The Problem of Nonlinear Filtering

Dan Ovidiu Crișan

Ph. D.
University of Edinburgh
1996

The Problem Of Nonlinear Filtering

by

Dan Ovidiu Crişan

Submitted to the Department of Mathematics
on 15 July 1996, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Stochastic filtering theory studies the problem of estimating an unobservable 'signal' process X given the information obtained by observing an associated process Y (a 'noisy' observation) within a certain time window $[0, t]$. It is possible to explicitly describe the distribution of X given Y in the setting of linear/gaussian systems. Outside the realm of the linear theory, it is known that only a few very exceptional examples have explicitly described posterior distributions. We present in detail a class of nonlinear filters (Beneš filters) which allow explicit formulae. Using the explicit expression of the Laplace transform of a functional of Brownian motion we give a direct computation of the unnormalized conditional density of the signal for the Beneš filter and obtain the formula for the normalized conditional density of X for two particular filters.

In the case in which the signal X is a diffusion process and Y is given by the equation $dY_t = h(s, X_s)ds + dW_t$, where W is a Brownian motion independent of X, $Y_0 = 0$ and h satisfies certain conditions, the evolution of the conditional distribution of X is described by two stochastic partial differential equations: a linear equation - the *Zakai* equation - which describes the evolution of an unnormalised version of the conditional distribution of X and a nonlinear equation - the $Kushner-Stratonovitch$ equation - which describes the evolution of the conditional distribution of X itself.

We construct several measure valued processes, associated with the two equations, whose values give the conditional distribution of X (in the first case unnormalised). We do this by means of converging sequences of branching particle systems. The particles evolve independently, moving with the same law as X, and branch according to a mechanism that depends on their locations and the observation Y. The result is a cloud of paths, with those surviving to the current time providing an estimate for the conditional distribution of X.

The construction of these measure valued processes is new, since it involves wildly varying branching mechanisms. But their true value stems from the fact that we can successfully use them to solve numerically the problem of nonlinear filtering. We prove the validity of the algorithm and show the numerical computation for several examples.

Thesis Supervisor: Terry J. Lyons
Title: Professor

Thesis Supervisor: Alison Etheridge
Title: Dr.

Contents

List of Figures

Acknowledgements

I would like to express my deepest thanks to Terry Lyons, for everything he has done to help me throughout my Ph.D. and beyond. He has been a good friend and mentor.

I would also like to thank Alison Etheridge for the many illuminating discussions we had, for (proof)reading the manuscript and offering extensive comments and corrections.

I acknowledge the support of the whole staff of the Department of Mathematics, University of Edinburgh, particularly of Allan Sinclair, David Parker and Alan Forest. I have been deeply impressed by the hospitality at Edinburgh.

I thank Jessica Gaines for giving me so much of her time, helping me with C programming, wordprocessing and many others. Her numerical simulations have stimulated a lot this work.

I am also indebted to my professors in Romania, especially to Ioan Cuculescu, Gabriel Pripoae, Constantin Caragea and Hanzu Doina.

I would like to thank my colleagues Mehmet Firat, Lorenzo Bertini and David Williams for many enjoyable hours spent discussing mathematics and others.

There are no words good enough to thank my wife for her constant support and understanding, for patiently listening to my babbling, for sharing my moments of joy and weeks of disappointments.

Finally, I would like to thank my parents, Maria and Petru whose sole purpose in life was to raise me and my brother, and I hope this thesis is a sign that their efforts have not been wasted.

July 1996

Declaration

This thesis has been composed by myself and has not been submitted for any other degree or professional qualification. The work is my own, except where I have indicated to the contrary within the thesis.

DEDICATED TO MY PARENTS

Chapter 1

Introduction

1.1 Background

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space endowed with the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and X be a process defined on this space with values in E, where E is a locally compact complete separable metric space (in particular, it is a locally compact Polish space). Let $\mathcal{B}(E)$ be the space of bounded, Borel measurable functions on E and $\mathcal{P}(E)$ is the set of probability measures on E. We assume that $\{X_t, \mathcal{F}_t; t \geq 0\}$ is a diffusion process (called the 'signal' process), which solves the martingale problem associated with the infinitesimal generator $A : \mathcal{D}(A) \subset \mathcal{B}(E) \rightarrow \mathcal{B}(E)$ and the initial distribution $\pi_0 \in \mathcal{P}(E)$. Let also $h : [0, \infty) \times E \rightarrow \mathbb{R}^m$ be a continuous Borel measurable function and $W = \{W_t, \mathcal{F}_t; t \geq 0\}$ be an m-dimensional standard Brownian Motion independent of X. We define Y to be the following stochastic process

$$Y_t = \int_0^t h(s, X_s) ds + W_t, \qquad t \geq 0$$

with the initial condition $Y_0 = 0$. The process Y is usually called the 'observation' process. We denote $\mathcal{Y}_t = \sigma(Y_s, 0 \leq s \leq t)$.

The filtering problem (within the time frame $[0,T]$) consists in determining the conditional law of the signal given the observation process, i.e., in computing

$$\pi_t(\varphi) \overset{\text{def}}{=} E[\varphi(X_t) | \mathcal{Y}_t], \quad \forall t \in [0,T], \ \varphi \in \mathcal{B}(E).$$

We observe here that π_0 - the initial distribution of X - is identical with the conditional distribution of X_0, given $\mathcal{Y}_0$, and that is why we use the same notation for both. To solve the problem, first one changes the underlying measure so that Y_t becomes a Brownian motion. By assuming that

$$E\left[\int_0^T |h(s,X_s)|^2 ds\right] < \infty$$

and using the definition of Y , the formula

$$\left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_T} = \exp\left(-\int_0^T h^*(s,X_s)dW_s - \frac{1}{2}\int_0^T |h(s,X_s)|^2 ds\right)$$

defines a new probability measure $\tilde{P}$ absolutely continuous with respect to P and with respect to which Y is a Brownian motion. The Kallianpur-Striebel formula tells us that

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(1)}, \quad \mathrm{P-a.s.}, \tag{1.1}$$

where

$$\rho_t(\varphi) \overset{\text{def}}{=} \tilde{E}\left[\varphi(X_t)\exp\left(\int_0^t h^*(s,X_s)dY_s - \frac{1}{2}\int_0^t |h(s,X_s)|^2 ds\right)|\mathcal{Y}_t\right]$$

and $\tilde{E}$ is the expectation with respect to $\tilde{P}$. ρ_t is usually called the *unnormalised conditional distribution* on X (one can see why from (1.1)). Under certain conditions, to be described in detail in the next chapter, one proves that ρ_t uniquely satisfies the following evolution equation, called the *Zakai* equation

$$\rho_t(\varphi) = \pi_0(\varphi) + \int_0^t \rho_s(A\varphi)ds + \int_0^t \rho_s(h^*\varphi)dY_s, \quad a.s.\ \forall t, \tag{1.2}$$

where φ is in the domain of the infinitesimal generator A. From (1.1) and (1.2) one obtains that $\pi_t(\varphi)$ satisfies the following evolution equation, called the *Kushner-Stratonovitch* equation

$$\pi_t(\varphi) = \pi_0(\varphi) + \int_0^t \pi_s(A\varphi)ds + \int_0^t (\pi_s(h^*\varphi) - \pi_s(h^*)\pi_s(\varphi))(dY_s - \pi_s(h)ds),$$

where φ is, again, in the domain of the infinitesimal generator A.

As we have set out above, the essential problem of stochastic filtering is to find the conditional distribution of X_t given the information obtained by measuring Y_s for s in the time window $[0, t]$. The problem has considerable importance, but its usefulness is limited to those cases *where numerical solution is feasible.* In the special case where the evolution of X_t is given by a linear stochastic differential equation, h is also linear and X_0 has a Gaussian distribution one has the very nice property that the conditional distribution of X_t is always Gaussian, and in consequence can be described by a finite number of parameters (its mean and covariance). This remark has enormous computational significance: the conditional distribution can be obtained by solving an ordinary differential equation for the covariance and a stochastic differential equation for the mean. This approach is the well known Kalman filter ([23], [24]). One can say that the linear case is completely understood. However, there are many situations, where the linear/Gaussian assumptions of this model are inappropriate; that is why we concentrate on *nonlinear* filtering.

In the nonlinear case it would seem attractive to apply the Zakai equation which gives a linear stochastic PDE for the measure (or its density) describing the (unnormalised) conditional distribution of X_t. Unfortunately, in real applications X_t is often a multidimensional variable, even in four dimensions it can be a serious problem to solve a PDE and more difficult to accurately solve an SPDE, in fifty dimensions it is utterly hopeless. This has lead to attempts to find wider classes of models where the conditional distribution lies in a finite dimensional manifold (the Benes and Ocone filters, see Chapter 3), but these represent a very small class. More practical have been the approaches were linearisation can be applied recursively using the so called *extended* Kalman filter ([36]). But approaches via linearisation have strong limitations if there is significant uncertainty in the observations. It has remained a serious problem to find good ways to approximate π_t in the general case. It is this problem that we try to address in this thesis.

1.2 The main results

In high dimensions, one of the most convenient ways to describe a measure is to generate a sample from it; in other words a sequence of points chosen at random according to its distribution. This fact has been realised by Statisticians for many years and explains the popularity of Gibbs Sampling (cf [18], [19], [20]). The reason is that one is often interested in some low dimensional marginal distribution and not the measure itself. Obtaining this directly from a density function in high dimensions is not computationally feasible, as it involves a numerical integration over the whole space. On the other hand, the projection of a sample can quickly be computed, and non-parametric approaches can be used effectively to construct approximate marginal distributions. Our idea is that it might be possible to approach the Zakai equation and the Kushner-Stratonovitch equation by creating a *sample* from the posterior measure. We do not quite succeed, but we are able to produce arbitrarily good approximations.

Let $M_F(\mathbb{R}^d)$ be the space of finite measures over $\mathbb{R}^d$ and $\mathcal{X}$ (calligraphic X) be a measure-valued branching process defined on $(\Omega, \mathcal{F}, \tilde{P})$, with the property that, for every φ in the domain of A, the process

$$M^{\varphi}(t) \overset{\text{def}}{=} (\mathcal{X}(t), \varphi) - (\mathcal{X}(0), \varphi) - \int_0^t (\mathcal{X}(s), A\varphi) ds - \int_0^t (\mathcal{X}(s), h^*(s)\varphi) dY_s \tag{1.3}$$

is a square integrable martingale with respect to the filtration $\mathcal{F}_t \vee \mathcal{Y}$ with quadratic variation

$$< M^{\varphi}(t) >= \int_0^t (\mathcal{X}_s, v_s \varphi^2) ds, \qquad P - a.s. \tag{1.4}$$

where v_s is any given bounded, positive function, continuous in time for which $v_t \geq \frac{1}{4}$, $\forall t \in [0, 1]$. We make an abuse of notation here since, in fact, the process $\mathcal{X}$ will be constructed on $(\Omega', \mathcal{F}', \tilde{P}')$ a probability space larger than the initial one with $\Omega \subset \Omega'$, $\mathcal{F}' |_{\Omega} = \mathcal{F}$, and $\tilde{P}' |_{\mathcal{F}} = \tilde{P}$. We construct $\mathcal{X}$ using a sequence of branching particle systems $\mathcal{X}_n$ with wildly varying (space and time dependent) branching mechanisms. The particles constituting the systems $\mathcal{X}_n$ will each have the same weight, will move independently and according to the same law as the signal and will branch with a mean depending on their recent past and the observation Y and variance v_s if the branching occurs at time s. This sequence is convergent to $\mathcal{X}$. One shows that $\widetilde{E}\left[(\mathcal{X}(t),\varphi)\,|\mathcal{Y}_t\right] = \rho_t(\varphi)$ and

$$\lim_{n\to\infty} \widetilde{E}\left[(\mathcal{X}_n(t),\varphi)\,|\mathcal{Y}_t\right] = \widetilde{E}\left[(\mathcal{X}(t),\varphi)\,|\mathcal{Y}_t\right]$$

Therefore one can use the systems $\mathcal{X}_n$ in order to compute numerically ρ_t, hence the solution of the Zakai equation. The processes $\mathcal{X}_n$ are easy to simulate and, since the particles involved move independently, one can use the powerful tools of parallel computing to do this. This way, the dimension of the system is not so important since the complexity of the numerical algorithms depends only polynomially on the dimension. If we take independent copies of $\mathcal{X}_n$ (while keeping the observation path $Y_\cdot$ fixed) and then integrate φ against the resulting measure, we can compute $\widetilde{E}\left[(\mathcal{X}_n(t),\varphi)\,|\mathcal{Y}_t\right]$, and hence approximate $\rho_t(\varphi)$.

By slightly varying the previous construction, we produce the measure-valued branching process U whose distribution at time t is exactly ρ_t. We do this, once again, by using a sequence of branching particle systems U_n convergent to U. This time, the evolution equation satisfied by the limiting measure-valued process will be

$$(U(t),\varphi) = (U(0),\varphi) + \int_0^t (U(s), A\varphi)ds + \int_0^t (U(s), h^*(s)\varphi)dY_s \tag{1.5}$$

The algorithm based on this result can be used, too, to solve numerically the Zakai equation and, hence, the filtering problem. This result is an improvement of the previous one, because it eliminates the extra degree of randomness introduced there. The system of particles will move according to the law of the signal, independently of each other and after fixed-length intervals will branch. The mean number of offspring of a particle will depend on the last part of its trajectory and on the observation process, but the variance of the branching mechanism will be the minimum possible one. We can use these particle systems to solve numerically the filtering problem since, as the number of particles is increased, the empirical measure associated to the cloud of particles converges to the solution of the Zakai equation. By starting with a number of particles constituting a sample of the initial distribution of X, we allow the system to evolve up to time t and then use the empirical law to estimate the required statistic φ. In the previous model, the variance of the branching mechanism was a-priori given and, in this way, we introduced an extra degree of randomness, so only the (conditional) expectation of that sequence converged to ρ_t. Therefore, we needed a whole set of copies of the particles system in order to obtain a good approximation to the solution of the Zakai equation. Since $\lim_{n\to\infty}(U_n(t),\varphi) = (U(t),\varphi) = \rho_t(\varphi)$, here we don't need to estimate an average.

One may ask at this point why we present both constructions in the thesis and not only one of them, possibly the second one since it seems to be better. First of all, we point out that the small variation in the construction creates quite a big one in the analysis of the processes, so one cannot produce a general construction which will contain as particular cases both the fixed variance case and the minimal variance case. Also, it shows the existence of a new class of measure valued branching processes. The construction of $\mathcal{X}$ was 'historically' the first one and, although U does a better job, we feel closer to our first 'child'.

The individual processes are easy to simulate (particularly on parallel machines) and the complexity of computing a single trajectory grows only polynomially with the dimension of the state space. Combining these remarks, it is therefore feasible to produce an approximate sample in any reasonable dimension and with increasing effort an arbitrarily accurate and arbitrarily large sample.

The last measure valued-process that we consider is the one that solves the other important equation in nonlinear filtering, the Kushner - Stratonovitch equation. We construct the probability measure-valued process α which satisfies the following evolution equation

$$\begin{aligned}(\alpha(t),\varphi) &= (\alpha(0),\varphi) + \int_0^t (\alpha(s),A\varphi)ds + \int_0^t ((\alpha(s),h^*\varphi) - (\alpha(s),h^*)(\alpha(s),\varphi)) \\ &\quad \times (dY_s - (\alpha(s),h)ds). \qquad (1.6)\end{aligned}$$

We succeed in doing this by using a sequence α_n of branching particle systems, which converges to α and the limit process has the same law as the conditional law of the signal, i.e.,

$$\lim_{n\to\infty}(\alpha_n(t),\varphi) = (\alpha(t),\varphi) = \pi_t(\varphi).$$

This way we manage to produce a sample approximation of the conditional distribution of X and therefore, based on this, we can construct a numerical algorithm for solving directly the Kushner - Stratonovitch equation and, implicitly, the nonlinear filtering problem. This way we successfully solve the nonlinear SPDE (1.6) in a direct fashion and not via its linearised form - the Zakai equation. Unfortunately, this time the processes α_n will no longer consist of independent particles. The mean number of offsprings of these particles are correlated: at each branching time we compute the branching means in a similar way to those for U_n, but then we normalize them so as to keep their sum a martingale. In this case the limiting process α will be a random *probability* measure. This makes the use of parallel computing for numerical simulations less effective than before, but eliminates the error resulting from normalising at the end only. Also, this way we keep the mass under control, hence preventing it from exploding or vanishing.

If we contrast this approach with the one where particles are weighted with exponentials (the classical Monte - Carlo method), we see two important advantages over this (largely disastrous) method. Firstly, all computations are associated with particles that carry the same weight - one never finds oneself computing a trajectory that will obviously have a smaller weight than another. Secondly, the paths exploring unfruitful directions of exploration are rapidly killed. This again suggests a sifting out of potentially unhelpful computation. The algorithm is feasible in the sense that one can carry it out and get a return directly related to the amount of computational effort invested. However, it has to be said that the convergence could still be quite slow. We have no definite result on the rates of convergence at this point in time.

We believe that this is novel and potentially important work. Particle systems have been used before to solve PDE's (cf. [5], [7], [8], [40]) - the novelty here is that (i) we use them to construct a measure and this is more useful that the density and (ii) we are looking at technically more complicated situations than previous work. The algorithm developed by us is different from the genetic algorithms ([11], [13], [21]). One difference is that the breeding pattern of our algorithm is influenced by a noisy external process which affects the analysis significantly; the algorithm is appropriate for solving SPDE's. This is an attempt to consolidate our understanding of the bridge between two important parts of stochastic analysis: measure valued processes and parabolic SPDE's, particularly to focus on those aspects of the theory that are relevant to computational efficiency in the solution of SPDE's of the type that arise in non-linear filtering. This opens up possibilities for a wider study of various features (existence, support, blow-up times, etc.) for different SPDE's using the tools that come from the theory of measure valued processes. The Zakai equation and the Kushner - Stratonovitch equation are only two of the many SPDE's which do not have reliable numerical algorithms for solving them. We therefore hope that our results will be extended to other equations of this type.

Acknowledgement of colaboration

I worked jointly with my supervisor, Terry Lyons, in developing the results of this thesis. However, the detail, presentation and proofs given here are my own.

1.3 Summary

We start Chapter 2 by setting up the notation and assumptions for defining rigorously the stochastic filtering problem. We also present briefly the linear/Gaussian filter - the *Kalman-Bucy* filter. We then go on to establish several preparatory results used in proving that the conditional distribution of X satisfies the *Kushner - Stratonovitch* equation and its unnormalised form satisfies the *Zakai* equation. We look also at the uniqueness of the solution of the two equations.

Chapter 3 deals with one of the few nonlinear filters which admits explicit solutions - the *Beneš* filter. After the description of the filter, we explicitly compute an exponential functional of Brownian motion. The explicit formulae help us to solve the *Beneš* filter in a different and, we think, a more direct way. We then observe the long time behaviour for two particular filters.

Chapter 4 contains results on convergence in distribution for processes with cádlág paths with values in the space of measures over a locally compact metric space.

The main results of the thesis are contained in Chapter 5 and Chapter 6. Here we put together the knowledge of nonlinear filtering and of measure valued branching processes and construct three measure valued processes - $\mathcal{X}$, U, α - which in turn help us solve numerically the nonlinear filtering problem. For each of them we use a sequence of approximating particle systems, prove that the sequence is tight and show that it converges to the desired limit.

Using the particle systems described in Chapter 5 and 6 one can construct numerical algorithms to solve the filtering problem. Chapter 7 contains some numerical applications based on these algorithms. The Appendix contains a few results we use in the proofs.

Chapter 2 and Chapter 4 contain well known results, but some of the proofs might be new. Chapter 3, Chapter 5 and Chapter 6 contain original work.

Chapter 2

The Nonlinear Filtering Problem

2.1 Notation and assumptions

Let (E, d) be a locally compact complete separable metric space (in particular, it is a locally compact Polish space) and Ω^1 be the space of E-valued continuous functions defined on $[0, \infty)$, endowed with the topology of uniform convergence on compact intervals; let the associated Borel σ-field be denoted by $\mathcal{F}^1$, i.e.,

$$\Omega^1 = C([0, \infty), E), \qquad \mathcal{F}^1 = \mathcal{B}(\Omega^1). \tag{2.1}$$

Let X be the canonical E-valued process defined on this space, $X_t(\omega^1) = \omega^1(t)$, $\omega^1 \in \Omega^1$. We observe that X_t is measurable with respect to the σ-field $\mathcal{F}^1$ and consider the filtration associated with the process X

$$\mathcal{F}_t^1 = \sigma(X_s; s \in [0, t]). \tag{2.2}$$

Let $C_b(E)$ be the space of bounded continuous functions on E endowed with the supremum norm $\|\cdot\|$ and $\mathcal{P}(E)$ be the set of probability measures on E. Let also $A : C_b(E) \to C_b(E)$ be a markovian infinitesimal generator with domain $\mathcal{D}(A)$ with $1 \in \mathcal{D}(A)$ and $A1 = 0$, and P^1 be a probability measure solving the martingale problem associated with A and initial distribution $\pi_0 \in \mathcal{P}(E)$, i.e., under P^1, the distribution of X_0 is π_0 and

$$M_t^{\varphi} = \varphi(X_t) - \varphi(X_0) - \int_0^t A\varphi(X_s)ds, \quad \mathcal{F}_t^1, \quad 0 \le t < \infty, \tag{2.3}$$

is a martingale for any $\varphi \in \mathcal{D}(A)$. Let also Ω^2 be defined similarly to Ω^1, but with $E = \mathbb{R}^m$. Hence

$$\Omega^2 = C([0,\infty),\mathbb{R}^m), \qquad \mathcal{F}^2 = \mathcal{B}(\Omega^2). \tag{2.4}$$

We denote by V the canonical process in Ω^2, i.e., $Y_t(\omega^2) = \omega^2(t)$, $\omega^2 \in \Omega^2$ and take P^2 to be a probability measure under which V is an m-dimensional standard Brownian motion on $(\Omega^2, \mathcal{F}^2)$. Consider now the following:

$$\begin{aligned}
\Omega &\stackrel{\text{def}}{=} \Omega^1 \times \Omega^2,\\
\mathcal{F}' &\stackrel{\text{def}}{=} \mathcal{F}^1 \bigotimes \mathcal{F}^2,\\
P &\stackrel{\text{def}}{=} P^1 \bigotimes P^2,\\
\mathcal{F} &\stackrel{\text{def}}{=} \mathcal{F}' \vee \mathcal{N},\\
\mathcal{N} &\stackrel{\text{def}}{=} \{A \in \mathcal{F}'; P(A) = 0\}.
\end{aligned}$$

$(\Omega, \mathcal{F}, P)$ is a complete probability space and, under P, X and V are two independent processes. They can be viewed as processes on the product space $(\Omega, \mathcal{F}, P)$ in the usual way: If W is the canonical process on Ω, then

$$\begin{aligned}
W(t) &= \omega(t) = (\omega^1(t), \omega^2(t)),\\
X &= p^1(\omega), \text{ where } p^1 : \Omega \to \Omega^1,\ p^1(\omega) = \omega^1,\\
V &= p^2(\omega), \text{ where } p^2 : \Omega \to \Omega^2,\ p^2(\omega) = \omega^2.
\end{aligned}$$

M_t^φ is also a martingale with respect to the larger filtration $\mathcal{F}_t$, where

$$\mathcal{F}_t = \sigma(X_s, V_s; s \in [0,t]) \vee \mathcal{N}. \tag{2.5}$$

Let now $h : [0,\infty) \times E \to \mathbb{R}^m$ be a continuous Borel measurable function with the property that

$$E\left[\int_0^T |h(s,X_s)|^2\, ds\right] < \infty \qquad \forall T > 0, \tag{2.6}$$

and take Y to be the following $\mathcal{F}_t$-adapted process (usually called the 'observation' process)

$$Y_t = \int_0^t h(s,X_s)ds + V_t, \qquad t \geq 0. \tag{2.7}$$

If we also introduce the following filtration

$$\mathcal{Y}_t = \sigma(Y_s, s \in [0,t]) \vee \mathcal{N} \subset \mathcal{F}_t, \tag{2.8}$$

$$\mathcal{Y} = \bigcup_{t \in \mathbb{R}_+} \mathcal{Y}_t, \tag{2.9}$$

then the filtering problem (within the time frame $[0,T]$) consists in determining the conditional law of the signal given the observation process, i.e., in computing

$$\pi_t(\varphi) \overset{\text{def}}{=} E[\varphi(X_t)|\mathcal{Y}_t], \quad \forall t \in [0,T],\ \varphi \in \mathcal{B}(E). \tag{2.10}$$

We observe here that π_0 - the initial distribution of X - is identical with the conditional distribution of X_0, given $\mathcal{Y}_0$, and that is why we used the same notation for both

$$\pi_0(\varphi) = \int \varphi(x) d(P \circ X_0^{-1}).$$

To summarise, on a complete probability space $(\Omega, \mathcal{F}, P)$ we have defined a pair $\{(X_t, Y_t), \mathcal{F}_t; t \geq 0\}$ of processes as follows:

1. X is the solution of the martingale problem associated with the infinitesimal generator A and initial condition π_0 (the signal).

2. $Y_t = \int_0^t h(s, X_s) ds + V_t$, where V is a Brownian motion independent of X and h satisfies condition 2.6 (the observation).

Given these processes, we want to find out π_t the conditional law of the signal given the observation.

For the nonlinear case, one changes first the underlying measure, so that Y_t becomes a Brownian motion. Next we present the solution of the filtering problem for the linear case - the Kalman-Bucy filter.

