta \left(1 \wedge d(X_t^n, X_{t-v}^n)\right)^\beta |\mathcal{F}_t\right] \leq \tilde{\mathbb{E}}\left[\gamma^n(\Delta)|\mathcal{F}_t\right] \tag{5.7}$$

  for $0 \leq t \leq T'$, $0 \leq u \leq \Delta$ and $0 \leq v \leq \Delta \wedge t$;

- *For $\gamma^n(\Delta)$, we have*

$$\lim_{\Delta\to 0}\limsup_{n\to\infty} \tilde{\mathbb{E}}\left[\gamma^n(\Delta)\right] = 0; \tag{5.8}$$

- *At the initial time*

$$\lim_{\Delta\to 0}\limsup_{n\to\infty} \tilde{\mathbb{E}}\left[\left(1 \wedge d(X_\Delta^n, X_0^n)\right)^\beta\right] = 0. \tag{5.9}$$

To justify (5.6), we need to prove the following lemma:

Lemma 5.1.5. *For all $\eta > 0$, there exists a constant β such that for the associated probabilities $\{\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n\}$ of $\{\pi_{\tilde{f}_k}U^n\}$ and $A = \{x \in D_{\overline{\mathbb{R}}}[0,T] : \sup_{t\in[0,T]}|x(t)| > \beta\}$, we have*

$$\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n(A) \leq \eta. \tag{5.10}$$

Proof. Note that $\pi_{\tilde{f}_k}U_t^n = U_t^n(\tilde{f}_k)$, so that

$$\begin{aligned}
\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n(A) &= \tilde{\mathbb{P}}_n\pi_{\tilde{f}_k}^{-1}(A) \\
&= \tilde{\mathbb{P}}_n\left(U^n \in D_{\mathcal{M}_F}[0,T] : \sup_t |U_t^n(\tilde{f}_k)| > \beta\right) \\
&= \tilde{\mathbb{P}}_n\left(U^n \in D_{\mathcal{M}_F}[0,T] : \sup_t |\sqrt{n}(\rho_t^n(\tilde{f}_k) - \rho_t(\tilde{f}_k))| > \beta\right) \\
&\leq \frac{\Lambda_T^n(\tilde{f}_k)}{\beta^2},
\end{aligned} \tag{5.11}$$

where $\Lambda_T^n(\tilde{f}_k) = \tilde{\mathbb{E}}\left[\sup_t \left(\sqrt{n}(\rho_t^n(\tilde{f}_k) - \rho_t(\tilde{f}_k))\right)^2\right]$.

It suffices to show that $\Lambda_T^n(\tilde{f}_k)$ is bounded above by a constant independent of n, which is an immediate consequence of Jensen's inequality and Theorem 4.2.11. Then we choose

$$\beta^2 = \frac{\eta}{\Lambda_T^n(\tilde{f}_k)}$$

and the proof is complete. ■

In order to prove the tightness of $\{U^n(\tilde{f}_k)\}$, we still need to show that (5.7), (5.8) and (5.9) are satisfied by $\{U^n(\tilde{f}_k)\}$. We prove these by showing that each of the increments of the process appearing on the right hand side of (5.4) satisfy similar bounds.

In the following we will choose Δ to be sufficiently small. To be specific, we let $\Delta < \frac{\delta}{2}$, where δ is the time length between two resampling events. This ensures that either $[t-\Delta, t]$ or $[t, t+\Delta]$ does not contain a resampling event, in other words, there is at most one resampling event in $[t, t+u]$ and $[t-v, t]$, where $0 \le u \le \Delta$ and $0 \le v \le \Delta \wedge t$.

If the resampling happens only in the interval $[t-v, t]$, and obtain

$$\tilde{\mathbb{E}}\left[\left(1 \wedge d(X_{t+u}^n, X_t^n)\right)^\beta \left(1 \wedge d(X_t^n, X_{t-v}^n)\right)^\beta |\mathcal{F}_t\right] \le \tilde{\mathbb{E}}\left[\left(1 \wedge d(X_{t+u}^n, X_t^n)\right)^\beta |\mathcal{F}_t\right].$$

Therefore in order to determine $\gamma^n(\Delta)$ and shows that (5.7) is satisfied by $\{U^n(\tilde{f}_k)\}_n$, it suffices to find an appropriate $\gamma^n(\Delta)$ and show

$$\tilde{\mathbb{E}}\left[\left(1 \wedge d(U_{t+u}^n(\tilde{f}_k), U_t^n(\tilde{f}_k)\right)^2 |\mathcal{F}_t\right] \le \tilde{\mathbb{E}}\left[\gamma^n(\Delta)|\mathcal{F}_t\right]. \tag{5.12}$$

This will be done in the following proposition.

Proposition 5.1.6. *Let $k \in \mathbb{N}$, and we further assume that $\tilde{f}_k \in C_b^6(\mathbb{R})$, and Assumption (A) holds. Let the length bewtween two resampling events δ to be fixed and let $\alpha \propto \frac{1}{\sqrt{n}}$. Define the family $\{\gamma_u^n(\Delta) : 0 < \Delta < 1\}$ of non-negative random variables*

$$\gamma^n(\Delta) \triangleq 3n\Delta^2 \sup_{s\in[t,t+u]} \left(\rho_s^n(A\tilde{f}_k) - \rho_s(A\tilde{f}_k)\right)^2 + 3n\Delta \sup_{s\in[t,t+u]} \left(\rho_s^n(h\tilde{f}_k) - \rho_s(h\tilde{f}_k)\right)^2$$
$$+\frac{3\Delta}{n} C_\gamma \|\tilde{f}_k\|_{6,\infty}^2 \sum_{j=1}^n \sup_{s\in[t,t+u]} \left(\xi_{i\delta}^n a_j^n(s)\right)^2, \tag{5.13}$$

where C_γ is a constant independent of n. By Theorem 4.2.11, we know that

$$\sup_{s\in[t,t+u]} n\left(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k)\right)^2 \text{ and } \sup_{s\in[t,t+u]} n\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)^2$$

are bounded and independent of Δ. Then we have

$$\tilde{\mathbb{E}}\left[1\wedge d(U_{t+u}^n(\tilde{f}_k),U_t^n(\tilde{f}_k))^2|\mathcal{F}_t\right]\le \tilde{\mathbb{E}}\left[\gamma^n(\Delta)|\mathcal{F}_t\right]. \tag{5.14}$$

Proof. Bearing in mind that there is no resampling event within $[t,t+u]$, thus $[(t+u)/\delta]=[t/\delta]$ and

$$M_{[(t+u)/\delta]}^{n,\tilde{f}_k}-M_{[t/\delta]}^{n,\tilde{f}_k}=0.$$

Therefor we have that

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[1\wedge d(U_{t+u}^n(\tilde{f}_k),U_t^n(\tilde{f}_k))^2\middle|\mathcal{F}_t\right]\\
\le{}&\tilde{\mathbb{E}}\left[d(U_{t+u}^n(\tilde{f}_k),U_t^n(\tilde{f}_k))^2\middle|\mathcal{F}_t\right]\\
={}&\tilde{\mathbb{E}}\left[|U_{t+u}^n(\tilde{f}_k)-U_t^n(\tilde{f}_k)|^2\middle|\mathcal{F}_t\right]\\
={}&\tilde{\mathbb{E}}\left[\left|\sqrt{n}\left(\left(\rho_{t+u}^n(\tilde{f}_k)-\rho_{t+u}(\tilde{f}_k)\right)-\left(\rho_t^n(\tilde{f}_k)-\rho_t(\tilde{f}_k)\right)\right)\right|^2\middle|\mathcal{F}_t\right]\\
\le{}&3n\Bigg\{\tilde{\mathbb{E}}\left[\left(\int_t^{t+u}(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k))ds\right)^2\middle|\mathcal{F}_t\right]\\
&+\tilde{\mathbb{E}}\left[\left(\int_t^{t+u}(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k))dY_s\right)^2\middle|\mathcal{F}_t\right]\\
&+\frac{1}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^n\int_t^{t+u}\xi_{i\delta}^n a_j^n(s)\left[R_{s,j}^1(\tilde{f}_k)ds+R_{s,j}^2(\tilde{f}_k)dY_s+R_{s,j}^3(\tilde{f}_k)dV_s^{(j)}\right]\right)^2\middle|\mathcal{F}_t\right]\Bigg\}.
\end{aligned} \tag{5.15}$$

We examine each of the terms in (5.15) and observe the following:

For the first term in (5.15), by Jensen's inequality, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sqrt{n}\int_t^{t+u}(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k))ds\right)^2\middle|\mathcal{F}_t\right]\\
\leq&\tilde{\mathbb{E}}\left[nu\int_t^{t+u}\left(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k)\right)^2ds\middle|\mathcal{F}_t\right]\\
=&nu\int_t^{t+u}\tilde{\mathbb{E}}\left[\left(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]ds\\
\leq&nu\int_t^{t+u}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]ds\\
=&nu^2\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\rho_s^n(A\tilde{f}_k)-\rho_s(A\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]. \qquad (5.16)
\end{aligned}$$

For the second term in (5.15),

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\int_t^{t+u}\sqrt{n}\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)dY_s\right)^2\middle|\mathcal{F}_t\right]\\
\leq&\tilde{\mathbb{E}}\left[\left\langle\int_t^{\cdot}\sqrt{n}\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)dY_s\right\rangle_{t+u}\middle|\mathcal{F}_t\right]\\
=&n\tilde{\mathbb{E}}\left[\int_t^{t+u}\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)^2ds\middle|\mathcal{F}_t\right]\\
=&n\int_t^{t+u}\tilde{\mathbb{E}}\left[\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]ds\\
\leq&n\int_t^{t+u}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]ds\\
=&un\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\rho_s^n(h\tilde{f}_k)-\rho_s(h\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]. \qquad (5.17)
\end{aligned}$$

For the remaining terms in (5.15), note that

$$R^1_{s,j}(\tilde{f}_k)\leq C_1\alpha\delta\|\tilde{f}_k\|_{6,\infty}\leq\frac{C_1}{n}\|\tilde{f}_k\|_{6,\infty},$$

we then have

$$
\begin{aligned}
& n\frac{1}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\int_t^{t+u}\xi_{i\delta}^n a_j^n(s)\left[R_{s,j}^1(\tilde{f}_k)ds\right]\right)^2\middle|\mathcal{F}_t\right] \\
\leq & \frac{1}{n}\tilde{\mathbb{E}}\left[u\int_t^{t+u}\left(\sum_{j=1}^{n}\xi_{i\delta}^n a_j^n(s)R_{s,j}^1(\tilde{f}_k)\right)^2 ds\middle|\mathcal{F}_t\right] \\
\leq & \frac{1}{n}\tilde{\mathbb{E}}\left[u\int_t^{t+u} n\sum_{j=1}^{n}\left[\xi_{i\delta}^n a_j^n(s)R_{s,j}^1(\tilde{f}_k)\right]^2 ds\middle|\mathcal{F}_t\right] \\
= & u\sum_{j=1}^{n}\int_t^{t+u}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)R_{s,j}^1(\tilde{f}_k)\right)^2\middle|\mathcal{F}_t\right]ds \\
\leq & u^2\frac{C_1^2}{n}\|\tilde{f}_k\|_{6,\infty}^2\sum_{j=1}^{n}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)\right)^2\middle|\mathcal{F}_t\right] \\
\leq & u\frac{C_1^2}{n}\|\tilde{f}_k\|_{6,\infty}^2\sum_{j=1}^{n}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)\right)^2\middle|\mathcal{F}_t\right];
\end{aligned}
\tag{5.18}
$$

and also note that

$$
R_{s,j}^2(\tilde{f}_k)\leq C_2\alpha\delta\|\tilde{f}_k\|_{4,\infty}\leq\frac{C_2}{n}\|\tilde{f}_k\|_{4,\infty},
$$

then we have

$$
\begin{aligned}
& n\frac{1}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\int_t^{t+u}\xi_{i\delta}^n a_j^n(s)R_{s,j}^2(\varphi)dY_s\right)^2\middle|\mathcal{F}_t\right] \\
\leq & \frac{1}{n}\tilde{\mathbb{E}}\left[n\sum_{j=1}^{n}\left(\int_t^{t+u}\xi_{i\delta}^n a_j^n(s)R_{s,j}^2(\varphi)dY_s\right)^2\middle|\mathcal{F}_t\right] \\
\leq & \sum_{j=1}^{n}\tilde{\mathbb{E}}\left[\int_t^{t+u}\left[\xi_{i\delta}^n a_j^n(s)R_{s,j}^2(\varphi)\right]^2 ds\middle|\mathcal{F}_t\right] \\
= & \sum_{j=1}^{n}\int_t^{t+u}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)R_{s,j}^2(\varphi)\right)^2\middle|\mathcal{F}_t\right]ds \\
\leq & \sum_{j=1}^{n}\int_t^{t+u}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)R_{s,j}^2(\varphi)\right)^2\middle|\mathcal{F}_t\right]ds \\
= & u\frac{C_2^2}{n}\|\tilde{f}_k\|_{4,\infty}^2\sum_{j=1}^{n}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi_{i\delta}^n a_j^n(s)\right)^2\middle|\mathcal{F}_t\right];
\end{aligned}
\tag{5.19}
$$

and finally since

$$R^3_{s,j}(\tilde{f}_k) \leq (C_0 + C_3\alpha\delta)\|\tilde{f}_k\|_{5,\infty} \leq (C_0 + C_3)\|\tilde{f}_k\|_{5,\infty},$$

we have that

$$\begin{aligned}
& n\frac{1}{n^2}\tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n}\int_t^{t+u} \xi^n_{i\delta}a^n_j(s)R^3_{s,j}(\varphi)dV^{(j)}_s\right)^2\middle|\mathcal{F}_t\right] \\
\leq & \frac{1}{n}\tilde{\mathbb{E}}\left[\sum_{j=1}^{n}\left(\int_t^{t+u} \xi^n_{i\delta}a^n_j(s)R^3_{s,j}(\varphi)dV^{(j)}_s\right)^2\middle|\mathcal{F}_t\right] \\
\leq & \frac{1}{n}\tilde{\mathbb{E}}\left[\sum_{j=1}^{n}\int_t^{t+u}\left(\xi^n_{i\delta}a^n_j(s)R^3_{s,j}(\varphi)\right)^2 ds\middle|\mathcal{F}_t\right] \\
= & \frac{1}{n}\sum_{j=1}^{n}\int_t^{t+u}\tilde{\mathbb{E}}\left[\left(\xi^n_{i\delta}a^n_j(s)R^3_{s,j}(\varphi)\right)^2\middle|\mathcal{F}_t\right] ds \\
\leq & \frac{1}{n}\sum_{j=1}^{n}\int_t^{t+u}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi^n_{i\delta}a^n_j(s)R^3_{s,j}(\varphi)\right)^2\middle|\mathcal{F}_t\right] ds \\
= & \frac{u}{n}(C_0 + C_3)^2\|\tilde{f}_k\|^2_{5,\infty}\sum_{j=1}^{n}\sup_{s\in[t,t+u]}\tilde{\mathbb{E}}\left[\left(\xi^n_{i\delta}a^n_j(s)\right)^2\middle|\mathcal{F}_t\right]. \qquad (5.20)
\end{aligned}$$

Therefore, considering the bounds in the right hand sides of (5.16), (5.17), (5.18), (5.19), and (5.20); we can define $\gamma^n(\Delta)$ as in (5.13) by letting

$$C_\gamma = C_1^2 + C_2^2 + (C_0 + C_3)^2.$$

By virtue of (5.15), we know that (5.14) is satisfied. ■

The above discussion defines $\gamma^n(\Delta)$ and shows that (5.7) is satisfied for $\{U^n(\tilde{f}_k)\}_n$.

The following proposition shows that $\gamma^n(\Delta)$ defined in (5.13) satisfies (5.8).

Proposition 5.1.7. $\gamma^n(\Delta)$ *defined in* (5.13) *has the following property*

$$\lim_{\Delta\to 0}\limsup_{n\to\infty}\tilde{\mathbb{E}}\left[\gamma^n(\Delta)\right] = 0. \qquad (5.21)$$

Proof. We show this by looking at the expectation of each term in (5.13).

For the first term, by Theorem 4.2.11

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[n\Delta^2 \sup_{s\in[t,t+u]} \left(\rho_s^n(A\tilde{f}_k) - \rho_s(A\tilde{f}_k\right)^2\right] \\
=&n\Delta^2\tilde{\mathbb{E}}\left[\sup_{s\in[t,t+u]} \left(\rho_s^n(A\tilde{f}_k) - \rho_s(A\tilde{f}_k\right)^2\right] \\
\leq&n\Delta^2\frac{c^T}{n}\left\|A\tilde{f}_k\right\|^2_{m+2,\infty} = \Delta^2 c^T \left\|A\tilde{f}_k\right\|^2_{m+2,\infty} \to 0, \quad \text{as } \Delta \to 0. \qquad (5.22)
\end{aligned}$$

Similarly, for the second term,

$$\tilde{\mathbb{E}}\left[n\Delta \sup_{s\in[t,t+u]} \left(\rho_s^n(h\tilde{f}_k) - \rho_s(h\tilde{f}_k)\right)^2\right] \leq \Delta\tilde{c}^T \left\|h\tilde{f}_k\right\|^2_{m+2,\infty} \to 0, \quad \text{as } \Delta \to 0. \quad (5.23)$$

For the remaining term, again note that $(\alpha\delta)^2 \sim 1/n$, and

$$\tilde{\mathbb{E}}\left[\sum_{j=1}^n \sup_{s\in[t,t+u]} \left(\xi^n_{i\delta} a^n_j(s)\right)^2\right] = \sum_{j=1}^n \tilde{\mathbb{E}}\left[\sup_{s\in[t,t+u]} \left(\xi^n_{i\delta} a^n_j(s)\right)^2\right] \leq nc_2^{t,2}. \qquad (5.24)$$

Thus

$$\begin{aligned}
&\frac{\Delta}{n}C_{\gamma^n}\|\tilde{f}_k\|^2_{6,\infty}\sum_{j=1}^n \tilde{\mathbb{E}}\left[\sup_{s\in[t,t+u]} \left(\xi^n_{i\delta} a^n_j(s)\right)^2\right] \\
\leq&\frac{\Delta}{n}\|\tilde{f}_k\|^2_{6,\infty} nc_2^{t,2} = \Delta c_2^{t,2}\|\tilde{f}_k\|^2_{6,\infty} \to 0, \quad \text{as } \Delta \to 0. \qquad (5.25)
\end{aligned}$$

This completes the proof. ■

The following proposition shows that (5.9) holds for $\{U^n(\tilde{f}_k)\}$.

Proposition 5.1.8. *For each $k \in \mathbb{N}$, we have*

$$\lim_{\Delta\to 0}\limsup_{n\to\infty} \tilde{\mathbb{E}}\left[\left(1 \wedge d(U^n_\Delta(\tilde{f}_k), U^n_0(\tilde{f}_k))\right)^2\right] = 0. \qquad (5.26)$$

Proof. The result follows immediately by continuity of $\{U^n(\tilde{f}_k)\}_n$ at the initial time 0. ■

Finally, Lemma 5.1.5, Proposition 5.1.6, Proposition 5.1.7, together with Proposition 5.1.8 state that all the conditions in Theorem 5.1.4 are satisfied. Then by Theorem 5.1.4 we know that $\{\pi_{\tilde{f}_k} U^n\}_n$ is tight, which implies that $\{\pi_{\tilde{f}_k}\tilde{\mathbb{P}}_n\}_n$ forms a tight sequence on $\mathcal{P}(D_{\mathbb{R}}[0,T])$; then by Theorem 5.1.3 we know $\{\tilde{\mathbb{P}}_n\}$ forms a tight sequence on $\mathcal{P}(D_{\mathcal{M}_F(\mathbb{R}^d)}[0,T])$. By definition we can then conclude the following tightness result.

Theorem 5.1.9. *The measure-valued processes $\{U_t^n : t \in [0,T]\}_{n\geq 1}$ forms a tight sequence.*

Remark 5.1.10. *If we assume that the resampling happens only in $[t, t+u]$, then by exactly the same discussion as above (except that we replace $s \in [t, t+u]$ by $s \in [t-v, u]$), we can also obtain the tightness for the process $\{U_t^n\}_{n\geq 1}$.*

5.2 Convergence in Distribution

In this section we show that $\{U^n\}_n$ converges in distribution to a uniquely determined process U. The strategy of the proof of the convergence in distribution is as follows: Since the sequence of the measure-valued process $\{U^n\}_n$ is tight, then any subsequence $\{U^{n_k}\}_k$ of $\{U^n\}_n$ contains a convergent sub-subsequence $\{U^{n_{k_l}}\}_l$. We will prove that any convergent subsequence has a weak limit U which is the unique solution of (5.31). This ensures that the entire sequence $\{U^n\}_n$ is convergent and its weak limit is the solution U of (5.31).

We need the following preliminary result.

Lemma 5.2.1. *Let $\varphi \in C_b^m(\overline{\mathbb{R}})$ $(m \geq 6)$ be a test function, and define the measure-valued processes*

$$\tilde{\rho}_t^{n.1} \triangleq \frac{1}{n}\sum_{j=1}^n \xi_{i\delta}^n a_j^n(t)\delta_{v_j^n(t)} = \sum_{j=1}^n \xi_t^n \bar{a}_j^n(t)\delta_{v_j^n(t)},$$
$$\tilde{\rho}_t^{n.2} \triangleq \frac{1}{n}\sum_{j=1}^n \left\{\xi_{i\delta}^n a_j^n(t)\right\}^2 \delta_{v_j^n(t)} = n\sum_{j=1}^n \left\{\xi_t^n \bar{a}_j^n(t)\right\}^2 \delta_{v_j^n(t)}. \tag{5.27}$$

then for any $t \in [0,T]$,

$$\tilde{\rho}_t^{n,1} \to \tilde{\rho}_t^1, \quad \tilde{\rho}_t^{n,2} \to \tilde{\rho}_t^2, \qquad \tilde{\mathbb{P}} - a.s.,$$

where $\tilde{\rho}^1$ and $\tilde{\rho}^2$ are two measure-valued processes satisfying, for any $\varphi \in \mathcal{D}(A)$,

$$\begin{aligned}\tilde{\rho}_t^1(\varphi) = \pi_0(\varphi) + \int_0^t &\Big\{\rho_s(A\varphi) + \pi_s(h)\left[\pi_s(h)\rho_s(\varphi) - \rho_s(h\varphi)\right] \\ &\quad + \rho_s(h)\left[\pi_s(h\varphi) - \pi_s(h)\pi_s(\varphi)\right]\Big\}ds \\ &+ \int_0^t \Big\{\rho_s(h\varphi) - \pi_s(h)\rho_s(\varphi) + \pi_s(\varphi)\rho_s(h)\Big\}dY_s;\end{aligned} \tag{5.28}$$

$$
\begin{aligned}
\tilde{\rho}_t^2(\varphi) = \pi_0(\varphi) + \int_0^t \Big\{ &\rho_s(\mathbf{1})\rho_s(A\varphi) - \left[\rho_s(\mathbf{1})\rho_s(h\varphi) - \rho_s(h)\rho_s(\varphi)\right]\pi_s(h) \\
&+ \pi_s(\varphi)(\rho_s(h))^2 + 2\left[\rho_s(h)\rho_s(h\varphi) - (\rho_s(h))^2\pi_s(\varphi)\right] \Big\} ds \\
+ \int_0^t &\left\{\rho_s(\mathbf{1})\rho_s(h\varphi) + \rho_s(h)\rho_s(\varphi)\right\} dY_s. \qquad (5.29)
\end{aligned}
$$

Proof. See Appendix C.2. ■

Proposition 5.2.2. *For any $\varphi \in C_b^6(\overline{\mathbb{R}})$, let Λ^φ be the process defined by*

$$
\begin{aligned}
\Lambda_t^\varphi = &\sum_{i=1}^{[t/\delta]} \rho_{i\delta}(\mathbf{1})\sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\,\Upsilon_i + c_\omega \int_0^t \tilde{\rho}_s^1(\Psi\varphi)ds \\
&+ c_\omega \int_0^t \left(\tilde{\rho}_s^1(h\varphi'' - (h\varphi)'')\right) dB_s^{(2)} + \int_0^t \sqrt{\tilde{\rho}_s^2\left((\sigma\varphi')^2\right)}dB_s^{(3)} \qquad (5.30)
\end{aligned}
$$

for $t \in [0,T]$. In (5.30), $\{\Upsilon_i\}_{i\in\mathbb{N}}$ is a sequence of independent identically distributed, standard normal random variables, and $\left\{\sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\,\Upsilon_i\right\}_i$ are mutually independent given the σ-algebra $\mathcal{Y}$. c_ω is a constant independent of n, and the operator Ψ is defined by

$$
\Psi\varphi = \frac{f\varphi'''}{2} + \frac{\sigma\varphi^{(4)}}{4} - \frac{3(A\varphi)''}{2}.
$$

$B^{(2)}$ and $B^{(3)}$ are two independent standard Brownian motion and both independent of the observation Y.

If U is a $\mathcal{D}_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$-valued process such that for $\varphi \in \mathcal{C}_b^6(\overline{\mathbb{R}})$

$$
U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)ds + \int_0^t U_s(h\varphi)dY_s + \Lambda_t^\varphi, \qquad (5.31)
$$

then U is pathwise unique. That is, for any two strong solutions U and $\bar{U}$ of (5.31) and with common initial value i.e. $\mathbb{P}\left[U_0 = \bar{U}_0\right] = 1$, the two processes are indistinguishable, i.e. $\mathbb{P}\left[U_t = \bar{U}_t; t \in [0,T]\right] = 1$.

Proof. Firstly, it can be seen that the first three terms of (5.30) are martingales while the final term is not a martingale.