2.2 The Kalman - Bucy Filter

Let A be the following second order differential operator

$$A(t)\varphi(x) = \sum_{i=1}^{d} (F_i(t)x + f_i(t)) \frac{\partial \varphi(x)}{\partial x_i} + \sum_{i,j=1}^{d} Q_{ij}(t) \frac{\partial^2 \varphi(x)}{\partial x_i \partial x_j},$$

where F_i, $f_i, Q_{ij} \in C([0,T])$ and $Q_{ij} \geq 0$. Under these conditions the solution X of the martingale problem associated with the infinitesimal operator A satisfies the following stochastic differential equation (see [38], Chapter II, Th. 2)

$$X_t = X_0 + \int_0^t (F_i(t)X_s + f_i(t))\, ds + \int_0^t Q(s) dB_s, \tag{2.11}$$

where $\{B_t, \mathcal{F}_t, t \in [0,T]\}$ is a d-dimensional standard Brownian motion. We also assume that the initial data π_0 is Gaussian with mean x_0 and covariance matrix P_0, i.e.,

$$\pi_0(\varphi) = \int_{\mathbb{R}^d} \varphi\left(x_0 + P_0^{\frac{1}{2}}\xi\right) \frac{\exp\left(-\frac{1}{2}|\xi|^2\right)}{(2n)^{\frac{n}{2}}} d\xi \text{ for all } \varphi \in B(\mathbb{R}^d). \tag{2.12}$$

Let the observation process Y_t be defined as

$$Y_t = \int_0^t (H(s)X_s + h(s))\, ds + V_t, \quad t \geq 0, \tag{2.13}$$

with H and h continuous in time. Under these assumptions we have the following

Theorem 1 (Kalman-Bucy) *The conditional distribution of the signal given the observation is Gaussian with mean $\hat{x}_t$ and covariance matrix P_t, where $\hat{x}_t$ is the solution of the SDE*

$$\begin{cases} d\hat{x}_t & = (F(t)\hat{x}_t + f(t))\, dt + P_t H^*(t) \left[dY_t - (H(t)\hat{x}_t + h(t))\, dt\right] \\ \hat{x}_0 & = x_0 \end{cases} \tag{2.14}$$

and P_t is the solution of the Ricatti equation

$$\frac{dP_t}{dt} = -P_t H^*(t) H(t) P_t + Q(t) + F(t)P_t + P_t F(t). \tag{2.15}$$

with the initial condition P_0.

Remark 1 *In other words, for all $\varphi \in B(\mathbb{R}^d)$, $t \geq 0$, we have*

$$\pi_t(\varphi) = \int_{\mathbb{R}^d} \varphi\left(\hat{x}_t + P_t^{\frac{1}{2}}\xi\right) \frac{\exp\left(-\frac{1}{2}|\xi|^2\right)}{(2n)^{\frac{n}{2}}} d\xi.$$

Theorem 2 *The unnormalised conditional distribution of the signal given the observation is given by*

$$\rho_t(\varphi) = \left[\int_{\mathbb{R}^d} \varphi\left(\hat{x}_t + P_t^{\frac{1}{2}} \xi\right) \frac{\exp\left(-\frac{1}{2}|\xi|^2\right)}{(2n)^{\frac{n}{2}}} d\xi \right] s_t,$$

where $\hat{x}_t$ is the solution of the SDE (2.14), P_t is the solution of the Ricatti equation (2.15) and

$$s_t = \exp\left[\int_0^t (H^*(s)\hat{x}_s + h^*(s))\, dY_s - \frac{1}{2}\int_0^t (H^*(s)\hat{x}_s + h^*(s))\,(H(t)\hat{x}_t + h(t))\, ds \right].$$

2.3 Preparatory Results

As we indicated in the first section, first we intend to modify the probability on Ω, in order to transform the process Y_t into a Brownian motion. We use Girsanov's theorem. For this, we introduce Z_t defined by

$$Z_t = \exp\left(- \int_0^t h^*(s, X_s) dV_t - \frac{1}{2}\int_0^t |h(s, X_s)|^2\, ds \right) \tag{2.16}$$

To apply the Girsanov Theorem, we need Z_t to be a martingale. For this we need to impose the following condition:

$$E\left[Z_t |h(t, X_t)|^2\right] < C, \text{ for all } t > 0. \tag{2.17}$$

Remark 2 *If h is bounded, then (2.17) holds.*

We prove this in the next proposition

Proposition 1 *The process $\{Z_t, \mathcal{F}_t; t \geq 0\}$ is a martingale.*

Proof. By differentiating (2.16) with respect to time, we obtain

$$\begin{aligned} dZ_t &= -Z_t h^*(t, X_t) dV_t \\ Z_t &= 1 - \int_0^t Z_s h^*(s, X_s) dV_s \end{aligned}$$

hence Z_t is a positive, continuous, local martingale, and therefore a (continuous) supermartingale. To prove that Z_t is a martingale it is enough to show that it has constant expectation. First, since Z_t is a (continuous) local martingale, there exists a sequence T_n of stopping times such that $\lim_{n\to\infty} T_n = \infty$ and $Z_{t\wedge T_n}$ is a martingale (by the definition of a local martingale). Using Fatou's lemma (see Appendix) we obtain that

$$EZ_t = E\lim_{n\to\infty} Z_{t\wedge T_n} \leq \liminf EZ_{t\wedge T_n} = 1,$$

hence Z_t is integrable and $EZ_t \leq 1$. Using Itô's formula we have that

$$\frac{Z_t}{1+\varepsilon Z_t} = \frac{1}{1+\varepsilon} + \int_0^t \frac{Z_s h^*(s,X_s)}{(1+\varepsilon Z_s)^2} dV_s - \int_0^t \frac{\varepsilon Z_s^2 \left|h(s,X_s)\right|^2}{(1+\varepsilon Z_s)^3} ds \tag{2.18}$$

Note that
$$E\left[\int_0^t \left(\frac{Z_s h^*(s,X_s)}{(1+\varepsilon Z_s)^2}\right)^2 ds\right] \leq \frac{1}{\varepsilon^2} E\left[\int_0^T \left|h(s,X_s)\right|^2 ds\right]$$

and from (2.6) the integral is finite. Hence the second term in (2.18) is a martingale with null expectation. By taking expectation in (2.18), we obtain

$$E\left[\frac{Z_t}{1+\varepsilon Z_t}\right] = \frac{1}{1+\varepsilon} - E\left[\int_0^t \frac{\varepsilon Z_s^2 \left|h(s,X_s)\right|^2}{(1+\varepsilon Z_s)^3} ds\right] \tag{2.19}$$

We take now the limit in (2.19) as ε tends to 0 and using the dominated convergence theorem (based on (2.17)), we obtain our claim.

■

We can now define $\tilde{P}$ a new probability measure such that the Radon-Nikodym derivative with respect to P is

$$\left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_T} = Z_T$$

Proposition 2 *Under $\tilde{P}$, the observation process Y is a Brownian motion independent of X.*

Proof. By Girsanov's theorem (see the Appendix), the process $Y_t = V_t + \int_0^t h(s,X_s)ds$ is a Brownian motion with respect to the new probability measure. Also, for each $T > 0$, the law of the process consisting of the pair (X,Y) on the interval $[0,T]$ is absolutely continuous with respect to the law of the process (X,V) on the interval $[0,T]$ (since the latter is equal to the former plus a drift term) and its Radon-Nikodym derivative is exactly Z_T, i.e., for any bounded measurable function f

$$E[f(X,Y)Z_T] = E[f(X,V)], \tag{2.20}$$

where in (2.20) both processes are regarded up to time T. Hence $\widetilde{E}[f(X,Y)] = E[f(X,V)]$ and therefore X and Y are independent under $\tilde{P}$.

■

Proposition 3 *Let U be an integrable $\mathcal{F}_t$ -measurable random variable. Then we have*

$$\tilde{E}[U|\mathcal{Y}_t] = \tilde{E}[U|\mathcal{Y}] \tag{2.21}$$

Proof. Let us denote by

$$\mathcal{Y}_t' = \sigma(Y_{t+u} - Y_t; u \geq 0), \tag{2.22}$$

then $\mathcal{Y} = \mathcal{Y}_t \bigvee \mathcal{Y}_t'$. Under the new probability measure $\mathcal{Y}_t' \subset \mathcal{Y}$ is independent of $\mathcal{F}_t$ because Y is an $\mathcal{F}_t$-adapted Brownian motion. Hence

$$\tilde{E}[U|\mathcal{Y}_t] = \tilde{E}[U|\mathcal{Y}_t \bigvee \mathcal{Y}_t'] = \tilde{E}[U|\mathcal{Y}]$$

■

It is the right time to introduce a new player into the game - ρ_t - the *unnormalised conditional distribution* of X. First we introduce the notation $\tilde{Z}_t \overset{\text{def}}{=} Z_t^{-1}$. Under $\tilde{P}$, $\tilde{Z}_t$ satisfies the following stochastic differential equation:

$$d\tilde{Z}_t = \tilde{Z}_t h^*(t, X_t) dY_t \tag{2.23}$$

and

$$\tilde{Z}_t = \exp\left(\int_0^t h^*(s, X_s) dY_t - \frac{1}{2}\int_0^t |h(s, X_s)|^2 ds\right) \tag{2.24}$$

then $\tilde{E}\tilde{Z}_t = E\tilde{Z}_t Z_t = 1$, so $\tilde{Z}_t$ is a martingale under $\tilde{P}$ and we have $\frac{dP}{d\tilde{P}} = \tilde{Z}_t$ on $\mathcal{F}_t$, $t \geq 0$. For every φ bounded Borel measurable function, we define $\rho_t(\varphi)$ to be

$$\rho_t(\varphi) \overset{\text{def}}{=} \tilde{E}[\varphi(X_t)\, \tilde{Z}_t | \mathcal{Y}_t] \tag{2.25}$$

and $\tilde{E}$ is the expectation with respect to $\tilde{P}$. We observe that, due to the previous proposition, in (2.25) we can take $\mathcal{Y}$ instead of $\mathcal{Y}_t$.

The following formula, called the Kallianpur-Striebel formula gives us the basic identity used in the Monte-Carlo method.

Proposition 4 (Kallianpur-Striebel) *For every φ bounded Borel measurable function we have*

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(1)}, \quad \widetilde{P} - \text{a.s.}. \tag{2.26}$$

Proof.

We show that $\pi_t(\varphi)\rho_t(1) = \rho_t(\varphi)$, $\widetilde{P}$-a.s., which is equivalent to

$$\pi_t(\varphi)\tilde{E}[\tilde{Z}_t|\mathcal{Y}_t] = \tilde{E}[\varphi(X_t)\tilde{Z}_t|\mathcal{Y}_t], \quad \widetilde{P} - \text{a.s.}.$$

Let b be an arbitrary $\mathcal{Y}_t$-measurable, bounded random variable. We have the following sequence of identities

$$\begin{aligned} E\pi_t(\varphi)b &= E\varphi(X_t)b \\ \tilde{E}\pi_t(\varphi)b\tilde{Z}_t &= \tilde{E}\varphi(X_t)b\tilde{Z}_t \\ \tilde{E}\left[\pi_t(\varphi)\tilde{E}[\tilde{Z}_t|\mathcal{Y}_t]b\right] &= \tilde{E}\left[\tilde{E}[\varphi(X_t)\tilde{Z}_t|\mathcal{Y}_t]b\right], \end{aligned}$$

which prove the required identity and hence the proposition.

■

Remark 3 *Since $\widetilde{P}$ and P are absolutely continuous with respect to each other, we also have*

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(1)}, \quad P - \text{a.s.}.$$

The Kallianpur-Striebel formula explains the usage of the term *unnormalised* in the definition of ρ_t.

We end this section with several results, which we will use in the proofs in the next section, but are also interesting in their own right.

Lemma 1 *Let*

$$S_t = \left\{\varepsilon_t = \exp\left(i\int_0^t r_s^* dY_s + \frac{1}{2}\int_0^t r_s^* r_s ds\right); r_s \in L^\infty\left([0,t], \mathbb{R}^m\right)\right\} \tag{2.27}$$

Then S_t is a total set in $L^1(\Omega, \mathcal{Y}_t, \tilde{P})$, i.e., if $a \in L^1(\Omega, \mathcal{Y}_t, \tilde{P})$ and $\tilde{E}[a\varepsilon_t] = 0$, $\forall \varepsilon_t \in S_t$, then $a = 0$.

Proof. It is enough to show this for the set

$$S_t' = \left\{ \varepsilon_t = \exp\left(i \int_0^t r_s^* dY_s \right) \right\}$$

since any element from S_t is equal with an element from S_t' multiplied by a constant $(\exp(\frac{1}{2}\int_0^t r_s^* r_s ds))$. To do this, we take $t_1, t_2, ..., t_p \in (0, t)$; $t_1 < t_2 < ... < t_p$, then

$$\sum_{h=1}^{p} l_h^* Y_{t_h} = \sum_{h=1}^{p} \mu_h^* (Y_{t_h} - Y_{t_{h-1}}) = \int_0^t \beta_s^* dY_s,$$

where $l_1, l_2, ..., l_n$ are given, $t_0 = 0$ and

$$\mu_p = l_p, \mu_{p-1} = l_p + l_{p-1}, ..., \mu_1 = l_p + ... + l_1$$

$$\beta(t) = \begin{cases} \mu_h, & t \in (t_{h-1}, t_h),\ h = 1, ..., p \\ 0, & t \in (t_p, T) \end{cases}.$$

We have then

$$\tilde{E}\left[a \cdot \exp\left(i \sum_{h=1}^{p} l_h^* Y_{t_h} \right) \right] = 0 \tag{2.28}$$

and also

$$\tilde{E}\left[a \sum_{k=1}^{K} c_k \exp\left(i \sum_{h=1}^{p} l_{h,k}^* Y_{t_h} \right) \right] = 0 \tag{2.29}$$

$\forall K, \forall c_1, ..., c_K \in \mathbb{C}, l_{h,k} \in \mathbb{R}$. Let $F(x_1, ..., x_p)$ be a bounded continuous complex valued function defined on $((\mathbb{R}^m)^p$. By Weierstrass' theorem, there exists an uniformly bounded sequence of functions of the form

$$P^{\varsigma}(x_1, ..., x_p) = \sum_{k=1}^{K^{\varsigma}} c_k^{\varsigma} exp \left(i \sum_{h=1}^{p} (l_{h,k}^{\varsigma})^* x_h \right)$$

such that

$$\lim_{\varsigma \to \infty} P^{\varsigma}(x_1, ..., x_p) = F(x_1, ..., x_p).$$

Hence we have $\tilde{E}(aF(Y_{t_1}, ..., Y_{t_p})) = 0$ for every F bounded continuous function and by a further approximation argument, we can take F to be a bounded Borel function, measurable with respect to the σ-field $\sigma(Y_{t_1},...,Y_{t_p})$. Since $t_1,t_2,...,t_p$ were taken arbitrarily, we obtain that $\tilde{E}(ab)=0$, for any bounded $\mathcal{Y}_t-$ measurable function b, which in particular gives $\tilde{E}(a^2\wedge n)=0$ for arbitrary n, hence $a=0$, $\widetilde{P}-$a.s.. ■

Lemma 2 *Let* $\{U_t;t\geq 0\}$ *be a càdlàg,* $\mathcal{F}_t$*-adapted process such that*

$$\tilde{E}\int_0^T U_t^2dt<\infty,\ \forall T\geq 0, \tag{2.30}$$

then

$$\tilde{E}\left[\int_0^t U_s dY_s^j|\mathcal{Y}_t\right]=\int_0^t \tilde{E}[U_s|\mathcal{Y}_s]dY_s^j;\ t\geq 0,\ j=1,...,m. \tag{2.31}$$

Proof.

Every ε_t from the set S_t (2.27)satisfies the following stochastic differential equation

$$\varepsilon_t=1+\int_0^t i\varepsilon_s r_s^* dY_s$$

Then

$$\begin{aligned}\tilde{E}\left[\varepsilon_t\tilde{E}[\int_0^t U_s dY_s^j|\mathcal{Y}_t]\right] &= \tilde{E}\left[\varepsilon_t\int_0^t U_s dY_s^j\right]\\ &= \tilde{E}\left[\int_0^t U_s dY_s^j\right]+\tilde{E}\left[\int_0^t i\varepsilon_s r_s^j U_s ds\right]\\ &= \tilde{E}\left[\tilde{E}[\int_0^t i\varepsilon_s r_s^j U_s ds|\mathcal{Y}_t]\right]\\ &= \tilde{E}\left[\int_0^t i\varepsilon_s r_s^j\tilde{E}[U_s|\mathcal{Y}_s]ds\right]\\ &= \tilde{E}\left[\varepsilon_t\int_0^t\tilde{E}[U_s|\mathcal{Y}_s]dY_s^j\right]\end{aligned}$$

completes the proof of lemma. ■

Since for all $\varphi\in D(A)$, $\{M_t^\varphi,\mathcal{F}_t;t\geq 0\}$ (2.3) is a bounded martingale (hence square integrable), we can define the Itô integral with respect to it. We have the following lemma

Lemma 3 *Let* $\{U_t;t\geq 0\}$ *be a càdlàg,* $\mathcal{F}_t$*-adapted process such that*

$$\tilde{E}\int_0^T U_t^2dt<\infty,\forall T\geq 0$$

then

$$\tilde{E}[\int_0^t U_s dM_t^\varphi | \mathcal{Y}_t] = 0 \tag{2.32}$$

Proof. The proof is similar to the previous one. We take once again ε_t from the set S_t and obtain the following sequence of identities (we use the fact that the bracket between $M_t^\varphi$and Y is 0)

$$\begin{aligned}\tilde{E}\left[\varepsilon_t \tilde{E}[\int_0^t U_s dM_t^\varphi | \mathcal{Y}_t]\right] &= \tilde{E}\left[\varepsilon_t \int_0^t U_s dM_t^\varphi\right] \\ &= \tilde{E}\left[\int_0^t \varepsilon_s U_s dM_t^\varphi\right] \\ &= 0.\end{aligned}$$

■

2.4 The Zakai and Kushner - Stratonovitch Equations

We now have all the tools to prove that ρ_t satisfies the Zakai equation and π_t the Kushner - Stratonovitch equation. We do this in the next two theorems. For our purposes we restrict attention to the case when h is uniformly bounded, i.e.,

$$\|h\| = \max_{i=1,..,m} \sup_{x \in E} |h_i(x)| < \infty$$

Theorem 3 *The process ρ_t satisfies the following evolution equation, called the Zakai equation*

$$\rho_t(\varphi) = \pi_0(\varphi) + \int_0^t \rho_s(A\varphi)ds + \int_0^t \rho_s(h^*\varphi)dY_s, \quad a.s. \ \forall t, \tag{2.33}$$

for all φ in the domain of the infinitesimal generator A.

Proof.

The argument follows closely the one in [3] pp. 83-87. We first approximate $\tilde{Z}_t$ by $\tilde{Z}_t^\varepsilon$ given by

$$\tilde{Z}_t^\varepsilon = \frac{\tilde{Z}_t}{1 + \varepsilon \tilde{Z}_t}$$

Using Itô's rule and integration by parts, we find

$$d\tilde{Z}_t^\varepsilon \varphi(X_t) = \left[\tilde{Z}_t^\varepsilon A\varphi(X_t) - \varepsilon\varphi(X_t)(1+\varepsilon\tilde{Z}_t)^{-3}\tilde{Z}_t^2 |h(t, X_t)|^2\right] dt$$

$$+\tilde{Z}_t^\varepsilon dM_t^\varphi + \varphi(X_t)(1+\varepsilon\tilde{Z}_t)^{-2}\tilde{Z}_t h^*(t,X_t)dY_t$$

By taking conditional expectation and applying (2.31) and (2.32), we get

$$\begin{aligned}\tilde{E}[\tilde{Z}_t^\varepsilon \varphi(X_t)|\mathcal{Y}_t] &= \frac{\Pi_0(\varphi)}{1+\varepsilon}+ \\ &+\int_0^t \tilde{E}\left[\tilde{Z}_s^\varepsilon A\varphi(X_s) - \varepsilon\varphi(X_s)(\tilde{Z}_s^\varepsilon)^2(1+\varepsilon\tilde{Z}_s)^{-1}|h(s,X_s)|^2\,|\mathcal{Y}_s\right]ds \\ &+\int_0^t \tilde{E}\left[\varphi(X_s)\tilde{Z}_s^\varepsilon(1+\varepsilon\tilde{Z}_s)^{-1}h^*(X_s)\,|\mathcal{Y}_s\right]dY_s \qquad (2.34)\end{aligned}$$

Now let ε tend to 0. We have

$$\begin{aligned}\tilde{Z}_t^\varepsilon &\nearrow \tilde{Z}_t \\ \tilde{E}[\tilde{Z}_t^\varepsilon\varphi(X_t)|\mathcal{Y}_t] &\rightarrow \rho_t(\varphi) \qquad a.s. \\ \text{For almost every } s\in[0,t]\ \tilde{E}[\tilde{Z}_s^\varepsilon A\varphi(X_s)|\mathcal{Y}_s] &\rightarrow \rho_s(A(\varphi)) \qquad a.s.\end{aligned}$$

and, because the last sequence remains bounded by the integrable random variable $\|A\varphi\|\,\tilde{E}[\tilde{Z}_t|\mathcal{Y}_t]$, which is in $L^1((0,T)\times\Omega; dt\otimes d\tilde{P})$ we get

$$\int_0^t \tilde{E}[\tilde{Z}_s^\varepsilon A(\varphi)|\mathcal{Y}_s]ds \rightarrow \int_0^t \rho_s(A(\varphi))ds \qquad a.s.$$

Also

$$\lim_{\varepsilon\rightarrow 0}\varepsilon\varphi(X_s)(\tilde{Z}_s^\varepsilon)^2(1+\varepsilon\tilde{Z}_s)^{-1}|h(s,X_s)|^2 = 0\ \tilde{P}-a.s.,\ dt-a.e.$$

$$\varepsilon\left|\varphi(X_s)(\tilde{Z}_s^\varepsilon)^2(1+\varepsilon\tilde{Z}_s)^{-1}\right||h(X_s)|^2 \le \|h\|^2\,\|\varphi\|\,\tilde{Z}_s$$

which is integrable over the space $\Omega\times[0,t]$ with respect to $\tilde{P}\times dt$. Thus using the Dominated Convergence Theorem we obtain that the integral $\int_0^t\varepsilon\tilde{E}[\varphi(X_s)(\tilde{Z}_s^\varepsilon)^2(1+\varepsilon\tilde{Z}_s)^{-1}|h(s,X_s)|^2|\mathcal{Y}_s]ds$ tends to 0 as ε tends to 0.

It only remains to show that $\int_0^t\tilde{E}[\varphi(X_s)\tilde{Z}_s^\varepsilon(1+\varepsilon\tilde{Z}_s)^{-1}h^*(s,X_s)|\mathcal{Y}_s]dY_s \rightarrow \int_0^t\rho_s(h^*\varphi)dY_s$ to complete the proof of the Zakai theorem.

First note that $\int_0^t\tilde{E}[\varphi(X_s)\tilde{Z}_s^\varepsilon(1+\varepsilon\tilde{Z}_s)^{-1}h^*(s,X_s)|\mathcal{Y}_s]dY_s$ is a well defined square integrable martingale since

$$\tilde{E}\left[\int_0^t\left(\tilde{E}[\varphi(X_s)\tilde{Z}_s^\varepsilon(1+\varepsilon\tilde{Z}_s)^{-1}h^*(X_s)|\mathcal{Y}_s]\right)^2ds\right] \le \frac{t}{\varepsilon}\,\|\varphi\|^2\,\|h\|^2$$

We show now that $\int_0^t \rho_s(h^*\varphi)dY_s$ is well defined. The process $\tilde{E}[\tilde{Z}_s|\mathcal{Y}_s]$ is a martingale with respect to the Brownian filtration $\mathcal{Y}_s$ hence it has a continuous version (see Corollary 2.3.2 in [36]). This implies that $\rho_s(1)$ is bounded a.s. on compact intervals and also $\rho_s(h^*\varphi)$ is bounded on compact intervals. Hence

$$\tilde{P}\left[\int_0^t (\rho_s(h^*\varphi))^2 ds < \infty\right] = 1$$

and therefore the stochastic integral is well defined and is a local martingale.

$$\int_0^t \tilde{E}[\varphi(X_s)\tilde{Z}_s^\varepsilon(1+\varepsilon\tilde{Z}_s)^{-1}h^*(s,X_s)|\mathcal{Y}_s]dY_s - \int_0^t \rho_s(h^*\varphi)dY_s =$$

$$\int_0^t \tilde{E}[\varphi(X_s)\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}h^*(s,X_s)|\mathcal{Y}_s]dY_s.$$

In order to show that the term on the right hand side of the above inequality converges to 0, at least for a convenient subsequence ε_n, we use the following property of stochastic integrals (see, for instance [4], pp. 34)

Proposition 5 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{B_t, \mathcal{F}_t\}$ be a standard n-dimensional Brownian motion defined on this space and Ψ_n, Ψ be $\mathcal{F}_t$-adapted process such that $\int_0^t \Psi_n^2 ds < \infty$, $\int_0^t \Psi^2 ds < \infty$, P-a.s. and*

$$\lim_{n\to\infty}\int_0^t (\Psi_n - \Psi)^2 ds = 0$$

in probability, then

$$\lim_{n\to\infty}\sup_{t\in[0,T]}\left|\int_0^t (\Psi_n - \Psi)dB_s\right| = 0$$

in probability.

We show that

$$\int_0^t \left(\tilde{E}[\varphi(X_s)\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}h^*(s,X_s)|\mathcal{Y}_s]\right)^2 ds \to 0,\ \tilde{P}-a.s.,$$

at least for a convenient sequence ε_n. Obviously it is enough to show that

$$\int_0^t \left(\tilde{E}[\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}|\mathcal{Y}_s]\right)^2 ds \to 0\ \tilde{P}-a.s.. \tag{2.35}$$

Since

$$\tilde{E}\int_0^t \tilde{E}[\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}|\mathcal{Y}_s]ds$$

$$= \tilde{E}\int_0^t \varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}ds \to 0$$

for a convenient subsequence

$$\tilde{E}[\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}|\mathcal{Y}_s] \to 0,\ \widetilde{P}-a.s.,\ dt-a.e.$$

which implies, that $\widetilde{P}-a.s.$

$$\tilde{E}[\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}|\mathcal{Y}_s] \to 0, \text{ for almost every } s \in [0,t]$$

and, for the same subsequence

$$\left(\tilde{E}[\varepsilon\tilde{Z}_s^2(2+\varepsilon\tilde{Z}_s)(1+\varepsilon\tilde{Z}_s)^{-2}|\mathcal{Y}_s]\right)^2 \to 0, \text{ for almost every } s \in [0,t]$$

Moreover, this is bounded by $4(\tilde{E}[\tilde{Z}_s|\mathcal{Y}_s])^2$, which is continuous in time (see Corollary 2.3.2 in [36]), hence bounded on compact intervals, and (2.35) is proved.

■

From (2.26) and (2.33) one obtains that

Theorem 4 *The conditional distribution of X given the observation process Y, i.e., π_t, satisfies the following evolution equation (called the Kushner-Stratonovitch equation)*

$$\pi_t(\varphi) = \pi_0(\varphi) + \int_0^t \pi_s(A\varphi)ds + \int_0^t (\pi_s(h^*\varphi) - \pi_s(h^*)\pi_s(\varphi))(dY_s - \pi_s(h)ds) \quad (2.36)$$

where φ is in the domain of the infinitesimal generator A.

Proof.

From (2.33), one obtains that $\rho_t(1)$ satisfies the following equation

$$\rho_t(1) = 1 + \int_0^t \rho_s(h^*)dY_s$$

which gives

$$\rho_t(1) = 1 + \int_0^t \rho_s(1)\pi_s(h^*)dY_s$$

and using Itô's rule one can prove that $\rho_t(1)$ is explicitly given by

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

and, hence

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

$$d\frac{1}{\rho_t(1)} = \frac{1}{\rho_t(1)}\left[-\pi_t(h^*)dY_t + \pi_t(h^*)\pi_t(h)dt\right] \tag{2.37}$$

By using (stochastic) integration by parts, (2.37), the Zakai equation for $\rho_t(\varphi)$ and the Kallianpur-Striebel formula, we obtain the stochastic differential equation satisfied by π_t

$$\begin{aligned}
\pi_t(\varphi) &= \rho_t(\varphi)\cdot\frac{1}{\rho_t(1)} \\
d\pi_t(\varphi) &= \pi_t(A(\varphi))dt + \pi_t(h^*\varphi)dY_t + \pi_t(\varphi)\left[-\pi_t(h^*\varphi)dY_t + \pi_t(h^*)\pi_t(h)dt\right] \\
&\quad -\pi_t(h^*\varphi)\pi_t(h)dt
\end{aligned}$$

which gives us the result.

■

The two theorems are valid for h unbounded and only satisfying (2.6), proofs for the two theorems under wider conditions can be found in [3], [4], [36] and the references therein.

2.5 Uniqueness Results

This section contains two similar results: if a process (belonging to a class to be specified) satisfies the equation (2.33) (respectively (2.36)), then it is almost surely equal to ρ_t (respectively to π_t). Let $\mathcal{B}^w(E)$ be the space of bounded Borel functions endowed with the *bp-convergence* (bounded and pointwise) topology, i.e., $a_n \in \mathcal{B}^w(E)$ converges to $a \in \mathcal{B}^w(E)$, if $\sup_{x\in E}|a_n(x)|$ remains bounded and $a_n(x) \to a(x)$, for all $x \in E$. The bp-convergence topology is also known as the weak topology on $\mathcal{B}(E)$ (the space of bounded Borel functions). Let also $(L^1(\Omega,\mathcal{Y}_t,\tilde{P}))^w$ be the space of $\mathcal{Y}_t$-measurable, integrable random variables endowed with the weak topology, $(L^\infty([0,T],\mathcal{B}(E)))^w$ be the space of uniformly bounded functions defined on $[0,t]$ with values in $\mathcal{B}(E)$ endowed with the weak topology and

$$\widetilde{L}^1_{\mathcal{Y}}(0,T) = \left\{\theta(t,\omega) \in L^1\left((0,T)\times\Omega; dt\otimes d\widetilde{P}\right), \text{for almost all } t, \theta(t,\cdot) \in L^1(\Omega,\mathcal{Y}_t,\widetilde{P})\right\}$$

Remark 4 *The unnormalised conditional probability* ρ *belongs to the class* $\mathcal{L}((L^\infty([0,T],\mathcal{B}(E)))^w,(\widetilde{L}^1_{\mathcal{Y}}(0,T))^w)$ *and for all* t, $\rho_t \in \mathcal{L}(\mathcal{B}^w(E),(L^1(\Omega,\mathcal{Y}_t,\widetilde{P}))^w)$.

We want to prove that if ρ' (respectively, π') belongs to same class as ρ (respectively, π) and satisfies the Zakai equation (respectively, the Kushner-Stratonovitch equation) then ρ and ρ' coincide. For this to hold we need to introduce an extra assumption:

Assumption **A**. For any smooth bounded function $r : [0,T] \to \mathbb{R}^m$ there exists a set $\mathcal{M}$ of functions bp-dense in $C_b(E)$ such that, for every $\varphi \in \mathcal{M}_{T,r}$ the equation

$$\begin{cases} \frac{\partial\psi(t,x)}{\partial t} - A\psi(t,x) + i\psi(t,x)h^*(t,x)r(t) = 0 \\ \psi(T,x) = \varphi(x) \end{cases} \tag{2.38}$$

has a unique solution ψ such that $\psi(t,\cdot) \in \mathcal{D}(A)$ and $\frac{\partial\psi(t,.)}{\partial t} \in C_b(E)$ for all $t \in [0,T]$.

Next we present a generic case where the property **A** holds. Let $E = \mathbb{R}^d$ and A be the second order elliptic differential operator

$$A = \sum_{i=1}^d f_i \frac{\partial}{\partial x_i} + \sum_{i,j=1}^d a_{ij}\frac{\partial^2}{\partial x_i \partial x_j}$$

where $f_i, a_{ij} : \mathbb{R}^d \to \mathbb{R}$, $i,j = 1...d$ are continuous bounded functions and $\mathcal{D}(A)$ consists of the set of bounded continuous functions with bounded continuous first and second derivatives. Then we have the following theorem (with a proof similar to that of Proposition 4.2.1, pp. 90, [3]),

Theorem 5 *Let us assume that* $\varphi : \mathbb{R}^d \to \mathbb{C}$ *is a bounded continuous twice differentiable function such that* $\frac{\partial\varphi}{\partial x_i}$, $\frac{\partial^2\varphi}{\partial x_i\partial x_j}$, $i,j = 1...d$ *are bounded and the following two conditions are satisfied by* h *and the coefficients of* A

i. h_i, $i = 1...m$ *are bounded continuous functions differentiable in the time variable and twice differentiable in the space variable and* $\frac{\partial h_i}{\partial t}$, $\frac{\partial h_i}{\partial x_i}$, $\frac{\partial^2 h_i}{\partial x_i x_j}$, $i = 1...d$ *are uniformly* *bounded.*

ii. f_i, a_{ij}, $i,j = 1...d$, h_i *are continuous bounded functions, twice differentiable with the property that* $\frac{\partial f_i}{\partial x_i}$, $\frac{\partial^2 f_i}{\partial x_i x_j}$, $\frac{\partial a_{ij}}{\partial x_k}$, $\frac{\partial^2 a_{ij}}{\partial x_k x_l}$, $i,j,k,l = 1...d$ *are uniformly bounded.*

then the system (2.38) has a unique solution ψ *such that* $\frac{\partial \psi}{\partial t}$, $\frac{\partial \psi}{\partial x_i}$, $\frac{\partial^2 \psi}{\partial x_i x_j}$ *are bounded.*

Let Θ denote the class of functions where we seek solutions to the two equations, i.e.,

$$\begin{aligned}\Theta &= \Big\{p(\cdot) \in \mathcal{L}((L^\infty([0,T],\mathcal{B}(E)))^w, (\tilde{L}^1_{\mathcal{Y}}(0,T))^w), \\ &\quad \forall t, p(\cdot) \in \mathcal{L}(\mathcal{B}^w(E), (L^1(\Omega, \mathcal{Y}_t, \tilde{P}))^w)\Big\}.\end{aligned}$$

Theorem 6 *Assuming that A holds, then there exists a unique solution of the Zakai equation (2.33) in the class* Θ.

Proof. We want to show that for arbitrary T , if we have two solutions of the Zakai equation ρ and ρ' then

$$\rho_T(\varphi) = \rho'_T(\varphi),\ \forall \varphi \in \mathcal{B}^w(E)$$

In fact, it is enough to show the above identity for φ belonging to a bp-dense subset of $\mathcal{B}^w(E)$. If $\varepsilon_T \in S_T$ (see (2.27)), then, using stochastic integration by parts and the Zakai equation, we obtain

$$\tilde{E}[\rho_T(v(T))\varepsilon_T] = \pi_0(v(0)) + \tilde{E}\left[\int_0^T \rho_s(v(s))\varepsilon_s \left(\frac{\partial v(s)}{\partial s} - Av(s) + iv(s)h^*(s)r(s)\right) ds\right] \tag{2.39}$$

where v is a reasonably smooth bounded function. The analogous equation holds for ρ'. Now if we take v to be the solution of (2.38) for $\varphi \in \mathcal{M}_{T,r}$ and r smooth and bounded, then

$$\tilde{E}[\rho_T(\varphi)\varepsilon_T] = \pi_0(\psi(0,\cdot)) = \tilde{E}[\rho'_T(\varphi)\varepsilon_T]$$

which implies that $\tilde{E}[\rho_T(\varphi)\varepsilon_T] = \tilde{E}[\rho'_T(\varphi)\varepsilon_T]$ for any Borel bounded r and using Lemma 1 we get that $\rho_T(\varphi) = \rho'_T(\varphi)$ for all $\varphi \in \mathcal{M}_{T,r}$. Since $\mathcal{M}_{T,r}$ is bp-dense in $C_b(E)$ thus also in $\mathcal{B}^w(E)$, we get that $\rho_T = \rho'_T$ as elements in $(L^1(\Omega, \mathcal{Y}_t, \tilde{P}))^w$. Since T was arbitrarily taken, Zakai equation has a unique solution in any interval $[0,T]$.

■

Before proceeding to prove the uniqueness of the solution of the Kushner-Stratonovitch equation we need the following simple lemma

Lemma 4 *For all $t \in [0,T]$, we have*

$$E[Z_t|\mathcal{Y}_t]\widetilde{E}[\widetilde{Z}_t|\mathcal{Y}_t] = 1, \ \widetilde{P} - \text{a.s.} \tag{2.40}$$

Proof. Let a be a bounded $\mathcal{Y}_t$-measurable function, then

$$E[E[Z_t|\mathcal{Y}_t]a] = E[Z_t a]$$

Hence

$$\widetilde{E}[E[Z_t|\mathcal{Y}_t]a\widetilde{Z}_t] = \widetilde{E}[Z_t a\widetilde{Z}_t] = \widetilde{E}[1 \cdot a]$$

which in turn implies (2.40).

■

Remark 5 *Since P is absolutely continuous with respect to $\widetilde{P}$, (2.40) holds also* $P-$a.s..

Theorem 7 *Assuming that A holds, there exists a unique solution of the Kushner-Stratonovitch equation (2.36) in the class Θ.*

Proof. The key point of the proof is that the integrated forms of the Zakai and Kushner Stratonovitch equation coincide, except that P plays the rôle of $\tilde{P}$. But first observe that using the notation of the previous theorem

$$\begin{aligned}
\widetilde{E}[\rho_T(v(T))\varepsilon_T] &= E[\rho_T(v(T))\varepsilon_T E[Z_t|\mathcal{Y}_t]] \\
&= E\left[\rho_T(v(T))\varepsilon_T \frac{1}{\widetilde{E}[\widetilde{Z}_t|\mathcal{Y}_t]}\right] \\
&= E\left[\pi_T(v(T))\varepsilon_T\right]
\end{aligned}$$

and proving the analogous formula from the right hand side of (2.39) we obtain that

$$E[\pi_T(v(T))\varepsilon_T] = \pi_0(v(0)) + E\left[\int_0^T \pi_s(v(s))\varepsilon_s\left(\frac{\partial v(s)}{\partial s} - Av(s) + iv(s)h^*(s)r(s)\right)ds\right]$$

and uniqueness follows from a similar argument to before.

■

2.6 The Measure Valued Processes asociated to the Filtering Problem

This section contains results concerning the existence and uniqueness of the solution of Zakai equation and Kushner Stratonovitch equation, regarded as measure valued processes. The proof of the results can be found in [30] and the references therein.

Proposition 6 *There exists a $\mathcal{P}(E)$-valued, càdlàg, $\mathcal{Y}_t$-adapted process α such that*

$$(\alpha(t), \varphi) = \pi_t(\varphi), \text{ for all } t \geq 0, \tag{2.41}$$

for all $\varphi \in \mathcal{B}(E)$, and the identity (2.41) means that the two processes are indistinguishable.

We define the $M_F(E)$-valued process U as follows

$$U(t) = \exp\left(\int_0^t (\alpha(s), h^*) dY_s - \frac{1}{2}\int_0^t |(\alpha(s), h)|^2\, ds\right) \alpha(t) \tag{2.42}$$

It can be shown that $U(t)$ satisfies

$$(U(t), \varphi) = \rho_t(\varphi) \tag{2.43}$$

for all $\varphi \in \mathcal{B}(E)$ and the identity (2.43) means that the two processes are indistinguishable. From (2.33) and (2.36), we have that U satisfies the equation

$$(U(t), \varphi) = (\alpha(0), \varphi) + \int_0^t (U(s), A\varphi) ds + \int_0^t (U(s), h^*\varphi) dY_s$$

and α satisfies the equation

$$\begin{aligned}(\alpha(t), \varphi) \quad &= \quad (\alpha(0), \varphi) + \int_0^t (\alpha(s), A\varphi) ds + \\ &\int_0^t \left((\alpha(s), h^*\varphi) - (\alpha(s), h^*)(\alpha(s), \varphi)\right)(dY_s - (\alpha(s), h) ds)\end{aligned}$$

for every $\varphi \in \mathcal{D}(A)$, $t > 0$.

Let A_0 be the restriction of A to $C_0(E)$. We assume that the domain of A_0 is a *dense algebra* in $C_0(E)$ (in particular that if $f \in \mathcal{D}(A_0)$, then $f^2 \in \mathcal{D}(A_0)$) and that the martingale problem for A_0 is well posed.

Theorem 8 *Assume that* A_0 *satisfies the conditions from above, i.e.,* $A_0 : C_0(E) \longrightarrow C_0(E)$, $\mathcal{D}(A_0)$ *is a dense algebra of* $C_0(E)$*, the martingale problem for* A_0 *is well posed and* $fh_i \in C_0(E)$ *for all* $f \in \mathcal{D}(A_0)$, $1 \le i \le m$*. Let* α' *be an* $\mathcal{Y}_t$*-adapted cadlag* $\mathcal{P}(E)$*-valued process such that*

$$\begin{aligned}(\alpha'(t), \varphi) &= (\alpha(0), \varphi) + \int_0^t (\alpha'(s), A\varphi) ds + \\ &\quad \int_0^t \left((\alpha'(s), h^*\varphi) - (\alpha'(s), h^*)(\alpha'(s), \varphi)\right) \\ &\quad \times \left(dY_s - (\alpha'(s), h) ds\right) \end{aligned} \tag{2.44}$$

for every $\varphi \in \mathcal{D}(A_0)$, $t \le T$*. Then* $\alpha'(t) = \alpha(t)$ *for all* $t \le T$ *a.s..*

Theorem 9 *Assume that* A_0 *satisfies the conditions from above, i.e.,* $A_0 : C_0(E) \longrightarrow C_0(E)$, $\mathcal{D}(A_0)$ *is a dense algebra of* $C_0(E)$*, the martingale problem for* A_0 *is well posed and* $fh_i \in C_0(E)$ *for all* $f \in \mathcal{D}(A_0)$, $1 \le i \le m$*. Let* U' *be an* $\mathcal{Y}_t$*-adapted cadlag* $\mathcal{P}(E)$*-valued process such that*

$$(U'(t), \varphi) = (\alpha(0), \varphi) + \int_0^t (U'(s), A\varphi) ds + \int_0^t (U'(s), h^*\varphi) dY_s \tag{2.45}$$

for every $\varphi \in \mathcal{D}(A_0) \cup \{1\}$, $t \le T$*. Then* $U'(t) = U(t)$ *for all* $t \le T$ *a.s..*

Chapter 3

Finite dimensional Filters

3.1 The Benes Filter

The Beneš filter (see [2]) is one of the few nonlinear filters that admit explicit solutions. In the following, we give a direct computation of the unnormalised conditional density for the Beneš filter using an explicit expression of the Laplace transform of a functional of Brownian motion. The same functional gives us the formula for the normalized conditional density for two particular filters.

Here is how we will proceed. We start by presenting the characteristics of the Beneš filter and then, in the next section we compute the Laplace transform of a functional of Brownian motion of the form

$$E\left[\exp\left(\alpha(B_t)+\int_0^t B_s^*\beta(s)ds-\frac{1}{2}\int_0^t \mid \Gamma B_s \mid^2 ds\right)\ |B_t=\delta\right],$$

where α, β are functions and Γ is a matrix. Using this computation we find in the last section an explicit form of the unnormalised conditional density for the Beneš filter and normalise in two particular cases.

We presume all the assumptions made in the previous chapter only this time $E \equiv \mathbb{R}$, $X_0 \equiv x_0 \in \mathbb{R}$ and the drift in (2.7) is linear, i.e., $h(x) = ax + b$, where $a, b \in \mathbb{R}$. Because we don't want too many indices, we will treat only the one dimensional case. An extension to the n-dimensional case can be made along the same lines as in [2]. Let $\mathcal{F}_t = \sigma(X_s; s \leq t)$. We assume that X is the solution of the martingale problem associated with the infinitesimal generator

$$A\varphi = \frac{1}{2}\varphi'' + f\varphi',$$

that is, for any $\varphi \in \mathcal{D}(A)$, the process

$$M_t^{\varphi} = \varphi(X_t) - \varphi(X_0) - \int_0^t A\varphi(X_s)ds, \qquad \mathcal{F}_t, \qquad 0 \le t < \infty,$$

is a martingale. We also assume that the drift coefficient satisfies the Beneš condition

$$f'(x) + f^2(x) + (ax)^2 = (px)^2 + 2xq + r, \quad p, q, r \in \mathbb{R}, \quad x \in \mathbb{R} \tag{3.1}$$

(f' is the derivative of f). Under these conditions (see [38], Chapter II, Th. 2), the process defined by the relation

$$M_t^x = X_t - x_0 - \int_0^t f(X_s)ds$$

is an $\mathcal{F}_t$-adapted local martingale with quadratic variation

$$< M^x >_t = t$$

hence a standard Brownian motion. In consequence, the process X is the solution of the following stochastic differential equation

$$X_t = x_0 + \int_0^t f(X_s)ds + V_t$$

where $\{V_t, \mathcal{F}_t; t \ge 0\}$ is a standard Brownian motion and, in fact we can consider the pair $\{(X_t, Y_t); t \ge 0\}$ as the solution of the following stochastic differential system

$$\begin{aligned} dX_t &= f(X_t)dt + dV_t, && (3.2) \\ dY_t &= (aX_t + b)dt + dW_t, && (3.3) \end{aligned}$$

where V and W are independent processes, $X_0 = x_0$ and $Y_0 = 0$. After the change of measure presented in the previous chapter, for a given $\omega \in \Omega$ and $Y_.(\omega)$ the corresponding path, we can express π_t using the Kallianpur-Striebel formula:

$$\pi_t(\varphi)(\omega) = \frac{\tilde{E}\left[\varphi(X_t)\exp\left(\int_0^t (aX_s + b)dY_s(\omega) - \frac{1}{2}\int_0^t (aX_s + b)^2 ds\right)\right]}{\tilde{E}\left[\exp\left(\int_0^t (aX_s + b)dY_s(\omega) - \frac{1}{2}\int_0^t (aX_s + b)^2 ds\right)\right]}.$$

In this case we are able to compute the density of the unnormalised conditional distribution of the signal

$$\rho_t(z)(\omega)dz = \tilde{E}\left[\mathbf{1}_{\{X_t \in dz\}} \exp\left(\int_0^t (aX_s + b)dY_s(\omega) - \frac{1}{2}\int_0^t (aX_s + b)^2 ds\right)\right] \tag{3.4}$$

and then, by normalising it, we obtain the density of π_t .

3.2 The Computation of an Exponential Functional of Brownian Motion

Let $\{B_t; t \geq 0\}$ be a d-dimensional Brownian motion, starting at the origin. Let also $\alpha : \mathbb{R}^d \to \mathbb{R}$ be a function, $\beta : [0, t] \to \mathbb{R}^d$ be a continuous function, Γ a $d \times d$ real matrix and $\delta \in \mathbb{R}^d$. In this section, we compute the following functional of B

$$E\left[\exp\left(\alpha(B_t) + \int_0^t B_s^* \beta(s)ds - \frac{1}{2}\int_0^t \mid \Gamma B_s \mid^2 ds\right) \quad | B_t = \delta\right]. \tag{3.5}$$

To obtain a closed formula for (3.5), we use Lévy's *diagonalisation procedure*, a powerful tool for deriving explicit formulae. Other results and techniques of this kind can be found in [44] and the references therein. For $s \leq t$ the orthogonal decomposition of B_s with respect to B_t is $B_s = \frac{s}{t}B_t + (B_s - \frac{s}{t}B_t)$ and using the Fourier decomposition of the Brownian motion (Wiener's construction of the Brownian motion)

$$B_s = \frac{s}{t}B_t + \sum_{k \geq 1} \sqrt{\frac{2}{t}} \frac{\sin \frac{ks\pi}{t}}{\frac{k\pi}{t}} \xi_k, \quad 0 \leq s \leq t, \tag{3.6}$$

where ξ_k are standard normal random vectors with independent entries and independent of B_t and the infinite sum has a subsequence of its partial sums uniformly convergent almost surely (cf [32], pp. 22). Using (3.6), the expression (3.5) becomes

$$e^{\nu}E\left[\exp\left(\sqrt{\frac{2}{t}}\sum_{k\geq 1}[\mu_k^* + (-1)^k\frac{t^2}{k^2\pi^2}\delta^*\Gamma^*\Gamma]\xi_k - \frac{1}{2}\int_0^t\left|\Gamma\left(\sum_{k\geq 1}\sqrt{\frac{2}{t}}\frac{\sin\frac{ks\pi}{t}}{\frac{k\pi}{t}}\xi_k\right)\right|^2 ds\right)\right], \tag{3.7}$$

where

$$\nu = \alpha(\delta) + \frac{\delta^*}{t}\int_0^t s\beta(s)ds - \frac{\mid\Gamma\delta\mid^2}{6}t$$

and

$$\mu_k = \int_0^t \beta(s)\frac{\sin\frac{ks\pi}{t}}{\frac{k\pi}{t}}ds.$$

Using the fact that

$$\int_0^t \sin\frac{k_1 s\pi}{t}\sin\frac{k_2 s\pi}{t}ds = 0, \quad \forall k_1, k_2 \geq 1, \quad k_1 \neq k_2$$

and integrating from 0 to t, the functional becomes

$$e^{\nu}E\left[\exp\sum_{k\geq 1}\left(\sqrt{\frac{2}{t}}(\mu_k^* + (-1)^k\frac{t^2}{k^2\pi^2}\delta^*\Gamma^*\Gamma)\xi_k - \frac{t^2}{2k^2\pi^2}\mid\Gamma\xi_k\mid^2\right)\right]. \tag{3.8}$$

Without loss of generality, we consider the case in which $\Gamma^*\Gamma$ is diagonal (one can choose the appropriate metric for this).