Suppose there exists two solutions U^1 and U^2 of (5.31). Then take $\varphi \in C_b^6(\mathbb{R})$, we have

$$
U_t^1(\varphi) = U_0^1(\varphi) + \int_0^t U_s^1(A\varphi)ds + \int_0^t U_s^1(h\varphi)dY_s + \Lambda_t^\varphi, \qquad (5.32)
$$

$$U_t^2(\varphi) = U_0^2(\varphi) + \int_0^t U_s^2(A\varphi)ds + \int_0^t U_s^2(h\varphi)dY_s + \Lambda_t^{\varphi}. \tag{5.33}$$

For $i, j = \{1, 2\}$ let $\bar{U}^{ij}(\varphi_1, \varphi_2) \triangleq \tilde{\mathbb{E}}[U^i(\varphi_1)U^j(\varphi_2)]$, for $\varphi_1, \varphi_2 \in C_b^6(\overline{\mathbb{R}})$.

By Itô's formula we have

$$\begin{aligned}
\bar{U}^{12}(\varphi_1, \varphi_2) = & \int_0^t \bar{U}^{12}(\varphi_1, A\varphi_2)ds + \int_0^t \bar{U}^{12}(A\varphi_1, \varphi_2)ds + \int_0^t \bar{U}^{12}(h\varphi_1, h\varphi_2)ds \\
& + \int_0^t \tilde{\mathbb{E}}\Big[U_s^1(\varphi_1)\tilde{\rho}_s^1(\Psi\varphi_2) + U_s^2(\varphi_2)\tilde{\rho}_s^1(\Psi\varphi_1)\Big]ds \\
& + \int_0^t \tilde{\mathbb{E}}\left[\sqrt{\tilde{\rho}_s^2((\sigma\varphi_1)^2)\tilde{\rho}_s^2((\sigma\varphi_2)^2)} + \tilde{\rho}_s^1(h\varphi_1'' - (h\varphi_1)'')\tilde{\rho}_s^1(h\varphi_2'' - (h\varphi_2)'')\right] ds \\
& + \tilde{\mathbb{E}}\left[\sum_{i=0}^{[t/\delta]} \tilde{\mathbb{E}}\left[(\rho_{i\delta}(\mathbf{1}))^2 \left(\pi_{i\delta-}(\varphi_1\varphi_2) - \pi_{i\delta-}(\varphi_1)\pi_{i\delta-}(\varphi_2)\right) \big| \mathcal{F}_{i\delta-}\right]\right];
\end{aligned}$$

$$\begin{aligned}
\bar{U}^{11}(\varphi_1, \varphi_2) = & \int_0^t \bar{U}^{11}(\varphi_1, A\varphi_2)ds + \int_0^t \bar{U}^{11}(A\varphi_1, \varphi_2)ds + \int_0^t \bar{U}^{11}(h\varphi_1, h\varphi_2)ds \\
& + \int_0^t \tilde{\mathbb{E}}\Big[U_s^1(\varphi_1)\tilde{\rho}_s^1(\Psi\varphi_2) + U_s^1(\varphi_2)\tilde{\rho}_s^1(\Psi\varphi_1)\Big]ds \\
& + \int_0^t \tilde{\mathbb{E}}\left[\sqrt{\tilde{\rho}_s^2((\sigma\varphi_1)^2)\tilde{\rho}_s^2((\sigma\varphi_2)^2)} + \tilde{\rho}_s^1(h\varphi_1'' - (h\varphi_1)'')\tilde{\rho}_s^1(h\varphi_2'' - (h\varphi_2)'')\right] ds \\
& + \tilde{\mathbb{E}}\left[\sum_{i=0}^{[t/\delta]} \tilde{\mathbb{E}}\left[(\rho_{i\delta}(\mathbf{1}))^2 \left(\pi_{i\delta-}(\varphi_1\varphi_2) - \pi_{i\delta-}(\varphi_1)\pi_{i\delta-}(\varphi_2)\right) \big| \mathcal{F}_{i\delta-}\right]\right];
\end{aligned}$$

$$\begin{aligned}
\bar{U}^{21}(\varphi_1, \varphi_2) = & \int_0^t \bar{U}^{21}(\varphi_1, A\varphi_2)ds + \int_0^t \bar{U}^{21}(A\varphi_1, \varphi_2)ds + \int_0^t \bar{U}^{21}(h\varphi_1, h\varphi_2)ds \\
& + \int_0^t \tilde{\mathbb{E}}\Big[U_s^2(\varphi_1)\tilde{\rho}_s^1(\Psi\varphi_2) + U_s^1(\varphi_2)\tilde{\rho}_s^1(\Psi\varphi_1)\Big]ds \\
& + \int_0^t \tilde{\mathbb{E}}\left[\sqrt{\tilde{\rho}_s^2((\sigma\varphi_1)^2)\tilde{\rho}_s^2((\sigma\varphi_2)^2)} + \tilde{\rho}_s^1(h\varphi_1'' - (h\varphi_1)'')\tilde{\rho}_s^1(h\varphi_2'' - (h\varphi_2)'')\right] ds \\
& + \tilde{\mathbb{E}}\left[\sum_{i=0}^{[t/\delta]} \tilde{\mathbb{E}}\left[(\rho_{i\delta}(\mathbf{1}))^2 \left(\pi_{i\delta-}(\varphi_1\varphi_2) - \pi_{i\delta-}(\varphi_1)\pi_{i\delta-}(\varphi_2)\right) \big| \mathcal{F}_{i\delta-}\right]\right];
\end{aligned}$$

and

$$
\begin{aligned}
\bar{U}^{22}(\varphi_1,\varphi_2) = & \int_0^t \bar{U}^{22}(\varphi_1, A\varphi_2)ds + \int_0^t \bar{U}^{22}(A\varphi_1,\varphi_2)ds + \int_0^t \bar{U}^{22}(h\varphi_1, h\varphi_2)ds \\
& + \int_0^t \tilde{\mathbb{E}}\Big[U_s^2(\varphi_1)\tilde{\rho}_s^1(\Psi\varphi_2) + U_s^2(\varphi_2)\tilde{\rho}_s^1(\Psi\varphi_1)\Big]ds \\
& + \int_0^t \tilde{\mathbb{E}}\left[\sqrt{\tilde{\rho}_s^2((\sigma\varphi_1)^2)\tilde{\rho}_s^2((\sigma\varphi_2)^2)} + \tilde{\rho}_s^1(h\varphi_1'' - (h\varphi_1)'')\tilde{\rho}_s^1(h\varphi_2'' - (h\varphi_2)'')\right] ds \\
& + \tilde{\mathbb{E}}\left[\sum_{i=0}^{[t/\delta]} \tilde{\mathbb{E}}\left[(\rho_{i\delta}(\mathbf{1}))^2\left(\pi_{i\delta-}(\varphi_1\varphi_2) - \pi_{i\delta-}(\varphi_1)\pi_{i\delta-}(\varphi_2)\right)\big|\mathcal{F}_{i\delta-}\right]\right].
\end{aligned}
$$

Let

$$
v_t = \left(\bar{U}_t^{12} - \bar{U}_t^{11}\right) + \left(\bar{U}_t^{21} - \bar{U}_t^{22}\right), \tag{5.34}
$$

it then follows that

$$
v_t(\varphi_1,\varphi_2) = \int_0^t v_s(\varphi_1, A\varphi_2)ds + \int_0^t v_s(A\varphi_1,\varphi_2)ds + \int_0^t v_s(h\varphi_1, h\varphi_2)ds; \tag{5.35}
$$

and $v_0(\varphi_1,\varphi_2) = 0$.

It follows by Theorem 2.21(i) and Remark 3.4 in [55] that (5.35) has a unique solution and since (5.35) is a homogeneous equation beginning at 0. Then we have $v_t(\varphi_1,\varphi_2) \equiv 0$, which implies

$$
\left(\bar{U}_t^{11} - \bar{U}_t^{12}\right) + \left(\bar{U}_t^{22} - \bar{U}_t^{21}\right) = 0,
$$

that is to say, for $\varphi_1 = \varphi_2 = \varphi$

$$
\begin{aligned}
& \tilde{\mathbb{E}}\left[U_t^1(\varphi_1)U_t^1(\varphi) - U_t^1(\varphi_1)U_t^2(\varphi)\right] + \tilde{\mathbb{E}}\left[U_t^2(\varphi_1)U_t^2(\varphi) - U_t^2(\varphi_1)U_t^1(\varphi)\right] \\
= & \tilde{\mathbb{E}}\left[\left(U_t^1(\varphi) - U_t^2(\varphi)\right)^2\right] = 0;
\end{aligned} \tag{5.36}
$$

and thus $U^1(\varphi) = U^2(\varphi)$ for $\varphi \in C_b^6(\overline{\mathbb{R}})$, which in turn implies that the solution U of (5.31) is unique (See Exercise 4.1 in [3]). ■

The following Theorem 5.2.3 states that unique solution $\{U\}$ of (5.31) is indeed the weak limit of the measure-valued process $\{U^n\}_n$, in other words, $\{U^n\}_n$ converges in distribution to $\{U\}$.

Theorem 5.2.3. *$\{U^n\}_n$ converges in distribution to a unique $\mathcal{D}_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$-valued process U such that for $\varphi \in C_b^6(\overline{\mathbb{R}})$,*

$$U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)ds + \int_0^t U_s(h\varphi)dY_s + \Lambda_t^\varphi, \tag{5.37}$$

where Λ_t^φ is defined as in (5.30).

Proof. From Proposition 5.3.20 in [44] and its extension to stochastic partial differential equation and infinitely dimensional stochastic differential equations, it follows that for solutions of stochastic partial differential equations, pathwise uniqueness implies uniqueness in law. The extension to stochastic PDE was done by Ondreját (see [57]) and Röckner, Schmuland and Zhang (see [61]).

Thus by Proposition 5.2.2 the solution U of (5.31) is unique in distribution.

Now let $\{U^{n_k}\}_k$ be any convergent (in distribution) subsequence of $\{U^n\}_n$ to a process U. We then verify that this process U solves (5.31), and then the uniqueness of solution of (5.31) implies that the original sequence $\{U^n\}_n$ converges to U as well.

Bearing in mind that U^{n_k} satisfies (5.4), it then essentially suffices to show that Λ_t^φ in (5.37), which is given by the weak limits of $\sqrt{n}M_{[t/\delta]}^{n,\varphi}$ and $\sqrt{n}B_t^{n,\varphi}$ in (5.4), does satisfy (5.30). We first denote by

$$\bar{\Lambda}_t^\varphi \triangleq \Lambda_t^\varphi - \int_0^t \tilde{\rho}_s^1(\Psi\varphi)ds$$

the martingale part of Λ_t^φ. Then we only need to show that $\bar{\Lambda}^\varphi$ has the quadratic variation which is the same as that of Λ^φ in (5.30). In order to do so, we show that for all $d, d' \geq 0$, $0 \leq t_1 < t_2 < \cdots < t_d \leq s \leq T$, $0 \leq t'_1 < t'_2 < \cdots < t'_{d'} \leq t \leq T$, continuous bounded functions $\alpha_1, \ldots, \alpha_d$ on $\mathcal{M}_F(\overline{\mathbb{R}})$ and continuous functions $\alpha'_1, \ldots, \alpha'_{d'}$ on $\overline{\mathbb{R}}$; we have:

$$\tilde{\mathbb{E}}\left[\left(\bar{\Lambda}_t^\varphi - \bar{\Lambda}_s^\varphi\right)\prod_{i=1}^d \alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right] = 0, \tag{5.38}$$

and

$$\begin{aligned}\tilde{\mathbb{E}}\Bigg[\Bigg((\bar{\Lambda}_t^\varphi - \bar{\Lambda}_s^\varphi)^2 - \int_s^t \left\{\tilde{\rho}_r^2\left((\sigma\varphi')^2\right) + \left(\tilde{\rho}_r^1\left(h\varphi'' - (h\varphi)''\right)\right)^2\right\}dr \\ - \sum_{i=[s/\delta]+1}^{[t/\delta]} (\rho_{i\delta}(\mathbf{1}))^2\left[\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2\right]\Bigg)\prod_{i=1}^d \alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg] = 0.\end{aligned} \tag{5.39}$$

To prove (5.38), we first observe the following:

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \xi^n_{[s/\delta]\delta}a^n_j(s)R^1_{s,j}(\varphi)ds \triangleq \lim_{n\to\infty}\Lambda^{n,R^1,\varphi}_t = \int_0^t \tilde{\rho}^1_s(\Psi\varphi)ds, \tag{5.40}$$

the proof can be found in Appendix C.3. Then note that

$$\bar{\Lambda}^\varphi_t - \bar{\Lambda}^\varphi_s = U_t(\varphi) - U_s(\varphi) - \int_s^t U_r(A\varphi)dr - \int_s^t U_r(h\varphi)dY_r - \int_s^t \tilde{\rho}^1_r\left(\Psi\varphi\right)dr,$$

thus showing (5.38) is equivalent to showing

$$\tilde{\mathbb{E}}\Bigg[\left(U_t(\varphi) - U_s(\varphi) - \int_s^t U_r(A\varphi)dr - \int_s^t U_r(h\varphi)dY_r - \int_s^t \tilde{\rho}^1_r(\Psi\varphi)dr\right)$$
$$\times\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg] = 0. \tag{5.41}$$

This equality will follow by virtue of the martingale property of $\bar{\Lambda}^\varphi_t - \bar{\Lambda}^\varphi_s$.

By virtue of the existence of $\Lambda^n_T(\tilde{f}_k)$ in Lemma 5.1.5, it follows , for $n' \in \mathbb{N}$, that

$$\sup_{n'}\tilde{\mathbb{E}}\left[(U^{n'}(\varphi))^2\right] < \infty,$$

which implies that $\{U^{n_k}\}$ is uniformly integrable (see II.20, Lemma 20.5 in [63]). Therefore we have that

$$\lim_{k\to\infty}\tilde{\mathbb{E}}\left[U^{n_k}_t(\varphi)\prod_{i=1}^{d}\alpha_i(U^{n_k}_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right] = \tilde{\mathbb{E}}\left[U_t(\varphi)\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right],$$
$$\lim_{k\to\infty}\tilde{\mathbb{E}}\left[U^{n_k}_s(\varphi)\prod_{i=1}^{d}\alpha_i(U^{n_k}_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right] = \tilde{\mathbb{E}}\left[U_s(\varphi)\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right].$$

By Burkholder-Davis-Gundy inequality, we know that

$$\sup_{n'}\tilde{\mathbb{E}}\left[\left(\int_0^t U^{n'}_r(A\varphi)dr\right)^2\right] < \infty;$$

thus we have

$$\lim_{k\to\infty}\tilde{\mathbb{E}}\left[\int_s^t U^{n_k}_r(A\varphi)dr\prod_{i=1}^{d}\alpha_i(U^{n_k}_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right] = \tilde{\mathbb{E}}\left[\int_s^t U_r(A\varphi)dr\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right].$$

Similarly, by Burkholder-Davis-Gundy inequality, we can show that

$$\sup_{n'} \tilde{\mathbb{E}}\left[\left(\int_s^t U_r^{n'}(h\varphi)dY_r\right)^2\right] < \infty,$$

we therefore have that (by Theorem 2.2 in [48]), since (U^{n_k}, Y) converges in distribution to (U, Y), then $(U^{n_k}, Y, \int_s^t U_r^{n_k}(h\varphi)dY_r)$ also converges in distribution to $(U, Y, \int_s^t U_r(h\varphi)dY_r)$, thus we have

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\int_s^t U_r^{n_k}(h\varphi)dY_r \prod_{i=1}^{d} \alpha_i(U_{t_i}^{n_k}) \prod_{j=1}^{d'} \alpha'_j(Y_{t'_j})\right] = \tilde{\mathbb{E}}\left[\int_s^t U_r(h\varphi)dY_r \prod_{i=1}^{d} \alpha_i(U_{t_i}) \prod_{j=1}^{d'} \alpha'_j(Y_{t'_j})\right].$$

For $\int_s^t \tilde{\rho}_r^1\left(\Psi\varphi\right) dr$, we have

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\Lambda_t^{n_k, R^1, \varphi} \prod_{i=1}^{d} \alpha_i(U_{t_i}^{n_k}) \prod_{j=1}^{d'} \alpha'_j(Y_{t'_j})\right] = \tilde{\mathbb{E}}\left[\int_s^t \tilde{\rho}_r^1\left(\Psi\varphi\right) dr \prod_{i=1}^{d} \alpha_i(U_{t_i}) \prod_{j=1}^{d'} \alpha'_j(Y_{t'_j})\right].$$

Now we have shown (5.41), and hence (5.38).

In order to show the second equality (5.39), we firstly make the following observations about the limits of the terms in (5.4):

- We have

$$\lim_{n\to\infty} \left\langle \sqrt{n} A_\cdot^{n,\varphi} \right\rangle_t = \sum_{i=1}^{[t/\delta]} (\rho_{i\delta}(\mathbf{1}))^2 \left[\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2\right].$$

  If we let

$$\bar{A}_t^{\varphi} \triangleq \sum_{i=1}^{[t/\delta]} \rho_{i\delta}(\mathbf{1}) \sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\, \Upsilon_i, \tag{5.42}$$

  where $\{\Upsilon_i\}_{i\in\mathbb{N}}$ is a sequence of independent identically distributed, standard normal random variables, and $\left\{\sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\,\Upsilon_i\right\}_i$ are mutually independent given the σ-algebra $\mathcal{Y}$; then we have $\langle \bar{A}_\cdot^{\varphi}\rangle_t = \lim_{n\to\infty} \left\langle \sqrt{n} A_\cdot^{n,\varphi} \right\rangle_t$.

- For $G_{[t/\delta]}^{n,\varphi}$, we have

$$\lim_{n\to\infty} \left|\sqrt{n} G_{[t/\delta]}^{n,\varphi}\right| = 0 \quad \text{a.s..}$$

- We have

$$\lim_{n\to\infty}\left\langle \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^{\cdot}\xi^n_{[s/\delta]\delta}a^n_j(s)R^2_{s,j}(\varphi)dY_s\right\rangle_t \triangleq \lim_{n\to\infty}\left\langle \Lambda^{n,R^2,\varphi}_{\cdot}\right\rangle_t = \left\langle \Lambda^{R^2,\varphi}_{\cdot}\right\rangle_t,$$

  where

$$\Lambda^{R^2,\varphi}_t = c_\omega\int_0^t \left(\tilde{\rho}^1_s(h\varphi'' - (h\varphi)'')\right)dB^{(2)}_s, \tag{5.43}$$

  c_ω is a constant and $B^{(2)}$ is a Brownian motion independent of Y.

- We have that

$$\lim_{n\to\infty}\left\langle \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^{\cdot}\xi^n_{[s/\delta]\delta}a^n_j(s)R^3_{s,j}(\varphi)dV^{(j)}_s\right\rangle_t \triangleq \lim_{n\to\infty}\left\langle \Lambda^{n,R^3,\varphi}_{\cdot}\right\rangle_t = \left\langle \Lambda^{R^3,\varphi}_{\cdot}\right\rangle_t,$$

  where

$$\Lambda^{R^3,\varphi}_t = \int_0^t \sqrt{\tilde{\rho}^2_s\left((\sigma\varphi')^2\right)}dB^{(3)}_s, \tag{5.44}$$

  $B^{(3)}$ is a Brownian motion independent of $B^{(2)}$ and Y.

The proofs of these observations can be found in Appendix C.3.

From the above observations, we obtain that

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[(\bar{\Lambda}^\varphi_t - \bar{\Lambda}^\varphi_s)^2\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right]\\
&= \lim_{k\to\infty}\tilde{\mathbb{E}}\Bigg[\Bigg(\left(\left\langle\sqrt{n}A^{n_k,\varphi}_{\cdot}\right\rangle_t - \left\langle\sqrt{n}A^{n_k,\varphi}_{\cdot}\right\rangle_s\right) + \left(\left\langle\Lambda^{n_k,R^2,\varphi}_{\cdot}\right\rangle_t - \left\langle\Lambda^{n_k,R^2,\varphi}_{\cdot}\right\rangle_s\right)\\
&\qquad + \left(\left\langle\Lambda^{n_k,R^3,\varphi}_{\cdot}\right\rangle_t - \left\langle\Lambda^{n_k,R^3,\varphi}_{\cdot}\right\rangle_s\right)\Bigg)\times\prod_{i=1}^{d}\alpha_i(U^{n_k}_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg]\\
&= \lim_{k\to\infty}\tilde{\mathbb{E}}\Bigg[\Bigg(\sum_{i=[s/\delta]+1}^{[t/\delta]}(\rho^{n_k}_{i\delta}(\mathbf{1}))^2\left[\pi^{n_k}_{i\delta-}(\varphi^2) - \left(\pi^{n_k}_{i\delta-}(\varphi)\right)^2\right]\\
&\qquad + \int_s^t\left(\tilde{\rho}^{n_k,1}_r(h\varphi'' - (h\varphi)'')\right)^2dr + \int_s^t\tilde{\rho}^{n_k,2}_r\left((\sigma\varphi')^2\right)dr\Bigg)\times\prod_{i=1}^{d}\alpha_i(U^{n_k}_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg]
\end{aligned}$$

$$
\begin{aligned}
=&\tilde{\mathbb{E}}\Bigg[\Bigg(\sum_{i=[s/\delta]+1}^{[t/\delta]}(\rho_{i\delta}(\mathbf{1}))^2\left[\pi_{i\delta-}(\varphi^2)-(\pi_{i\delta-}(\varphi))^2\right]\\
&\qquad+\int_s^t\left(\tilde{\rho}_r^1(h\varphi''-(h\varphi)'')\right)^2dr+\int_s^t\tilde{\rho}_r^2\left((\sigma\varphi')^2\right)dr\Bigg)\times\prod_{i=1}^d\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg]\\
=&\tilde{\mathbb{E}}\left[\left(\langle\bar{\Lambda}.^{\varphi}\rangle_t-\langle\bar{\Lambda}.^{\varphi}\rangle_s\right)\prod_{i=1}^d\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right]; \qquad (5.45)
\end{aligned}
$$

and (5.39) follows from this identity. ■

Corollary 5.2.4. *For and $t \geq 0$ and test function $\varphi \in C_b^6(\overline{\mathbb{R}})$, let*

$$\bar{U}_t^n(\varphi) \triangleq \sqrt{n}\left(\pi_t^n(\varphi)-\pi_t(\varphi)\right).$$

Then $\{\bar{U}^n\}_n$ converges in distribution to a unique $D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$-valued process $\bar{U} = \{\bar{U}_t : t \geq 0\}$, such that

$$\bar{U}_t(\varphi) = \frac{1}{\rho_t(\mathbf{1})}\left(U_t(\varphi)-\pi_t(\varphi)U_t(\mathbf{1})\right), \qquad (5.46)$$

where U satisfies (5.31)*.*

Proof. By the fact that

$$\pi_t^n(\varphi)-\pi_t(\varphi) = \frac{1}{\rho_t(\mathbf{1})}(\rho_t^n(\varphi)-\rho_t(\varphi)) - \frac{\pi_t^n(\varphi)}{\rho_t(\mathbf{1})}(\rho_t^n(\mathbf{1})-\rho_t(\mathbf{1})),$$

and $\rho_t^n(\varphi) \to \rho_t(\varphi)$, a.s. and $\pi_t^n(\varphi) \to \pi_t(\varphi)$ a.s. (see Remark B.1.3 in Appendix C), we have the result. ■

Remark 5.2.5. *In this chapter we view $\{U^n\}_{n\in\mathbb{N}}$ and its weak limit $\{U\}$ as processes with sample paths in $D_{\mathcal{M}_F(\overline{\mathbb{R}})}[0,\infty)$, which is complete and separable. In the following we show that U actually takes value in a smaller space $\mathcal{M}_F(\mathbb{R})$ (i.e. U is a $D_{\mathcal{M}_F(\mathbb{R})}[0,\infty)$-valued random variable). In other words, U has no mass 'escaping' to infinity. This is done by using the same approach as in Section 5 in [16].*

Since the weak topology on $\mathcal{M}_F(\mathbb{R})$ coincides with the trace topology from $\mathcal{M}_F(\overline{\mathbb{R}})$ to $\mathcal{M}_F(\mathbb{R})$, it follows that U has sample paths in $D_{\mathcal{M}_F(\mathbb{R})}[0,\infty)$. It then suffices to show that that for arbitrary t, there exists a sequence of compact sets $\{K_p\}_{p>0} \in \mathbb{R}$ (possibly depending on t) which exhaust $\mathbb{R}$ such that for all $\varepsilon > 0$,

$$\lim_{p\to\infty}\tilde{\mathbb{P}}\left[\sup_{s\in[0,t]}\left(U_s(\mathbf{1}_{K_p^c})\right) \geq \varepsilon\right] = 0, \qquad (5.47)$$

where K_p^c denotes the compliment of K_p. The proof of (5.47) *can be found in Section 5 in [16].*

Therefore, from now on, when discussing the tightness or convergence in distribution results of this chapter, our convention will be that $\mathcal{M}_F(\mathbb{R})$ is endowed with the weak topology generated by the metric $d_{\mathcal{M}}$. The discussion on the compactification $\overline{\mathbb{R}}$ is no longer required.

5.3 Discussion

To be able to obtain the tightness (and hence convergence in distribution) results, multinomial branching algorithm was selected. From Chapter 4 we see that L^2-convergence results for ρ^n and π^n can be obtained under both tree based branching algorithm (TBBA) and multinomial branching algorithm. The main advantage of tree based branching algorithm over multinomial algorithm is that the TBBA has conditional minimal variance property. In other words, it produces the offspring (generalised) particles with a probability distribution that minimises their conditional variance. This is a very attractive feature for resampling algorithms because it is the variance of offspring that determines the speed of convergence.

As can be see from this chapter, the central limit type result, however, can only be obtained under the multinomial algorithm. We cannot obtain the corresponding central limit result for the generalised particle filters involving branching procedure based on the TBBA as the limiting process correspondent to the sequence of quadratic variations $\langle M^{n,\varphi}\rangle_t$ can not be identified explicitly. Therefore we can not describe the evolution equation of the limit U of U^n (we can, however, prove that the process is tight). This is left for future research.