Let ζ be the exponential in (3.8) and γ_i be the i-th entry on the diagonal of $\Gamma^*\Gamma$. Let also μ_k^i, δ_i and β_i be the i-th coordinate of μ_n, δ and, β, respectively. We define the σ-fields

$$\mathcal{G}_k = \sigma(\xi_p \mid\ p \geq k),$$

$$\mathcal{G} = \bigcap_{k\geq 1}\mathcal{G}_k.$$

Using the independence of $\xi_1, ..., \xi_n, ...$ and the 0-1 -Law, we get that

$$\tilde{E}[\zeta] = \tilde{E}\left[\zeta\,\middle|\,\bigcap_{k\geq 1}\mathcal{G}_k\right].$$

Since $\mathcal{G}_k$ is a decreasing sequence of σ-fields, Lévy's 'Downward' theorem (cf [43], pp

136) tells us that

$$\tilde{E}\left[\zeta \middle| \bigcap_{k\geq 1} \mathcal{G}_k\right] = \lim_{k\to\infty} \tilde{E}\left[\zeta \,|\mathcal{G}_k\right].$$

Hence we can first determine $E[\zeta\,|\mathcal{G}_k]$ and then take the limit to obtain the expectation in (3.8). The result is

$$e^{\nu}\prod_{i=1}^{d}\frac{1}{\sqrt{\prod_{k\geq 1}\left[\frac{\gamma_i t^2}{k^2\pi^2}+1\right]}}\exp\sum_{k\geq 1}\frac{\left(\mu_k^i+(-1)^k\frac{\gamma_i\delta_i t^2}{k^2\pi^2}\right)^2}{t(\frac{\gamma_i t^2}{k^2\pi^2}+1)}. \tag{3.9}$$

The expression (3.9) contains one infinite product and three infinite sums which we compute using the following classical identities

$$\prod_{k\geq 1}\left[1+\frac{l^2}{k^2}\right]=\frac{\sinh(\pi l)}{\pi l}, \quad \sum_{k\geq 1}\frac{1}{z^2+k^2\pi^2}=\frac{1}{2z}(\coth z-\frac{1}{z}),$$

$$\sum_{k\geq 1}\frac{\cos kr}{z^2+k^2}=\frac{\pi}{2z}\frac{e^{(r-\pi)z}+e^{-(r-\pi)z}}{e^{\pi z}-e^{-\pi z}}-\frac{1}{2z^2}, \quad \forall r\in[0,2\pi],$$

$$\sum_{k\geq 1}(-1)^k\frac{\cos kr}{z^2+k^2}=\frac{\pi}{2z}\frac{e^{rz}+e^{-rz}}{e^{\pi z}-e^{-\pi z}}-\frac{1}{2z^2}, \quad \forall r\in[-\pi,\pi]$$

and $\sum_{k\geq 1}\frac{1}{k^2}=\frac{\pi^2}{6}$ (cf [34]); we finally find the closed formula for the Brownian functional (3.5)

$$e^{\alpha(\delta)}\prod_{i=1}^{d}\sqrt{\frac{t\sqrt{\gamma_i}}{\sinh(t\sqrt{\gamma_i})}}\exp\left(\int_0^t\int_0^s\frac{\sinh((s-t)\sqrt{\gamma_i})\sinh(s'\sqrt{\gamma_i})}{\sinh(t\sqrt{\gamma_i})\sqrt{\gamma_i}}\beta_i(s)\beta_i(s')ds'ds\right.$$

$$\left.+\delta_i\int_0^t\frac{\sinh(s\sqrt{\gamma_i})}{\sinh(t\sqrt{\gamma_i})}\beta_i(s)ds-\frac{\sqrt{\gamma_i}\coth(t\sqrt{\gamma_i})}{2}\delta_i^2+\frac{\delta_i^2}{2t}\right). \tag{3.10}$$

3.3 Application to the Filtering Problem

We use the explicit form (3.10) of (3.5) to compute the unnormalised conditional density of the signal (3.4). We note first that, the Itô integral $\int_0^t(aX_s+b)dY_s$ coincides with the Stratonovitch integral $\int_0^t(aX_s+b)dY_s$, so we can view this integral as if it were a Stieltjes integral with respect to a smooth function. Next we change the probability so that X becomes a Brownian motion. For this we introduce the process

$$U_t \stackrel{\text{def}}{=} \exp\left(-\int_0^t f(X_s)dV_s - \frac{1}{2}\int_0^t f^2(X_s)ds\right),$$

which is a martingale if we impose some restrictive conditions on f (linear growth for instance). We introduce the probability measure $\hat{P}$ such that

$$\frac{d\tilde{P}}{d\hat{P}}|_{\mathcal{F}_t} = U_t, \quad \forall t \geq 0.$$

Again Girsanov's theorem tells as that under $\hat{P}$

$$\widehat{V}_t \stackrel{\text{def}}{=} V_t - \left(-\int_0^t f(X_s)ds\right)$$

is a standard Brownian motion and using again the Kallianpur-Striebel formula and the independence of the processes X and Y, the expectation in (3.4) becomes

$$\hat{E}\mathbf{1}_{\{x_0+\hat{W}_t\in dz\}} \exp\left(\int_0^t (a(x_0+\widehat{V}_s)+b)dY_s() + \int_0^t (f(x_0+\widehat{V}_s))d\hat{W}_s\right.$$

$$\left. -\frac{1}{2}\int_0^t \left(a(x_0+\widehat{V}_s)+b\right)^2 + \left(f(x_0+\widehat{V}_s)\right)^2 ds\right) =$$

$$= \frac{e^{F(z)-F(x_0)+bY_t+x_0aY_t-x_0abt-\frac{b^2t}{2}-\frac{(z-x_0)^2}{2t}}}{\sqrt{2\pi t}}\hat{E}\left[\exp\left(\int_0^t a\hat{W}_s(dY_s(\omega)-bds)\right.\right.$$

$$\left.\left.-\frac{1}{2}\int_0^t (f'(x_0+\widehat{V}_s)+f^2(x_0+\widehat{V}_s)+(a(x_0+\widehat{V}_s))^2)ds\right)\middle|\widehat{V}_t = z-x_0\right], \tag{3.11}$$

where F is an antiderivative of f. By imposing the Beneš condition (3.1), (3.11) can be written as

$$\frac{e^{F(z)-\frac{(z-x_0)^2}{2t}+v}}{\sqrt{2\pi t}}\hat{E}\left[\exp\left(\int_0^t \widehat{V}_s(adY_s(\omega)-uds) - \frac{1}{2}\int_0^t (p\hat{W}_s)^2 ds\right)|\widehat{V}_t = z-x_0\right],$$

where $v = bY_t - F(x_0) + x_0(aY_t - abt - q) - \frac{(b^2+(px_0)^2+r)t}{2}$ and $u = ab+q+p^2x_0$. Finally, using (3.10), we obtain the explicit form of the unnormalised conditional density

$$\rho_t(z)(\omega) = \sqrt{\frac{p}{2\pi\sinh(pt)}}\exp\left(\int_0^t\int_0^s \frac{\sinh(p(s-t))\sinh(ps')}{p\sinh(pt)}(adY_{s'}(\omega)-uds')(adY_s(\omega)-uds)\right.$$

$$+v+F(z)+(z-x_0)\int_0^t \frac{\sinh(ps)}{\sinh(pt)}(adY_s-uds)-\frac{p\coth(pt)}{2}(z-x_0)^2\Big). \tag{3.12}$$

We observe that

$$\sqrt{\frac{p}{2\pi\sinh(pt)}}\exp\Bigg(\int_0^t\int_0^s \frac{\sinh(p(s-t))\sinh(ps')}{p\sinh(pt)}(adY_{s'}(\omega)-uds')(adY_s(\omega)-uds)$$

$$+v-x_0\int_0^t \frac{\sinh(ps)}{\sinh(pt)}(adY_s(\omega)-uds)-\frac{p\coth(pt)}{2}x_0^2\Bigg)$$

is independent of z, hence the relevant part of $\rho_t(z)(\omega)$ is

$$\exp\left(F(z)+z\left(a\int_0^t \frac{\sinh(ps)}{\sinh(pt)}dY_s(\omega)+\frac{px_0}{\sinh(pt)}-\frac{ab+q}{p}\tanh\frac{tp}{2}\right)-\frac{p\coth(pt)}{2}z^2\right). \tag{3.13}$$

Remark 6 *We observe that, for large t, (3.13) is approximately equal to*

$$\exp\left(F(z)+z\left(a\int_{t'}^t \frac{\sinh(ps)}{\sinh(pt)}dY_s(\omega)-\frac{ab+q}{p}\right)-\frac{p}{2}z^2\right), \quad t'<<t$$

hence the past observations become quickly (exponentially) irrelevant.

If $a=p$, $q=0$ and $r>0$ in (3.1), then the drift in the equation (3.2) satisfies the particular Ricatti equation

$$f'(x)+f^2(x)=r, \quad x\in\mathbb{R},$$

which has the solution

$$f(x)=\sqrt{r}\frac{\kappa e^{\sqrt{r}x}-e^{-\sqrt{r}x}}{\kappa e^{\sqrt{r}x}+e^{-\sqrt{r}x}}, \quad \kappa\in\mathbb{R}.$$

In this case, we can explicitly normalize $\rho(t)$. By normalizing, we obtain the density $\pi_t(z)(\omega)$ of the conditional measure π_t. Namely, π_t will appear as

$$\pi_t(\phi)(\omega)=\int_{\mathbb{R}}\phi(z)\pi_t(z)(\omega)dz.$$

If we denote with $\iota_t(Y_\cdot(\omega))$ the quantity

$$\frac{a\int_0^t \frac{\sinh(as)}{\sinh(at)}dY_s(\omega) + \frac{ax_0}{\sinh(at)} - b\tanh\frac{at}{2}}{a\coth(at)},$$

then the normalized conditional density has the form

$$\pi_t(z)(\omega) = \sqrt{\frac{a\coth(at)}{2\pi}}\frac{\kappa e^{z\sqrt{r}} + e^{-z\sqrt{r}}}{\kappa e^{\iota_t(Y_\cdot(\omega))\sqrt{r}} + e^{-\iota_t(Y_\cdot(\omega))\sqrt{r}}}e^{-\frac{r}{2a\coth(at)} - \frac{a\coth(at)}{2}(z-\iota_t(Y_\cdot(\omega)))^2}$$

$$= \sqrt{\frac{a\coth(at)}{2\pi}}\exp\left(F(z) - F(\iota_t(Y_\cdot(\omega))) - \frac{r}{2a\coth(at)} - \frac{a\coth(at)}{2}(z - \iota_t(Y_\cdot(\omega)))^2\right). \tag{3.14}$$

Remark 7 *In this case, the conditional density of the signal, when t is very large is*

$$\pi_t(z)(\omega) \cong \sqrt{\frac{a}{2\pi}}\exp\left(F(z) - F\left(\iota_{t,t'}(Y_\cdot(\omega))\right) - \frac{r}{2a} - \frac{a}{2}\left(z - \iota_{t,t'}(Y_\cdot(\omega))\right)^2\right)$$

where $\iota_{t,t'}(Y_\cdot(\omega)) = \int_{t'}^t \frac{\sinh(as)}{\sinh(at)}dY_s(\omega) + x_0 - \frac{b}{a}$ and $t' \ll t$.

Another particular case when one can normalize is when the drift in (3.2) is linear, that is,

$$f(x) = cx + d, \quad c, d \in \mathbb{R}.$$

In this case, if $X_0 = 0$, then the conditional law $\pi_t()$ is Gaussian with mean

$$\frac{d + a\int_0^t \frac{\sinh(s\sqrt{a^2+c^2})}{\sinh(t\sqrt{a^2+c^2})}dY_s(\omega) - \frac{\tanh(\frac{t}{2}\sqrt{c^2+a^2})}{\sqrt{a^2+c^2}}(ab+cd)}{\sqrt{a^2+c^2}\coth(t\sqrt{a^2+c^2}) - c}$$

and variance

$$\frac{1}{\sqrt{a^2+c^2}\coth(t\sqrt{a^2+c^2}) - c},$$

as one would expect from computing the classical Ricatti equation for the variance and the stochastic differential equation for the mean.

Remark 8 *For large t, the conditional distribution of the signal is normal with mean approximately equal to*

$$\frac{d + a\int_{t'}^t \frac{\sinh(s\sqrt{a^2+c^2})}{\sinh(t\sqrt{a^2+c^2})}dY_s(\omega) - \frac{ab+cd}{\sqrt{a^2+c^2}}}{\sqrt{a^2+c^2} - c}$$

and variance roughly $\frac{1}{\sqrt{a^2+c^2}-c}$.

Chapter 4

Convergence Results

In this chapter we present a series of results which will be used in the next two chapters to construct several measure valued processes arising as limits of branching particle systems (Dawson - Watanabe processes). There are a good number of construction and characterisations of these processes, such as [12], [22], [28], [29], [41] [42]. All the proofs of the results presented here can be found in [1], [15] and [39]. We start with several background results on convergence in distribution.

4.1 Convergence in distribution

Let $\{E, d\}$ be a separable metric space and let $\mathcal{B}(E)$ denote the space of bounded Borel functions on E, let $C_b(E)$ denote the space of bounded continuous functions on E, $C_0(E)$ denote the space of continuous functions vanishing at infinity (we define this set in the case in which E is locally compact), $C_K(E)$ the space of continuous functions with compact support, $M_F(E)$ the space of positive finite Borel measures, $\mathcal{P}(E)$ the space of Borel probability measures on E.

We endow $C_b(E)$, $C_0(E)$, $C_K(E)$ with the topology generated by the supremum norm $\|\cdot\|$, where

$$\|f\| \stackrel{\text{def}}{=} \sup_{x\in E} |f(x)|$$

We remark that, with this norm, $C_b(E)$, $C_0(E)$ and $C_K(E)$ become Banach space and, since E is separable, $C_0(E)$ and $C_K(E)$ are also separable, but $C_b(E)$ is not necessarily.

We also endow $M_F(E)$ and $\mathcal{P}(E)$ with the weak topology, i.e., $\mu_n \in M_F(E)$ (respectively, $\mathcal{P}(E)$) converges weakly to $\mu \in M_F(E)$ (respectively, $\mathcal{P}(E)$) if for all $f \in C_b(E)$, $\lim_{n\to\infty}(\mu_n, f) = (\mu, f)$ (for $\nu \in M_F(E)$ (respectively, $\mathcal{P}(E)$) and $\varphi \in C_b(E)$, (ν, f) denotes the integral of f with respect to ν). We denote weak convergence by $\mu_n \Rightarrow \mu$. The *distribution* of an E-valued random variable X, with respect to a reference probability P, denoted by PX^{-1}, is the probability measure given by $PX^{-1}(B) \overset{\text{def}}{=} P(X \in B)$. A sequence X_n of E-valued random variables is said to converge in distribution to the E-valued random variable X, if PX_n^{-1} converges weakly to PX^{-1}, or equivalently, if

$$\lim_{n\to\infty} E[f(X_n)] = E[f(X)], \text{ for all } f \in C_b(E),$$

where E denotes the expectation with respect to the probability measure P. We will denote convergence in distribution by $X_n \Rightarrow X$.

We introduce the *Prohorov* metric on $\mathcal{P}(E)$

$$\rho(P, Q) = \inf\{\varepsilon > 0; P(F) \leq Q(F^\varepsilon) + \varepsilon \quad \text{for all} \quad F \subset E \quad \text{closed}\}, \tag{4.1}$$

where $F^\varepsilon = \{x \in E; d(x, F) < \varepsilon\}$.

Remark 9 *The Prohorov metric is uniformly bounded by 1, i.e.,*

$$\rho(P, Q) \leq 1, \textit{ for all } P, Q \in \mathcal{P}(E).$$

The next theorem shows us that weak convergence topology is metrisable; weak convergence is equivalent to convergence in the Prohorov metric.

Theorem 10 *Let (E, d) be a separable metric space and let $\{P_n\} \subset \mathcal{P}(E)$ and $P \in \mathcal{P}(E)$. Then the following six conditions are equivalent:*

a. $\lim_{n\to\infty} \rho(P_n, P) = 0$.

b. $P_n \Rightarrow P$.

c. $\lim_{n\to\infty} \int f dP_n = \int f dP$ *for all uniformly continuous* $f \in C_b(E)$.

d. $\limsup_{n\to\infty} P_n(F) \leq P(F)$ *for all closed sets* $F \subset E$.

e. $\liminf_{n\to\infty} P_n(G) \geq P(G)$ *for all open sets* $G \subset E$.

f. $\lim_{n\to\infty} P_n(A) = P(A)$ *for all P-continuity sets $A \subset E$ (A is a P-continuity set if $P(\partial A) = 0$).* .

Corollary 1 *Let P_n, $n = 1, 2, ...$, and P belong to $\mathcal{P}(E)$ and let S' be a Borel subset of S. For $n = 1, 2, ...$, suppose that $P_n(S') = P(S') = 1$ and let P'_n and P' be the restrictions of P_n and P to S'. Then $P_n \Rightarrow P$ on S if and only if $P'_n \Rightarrow P'$ on S'.*

Corollary 2 *Let (E, d) be a metric space and let $\{X_n\}$, $\{Y_n\}$, $n = 1, 2, ...$ and X be E-valued random variables. If $X \Rightarrow X$ and $d(X_n, Y_n) \to 0$ in probability, then $Y_n \to X$.*

The next theorem is actually Theorem 1.7, pp. 101 from [15].

Theorem 11 *If E is separable, then $\mathcal{P}(E)$ is separable. If, in addition, (E, d) is complete, then $(\mathcal{P}(E), \rho)$ is complete.*

For the next two chapters, we need a good criterion of convergence in distribution of a sequence of E-valued random variables $\{X_n\}$, or, equivalently, weak convergence of the distribution of $\{X_n\}$. A common approach for verifying the convergence of a sequence $\{x_n\}$ of elements of a metric (metrisable) space is to first show that $\{x_n\}$ is contained in some compact set and then to show that every convergent subsequence must converge to the same element x. Since we now know that $\mathcal{P}(E)$ is a metric (metrisable) space, we can use this approach to prove that the distribution of $\{X_n\}$ is convergent. So we would need to show that the sequence x_n is relatively compact. Consequently, we need criteria for relative compactness in $\mathcal{P}(E)$.

A probability measure $P \in \mathcal{P}(E)$ is said to be tight if for each $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $P(K) \geq 1 - \varepsilon$. A family of probability measures $\mathcal{M} \subset \mathcal{P}(E)$ is said to be tight if for each $\varepsilon > 0$ there exists a compact set $K \in E$ such that

$$\inf_{P \in \mathcal{M}} P(K) \geq 1 - \varepsilon$$

Proposition 7 *If (E, d) is complete and separable, then each $P \in \mathcal{P}(E)$ is tight.*

Theorem 12 (Prohorov) *Let (E, d) be complete and separable, and let $\mathcal{M} \subset \mathcal{P}(E)$. Then $\mathcal{M}$ is relatively compact if and only if $\mathcal{M}$ is tight.*

Tightness is a crucial concept since it gives us a convenient characterisation of relative compactness. Our goal is to look at convergence in distribution for processes.

Next we apply these results to the space $D_E[0, \infty)$ consisting of all right continuous functions $x : [0, \infty) \to E$ with left limits (càdlàg functions), i.e., for each $t \geq 0$,

$\lim_{s\to t+} x(s) = x(t)$ and $\lim_{s\to t-} x(s)$ exists. In order to apply the previous results we need to define a metric on $D_E[0,\infty)$ under which $D_E[0,\infty)$ is a complete separable metric space.

Let (E,r) be a metric space. Let Λ' be the collection of (strictly) increasing functions λ mapping $[0,\infty)$ onto $[0,\infty)$ (in particular, $\lambda(0) = 0$ and $\lim_{t\to\infty} \lambda(t) = \infty$, and λ is continuous). Let Λ be the set of Lipschitz continuous functions $\lambda \in \Lambda'$ such that

$$\gamma(\lambda) \overset{\text{def}}{=} \sup_{s>t\geq 0} \left|\log \frac{\lambda(s)-\lambda(t)}{s-t}\right| < \infty$$

For $x, y \in D_E[0,\infty)$, define

$$d(x,y) = \inf_{\lambda\in\Lambda} \left[\gamma(\lambda) \vee \int_0^\infty e^{-u} d(x,y,\lambda,u)du\right] \tag{4.2}$$

where

$$d(x,y,\lambda,u) = \sup_{t\geq 0} r(x(t\wedge u), y(\lambda(t)\wedge u)) \wedge 1$$

Remark 10 *The function d is a metric on $D_E[0,\infty)$ and the topology induced on $D_E[0,\infty)$ by d is called the Skorohod topology.*

Theorem 13 *If E is separable, then $D_E[0,\infty)$ is separable. If (E,r) is complete, then $(D_E[0,\infty), d)$ is complete.*

In order to apply Prohorov's theorem to $\mathcal{P}(D_E[0,\infty))$ we need to characterise the compact sets in $D_E[0,\infty)$. The conditions for compactness are stated in terms of the following modulus of continuity. For $x \in D_E[0,\infty)$, $\delta > 0$, and $T > 0$, define

$$w'(x,\delta,T) = \inf_{\{t_i\}_1^n} \max_i \sup_{s,t\in[t_{i-1},t_i)} r(x(s),x(t)), \tag{4.3}$$

where $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < t_1 < ... < t_{n-1} < T \leq t_n$ with $\min_{1\leq i\leq n}(t_i - t_{i-1}) > \delta$ and $n \geq 1$. Note that $w'(x,\delta,T)$ is nondecreasing in δ and in T, and that $n \geq 1$.

Theorem 14 *Let (E,r) be complete. Then the set $A \subset D_E[0,\infty)$ relatively compact if and only if the following two conditions hold:*

a. For every rational $t \geq 0$, there exists a compact set $\Gamma_t \subset E$ such that $x(t) \in \Gamma_t$ for *all* $x \in A$.

b. For each $T > 0$,

$$\lim_{\delta \to 0} \sup_{x \in A} w'(x, \delta, T) = 0$$

Remark 11 *In Theorem 14 it is actually necessary that for each* $T > 0$ *there exists a compact set* $\Gamma_T \subset E$ *such that* $x(t) \in \Gamma_T$ *for all* $0 \leq t \leq T$ *and all* $x \in A$.

Theorem 15 *Let* (E, r) *be complete and separable, and let* $\{X_n\}_{n=1}^{\infty}$ *be a sequence of processes with sample paths in* $D_E[0, \infty)$. *Then* $\{X_n\}_{n=1}^{\infty}$ *is relatively compact if and only if the following conditions hold:*

a. For every $\eta > 0$ *and rational* $t \leq 0$, *there exists a compact set* $\Gamma_{\eta,t} \subset E$ *such that*

$$\liminf_{n \to \infty} P(X_n(t) \in \Gamma_{\eta,t}) \geq 1 - \eta.$$

b. For every $\eta > 0$ *and* $T > 0$, *there exists* $\delta > 0$ *such that*

$$\limsup_{n \to \infty} P(w'(X_n, \delta, T) \geq \eta) \leq \eta.$$

Remark 12 *In fact, if* $\{X_n\}_{n=1}^{\infty}$ *is relatively compact, then the stronger compact containment condition holds; that is, for every* $\eta > 0$ *and* $T > 0$ *there is a compact set* $\Gamma_{\eta,T} \in E$ *such that*

$$\inf_n P(X_n(t) \in \Gamma_{\eta,t} \text{ for } 0 \leq t \leq T) \geq 1 - \eta$$

Theorem 16 *Let* E *be separable and let* X_n, $n = 1, 2, ...$, *and* X *be processes with sample paths in* $D_E[0, \infty)$.

a. If X_n *converges in distribution to* X, *then*

$$(X_n(t_1), ..., X_n(t_k)) \Rightarrow (X(t_1), ..., X(t_k)) \tag{4.4}$$

for every finite set $\{t_1, ..., t_k\} \subset [0, \infty)$, $k > 0$.

b. If $\{X_n\}$ *is relatively compact and there exists a dense set* $D \in [0, \infty)$ *such that (4.4) holds for every finite set* $\{t_1, ..., t_k\}$, *then* X_n *converges in distribution to* X.

Theorem 17 *Let* (E, r) *be complete and separable, and let* $\{X_n\}$ *be a sequence of* *processes with sample paths in* $D_E[0,\infty)$. *Suppose that the compact containment condition holds. Let* H *be a dense subset of* $C_b(E)$ *in the topology of uniform convergence on compact sets. Then* $\{X_n\}$ *is relatively compact if and only if* $f \circ \{X_n\}$ *is relatively compact as a family of processes with sample paths in* $D_{\mathbb{R}}[0,\infty)$ *for each* $f \in H$.

We end this section with the following result which is contained in the proof of Theorem 4.8.2 from [15].

Theorem 18 *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space on which we have defined the filtration* $\mathcal{F}_t$ *and* $\{M_t, t \geq 0\}$ *be an* $\mathcal{F}_t$*-adapted process. We assume that*

$$\mathcal{F}_t = \sigma(X_s, Y_u; s \in [0,t], u \in [0,\infty))$$

where X_t, Y_t *are processes with càdlàg paths and have values in some separable complete metric space* E. *Then* M_t *is a martingale with respect to the filtration* $\mathcal{F}_t$ *iff*

$$\int (M(t+s) - M(t))\Pi_{i=1}^{k} f_i(X(t_i))\Pi_{i'=1}^{k'} f'_{i'}(Y(t'_{i'}))dP = 0 \tag{4.5}$$

for all $k, k' \geq 0$, $0 \leq t_1 < t_2 < \ldots < t_k \leq t$, $0 \leq t'_1 < t'_2 < \ldots < t'_{k'} < \infty$, $s \geq 0$, $f_1, f_2, \ldots, f_k, f'_1, f'_2, \ldots, f'_{k'} \in C_b(E)$.

4.2 Convergence in distribution for measure valued processes

We now take the general results in the previous section and apply them to processes with values in $M_F(E)$ with E is a locally compact complete metric space. We endow $M_F(E)$ with the weak topology. To be able to apply the previous results, the weak topology on $M_F(E)$ should be metrisable. But $M_F(E)$ is homeomorphic with $[0,\infty) \times \mathcal{P}(E)$ with all the elements of type $0 \times P$ identified with a generic null element $\{0\}$, the homeomorphism being

$$\mu \in M_F(E) \longrightarrow (\mu(E), \frac{\mu}{\mu(E)}) \in (0,\infty) \times \mathcal{P}(E) \tag{4.6}$$

for μ non-trivial and the null measure corresponding to 0. We introduce the following metric on $M_F(E)$

$$d(\mu,\nu) = \rho(\frac{\mu}{\mu(E)}, \frac{\nu}{\nu(E)}) \times \min(\mu(E), \nu(E)) + |\mu(E) - \nu(E)|$$

if μ, ν are non-trivial and $d(\mu, 0) = \mu(E)$ for μ non-trivial finite measures.

Remark 13 *The function d is a distance and it generates the weak topology, i.e., $\lim_{n\to\infty} d(\mu_n, \mu) = 0$ iff $\lim_{n\to\infty}(\mu_n, f) = (\mu, f)$ for all $f \in C_b(E)$.*

Proof. In order to show that d is a distance, only the triangle rule is not be self-evident, but it is still simple algebra. Here are the generic cases

1. $0 < \mu(E) \leq \nu(E) \leq \xi(E)$ then

$$\begin{aligned} d(\mu,\nu) &\leq \left(\rho(\frac{\mu}{\mu(E)}, \frac{\xi}{\xi(E)}) + \rho(\frac{\nu}{\nu(E)}, \frac{\xi}{\xi(E)})\right) \mu(E) + \xi(E) - \nu(E) + \xi(E) - \mu(E) \\ &\leq d(\mu,\xi) + d(\xi,\nu) \end{aligned}$$

2. $0 < \mu(E) \leq \xi(E) \leq \nu(E)$ then

$$\begin{aligned} d(\mu,\nu) &\leq \left(\rho(\frac{\mu}{\mu(E)}, \frac{\xi}{\xi(E)}) + \rho(\frac{\nu}{\nu(E)}, \frac{\xi}{\xi(E)})\right) \mu(E) + \nu(E) - \xi(E) + \xi(E) - \mu(E) \\ &\leq d(\mu,\xi) + d(\xi,\nu) \end{aligned}$$

3. $0 < \xi(E) \leq \mu(E) \leq \nu(E)$ then

$$\begin{aligned} d(\mu,\nu) &\leq \rho(\frac{\mu}{\mu(E)}, \frac{\nu}{\nu(E)})(\mu(E) - \xi(E)) + \left(\rho(\frac{\mu}{\mu(E)}, \frac{\xi}{\xi(E)}) + \rho(\frac{\nu}{\nu(E)}, \frac{\xi}{\xi(E)})\right) \xi(E) \\ &\quad + \nu(E) - \mu(E) \\ &\leq d(\mu,\xi) + d(\xi,\nu) + \rho(\frac{\mu}{\mu(E)}, \frac{\nu}{\nu(E)})(\mu(E) - \xi(E)) - (\mu(E) - \xi(E)) \\ &\leq d(\mu,\xi) + d(\xi,\nu) \end{aligned}$$

4. $0 = \mu(E) < \nu(E) \leq \xi(E)$ then

$$d(\mu,\nu) = \nu(E) \leq \xi(E) + d(\xi,\nu) \leq d(\mu,\xi) + d(\xi,\nu)$$

5. $0 = \mu(E) < \xi(E) \leq \nu(E)$ then

$$d(\mu,\nu) = \nu(E) \leq \xi(E) + d(\xi,\nu) \leq d(\mu,\xi) + d(\xi,\nu)$$

6. $0 = \xi(E) < \mu(E) \leq \nu(E)$ then

$$d(\mu,\nu) \leq \nu(E) \leq \nu(E) + \mu(E) = d(\mu,\xi) + d(\xi,\nu)$$

We prove now the second assertion in the Remark. If μ is trivial, then

$$\lim_{n\to\infty} d(\mu_n,\mu) = 0 \Leftrightarrow \lim_{n\to\infty} \mu_n(E) = 0 \Leftrightarrow \lim_{n\to\infty} (\mu_n, f) = 0\, \forall f \in C_b(E).$$

If μ is non-trivial, there exists $N > 0$ such that for all $n \geq N$ the measure μ_n is non-trivial. Obviously

$$\begin{aligned}\lim_{n\to\infty}(\mu_n, f) = (\mu, f) &\Leftrightarrow \lim_{n\to\infty}\mu_n(E) = \mu(E) \text{ and } \lim_{n\to\infty}\rho\left(\frac{\mu_n}{\mu_n(E)}, \frac{\mu}{\mu(E)}\right) = 0\\ &\Leftrightarrow \lim_{n\to\infty}\mu_n(E) = \mu(E) \text{ and }\\ & \lim_{n\to\infty}\left(\frac{\mu_n}{\mu_n(E)}, f\right) = \left(\frac{\mu}{\mu(E)}, f\right)\ \forall f \in C_b(E).\\ &\Leftrightarrow \lim_{n\to\infty}(\mu_n, f) = (\mu, f),\ \forall f \in C_b(E).\end{aligned}$$

■

Remark 14 *The simpler 'distance'* $\rho(\frac{\mu}{\mu(E)}, \frac{\nu}{\nu(E)}) + |\mu(E) - \nu(E)|$ *would not be suitable, because of the above identification. Let* P *and* Q *be two different probability measures. Then* $\lim_{n\to\infty} \frac{1}{n}P = \lim_{n\to\infty}\frac{1}{n}Q = 0$. *Hence the distance between the terms of the sequence should tend to* 0; *instead it is constant and equal to* $\rho(P, Q)$.

From the identification (4.6) and Theorem 11 we have the following

Proposition 8 *If* E *is separable, then* $\mathcal{P}(E)$ *is separable. If, in addition,* (E, r) *is complete, then* $(M_F(E), d)$ *is complete.*

Let $\{f_k\}_{k>0}$ be a sequence uniformly bounded by 1 such that $\text{sp}\{f_k; k > 0\}$ is dense in $C_K(E)$ and we denote $f_0 \equiv 1$. Then $\{f_k\}_{k>0}$ is convergence determining and the function d' defined by

$$d'(\mu,\nu) = \sum_{k=0}^{\infty}\frac{1}{2^k}((\mu, f_k) - (\nu, f_k)) \tag{4.7}$$

is a metric equivalent to d. This fact leads to the following result (see [39] for a proof). If $\pi_{f_k}: M_F(E) \to \mathbb{R}$ is defined as follows

$$\pi_{f_k}(\mu) = (\mu, f_k)$$

then we have the following

Theorem 19 *Let* $\{X_n\}_{n>0}$ *be a sequence of processes with sample paths in* $D_{M_F(E)}[0,\infty)$. *If, for each* $k \in \mathbb{N}$, $(\pi_{f_k}(X_n))_{n>0}$ *forms a tight sequence of processes with sample paths in* $D_{\mathbb{R}}[0,\infty)$, *then* $\{X_n\}_{n>0}$ *is tight in* $D_{M_F(E)}[0,\infty)$.

We present now a result from [1] which will be used subsequently to prove that a sequence of processes satisfies condition b. from either Theorem 14 or Theorem 15. The theorem is stated for $E = \mathbb{R}$. Let then $\{X_n\}$ be processes with paths in $D_{\mathbb{R}}[0,\infty)$ and let $\{\tau_n, \delta_n\}$ be such that

a. for each n, τ_n is a stopping time to the process $\{X_n(t); 0 \leq t \leq 1\}$, with respect to the natural σ-field and τ_n takes only finitely many values;

b. for each n, δ_n is a constant, $0 \leq \delta_n \leq 1$, and $\delta_n \to 0$ as $n \to \infty$.

We are interested in the following condition on $\{X_n\}$:

$$X_n(\tau_n + \delta_n) - X_n(\tau_n) \to 0 \text{ for all sequences } \{\tau_n, \delta_n\} \text{ satisfying } \mathbf{a} \text{ and } \mathbf{b} \tag{4.8}$$

where the convergence is in probability.

Theorem 20 (Aldous) *Suppose that* $\{X_n\}$ *satisfies condition (4.8), and* $\{X_n(t)\}$ *is tight on the line, for each* $t \in [0,1]$. *Then* $\{X_n(t)\}$ *is tight in* $D_{\mathbb{R}}[0,\infty)$.

By combining Theorem 19 and Theorem 20 we obtain the following criterion for tightness (relative compactness) of sequences of processes with sample paths in $D_{M_F(E)}[0,\infty)$.

Theorem 21 *Let* $\{X_n\}_{n>0}$ *be a sequence of processes with sample paths in* $D_{M_F(E)}[0,\infty)$. *If the following two conditions are satisfied*

a. The mass process $\{(X_n(t),1)\}_{n>0}$ *satisfies the compact containment condition, i.e., for all* ε *there exists* M_ε *such that*

$$\inf_n P\left((X_n(t),1) < M_\varepsilon \;\forall\; t \in [0,\infty)\right) \geq 1 - \varepsilon,$$

b. For all $k \geq 0$ *the sequence* $\{(X_n(t), f_k)\}_{n>0}$ *satisfies condition (4.8),*

then the sequence is tight.

We present next a transcription of Theorem 16. Again its proof can be found in [39]

Theorem 22 *Let* $\{X_n\}_{n=1}^{\infty}, X$ *be processes with sample paths in* $D_{M_F(E)}[0,\infty)$ *and* $\{f_k\}_{k>0}$ *as above. If* $\{X_n\}_{n=1}^{\infty}$ *form a tight sequence and for all* $m > 0$ *and each* $t_1, t_2, ..., t_m \in \mathbb{R}_+$

$$((X_n(t_1), f_{k_1}), (X_n(t_2), f_{k_2}), ..., (X_n(t_m), f_{k_m}),)$$

$$\Rightarrow ((X(t_1), f_{k_1}), (X(t_2), f_{k_2}), ..., (X(t_m), f_{k_m}))$$

then $X_n \Rightarrow X$.

We end the section with a weak limit theorem for stochastic integrals from [31]. Let $\mathbb{M}^{km}$ denote the real-valued, $k \times m$ matrices.

Theorem 23 (Kurtz-Protter) *For each* n, *let* (X_n, Y) *be an* $\{F_t^n\}$*-adapted process with sample paths in* $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$ *and let* Y *be a standard* m*-dimensional Brownian motion. If* $(X_n, Y) \Rightarrow (X, Y)$ *in the Skorohod topology on* $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$, *then* $(X_n, Y, \int X_n dY_n) \Rightarrow (X, Y, \int X dY)$ *in the Skorohod topology on* $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. *If* $(X_n, Y) \to (X, Y)$ *in probability, then the triple converges in probability.*

Chapter 5

Measure Valued Processes Associated with the Zakai equation

In this section we will define two sequences of branching particle systems and then check that they are tight. We will then show that the first one converges to the solution of a 'filtered' martingale problem and its (conditional) expectation to ρ_t, the unnormalised conditional distribution of the signal and that the second one converges as well to ρ_t.

5.1 The Measure - Valued Process $\mathcal{X}$

In this section, we restrict ourselves to $E = \mathbb{R}^d$. So, once again, let $C(\mathbb{R}^d)$ be the set of continuous functions on $\mathbb{R}^d$, $C_b(\mathbb{R}^d)$ be the space of continuous bounded functions on $\mathbb{R}^d$, $C_K(\mathbb{R}^d)$ be the space on continuous functions with compact support and $C_0(\mathbb{R}^d)$ be the space of continuous functions which vanish at infinity. Let $M_F(\mathbb{R}^d)$ be the space of finite measures over $\mathbb{R}^d$ and $M_F'(\mathbb{R}^d)$ be the subspace of $M_F(\mathbb{R}^d)$ comprising finitely atomic measures:

$$M_F'(\mathbb{R}^d) \stackrel{\text{def}}{=} \{\mu \in M_F(\mathbb{R}^d) | \mu = \sum_{i=1}^{n} q_i \delta_{x_i}, \;\; x_i \in \mathbb{R}^d, \;\; q_i \in \mathbb{R}_+, \;\; i = 1, 2, ..., n\}$$

Let A be the following second order differential operator

$$A(t)\varphi(x) = \sum_{i=1}^{d} f_i(t,x)\frac{\partial\varphi(x)}{\partial x_i} + \sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}$$

where $a_{ij}(t,\cdot)$, $f_i(t,\cdot) \in C(\mathbb{R}^d)$ and

$$(a_{ij}(t,x)\xi,\xi) > 0 \,\forall t \in [0,T];\; x,\xi \in \mathbb{R}^d. \tag{5.1}$$

From (5.1) we deduce that there exists σ such that $a = \sigma\sigma^*$. Under these conditions the solution X of the martingale problem associated with the infinitesimal operator A satisfies the following stochastic differential equation (see [38], Chapter II, Th. 2)

$$X_t = X_0 + \int_0^t f(s,X_s)ds + \int_0^t \sigma(s,X_s)dB_s \tag{5.2}$$

where $\{B_t, \mathcal{F}_t, t \in [0,T]\}$ is a standard Brownian motion. The observation process Y_t was defined in Chapter 2 as

$$Y_t = \int_0^t h(s,X_s)ds + V_t, \qquad t \geq 0. \tag{5.3}$$

We assume that the coefficients of the system (5.2)+(5.3) satisfy sufficient conditions for the existence and uniqueness of the solution of the Zakai equation (see Chapter 2) and that h is bounded. Let $\|h\|$ be the quantity

$$\|h\| = \sup_{(t,x)\in[0,\infty)\times\mathbb{R}^d} \|h(t,x)\|.$$

We will assume that the domain $\mathcal{D}(A) \subset C_b(\mathbb{R}^d)$ of the infinitesimal generator $A(t)$ has the following property:

Assumption I. For every $f \in \mathcal{D}(A)$, there exists a sequence $f_n \in \mathcal{D}(A)$ such that $f_n^2 \in \mathcal{D}(A)$, $\sup \|f_n\| < \infty$ and $\lim_{n\to\infty} f_n = f$ pointwise and also $\sup \|Af_n\| < \infty$ and $\lim_{n\to\infty} Af_n = Af$ pointwise.

For example, if the coefficients of (5.2) are bounded, and $\mathcal{D}(A)$ is the set of bounded continuous functions with bounded first and second partial derivatives, then $\mathcal{D}(A)$ satisfies **I**. If the coefficients of (5.2) are continuous but not bounded, and $\mathcal{D}(A)$ is the set of twice continuously differentiable functions with compact support, then, also in this case, $\mathcal{D}(A)$ satisfies **I**. From now on, we work under the new probability measure $\tilde{P}$ and all the expectations and conditional expectations will be considered with respect to $\tilde{P}$.

5.1.1 The Construction of the Sequence of the Particle Systems $\mathcal{X}_n$

Let $\{(\mathcal{X}_n(t), \mathcal{F}_t),\ 0 \le t \le 1\}$ be a sequence of branching particle systems on $(\Omega, \mathcal{F}, \tilde{P})$ with values in $M'_F(\mathbb{R}^d)$ defined as follows:

a. Initial condition

1. $\mathcal{X}_n(0)$ is the occupation measure of n particles (we will denote the number of particles alive at time t by $N_n(t)$) of mass $\frac{1}{n}$, i.e.,

$$\mathcal{X}_n(0) = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i^n},$$

where $x_i^n \in \mathbb{R}^d$, for every $i, n \in \mathbb{N}$.

2. Assume that the occupation measure of the particles converges weakly to the initial distribution of the signal, i.e.,

$$\lim_{n\to\infty} (\mathcal{X}_n(0), \varphi) = \tilde{E}[\varphi(\xi)] = \pi_0(\varphi), \quad \forall \varphi \in C_b(\mathbb{R}^d).$$

b. Evolution in time

We describe the evolution of the processes in the interval $\left[\frac{i}{n}, \frac{i+1}{n}\right]$, $i = 0, 1, ..., n-1$.

1. At the time $\frac{i}{n}$, the process consists of the occupation measure of $N_n(\frac{i}{n})$ particles of mass $\frac{1}{n}$.

2. During the interval the particles move independently with the same law as the signal (5.2). Let $V(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$ be the trajectory of a generic particle in this interval.

3. At the end of the interval, each particle branches into a random number of particles with a mechanism depending on its trajectory in the interval. The mechanism is chosen so that it has finite second moment and the mean number of offspring for a particle given

the σ-field $\mathcal{F}_{\frac{i+1}{n}-} = \sigma(\mathcal{F}_s,\ s < \frac{i+1}{n})$ of events up to time $\frac{i+1}{n}$ is

$$\exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t, V(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t, V(t))dt\right) \tag{5.4}$$

and the variance is equal to $v_{\frac{i+1}{n}}$. The particles branch independently of each other.

In the description above v_s is an arbitrary bounded, positive function, continuous in time and $v_t \geq \frac{1}{4}$, $\forall t \in [0,1]$ (see the Appendix for this). The last condition ensures the existence of the required branching mechanism. We denote by $\|v\|$ the supremum of v over the interval $[0,1]$, i.e.,

$$\|v\| = \sup_{t\in[0,1]} v_t.$$

Just before the $(i+1)$-th branching, we will have $N_n(\frac{i}{n})$ particles. Let us denote by $\mathcal{X}_n(\frac{i+1}{n}-)$ the state of the process just before the $(i+1)$-th branching and by $V_n^j(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$ the trajectory of the j-th particle alive during the interval $(1 \leq j \leq N_n(\frac{i}{n}))$. Let also $q_n^j(\frac{i+1}{n})$ be the number of offspring of the j-th particle at time $\frac{i+1}{n}$.

Lemma 5 *We have the following relations:*

i. $\tilde{E}[N_n(t)] = N_n(0) = n,\ \forall n \geq 0,\ t \in [0,1]$.

ii. $\tilde{E}[N_n^2(t)] \leq e^{\|h\|^2\frac{[nt]}{n}}n^2 + \sum_{k\leq[nt]} v_{\frac{k}{n}} e^{\|h\|^2\frac{[nt]-k}{n}}$, $\forall n \geq 0,\ t \in [0,1]$ *(*$[x]$ *is the largest integer smaller that* x*).*

Proof.

i. N_n does not change during the intervals $(\frac{k}{n}, \frac{k+1}{n})$, $k = 1, ..., n-1$ so $N_n(t) = N_n(\frac{[tn]}{n})$. Therefore it suffices to prove that $\tilde{E}[N_n(\frac{i}{n})] = \tilde{E}[N_n(\frac{i+1}{n})]$ for $0 \leq i < n$. Using (5.4), we have

$$\begin{aligned}
\tilde{E}\left[N_n\left(\frac{i+1}{n}\right)\right] &= \tilde{E}\left[\sum_{j=1}^{N_n(\frac{i}{n})} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t, V_n^j(t))dt\right)\right] \\
&= \tilde{E}\left[\tilde{E}\left[\sum_{j=1}^{N_n(\frac{i}{n})} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t, V_n^j(t))dt\right)\middle|\mathcal{F}_{\frac{i}{n}}\right]\right] \\
&= \tilde{E}\left[\sum_{j=1}^{N_n(\frac{i}{n})} \tilde{E}\left[\exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t, V_n^j(t))dt\right)\middle|\mathcal{F}_{\frac{i}{n}}\right]\right] \\
&= E\left[N_n\left(\frac{i}{n}\right)\right]
\end{aligned}$$

since $s \to \exp\left(\int_{\frac{i}{n}}^{s} h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{s} h^*h(t, V_n^j(t))dt\right)$ is an $\mathcal{F}_s$-adapted martingale.

ii. From the construction of the branching mechanism of the particles we have that

$$\begin{aligned}
&\tilde{E}[(q_n^j(\frac{(i+1)}{n}))^2|\mathcal{F}_{\frac{i+1}{n}-}] \\
&= v_{\frac{i+1}{n}} + \left(\exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t, V_n^j(t))dt\right)\right)^2 \\
&\leq v_{\frac{i+1}{n}} + e^{\frac{\|h\|^2}{n}} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} 2h^*(t, V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} (2h)^*2h(t, V_n^j(t))dt\right)
\end{aligned}$$

This inequality and the independence of the particles implies (as in i.)

$$\begin{aligned}
\tilde{E}[(N_n(\frac{i+1}{n}))^2] &= \tilde{E}[\sum_{j=1}^{N_n(\frac{i}{n})} \tilde{E}[(q_n^j(\frac{i+1}{n}))^2|\mathcal{F}_{\frac{i+1}{n}-}]\mathcal{F}_{\frac{i}{n}}]] \\
&\quad + \tilde{E}\left[2\sum_{1\leq j_1<j_2\leq l}^{N_n(\frac{i}{n})} \tilde{E}\left[\tilde{E}\left[q_n^{j_1}\left(\frac{i+1}{n}\right)|\mathcal{F}_{\frac{i+1}{n}-}\right]\right.\right. \\
&\quad \left.\left.\times\tilde{E}\left[q_n^{j_2}\left(\frac{i+1}{n}\right)|\mathcal{F}_{\frac{i+1}{n}-}\right]\Big|\mathcal{F}_{\frac{i}{n}}\right]\right] \\
&\leq v_{\frac{i+1}{n}} + e^{\frac{\|h\|^2}{n}}\tilde{E}[N_n(\frac{i}{n})] + e^{\frac{\|h\|^2}{n}}\tilde{E}[N_n(\frac{i}{n})(N_n(\frac{i}{n}) - 1)].
\end{aligned}$$

It follows that

$$\tilde{E}[(N_n(\frac{i+1}{n}))^2] \leq e^{\frac{\|h\|^2}{n}}\tilde{E}[(N_n(\frac{i}{n}))^2] + v_{\frac{i+1}{n}} \tag{5.5}$$

hence,

$$\begin{aligned}
\tilde{E}[(N_n(\frac{t}{n}))^2] &= \tilde{E}[(N_n(\frac{[tn]}{n}))^2] \\
&\leq e^{\|h\|^2\frac{[nt]}{n}}n^2 + \sum_{k\leq[nt]} v_{\frac{k}{n}} e^{\|h\|^2\frac{[nt]-k}{n}}
\end{aligned}$$

where the second inequality was obtain from (5.5). This completes the proof of the lemma. ■

Let φ be a continuous bounded function. Using the Lemma 5 we see that $(\mathcal{X}_n(t), \varphi)$ is square integrable and

$$\tilde{E}[(\mathcal{X}_n(\frac{i+1}{n}),\varphi)|\mathcal{F}_{\frac{i+1}{n}-}]$$

$$= \frac{1}{n}\sum_{i=1}^{N_n(\frac{i}{n})}\varphi(V_n^j(\frac{i+1}{n}))\exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(t,V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(t,V_n^j(t))dt\right) \tag{5.6}$$

and also

$$\tilde{E}[(\mathcal{X}_n(\frac{i+1}{n}),\varphi)^2|\mathcal{F}_{\frac{i+1}{n}-}]-(\tilde{E}[(\mathcal{X}_n(\frac{i+1}{n}),\varphi)|\mathcal{F}_{\frac{i+1}{n}-}])^2 = \frac{1}{n}(\mathcal{X}_n(\frac{i+1}{n}-),v_{\frac{i+1}{n}}\varphi^2). \tag{5.7}$$

In between two branches the particles move according to the prescribed SDE (5.2), hence for t in the interval $[\frac{i}{n},\frac{i+1}{n})$, such that $\varphi\in\mathcal{D}(A)$

$$(\mathcal{X}_n(t),\varphi) = (\mathcal{X}_n(\frac{i}{n}),\varphi) + \int_{\frac{i}{n}}^{t}(\mathcal{X}_n(s),A(s)\varphi)ds + S_n^{\varphi,i}(t), \tag{5.8}$$

where $\{(S_n^{\varphi,i}(t),\mathcal{F}_t),\ t\in[\frac{i}{n},\frac{i+1}{n}]\}$ is a square integrable local martingale (we use again Lemma 5) with the quadratic variation

$$< S_n^{\varphi,i} > (t) = \frac{1}{n}\int_{\frac{i}{n}}^{t}(\mathcal{X}_n(s),\|\sigma^*D\varphi\|^2)ds. \tag{5.9}$$

It follows that

$$\begin{aligned}(\mathcal{X}_n(t),\varphi) &= (\mathcal{X}_n(0),\varphi) + \int_0^t(\mathcal{X}_n(s),A(s)\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt])\\ &\quad+\sum_{i=1}^{[nt]}\left(\tilde{E}[(\mathcal{X}_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] - (\mathcal{X}_n(\frac{i}{n}-),\varphi)\right),\end{aligned} \tag{5.10}$$

where $\{(S_n^{\varphi}(t),\ \mathcal{F}_t),\ t\in[0,1]\}$ is a square integrable local martingale

$$S_n^{\varphi}(t) \stackrel{\text{def}}{=} S_n^{\varphi,[nt]}(t) + \sum_{i=0}^{[nt]-1} S_n^{\varphi,i}(\frac{i+1}{n})$$

which has the quadratic variation

$$< S_n^{\varphi} > (t) = \frac{1}{n}\int_0^t(\mathcal{X}_n(s),\|\sigma^*D\varphi\|^2)ds \tag{5.11}$$

and $\{(M_n^\varphi(l),\ \mathcal{F}_{\frac{l+1}{n}-}),\ l=0,1,...,n\}$ is a discrete martingale defined by

$$M_n^\varphi(0) \overset{\text{def}}{=} 0,$$

$$M_n^\varphi(l) \overset{\text{def}}{=} \sum_{i=1}^{l} \left((\mathcal{X}_n(\frac{i}{n}),\varphi) - \tilde{E}[(\mathcal{X}_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] \right)$$

and has conditional quadratic variation

$$\begin{aligned} < M_n^\varphi > (l) &= \sum_{i=1}^{l} \tilde{E}\left[\left((\mathcal{X}_n(\frac{i}{n}),\varphi) - \tilde{E}[(\mathcal{X}_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] \right)^2 |\mathcal{F}_{\frac{i}{n}-} \right] \\ &= \frac{1}{n}\sum_{i=1}^{l} (\mathcal{X}_n(\frac{i}{n}-), v_{\frac{i}{n}}\varphi^2) \qquad (5.12) \end{aligned}$$

Remark 15 *The process $M_n^\varphi(l)$ is a martingale also with respect to the larger filtration $\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}$.*

Proof. We have

$$\begin{aligned} &\tilde{E}\left[M_n^\varphi(l+1)\left|\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}\right.\right] - M_n^\varphi(l) \\ = \ &\tilde{E}\left[(\mathcal{X}_n(\frac{l+1}{n}),\varphi) - \tilde{E}[(\mathcal{X}_n(\frac{l+1}{n}),\varphi)|\mathcal{F}_{\frac{l+1}{n}-}]\left|\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}\right.\right] \\ = \ &\tilde{E}\left[(\mathcal{X}_n(\frac{l+1}{n}),\varphi)\left|\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}\right.\right] - \tilde{E}[(\mathcal{X}_n(\frac{l+1}{n}),\varphi)|\mathcal{F}_{\frac{l+1}{n}-}] \end{aligned}$$

But $(\mathcal{X}_n(\frac{l+1}{n}),\varphi)$ is independent of the 'future' observations, hence

$$\tilde{E}\left[(\mathcal{X}_n(\frac{l+1}{n}),\varphi)\left|\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}\right.\right] = \tilde{E}\left[(\mathcal{X}_n(\frac{l+1}{n}),\varphi)\left|\mathcal{F}_{\frac{l+1}{n}-} \vee \mathcal{Y}_{\frac{l+1}{n}}\right.\right]$$

and since $\mathcal{Y}_t$ is a right continuous filtration $\mathcal{Y}_{\frac{l+1}{n}} = \mathcal{Y}_{\frac{l+1}{n}-} \subset \mathcal{F}_{\frac{l+1}{n}-}$ hence the above difference is 0.