Chapter 6

Suggestions for Possible Areas of Future Research

In Chapters 3, 4, and 5, we did a comprehensive study on the Gaussian mixtures approximation to the solution of the nonlinear filtering problem. We set up the approximating algorithm, and proved a law of large number type result and a central limit type result. These three chapters form the core part of the thesis.

It should be noted that Gaussian mixtures approximation is a natural generalisation of the classic particle filters; it is, however, by no means the unique way of generalising the classic particle filters. In Section 6.1, we introduce the basic ideas of using wavelets, orthomormal polynomials, and finite elements to construct the generalised particle approximations. The ultimate aim is to integrate within the framework of generalised particle filters a wide variety of numerical methods, and develop a common approach to analyse and compare these methods. At the time of this thesis, however, only basic ideas are presented; the rigorous analysis and comparison between these methods are left as future work.

In addition to these, being able to apply generalised particle filters to solve practical problems is of essential importance. An important application of generalised particle filters, especially the Gaussian mixtures approximations, is the problem of filtering the Navier-Stokes equation, whose idea is described in Section 6.2. The content in that section is initiative and more rigorous working, in particular the advantage of Gaussian mixtures over Dirac mixtures, will be the author's ongoing and future work.

6.1 Other Possible Forms of Generalised Particle Filters

In this section, we introduce the basic ideas of constructing the generalised particle filters systems by using different numerical methods. These numerical methods can include

- *Classical Particle Filters*: as explained above, in this case the particles carry information about their weights and positions.
- *Gaussian Mixtures*: the particles are in this case characterised by Gaussian measures. They are parameterised by their weights, mean values and the corresponding covariance matrices.
- *Wavelets*: an orthonormal wavelets series with properly selected dilation and translation parameters is chosen to characterise the particles. The transition centres are viewed as positions; and the weights of the particles are the inner products of the wavelets and a certain chosen density function.
- *Orthonormal Polynomials*: similar to wavelets method, an orthonormal basis of a Hilbert space with properly selected dilation and shifting parameters is chosen to characterise the generalised particles; the Hermite basis is a particular example.
- *Finite Element methods*: the shape functions of a finite element are considered as the positions of the generalised particles, and the nodal variables should act as the generalised weights.

The first two methods have already been rigorously studied. In what follows we will discuss the general ideas of the remaining three methods.

6.1.1 Wavelets Method

In this subsection we describe the idea of using the wavelets method, the evolution equations satisfied by the signal X and the observation Y are assumed to be the same as in the previous section. It is proved in Chapter 7 of [3] that, if the matrix-valued function $a = \frac{1}{2}\sigma\sigma^\top$ is uniformly strictly elliptic, then the density p_t of ρ_t (the solution of the Zakai equation) with respect to the Lebesgue measure exists and is smooth. We can therefore consider the approximation of the density p_t, which is denoted by $p^n = \{p_t^n, t \geq 0\}$. We further assume that p_t^ns exist and are chosen to be smooth functions. Then for any $\varphi \in \mathcal{D}(A)$ we can construct the approximation ρ^n of ρ as

$$\rho_t^n(\varphi) = \int_{\mathbb{R}^d} \varphi(x) p_t^n(x) dx. \tag{6.1}$$

Similarly, for $A\varphi$ and $h\varphi$ we have

$$\rho_t^n(A\varphi) = \int_{\mathbb{R}^d} (A\varphi)(x) p_t^n(x) dx, \quad \rho_t^n(h\varphi) = \int_{\mathbb{R}^d} h(x)\varphi(x) p_t^n(x) dx. \tag{6.2}$$

Consider a continuously differentiable function ψ with compact support chosen as the mother wavelet, and consider the discrete wavelet transform

$$\psi_{j,k}(x) = a^{-\frac{j}{2}} \psi(a^{-j}x - kb); \tag{6.3}$$

We know from the appendix that, by properly choosing dilation parameter a and translation parameter b, and the mother wavelet ψ, the wavelet series $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ can be constructed to form an orthonormal basis for the Hilbert space $L^2(\mathbb{R})$, that is, for any $f \in L^2(\mathbb{R})$,

$$f = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k}$$

where the inner product $\langle f, \psi_{j,k} \rangle = \int_{\mathbb{R}} f(x) \psi_{j,k}(x) dx$.

Instead of the infinite sum in the above equation, we are looking at particles/wavelets with a finite number of elements. Therefore, given the function $p_t^n(x)$, we aim to have:

$$p_t^n = \sum_{j=0}^{m_1} \sum_{k=0}^{m_2} \langle p_t^n, \psi_{j,k} \rangle \psi_{j,k} \tag{6.4}$$

for some $m_1,\ m_2 \in \mathbb{N}$.

By formula (6.4), we can rewrite (6.1) and (6.2) as follows:

$$\rho_t^n(\varphi) = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \int_{\mathbb{R}} \varphi(x) \langle p_t^n, \psi_{j,k}\rangle \psi_{j,k}(x)dx = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \langle p_t^n, \psi_{j,k}\rangle \langle \varphi, \psi_{j,k}\rangle ;$$

$$\rho_t^n(A\varphi) = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \int_{\mathbb{R}} (A\varphi)(x) \langle p_t^n, \psi_{j,k}\rangle \psi_{j,k}(x)dx = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \langle p_t^n, \psi_{j,k}\rangle \langle A\varphi, \psi_{j,k}\rangle ;$$

$$\rho_t^n(h\varphi) = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \int_{\mathbb{R}} (h\varphi)(x) \langle p_t^n, \psi_{j,k}\rangle \psi_{j,k}(x)dx = \sum_{j=0}^{m_1}\sum_{k=0}^{m_2} \langle p_t^n, \psi_{j,k}\rangle \langle h\varphi, \psi_{j,k}\rangle .$$

By properly choosing the wavelet ψ and the dilation and translation parameters, we hope to obtain an equation of the form

$$d\rho_t^n(\varphi) = \rho_t^n(A\varphi)dt + \rho_t^n(h\varphi)dY_t + \text{“small terms”}.$$

At this stage it is not possible to say more about the “small terms” in this equation. The idea is to be able to control the additional terms with bounds depending on the number of wavelets.

The above described work is done within each interval of the partition $[i\delta, (i+1)\delta)$, we should obtain the “small terms” explicitly before we can determine whether or not we need the branching procedure (i.e. $\delta = \infty$ or finite) and what the algorithm should be used if branching is required.

6.1.2 Orthonormal Polynomials Method

We discuss in this subsection the idea of using orthonormal polynomials to characterise the generalised particles, with the emphasis of Hermite polynomials. Instead of unnormalised conditional distribution ρ_t, we consider its smooth density p_t with respect to the Lebesgue measure (see Chapter 7 in [3] for the existence and smoothness of p_t).

The (one-dimensional) Hermite polynomials are defined as follows

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}; \tag{6.5}$$

and the corresponding Hermite functions are

$$\psi_n(x) = (2^n n!\sqrt{\pi})^{-1/2} e^{-x^2/2} H_n(x). \tag{6.6}$$

Hermite functions $\{\psi_n\}_{n=1}^{\infty}$ form an orthomormal basis of $L^2(\mathbb{R})$, and the corresponding inner product is

$$\langle \psi_m, \psi_n \rangle = \int_{\mathbb{R}} \psi_m(x)\psi_n(x)dx = \delta_{mn} = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}. \tag{6.7}$$

From the above definitions, it can be seen that the Hermite functions satisfy the following recursive relations:

$$\begin{aligned} \frac{d}{dx}\psi_n(x) &= \sqrt{\frac{n}{2}}\psi_{n-1}(x) - \sqrt{\frac{n+1}{2}}\psi_{n+1}(x) \\ x\psi_n(x) &= \sqrt{\frac{n}{2}}\psi_{n-1}(x) + \sqrt{\frac{n+1}{2}}\psi_{n+1}(x). \end{aligned} \tag{6.8}$$

We can then conclude, from these relations, that for $m \geq 1$ and $m \in \mathbb{N}$,

$$\frac{d^m}{dx^m}\psi_n(x),\ x^m\psi_n(x) \in \text{span}\{\psi_0, \psi_1, \ldots, \psi_{n+m}\},$$

in other words, we are able to represent $\psi_n^{(m)}(x)$ and $x^m\psi_n(x)$ as (finite) linear combinations of $\{\psi_j(x)\}_{j=0}^{n+m}$. The density p_t can therefore be decomposed as

$$p_t(x) = \sum_{n=0}^{\infty} \langle p_t, \psi_n \rangle \psi_n(x). \tag{6.9}$$

From Chapter 7 in [3] we know that

$$p_t(x) = p_0(x) + \int_0^t A^* p_s(x) ds + \int_0^t h(x)p_s(x)dY_s,$$

where A^* is the adjoint operator of A defined in (2.10), i.e.

$$A^*\varphi = \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}(a^{ij}\varphi) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(f^i\varphi).$$

Then we would like to approximate $p_t(x)$ by a using a finite number of elements in $\{\psi_n\}$ as follows

$$p_t^N(x) = \sum_{n=0}^{N} C_n(t)\psi_n\left(\frac{x_t - \mu_t}{\sqrt{\omega_t}}\right), \tag{6.10}$$

and by determining $C_n(t)$, μ_t and ω_t, we aim to show p_t^N satisfies

$$p_t^N(x) = p_0^N(x) + \int_0^t A^* p_s^N(x) ds + \int_0^t h(x) p_s^N(x) dY_s + \text{“small terms”}. \qquad (6.11)$$

The reason for choosing Hermite polynomial to form the orthonormal basis is that, in the linear case of Kalman-Bucy filter (see, for example, Chapter 6 in [3]), the signal X and the observation Y satisfy the following evolution equations respectively:

$$\begin{aligned} X_t =& X_0 + \int_0^t (F_s X_s + f_s) ds + \int_0^t \sigma_s dV_s, \\ Y_t =& Y_0 + \int_0^t (H_s X_s + h_s) ds + W_t; \end{aligned}$$

where, for and $s \geq 0$, F_s is a $d \times d$ matrix, σ_s is a $d \times p$ matrix, f_s is a d-dimensional vector; H_s is a $m \times d$ matrix, and h_s is a m-dimensional vector. In this case, the density $p_t(x)$ has the following explicit expression (see, for example, [3])

$$p_t(x) = \frac{\hat{Z}_t}{\sqrt{2\pi R_t}} \exp\left(-\frac{(x - \hat{x}_t)^2}{2R_t} \right) \triangleq C_0(t) \psi_0 \left(\frac{x - \hat{x}_t}{\sqrt{R_t}} \right), \qquad (6.12)$$

where $\hat{x}$, $\hat{Z}$ and R satisfy the following evolution equation:

$$\begin{aligned} d\hat{x}_t &= (F_t \hat{x}_t + f_t) dt + R_t H_t^\top (dY_t - (H_t \hat{x}_t + h_t) dt), \\ dR_t &= (\sigma_t \sigma_t^\top + F_t R_t + R_t F_t^\top - R_t H_t^\top H_t R_t) dt, \\ \hat{Z}_t &= \exp\left(\int_0^t (H\hat{x}_s + h)^\top dY_s - \int_0^t \|H\hat{x}_s + h\|^2 ds \right). \end{aligned}$$

From (6.12) we can see that for this Kalman-Bucy filter, we can explicitly represent the density p_t using the one-dimensional subspace of the (infinite dimensional) space $L^2(\mathbb{R})$ with orthonormal basis $\{\psi_n\}_{n=0}^\infty$. Then for general non-linear filtering problem, as a natural extension, it may be possible to use finite dimensional (N) subspace of $L^2(\mathbb{R})$ to characterise the approximation p_t^N.

6.1.3 Finite Element Method

The idea of using finite element analysis is similar to the wavelet method. Again we assume that the signal X and observation Y satisfy the same evolution equations as before, and we know the existence and smoothness of the density p_t of ρ_t with respect to the Lebesgue measure. We consider the approximation sequence $p^n = \{\rho_t^n,\ t \geq 0\}$

of $p = \{p_t, t \geq 0\}$. For any $\varphi \in \mathcal{D}(A)$, we can construct the approximation ρ^n of the solution of the Zakai equation ρ as

$$\rho_t^n(\varphi) = \int_{\mathbb{R}^d} \varphi(x) p_t^n(x) dx. \tag{6.13}$$

Similarly, we have for $A\varphi$ and $h\varphi$ that

$$\rho_t^n(A\varphi) = \int_{\mathbb{R}^d} (A\varphi)(x) p_t^n(x) dx, \qquad \rho_t^n(h\varphi) = \int_{\mathbb{R}^d} h(x)\varphi(x) p_t^n(x) dx. \tag{6.14}$$

By using finite element method, we are essentially approximating functions on individual element domain, or on a collection of element domains (mesh) which subdivide a larger domain. The interpolation error $|f - \mathcal{I}_{\mathcal{T}} f|$ (see Definitions A.2.5 and A.2.6 in Appendix A.2 for details) should also be a major concern.

To be specific, for example, the global interpolation of p_t^n can be written as

$$\mathcal{I}_{\mathcal{T}} p_t^n = \sum_{m=1}^{N} \mathcal{I}_{K_m} p_t^n = \sum_{m=1}^{N} \sum_{i=1}^{k_m} N_i^m(p_t^n) \phi_i^m. \tag{6.15}$$

Then obviously p_t^n has the following expression

$$p_t^n = \mathcal{I}_{\mathcal{T}} p_t^n + Err(p_t^n) = \sum_{m=1}^{N} \sum_{i=1}^{k_m} N_i^m(p_t^n) \phi_i^m + Err(p_t^n), \tag{6.16}$$

where $Err(p_t^n)$ is the approximating error of the interpolation. Therefore we have

$$\rho_t^n(\varphi) = \int_{\mathbb{R}^d} \varphi(x) p_t^n(x) dx = \sum_{m=1}^{N} \sum_{i=1}^{k_m} \int_{\mathbb{R}^d} \varphi(x) \left(N_i^m(p_t^n)\phi_i^m\right)(x) dx + \int_{\mathbb{R}^d} \varphi(x) Err(p_t^n)(x) dx.$$

Similarly we can obtain that

$$\rho_t^n(A\varphi) = \sum_{m=1}^{N} \sum_{i=1}^{k_m} \int_{\mathbb{R}^d} (A\varphi)(x) \left(N_i^m(p_t^n)\phi_i^m\right)(x) dx + \int_{\mathbb{R}^d} (A\varphi)(x) Err(p_t^n)(x) dx;$$

$$\rho_t^n(h\varphi) = \sum_{m=1}^{N} \sum_{i=1}^{k_m} \int_{\mathbb{R}^d} (h\varphi)(x) \left(N_i^m(p_t^n)\phi_i^m\right)(x) dx + \int_{\mathbb{R}^d} (h\varphi)(x) Err(p_t^n)(x) dx.$$

Then following the same procedure as we did for wavelets method, we should choose appropriate forms of finite elements and their corresponding basis, so that the equation of the following form can be obtained:

$$d\rho_t^n(\varphi) = \rho_t^n(A\varphi)dt + \rho_t^n(h\varphi)dY_t + \text{“small terms”}.$$

6.2 Filtering the Solution of the Stochastic Navier-Stokes Equation

In this section we will look at an application of the generalised particle filters discussed in the previous chapters. We consider the Navier-Stokes equation as an example. Instead of solving the problem, we only present the filtering model in this section. The content here is motivated by the work in [6].

6.2.1 Problem Setting

We consider the 2D Stochastic Navier-Stokes equation on the torus $\mathbb{T}^2 \triangleq [0, L) \times [0, L)$ with periodic boundary conditions:

$$\begin{aligned} \frac{\partial u}{\partial t} - \nu\Delta u + u \cdot \nabla u + \nabla p &= f + W(t,x) && \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty), \qquad (6.17) \\ \nabla u &= 0 && \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty), \\ u(x,0) &= u_0(x) && \text{for all } x \in \mathbb{T}^2. \end{aligned}$$

Here $u : \mathbb{T}^2 \times [0,\infty) \to \mathbb{R}^2$ is a time-dependent vector field representing the velocity, $p : \mathbb{T}^2 \times [0,\infty) \to \mathbb{R}^2$ is a time-dependent scalar field representing the pressure, $f : \mathbb{T}^2 \to \mathbb{R}^2$ is a time-independent vector filed representing the forcing, and ν is the viscosity, and $W(t,x)$ is a coloured noise which will be described below. We define

$$\mathcal{H} \triangleq \Bigg\{ L\text{−periodic trigonometirc polynomials } u : [0,L)^2 \to \mathbb{R}^2 \Bigg| \nabla \cdot u = 0, \int_{\mathbb{T}^2} u(x)dx = 0 \Bigg\}$$

and H as the closure of $\mathcal{H}$ with respect to the $(L^2(\mathbb{T}^2))^2$ norm. We then define $P : (L^2(\mathbb{T}^2))^2 \to H$ to be the Leray-Helmholtz orthogonal projector.

Given $k = (k_1, k_2)^\top$, define $k^\perp = (k_2, -k_1)^\top$. Then an orthonormal basis for H is given by $\psi_k : \mathbb{R}^2 \to \mathbb{C}^2$, where

$$\psi_k(x) \triangleq \frac{k^\perp}{|k|} \exp\left(\frac{2\pi i k \cdot x}{L}\right)$$

for $k \in \mathbb{Z}^2 \setminus \{0\}$ and $i = \sqrt{-1}$. Thus for $u \in H$ we may write

$$u = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} u_k(t)\psi_k(x)$$

where, since u is a real-valued function, we have the reality constraint $u_{-k} = -\overline{u_k}$.

We choose the coloured noise $W(t,x)$ to be of the form

$$W(t,x) = \sum_{k \in \mathbb{Z}^2 \backslash \{0\}} \varepsilon_k \psi_t(x) W_t^k, \tag{6.18}$$

where $\{W_t^k\}_{(t \geq 0,\ k \in \mathbb{Z}^2 \backslash \{0\})}$ are mutually independent one-dimensional Brownian motions, and ε_ks are chosen to have the following property

$$\sum_{k \in \mathbb{Z}^2 \backslash \{0\}} (4\pi^2 |k|^2)^s \varepsilon_k^2 < \infty \qquad \text{for } s \in \mathbb{R} \text{ and } s \geq 1,$$

and then $W(t, \cdot) \in H$.

Using the above Fourier decomposition of u, we can define the fractional Sobolev space

$$H^s \triangleq \left\{ u \in H : \sum_{k \in \mathbb{Z}^2 \backslash \{0\}} (4\pi^2 |k|^2)^s |u_k|^2 < \infty \right\}$$

with norm $\|u\|_s = \left(\sum_k (4\pi^2 |k|^2)^s |u_k|^2\right)^{1/2}$ and $s \in \mathbb{R}$.

The following proposition shows the stochastic Navier-Stokes equation can be written as an stochastic ordinary differential equation by applying the projection P in H.

Proposition 6.2.1. *The Stochastic Navier-Stokes equation can be written as*

$$\frac{du}{dt} + \nu A u + \mathcal{B}(u,u) = f + W(t,x). \tag{6.19}$$

Here $A = -P\Delta$ is the Stokes operator, the term $\mathcal{B}(u,u) = P(u \cdot \nabla u)$ is the bilinear form found by projecting the nonlinear term $u \cdot \nabla u$ into H, and f is the original forcing projected into H.

Proof. Without loss of generality, we assume that $u, f \in H$. We take the inner product of this equation with an element $v \in H$, to obtain

$$\left(\frac{\partial u}{\partial t}, v\right) - \nu \int_{\mathbb{T}^2} (\Delta u) \cdot v dx + \int_{\mathbb{T}^2} (u \cdot \nabla u) \cdot v dx + \int_{\mathbb{T}^2} (\nabla p) \cdot v dx = \int_{\mathbb{T}^2} f \cdot v dx + \int_{\mathbb{T}^2} W \cdot v dx \tag{6.20}$$

By integrating the p term by parts we obtain

$$\int_{\mathbb{T}^2} (\nabla p) \cdot v dx = \int_{\mathbb{T}^2} p (\nabla \cdot v) dx = 0.$$

Applying the projector P to both sides of (6.20), note that $W \in H$, we have for all $v \in H$

$$\left(\frac{\partial u}{\partial t}, v\right) + \nu \int_{\mathbb{T}^2} ((-P\Delta)u) \cdot v dx + \int_{\mathbb{T}^2} P(u \cdot \nabla u) \cdot v dx = \int_{\mathbb{T}^2} f \cdot v dx + \int_{\mathbb{T}^2} W \cdot v dx,$$

therefore by letting $A = -P\Delta$ and $\mathcal{B}(u,u) = P(u \cdot \nabla u)$, we can rewrite this equation as

$$\frac{du}{dt} + \nu A u + \mathcal{B}(u,u) = f + W(t,x);$$

which is exactly (6.19). ■

Remark 6.2.2. *E, Mattingly, and Sinai (see [33]) studies the stochastically forced Navier-Stokes equation with similar random forcing term. They proved the uniqueness of the stationary measure under the condition that all "determining modes" are forced by studying the Gibbsian dynamics of the low modes obtained by representing the high modes as functionals of the time-history of the low modes.*

Recall that

$$u = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} u_k(t) \psi_k(x). \tag{6.21}$$

The following theorem gives the evolution equations satisfied by $u_k(t)$.

Theorem 6.2.3. *Let* (6.21) *be the decomposition of the solution of the stochastic Navier-Stokes equation. Then for each* $k \in \mathbb{Z}^2/\{0\}$, $u_k(t)$ *satisfies*

$$du_k(t) = \left(-\nu \lambda_k u_k(t) - \alpha_k^{l,j} \sum_{l+j=k} u_l(t) u_j(t) + f_k\right) dt + \varepsilon_k dW_t^k, \tag{6.22}$$

where

$$\alpha_k^{l,j} = \begin{cases} \frac{2\pi i (l_2 j_1 - l_1 j_2)(k_1 j_1 + k_2 j_2)}{L \; |k||l||j|} & \text{if} \quad k = l + j, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{6.23}$$

Proof. From (6.21) we know that

$$du = \sum_{k \in \mathbb{Z}^2/\{0\}} \psi_k(x) du_k(t). \tag{6.24}$$

Recalling (6.19), since the Stokes operator A can be diagonalised in the basis comprised of the $\{\psi_k\}_{k \in \mathbb{Z}^2 \setminus \{0\}}$ on H, and the eigenvalues of A are $\lambda_k = 4\pi^2 |k|^2 / L^2$, we know that

$$Au = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \lambda_k u_k(t) \psi_k(x). \tag{6.25}$$

The bilinearity of $\mathcal{B}(u,u)$ implies

$$\begin{aligned}\mathcal{B}(u,u) &= \sum_{l,j\in\mathbb{Z}^2\setminus\{0\}} u_l(t)u_j(t)\mathcal{B}(\psi_l,\psi_j) = \sum_{l,j\in\mathbb{Z}^2\setminus\{0\}} u_l(t)u_j(t)P(\psi_l\cdot\nabla\psi_j)\\ &= \sum_{l,j\in\mathbb{Z}^2\setminus\{0\}} u_l(t)u_j(t)\sum_{k\in\mathbb{Z}^2\setminus\{0\}} \alpha_k^{l,j}\psi_k(x)\\ &= \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\left(\sum_{l,j\in\mathbb{Z}^2\setminus\{0\}} u_l(t)u_j(t)\right)\alpha_k^{l,j}\psi_k(x); \end{aligned} \tag{6.26}$$

where $\alpha_k^{i,j}$ is the inner product $\langle P(\psi_l\cdot\nabla\psi_j),\psi_k\rangle$ written as (see Appendices D.1 and D.2)

$$\alpha_k^{l,j} = \frac{1}{L^2}\int_{\mathbb{T}^2}\left((\psi_l\cdot\nabla\psi_j)\cdot\overline{\psi_k}\right)(x)dx. \tag{6.27}$$

By (D.11) (see Appendix D.2) we know $\alpha_k^{l,j}$ has the expression as in (6.23). Thus

$$\mathcal{B}(u,u) = \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\alpha_k^{l,j}\left(\sum_{l+j=k} u_l(t)u_j(t)\right)\psi_k(x); \tag{6.28}$$

As $f\in H$, we can write it as

$$f(x) = \sum_{k\in\mathbb{Z}^2\setminus\{0\}} f_k\psi_k(x). \tag{6.29}$$

Finally the decomposition of $W(t,x)$ comes from its definition (6.18).