■

Using (5.6) and (5.10), we can express the process $(\mathcal{X}_n(t),\varphi)$ as

$$(\mathcal{X}_n(t),\varphi) = (\mathcal{X}_n(0),\varphi) + \int_0^t (\mathcal{X}_n(s), A(s)\varphi)ds + S_n^\varphi(t) + M_n^\varphi([nt])$$

$$+\sum_{i=1}^{[nt]} \frac{1}{n} \sum_{j=1}^{N_n(\frac{i-1}{n})} \varphi(V_n^j(\frac{i}{n})) \left(\exp\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(s, V_n^j(s))dY_s - \frac{1}{2} \int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(s, V_n^j(s))ds\right) - 1 \right). \tag{5.13}$$

Then applying Itô's rule to the exponential in the last term of (5.13) and exploiting the fact that Y is a Brownian motion, we get

$$\begin{aligned} (\mathcal{X}_n(t), \varphi) &= (\mathcal{X}_n(0), \varphi) + \int_0^t (\mathcal{X}_n(s), A(s)\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt]) \\ &\quad + \frac{1}{n} \int_0^{\frac{[nt]}{n}} \sum_{j=1}^{N_n(\frac{[sn]}{n})} \varphi(V_n^j(\frac{[sn]+1}{n})) B_n^s(V_n^j, s) h^*(s, V_n^j(s)) dY_s \end{aligned} \tag{5.14}$$

where

$$B_n^s(V_n^j, p) = \exp\left(\int_{\frac{[sn]}{n}}^{p} h^*(r, V_n^j(r))dY_r - \frac{1}{2} \int_{\frac{[sn]}{n}}^{p} h^*h(r, V_n^j(r))dr \right) \tag{5.15}$$

5.1.2 The Existence of the Process

We show first that the sequence $\{\mathcal{X}_n\}_{n>0}$ is tight. For this we need to prove that $\{\mathcal{X}_n\}_{n>0}$ satisfies conditions **a** and **b** of Theorem 21. Condition **a** follows from the following proposition.

Proposition 9 *For every $t \in [0,1]$ we have*

$$\lim_{k\to\infty} \sup_{n\geq 0} \tilde{P}(\sup_{0\leq s\leq t} (\mathcal{X}_n(s), 1) > k) = 0. \tag{5.16}$$

Proof. Since

$$\tilde{P}(\sup_{0\leq s\leq t} (\mathcal{X}_n(s), 1) > k) \leq \frac{\tilde{E}[(\sup_{0\leq s\leq t}(\mathcal{X}_n(s), 1))^2]}{k^2}, \tag{5.17}$$

it is enough to prove that $\sup_{n\geq 0} \tilde{E}[(\sup_{0\leq s\leq t}(\mathcal{X}_n(s), 1))^2]$ is finite. Let us denote by

$$\psi_n(t) \overset{\text{def}}{=} \tilde{E}\left[\left(\sup_{0\leq s\leq t} (\mathcal{X}_n(s), 1) \right)^2 \right].$$

From (5.14) we obtain

$$\psi_n(t) \leq 3(\mathcal{X}_n(0), 1)^2 + 3\tilde{E}[(\sup_{0\leq i\leq [nt]} |M_n^1(i)|)^2]$$

$$+\frac{3}{n^2}\tilde{E}\left[\left(\sup_{0\le p\le \frac{\lfloor tn\rfloor}{n}}\left|\int_0^p \sum_{j=1}^{N_n(\frac{\lfloor sn\rfloor}{n})} B_n^s(V_n^j,s)h^*(s,V_n^j(s))dY_s\right|\right)^2\right] \tag{5.18}$$

We prove that ψ_n is bounded from above uniformly in n, by exploiting (5.18) and using Gronwall's inequality. For this we give an upper bound for each of the three terms of the right hand side of the inequality (5.18) of the form $\alpha+\beta\int_0^t\psi_n(s)ds$.

The first term

We have

$$(\mathcal{X}_n(0),1)^2=\frac{N_n(0)^2}{n^2}=1. \tag{5.19}$$

The second term

Doob's maximal inequality (cf [25], pp. 14) gives us the following upper bound:

$$\begin{aligned}\tilde{E}[(\sup_{0\le i\le [nt]}|M_n^1(i)|)^2] &\le 4\tilde{E}[(M_n^1([nt]))^2]=4\tilde{E}[(<M_n^1>([nt]))] \\ &\le \frac{4\|v\|}{n}\sum_{i=1}^{[nt]}\tilde{E}[(\mathcal{X}(\frac{i}{n}-),1)]\le 4\|v\|\frac{[nt]}{n}\le 4\|v\|. \end{aligned} \tag{5.20}$$

The third term

We find first an upper bound for $\tilde{E}[(\sum_{j=1}^{N_n(\frac{\lfloor sn\rfloor}{n})}B_n^s(V_n^j,p))^2]$. We have that

$$\begin{aligned}&\tilde{E}\left[\left(\sum_{j=1}^{N_n(\frac{\lfloor sn\rfloor}{n})}B_n^s(V_n^j,p)\right)^2\right]=\tilde{E}\left[\sum_{j_1,j_2=1}^{N_n(\frac{\lfloor sn\rfloor}{n})}\tilde{E}\left[B_n^s(V_n^{j_1},p)B_n^s(V_n^{j_2},p)\middle|\mathcal{F}_{\frac{\lfloor sn\rfloor}{n}}\right]\right] \\ &\le \tilde{E}\left[\sum_{j_1,j_2=1}^{N_n(\frac{\lfloor sn\rfloor}{n})}e^{\|h\|^2(p-\frac{\lfloor sn\rfloor}{n})}\tilde{E}\left[\exp\left(\int_{\frac{\lfloor sn\rfloor}{n}}^p(h^*(q,V_n^{j_1}(q))+h^*(q,V_n^{j_1}(q)))dY_q\right.\right.\right. \\ &\qquad\left.\left.\left.-\frac{1}{2}\int_{\frac{\lfloor sn\rfloor}{n}}^p\left|h(q,V_n^{j_1}(q))+h(q,V_n^{j_1}(q))\right|^2\right)\middle|\mathcal{F}_{\frac{\lfloor sn\rfloor}{n}}\right]\right]\end{aligned}$$

which gives us as in Lemma 5

$$\tilde{E}[(\sum_{j=1}^{N_n(\frac{\lfloor sn\rfloor}{n})}B_n^s(V_n^j,p))^2]\le e^{2\|h\|^2(p-\frac{\lfloor sn\rfloor}{n})}\tilde{E}[(N_n(\frac{[sn]}{n}))^2] \tag{5.21}$$

$$\tilde{E}[(\sum_{j=1}^{N_n(\frac{[sn]}{n})} B_n^s(V_n^j, s))^2] \quad \leq \quad e^{2\|h\|^2}\tilde{E}[(N_n(\frac{[sn]}{n}))^2] \tag{5.22}$$

Finally, using Doob's maximal inequality (see the Appendix) and (5.22), we find

$$\begin{aligned} \tilde{E}\left[\left(\sup_{0\leq p\leq \frac{[tn]}{n}}\left|\int_0^p \sum_{j=1}^{N_n(\frac{[sn]}{n})} B_n^s(V_n^j, s)h^*(s, V_n^j(s))dY_s\right|\right)^2\right] & \\ \leq \int_0^{\frac{[tn]}{n}} \tilde{E}\left[\left|\sum_{j=1}^{N_n(\frac{[sn]}{n})} B_n^s(V_n^j, s)h(s, V_n^j(s))\right|^2\right] ds & \\ \leq 4e^{2\|h\|^2}\|h\|^2 \int_0^t \tilde{E}[(N_n(\frac{[sn]}{n}))^2]ds & \end{aligned}$$

The last inequality gives the following upper bound on the third term of (5.18)

$$4e^{2\|h\|^2}\|h\|^2 \int_0^t \psi_n(s)ds \tag{5.23}$$

From (5.18), (5.19), (5.20) and (5.23) we obtain

$$\psi_n(t) \leq (3 + 12\|v\|) + 12e^{2\|h\|^2}\|h\|^2 \int_0^t \psi_n(s)ds$$

Using once again *Gronwall's inequality* we find that $\psi_n(t) \leq c(t)$, where

$$c(t) \overset{\text{def}}{=} (3 + 12\|v\|)e^{\frac{4\|h\|^2 e^{2\|h\|^2}}{1+4\|v\|}t}, \; t \in [0, 1].$$

So also $\sup_{n\geq 1} \tilde{E}[(\sup_{0\leq s\leq t}(\mathcal{X}_n(s), 1))^2] \leq c(t)$ which finishes the proof of the proposition.

■

Remark 16 *Using a similar argument one can prove that,* $\forall p \geq 1$, *there exists a function* $c_p : [0, 1] \longrightarrow \mathbb{R}_+$, *such that*

$$\sup_{n\geq 1} \tilde{E}[(\sup_{0\leq s\leq t} (\mathcal{X}_n(s), 1))^p] \leq c_p(t), \; t \in [0, 1]. \tag{5.24}$$

Condition b follows from the following

Proposition 10 *For any arbitrary sequence of stopping times* $\{\tau_n\}_{n\geq 0}$ *any real positive sequence* $\{\delta_n\}_{n\geq 0}$ *with* $\lim_{n\to\infty}\delta_n = 0$ *and* $\varphi \in \mathcal{D}(A)\cap C_K(R^d)$ *we have*

$$\lim_{n\to\infty}\tilde{E}[|(\mathcal{X}_n(\tau_n+\delta_n),\varphi)-(\mathcal{X}_n(\tau_n),\varphi)|^2]=0 \tag{5.25}$$

and hence

$$\lim_{n\to\infty}\tilde{P}(|(\mathcal{X}_n(\tau_n+\delta_n),\varphi)-(\mathcal{X}_n(\tau_n),\varphi)|\geq\epsilon)=0 \tag{5.26}$$

Proof. Let a and b be the following quantities

$$a \triangleq \sup_{\{(t,x)\in[0,1]\times\mathbb{R}^d\}} \|A(t)\varphi\| < \infty$$

$$b \triangleq \sup_{\{(t,x)\in[0,1]\times\mathbb{R}^d\}} \|\sigma(t,x)D\varphi\| < \infty.$$

Obviously, if φ is the constant function1,the $a=b=0$. Using (5.14) we see

$$\begin{aligned}\tilde{E}[|(\mathcal{X}_n(\tau_n+\delta_n),\varphi)-(\mathcal{X}_n(\tau_n),\varphi)|^2] &\leq 4\tilde{E}[(\int_{\tau_n}^{\tau_n+\delta_n}(\mathcal{X}_n(s),A(s)\varphi)ds)^2]\\ &+4\tilde{E}[(S_n^\varphi(\tau_n+\delta_n)-S_n^\varphi(\tau_n))^2]+4\tilde{E}[(M_n^\varphi([n(\tau_n+\delta_n)])-M_n^\varphi([n\tau_n]))^2]\\ &+4\tilde{E}[(\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}}\sum_{j=1}^{N_n(\frac{[sn]}{n})}\varphi(V_n^j(\frac{[sn]+1}{n}))B_n^s(V_n^j,s)h^*(V_n^j(s))dY_s)^2]\end{aligned} \tag{5.27}$$

We have, consecutively,

$$\begin{aligned}\tilde{E}[(\int_{\tau_n}^{\tau_n+\delta_n}(\mathcal{X}_n(s),A(s)\varphi)ds)^2] &\leq \delta_n\tilde{E}[\int_{\tau_n}^{\tau_n+\delta_n}(\mathcal{X}_n(s),A(s)\varphi)^2ds]\\ &\leq \delta_n^2a^2\tilde{E}[(\sup_{0\leq s\leq 1}(\mathcal{X}_n(s),1))^2((\tau_n+\delta_n)-\tau_n)]\\ &\leq \delta_n^2a^2c(1)\end{aligned} \tag{5.28}$$

$$\begin{aligned}\tilde{E}[(S_n^\varphi(\tau_n+\delta_n)-S_n^\varphi(\tau_n))^2] &\leq K\tilde{E}[<S_n^\varphi>(\tau_n+\delta_n)-<S_n^\varphi>(\tau_n)]\\ &\leq \frac{K}{n}\tilde{E}[\int_{\tau_n}^{\tau_n+\delta_n}(\mathcal{X}_n(s),\|\sigma(t)D\varphi)\|^2]ds\\ &\leq \frac{Kb^2}{n}\tilde{E}[\sup_{0\leq s\leq 1}(\mathcal{X}_n(s),1)((\tau_n+\delta_n)-\tau_n)]\\ &\leq \frac{Kb^2}{2n}(1+c(1))\delta_n\end{aligned} \tag{5.29}$$

$$\begin{aligned}\tilde{E}[(M_n^{\varphi}([n(\tau_n+\delta_n)])-M_n^{\varphi}([n\tau_n]))^2] &= \tilde{E}[(M_n^{\varphi}([n(\tau_n+\delta_n)])-M_n^{\varphi}([n\tau_n]))^2] \\ &= \frac{1}{n}\tilde{E}[\sum_{[n\tau_n]+1}^{[n(\tau_n+\delta_n)]}(\mathcal{X}_n(\frac{i}{n}-),v_{\frac{i}{n}}\varphi^2)] \\ &\leq \frac{\|v\|\|\varphi\|^2}{n}\tilde{E}[\sup_{0\leq s\leq 1}(\mathcal{X}_n(s),1)([n(\tau_n+\delta_n)]-[n\tau_n]-1)] \\ &\leq \|v\|\|\varphi\|^2\frac{1+c(1)}{2}\delta_n \qquad (5.30)\end{aligned}$$

$$\tilde{E}[(\frac{1}{n^2}\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n)]}{n}}\sum_{j=1}^{N_n(\frac{[sn]}{n})}\varphi(V_n^j(\frac{[sn]+1}{n}))B_n^s(V_n^j,s)h^*(s,V_n^j(s))dY_s)^2]$$

$$= \tilde{E}[\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n)]}{n}}\frac{1}{n^2}(\sum_{j=1}^{N_n(\frac{[sn]}{n})}\varphi(V_n^j(\frac{[sn]+1}{n}))B_n^s(V_n^j,s)h^*(s,V_n^j(s)))^2ds]$$

$$\leq \|h\|^2\|\varphi\|^2\tilde{E}[\frac{(\sup_{0\leq s\leq 1}\sum_{j=1}^{N_n(\frac{[sn]}{n})}B_n^s(V_n^j,s))^2}{n^2}(\frac{[n(\tau_n+\delta_n]}{n}-\frac{[n\tau_n]}{n})]$$

$$\leq \|\varphi\|^2\|h\|^2c'(1)(\delta_n+\frac{1}{n}) \qquad (5.31)$$

where $c'(1)$ is obtained in a similar way to $c(1)$ as an uniform upper bound for

$$\tilde{E}[\frac{(\sup_{0\leq s\leq 1}\sum_{j=1}^{N_n(\frac{[sn]}{n})}B_n^s(V_n^j,s))^2}{n^2}]$$

The inequalities (5.28), (5.29), (5.30), (5.31) imply that all the terms from the right hand side of (5.27) tend to 0 as n tends to ∞, hence $\tilde{E}[|(\mathcal{X}_n(\tau_n+\delta_n),\varphi)-(\mathcal{X}_n(\tau_n),\varphi)|^2]$ tends to 0 as well, which completes the proof of the proposition.

■

We know now that the sequence $\mathcal{X}_n$ is tight, hence relatively compact. Then $(\mathcal{X}_n,Y)$ is relatively compact. Let $(\mathcal{X},Y)$ be the limit process of one of its convergent subsequences (to avoid even more cumbersome notation we reindex this sequence as $\{(\mathcal{X}_n,Y)\}_{n\geq 0}$). We will show that $\mathcal{X}$ is a solution of the martingale problem (1.3)+(1.4), i.e., for every φ in the domain of A, the process

$$M^{\varphi}(t)\stackrel{\text{def}}{=}(\mathcal{X}(t),\varphi)-(\mathcal{X}(0),\varphi)-\int_0^t(\mathcal{X}(s),A\varphi)ds-\int_0^t(\mathcal{X}(s),h^*(s)\varphi)dY_s$$

is a square integrable martingale with respect to the filtration $\mathcal{F}_t \vee \mathcal{Y}$ with quadratic variation

$$< M^{\varphi}(t) >= \int_0^t (\mathcal{X}_s, v_s\varphi^2)ds, \qquad P-a.s.$$

We need first several preliminary results.

Proposition 11 *For each $p \geq 1$, $t \in [0,1]$ and bounded continuous function φ we have*

$$\tilde{E}[|(\mathcal{X}(t),\varphi)|^p] < \infty \tag{5.32}$$

Proof. Straightforward from (5.24).

■

Proposition 12 *Let φ be a continuous bounded function. Then, for all $p \geq 1$*

$$\lim_{n\to\infty} \tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p] = \tilde{E}[|(\mathcal{X}(t),\varphi)|^p]. \tag{5.33}$$

Proof. We prove that for all $\epsilon > 0$, there exists n_ϵ such that for every $n \geq n_\epsilon$ we have

$$|\tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p] - \tilde{E}[|(\mathcal{X}(t),\varphi)|^p]| \leq \epsilon. \tag{5.34}$$

Since $\lim_{k\to\infty} \tilde{E}[|(\mathcal{X}(t),\varphi)|^p \wedge k] = \tilde{E}[|(\mathcal{X}(t),\varphi)|^p]$ there exists k_1 such that for every $k \geq k_1$

$$|\tilde{E}[|(\mathcal{X}(t),\varphi)|^p \wedge k] - \tilde{E}[|(\mathcal{X}(t),\varphi)|^p]| \leq \frac{\epsilon}{3}. \tag{5.35}$$

Also

$$\begin{aligned} |\tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p \wedge k] - \tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p]| &\leq \tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p \mathrm{I}_{|(\mathcal{X}_n(t),\varphi)|>k^{\frac{1}{p}}}] \\ &\leq \sqrt{\tilde{E}[|(\mathcal{X}_n(t),\varphi)|^{2p}]P(|(\mathcal{X}_n(t),\varphi)| > k^{\frac{1}{p}})} \\ &\leq \frac{\|\varphi\|^{p+1}}{k^{\frac{1}{p}}}\sqrt{c_{2p}(1)c_2(1)} \end{aligned}$$

Thus we can choose k_2 such that for every $k \geq k_2$

$$|\tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p \wedge k] - \tilde{E}[|(\mathcal{X}_n(t),\varphi)|^p]| \leq \frac{\epsilon}{3}. \tag{5.36}$$

Let now $k = \max(k_1, k_2)$ and since $\lim_{n\to\infty} \tilde{E}[|(\mathcal{X}_n(t), \varphi\ )|^p \wedge k] = \tilde{E}[|(\mathcal{X}(t), \varphi)|^p \wedge k]$, there exists n_ϵ such that for every $n \geq n_\epsilon$

$$|\tilde{E}[|(\mathcal{X}_n(t), \varphi)|^p \wedge k] - \tilde{E}[|(\mathcal{X}(t), \varphi)|^p \wedge k]| \leq \frac{\epsilon}{3}. \tag{5.37}$$

From (5.35), (5.36) and (5.37) we obtain (5.34) and with it our claim.

■

Proposition 13 *Let φ be a bounded continuous function such that $\varphi, \varphi^2 \in \mathcal{D}(A)$. Then*

$$\lim_{n\to\infty} \tilde{E}\left[\left(\frac{1}{n}\int_{\frac{[ns]}{n}}^{\frac{[nt]}{n}} \sum_{j=1}^{N_n(\frac{[rn]}{n})} \varphi(V_n^j(\frac{[rn]+1}{n}))B_n^r(V_n^j, r)h^*(r, V_n^j(r))dY_r \right.\right.$$
$$\left.\left. - \int_s^t (\mathcal{X}_n(r), h^*(r)\varphi)dY_r\right)^2\right] = 0 \tag{5.38}$$

Proof. Firstly we observe that the last integral can be taken from $\frac{[ns]}{n}$ to $\frac{[nt]}{n}$ without changing the limit. Then, using (5.21), we get

$$\tilde{E}[(\frac{1}{n}\int_{\frac{[ns]}{n}}^{\frac{[nt]}{n}} \sum_{j=1}^{N_n(\frac{[rn]}{n})} \varphi(V_n^j(\frac{[rn]+1}{n}))(B_n^r(V_n^j, r) - 1)h^*(r, V_n^j(r))dY_r)^2]$$

$$= \frac{1}{n^2}\int_{\frac{[nr]}{n}}^{\frac{[nt]}{n}} \tilde{E}[|\sum_{j=1}^{N_n(\frac{[rn]}{n})} \varphi(V_n^j(\frac{[rn]+1}{n}))(B_n^r(V_n^j, r) - 1)h^*(r, V_n^j(r))|^2 dr]$$

$$\leq \|h\|^2\|\varphi\|^2 \int_{\frac{[ns]}{n}}^{\frac{[nt]}{n}} \tilde{E}[\frac{1}{n^2}(\int_{\frac{[rn]}{n}}^{r} \sum_{j=1}^{N_n(\frac{[rn]}{n})} (B_n^r(V_n^j, p)h^*(s, V_n^j(p))dY_p)^2]dr$$

$$\leq \frac{1}{2}\|h\|^2\|\varphi\|^2 c(1)\frac{[nt]-[ns]}{n}(e^{\frac{1}{n}} - 1). \tag{5.39}$$

Thus one can eliminate $B_n^r(V_n^j, r)$ from the first term of (5.38) without changing the limit. After these 2 transformations, (5.38) becomes

$$\lim_{n\to\infty} \tilde{E}\left[\left(\frac{1}{n}\int_{\frac{[ns]}{n}}^{\frac{[nt]}{n}} \sum_{j=1}^{N_n(\frac{[rn]}{n})} (\varphi(V_n^j(\frac{[rn]+1}{n})) - \varphi(V_n^j(r)))h^*(r, V_n^j(r))dY_r\right)^2\right]$$

Using once again Doob's inequality, we find the following upper bound for the sequence

$$2\|h\|^2\frac{[nt]-[ns]}{n}c(1)(\frac{\|A\varphi\|^2}{n^2}+\frac{C\|A\varphi^2-2\varphi A\varphi\|}{n})$$

which completes our proof.

■

We are now able to proof that $\mathcal{X}$ satisfies the desired martingale problem.