We then obtain the evolution equation for $\{u_k(t)\}_{t\geq 0}$ and each $k\in\mathbb{Z}\setminus\{0\}$ as

$$du_k(t) = \left(-\nu\lambda_k u_k(t) - \alpha_k^{l,j}\sum_{l+j=k} u_l(t)u_j(t) + f_k\right)dt + \varepsilon_k dW_t^k,$$

which is exactly (6.22). ■

Remark 6.2.4. *Figure 6.1 shows the values of $\alpha_k^{l,j}$ over different indices k and l.*

From Theorem 6.2.3, the evolution of each $u_k(t)$ depends on infinite number of $u_l(t)$s, which makes the analysis of this dynamic system difficult. We then define the projection operators $P_\lambda : H\to H$ and $Q_\lambda : H\to H$ by

$$P_\lambda u = \sum_{\substack{k\in\mathbb{Z}^2\setminus\{0\}\\ |2\pi k|^2<\lambda L^2}} u_k(t)\psi_k(x),\quad Q_\lambda = I - P_\lambda;$$

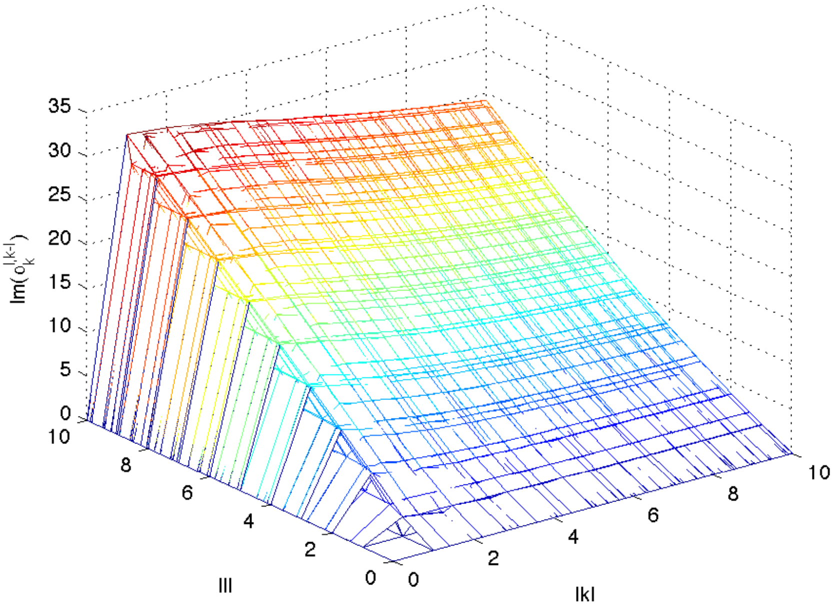

Figure 6.1: Values of $\alpha_k^{l,j}$

and consider the projected eigenvalues, we obtain the following evolution equation for the approximation of $u_k(t)$, which is denoted by $\tilde{u}_k(t)$, for each $k \in \mathbb{Z} \setminus \{0\}$ with $|2\pi k|^2 < \lambda L^2$:

$$d\tilde{u}_k(t) = \left(-\nu\lambda_k \tilde{u}_k(t) - \alpha_k^{l,j} \sum_{\Gamma} \tilde{u}_l(t)\tilde{u}_j(t) + f_k \right) dt + \varepsilon_k dW_t^k; \tag{6.30}$$

where the set $\Gamma \triangleq \left\{ (l,j) \middle| l+j=k \text{ and } |2\pi l|^2 < \lambda L^2 \text{ and } |2\pi j|^2 < \lambda L^2 \right\}$. The approximation $\tilde{u}$ of u is then given by

$$\tilde{u}_\lambda = \sum_{\substack{k \in \mathbb{Z}^2 \setminus \{0\} \\ |2\pi k|^2 < \lambda L^2}} \tilde{u}_k(t)\psi_k(x).$$

Figure 6.2 shows the magnitudes and angles of the complex valued $\tilde{u}_k(t)$ for different k over time t (assuming f_k and ε_k to be 0). It can be seen from the left hand side, which plots the magnitudes, that almost all $\tilde{u}_k$ decay to 0 very quickly (when $t \leq 20$). The right hand side are the angles of each $\tilde{u}_k(t)$, which range from $-\pi$ to π.

From the above simulation we can see that $\tilde{u}_k(t)$ (or $u_k(t)$) decays to 0 as k and t increase. Heuristically $\tilde{u}_k$ can converge to u_k as $k \to \infty$. Actually for the

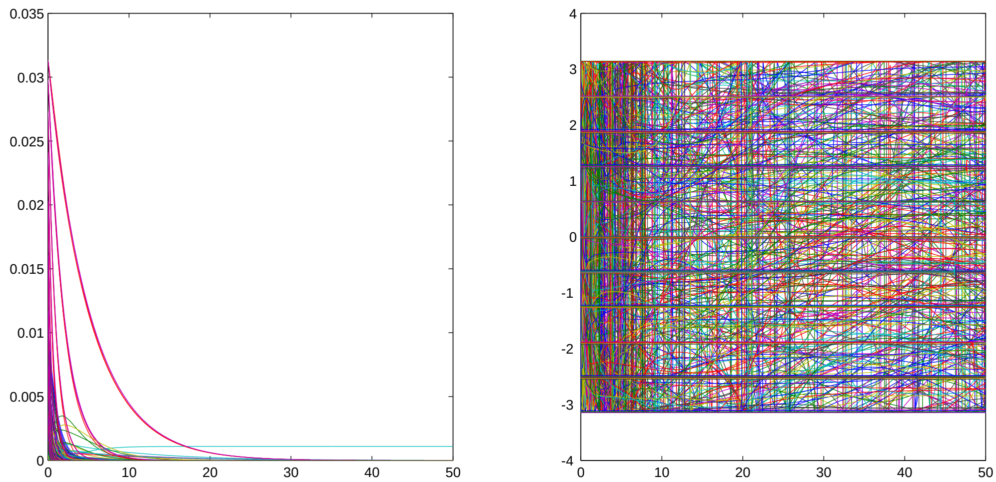

Figure 6.2: Magnitudes and angles of $\tilde{u}_k(t)$

deterministic Navier-Stokes equation, we have the following propositions, which was proved by Foias, Manley, Rosa and Temam (see [37]):

Proposition 6.2.5. *Suppose u satisfies (6.19) except that $W(t,x)=0$, and u is in the Gevrey space $D(\exp(\sigma A^s))$ for $\sigma>0$ and $s=1/2$; where A is the Stokes operator. Then u has the following decomposition*

$$u=\sum_{k\in\mathbb{Z}^d} u_k e^{2\pi i \frac{k}{L}\cdot x}, \quad u_k\in\mathbb{C}^d,\ u_{-k}=\bar{u}_k. \tag{6.31}$$

Then the Fourier coefficients u_k have upper bound

$$|u_k|^2\le \frac{M}{\sqrt{2\pi L}}\left|\frac{k}{L}\right| e^{-2\pi\delta_0\left|\frac{k}{L}\right|}. \tag{6.32}$$

Proof. See discussions in Appendix D.3. ∎

Proposition 6.2.6. *Assume u satisfies (6.19) with $W(t,x)=0$, and that $u_0\in H^1$, $f\in H$, then the equation satisfied by u has a unique strong solution on $t\in[0,T]$ for any $T>0$:*

$$u\in L^\infty((0,T);H^1)\cap L^\infty((0,T);D(A)), \quad \frac{du}{dt}\in L^2((0,T);H).$$

Furthermore, the equation has a global attractor $\mathcal{A}$ and there exists $K>0$ such that, if $u_0\in\mathcal{A}$, then

$$\sup_{t\ge 0}\|u(t)\|_{1,\infty}^2\le K.$$

Proof. See Theorems 9.5 and 12.5 in [60]. ■

For the solution u of stochastic Navier-Stokes equation (6.19), the author would like to prove the convergence as well as find the convergence rate of $\tilde{u}_k$ to u_k, and then obtain similar results for $\tilde{u}_k$ as in Propositions 6.2.5 and 6.2.6 (possibly) using the techniques adopted in Section D.3. Once this is done, we can know the exact error between u_k and $\tilde{u}_k$, and it suffies to focus on $\tilde{u}_k$, which is a finite dimensional system, to study various properties of $u(t)$.

6.2.2 Filtering the Navier-Stokes Equations

In some cases the flow modelled by the stochastic Navier-Stokes equations is not observable directly, which makes filtering a necessary tool to investigate the problem. From the discussions in the previous section, it is known that in order to study (6.17) under the filtering framework, it suffices to investigate (6.22) or its truncated version (6.30). We view (6.30) as the signal, and its observation will inevitably be perturbed by certain noises. Our interests are therefore to find the conditional distribution of the signal process $\tilde{u}_k(t)$ based on its noisy observations. In what follows we will build up the filtering model based on this idea.

Recall the system of $\{\tilde{u}_k(t)\}_k$ satisfies (6.30), viz

$$d\tilde{u}_k(t) = \left(-\nu\lambda_k\tilde{u}_k(t) - \alpha_k^{l,j}\sum_{\Gamma}\tilde{u}_l(t)\tilde{u}_j(t) + f_k\right)dt + \varepsilon_k dW_t^k;$$

where the set $\Gamma \triangleq \left\{(l,j)\middle| l+j=k \text{ and } |2\pi l|^2 < \lambda L^2 \text{ and } |2\pi j|^2 < \lambda L^2\right\}$. For simplicity, the corresponding system of the observation process $\{y_m(t)\}_m$ is, for the moment, modelled as linear:

$$dy_m(t) = h_{k,m}\tilde{u}_k(t)dt + d\tilde{W}_t^m, \tag{6.33}$$

where $h_{k,m} \in \mathbb{R}$, $\tilde{W}_t^m$ is a one-dimensional Brownian motion, and $m = 1, \ldots, M$. We further assume that M is much smaller than the number of elements in Γ. This is reasonable because in practice it is usually difficult to obtain and process the observation process with dimensions as large as the signal process.

From the simulation in the previous section, we can see that $\tilde{u}_k(t)$s decay to zero very quickly as k and t increase; and thus, although a rigorous proof is still required, it is reasonable to assume in the first instance that $\tilde{u}_k(t) \approx 0$ as $|k| \geq M$. This feature enables us to reduce the system of signal processes $\tilde{u}_k(t)$ into the following simplified one which has the same dimension as the system of observation processes:

$$d\tilde{\tilde{u}}_k(t) = \left(-\nu\lambda_k \tilde{\tilde{u}}_k(t) - \alpha_k^{l,j} \sum_{\Gamma_M} \tilde{\tilde{u}}_l(t)\tilde{\tilde{u}}_j(t) + f_k \right) dt + \varepsilon_k dW_t^k, \tag{6.34}$$

where the set $\Gamma_M \triangleq \left\{ (l,j) \middle| (l,j) \in \Gamma,\ |l+j|^2 = M^2 \right\}$. Then it can be seen that the number of elements in Γ_M is much smaller than that in Γ.

Now we have an idea of the construction of the filtering framework, and we will now pose some further questions which can be looked into to gain a further insight. First it is necessary to rigorously prove that the difference between $\tilde{u}_k$ and $\tilde{\tilde{u}}_k$ can be controlled by some small terms. After that, it will be interesting to see how generalised particle filters, especially the mixture of Gaussian measure, can be used to approximate the solution of this filtering problem. This work is still ongoing and the complete results will be obtained in the near future.

6.3 Suggestions for Future Research

Based on the discussion in Section 5.3, Sections 6.1 and 6.2, I would suggest the following three aspects as possible directions for future research.

- As we can see from Chapters 4 and 5, L^2-convergence results can be obtained for both tree based branching algorithm (TBBA) and multinomial branching algorithm; the central limit type result, however, can only be obtained for the multinomial algorithm. It is worth investigating, both from theoretical and practical point of view, how we can obtain convergence in distribution result under TBBA; because this procedure has conditional minimal variance property.

- As mentioned in Section 6.1, there are still several other possible tools to help construct the generalised particles, including wavelets, orthonormal polynomials, and finite elements. The key idea of constructing the approximations is similar to using the Gaussian mixtures. Again we denote the approximation of the ρ_t by ρ_t^n, and aim to make the approximation ρ_t^n satisfy

$$d\rho_t^n(\varphi_t) = \rho_t^n\left(\frac{\partial \varphi_t}{\partial t} + A\varphi_t\right)dt + \rho_t^n(\varphi_t h^\top)dY_t + R_t^n(\varphi_t). \qquad (6.35)$$

Comparing (6.35) with the Zakai equation (2.20), it can be seen that these two equations are “sufficiently” close to each other provided the remainder term $R_t^n(\varphi_t)$ in (6.35) is “sufficiently” small, in which case ρ_t^n will converge to ρ_t (see Appendix B.2 for the rigorous statement and its proof). Once this is done, a comparative theoretical analysis can be established in order to identify the optimal methods within the class of generalised particle filters for various classes of approximations.

- Filtering the (stochastic) Navier-Stokes equation is a relatively new area and the known knowledge about it is still quite limited. From the discussions in Section 6.2, we see that several gaps are waiting to be filled, and they are the author's ongoing and future work. To be specific, the convergence results of $\tilde{u}_k(t)$ to $u_k(t)$ and $\tilde{\tilde{u}}_k(t)$ to $\tilde{u}_k(t)$ (as $k \to \infty$) should be proved before it can be placed under the filtering framework with $\tilde{\tilde{u}}_k(t)$ being viewed as the signal. The following work is to apply the generalised particle filters, especially the mixture of Gaussian measures, to the established filtering model; and prove the corresponding convergence results.

Chapter 7

Conclusions

In this thesis we have analysed a class of approximations of the posterior distribution under continuous time framework. In particular, we investigate in details the case where Gaussian mixtures are used to approximate the posterior distribution.

The L^2-convergence rate and a central limit type result of such approximation are obtained. This method can be viewed as a natural extension of the classic particle filters, in the sense that the classic one is a special case of the generalised one. In general, the approximating measure has a smooth density with respect to the Lebesgue measure and this can enable us to study more properties of the posterior measures than the classic particle filters do; especially this makes it possible to study various properties about the density of ρ_t through its approximation ρ_t^n. Furthermore, the Gaussian mixture particle filters also reduces the computational efforts by carrying more information on each (generalised) particle. It can also be seen that the asymptotic behaviour ($n \to \infty$) of the Gaussian mixtures approximation is similar to the classic particle filters, which is not surprising. As the number of (generalised) particles increases, the quantisation of the posterior distribution becomes finer and finer. Therefore, *asymptotically*, the positions and the weights of the particles provide sufficient information to obtain a good approximation.

Finally, Chapter 6 outlined some possible directions for future research, which include three other forms of generalised particles as well as the application to the filtering of the Navier-Stokes equation.

Appendix A

Convergence Analysis

A.1 Preliminary Results

Proposition A.1.1 (Gronwall's inequality)**.** *Suppose that a continuous function $g(t)$ satisfies*

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s)ds, \qquad 0 \leq t \leq T,$$

with $\beta \leq 0$ and $\alpha : [0, T] \to \mathbb{R}$ integrable; then

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s) e^{\beta(t-s)} ds, \qquad 0 \leq t \leq T.$$

Proof. See, for example, in Chapter 5 in [44]. ■

Proposition A.1.2 (Jensen's inequality for definite integrals)**.** *Suppose $u \in L^p([0, T])$ is integrable for $p \geq 1$, then the following inequality holds*

$$\left(\int_0^t u_s ds \right)^p \leq t^{p-1} \int_0^t u_s^p ds.$$

Proof. See, for example, [70]. ■

Proposition A.1.3 (The Burkholder-Davis-Gundy inequality)**.** *Assume that M is a continuous local martingale, then for every $p > 0$, there exists a universal constant K_p such that*

$$\tilde{\mathbb{E}} \left[\left(\sup_{t \leq T} |M_T| \right)^p \right] \leq K_p \tilde{\mathbb{E}} \left[\langle M \rangle_T^{p/2} \right].$$

Proof. See, for example, Theorem 3.3.28 in [44]. ■

Definition A.1.4 (Markov Semigroup on $C_b(\mathbb{R}^d)$). *A one-parameter family $(P_t)_{t\geq 0}$ of bounded linear operators on $C_b(\mathbb{R}^d)$ with norm $\|\cdot\|$ is a semigroup if*

- *$P_0 = I$ (the identity operator),*
- *$P_{t+s}(f) = P_t(P_s(f))$ for any $f \in C_b(\mathbb{R}^d)$ (semigroup property).*

A Markov semigroup $(P_t)_{t\geq 0}$ on $\mathbb{R}^d$ is a semigroup associated to a Markov process

$$X = (\Omega,\ (\mathcal{F}_t)_{t\geq 0},\ (X_t)_{t\geq 0},\ (P_t)_{t\geq 0},\ \{\mathbf{P}^x : x \in E\})$$

where $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is a measurable space, $(\Omega, (\mathcal{F}_t)_{t\geq 0})$ is a filtered space and $\mathbf{P}^x$ is the probability law for each point $x \in \mathbb{R}^d$ such that for $0 \leq s \leq t$, $f \in \mathcal{B}(\mathbb{R}^d)$, the set of bounded $\mathcal{B}(\mathbb{R}^d)$-measurable functions, and $x \in \mathbb{R}^d$

$$\mathbb{E}^x[f(X_{s+t})|\mathcal{F}_s] = (P_t f)(X_s), \quad \mathbf{P}^x - a.s..$$

An example of $(P_t)_{t\geq 0}$ are the transition functions (see Definition 1.1 in Chapter III.1 of [63] for details).

A.2 Proof of Theorem 4.2.6

Theorem A.2.1. *Let $\mu^n = \{\mu_t^n :\ t \geq 0\}$ be a measure-valued process such that for any $\varphi \in C_b^m(\mathbb{R}^d)$, $m \geq 6$, any fixed $\alpha \geq 1$ and fixed $s > t$, we have*

$$\mu_t^n(\varphi) = \mu_0^n(a_s(\varphi)) + \sum_{l=1}^{\alpha} R_{t,l}^{n,\varphi} + \sum_{k=1}^{\beta} \int_0^t \mu_r^n(a_{s,r}^k(\varphi))dW_r^k, \tag{A.1}$$

where $W = (W^k)_{k=1}^{\beta}$ is an β-dimensional Brownian motion, and $a_s,\ a_{s,r}^k :\ C_b^m(\mathbb{R}^d) \to C_b^m(\mathbb{R}^d)$ are bounded linear operators with bounds c and C_k $(k = 1, \ldots, \beta)$ respectively, i.e., $\|a_s(\varphi)\|_{m,\infty} \leq c\|\varphi\|_{m,\infty}$ and $\|a_{s,r}^k(\varphi)\|_{m,\infty} \leq C_k\|\varphi\|_{m,\infty}$. If for any $T > 0$ there exist constants $\gamma_0, \gamma_1, \ldots, \gamma_\alpha$ such that for $t \in [0,T]$, $p \geq 2$ and $q_l > 0$ $(l = 0, 1, \ldots, \alpha)$,

$$\tilde{\mathbb{E}}\left[|\mu_0^n(a_s(\varphi))|^p\right] \leq \frac{\gamma_0}{n^{q_0}}\|\varphi\|_{m,\infty}^p, \qquad \tilde{\mathbb{E}}\left[|R_{t,l}^{n,\varphi}|^p\right] \leq \frac{\gamma_l}{n^{q_l}}\|\varphi\|_{m,\infty}^p,\ l = 1, \ldots, \alpha. \tag{A.2}$$

Then for any $t \in [0,T]$, we have

$$\tilde{\mathbb{E}}\left[|\mu_t^n(\varphi)|^p\right] \leq \frac{c_t}{n^q}\|\varphi\|_{m,\infty}^p, \tag{A.3}$$

where c_t is a constant independent of n and $q = \min(q_0, q_1, \ldots, q_\alpha)$.

Proof. We first show that for any $t \in [0,T]$

$$\|\mu_t^n(\mathbf{1})\|_p^p = \tilde{\mathbb{E}}\left[|\mu_t^n(\mathbf{1})|^p\right] < \infty.$$

Observe that for $\varphi = \mathbf{1}$ and $t \in [0,T]$

$$\mu_t^n(\mathbf{1}) = \mu_0^n(a_s(\mathbf{1})) + \sum_{l=1}^{\alpha} R_{t,l}^{n,\mathbf{1}} + \sum_{k=1}^{\beta} \int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k,$$

then Minkowski inequality and the fact that $\|\mathbf{1}\|_{m,\infty} = 1$ imply

$$\begin{aligned}
\|\mu_t^n(\mathbf{1})\|_p &\leq \|\mu_0^n(a_t(\mathbf{1}))\|_p + \sum_{l=1}^{\alpha} \|R_{t,l}^{n,\mathbf{1}}\|_p + \left[\tilde{\mathbb{E}}\left|\sum_{k=1}^{\beta} \int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k\right|^p\right]^{1/p} \\
&\leq (\alpha+1)\left(\frac{\gamma}{n^q}\right)^{1/p} + \left[\tilde{\mathbb{E}}\left|\sum_{k=1}^{\beta} \int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k\right|^p\right]^{1/p}
\end{aligned}$$

where $\gamma = \max(\gamma_0, \gamma_1, \ldots, \gamma_\alpha)$. Then Burkholder-Davis-Gundy and Jensen's inequalities we have that

$$\begin{aligned}
\|\mu_t^n(\mathbf{1})\|_p^p &\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\left[\tilde{\mathbb{E}}\left|\sum_{k=1}^{\beta} \int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k\right|^p\right] \\
&\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^{p-1} \sum_{k=1}^{\beta} \tilde{\mathbb{E}}\left[\left|\int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k\right|^p\right] \\
&\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^{p-1} K \sum_{k=1}^{\beta} \tilde{\mathbb{E}}\left[\left\langle \int_0^\cdot \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right) dW_r^k \right\rangle_t^{p/2}\right] \\
&= 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^{p-1} K \sum_{k=1}^{\beta} \tilde{\mathbb{E}}\left[\left(\int_0^t \mu_r^n\left(a_{s,r}^k(\mathbf{1})\right)^2 dr\right)^{p/2}\right] \\
&\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^{p-1} K t^{p/2-1} \sum_{k=1}^{\beta} \tilde{\mathbb{E}}\left[\int_0^t \left|\mu_r^n\left(a_{s,r}^k(\mathbf{1})\right)\right|^p dr\right] \\
&\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^{p-1} K t^{p/2-1} \sum_{k=1}^{\beta} C_k^p \int_0^t \tilde{\mathbb{E}}\left[|\mu_r^n(\mathbf{1})|^p\right] dr \\
&\leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} + 2^{p-1}\beta^p K t^{p/2-1} C^p \int_0^t \|\mu_r^n(\mathbf{1})\|_p^p dr \qquad \text{(A.4)}
\end{aligned}$$

where $C = \max(C_1, \ldots, C_\beta)$. Then from Gronwall's inequality we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[|\mu_s^n(\mathbf{1})|^p\right] = \|\mu_s^n(\mathbf{1})\|_p^p \\
\leq& 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q}\left(1 + 2^{p-1}\beta^p K t^{p/2-1} C^p \int_0^t \exp\left(2^{p-1}\beta^p K t^{p/2-1} C^p (t-r)\right) dr\right) \\
=& 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q} \exp\left(2^{p-1}\beta^p K t^{p/2} C^p\right) \\
\triangleq& D < \infty.
\end{aligned}
$$

Using a similar approach, with $\mathbf{1}$ replaced by φ, we can obtain a similar inequality as in the third-from-the-last inequality in (A.4):

$$
\|\mu_t^n(\varphi)\|_p^p \leq 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q}\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^{p-1} K t^{p/2-1} \sum_{k=1}^{\beta} \int_0^t \tilde{\mathbb{E}}\left[\left|\mu_r^n\left(a_{s,r}^k(\varphi)\right)\right|^p\right] dr.
$$

Now denote by

$$
A_{s,t}^k \triangleq \int_0^t \tilde{\mathbb{E}}\left[\left|\mu_r^n\left(a_{s,r}^k(\varphi)\right)\right|^p\right] dr = \int_0^t \|\mu_r^n\left(a_{s,r}^k(\varphi)\right)\|_p^p dr, \tag{A.5}
$$

and $\Delta = (\alpha+1)^p \frac{\gamma}{n^q}$, we have

$$
\|\mu_t^n(\varphi)\|_p^p \leq 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^{p-1} K t^{p/2-1} \sum_{k=1}^{\beta} A_{s,t}^k. \tag{A.6}
$$

Similar to the penultimate inequality in (A.4), we can have that

$$
\begin{aligned}
\|\mu_t^n(\varphi)\|_p^p \leq& 2^{p-1}(\alpha+1)^p \frac{\gamma}{n^q}\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^{p-1} K t^{p/2-1} \sum_{k=1}^{\beta} C_k^p \int_0^t \tilde{\mathbb{E}}\left[|\mu_r^n(\varphi)|^p\right] dr \\
\leq& 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^{p-1} K t^{p/2-1} C^p \sum_{k=1}^{\beta} \int_0^t \|\varphi\|_{m,\infty}^p \tilde{\mathbb{E}}\left[|\mu_r^n(\mathbf{1})|^p\right] dr \\
\leq& 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^{p-1} K t^{p/2-1} C^p \beta \int_0^t \|\varphi\|_{m,\infty}^p D dr \\
=& 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^p K t^{p/2} C^p D \|\varphi\|_{m,\infty}^p.
\end{aligned} \tag{A.7}
$$

Replacing φ by $a_{s,r}^k(\varphi)$ in (A.7), we get that

$$
\begin{aligned}
\|\mu_r^n\left(a_{s,r}^k(\varphi)\right)\|_p^p \leq& 2^{p-1}\Delta\|a_{s,r}^k(\varphi)\|_{m,\infty}^p + 2^{p-1}\beta^p K r^{p/2} C^p D \|a_{s,r}^k(\varphi)\|_{m,\infty}^p \\
\leq& 2^{p-1}\Delta C^p \|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^p K r^{p/2} C^{2p} D \|\varphi\|_{m,\infty}^p,
\end{aligned} \tag{A.8}
$$

Substituting into (A.5) and denote by $\kappa = p/2$, we have for $k = 1, \dots, \beta$

$$A_{s,t}^k \le 2^{p-1}\Delta C^p t\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^p K C^{2p} D \frac{t^{\kappa+1}}{\kappa+1}\|\varphi\|_{m,\infty}^p \tag{A.9}$$

and (A.6) becomes

$$\begin{aligned}\|\mu_t^n(\varphi)\|_p^p \le & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p \\ & + 2^{p-1}\beta^{p-1}Kt^{p/2-1}\sum_{k=1}^{\beta}\left[2^{p-1}\Delta C^p t\|\varphi\|_{m,\infty}^p + 2^{p-1}\beta^p K C^{2p} D\frac{t^{\kappa+1}}{\kappa+1}\|\varphi\|_{m,\infty}^p\right] \\ \le & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa \Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^{2p}K^2C^{2p}D\frac{t^{2\kappa}}{\kappa+1}\|\varphi\|_{m,\infty}^p.\end{aligned} \tag{A.10}$$

Repeat what was done in (A.8) and (A.9), and from (A.10), we have that

$$\begin{aligned}A_{s,t}^k \le & 2^{p-1}C^p\Delta t\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^{2p}\Delta\frac{t^{\kappa+1}}{\kappa+1}\|\varphi\|_{m,\infty}^p \\ & +2^{2(p-1)}\beta^{2p}K^2C^{3p}D\frac{t^{2\kappa+1}}{(\kappa+1)(2\kappa+1)}\|\varphi\|_{m,\infty}^p;\end{aligned}$$

and then (A.6) becomes

$$\begin{aligned}\|\mu_t^n(\varphi)\|_p^p \le & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa\Delta\|\varphi\|_{m,\infty}^p + 2^{3(p-1)}\beta^{2p}K^2C^{2p}\frac{t^{2\kappa}}{\kappa+1}\Delta\|\varphi\|_{m,\infty}^p \\ & +2^{3(p-1)}\beta^{3p}K^3C^{3p}D\frac{t^{3\kappa}}{(\kappa+1)(2\kappa+1)}\|\varphi\|_{m,\infty}^p.\end{aligned}$$