Theorem 24 *For $\varphi \in \mathcal{D}(A)$ the process $\{(M^\varphi(t), \mathcal{F}_t \vee \mathcal{Y}), t \in [0,1]\}$ where*

$$M^\varphi(t) \overset{\text{def}}{=} (\mathcal{X}(t),\varphi) - (\mathcal{X}(0),\varphi) - \int_0^t (\mathcal{X}(s), A(s)\varphi)ds - \int_0^t (\mathcal{X}(s), h^*(s)\varphi)dY_s$$

is a square integrable martingale with the quadratic variation

$$< M^\varphi > (t) = \int_0^t (\mathcal{X}(s), v_s^2\varphi)ds$$

Proof. We will use Theorem 18. We want to prove that for all $\varphi \in \mathcal{D}(A)$

$$\tilde{E}[(M^\varphi(t) - M^\varphi(s))\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))] = 0 \qquad (5.40)$$

and

$$\tilde{E}[((M^\varphi(t) - M^\varphi(s))^2 - \int_s^t (\mathcal{X}(r), v_r^2\varphi)dr)\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j'))] = 0 \qquad (5.41)$$

for all $m, m' \geq 0$, $0 \leq t_1 < t_2 < \ldots < t_m \leq s \leq t$, $0 \leq t_1' < t_2' < \ldots < t_m' \leq 1$, $k_1, \ldots, k_m \in C_b(\mathbb{R}^d)$ and $k_1', \ldots, k_m' \in C_b(\mathbb{R}^d)$. We prove only (5.40), since (5.41) can be done analogously. From the definition of M^φ, (5.40) is equivalent to

$$\tilde{E}[((\mathcal{X}(t),\varphi) - (\mathcal{X}(s),\varphi) - \int_s^t (\mathcal{X}(r), A(r)\varphi)ds - \int_s^t (\mathcal{X}(r), h^*(r)\varphi)dY_r)$$

$$\times \Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j'))] = 0 \qquad (5.42)$$

We only need to show (5.42) for φ with the property that $\varphi^2 \in \mathcal{D}(A)$ since using the assumption **I** and the dominated convergence theorem we can extend this to an arbitrary $\varphi \in \mathcal{D}(A)$. Using a proof analogous to the one used in Proposition 12 one shows, consecutively, that since $(\mathcal{X}_n, Y)$ converges in distribution to $(\mathcal{X}, Y)$

$$\lim_{n\to\infty} \tilde{E}[(\mathcal{X}_n(t),\varphi)\Pi_{i=1}^m k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))]$$

$$= \tilde{E}[(\mathcal{X}(t),\varphi)\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))], \qquad (5.43)$$

$$\lim_{n\to\infty} \tilde{E}[(\mathcal{X}_n(s),\varphi)\Pi_{i=1}^m k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))]$$

$$= \tilde{E}[(\mathcal{X}(s),\varphi)\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))], \qquad (5.44)$$

and

$$\lim_{n\to\infty} \tilde{E}[\int_s^t (\mathcal{X}_n(r), A(r)\varphi)dr\Pi_{i=1}^m k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))]$$

$$= \tilde{E}[\int_s^t (\mathcal{X}(r), A(r)\varphi)dr\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))]. \quad (5.45)$$

Using Theorem 23, we have that, since $(\mathcal{X}_n, Y)$ converges in distribution to $(\mathcal{X}, Y)$ also $(\mathcal{X}_n, Y, \int_0^t(\mathcal{X}_n(s), h^*(s)\varphi)dY_s)$ converges in distribution to $(\mathcal{X}, Y, \int_0^t(\mathcal{X}(s), h^*(s)\varphi)dY_s)$ and using (5.38) and, once again, an argument similar to the one used in Proposition 12, we have that

$$\lim_{n\to\infty} \tilde{E}\left[\frac{1}{n}\int_{\frac{[ns]}{n}}^{\frac{[nt]}{n}} \sum_{j=1}^{N_n(\frac{[rn]}{n})} \varphi(V_n^j(\frac{[rn]+1}{n}))B_n^r(V_n^j, r)h^*(r, V_n^j(r))dY_r \right.$$

$$\left. \times\Pi_{i=1}^m k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))\right]$$

$$= \tilde{E}\left[\int_0^t (\mathcal{X}(s), h^*(s)\varphi)dY_s\Pi_{i=1}^m k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))\right]. \qquad (5.46)$$

Since $\varphi^2 \in \mathcal{D}(A)$, we have that $\|\sigma^* D(\varphi)\|^2 = A\varphi^2 - 2\varphi A\varphi$ is a bounded function and hence S_n^φ is a square integrable martingale such that

$$E[(S_n^\varphi)^2(p)] = E[< S_n^\varphi >^2 (p)] \leq \frac{\|A\varphi^2 - 2\varphi A\varphi\|}{n}$$

and hence

$$\lim_{n\to\infty} \tilde{E}[(S_n^\varphi(t) - S_n^\varphi(s))\Pi_{i=1}^m k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))] = 0. \qquad (5.47)$$

From (5.43),(5.44),(5.45),(5.46) and (5.47) we obtain that

$$\begin{aligned}\tilde{E}[(M^{\varphi}(t) - M^{\varphi}(s))\Pi_{i=1}^{m} k_i(X(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))] \\ &= \lim_{n\to\infty} \tilde{E}[(M_n^{\varphi}(t) - M_n^{\varphi}(s))\Pi_{i=1}^{m} k_i(X_n(t_i))\Pi_{j=1}^{m'} k_i'(Y(t_j))] \\ &= 0. \qquad (5.48)\end{aligned}$$

■

Remark 17 *The martingale M^{φ} is also a martingale with respect to the initial filtration $\mathcal{F}_t$ and its conditional expectation with respect to $\mathcal{Y}$ is* 0.

With this we conclude the existence of the process with the properties described in the introduction. One can also prove that the martingale problem (1.3)+(1.4) has a unique solution. Since it is not central to our objective we will not do it here.

Proposition 14 *The sequence $\{\mathcal{X}_n\}_{n\geq 0}$ is convergent to the unique solution of the martingale problem (1.3)+ (1.4).*

Proof.

From the tightness of the sequence, we have that every subsequence contains a weakly convergent subsequence to $\mathcal{X}$, the unique solution of the martingale problem (1.3)+ (1.4), hence the whole sequence is convergent to $\mathcal{X}$.

■

5.1.3 Application to the Nonlinear Filtering

The process $\mathcal{X}$ is the solution of the desired 'filtered' martingale problem. Hence for $\varphi \in \mathcal{D}(A)$

$$\tilde{E}[(\mathcal{X}(t),\varphi)|\mathcal{Y}_t] = \tilde{E}[(\mathcal{X}(0),\varphi)|\mathcal{Y}_0] + \int_0^t \tilde{E}[(\mathcal{X}(s),A\varphi)|\mathcal{Y}_s]ds.$$

$$+ \int_0^t \tilde{E}[(\mathcal{X}(s),h^*\varphi)|\mathcal{Y}_s]dY_s \qquad (5.49)$$

In establishing (5.49), we used the fact that for every integrable $\mathcal{F}_t$-measurable random variable $\mathcal{A}$ we have $\tilde{E}[\mathcal{A}|\mathcal{Y}] = \tilde{E}[\mathcal{A}|\mathcal{Y}_t]$ (Proposition 3) and if $\{U_t; t \geq 0\}$ is a càdlàg, $\mathcal{F}_t$-adapted process such that $\tilde{E}\int_0^t U_s^2 dt < \infty, \forall s \geq 0$, then $\tilde{E}[\int_0^t U_s dY_s|\mathcal{Y}_t] = \int_0^t \tilde{E}[U_s|\mathcal{Y}_s]dY_s$

(Proposition 2) and $\tilde{E}[\int_0^t U_s ds|\mathcal{Y}_t] = \int_0^t \tilde{E}[U_s|\mathcal{Y}_s]ds$ (Fubini's theorem). One can also obtain the corresponding evolution equation for time dependent φ.

Let $\mathcal{X}_n^{Y(\omega)}$ and $\mathcal{X}^{Y(\omega)}$ be the processes $\mathcal{X}_n$ and, respectively, $\mathcal{X}$ given the observation path $Y(\omega)$. Let also $\tilde{E}_\omega$ be the corresponding expectations given $Y(\omega)$, $\mathcal{Z}_n^\omega(t) = \tilde{E}_\omega[\mathcal{X}_n^{Y(\omega)}(t)]$, i.e., the measure obtained by integrating the measure valued random variable $\mathcal{X}_n^{Y(\omega)}(t)$ (this is, actually, what we are computing in the computer algorithm) and $\mathcal{Z}^\omega(t) = \tilde{E}_\omega[\mathcal{X}^{Y(\omega)}(t)]$. Using Fubini's theorem, we have

$$(\mathcal{Z}^\omega(t), \varphi) = \tilde{E}_\omega[(\mathcal{X}^{Y(\omega)}(t), \varphi)] = \tilde{E}[(\mathcal{X}(t), \varphi)|\mathcal{Y}](\omega) \tag{5.50}$$

$$(\mathcal{Z}_n^\omega(t), \varphi) = \tilde{E}_\omega[(\mathcal{X}_n^{Y(\omega)}(t), \varphi)] = \tilde{E}[(\mathcal{X}_n(t), \varphi)|\mathcal{Y}](\omega) \tag{5.51}$$

Using (5.50), the evolution equation (5.49) becomes

$$(\mathcal{Z}(t), \varphi) = (\mathcal{Z}(0), \varphi) + \int_0^t (\mathcal{Z}(s), A\varphi)ds + \int_0^t (\mathcal{Z}(t), h^*\varphi)dY_s \tag{5.52}$$

From (5.52) and the fact that we assumed from the beginning that the solution of the Żakai equation is unique, we deduce the following

Theorem 25 *The unnormalised conditional distribution of the signal X given the observation Y coincides with the conditional expectation of $\mathcal{X}$ given the observation Y.*

The next theorem is the cornerstone of the numerical algorithm. It shows that, in order to approximate the unnormalised conditional distribution ρ_t, we can construct the process $\mathcal{X}_n$ up to time t (where n is taken so that the error is as small as we want), keeping the observation path fixed, and then compute its (conditional) expectation.

Theorem 26 *There exists $\bar{\Omega} \in \Omega$ with $\tilde{P}(\bar{\Omega}) = 1$ such that for every $\omega \in \bar{\Omega}$ we have $\lim_{n\to\infty} \mathcal{Z}_n^\omega(t) = \rho_t^{Y(\omega)}$, i.e.,*

$$\lim_{n\to\infty} (\mathcal{Z}_n^\omega(t), \varphi) = \rho_t^{Y(\omega)}(\varphi) \tag{5.53}$$

for every continuous bounded function φ ($\rho_t^{Y(\omega)}$ is the unnormalised distribution of the signal given the observation path $Y(\omega)$).

Proof. Let M be a set containing a countable collection of $C_0^\infty(\mathbb{R}^d)$ uniformly dense in $C_0^\infty(\mathbb{R}^d)$ and the constant function 1. To prove the theorem we only need to show that, for every function in M

$$\lim_{n\to\infty}(\mathcal{Z}_n^t(\omega),\varphi) = \rho_t^{Y(\omega)}(\varphi) \qquad \tilde{P}-a.s. \tag{5.54}$$

(to simplify the notation we will omit the ω variable from now on). For this we use the solution of the following backward Itô equation

$$\begin{aligned} d\psi_s(x) &= -A(s)\psi_s(x) - h^*(s,x)\psi_s(x)dY_s \\ \psi_t(x) &= \varphi(x) \end{aligned} \tag{5.55}$$

From [3], pp. 126-134 or [36], we obtain that equation (5.55) has a unique solution and $\rho_t(\psi_t) = \rho_0(\psi_0)$. Since ψ_0 is continuous and bounded $\tilde{P}-a.s.$, it follows that $\lim_{n\to\infty}(\mathcal{Z}_n(0),\psi_0) = (\pi_0,\psi_0) = \rho_0(\psi_0)$, $\tilde{P}$-a.s.. Hence, in order to show (5.54), we need to prove that

$$\lim_{n\to\infty}(\mathcal{Z}_n(t),\psi_t) - (\mathcal{Z}_n(0),\psi_0) = 0 \qquad \tilde{P}-a.s. \tag{5.56}$$

The first step is to prove that

$$(\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]}) = (\mathcal{Z}_n(0),\psi_0) \qquad \tilde{P}-a.s. \tag{5.57}$$

and then that

$$\lim_{n\to\infty}(\mathcal{Z}_n(t),\psi_t) - (\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]}) = 0 \qquad \tilde{P}-a.s.. \tag{5.58}$$

We have that

$$(\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]}) - (\mathcal{Z}_n(0),\psi_0) = \sum_{i=1}^{[\frac{nt}{n}]} \tilde{E}[(\mathcal{X}_n(\frac{i}{n}),\psi_{\frac{i}{n}})|\mathcal{Y}] - \tilde{E}[(\mathcal{X}_n(\frac{i-1}{n}),\psi_{\frac{i-1}{n}})|\mathcal{Y}]$$

and

$$\begin{aligned} &\tilde{E}[(\mathcal{X}_n(\frac{i}{n}),\psi_{\frac{i}{n}})|\mathcal{Y}] - \tilde{E}[(\mathcal{X}_n(\frac{i-1}{n}),\psi_{\frac{i-1}{n}})|\mathcal{Y}] \\ &= \tilde{E}[\sum_{j=1}^{N_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))q_n^j(\frac{i}{n}) - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y}] \end{aligned}$$

Since the number of offspring $q_n^j(\frac{i}{n})$ of the particle V_n^j is independent of the 'future' of $Y_{\frac{i}{n}}$, we have that

$$\begin{aligned}\tilde{E}[q_n^j(\frac{i}{n})|\mathcal{F}_{\frac{i}{n}} \vee \mathcal{Y}] &= \tilde{E}[q_n^j(\frac{i}{n})|\mathcal{F}_{\frac{i}{n}} \vee \mathcal{Y}_{\frac{i}{n}}] \\ &= \tilde{E}[q_n^j(\frac{i}{n})|\mathcal{F}_{\frac{i}{n}}] \\ &= e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds}\end{aligned}$$

Hence

$$\begin{aligned}&\tilde{E}[\sum_{j=1}^{N_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))q_n^j(\frac{i}{n}) - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y}] \\ &= \tilde{E}[\sum_{j=1}^{N_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))\tilde{E}[q_n^j(\frac{i}{n})|\mathcal{F}_{\frac{i}{n}} \vee \mathcal{Y}] - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y}] \\ &= \tilde{E}[\sum_{j=1}^{N_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds} - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y}] \\ &= \tilde{E}[\sum_{j=1}^{N_n(\frac{i-1}{n})} \tilde{E}[\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds} \\ &\quad -\psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y} \vee \mathcal{F}_{\frac{i}{n}}]|\mathcal{Y}] \qquad (5.59)\end{aligned}$$

We prove that

$$\tilde{E}\left[\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds} - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}\right] = 0 \qquad (5.60)$$

Since V_n^j is a Markov process, we have that

$$\begin{aligned}&\tilde{E}\left[\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}\right] \\ &= \tilde{E}\left[\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \sigma(V_n^j(\frac{i-1}{n}))\right] \qquad (5.61)\end{aligned}$$

We compute first

$$R(x) = \tilde{E}_{\frac{i-1}{n},x}\left[\varphi(V_n^j(t))e^{\int_{\frac{i-1}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{t} h^*h(V_n^j(s))ds}|\mathcal{Y}\right]$$

where the expectation $\tilde{E}_{\frac{i-1}{n},x}$ is taken with respect to the probability $\tilde{P}_{\frac{i-1}{n},x}$ and $\tilde{P}_{\frac{i-1}{n},x}$ is taken so that V_n^j start at time $\frac{i-1}{n}$ from x. This will imply that the conditional expectation of $\varphi(V_n^j(t))e^{\int_{\frac{i-1}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{t} h^*h(V_n^j(s))ds}$ given $\mathcal{Y} \vee \sigma(V_n^j(\frac{i-1}{n}))$ (and, consequently, given $\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}$) is $R(V_n^j(\frac{i-1}{n}))$. Using the fact that $\rho_t(\varphi) = p_{\frac{i-1}{n}}(\psi_{\frac{i-1}{n}})$, we find

$$\begin{aligned}
&\tilde{E}_{\frac{i-1}{n},x}\left[\varphi(V_n^j(t))e^{\int_{\frac{i-1}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{t} h^*h(V_n^j(s))ds}|\mathcal{Y}\right] \\
&\qquad = \tilde{E}_{\frac{i-1}{n},x}\left[\psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))|\mathcal{Y}\right] = \psi_{\frac{i-1}{n}}(x)
\end{aligned}$$

Hence

$$\tilde{E}\left[\varphi(V_n^j(t))e^{\int_{\frac{i-1}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{t} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}\right] = \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n})) \tag{5.62}$$

Similarly

$$\tilde{E}\left[\varphi(V_n^j(t))e^{\int_{\frac{i}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i}{n}}^{t} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i}{n}}\right] = \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n})) \tag{5.63}$$

From (5.62) and (5.63) we get that

$$\begin{aligned}
&\tilde{E}\left[\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}\right] \\
&\quad = \tilde{E}\left[\tilde{E}\left[\varphi(V_n^j(t))e^{\int_{\frac{i}{n}}^{t} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i}{n}}^{t} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i}{n}}\right]\right. \\
&\qquad \left.\times\, e^{\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*(V_n^j(s))dY_s - \frac{1}{2}\int_{\frac{i-1}{n}}^{\frac{i}{n}} h^*h(V_n^j(s))ds}|\mathcal{Y} \vee \mathcal{F}_{\frac{i-1}{n}}\right] \\
&\quad = \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))
\end{aligned}$$

which proves (5.60). The identity (5.57) follows now from (5.59) and (5.60). In the analysis above we considered V_n^j defined up to time t, although in the description of the branching system it is not, but obviously we can attach 'an extension' from $\frac{i}{n}$ to

t, satisfying the same SDE and independent of Y. We prove now (5.58). Using Itô's formula we have that

$$
\begin{aligned}
&(\mathcal{Z}_n(t),\psi_t) - (\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]})\\
&= \frac{1}{n}\tilde{E}[\sum_{j=1}^{N_n([\frac{nt}{n})])}\int_{[\frac{nt}{n}]}^{t} h^*(V_n^j(s))\psi(V_n^j(s))dY_s|\mathcal{Y}]\\
&= \frac{1}{n}\tilde{E}[\sum_{j=1}^{N_n([\frac{nt}{n})])}\int_{[\frac{nt}{n}]}^{t} \tilde{E}[h^*(V_n^j(s))\psi(V_n^j(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]dY_s|\mathcal{Y}] \qquad (5.64)
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\tilde{E}[((\mathcal{Z}_n(t),\psi_t) - (\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]}))^2]\\
&\leq \frac{1}{n^2}\tilde{E}[(\sum_{j=1}^{N_n([\frac{nt}{n})])}\int_{[\frac{nt}{n}]}^{t} \tilde{E}[h^*(V_n^j(s))\psi(V_n^j(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]dY_s)^2]\\
&= \frac{1}{n^2}\tilde{E}\left[\sum_{j_1,j_2=1}^{N_n([\frac{nt}{n})])}\tilde{E}[\int_{[\frac{nt}{n}]}^{t} \tilde{E}[h^*(V_n^{j_1}(s))\psi(V_n^{j_1}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]dY_s\right.
\end{aligned}
$$

$$
\begin{aligned}
&\left.\times\int_{[\frac{nt}{n}]}^{t} \tilde{E}[h^*(V_n^{j_2}(s))\psi(V_n^{j_2}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]dY_s|\mathcal{F}_{[\frac{nt}{n}]}]\right]\\
=\ &\frac{1}{n^2}\tilde{E}\left[\sum_{j_1,j_2=1}^{N_n([\frac{nt}{n})])}\tilde{E}[\int_{[\frac{nt}{n}]}^{t} \tilde{E}[h^*(V_n^{j_1}(s))\psi(V_n^{j_1}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]\right.\\
&\left.\times\tilde{E}[h^*(V_n^{j_2}(s))\psi(V_n^{j_2}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]ds|\mathcal{F}_{[\frac{nt}{n}]}]\right]\\
\leq\ &\frac{\|h\|^2}{n^2}\tilde{E}\left[\sum_{j_1,j_2=1}^{N_n([\frac{nt}{n})])}\int_{[\frac{nt}{n}]}^{t}\tilde{E}[\tilde{E}[\psi(V_n^{j_1}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]\tilde{E}[\psi(V_n^{j_2}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}]|\mathcal{F}_{[\frac{nt}{n}]}]ds\right]\\
\leq\ &\frac{\|h\|^2}{n^2}\tilde{E}\left[\sum_{j_1,j_2=1}^{N_n([\frac{nt}{n})])}\int_{[\frac{nt}{n}]}^{t}\sqrt{\tilde{E}[(\tilde{E}[\psi(V_n^{j_1}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}])^2|\mathcal{F}_{[\frac{nt}{n}]}]}\right.\\
&\left.\times\sqrt{\tilde{E}(\tilde{E}[\psi(V_n^{j_2}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}])^2|\mathcal{F}_{[\frac{nt}{n}]}]}ds\right] \qquad (5.65)
\end{aligned}
$$

Using once again an argument based on the Gronwall inequality we obtain that

$$\tilde{E}(\tilde{E}[\psi(V_n^{j_2}(s))|\mathcal{Y}\vee\mathcal{F}_{[\frac{nt}{n}]}])^2|\mathcal{F}_{[\frac{nt}{n}]}] \leq M\|\varphi\|^2$$

and plugging this in (5.65), we get that

$$\tilde{E}[((\mathcal{Z}_n(t),\psi_t) - (\mathcal{Z}_n(\left[\frac{nt}{n}\right]),\psi_{[\frac{nt}{n}]}))^2] \leq \frac{Mc(t)\|h\|^2\|\varphi\|^2}{n} \tag{5.66}$$

where c is the function introduced in section 3.

■

In conclusion, we have proved in this section the *existence* of a solution to the filtered martingale problem (1.3)+ (1.4). This is an extension of the classical Dawson-Watanabe construction. Averaging the particle approximations over independent evolutions leads to the numerical approximation of the Zakai equation.

5.2 The Measure - Valued Process U

In the following we will construct a sequence of branching particle systems U_n, whose laws will approximate ρ_t, i.e.

$$\lim_{n\to\infty}(U_n,\varphi) = (U,\varphi) = \rho_t(\varphi).$$

The particles will move according to the law of the signal, independently of each other and after fixed-length intervals will branch. The mean number of offspring of a particle will depend on the last part of its trajectory and on the observation process and the variance of the branching mechanism will be the minimum possible one.

We can use these particle systems to solve numerically the filtering problem since, as the number of particles is increased, the empirical measure associated to the cloud of particles converges to the solution of the Zakai equation. By starting with a number of particles constituting a sample approximating the initial distribution of X, we allow the system to evolve up to time t and then use the empirical law to estimate the distribution of the required statistic φ.

In the previous section, we constructed a sequence $\mathcal{X}_n$ of similar branching particle systems also related to the Zakai equation. In that case, the variance of the branching mechanism was a-priori given. In this way, we introduced an extra degree of randomness and only the (conditional) expectation of that sequence tended to ρ_t. Therefore, we needed a whole set of copies of the particle system in order to obtain a good approximation to the solution of the Zakai equation. Our new approach converges directly to ρ_t, we don't need to estimate an average.

In this section we revert to a general locally compact, separable metric space E and use the notation set up in Chapter 2 and 4.

We assume that the domain of A_0 - the restriction of A to $C_0(E)$ - is a dense algebra in $C_0(E)$ and that the martingale problem for A_0 is well posed so that the uniqueness of the solution of the measure valued equation (2.45) holds.

5.2.1 The Construction of the Sequence of the Particle Systems U_n

Once again, we work under the new probability measure $\tilde{P}$ and all the expectations will be considered with respect to $\tilde{P}$. We will construct the particle systems and, implicitly, the measure valued process U, up to a fixed horizon, i.e., on the fixed interval $[0,1]$, the construction being identical for any interval $[0,T]$. Then, using an argument based on the Caratheodory extension theorem, the construction can be extended for the whole positive axis.

Let $\{U_n(t), \mathcal{F}_t;\ 0 \leq t \leq 1\}$ be a sequence of branching particle systems on $(\Omega, \mathcal{F}, \tilde{P})$ with values in $M_F'(E)$ defined as follows:

Initial condition

1. $U_n(0)$ is the empirical measure of n particles of mass $\frac{1}{n}$, i.e., $U_n(0) = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i^n}$, where $x_i^n \in E$, for every $i, n \in \mathbb{N}$.
2. The empirical measure of the particles tends weakly to π_0.

Evolution in time

We describe the evolution of the processes on the interval $[\frac{i}{n}, \frac{i+1}{n}]$, $i = 0, 1, ..., n-1$.

1. At the time $\frac{i}{n}$, the process consists of the occupation measure of $m_n(\frac{i}{n})$ particles of mass $\frac{1}{n}$ (we will denote the number of particles alive at time t by $m_n(t)$).
2. During the interval the particles move independently with the same law as the signal X. Let $V(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$ be the trajectory of a generic particle in this interval.
3. At the end of the interval, each particle branches into a random number of particles

with a mechanism depending on its trajectory in the interval. The mechanism is chosen so that it has finite second moment, the mean number of offspring for a particle given the σ-field $\mathcal{F}_{\frac{i+1}{n}-} = \sigma(\mathcal{F}_s,\, s < \frac{i+1}{n})$ of events up to time $\frac{i+1}{n}$ is

$$\mu_n^i(V) \overset{\text{def}}{=} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(V(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(V(t))dt\right) \tag{5.67}$$

and so that the variance $\nu_n^i(V)$ is minimal consistent with the number of offspring being an integer. The particles branch independently of each other, given $\mathcal{F}_{\frac{i+1}{n}-}$.

We remark that $\nu_n^i(V) = (\mu_n^i(V) - [\mu_n^i(V)])([\mu_n^i(V)] + 1 - \mu_n^i(V))$ and so is always less than $\frac{1}{4}$ ($[x]$ is the largest integer smaller that x).

We now make some preliminary estimates before showing the convergence of U_n in the next section. Just before the $(i+1)$-th branching, we will have $m_n(\frac{i}{n})$ particles. Let us denote by $U_n(\frac{i+1}{n}-)$ the state of the process just before the $(i+1)$-th branching and by $V_n^j(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$ the trajectory of the j-th particle alive during the interval. Let also $q_n^j(\frac{i+1}{n})$ be the number of offspring of the j-th particle with $1 \le j \le m_n(\frac{i}{n})$ at time $\frac{i+1}{n}$.

Proposition 15 *We have the following trivial a priori upper bounds:*

i. $\tilde{E}[m_n(t)] = n$, $\forall n \ge 0$, $t \in [0,1]$.

ii. $\tilde{E}[m_n^2(t)] \le n^2 e^{\|h\|^2\frac{[nt]}{n}} + \frac{1}{4}\sum_{k\le[nt]} e^{\frac{k\|h\|^2}{n}}$, $\forall n \ge 0$, $t \in [0,1]$.