Repeat the iteration process again, we have that

$$\begin{aligned}A_{s,t}^k \le & 2^{p-1}C^p\Delta t\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^{2p}\Delta\frac{t^{\kappa+1}}{\kappa+1}\|\varphi\|_{m,\infty}^p \\ & +2^{3(p-1)}\beta^{2p}K^2C^{3p}\Delta\frac{t^{2\kappa+1}}{(\kappa+1)(2\kappa+1)}\|\varphi\|_{m,\infty}^p \\ & +2^{3(p-1)}\beta^{3p}K^3C^{4p}D\frac{t^{3\kappa+1}}{(\kappa+1)(2\kappa+1)(3\kappa+1)}\|\varphi\|_{m,\infty}^p;\end{aligned}$$

and that

$$\begin{aligned}\|\mu_t^n(\varphi)\|_p^p \le & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa\Delta\|\varphi\|_{m,\infty}^p + 2^{3(p-1)}\beta^{2p}K^2C^{2p}\frac{t^{2\kappa}}{\kappa+1}\Delta\|\varphi\|_{m,\infty}^p \\ & +2^{4(p-1)}\beta^{3p}K^3C^{3p}\frac{t^{3\kappa}}{(\kappa+1)(2\kappa+1)}\Delta\|\varphi\|_{m,\infty}^p \\ & +2^{4(p-1)}\beta^{4p}K^4C^{4p}D\frac{t^{4\kappa}}{(\kappa+1)(2\kappa+1)(3\kappa+1)}\|\varphi\|_{m,\infty}^p.\end{aligned}$$

Once again we have

$$\begin{aligned}\|\mu_t^n(\varphi)\|_p^p \leq & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa \Delta\|\varphi\|_{m,\infty}^p + 2^{3(p-1)}\beta^{2p}K^2C^{2p}\frac{t^{2\kappa}}{\kappa+1}\Delta\|\varphi\|_{m,\infty}^p \\ & + 2^{4(p-1)}\beta^{3p}K^3C^{3p}\frac{t^{3\kappa}}{(\kappa+1)(2\kappa+1)}\Delta\|\varphi\|_{m,\infty}^p \\ & + 2^{5(p-1)}\beta^{4p}K^4C^{4p}\frac{t^{4\kappa}}{(\kappa+1)(2\kappa+1)(3\kappa+1)}\Delta\|\varphi\|_{m,\infty}^p \\ & + 2^{5(p-1)}\beta^{5p}K^5C^{5p}D\frac{t^{5\kappa}}{(\kappa+1)(2\kappa+1)(3\kappa+1)(4\kappa+1)}\|\varphi\|_{m,\infty}^p.\end{aligned}$$

In general after $k^{th}-$iteration, we have that

$$\begin{aligned}\|\mu_t^n(\varphi)\|_p^{p,k} \triangleq & \|\mu_t^n(\varphi)\|_p^p \\ \leq & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa \Delta\|\varphi\|_{m,\infty}^p + 2^{3(p-1)}\beta^{2p}K^2C^{2p}\frac{t^{2\kappa}}{\kappa+1}\Delta\|\varphi\|_{m,\infty}^p \\ & + \cdots + \frac{2^{k(p-1)}\beta^{(k-1)p}K^{k-1}C^{(k-1)p}t^{(k-1)\kappa}}{(\kappa+1)(2\kappa+1)(3\kappa+1)\cdots((k-2)\kappa+1)}\Delta\|\varphi\|_{m,\infty}^p \\ & + 2^{k(p-1)}\beta^{kp}K^kC^{kp}D\frac{t^{r\kappa}}{(\kappa+1)(2\kappa+1)\cdots((k-2)\kappa+1)((k-1)\kappa+1)}\|\varphi\|_{m,\infty}^p.\end{aligned}$$

Letting $k \to \infty$, we get that[1]

$$\begin{aligned}\tilde{\mathbb{E}}\left[|\mu_t^n(\varphi)|^p\right] = & \|\mu_t^n(\varphi)\|_p^p \\ \leq & 2^{p-1}\Delta\|\varphi\|_{m,\infty}^p + 2^{2(p-1)}\beta^p K C^p t^\kappa \Delta\|\varphi\|_{m,\infty}^p + 2^{3(p-1)}\beta^{2p}K^2C^{2p}\frac{t^{2\kappa}}{\kappa+1}\Delta\|\varphi\|_{m,\infty}^p \\ & + \cdots + \frac{2^{k(p-1)}\beta^{(k-1)p}K^{k-1}C^{(k-1)p}t^{(k-1)\kappa}}{(\kappa+1)(2\kappa+1)(3\kappa+1)\cdots((k-2)\kappa+1)}\Delta\|\varphi\|_{m,\infty}^p \\ & + \cdots\cdots \\ = & 2^{p-1}(\alpha+1)^p\frac{\gamma}{n^q}\|\varphi\|_{m,\infty}^p\sum_{k=1}^{\infty}\left[2^{(k-1)(p-1)}\beta^{(k-1)p}K^{k-1}C^{(k-1)p}\frac{t^{(k-1)\kappa}}{\prod_{j=0}^{k-2}(j\kappa+1)}\right].\end{aligned}$$

Let $\eta_{t,k} = 2^{(k-1)(p-1)}\beta^{(k-1)p}K^{k-1}C^{(k-1)p}\frac{t^{(k-1)\kappa}}{\prod_{j=0}^{k-2}(j\kappa+1)}$, we know $\xi_t \triangleq \sum_{k=1}^{\infty}\eta_{t,k}$ exists by the following ratio test

$$\lim_{k\to\infty}\frac{\eta_{t,k+1}}{\eta_{t,k}} = 2^{p-1}\beta^pKC^p\frac{t^\kappa}{(k-1)\kappa+1} = 0 < 1.$$

Finally the result (A.3) follows by setting $c_t = 2^{p-1}(\alpha+1)^p\gamma\xi_t$. ■

[1] We use the convention that $\prod_{j=0}^{-1} = 1$.

Appendix B

Central Limit Theorem

B.1 Limits of π^n and ρ^n

Lemma B.1.1. *If the approximation π^n is defined by* (3.7)*, in other words,*

$$\pi_t^n(\varphi) = \sum_{j=1}^n \bar{a}_j^n(t) \int_{\mathbb{R}} \varphi\left(v_j^n(t) + y\sqrt{\omega_j^n(t)}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy;$$

then we have

$$\pi_t(\varphi) = \lim_{n\to\infty} \pi_t^n(\varphi) = \lim_{n\to\infty} \sum_{j=1}^n \bar{a}_j^n(t)\varphi(v_j^n(t)). \tag{B.1}$$

That is to say, asymptotically, the variances of the Gaussian measures do not contribute to the approximation, and the combination of positions and weights provide a good approximation.

Proof. Since

$$\begin{aligned}\pi_t^n(\varphi) =& \sum_{j=1}^n \bar{a}_j^n(t) \int_{\mathbb{R}} \varphi\left(v_j^n(t) + y\sqrt{\omega_j^n(t)}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ =& \sum_{j=1}^n \bar{a}_j^n(t) \left\{\varphi(v_j^n(t)) + \frac{1}{2}\omega_j^n(t)\varphi''(v_j^n(t)) + \mathcal{O}\left((\omega_j^n(t))^2\right)\right\},\end{aligned}$$

and $\omega_j^n(t) \sim 1/\sqrt{n}$, it suffices to show that for $k \in \mathbb{N}$ and $k \geq 1$

$$\lim_{n\to\infty} \left|\sum_{j=1}^n \frac{1}{n^{k/2}} \bar{a}_j^n(t)\varphi(v_j^n(t))\right| = 0. \tag{B.2}$$

We know

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|^4\right]\\
=&\frac{1}{n^{2k}}\left\{\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n}\tilde{\mathbb{E}}\left[\bar{a}_{j_1}^n(t)\bar{a}_{j_2}^n(t)\bar{a}_{j_3}^n(t)\bar{a}_{j_4}^n(t)\varphi(v_{j_1}^n(t))\varphi(v_{j_2}^n(t))\varphi(v_{j_3}^n(t))\varphi(v_{j_4}^n(t))\right]\right\}\\
\leq&\frac{\|\varphi\|_{0,\infty}^4}{n^{2k}}\left\{\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n}\sqrt{\tilde{\mathbb{E}}\left[\left(\bar{a}_{j_1}^n(t)\bar{a}_{j_2}^n(t)\right)^2\right]\tilde{\mathbb{E}}\left[\left(\bar{a}_{j_3}^n(t)\bar{a}_{j_4}^n(t)\right)^2\right]}\right\}\\
\leq&\frac{\|\varphi\|_{0,\infty}^2}{n^{2k}}\left\{\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n}\frac{1}{n^4}e^{4c_2t}\right\}=\frac{\|\varphi\|_{0,\infty}^2e^{4c_2t}}{n^{2k}};
\end{aligned}
$$

and thus, for $\varepsilon\in\left(0,\frac{1}{4}\right)$, we have

$$
\tilde{\mathbb{E}}\left[\sum_{n=1}^{\infty}n^{4\varepsilon}\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|^4\right]\leq\|\varphi\|_{0,\infty}^2e^{4c_2t}\sum_{n=1}^{\infty}\frac{1}{n^{2k-4\varepsilon}}<\infty. \tag{B.3}
$$

Let

$$
\psi_\varepsilon^t=\left[\sum_{n=1}^{\infty}n^{4\varepsilon}\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|^4\right]^{\frac{1}{4}},
$$

then ψ_ε^t is integrable, and by (B.3), it is finite a.s.. Also note that for every $n\geq1$,

$$
\left[n^\varepsilon\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|\right]^4\leq\sum_{n=1}^{\infty}\left[n^\varepsilon\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|\right]^4=\left(\psi_\varepsilon^t\right)^4,
$$

therefore

$$
\left|\sum_{j=1}^{n}\frac{1}{n^{k/2}}\bar{a}_j^n(t)\varphi(v_j^n(t))\right|\leq\frac{\psi_\varepsilon^t}{n^\varepsilon}
$$

and (B.2) follows immediately. ■

As a direct consequence, we have the following corollary for the unnormalised approximation ρ^n:

Corollary B.1.2. *If the approximation ρ^n is defined as in* (4.3)*, i.e.*

$$
\rho_t^n(\varphi)=\xi_t^n\pi_t^n(\varphi)=\xi_t^n\sum_{j=1}^{n}\bar{a}_j^n(t)\int_{\mathbb{R}}\varphi\left(v_j^n(t)+y\sqrt{\omega_j^n(t)}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy;
$$

then we have

$$\rho_t(\varphi) = \lim_{n\to\infty} \rho_t^n(\varphi) = \lim_{n\to\infty} \xi_t^n \sum_{j=1}^n \bar{a}_j^n(t)\varphi(v_j^n(t)). \tag{B.4}$$

Remark B.1.3. *By Lemma B.1.1 we know asymptotically as $n \to \infty$, the Gaussian mixture approximation performs just as good as the classic particle filters. Furthermore, from Chapter 8 in [3] and Lemma B.1.1, we know that*

$$\rho_t^n(\varphi) \to \rho_t(\varphi) \quad \textit{and} \quad \pi_t^n(\varphi) \to \pi_t(\varphi) \quad \textit{almost surely.}$$

B.2 Limits of the terms in ρ^n

Lemma B.2.1. *Let $\varphi \in C_0^6(\mathbb{R})$ be a test function, and define the measure-valued processes*

$$\tilde{\rho}_t^{n.1} \triangleq \frac{1}{n}\sum_{j=1}^n \xi_{i\delta}^n a_j^n(t)\delta_{v_j^n(t)}, \quad \tilde{\rho}_t^{n.2} \triangleq \frac{1}{n}\sum_{j=1}^n \left\{\xi_{i\delta}^n a_j^n(t)\right\}^2 \delta_{v_j^n(t)} \tag{B.5}$$

then for any $t \in [0, T]$,

$$\tilde{\rho}_t^{n,1} \to \tilde{\rho}_t^1, \quad \tilde{\rho}_t^{n,2} \to \tilde{\rho}_t^2, \qquad \tilde{\mathbb{P}} - a.s.,$$

where $\tilde{\rho}^1$ and $\tilde{\rho}^2$ are two measure-valued processes satisfying

$$\begin{aligned}
\tilde{\rho}_t^1(\varphi) = \pi_0(\varphi) &+ \int_0^t \Big\{\rho_s(A\varphi) + \pi_s(h)\left[\pi_s(h)\rho_s(\varphi) - \rho_s(h\varphi)\right] \\
&\qquad + \rho_s(h)\left[\pi_s(h\varphi) - \pi_s(h)\pi_s(\varphi)\right]\Big\}ds \\
&+ \int_0^t \Big\{\rho_s(h\varphi) - \pi_s(h)\rho_s(\varphi) + \pi_s(\varphi)\rho_s(h)\Big\}dY_s; \qquad \text{(B.6)}\\
\tilde{\rho}_t^2(\varphi) = \pi_0(\varphi) &+ \int_0^t \Big\{\rho_s(\mathbf{1})\rho_s(A\varphi) - \left[\rho_s(\mathbf{1})\rho_s(h\varphi) - \rho_s(h)\rho_s(\varphi)\right]\pi_s(h) \\
&\qquad + \pi_s(\varphi)(\rho_s(h))^2 + 2\left[\rho_s(h)\rho_s(h\varphi) - (\rho_s(h))^2\pi_s(\varphi)\right]\Big\}ds \\
&+ \int_0^t \left\{\rho_s(\mathbf{1})\rho_s(h\varphi) + \rho_s(h)\rho_s(\varphi)\right\}dY_s. \qquad \text{(B.7)}
\end{aligned}$$

Proof. We begin by noting that for $t \in [i\delta, (i+1)\delta)$

$$\tilde{\rho}_t^{n,1}(\varphi) = \frac{1}{n}\sum_{j=1}^n \xi_{i\delta}^n a_j^n(t)\varphi(v_j^n(t)),$$

and that

$$
\begin{aligned}
\tilde{\rho}_t^{n,1}(\varphi) =& \tilde{\rho}_0^{n,1}(\varphi) + \sum_{i=0}^{[t/\delta]-1} \left(\tilde{\rho}_{(i+1)\delta}^{n,1}(\varphi) - \tilde{\rho}_{(i+1)\delta-}^{n,1}(\varphi) \right) \\
&+ \sum_{i=0}^{[t/\delta]-1} \left(\tilde{\rho}_{(i+1)\delta-}^{n,1}(\varphi) - \tilde{\rho}_{i\delta}^{n,1}(\varphi) \right) + \left(\tilde{\rho}_t^{n,1}(\varphi) - \tilde{\rho}_{[t/\delta]\delta}^{n,1}(\varphi) \right). \qquad \text{(B.8)}
\end{aligned}
$$

By the fact that

$$
\begin{aligned}
d\left(a_j^n(t)\varphi(v_j^n(t))\right) =& a_j^n(t)d\varphi(v_j^n(t)) + \varphi(v_j^n(t))da_j^n(t) + d\langle a_j^n(\cdot), \varphi(v_j^n(\cdot))\rangle_t \\
=& a_j^n(t)\left[(A\varphi)(v_j^n(t))dt + (h\varphi)(v_j^n(t))dY_t + (\varphi'\sigma)(v_j^n(t))dV_t^{(j)}\right] \\
&- \alpha a_j^n(t)\left[(\varphi'\sigma)(v_j^n(t))dV_t^{(j)} + (1-\alpha/2)(\varphi''\sigma^2)(v_j^n(t))dt\right]; \qquad \text{(B.9)}
\end{aligned}
$$

where $\alpha \propto \frac{1}{\sqrt{n}}$. Then we have that, for $t \in [i\delta, (i+1)\delta)$,

$$
\begin{aligned}
&\tilde{\rho}_t^{n,1}(\varphi) - \tilde{\rho}_{i\delta}^{n,1}(\varphi) \\
=& \int_{i\delta}^t d\tilde{\rho}_s^{n,1}(\varphi) = \int_{i\delta}^t \xi_{i\delta}^n \frac{1}{n}\sum_{j=1}^n d\left(a_j^n(s)\varphi(v_j^n(s))\right) \\
=& \int_{i\delta}^t \xi_{i\delta}^n \frac{1}{n}\sum_{j=1}^n \Bigg\{ a_j^n(s)\left[(A\varphi)(v_j^n(s))ds + (h\varphi)(v_j^n(s))dY_s + (\varphi'\sigma)(v_j^n(s))dV_s^{(j)}\right] \\
&\qquad - \alpha a_j^n(s)\left[(\varphi'\sigma)(v_j^n(s))dV_s^{(j)} + (1-\alpha/2)(\varphi''\sigma^2)(v_j^n(s))ds\right] \Bigg\} \\
=& \int_{i\delta}^t \tilde{\rho}_s^{n,1}(A\varphi)ds + \int_{i\delta}^t \tilde{\rho}_s^{n,1}(h\varphi)dY_s + \sqrt{1-\alpha}\frac{1}{n}\sum_{j=1}^n \int_{i\delta}^t \xi_{i\delta}^n a_j^n(s)(\varphi'\sigma)(v_j^n(s))dV_s^{(j)}.
\end{aligned}
$$

Hence we have that, for any $t \in [0,T]$,

$$
\begin{aligned}
&\sum_{i=0}^{[t/\delta]-1} \left(\tilde{\rho}_{(i+1)\delta-}^{n,1}(\varphi) - \tilde{\rho}_{i\delta}^{n,1}(\varphi) \right) + \left(\tilde{\rho}_t^{n,1}(\varphi) - \tilde{\rho}_{[t/\delta]\delta}^{n,1}(\varphi) \right) \\
=& \sum_{i=0}^{\infty} \Bigg\{ \int_{i\delta\wedge t}^{(i+1)\delta\wedge t} \tilde{\rho}_s^{n,1}(A\varphi)ds + \int_{i\delta\wedge t}^{(i+1)\delta\wedge t} \tilde{\rho}_s^{n,1}(h\varphi)dY_s \\
&+ \frac{\sqrt{1-\alpha}}{n}\sum_{j=1}^n \int_{i\delta\wedge t}^{(i+1)\delta\wedge t} \xi_{i\delta}^n a_j^n(s)(\varphi'\sigma)(v_j^n(s))dV_s^{(j)} \Bigg\}. \qquad \text{(B.10)}
\end{aligned}
$$

We now let

$$
\phi_t^n \triangleq \frac{1}{n}\sum_{j=1}^n a_j^n(t)\delta_{v_j^n(t)},
$$

and consider

$$\sum_{i=0}^{[t/\delta]-1}\left(\tilde{\rho}^{n,1}_{(i+1)\delta}(\varphi)-\tilde{\rho}^{n,1}_{(i+1)\delta-}(\varphi)\right)=\sum_{i=0}^{[t/\delta]-1}\xi^n_{i\delta}\left(\phi^n_{(i+1)\delta}(\varphi)-\phi^n_{(i+1)\delta-}(\varphi)\right)$$

$$\begin{aligned}=\sum_{i=0}^{[t/\delta]-1}&\Big\{\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)-\tilde{\mathbb{E}}\left[\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right]\\ &+\tilde{\mathbb{E}}\left[\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right]-\xi^n_{i\delta}\phi^n_{(i+1)\delta-}(\varphi)\Big\}\end{aligned}\tag{B.11}$$

We now consider the two terms in the right hand side of (B.11).

For the first term

$$M^{n,\tilde{\rho}^{n,1},\varphi}_t\triangleq\sum_{i=0}^{[t/\delta]-1}\left\{\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)-\tilde{\mathbb{E}}\left[\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right]\right\},\tag{B.12}$$

by Proposition A.4 and Lemma 4.7 in [56], it is a $\mathcal{F}_t$-adapted martingale and

$$\tilde{\mathbb{E}}\left[\sup_{t\in[0,T]}\left|M^{n,\tilde{\rho}^{n,1},\varphi}_t\right|^4\right]\le\frac{C^T}{n^2}\|\varphi\|^4_{m,\infty};\tag{B.13}$$

and then by virtue of the proof of Lemma B.1.1, we know that

$$\sup_{t\in[0,T]}\left|M^{n,\tilde{\rho}^{n,1},\varphi}_t\right|\to 0\text{ as }n\to\infty.$$

Then, for the second term in (B.11), since $a^n_j((i+1)\delta)=1$, we have

$$\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)=\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n a^n_j((i+1)\delta)\varphi(v^n_j((i+1)\delta))=\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n\varphi(v^n_j((i+1)\delta)),$$

and thus

$$\begin{aligned}\tilde{\mathbb{E}}\left[\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right]=&\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n\tilde{\mathbb{E}}\left[\varphi(v^n_j((i+1)\delta))\middle|\mathcal{F}_{(i+1)\delta-}\right]\\ =&\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^n\left(\sum_{j'=1}^n\bar{a}^n_{j'}((i+1)\delta)\varphi(v^n_{j'}((i+1)\delta))\right)\\ =&\xi^n_{i\delta}\sum_{j'=1}^n\bar{a}^n_{j'}((i+1)\delta)\varphi(v^n_{j'}((i+1)\delta)).\end{aligned}$$

Furthermore,

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left[\xi_{i\delta}^n \phi_{(i+1)\delta}^n(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right] - \xi_{i\delta}^n \phi_{(i+1)\delta-}^n(\varphi)\\
=&\xi_{i\delta}^n \sum_{j=1}^n \bar{a}_j^n((i+1)\delta)\varphi\left(v_j^n((i+1)\delta)\right) - \frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n((i+1)\delta-)\varphi\left(v_j^n((i+1)\delta-)\right)\\
=&\xi_{i\delta}^n\left(\frac{1}{n}\sum_{k=1}^n a_j^n((i+1)\delta)\right)\sum_{j=1}^n \bar{a}_j^n((i+1)\delta)\varphi\left(v_j^n((i+1)\delta)\right)\\
&-\frac{\xi_{i\delta}^n}{n}\sum_{j=1}^n a_j^n((i+1)\delta-)\varphi\left(v_j^n((i+1)\delta-)\right)
\end{aligned}
\tag{B.14}
$$

We now obtain the stochastic differential equation for the terms in (B.14).

First notice that, if let $S_t = \sum_{k=1}^n a_k^n(t)$, then

$$
dS_t = \sum_{k=1}^n a_k^n(t)h(v_k^n(t))dY_t;
$$

and

$$
\begin{aligned}
dS_t^{-1} =& - S_t^{-2}dS_t + S_t^{-3}d\langle S_\cdot\rangle_t\\
=& - S_t^{-2}\sum_{k=1}^n a_k^n(t)h(v_k^n(t))dY_t + S_t^{-3}\left(\sum_{k=1}^n a_k^n(t)h(v_k^n(t))\right)^2 dt.
\end{aligned}
$$

It thus follows that

$$
\begin{aligned}
d\bar{a}_j^n(t) =& d\left(a_j^n(t)S_t^{-1}\right) = a_j^n(t)dS_t^{-1} + S_t^{-1}da_j^n(t) + d\left\langle a_j^n(\cdot), S_\cdot^{-1}\right\rangle_t\\
=&a_j^n(t)\left[-S_t^{-2}\sum_{k=1}^n a_k^n(t)h(v_k^n(t))dY_t + S_t^{-3}\left(\sum_{k=1}^n a_k^n(t)h(v_k^n(t))\right)^2 dt\right]\\
&+ S_t^{-1}a_j^n(t)h(v_j^n(t))dY_t - S_t^{-2}\left[\sum_{k=1}^n a_k^n(t)h(v_k^n(t))\right]a_j^n(t)h(v_j^n(t))dt\\
=&\bar{a}_j^n(t)\left[\left(\sum_{k=1}^n \bar{a}_k^n(t)h(v_k^n(t))\right)^2 - h(v_j^n(t))\left(\sum_{k=1}^n \bar{a}_k^n(t)h(v_k^n(t))\right)\right]dt\\
&+ \bar{a}_j^n(t)\left[h(v_j^n(t)) - \sum_{k=1}^n \bar{a}_k^n(t)h(v_k^n(t))\right]dY_t\\
=&\bar{a}_j^n(t)\left[\eta_t^n dt + \zeta_t^n dY_t\right],
\end{aligned}
\tag{B.15}
$$

where

$$\eta_t^n = \left(\sum_{k=1}^n \bar{a}_k^n(t) h(v_k^n(t))\right)^2 - h(v_j^n(t)) \left(\sum_{k=1}^n \bar{a}_k^n(t) h(v_k^n(t))\right),$$

$$\zeta_t^n = h(v_j^n(t)) - \sum_{k=1}^n \bar{a}_k^n(t) h(v_k^n(t)).$$

Also notice that

$$d\varphi(v_j^n(t)) = \Big[(A\varphi)(v_j^n(t)) + (\alpha/2)(\varphi''\sigma^2)(v_j^n(t))\Big]dt + \sqrt{1-\alpha}(\varphi'\sigma)(v_j^n(t))dV_t^{(j)}.$$

Then we have

$$\begin{aligned}
&d\left[\bar{a}_j^n(t)\varphi(v_j^n(t))\right] \\
=&\bar{a}_j^n(t)d\varphi(v_j^n(t)) + \varphi(v_j^n(t))d\bar{a}_j^n(t) + d\left\langle \bar{a}_j^n(\cdot), \varphi(v_j^n(\cdot))\right\rangle_t \\
=&\bar{a}_j^n(t)\left\{\Big[(A\varphi)(v_j^n(t)) + (\alpha/2)(\varphi''\sigma^2)(v_j^n(t))\Big]dt + \sqrt{1-\alpha}(\varphi'\sigma)(v_j^n(t))dV_t^{(j)}\right\} \\
&+ \varphi(v_j^n(t))\left\{\bar{a}_j^n(t)\left[\eta_t^n dt + \zeta_t^n dY_t\right]\right\} \\
=&\bar{a}_j^n(t)\Bigg\{\Big[(A\varphi)(v_j^n(t)) + (\alpha/2)(\varphi''\sigma^2)(v_j^n(t)) + \varphi(v_j^n(t))\eta_t^n\Big]dt \\
&\qquad + \varphi(v_j^n(t))\zeta_t^n dY_t + \sqrt{1-\alpha}(\varphi'\sigma)(v_j^n(t))dV_t^{(j)}\Bigg\};
\end{aligned} \tag{B.16}$$

and

$$\begin{aligned}
&d\left(S_t\bar{a}_j^n(t)\varphi(v_j^n(t))\right) \\
=&S_t d\left(\bar{a}_j^n(t)\varphi(v_j^n(t))\right) + \bar{a}_j^n(t)\varphi(v_j^n(t))dS_t + \left\langle \bar{a}_j^n(\cdot)\varphi(v_j^n(\cdot)), S_\cdot\right\rangle_t \\
=&S_t\bar{a}_j^n(t)\Bigg\{\Big[(A\varphi)(v_j^n(t)) + (\alpha/2)(\varphi''\sigma^2)(v_j^n(t)) + \varphi(v_j^n(t))\eta_t^n\Big]dt \\
&\qquad + \varphi(v_j^n(t))\zeta_t^n dY_t + \sqrt{1-\alpha}(\varphi'\sigma)(v_j^n(t))dV_t^{(j)}\Bigg\} \\
&+ \bar{a}_j^n(t)\varphi(v_j^n(t))\sum_{k=1}^n a_k^n(t)h(v_k^n(t))dY_t \\
&+ \bar{a}_j^n(t)\varphi(v_j^n(t))\left(h(v_j^n(t)) - \sum_{k=1}^n \bar{a}_k^n(t)h(v_k^n(t))\right)\sum_{k=1}^n a_k^n(t)h(v_k^n(t))dt.
\end{aligned} \tag{B.17}$$

Therefore

$$
\begin{aligned}
&\frac{\xi^n_{i\delta}}{n} S_{(i+1)\delta} \sum_{j=1}^n \bar{a}^n_j((i+1)\delta)\varphi\left(v^n_j((i+1)\delta)\right) \\
=&\frac{\xi^n_{i\delta}}{n} S_{i\delta} \sum_{j=1}^n \bar{a}^n_j(i\delta)\varphi\left(v^n_j(i\delta)\right) + \frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n \int_{i\delta}^{(i+1)\delta} d\left[S_s \bar{a}^n_j(s)\varphi(v^n_j(s))\right] \\
=&\frac{\xi^n_{i\delta}}{n} S_{i\delta} \sum_{j=1}^n \bar{a}^n_j(i\delta)\varphi\left(v^n_j(i\delta)\right) + \frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n \int_{i\delta}^{(i+1)\delta} \sqrt{1-\alpha} S_s \bar{a}^n_j(s)(\varphi'\sigma)(v^n_j(s)) dV^{(j)}_s \\
&+ \frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n \int_{i\delta}^{(i+1)\delta} S_s \bar{a}^n_j(s)\varphi(v^n_j(s)) \left[h(v^n_j(s)) - \sum_{k=1}^n \bar{a}^n_k(s) h(v^n_k(s))\right] dY_s \\
&+ \frac{\xi^n_{i\delta}}{n} \int_{i\delta}^{(i+1)\delta} \sum_{k=1}^n a^n_k(s) h(v^n_k(s)) \sum_{j=1}^n \bar{a}^n_j(s)\varphi(v^n_j(s)) dY_s \\
&+ \frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n \int_{i\delta}^{(i+1)\delta} S_s \bar{a}^n_j(s) \Bigg\{ (A\varphi)(v^n_j(s)) + (\alpha/2)(\varphi''\sigma^2)(v^n_j(s)) \\
&\qquad + \varphi(v^n_j(s)) \left[\left(\sum_{k=1}^n \bar{a}^n_k(s) h(v^n_k(s))\right)^2 - h(v^n_j(s)) \left(\sum_{k=1}^n \bar{a}^n_k(s) h(v^n_k(s))\right)\right] \Bigg\} ds \\
&+ \frac{\xi^n_{i\delta}}{n} \int_{i\delta}^{(i+1)\delta} \sum_{k=1}^n a^n_k(s) h(v^n_k(s)) \sum_{j=1}^n \bar{a}^n_j(s)\varphi(v^n_j(s)) \left(h(v^n_j(s)) - \sum_{k=1}^n \bar{a}^n_k(s) h(v^n_k(s))\right) ds.
\end{aligned}
\tag{B.18}
$$

Because of (B.9),

$$
\begin{aligned}
&\frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n a^n_j((i+1)\delta-)\varphi\left(v^n_j((i+1)\delta-)\right) \\
=&\frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n a^n_j(i\delta)\varphi\left(v^n_j(i\delta)\right) + \frac{\xi^n_{i\delta}}{n} \int_{i\delta}^{(i+1)\delta} a^n_j(t) \Big[(A\varphi)(v^n_j(t)) + (\alpha/2)(\varphi''\sigma^2)(v^n_j(t)\Big] dt \\
&+ \frac{\xi^n_{i\delta}}{n} \int_{i\delta}^{(i+1)\delta} a^n_j(t)\sqrt{1-\alpha}(\varphi'\sigma)(v^n_j(t)) dV^{(j)}_t + \frac{\xi^n_{i\delta}}{n} \int_{i\delta}^{(i+1)\delta} a^n_j(t)(h\varphi)(v^n_j(t)) dY_t. \\
=&\frac{\xi^n_{i\delta}}{n} \sum_{j=1}^n a^n_j(i\delta)\varphi\left(v^n_j(i\delta)\right) + \int_{i\delta}^{(i+1)\delta} \tilde{\rho}^{n,1}_s(A\varphi) ds + \int_{i\delta}^{(i+1)\delta} \tilde{\rho}^{n,1}_s(h\varphi) dY_s \\
&+ (\alpha/2) \int_{i\delta}^{(i+1)\delta} \tilde{\rho}^{n,1}_s(\varphi''\sigma^2) ds + \frac{\sqrt{1-\alpha}}{n} \sum_{j=1}^n \int_{i\delta}^{(i+1)\delta} \xi^n_{i\delta} a^n_j(s)(\varphi'\sigma)(v^n_j(s)) dV^{(j)}_s.
\end{aligned}
$$

Now we can see that

$$\begin{aligned}
&\sum_{i=0}^{[t/\delta]-1}\left(\tilde{\rho}^{n,1}_{(i+1)\delta}(\varphi)-\tilde{\rho}^{n,1}_{(i+1)\delta-}(\varphi)\right)\\
=&M_t^{n,\tilde{\rho}^{n,1},\varphi}+\sum_{i=0}^{[t/\delta]-1}\left\{\tilde{\mathbb{E}}\left[\xi^n_{i\delta}\phi^n_{(i+1)\delta}(\varphi)\middle|\mathcal{F}_{(i+1)\delta-}\right]-\xi^n_{i\delta}\phi^n_{(i+1)\delta-}(\varphi)\right\}\\
=&M_t^{n,\tilde{\rho}^{n,1},\varphi}+\sum_{i=0}^{\infty}\left\{\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^{n}\int_{i\delta}^{(i+1)\delta}\sqrt{1-\alpha}S_s\bar{a}^n_j(s)(\varphi'\sigma)(v^n_j(s))dV^{(j)}_s\right.\\
&+\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^{n}\int_{i\delta}^{(i+1)\delta}S_s\bar{a}^n_j(s)\varphi(v^n_j(s))\left[h(v^n_j(s))-\sum_{k=1}^{n}\bar{a}^n_k(s)h(v^n_k(s))\right]dY_s\\
&+\frac{\xi^n_{i\delta}}{n}\int_{i\delta}^{(i+1)\delta}\sum_{k=1}^{n}a^n_k(s)h(v^n_k(s))\sum_{j=1}^{n}\bar{a}^n_j(s)\varphi(v^n_j(s))dY_s\\
&+\frac{\xi^n_{i\delta}}{n}\sum_{j=1}^{n}\int_{i\delta}^{(i+1)\delta}S_s\bar{a}^n_j(s)\left\{(A\varphi)(v^n_j(s))+(\alpha/2)(\varphi''\sigma^2)(v^n_j(s))\right.\\
&\qquad\left.+\varphi(v^n_j(s))\left[\left(\sum_{k=1}^{n}\bar{a}^n_k(s)h(v^n_k(s))\right)^2-h(v^n_j(s))\left(\sum_{k=1}^{n}\bar{a}^n_k(s)h(v^n_k(s))\right)\right]\right\}ds\\
&\left.+\frac{\xi^n_{i\delta}}{n}\int_{i\delta}^{(i+1)\delta}\sum_{k=1}^{n}a^n_k(s)h(v^n_k(s))\sum_{j=1}^{n}\bar{a}^n_j(s)\varphi(v^n_j(s))\left(h(v^n_j(s))-\sum_{k=1}^{n}\bar{a}^n_k(s)h(v^n_k(s))\right)ds\right\}\\
&-\sum_{i=0}^{\infty}\left\{\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\tilde{\rho}^{n,1}_s(A\varphi)ds+\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\tilde{\rho}^{n,1}_s(h\varphi)dY_s\right.\\
&\qquad\left.+(\alpha/2)\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\tilde{\rho}^{n,1}_s(\varphi''\sigma^2)ds+\frac{1-\alpha}{n}\sum_{j=1}^{n}\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi^n_{i\delta}a^n_j(s)(\varphi'\sigma)(v^n_j(s))dV^{(j)}_s\right\}.
\end{aligned}\tag{B.19}$$

Note the fact that

$$\frac{\xi^n_{i\delta}}{n}S_t=\xi^n_t,$$

then from (B.10) and (B.19), (B.8) becomes

$$
\begin{aligned}
&\tilde{\rho}_t^{n,1}(\varphi) \\
=&\tilde{\rho}_0^{n,1}(\varphi) + M_t^{n,\tilde{\rho}^{n,1},\varphi} \\
+&\sum_{i=0}^{\infty}\Bigg\{\sum_{j=1}^{n}\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\sqrt{1-\alpha}\xi_s^n\bar{a}_j^n(s)(\varphi'\sigma)(v_j^n(s))dV_s^{(j)} \\
&+\sum_{j=1}^{n}\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi_s^n\bar{a}_j^n(s)\varphi(v_j^n(s))\left[h(v_j^n(s))-\sum_{k=1}^{n}\bar{a}_k^n(s)h(v_k^n(s))\right]dY_s \\
&+\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\frac{\xi_{i\delta}^n}{n}\sum_{k=1}^{n}a_k^n(s)h(v_k^n(s))\sum_{j=1}^{n}\bar{a}_j^n(s)\varphi(v_j^n(s))dY_s \\
&+\sum_{j=1}^{n}\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\xi_s^n\bar{a}_j^n(s)\Bigg\{(A\varphi)(v_j^n(s))+(\alpha/2)(\varphi''\sigma^2)(v_j^n(s)) \\
&\qquad+\varphi(v_j^n(s))\left[\left(\sum_{k=1}^{n}\bar{a}_k^n(s)h(v_k^n(s))\right)^2-h(v_j^n(s))\left(\sum_{k=1}^{n}\bar{a}_k^n(s)h(v_k^n(s))\right)\right]\Bigg\}ds \\
&+\int_{i\delta\wedge t}^{(i+1)\delta\wedge t}\frac{\xi_{i\delta}^n}{n}\sum_{k=1}^{n}a_k^n(s)h(v_k^n(s))\sum_{j=1}^{n}\bar{a}_j^n(s)\varphi(v_j^n(s))\left(h(v_j^n(s))-\sum_{k=1}^{n}\bar{a}_k^n(s)h(v_k^n(s))\right)ds\Bigg\}.
\end{aligned}
\tag{B.20}
$$

Similar to the proof of Lemma B.1.1, we have

$$\sup_{t\in[0,T]}\left|\frac{1}{n}\sum_{j=1}^{n}\int_{i\delta}^{t}\xi_{i\delta}^n a_j^n(s)(\varphi'\sigma)(v_j^n(s))dV_s^{(j)}\right|\to 0\quad\text{as } n\to\infty; \tag{B.21}$$

and

$$\sup_{t\in[0,T]}\left|M_t^{n,\tilde{\rho}^{n,1},\varphi}\right|\to 0\quad\text{as } n\to\infty. \tag{B.22}$$

For $\tilde{\rho}_0^{n,1}(\varphi)$, since $a_j^n(0)=1$, then

$$\tilde{\rho}_0^{n,1}(\varphi)=\frac{1}{n}\sum_{j=1}^{n}\xi_0^n a_j^n(0)\varphi(v_j^n(0))=\frac{1}{n}\sum_{j=1}^{n}\varphi(v_j^n(0));$$

by Lemma B.1.1 we obtain

$$\lim_{n\to\infty}\tilde{\rho}_0^{n,1}(\varphi)=\pi_0(\varphi).$$

Finally, using Lemma B.1.1 and Corollary B.1.2, the limiting process

$$\tilde{\rho}_t^1 \triangleq \lim_{n\to\infty} \tilde{\rho}_t^{n,1}$$

satisfies

$$\begin{aligned}\tilde{\rho}_t^1(\varphi) = \pi_0(\varphi) + \int_0^t \Big\{&\rho_s(A\varphi) + \pi_s(h)\left[\pi_s(h)\rho_s(\varphi) - \rho_s(h\varphi)\right] \\ &+ \rho_s(h)\left[\pi_s(h\varphi) - \pi_s(h)\pi_s(\varphi)\right]\Big\} ds \\ + \int_0^t &\Big\{\rho_s(h\varphi) - \pi_s(h)\rho_s(\varphi) + \pi_s(\varphi)\rho_s(h)\Big\} dY_s. \end{aligned} \tag{B.23}$$

Similarly, for the second process

$$\tilde{\rho}_t^{n,2}(\varphi) = \frac{1}{n}\sum_{j=1}^n (\xi_{i\delta}^n)^2 \left(a_j^n(t)\right)^2 \varphi(v_j^n(t));$$

by using exactly the same approach as for $\{\tilde{\rho}^{n,1}\}_n$, we obtain the equation satisfied by its limiting process

$$\tilde{\rho}_t^2(\varphi) \triangleq \lim_{n\to\infty} \tilde{\rho}_t^{n,2}$$

to be

$$\begin{aligned}\tilde{\rho}_t^2(\varphi) = \pi_0(\varphi) + \int_0^t \Big\{&\rho_s(\mathbf{1})\rho_s(A\varphi) - \left[\rho_s(\mathbf{1})\rho_s(h\varphi) - \rho_s(h)\rho_s(\varphi)\right]\pi_s(h) \\ &+ \pi_s(\varphi)(\rho_s(h))^2 + 2\left[\rho_s(h)\rho_s(h\varphi) - (\rho_s(h))^2\pi_s(\varphi)\right]\Big\} ds \\ + \int_0^t &\left\{\rho_s(\mathbf{1})\rho_s(h\varphi) + \rho_s(h)\rho_s(\varphi)\right\} dY_s. \end{aligned} \tag{B.24}$$

The proof is now completed. ■

B.3 Limits of $\sqrt{n}M_{[t/\delta]}^{n,\varphi}$ and $\sqrt{n}B_t^{n,\varphi}$

In order to prove that $\{U^n\}_n$ converges in distribution to a unique process, we should first investigate the limiting processes in the right hand side of

$$\begin{aligned}U_t^n(\varphi) =& U_0^n(\varphi) + \int_0^t U_s^n(A\varphi)ds + \int_0^t U_s^n(h\varphi)dY_s + \sqrt{n}A_{[t/\delta]}^{n,\varphi} + \sqrt{n}G_{[t/\delta]}^{n,\varphi} \\ &+ \frac{1}{\sqrt{n}}\sum_{j=1}^n \int_0^t \xi_{[s/\delta]\delta}^n a_j^n(s)\Big[R_{s,j}^1(\varphi)ds + R_{s,j}^2(\varphi)dY_s + R_{s,j}^3(\varphi)dV_s^{(j)}\Big]; \end{aligned} \tag{B.25}$$

which are what the following lemmas have done.

Lemma B.3.1. *Assume the conditions in Proposition 5.1.6 hold, then*

$$\lim_{n\to\infty} \left\langle \sqrt{n} A^{n,\varphi}_{\cdot} \right\rangle_t = \sum_{i=1}^{[t/\delta]} \left(\rho_{i\delta}(\mathbf{1})\right)^2 \left[\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2\right]. \tag{B.26}$$

If we let

$$\bar{A}^{\varphi}_t \triangleq \sum_{i=1}^{[t/\delta]} \rho_{i\delta}(\mathbf{1}) \sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\, \Upsilon_i, \tag{B.27}$$

where $\{\Upsilon_i\}_{i\in\mathbb{N}}$ is a sequence of independent identically distributed, standard normal random variables, and $\left\{\sqrt{\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2}\,\Upsilon_i\right\}_i$ are mutually independent given the σ-algebra $\mathcal{Y}$; then we have $\langle \bar{A}^{\varphi}_{\cdot}\rangle_t = \lim_{n\to\infty} \langle \sqrt{n} A^{n,\varphi}_{\cdot}\rangle_t$.

Proof. Note that $A^{n,\varphi}$ is a discrete time martingale, then

$$\begin{aligned}
&\lim_n \left\langle \sqrt{n} A^{n,\varphi}_{\cdot} \right\rangle_t \\
=&\lim_n \left\langle \sum_{i=1}^{[\cdot/\delta]} \xi^n_{i\delta} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left[\left(o^{n,i\delta}_j - n\bar{a}^n_j(i\delta-)\right) \varphi(X^n_j(i\delta))\right] \right\rangle_t \\
=&\lim_n \frac{1}{n} \sum_{i=1}^{[t/\delta]} (\xi^n_{i\delta})^2 \tilde{\mathbb{E}}\left[\left(\sum_{j=1}^{n} \left[\left(o^{n,i\delta}_j - n\bar{a}^n_j(i\delta-)\right) \varphi(X^n_j(i\delta))\right]\right)^2 \middle| \mathcal{F}_{i\delta-}\right] \\
=&\lim_n \frac{1}{n} \sum_{i=1}^{[t/\delta]} (\xi^n_{i\delta})^2 \tilde{\mathbb{E}}\Bigg[\sum_{j=1}^{n} \left(o^{n,i\delta}_j - n\bar{a}^n_j(i\delta-)\right)^2 \left(\varphi(X^n_j(i\delta))\right)^2 \\
&\quad + \sum_{l\neq j} \left(o^{n,i\delta}_l - n\bar{a}^n_l(i\delta-)\right)\left(o^{n,i\delta}_j - n\bar{a}^n_j(i\delta-)\right) \varphi(X^n_l(i\delta))\varphi(X^n_j(i\delta)) \Bigg| \mathcal{F}_{i\delta-}\Bigg] \\
=&\lim_n \frac{1}{n} \sum_{i=1}^{[t/\delta]} (\xi^n_{i\delta})^2 \Bigg[\sum_{j=1}^{n} n\bar{a}^n_j(i\delta-)\left(1 - \bar{a}^n_j(i\delta-)\right)\left(\varphi(X^n_j(i\delta))\right)^2 \\
&\qquad - \sum_{l\neq j} n\bar{a}^n_l(i\delta-)\bar{a}^n_j(i\delta-)\varphi(X^n_l(i\delta))\varphi(X^n_j(i\delta))\Bigg] \\
=&\lim_n \sum_{i=1}^{[t/\delta]} (\rho^n_{i\delta}(\mathbf{1}))^2 \left[\sum_{j=1}^{n} \bar{a}^n_j(i\delta-)\left(\varphi(X^n_j(i\delta))\right)^2 - \left(\sum_{j=1}^{n} \bar{a}^n_j(i\delta-)\varphi(X^n_j(i\delta))\right)^2\right] \\
=&\sum_{i=1}^{[t/\delta]} \left(\rho_{i\delta}(\mathbf{1})\right)^2 \left[\pi_{i\delta-}(\varphi^2) - (\pi_{i\delta-}(\varphi))^2\right],
\end{aligned}$$

here we made use of Lemma B.1.1 and Remark B.1.3 in Appendix C.

The second part of the lemma is obvious. ∎

Lemma B.3.2. *Assume the conditions in Proposition 5.1.6 hold, then*

$$\lim_{n\to\infty}\left|\sqrt{n}G^{n,\varphi}_{[t/\delta]}\right| = 0 \quad a.s.. \tag{B.28}$$

Proof. For $G^{n,\varphi}$, we know that

$$\sqrt{n}G^{n,\varphi}_{[t/\delta]} = \sum_{i=1}^{[t/\delta]}\sum_{j=1}^{n}\sqrt{n}\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(i\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(i\delta))\right)\right],$$

first note that $X^n_j(i\delta)\sim N\left(v^n_j(i\delta),\omega^n_j(i\delta)\right)$ and X^n_js are mutually independent ($j = 1,\ldots,n$), also not the fact that $\omega\sim\mathcal{O}(1/\sqrt{n})$; if we let $Z^n_j(i\delta)\triangleq X^n_j(i\delta)-\tilde{\mathbb{E}}\left(X^n_j(i\delta)\right)$ then $Z^n_j(t)\sim\mathcal{N}(0,\omega^n_j(t))$, and then by making use of the central moments of Gaussian random variables, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\left(\sum_{i=1}^{[t/\delta]}\sum_{j=1}^{n}\sqrt{n}\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\left[\varphi(X^n_j(i\delta)) - \tilde{\mathbb{E}}\left(\varphi(X^n_j(i\delta))\right)\right]\right)^{12}\middle|\mathcal{Y}_{i\delta-}\right]\\
\le& 2\|\varphi'\|^{12}_{0,\infty}\tilde{\mathbb{E}}\left[\left(\sum_{i=1}^{[t/\delta]}\sum_{j=1}^{n}\sqrt{n}\xi^n_{i\delta}\bar{a}^n_j(i\delta-)Z^n_j(i\delta)\right)^{12}\middle|\mathcal{Y}_{i\delta-}\right]\\
\le& C^T\left(\|\varphi'\|_{0,\infty}\xi^n_{i\delta}\right)^{12}n^6n^6\sum_{j=1}^{n}\left(\bar{a}^n_j(i\delta-)\right)^8\left(\omega^n_j(i\delta-)\right)^6\\
\le& C^T\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6\alpha^6n^{12}\sum_{j=1}^{n}\left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^{12}\\
=& C^T\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6n^9\sum_{j=1}^{n}\left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^{12};
\end{aligned}$$

then by taking the expectation on both sides, we have

$$\begin{aligned}
\tilde{\mathbb{E}}\left[\left(\sqrt{n}G^{n,\varphi}_{[t/\delta]}\right)^{12}\right] \le& C^T\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6n^9\sum_{j=1}^{n}\tilde{\mathbb{E}}\left[\left(\xi^n_{i\delta}\bar{a}^n_j(i\delta-)\right)^{12}\right]\\
\le& C^T\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6n^9\sum_{j=1}^{n}\sqrt{\tilde{\mathbb{E}}\left[(\xi^n_{i\delta})^{24}\right]\tilde{\mathbb{E}}\left[\left(\bar{a}^n_j(i\delta-)\right)^{24}\right]}\\
\le& C^T\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6n^9\sum_{j=1}^{n}\sqrt{c^{t,24}_1\frac{e^{c_{24}t}}{n^{24}}}\\
=& \frac{\beta^T_{\varphi,\sigma,\delta}}{n^2},
\end{aligned}$$

where

$$\beta^T_{\varphi,\sigma,\delta} = C^T\sqrt{c_1^{T,24}e^{c_{24}T}}\|\varphi\|^{12}_{1,\infty}\|\sigma\|^{12}_{0,\infty}\delta^6$$

is a constant independent of n. Then similar to the proof of Lemma B.1.1 in Appendix C, we have the result. ■

Lemma B.3.3. *Assume the conditions in Proposition 5.1.6 hold, then*

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \xi^n_{[s/\delta]\delta}a^n_j(s)R^1_{s,j}(\varphi)ds = \Lambda^{R^1,\varphi}_t, \tag{B.29}$$

where

$$\Lambda^{R^1,\varphi}_t = c_\omega\int_0^t \tilde{\rho}^1_s(\Psi\varphi)ds; \tag{B.30}$$

c_ω is a constant, and the operator Ψ is defined by

$$\Psi\varphi = \frac{f\varphi'''}{2} + \frac{\sigma\varphi^{(4)}}{4} - \frac{3(A\varphi)''}{2}.$$

Proof. Since

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \xi^n_{[s/\delta]\delta}a^n_j(s)R^1_{s,j}(\varphi)ds\\
&=\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \xi^n_{[s/\delta]\delta}a^n_j(s)\left\{\omega^n_j(s)\left[\left(\frac{f\varphi'''}{2}+\frac{\sigma\varphi^{(4)}}{4}\right)(v^n_j(s)) - I_j(A\varphi)\right]\right\}ds\\
&=\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^t \xi^n_{[s/\delta]\delta}a^n_j(s)\left\{\omega^n_j(s)\left[\left(\frac{f\varphi'''}{2}+\frac{\sigma\varphi^{(4)}}{4}-\frac{3(A\varphi)''}{2}\right)(v^n_j(s))\right]\right\}ds\\
&=\lim_{n\to\infty}\frac{c_\omega}{n}\int_0^t\sum_{j=1}^{n} \xi^n_{[s/\delta]\delta}a^n_j(s)\left\{\left(\frac{f\varphi'''}{2}+\frac{\sigma\varphi^{(4)}}{4}-\frac{3(A\varphi)''}{2}\right)(v^n_j(s))\right\}ds\\
&=\lim_{n\to\infty}c_\omega\int_0^t\tilde{\rho}^{n,1}_s(\Psi\varphi)ds = c_\omega\int_0^t\tilde{\rho}^1_s(\Psi\varphi)ds,
\end{aligned}$$

we have the required result. ■

Lemma B.3.4. *Assume the conditions in Proposition 5.1.6 hold, then*

$$\lim_{n\to\infty}\left\langle\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^\cdot \xi^n_{[s/\delta]\delta}a^n_j(s)R^2_{s,j}(\varphi)dY_s\right\rangle_t = \left\langle\Lambda^{R^2,\varphi}_\cdot\right\rangle_t, \tag{B.31}$$

where

$$\Lambda^{R^2,\varphi}_t = c_\omega\int_0^t\left(\tilde{\rho}^1_s(h\varphi''-(h\varphi)'')\right)dB^{(2)}_s, \tag{B.32}$$

c_ω is a constant and $B^{(2)}$ is a Brownian motion independent of Y.

Proof. Observe that

$$
\begin{aligned}
&\lim_{n\to\infty}\left\langle \int_0^{\cdot} \frac{1}{\sqrt{n}}\sum_{j=1}^n \xi^n_{[s/\delta]\delta} a^n_j(s) R^2_{s,j}(\varphi) dY_s \right\rangle_t \\
=&\lim_{n\to\infty}\int_0^t \left(\frac{1}{\sqrt{n}}\sum_{j=1}^n \xi^n_{[s/\delta]\delta} a^n_j(s) R^2_{s,j}(\varphi)\right)^2 ds \\
=&\lim_{n\to\infty}\int_0^t \left(\frac{1}{\sqrt{n}}\sum_{j=1}^n \xi^n_{[s/\delta]\delta} a^n_j(s)\omega^n_j(s)\Big(\frac{h\varphi''}{2}(v^n_j(s)) - I_j(h\varphi)\Big)\right) ds \\
=&\lim_{n\to\infty}\int_0^t \left(\frac{1}{2\sqrt{n}}\sum_{j=1}^n \xi^n_{[s/\delta]\delta} a^n_j(s)\omega^n_j(s)\left[(h\varphi'' - (h\varphi)'')(v^n_j(s))\right]\right)^2 ds \\
=&\lim_{n\to\infty} c_\omega^2 \int_0^t \left(\frac{1}{n}\sum_{j=1}^n \xi^n_{[s/\delta]\delta} a^n_j(s)\left[(h\varphi'' - (h\varphi)'')(v^n_j(s))\right]\right)^2 ds \\
=&\lim_{n\to\infty} c_\omega^2 \int_0^t \left(\tilde{\rho}^{n,1}_s\left(h\varphi'' - (h\varphi)''\right)\right)^2 ds \\
=&c_\omega^2 \int_0^t \left(\tilde{\rho}^{1}_s\left(h\varphi'' - (h\varphi)''\right)\right)^2 ds = \left\langle \Lambda^{R^2,\varphi}_{\cdot}\right\rangle_t;
\end{aligned}
$$

and then we have the result. ∎

Lemma B.3.5. *Assume the conditions in Proposition 5.1.6 hold, then*

$$
\lim_{n\to\infty}\left\langle \frac{1}{\sqrt{n}}\sum_{j=1}^n \int_0^{\cdot} \xi^n_{[s/\delta]\delta} a^n_j(s) R^3_{s,j}(\varphi) dV^{(j)}_s \right\rangle_t = \left\langle \Lambda^{R^3,\varphi}_{\cdot}\right\rangle_t, \tag{B.33}
$$

where

$$
\Lambda^{R^3,\varphi}_t = \int_0^t \sqrt{\tilde{\rho}^2_s\left((\sigma\varphi')^2\right)} dB^{(3)}_s,
$$

$B^{(3)}$ *is a Brownian motion independent of* $B^{(2)}$ *and* Y*.*

Proof. Bearing in mind that $\omega^n_j \propto \frac{1}{\sqrt{n}}$, then using exactly the same approach as in the proof of Lemma B.1.1 in Appendix C, we have

$$
\begin{aligned}
&\lim_{n\to\infty} \left\langle \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \int_0^{\cdot} \xi^n_{[s/\delta]\delta} a^n_j(s) R^3_{s,j}(\varphi) dV^{(j)}_s \right\rangle_t \\
&= \lim_{n\to\infty} \int_0^t \frac{1}{n} \sum_{j=1}^{n} \left(\xi^n_{[s/\delta]\delta} a^n_j(s)\right)^2 \left(R^3_{s,j}(\varphi)\right)^2 ds \\
&= \lim_{n\to\infty} \int_0^t \frac{1}{n} \sum_{j=1}^{n} \left(\xi^n_{[s/\delta]\delta} a^n_j(s)\right)^2 \left((\sigma\varphi')(v^n_j(s))\right)^2 ds \\
&= \lim_{n\to\infty} \int_0^t \tilde{\rho}^{n,2}_s \left((\sigma\varphi')^2\right) ds = \int_0^t \tilde{\rho}^2_s \left((\sigma\varphi')^2\right) ds = \left\langle \Lambda^{R^3,\varphi}_{\cdot} \right\rangle_t .
\end{aligned}
\tag{B.34}
$$

We then have the result. ■

Remark B.3.6. *From the above arguments we can see that, asymptotically, the variances* $\{\omega^n_j\}_{j=1}^n$ *do not contribute to the approximating system.*

Appendix C

Generalised Particle Filters

C.1 Wavelets

A wavelet is a wave-like oscillation with an amplitude that starts out at zero, increases, and then decreases back to zero. Mathematically speaking, a *wavelet* is a function used to divide a given function or continuous-time signal into different scale components. Usually one can assign a frequency range to each scale component. Each scale component can then be studied with a resolution that matches its scale. A *wavelet transform* is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as *daughter wavelets*) of a finite-length or fast-decaying oscillating waveform (known as the *mother wavelet*). The main purpose of the mother wavelet is to provide a source function to generate daughter wavelets which are translated and dilated versions of the mother wavelet.

In formal terms, this representation is a wavelet series representation of a square-integrable function by a complete and orthonormal set of basis functions for the Hilbert space of square integrable functions; and this orthonormal series is generated by the mother wavelet. Wavelet transforms are classified into *discrete wavelet transforms* (DWTs) and *continuous wavelet transforms* (CWTs). Note that both DWT and CWT are continuous-time transforms, and thus can both be used to represent continuous-time signals.

Definition C.1.1. *The continuous wavelet transform of a function $f \in L^2(\mathbb{R})$ with* *respect to some mother wavelet ψ is defined as*

$$W_\psi f(a,b) = \int_{-\infty}^{\infty} f(t)\overline{\psi_{a,b}(t)}dt, \tag{C.1}$$

where

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right). \tag{C.2}$$

The parameters a and b are called dilation and translation parameters respectively. In order to reconstruct the original function $f(x)$, inverse continuous wavelet transform is given as

$$f(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{a^2}\left[W_\psi f(a,b)\right]\frac{1}{\sqrt{|a|}}\tilde{\psi}\left(\frac{t-b}{a}\right) db\ da, \tag{C.3}$$

where $\tilde{\psi}$ is the dual of ψ satisfying

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{|a^3|}\psi\left(\frac{t_1-b}{a}\right)\tilde{\psi}\left(\frac{t-b}{a}\right) db\ da = \delta(t-t_1).$$

Usually $\tilde{\psi}(t) = C_\psi^{-1}\psi(t)$, where the constant C_ψ satisfying

$$C_\psi = \int_{-\infty}^{\infty} \frac{\left|\hat{\psi}(\xi)\right|^2}{\xi} d\xi < \infty$$

is called the admissibility condition, and $\hat{\psi}$ is the Fourier transform of ψ.

Definition C.1.2. *A discrete wavelet transform is any wavelet transform for which the wavelets are discretely sampled. Specially, this is done by modifying the wavelets* (C.2) *into the following expression*

$$\psi_{j,k}(t) = \frac{1}{\sqrt{|a|^j}}\psi\left(\frac{t-kba^j}{a^j}\right) = |a|^{-\frac{j}{2}}\psi(a^{-j}t-kb) \tag{C.4}$$

where a is the fixed dilation step and b is the translation step which depends on the dilation step a. The discrete wavelet series can be made orthonormal by the choices of dilation and translation parameters and the mother wavelet. By orthonormality we mean that it can be used to define a Hilbert basis:

$$\left\langle \psi_{j,k}, \overline{\psi_{m,n}} \right\rangle = \int_{-\infty}^{\infty} \psi_{j,k}(x)\psi_{m,n}(x)dx = \begin{cases} 1 & if\ j=m\ and\ k=n \\ 0 & otherwise \end{cases} \tag{C.5}$$

Now an arbitrary function $f(x)$ can be reconstructed in the following way

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \left\langle f, \overline{\psi_{j,k}} \right\rangle \psi_{j,k}(t). \tag{C.6}$$

This expression is the inverse wavelet transform for discrete wavelets.

In the following, I will list some examples of (mother) wavelet functions.

Example C.1.3 (Haar wavelet)**.** *The Haar wavelet is the simplest possible wavelet, its mother wavelet function $\psi(x)$ is defined as*

$$\psi(x) = \begin{cases} 1, & if\ 0 < x \leq \frac{1}{2} \\ -1, & if\ \frac{1}{2} < x \leq 1 \\ 0, & otherwise \end{cases} \tag{C.7}$$

The support of ψ is $[0,1]$. We dilate ψ by powers of 2, and translate the dilate by 2^{-j} times an integer, in order to get the daughter wavelets

$$\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k).$$

It is easy to show that $\{\psi_{j,k} : j,k \in \mathbb{Z}\}$ is an orthonormal family. However, the discontinuity, and therefore the non-differentiability, is the main technical disadvantage of the Haar wavelet.

Example C.1.4 (Daubechies wavelets)**.** *The system of Daubechies wavelets is an expansion of Haar wavelets with the scaling function $\varphi(x) = \mathbf{1}_{[0,1]}(x)$. Note that for the wavelet function defined in* (C.7)*, we have*

$$\psi(x) = \varphi(2x) - \varphi(2x-1), \qquad \varphi(x) = \varphi(2x) + \varphi(2x-1).$$

It is observed that φ on a larger scale is essentially the same as φ on a smaller scale. The scaling function for the Daubechies wavelets satisfies a more complicated scaling function:

$$\varphi(x) = \sum_{k=0}^{N} a_k \varphi(2x-k), \tag{C.8}$$

where the coefficients a_k must be chosen with great care. Once the values of φ on integers are known, the values on half-integers can be obtained; and inductively so *are the values of φ at dyadic rationals $m2^{-j}$. This, by continuity, determines $\varphi(x)$ for all x. Hence the wavelet ψ is determined by the identity*

$$\psi(x) = \sum_{k=0}^{N} (-1)^k a_{N-k} \varphi(2x - k). \tag{C.9}$$

Example C.1.5 (Harmonic wavelets). *The harmonic wavelet is defined as:*

$$\psi(x) = \frac{\exp(i4\pi x) - \exp(i2\pi x)}{i2\pi x}, \tag{C.10}$$

and we can conclude that the wavelets series defined by

$$\psi(2^j x - k) = \frac{\exp(i4\pi(2^j x - k)) - \exp(i2\pi(2^j x - k))}{i2\pi(2^j x - k)} \tag{C.11}$$

forms an orthogonal set.

C.2 Finite Elements

The definition of a finite element was initially given by Ciarlet in 1978.

Definition C.2.1. *A **finite element** consists of a triple $(K, \mathcal{P}, \mathcal{N})$ where:*

1. *K is a compact, connected, Lipschitz subset of $\mathbb{R}^d$ with non-empty interior;*
2. *$\mathcal{P}$ is a finite-dimensional space of functions on K, i.e. $p \in \mathcal{P} : K \mapsto \mathbb{R}^m$;*
3. *$\mathcal{N} = \{N_1, N_2, \ldots, N_k\}$ is a basis for $\mathcal{P}'$, where $\mathcal{P}'$ is the dual space of $\mathcal{P}$, that is, a set of linear functionals on the Banach space $\mathcal{P}$.*

In the definition, K is called the element domain, $\mathcal{P}$ is called the space of shape functions; and $\mathcal{N}$ is the set of nodal variables.

Definition C.2.2. *Let $(K, \mathcal{P}, \mathcal{N})$ be a finite element. Then the basis $\{\phi_1, \phi_2, \ldots, \phi_k\}$ of $\mathcal{P}$ dual to $\mathcal{N}$ (i.e. $N_i(\phi_j) = \delta_{ij}$) is called the **nodal basis** of $\mathcal{P}$.*

The following proposition simplifies the verification of the third condition of Definition C.1.

Proposition C.2.3. *Let $\mathcal{P}$ be a d-dimensional vector space and let $\{N_1, N_2, \ldots, N_d\}$ be a subset of the dual space $\mathcal{P}'$, then the following two statements are equivalent.*

- $\{N_1, N_2, \ldots, N_d\}$ *is a basis for* $\mathcal{P}'$.
- *Given* $v \in \mathcal{P}$ *with* $N_i v = 0$ *for* $i = 1, 2, \ldots, d$, *then* $v \equiv 0$.

Proof. See Lemma 3.1.4 in [5]. ■

Definition C.2.4 (Mesh). *Let* Ω *be a domain in* $\mathbb{R}^d$. *A mesh is a union of a finite number (N) of compacted, Lipschitz sets* K_m *with non-empty interior such that* $\{K_1, K_2, \ldots, K_N\}$ *forms a partition of* Ω*,i.e.,*

$$\overline{\Omega} = \cup_{m=1}^{N} K_m \qquad and \qquad int K_m \cap int K_n = \phi \quad for \quad m \neq n.$$

The subsets K_m *are called* ***mesh cells*** *or* ***mesh elements****. A mesh* $\{K_1, K_2, \ldots, K_N\}$ *is denoted by* $\mathcal{T}_h$. *The subscript* $h = \max_{K \in \mathcal{T}_h} h_K$, *where* $\forall K \in \mathcal{T}_h$, $h_K = diam(K) = \max_{x_1, x_2 \in K} \|x_1 - x_2\|_d$.

Now we introduce the interpolants, we begin by defining the local interpolant.

Definition C.2.5. *Given a finite element* $(K, \mathcal{P}, \mathcal{N})$, *let the set* $\{\phi_1, \phi_2, \ldots, \phi_k\} \subseteq \mathcal{P}$ *be the basis dual to* $\mathcal{N}$. *If* v *is a function for which all* $N_i \in \mathcal{N}$, $i = 1, \ldots, k$, *are defined, then we define the* ***local interpolant*** *by*

$$\mathcal{I}_K v \triangleq \sum_{i=1}^{k} N_i(v)\phi_i \tag{C.12}$$

Definition C.2.6. *Suppose* Ω *is a domain with a mesh* $\mathcal{T}$. *Assume each element domain,* $K \in \mathcal{T}$, *is equipped with some type of shape functions* $\mathcal{P}$ *and nodal variables* $\mathcal{N}$, *such that* $(K, \mathcal{P}, \mathcal{N})$ *forms a finite element. Let* m *be the order of the highest partial derivatives involved in the nodal variables. For* $f \in C^m(\overline{\Omega})$, *the* ***global interpolant*** *is defined by*

$$\mathcal{I}_{\mathcal{T}} f = \sum_{m=1}^{N} \mathcal{I}_{K_m} f = \sum_{m=1}^{N} \sum_{i=1}^{k_m} N_i^m(f)\phi_i^m, \quad for\ all\ K_m \in \mathcal{T}. \tag{C.13}$$

In the applications of the global interpolation, it is essential to find a uniform bound for the norm of the local interpolation operator $\mathcal{I}_{\mathcal{T}}$. It is therefore necessary to compare the local interpolation operators on different elements. The following notion of *affine equivalent* can be shown as an equivalent relation (see Exercise (3.4.4) in [5]).

Definition C.2.7. *Let* $(K, \mathcal{P}, \mathcal{N})$ *be a finite element and let* $F(x) = \mathbf{A}x + \mathbf{b}$ *be an affine map. The finite element* $(\widehat{K}, \widehat{\mathcal{P}}, \widehat{\mathcal{N}})$ *is **affine equivalent** to* $(K, \mathcal{P}, \mathcal{N})$ *if*

- $F(K) = \widehat{K}$*;*
- $F^*\widehat{\mathcal{P}} = \mathcal{P}$*, where* $F^*(\hat{f}) \triangleq \hat{f} \circ F$*;*
- $F_*\mathcal{N} = \widehat{\mathcal{N}}$*, where* $(F_*N)(\hat{f}) \triangleq N(F^*(\hat{f}))$*.*

We write $(K, \mathcal{P}, \mathcal{N}) \simeq_F (\widehat{K}, \widehat{\mathcal{P}}, \widehat{\mathcal{N}})$ *if they are affine equivalent.*

Definition C.2.8. *The finite elements* $(K, \mathcal{P}, \mathcal{N})$ *and* $(K, \mathcal{P}, \widetilde{\mathcal{N}})$ *are **interpolation equivalent** if*

$$\mathcal{I}_{\mathcal{N}} f = \mathcal{I}_{\widetilde{\mathcal{N}}} f, \qquad \forall f \text{ sufficiently smooth}, \tag{C.14}$$

and is written as $(K, \mathcal{P}, \mathcal{N}) \simeq_{\mathcal{I}} (K, \mathcal{P}, \widetilde{\mathcal{N}})$*.*

Several examples of finite elements can be found in [5] and [34].

Appendix D

Navier-Stokes Equation

D.1 The Inner Product on H

Recall we have defined

$$\mathcal{H} \triangleq \Bigg\{ L - \text{periodic trigonometirc polynomials } u :$$
$$[0,L)^2 \to \mathbb{R}^2 \Bigg| \nabla \cdot u = 0, \int_{\mathbb{T}^2} u(x)dx = 0 \Bigg\}$$

and H as the closure of $\mathcal{H}$ with respect to the $(L^2(\mathbb{T}^2))^2$ norm. We also defined $P : (L^2(\mathbb{T}^2))^2 \to H$ to be the Leray-Helmholtz orthogonal projector.

Proposition D.1.1. *Given $u = (u_1, u_2) \in H, v = (v_1, v_2) \in H$, (i.e. $x \mapsto u(x) : [0,L)^2 \mapsto \mathbb{C}^2$ and $x \mapsto v(x) : [0,L)^2 \mapsto \mathbb{C}^2$; $x \mapsto u_1(x), u_2(x), v_1(x)$ or $v_2(x) : [0,L)^2 \mapsto \mathbb{C}$) we define the function $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ as follows:*

$$\langle u, v \rangle = \frac{1}{L^2} \int_{\mathbb{T}^2} (u \cdot \overline{v})(x)dx, \tag{D.1}$$

where $u \cdot v = u_1v_1 + u_2v_2$, $x = (x_1, x_2) \in \mathbb{T}^2 = [0,L)^2$. Then $\langle \cdot, \cdot \rangle$ is an inner product on H.

Proof. First we know that for $u, v, w \in H$ and $\alpha, \beta \in \mathbb{C}$

$$\begin{aligned}
\langle u, \alpha v + \beta w \rangle &= \frac{1}{L^2} \int_{\mathbb{T}^2} (u \cdot \overline{\alpha v + \beta w})(x)dx = \frac{1}{L^2} \int_{\mathbb{T}^2} (u \cdot (\overline{\alpha}\ \overline{v} + \overline{\beta}\ \overline{w}))(x)dx \\
&= \overline{\alpha} \frac{1}{L^2} \int_{\mathbb{T}^2} (u \cdot \overline{v})(x)dx + \overline{\beta} \frac{1}{L^2} \int_{\mathbb{T}^2} (u \cdot \overline{w})(x)dx \\
&= \overline{\alpha} \langle u, v \rangle + \overline{\beta} \langle u, w \rangle;
\end{aligned} \tag{D.2}$$

similarly we have

$$\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle. \tag{D.3}$$

It is obvious that

$$\overline{\langle u, v\rangle} = \frac{1}{L^2}\int_{\mathbb{T}^2} (\overline{u\cdot \overline{v}})(x)dx = \frac{1}{L^2}\int_{\mathbb{T}^2} (v\cdot \overline{u})(x)dx = \langle v, u\rangle; \tag{D.4}$$

and

$$\langle u, u\rangle \geq 0 \quad \text{and} \quad \langle u, u\rangle = 0 \Leftrightarrow u = (0,0). \tag{D.5}$$

Then by definition we have $\langle \cdot, \cdot\rangle$ is an inner product. ■

Given $k = (k_1, k_2)^\top$, define $k^\perp = (k_2, -k_1)^\top$. Then under the above defined inner product, an orthonormal basis for H is given by $\psi_k : \mathbb{R}^2 \to \mathbb{C}^2$, where

$$\psi_k(x) \triangleq \frac{k^\perp}{|k|}\exp\left(\frac{2\pi i k\cdot x}{L}\right)$$

for $k \in \mathbb{Z}^2 \setminus \{0\}$.

D.2 Calculation of $\alpha_k^{l,j}$

From the definition of the inner production, we should have

$$\alpha_k^{l,j} = \frac{1}{L^2}\int_{\mathbb{T}^2} \left(P(\psi_l\cdot\nabla\psi_j)\cdot\overline{\psi_k}\right)(x)dx,$$

but note that $\psi_l\cdot\nabla\psi_j \in H$ (see (D.9) below), so we can write $\alpha_k^{l,j}$ as

$$\alpha_k^{l,j} = \frac{1}{L^2}\int_{\mathbb{T}^2} \left((\psi_l\cdot\nabla\psi_j)\cdot\overline{\psi_k}\right)(x)dx, \tag{D.6}$$

where $k = (k_1, k_2)^\top$ and $x = (x_1, x_2)$. In order to calculate $\alpha_k^{l,j}$, we firstly write $\psi_k(x)$ as

$$\begin{aligned}\psi_k(x) &= \left(\frac{k_2}{|k|}\exp\left(\frac{2\pi i(k_1x_1 + k_2x_2)}{L}\right), -\frac{k_1}{|k|}\exp\left(\frac{2\pi i(k_1x_1 + k_2x_2)}{L}\right)\right)^\top \\ &\triangleq \left(\psi_k^1(x), \psi_k^2(x)\right)^\top;\end{aligned}$$

and therefore

$$
\begin{aligned}
&\overline{\psi_k(x)} \\
&= \left(\frac{k_2}{|k|} \exp\left(\frac{2\pi i(-k_1 x_1 - k_2 x_2)}{L} \right), - \frac{k_1}{|k|} \exp\left(\frac{2\pi i(-k_1 x_1 - k_2 x_2)}{L} \right) \right)^\top \\
&\triangleq \left(\psi^1_{-k}(x), \psi^2_{-k}(x) \right)^\top = \psi_{-k}(x);
\end{aligned}
$$

Letting $l = (l_1, l_2)^\top \in \mathbb{Z}^2 \setminus \{0\}$ and $j = (j_1, j_2)^\top \in \mathbb{Z}^2 \setminus \{0\}$, similarly for ψ_l and ψ_j we have

$$
\begin{aligned}
(\psi_l \cdot \nabla)\psi_j &= (\psi_l^1 \frac{\partial}{\partial x_1} + \psi_l^2 \frac{\partial}{\partial x_2}) \left(\psi_j^1, \psi_j^2 \right)^\top \\
&= \left(\psi_l^1 \frac{\partial \psi_j^1}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^1}{\partial x_2},\ \psi_l^1 \frac{\partial \psi_j^2}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^2}{\partial x_2} \right)^\top
\end{aligned}
\tag{D.7}
$$

and

$$
(\psi_l \cdot \nabla \psi_j) \cdot \overline{\psi_k} = \psi^1_{-k} \left(\psi_l^1 \frac{\partial \psi_j^1}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^1}{\partial x_2} \right) + \psi^2_{-k} \left(\psi_l^1 \frac{\partial \psi_j^2}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^2}{\partial x_2} \right). \tag{D.8}
$$

Simple calculation gives us

$$
\begin{aligned}
\frac{\partial \psi_j^1}{\partial x_1} &= \frac{2\pi i\ j_1 j_2}{L\ |j|} \exp\left(\frac{2\pi i j \cdot x}{L} \right), \qquad & \frac{\partial \psi_j^1}{\partial x_2} &= \frac{2\pi i\ j_2^2}{L\ |j|} \exp\left(\frac{2\pi i j \cdot x}{L} \right), \\
\frac{\partial \psi_j^2}{\partial x_1} &= -\frac{2\pi i\ j_1^2}{L\ |j|} \exp\left(\frac{2\pi i j \cdot x}{L} \right), \qquad & \frac{\partial \psi_j^2}{\partial x_2} &= -\frac{2\pi i\ j_1 j_2}{L\ |j|} \exp\left(\frac{2\pi i j \cdot x}{L} \right);
\end{aligned}
$$

and

$$
\begin{aligned}
(\psi_l \cdot \nabla \psi_j)(x) &= \frac{2\pi i(l_2 j_1 - l_1 j_2)}{L\ |l|} \exp\left(\frac{2\pi i l \cdot x}{L} \right) \exp\left(\frac{2\pi i j \cdot x}{L} \right) \left(\frac{j_2}{|j|}, \frac{-j_1}{|j|} \right)^\top \\
&= \frac{2\pi i(l_2 j_1 - l_1 j_2)}{L\ |l|} \exp\left(\frac{2\pi i l \cdot x}{L} \right) \psi_j(x);
\end{aligned}
\tag{D.9}
$$

and

$$
\begin{aligned}
\psi^1_{-k} \left(\psi_l^1 \frac{\partial \psi_j^1}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^1}{\partial x_2} \right) &= \frac{2\pi i(l_2 j_1 - l_1 j_2) k_2 j_2}{L\ |k||l||j|} \exp\left(\frac{2\pi i(l + j - k) \cdot x}{L} \right) \\
\psi^2_{-k} \left(\psi_l^1 \frac{\partial \psi_j^2}{\partial x_1} + \psi_l^2 \frac{\partial \psi_j^2}{\partial x_2} \right) &= \frac{2\pi i(l_2 j_1 - l_1 j_2) k_1 j_1}{L\ |k||l||j|} \exp\left(\frac{2\pi i(l + j - k) \cdot x}{L} \right).
\end{aligned}
$$

Then by (D.8) we know that

$$
\begin{aligned}
&(\psi_l \cdot \nabla \psi_j) \cdot \overline{\psi_k} \\
=&\frac{2\pi i(l_2 j_1 - l_1 j_2)(k_1 j_1 + k_2 j_2)}{L\ |k||l||j|} \exp\left(\frac{2\pi i(l+j-k)\cdot x}{L}\right). \\
=&\frac{2\pi i(l_2 j_1 - l_1 j_2)(k_1 j_1 + k_2 j_2)}{L\ |k||l||j|} \exp\left(\frac{2\pi i((l_1+j_1-k_1)x_1 + (l_2+j_2-k_2)x_2)}{L}\right)
\end{aligned}
\tag{D.10}
$$

We therefore obtain

$$
\alpha_k^{l,j} = \frac{1}{L^2}\int_{\mathbb{T}^2} \left((\psi_l \cdot \nabla \psi_j) \cdot \overline{\psi_k}\right)(x)dx = \frac{1}{L^2}\int_0^L \int_0^L \left((\psi_l \cdot \nabla \psi_j) \cdot \overline{\psi_k}\right)(x)dx_1 dx_2,
$$

in other words,

$$
\alpha_k^{l,j} = \begin{cases} \frac{2\pi i(l_2 j_1 - l_1 j_2)(k_1 j_1 + k_2 j_2)}{L\ |k||l||j|} & \text{if} \quad k = l + j; \\ \\ 0 & \text{otherwise.} \end{cases} \tag{D.11}
$$

D.3 The Decay of the Fourier Coefficients

In this section, we show in the periodic case (for deterministic Navier-Stokes equation), for an initial condition in the space V, the corresponding strong solution becomes analytic both in space and time. After establishing the space analyticity of the solutions in the 2-dimensional periodic case, we derive, as a consequence, the exponential decay of the Fourier coefficients with respect to their Fourier mode. The content of this section can be found in [37].

For each $\sigma, s > 0$, the Gevrey space $D(\exp(\sigma A^s))$ is defined as the domain of the exponential of σA^s, where A is the Stokes operator. We will give a precise characterisation of this space by means of Fourier series as follows. We know that a vector field $u \in H$ is characterised in terms of Fourier series as a function

$$
u = \sum_{k \in \mathbb{Z}^d} u_k e^{2\pi i \frac{k}{L}\cdot x}, \quad u_k \in \mathbb{C}^d,\ u_{-k} = \bar{u}_k, \tag{D.12}
$$

such that

$$
\frac{k}{L} \cdot u_k = 0 \quad \text{for all } k \in \mathbb{Z}^d; \qquad \text{and} \qquad |u|^2 = \sum_{k \in \mathbb{Z}^d} |u_k|^2 < \infty. \tag{D.13}
$$

For the Gevrey space, we can define the operator $\exp(\sigma A^s)$ in Fourier space by

$$\exp(\sigma A^s)u = \sum_{k\in\mathbb{Z}^d} \exp\left(\sigma\left(2\pi\frac{k}{L}\right)^{2s}\right) u_k \exp\left(2\pi i\frac{k}{L}\cdot x\right). \tag{D.14}$$

The domain $D(\exp(\sigma A^s))$ is defined as usual by

$$D(\exp(\sigma A^s)) = \left\{u\in H : e^{\sigma A^s}u\in H\right\}.$$

Therefore, a vector field $u\in D(\exp(\sigma A^s))$ can be characterised in terms of Fourier series representation by the divergence-free condition and by the condition that the Fourier coefficients decay exponentially fast in the sense that

$$\sum_{k\in\mathbb{Z}^d} e^{2\sigma\left|2\pi\frac{k}{L}\right|^{2s}}|u_k|^2 = |e^{\sigma A^s}u|^2 < \infty. \tag{D.15}$$

The norm in the space $D(\exp(\sigma A^s))$ is given by

$$|u|_{D(e^{\sigma A^s})} = |e^{\sigma A^s}u| \quad \text{for } u\in D(e^{\sigma A^s}). \tag{D.16}$$

The space $D(\exp(\sigma A^s))$ is actually a Hilbert space, and the associated inner product is given by

$$\langle u,v\rangle_{D(e^{\sigma A^s})} = \langle e^{\sigma A^s}u, e^{\sigma A^s}v\rangle \quad \text{for } u,v\in D(e^{\sigma A^s}). \tag{D.17}$$

In what follows, we will be mostly concerned with the case $s = 1/2$. Another Gevrey-type space that we will consider is $D(A^{1/2}\exp(\sigma A^{1/2}))$, which is also a Hilbert space; its inner product is given by

$$\langle u,v\rangle_{D(A^{1/2}\exp(\sigma A^{1/2}))} = \langle A^{1/2}e^{\sigma A^{1/2}}u, A^{1/2}e^{\sigma A^{1/2}}v\rangle \tag{D.18}$$

for $u,v\in D(A^{1/2}\exp(\sigma A^{1/2}))$; the associated norm is given by

$$|u|^2_{D(A^{1/2}e^{\sigma A^{1/2}})} = |A^{1/2}e^{\sigma A^{1/2}}u|^2 = 2\pi\sum_{k\in\mathbb{Z}^d}\left|\frac{k}{L}\right|^2 e^{4\pi\sigma\left|\frac{k}{L}\right|}|u_k|^2 \tag{D.19}$$

for $u\in D(A^{1/2}\exp(\sigma A^{1/2}))$.

The following inequality is satisfied by the bilinear term $\mathcal{B}(u,v)$ for u,v and w in $D(A^{1/2}\exp(\sigma A^{1/2}))$ with $\sigma>0$:

$$\begin{aligned}&|\langle e^{\sigma A^{1/2}}\mathcal{B}(u,v), e^{\sigma A^{1/2}}Aw\rangle|\\ \le{}& c_2|A^{1/2}e^{\sigma A^{1/2}}u||A^{1/2}e^{\sigma A^{1/2}}v||A^{1/2}e^{\sigma A^{1/2}}w|\left(1+\log\frac{|Ae^{\sigma A^{1/2}}u|^2}{\lambda_1|A^{1/2}e^{\sigma A^{1/2}}u|^2}\right)^{1/2}\end{aligned} \tag{D.20}$$

where c_2 depends only on the shape of the domain $\mathbb{T}^2$; and this inequality implies that the bilinear term $\mathcal{B}(u,v)$ belongs to $D(\exp(\sigma A^{1/2}))$.

Because we want to establish the analyticity in time of the solutions as functions with values in Gevrey space, we must assume that the forcing term f itself belongs to a Gevrey space. Hence, we assume that

$$f \in D(e^{\sigma_1 A^{1/2}})$$

for some $\sigma_1 > 0$. The NSE can then be written for complex times $\xi \in \mathbb{C}$ as

$$\frac{du}{d\xi} + \nu A u + \mathcal{B}(u,u) = f, \tag{D.21}$$

where $u = u(\xi)$.

In the 2-dimensional case, owing to the uniform bound on the enstrophy of the strong solutions, the domain of analyticity can be extended to a neighbourhood of the whole positive real axis. We fix $\theta \in [-\pi/4, \pi/4]$, $0 \le s \le T_0(\|u_0\|)$, where

$$\begin{aligned} T_0(\|u_0\|) &= T_0(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2) \\ &= \left[c_8 \nu \lambda_1 \left(1 + \frac{|f|_{\sigma_1}}{\nu^2 \lambda_1} + \frac{\|u_0\|^2}{\nu^2 \lambda_1} \right) \log c_9 \left(1 + \frac{|f|_{\sigma_1}}{\nu^2 \lambda_1} + \frac{\|u_0\|^2}{\nu^2 \lambda_1} \right) \right]^{-1} \end{aligned} \tag{D.22}$$

(c_8 and c_9 are constants depending only on the shape of the domain $\mathbb{T}^2$), and consider the time $\xi = s e^{i\theta}$ for $s > 0$; then the following estimate holds:

$$\|u(se^{i\theta})\|^2_{\varphi(s\cos\theta)} \le c_7 \lambda_1 \nu^2 + 2\|u_0\|^2, \tag{D.23}$$

where c_7 depends only on the shape of the domain $\mathbb{T}^2$. The function φ is chosen to be

$$\varphi(\xi) = \min(\nu \lambda_1^{1/2} \xi, \sigma_1)$$

for $\xi \ge 0$. Then we define the region

$$\begin{aligned} \Delta^0_{\sigma_1}(\|u_0\|) &= \Delta^0_{\sigma_1}(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2) \\ &= \left\{ \xi = s e^{i\theta} : |\theta| < \frac{\pi}{4}, 0 < s < T_0(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2), \nu \lambda_1^{1/2} s |\sin\theta| < \sigma_1 \right\} \end{aligned} \tag{D.24}$$

This set is a domain of analyticity of the solution $u = u(\xi)$ of the complex Navier-Stokes equations. The origin $\xi = 0$ belongs to the closure of $\Delta^0_{\sigma_1}(\|u_0\|)$. Moreover, on the closure of this domain we have

$$|u(\xi)|^2_{D(A^{1/2}e^{\varphi(s\cos\theta)A^{1/2}})} \le c_7\lambda_1\nu^2 + 2\|u_0\|^2 \quad \text{for } \xi \in \overline{\Delta^0_{\sigma_1}}(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2). \quad \text{(D.25)}$$

In the 2-dimensional case, the strong solutions exist for all positive time and their enstrophy is uniformly bounded. Hence, the domain of analyticity of the solutions can be extended to a neighbourhood of the positive real axis. Indeed, we know that for each $t \ge 0$,

$$\|u(t)\|^2 \le \|u_0\|^2 + \frac{1}{\nu^2\lambda_1}|f|^2. \quad \text{(D.26)}$$

Then at time $t_0 \ge 0$, we obtain the analyticity of the solution on the domain

$$t_0 + \Delta^0_{\sigma_1}((\|u_0\|^2 + |f|^2/(\nu^2\lambda_1))^{1/2}) \subset t_0 + \Delta^0_{\sigma_1}(\|u(t_0)\|).$$

By taking the union for all $t_0 > 0$ of the domains in the LHS of this expression, we obtain the analyticity in an open, pencil-like domain

$$\Delta^+_{\sigma_1}(\|u_0\|) = \bigcup_{t_0>0} \left\{t_0 + \Delta^0_{\sigma_1}((\|u_0\|^2 + |f|^2/(\nu^2\lambda_1))^{1/2})\right\}; \quad \text{(D.27)}$$

this is a neighbourhood of the positive real axis and has $\xi = 0$ on its boundary. Moreover, our estimates extend to all of $\Delta^+_{\sigma_1}(|u_0|)$ in the sense that

$$|u(\xi)|^2_{D(A^{1/2}e^{\varphi(s\cos\theta)A^{1/2}})} \le c_7\lambda_1\nu^2 + 2\|u_0\|^2 + |f|^2 \quad \text{(D.28)}$$

for $\xi = se^{i\theta} \in \overline{\Delta^+_{\sigma_1}}(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2)$.

From (D.22), (D.24), and (D.26), we can write the domain of analyticity as

$$\Delta^+_{\sigma_1}(\|u_0\|) = \Delta^+_{\sigma_1}(\|u_0\|, |f|_{\sigma_1}, \nu, \mathbb{T}^2) = \{\xi \in \mathbb{C}; |\text{Im } \xi| \le \min\{\text{Re } \xi, \delta_0\}\}, \quad \text{(D.29)}$$

where δ_0 is the largest width of the pencil-like domain $\Delta^+_{\sigma_1}$, estimated by

$$\delta_0 \ge \min\left\{\frac{\sigma_1}{\nu\lambda_1^{1/2}}, \left[c_{10}\nu\lambda_1\left(1 + \frac{|f|^2_{\sigma_1}}{\nu^4\lambda_1^2}\right)\log\left(c_{11}\left(1 + \frac{|f|^2_{\sigma_1}}{\nu^4\lambda_1^2}\right)\right)\right]^{-1}\right\}, \quad \text{(D.30)}$$

where c_{10} and c_{11} depend only on the shape of the domain $\mathbb{T}^2$.

An immediate consequence of the space analyticity of NSE solution in the 2-dimensional periodic case just derived is the exponential decrease of the Fourier coefficients of each solution with respect to the wave number. We have proven that for a forcing term f in the Gevrey space $D(\exp(\sigma_1 A^{1/2}))$ with $\sigma_1 > 0$, and for an initial velocity filed u_0 in H, the corresponding flow $u = u(t)$ is analytic in both space and time. Moreover, after some short transient time when the radius of analyticity of the solution $u(t)$ increases, we find $u(t)$ in the Gevrey space $D(\exp(\delta_0 A^{1/2}))$ with δ_0 as in (D.30). According to (D.28), the norm of $u(t)$ in this space is bounded uniformly in time:

$$|u(t)|^2_{D(A^{1/2}e^{\sigma A^{1/2}})} \leq c_7\lambda_1\nu^2 + 2\|u_0\|^2 + \frac{2}{\nu^2\lambda_1}|f|^2 \qquad \text{for } t \geq \delta_0. \tag{D.31}$$

From the Fourier series characterisation (D.19) of the space $D(\exp(\delta_0 A^{1/2}))$, we obtain

$$|u(t)|^2_{D(A^{1/2}e^{\sigma A^{1/2}})} = 2\pi \sum_{k\in\mathbb{Z}^d} \left|\frac{k}{L}\right| e^{4\pi\delta_0\left|\frac{k}{L}\right|}|u_k(t)|^2 \leq M^2, \tag{D.32}$$

where M^2 is the bound on the RHS of (D.31). Therefore, it is straightforward to deduce the following (crude) bound:

$$|u_k|^2 \leq \frac{M}{\sqrt{2\pi L}} \left|\frac{k}{L}\right| e^{-2\pi\delta_0\left|\frac{k}{L}\right|}. \tag{D.33}$$

Bibliography

[1] H. Akashi and H. Kumamoto, "Random sampling approach to state estimation in switching environments," *Automatica*, vol. 13, pp. 429-434, 1977.

[2] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50(2), pp. 174-188, Feb. 2002.

[3] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering*, Stochastic Modelling and Applied Probability, vol. 60, Springer, 2008.

[4] A. Beskos, D. Crisan, A. Jasra, and N. Whiteley, "Error bounds and normalizing constants for sequential monte carlo in high dimensions," *arXiv:1112.1544*, 2011.

[5] S.C. Brenner and L. Ridgway Scott, *The Mathematical Theory of Finite Element Methods*, Third ed., Texts in Applied Mathematics, vol. 15, Springer, 2008.

[6] C. Brett, A. Lam, K.J.H. Law, D. McCormick, M. Scott and A.M. Stuart, "Accuracy and stability of filters for the Navier-Stokes equation," submitted, *eprint, arXiv:1110.2527*, Oct. 2011.

[7] D. Brigo, B. Hanzon, and F. Le Gland, "A differential geometric approach to nonlinear filtering: the projection filter," *IEEE Trans. Automatic Control*, vol. 43(2), pp. 247-252, Feb. 1998.

[8] D.C. Brody, L.P. Hughston, and A. Macrina, "Information based asset pricing," *arXiv:0704.1976*, Apr. 2007.

[9] Z. Cai, F. Le Gland, and H. Zhang, "An adaptive local grid refinement method for nonlinear filtering," *Technical Report 2679, INRIA*, 1995.

[10] A. Carmi, F. Septier, and S.J. Godsill, "The Gaussian mixture MCMC particle algorithm for dynamic cluster tracking," in *Proc. The 12th International conference on information fusion*, Seattle, WA, USA, July 6-9 2009, pp. 1179-1186.

[11] J Carpenter, P. Clifford, P. Fearnhead, "Improved particle filter for nonlinear problems," *IEE Proceedings on Radar, Sonar and Navigation*, Vol. 146(1), pp. 2-7, Feb. 1999.

[12] H. Carvalho, P. Del Moral, A. Monin and G. Salut, "Optimal nonlinear filtering in gps/ins integration," *IEEE Trans. Aerospace and Electronic Systems,* vol. 33(3), pp. 835-850, Jul. 1997.

[13] D. Crisan, "Particle filters – a theoretical perspective," *Sequential Monte Carlo Methods in Practice, Stat. Eng. Inf. Sci.,* pp. 17-41, Springer, New York, 2001.

[14] D. Crisan, "Numerical methods for solving the stochastic filtering problem," *Numerical Methods and Stochastics (Toronto, ON, 1999),* vol. 34 of Fields Inst. Commun., pp. 1-20, American Mathematical Society, Providence, RI, 2002.

[15] D. Crisan, "Exact rates of convergence for a branching particle approximation to the solution of Zakai equation," *The Annals of Probability,* vol. 31(2), pp. 693-718, 2003.

[16] D. Crisan, "Superprocesses in a Brownian environment," *Proc. R. Soc. Lond. Ser. A. Math. Phys. Eng. Sci.,* vol. 460(2041), pp. 243-270, 2004.

[17] D. Crisan, P. Del Moral and T. Lyons, "Interacting particle systems approximations of the Kushner-Stratonovich equation," *Adv. in Appl. Probab.,* vol. 31(3), pp. 819-838, 1999.

[18] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners," *IEEE Trans. Signal Processing,* vol. 50(3), pp. 736-746, 2002.

[19] D. Crisan, J. Gaines and T. Lyons, "Convergence of a branching particle method to the solution of the Zakai equation," *SIAM J. Appl. Math.,* vol. 58(5), pp. 1568-1590 (electronic), 1998.

[20] D. Crisan and K. Li, "Generalised particle filters with Gaussian measures," *Proceedings of 19th European Signal Processing Conference*, pp. 659-663, 2011.

[21] D. Crisan and T. Lyons, "Nonlinear filtering and measure-valued processes," *Probability Theory and Related Fields,* vol. 109(2), pp. 217-244, 1997.

[22] D. Crisan and T. Lyons, "A particle approximation of the solution of the Kushner-Stratonovich equation," *Probability Theory and Related Fileds,* vol. 115(4), pp. 549-578, 1999.

[23] D. Crisan and T. Lyons, "Minimal entropy approximations and optimal algorithms," *Monte Carlo Methods Appl.,* vol. 8(4), pp. 343-355, 2002.

[24] D. Crisan and J. Miguez, "Particle approximation of the filtering density for state-space Markov models in discrete time," *arXiv:1111.5866*, 2012.

[25] D. Crisan and S. Ortiz-Latorre, "A KLV particle filter," *submitted*, Jun., 2012.

[26] D. Crisan, and B. Rozovsky, editors, *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press, 2011.

[27] P. Del Moral, "Nonlinear filtering using random particles," *Theory of Probability and its Applications,* vol. 40(4), pp. 690-701, 1995.

[28] P. Del Moral, "Nonlinear filtering: Interacting particle resolution," *Comptes Rendus del'Acadmie des Sciences - Series I - Mathematics,* vol. 325(6), pp. 653-658, 1997.

[29] P. Del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications.* New York: Springer, 2004.

[30] P. Del Moral and A. Guionnet, "Central limit theorem for nonlinear filtering and interacting particle systems," *Ann. Appl. Probab.,* vol. 9(2), pp. 275-297, 1999.

[31] P. Del Moral and L. Miclo, "Branching an interacting particle systems approximations of Feynman-Kac formula with applications to nonlinear filtering," *Séminaire de Probabilités, XXXIV,* vol. 1729 of Lecture Notes in Math., pp. 1-145, Springer, Berlin, 2000.

[32] A. Doucet, J.F.G. de Freitas, and N.J. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001.

[33] Weinan E, J. C. Mattingly, and Ya. Sinai, "Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation," *Communications in Mathematical Physics*, vol. 224(1), pp. 83-106, 2001.

[34] Alexandre Ern and Jean-Luc Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, vol. 159, Springer, 2004.

[35] S. N. Ethier and T. G. Kurtz, *Markov Processes,* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[36] M. Flament, G. Fleury and M.E. Davoust, "Particle filter and Gaussian-mixture filter efficiency evaluation for terrain-aided navigation," in *Proc. EUSIPCO 2004*, Vienna, Austria, September 6-10, 2004, pp. 605-608.

[37] C. Foias, O. Manley, R. Rosa and R. Temam, *Navier-Stokes equations and turbulence*, Cambridge University Press, 2001.

[38] N. Gordon, D. Salmond and C. Ewing, "Bayesian state estimation for tracking and guidance using the bootstrap filter," *Journal of Guidance, Control, and Dynamics,* vol. 18(6), pp. 1434-1443, 1995.

[39] N. Gordon, D. Salmond and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEEE Proc. Radar and Signal Processing,* vol. 140(2), pp. 107-113, Apr. 1993.

[40] J.E. Handschin and D.Q. Mayne, "Monte Carlo techniques to estimate the conditional expectation in multi-stage non-linear fltering," *International Journal of Control*, vol. 9(5), pp. 547-559, 1969.

[41] M. Iglesias, K. Law and A.M. Stuart, "The ensemble Kalman filter for inverse problems," *arXiv:1209.2736*, Sep. 2012.

[42] M. Iglesias, K. Law and A.M. Stuart, "Evaluating Gaussian approximations to Bayesian inversion for subsurface geophysical applications," in preparation.

[43] G. Kallianpur, *Stochastic Filtering Theory*, New York, Springer, 1980.

[44] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verleg, N.Y., 1991.

[45] G. Kitagawa, "Monte carlo filter and smoother for non-gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics,* vol. 5(1), pp. 1-25, 1996.

[46] J.H. Kotecha and P.M. Djurić, "Gaussian particle filtering," *IEEE Trans. Signal Processing*, vol. 51, pp. 2593-2602, Oct. 2003.

[47] J.H. Kotecha and P.M. Djurić, "Gaussian sum particle filtering," *IEEE Trans. Signal Processing*, vol. 51, pp. 2602-2612, Oct. 2003.

[48] T.G. Kurtz and P. Protter, "Weak limit theorems for stochastic integrals and stochastic differential equations," *Annals of Probability*, vol. 19(3), pp. 1035-1070, 1991.

[49] H.J. Kushner, "Approximations to optimal nonlinear flters," *IEEE Trans. Automatic Control,* vol. 12(5), pp. 546-556, Oct. 1967.

[50] K.J.H. Law and A.M. Stuart, "Evaluating data assimilation algorithms," to appear in *Monthly Weather Review*, *arXiv:1107.4118*, Mar. 2012.

[51] F. Le Gland, V. Monbet and V.D. Tran, "Large sample asymptotics for the ensamble Kalman filter," *Handbook on Nonlinear Filtering*, Oxford University Press, 2010.

[52] F. Le Gland and N. Oudjane, "Stability and uniform approximation of nonlinear filters using the Hilbert space metric and application to particle filters," *Annals of Applied Probability*, vol. 14(1), pp. 144-187, 2004.

[53] Wonjung Lee, "An adaptive approximation for the higher order statistics of the posterior," *Data Assimilation Conference*, Oxford, UK, Sep. 2012.

[54] S. Lototsky, R. Mikulevicius, and B. Rozovsky, "Nonlinear filtering revisited: A spectral approach," *SIAM J. Control Optim.*, vol. 35(2), pp. 435-461, 1997.

[55] V.M. Lucic and A.J. Heunis, "On uniqueness of solutions for the stochastic differential equations of nonlinear filtering," *Annals of Applied Probability*, vol. 11(1), pp. 182-209, 2001.

[56] O. Obanubi, "Particle filters with random resampling times," *PhD Thesis, Imperial College London*, Dec. 2010.

[57] M. Ondreját, "Uniqueness for stochastic evolution equations in Banach spaces," *Dissertationes Math. (Rozprawy Mat.)*, 2004.

[58] U. Orguner and F. Gustafsson, "Risk sensitive particle filters for mitigating sample impoverishment," *IEEE Trans. Signal Processing*, vol. 56(10), pp. 5001-5012, Oct. 2008.

[59] S. Reich, "A Gaussian mixture ensemble transform filter," submitted, *eprint arXiv:1102.3089*, Feb. 2011.

[60] J.C. Robinson, *Infinite-dimensional dynamic systems*, Cambridge Tex in Applied Mathematics, Cambridge University Press, 2001.

[61] M. Röckner, B. Schmuland, and X. Zhang, "Yamada-Watanabe theorem for stochastic evolution equations in infinite dimensions," *Condensed Matter Physics*, vol. 2(54), pp. 247-259, 2008.

[62] S. Roelly-Coppoletta, "A criterion of convergence of measure-valued processes: application to measure branching processes," *Stochastics*, vol. 17(1-2), pp. 43-65, 1986.

[63] C. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales: Volume I Foundations*, second ed., Cambridge, UK, Cambridge University Press, 2000.

[64] R. L. Stratonovich, "On the theory of optimal non-linear filtering of random functions," *Theory of Probability and its Applications,* vol. 4, pp. 223-225, 1959.

[65] R. L. Stratonovich, "Application of the Markov processes theory to optimal filtering," *Radio Engineering and Electronic Physics,* vol. 5(11), pp. 1-19, 1960.

[66] R. L. Stratonovich, "Conditional Markov Processes," *Theory of Probability and its Applications,* vol. 5, pp. 156-178, 1960.

[67] J. Szpirglas, "Sur l'équivalence d'équations différentielles stochastiques à valeurs measures intervenant dans le filtrageMarkovien non linéaire" [French], *Ann. Inst. H. Poincaré Sect. B (N.S.)*, vol 14(1), pp. 33-59, 1978.

[68] S. Thrun, J. Langford, and V. Verma, "Risk sensitive particle filters," in *Advances in Neural Information Processing Systems 14*, MIT Press, 2002.

[69] R. Van der Merwe and E. Wan, "Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models," in *Proc. ICASSP 2003*, pp. 701-704.

[70] D.Williams, *Probability with Martingales*, Cambridge University Press, 1991.

[71] M. Zakai, "On the optimal filtering of diffusion process," *Zeit. Wahrsch.*, vol. 11, pp. 230-243, 1969.