Proof. i. The number of particles does not change during the intervals $(\frac{k}{n}, \frac{k+1}{n})$, $k = 1, \dots, n-1$ so $m_n(t) = m_n(\frac{[tn]}{n})$. Therefore it suffices to prove that $\tilde{E}[m_n(\frac{i}{n})] = \tilde{E}[m_n(\frac{(i+1)}{n})]$ for $0 \le i < n$. Using (5.4), we have

$$\begin{aligned}
\tilde{E}[m_n(\frac{i+1}{n})] &= \tilde{E}[\sum_{j=1}^{m_n(\frac{i}{n})} q_n^j(\frac{i+1}{n})] \\
&= \tilde{E}[\sum_{j=1}^{m_n(\frac{i}{n})} \tilde{E}[q_n^j(\frac{i+1}{n})|\mathcal{F}_{\frac{i+1}{n}-}] \\
&= \tilde{E}[\sum_{j=1}^{m_n(\frac{i}{n})} \mu_n^i(V_n^j)]
\end{aligned}$$

$$
\begin{aligned}
&= \tilde{E}[\sum_{j=1}^{m_n(\frac{i}{n})} \tilde{E}[\mu_n^i(V_n^j)|\mathcal{F}_{\frac{i}{n}}]] \\
&= E[m_n(\frac{i}{n})]
\end{aligned}
$$

since $\tilde{E}[\mu_n^i(V_n^j)|\mathcal{F}_{\frac{i}{n}}] = 1$ ($\mu_n^i(V_n^j)$ is the value at time $\frac{i+1}{n}$ of an exponential martingale identically equal to 1 at time $\frac{i}{n}$).

ii. From the construction of the branching mechanism of the particles we have that

$$
\begin{aligned}
\tilde{E}[(q_n^j(\frac{(i+1)}{n}))^2|\mathcal{F}_{\frac{i+1}{n}-}] &= \nu_n^i(V_n^j) + (\mu_n^i(V_n^j))^2 \\
&\leq \frac{1}{4} + e^{\frac{\|h\|^2}{n}} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} 2h^*(V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} (2h)^*(2h)(t, V_n^j(t))dt\right)
\end{aligned}
$$

This inequality and the independence of the particles implies (as in i.)

$$
\begin{aligned}
\tilde{E}[(m_n(\frac{i+1}{n}))^2] &= \tilde{E}[\sum_{j=1}^{m_n(\frac{i}{n})} \tilde{E}[\tilde{E}[(q_n^j(\frac{i+1}{n}))^2|\mathcal{F}_{\frac{i+1}{n}-}]\mathcal{F}_{\frac{i}{n}}]] \\
&\quad +2\tilde{E}[\sum_{1\leq j_1<j_2\leq l}^{m_n(\frac{i}{n})} \tilde{E}[\mu_n^i(V_n^{j_1})\mu_n^i(V_n^{j_2})|\mathcal{F}_{\frac{i}{n}}] \\
&\leq \frac{1}{4} + e^{\frac{\|h\|^2}{n}}\tilde{E}[m_n(\frac{i}{n})] + e^{\frac{\|h\|^2}{n}}\tilde{E}[m_n(\frac{i}{n})(m_n(\frac{i}{n}) - 1]
\end{aligned}
$$

It follows that

$$
\tilde{E}[(m_n(\frac{i+1}{n}))^2] \leq e^{\frac{\|h\|^2}{n}}\tilde{E}[(m_n(\frac{i}{n}))^2] + \frac{1}{4}
$$

hence

$$
\tilde{E}[(m_n(t))^2] = \tilde{E}[(m_n(\frac{[tn]}{n}))^2] \leq e^{\|h\|^2\frac{[nt]}{n}} m_n^2(0) + \frac{1}{4}\sum_{k\leq[nt]} e^{\frac{k\|h\|^2}{n}}
$$

which completes the proof .

■

Remark 18 *For any $t \in [0,1]$ and $\varphi \in B(E)$, the processes $(U_n(t), \varphi)$ are square integrable.*

Proof. Let V_1, V_2, ..., $V_{m_n(t)}$ be the positions of the $m_n(t)$ particles alive at time t.

Using Proposition 3.1, we have

$$\tilde{E}[(U_n(t),\varphi)^2] = \tilde{E}[(\frac{1}{n}\sum_{j=1}^{m_n(t)}\varphi(V_j)^2] = \frac{\|\varphi\|}{n^2}\tilde{E}[m_n^2(t)] < \infty$$

■

Remark 19 *If $\varphi \in \mathcal{D}(A) \cup \{1\}$, then the process $(U_n(t),\varphi)$ satisfies the following evolution equation*

$$(U_n(t),\varphi) = (U_n(0),\varphi) + \int_0^t (U_n(s),A\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt])$$

$$+\sum_{i=1}^{[nt]}\frac{1}{n}\sum_{j=1}^{m_n(\frac{i-1}{n})}\varphi(V_n^j(\frac{i}{n}))(\mu_n^i(V_n^j)-1). \tag{5.68}$$

where $\{(S_n^{\varphi}(t),\ \mathcal{F}_t),\ t \in [0,1]\}$ is a square integrable martingale with quadratic variation

$$< S_n^{\varphi} > (t) = \frac{1}{n}\int_0^t (U_n(s),A\varphi^2 - 2\varphi A\varphi)ds \tag{5.69}$$

and $\{(M_n^{\varphi}(l),\ \mathcal{F}_{\frac{l+1}{n}-}),\ l = 0,1,...,n\}$ is a discrete martingale with (conditional) quadratic variation

$$< M_n^{\varphi} > (l) = \frac{1}{n}\sum_{i=1}^{l}(U_n(\frac{i+1}{n}-),\nu_n^i\varphi^2) \tag{5.70}$$

Proof. From the construction of the particle systems we have that

$$\tilde{E}[(U_n(\frac{i+1}{n}),\varphi)|\mathcal{F}_{\frac{i+1}{n}-}] = \frac{1}{n}\sum_{i=1}^{m_n(\frac{i}{n})}\varphi(V_n^j(\frac{i+1}{n}))\mu_n^i(V_n^j) \tag{5.71}$$

and also

$$\tilde{E}[(U_n(\frac{i+1}{n}),\varphi)^2|\mathcal{F}_{\frac{i+1}{n}-}] - (\tilde{E}[(U_n(\frac{i+1}{n}),\varphi)|\mathcal{F}_{\frac{i+1}{n}-}])^2 = \frac{1}{n}(U_n(\frac{i+1}{n}-),\nu_n^i\varphi^2). \tag{5.72}$$

In between two branches the particles move according to the prescribed law, hence for t in the interval $[\frac{i}{n},\frac{i+1}{n})$

$$(U_n(t),\varphi) = (U_n(\frac{i}{n}),\varphi) + \int_{\frac{i}{n}}^t (U_n(s),A\varphi)ds + S_n^{\varphi,i}(t), \tag{5.73}$$

where $\{(S_n^{\varphi,i}(t),\mathcal{F}_t),\ t\in[\frac{i}{n},\frac{i+1}{n}]\}$ is a square integrable martingale (we use Remark 3.2) with the quadratic variation

$$< S_n^{\varphi,i} > (t) = \frac{1}{n}\int_{\frac{i}{n}}^{t}(U_n(s), A\varphi^2 - 2\varphi A\varphi)ds. \tag{5.74}$$

In order to compute the evolution equation of $(U_n(t),\varphi)$, we need to add all the parts coming from the particles motion and all the parts coming from the particle branching, which gives

$$\begin{aligned}(U_n(t),\varphi) &= (U_n(0),\varphi) + \int_0^t (U_n(s), A\varphi)ds + S_n^{\varphi}(t)\\ &+\sum_{i=1}^{[nt]}\left((U_n(\frac{i}{n}),\varphi) - (U_n(\frac{i}{n}-),\varphi)\right)\end{aligned}$$

where $\{(S_n^{\varphi}(t),\ \mathcal{F}_t),\ t\in[0,1]\}$ is the square integrable martingale

$$S_n^{\varphi}(t) \stackrel{\text{def}}{=} S_n^{\varphi,[nt]}(t) + \sum_{i=0}^{[nt]-1} S_n^{\varphi,i}(\frac{i+1}{n})$$

with the quadratic variation presented in (5.11). We then split the term coming from the branching into a martingale part and a bounded variation part and obtain

$$\begin{aligned}(U_n(t),\varphi) &= (U_n(0),\varphi) + \int_0^t (U_n(s), A\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt])\\ &+\sum_{i=1}^{[nt]}\left(\tilde{E}[(U_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] - (U_n(\frac{i}{n}-),\varphi)\right),\end{aligned} \tag{5.75}$$

and $\{(M_n^{\varphi}(l),\ \mathcal{F}_{\frac{l+1}{n}-}),\ l=0,1,...,n\}$ is the square integrable martingale

$$M_n^{\varphi}(0) \stackrel{\text{def}}{=} 0,$$

$$M_n^{\varphi}(l) \stackrel{\text{def}}{=} \sum_{i=1}^{l}\left((U_n(\frac{i}{n}),\varphi) - \tilde{E}[(U_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}]\right).$$

with (conditional) quadratic variation

$$< M_n^{\varphi} > (l) = \sum_{i=1}^{l}\tilde{E}\left[\left((U_n(\frac{i}{n}),\varphi) - \tilde{E}[(U_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}]\right)^2 |\mathcal{F}_{\frac{i}{n}-}\right] \tag{5.76}$$

The remark follows now easily from (5.71), (5.72), (5.75) and (5.76).

■

We introduce also the notation

$$l_n^i(V,r) \overset{\text{def}}{=} \exp\left(\int_{\frac{i}{n}}^{r} h^*(V(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{s} h^*h(V(t))dt\right), \quad r \in [\frac{i}{n}, \frac{i+1}{n}]$$

where V is the trajectory of a particle alive in the interval $[\frac{i}{n}, \frac{i+1}{n}]$. Of course, $\mu_n^i(V) = l_n^i(V, \frac{i+1}{n})$. Applying Itô's rule and exploiting the fact that Y is a Brownian motion, we get from (5.13) that

$$(U_n(t),\varphi) = (U_n(0),\varphi) + \int_0^t (U_n(s), A\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt])$$

$$+\frac{1}{n}\int_0^{\frac{[nt]}{n}} \sum_{j=1}^{m_n(\frac{[sn]}{n})} \varphi(V_n^j(\frac{[sn]+1}{n}))l_n^{[sn]}(V_n^j, s)h^*(V_n^j(s))dY_s \qquad (5.77)$$

5.2.2 The Tightness of the Sequence U_n

Once again, we need to show that conditions **a** and **b** of Theorem 21 hold true. The two conditions follow from the following propositions.

Proposition 16 *For every $t \in [0,1]$, we have*

$$\lim_{k\to\infty} \sup_{n\geq 0} \tilde{P}(\sup_{0\leq s\leq t} (U_n(s), 1) > k) = 0. \qquad (5.78)$$

Proof. Since

$$\tilde{P}(\sup_{0\leq s\leq t} (U_n(s), 1) > k) \leq \frac{\tilde{E}[(\sup_{0\leq s\leq t}(U_n(s),1))^2]}{k^2}, \qquad (5.79)$$

it is enough to show that $\sup_{n\geq 0}\tilde{E}[(\sup_{0\leq s\leq t}(U_n(s),1))^2]$ is finite. Let us denote by $\psi_n(t) \overset{\text{def}}{=} \tilde{E}\left[(\sup_{0\leq s\leq t}(U_n(s),1)))^2\right]$. From (5.14) and the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}\psi_n(t) \quad &\leq \quad 3(U_n(0),1)^2 + 3\tilde{E}[(\sup_{0\leq i\leq [nt]} |M_n^1(i)|)^2] \\ &+\frac{3}{n^2}\tilde{E}\left[\left(\sup_{0\leq p\leq \frac{[tn]}{n}}\left|\int_0^p \sum_{j=1}^{m_n(\frac{[sn]}{n})} l_n^{[sn]}(V_n^j, s)h^*(V_n^j(s))dY_s\right|\right)^2\right] \qquad (5.80)\end{aligned}$$

The idea is to find an upper bound for each of the three terms from the right hand side of the inequality (5.80) of the form $\alpha + \beta \int_0^t \psi_n(s)ds$ and then use the Gronwall inequality.

The first term

Since we start with n particles, we have

$$(U_n(0), 1)^2 = 1, \quad \forall n \geq 1. \tag{5.81}$$

The second term

Doob's maximal inequality, the fact that $\nu_n^i \leq \frac{1}{4}$ and the proposition 3.1 gives us the following upper bound:

$$\tilde{E}[(\sup_{0 \leq i \leq [nt]} |M_n^1(i)|)^2] \leq 4\tilde{E}[(M_n^1([nt]))^2] = 4\tilde{E}[(< M_n^1 > ([nt]))^2]$$

$$= \frac{4}{n} \sum_{i=1}^{[nt]} \tilde{E}[(U(\frac{i}{n}-), \nu_n^i)] \leq \frac{1}{n} \sum_{i=1}^{[nt]} \tilde{E}[(U(\frac{i}{n}-), 1)] \leq \frac{1}{n} \sum_{i=1}^{[nt]} \tilde{E}[\frac{m_n(t)}{n}] \leq 1. \tag{5.82}$$

The third term

Using a standard technique similar with the one employed in the previous section, we find

$$\tilde{E}[(\sum_{j=1}^{m_n(\frac{[sn]}{n})} l_n^{[sn]}(V_n^j, s))^2] \leq e^{2\|h\|^2} \tilde{E}[(m_n(\frac{[sn]}{n}))^2]. \tag{5.83}$$

Then, we obtain the following upper bound on the third term of (5.80), using (5.83) and Burkholder's inequality

$$e^{2\|h\|^2} K_2 \|h\|^2 \int_0^t \psi_n(s)ds, \tag{5.84}$$

where K_2 is a constant independent of n.

From (5.80), (5.81), (5.82) and (5.84) we obtain

$$\psi_n(t) \leq 6 + 3e^{2\|h\|^2} K_2 \|h\|^2 \int_0^t \psi_n(s)ds$$

Using once again the Gronwall inequality we find that $\psi_n(t) \leq c(t)$, where

$$c(t) \overset{\text{def}}{=} 6e^{\frac{1}{2}e^{2\|h\|^2} K_2 \|h\|^2 t}, \ t \in [0,1] \tag{5.85}$$

and also $\sup_{n\geq 0}\tilde{E}[(\sup_{0\leq s\leq t}(U_n(s),1))^2]\leq c(t)$.

■

Proposition 17 *For any arbitrary sequence of stopping times* $\{\tau_n\}_{n\geq 0}$, *any real positive sequence* $\{\delta_n\}_{n\geq 0}$ *with* $\lim_{n\to\infty}\delta_n=0$ *and* $\varphi\in\mathcal{D}(A_0)$, *we have*

$$\lim_{n\to\infty}\tilde{E}[|(U_n(\tau_n+\delta_n),\varphi)-(U_n(\tau_n),\varphi)|^2]=0. \tag{5.86}$$

Proof. Using (5.14) and the Cauchy-Schwartz inequality we get

$$\begin{aligned}\tilde{E}[|(U_n(\tau_n+\delta_n),\varphi)-(U_n(\tau_n),\varphi)|^2] \leq\ & 4\tilde{E}[(\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)ds)^2]\\ & +4\tilde{E}[(S_n^{\varphi}(\tau_n+\delta_n)-S_n^{\varphi}(\tau_n))^2]\\ & +4\tilde{E}[(M_n^{\varphi}([n(\tau_n+\delta_n)])-M_n^{\varphi}([n\tau_n]))^2]\\ & +4\tilde{E}[(\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}\varphi(V_n^j(\frac{[sn]+1}{n}))\\ & \times l_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s)^2]\end{aligned} \tag{5.87}$$

We have, consecutively,

$$\begin{aligned}\tilde{E}[(\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)ds)^2] &\leq \delta_n\tilde{E}[\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)^2ds]\\ &\leq \delta_n^2\|A\varphi\|^2c(1) \qquad (5.88)\\ \tilde{E}[(S_n^{\varphi}(\tau_n+\delta_n)-S_n^{\varphi}(\tau_n))^2] &= \tilde{E}[<S_n^{\varphi}>(\tau_n+\delta_n)-<S_n^{\varphi}>(\tau_n)]\\ &\leq \frac{1}{n}\tilde{E}[\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi^2-\varphi A\varphi)ds]\\ &\leq \frac{(c(1)+1)(\|A\varphi^2\|+2\|\varphi\|\|A\varphi\|)}{2n}\delta_n \qquad (5.89)\\ \tilde{E}[(M_n^{\varphi}([n(\tau_n+\delta_n)])-M_n^{\varphi}([n\tau_n]))^2] &= \tilde{E}[<M_n^{\varphi}>([n(\tau_n+\delta_n)])-<M_n^{\varphi}>([n\tau_n])]\\ &= \frac{1}{n}\tilde{E}[\sum_{[n\tau_n]}^{[n(\tau_n+\delta_n)]}(U_n(\frac{i+1}{n}-),\nu_n^{i+1}\varphi^2)]\\ &\leq \frac{1}{8}(c(1)+1)\|\varphi\|^2(\delta_n+\frac{1}{n}) \qquad (5.90)\end{aligned}$$

$$\tilde{E}[(\frac{1}{n^2}\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}\varphi(V_n^j(\frac{[sn]+1}{n}))l_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s)^2]$$

$$
\begin{aligned}
&= \tilde{E}\Big[\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}} \frac{1}{n^2}\Big(\sum_{j=1}^{m_n(\frac{[sn]}{n})} \varphi\big(V_n^j(\frac{[sn]+1}{n})\big) l_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))\Big)^2 ds\Big] \\
&\leq \|\varphi\|^2\|h\|^2 c(1)e^{2\|h\|^2}\delta_n \qquad (5.91)
\end{aligned}
$$

where c is the function defined in (5.85). The inequalities (5.28), (5.29), (5.88), (5.91) imply that all the terms from the right hand side of (5.87) tend to 0 when n goes to ∞, hence $[\tilde{E}[|(U_n(\tau_n+\delta_n),\varphi)-(U_n(\tau_n),\varphi)|^2]$ tends to 0 as well.

■

5.2.3 Application to the Nonlinear Filtering

In this section we prove the following results:

Proposition 18 *Every subsequence of $U_n(t)$ contains a subsubsequence $U_{n_k}(t)$ such that for $\varphi \in \mathcal{M}\cup\{1\}$ we have*

$$\lim_{k\to\infty}(U_{n_k}(t),\varphi) = \rho_t(\varphi),\ \tilde{P}-a.s., \tag{5.92}$$

where $\mathcal{M}$ is a countable and uniformly dense set in $C_K(E)$ (the space of continuous functions with compact support).

Proposition 18 will imply the following theorem.

Theorem 27 *Every subsequence of $U_n(t)$ contains a subsubsequence $U_{n_k}(t)$ convergent $\tilde{P}$-a.s. to the unnormalised distribution of the signal ρ_t*

Remark 20 *Theorem 27 says much more that Proposition 18 since, for the theorem to be true, we need (5.92) to hold with the same null set for all $\varphi \in C_b(E)$, in other words we need to prove that there exists $\Omega' \subset \Omega$, with $\tilde{P}(\Omega')=1$ such that for all $\varphi \in C_b(E)$, we have*

$$\lim_{k\to\infty}(U_{n_k}(t),\varphi) = \rho_t(\varphi),\ \forall\omega\in\Omega' \tag{5.93}$$

Theorem 27 gives us the following obvious Corollary.

Corollary 3 *The sequence $U_n(t)$ is convergent in measure to ρ_t, i.e. to the law of $X(t)$ given the observation σ-field $\mathcal{Y}_t$.*

It is based on this result that one is able to use U_n to approximate numerically U. Theorem 27 (together with the tightness of the sequence) has another corollary, of theoretical importance.

Corollary 4 *The sequence U_n is convergent in distribution to the measure valued process that represents the unnormalised conditional law of X given the observation Y.*

Remark 21 *In Corollary 4, we look at U_n and p as having values in the space of càdlàg, $M_F(E)$-valued paths. The previous section proves that the sequence is tight over this space.*

To prove Proposition 18, we need the following lemma (the notations are those from Section 5.2.1).

Lemma 6 *For all $i = 0, \ldots, n-1$, we have*

$$\tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} \nu_n^i(V_n^j)\right] \leq \frac{M(h)}{\sqrt{n}}$$

where $M(h)$ is a constant depending only on h (the function appearing in the drift of the observation).

Proof. We remark that, if we have an integer random variable with mean μ and minimal variance ν, then $\nu \leq |\mu - 1|$. Hence

$$\begin{aligned}
\tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} \nu_n^i(V_n^j)\right] &\leq \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} |\mu_n^i(V_n^j) - 1|\right] \\
&\leq \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} E[|\mu_n^i(V_n^j) - 1| \,\big|\mathcal{F}_{\frac{i}{n}}]\right] \\
&\leq \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} \sqrt{E\left[\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \lambda_n^i(V_N^j, s)h^*(V_N^j(s))dY_s\right)^2 \Bigg|\mathcal{F}_{\frac{i}{n}}\right]}\right] \\
&\leq \frac{M(h)}{\sqrt{n}}
\end{aligned}$$

■

In the following we will denote by $h(r)$ the function $x \to h(r,x)$. We consider first the following backward stochastic partial differential equation

$$\begin{aligned} d\psi_s &= -A\psi_s ds - h^*(s)\psi_s d\bar{Y}_s \quad s \leq t \\ \psi_t &= \varphi \end{aligned} \tag{5.94}$$

which, written in the integral form, gives us

$$\psi_r = \psi_s - \int_s^r A\psi_p dp - \int_s^r h^*(p)\psi_p d\bar{Y}_p \; r,s \in [0,t]. \tag{5.95}$$

In (5.95), we took $\int_s^r h^*(p)\psi_p d\bar{Y}_p$ to be a backward Itô integral. Written in Stratonovitch form, (5.95) becomes

$$\psi_r = \psi_s - \int_s^p A\psi_p dp - \int_s^p h^*(p)\psi_p \circ dY_p + \frac{1}{2}\int_s^p h^*h(p)\psi_p dp \tag{5.96}$$

We will assume the following condition:

U. For all $t \in [0,1]$, there exists a countable set $\mathcal{M}$, uniformly dense in $C_K(E)$, such that for all $\varphi \in \mathcal{M} \cup \{1\}$ the SPDE (5.94) has a solution $\psi_s \in \mathcal{D}(A)$ which satisfies

$$B \stackrel{\text{def}}{=} \tilde{E}\left[\sup_{s\in[0,t]} \|\psi_s\|^2\right] < \infty. \tag{5.97}$$

and

$$C \stackrel{\text{def}}{=} \tilde{E}\left[\sup_{s\in[0,t]} \|A\psi_s^2 - 2\psi_s A\psi_s\|\right] < \infty. \tag{5.98}$$

See the Appendix for sufficient conditions on A and h, under which **U** holds. Let now $\{V_r, \mathcal{F}_r, r \in [s,t]\}$ be a process solution of the martingale problem associated with the infinitesimal generator A independent of $\mathcal{Y}_s^t \stackrel{\text{def}}{=} \sigma(Y_t - Y_r, r \in [s,t])$. From (5.96) we obtain that

$$\psi_r(V_r) = \psi_s(V_s) - \int_s^r h^*(p,V_p)\psi_p(V_p) \circ dY_p + \frac{1}{2}\int_s^r h^*h(p,V_p)\psi_p(V_p)dp + M_r^{\psi}$$

where $\{M_r^{\psi}, \mathcal{F}_r \vee \mathcal{Y}, r \in [s,t]\}$ is a square integrable martingale (due to (5.97)) with quadratic variation

$$< M_r^{\psi} >= \int_s^r A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p)dp.$$

Let

$$\xi_r = \exp(\int_s^r h^*(p,V_p) \circ dY_p - \frac{1}{2}\int_s^r h^*h(p,V_p)dp),$$

then

$$\xi_r = 1 + \int_s^r \xi_p h^*(p,V_p) \circ dY_p - \frac{1}{2}\int_s^r \xi_p h^*h(p,V_p)dp$$

and thus

$$\psi_r(V_r)\xi_r = \psi_s(V_s) + \overline{M}_r^{\psi} \tag{5.99}$$

where $\{\overline{M}_r^{\psi}, \mathcal{F}_r \vee \mathcal{Y}, r \in [s,t]\}$ is a square integrable martingale with quadratic variation

$$< \overline{M}_r^{\psi} >= \int_s^r \xi_p A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p)dp.$$

Hence

$$\begin{aligned} \tilde{E}[(\psi_r(V_r)\xi_r - \psi_s(V_s))^2|\mathcal{F}_r] &= \int_s^r \tilde{E}[\xi_p A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p)|\mathcal{F}_r]dp \\ &\leq \int_s^r \tilde{E}\left[\xi_p\|A\psi_p^2 - 2\psi_p A\psi_p\|\,|\mathcal{F}_r\right]dp \\ &\leq (r-s)C \end{aligned} \tag{5.100}$$

The last inequality holds true since $A\psi_p^2 - 2\psi_p A\psi_p$ is independent of ξ_p and $\mathcal{F}_r$ and $\tilde{E}[\xi_p\,|\mathcal{F}_r] = 1$. Armed now with the inequality (5.100) we can to prove Proposition 18 .

Proof of Proposition 18

Since $\overline{M}_r^{\psi}$ is a martingale with respect to the filtration $\mathcal{F}_r \vee \mathcal{Y}$, from (5.99), we get that for $\varphi \in \mathcal{M} \cup \{1\}$,

$$\rho_t(\varphi) = \tilde{E}[(\varphi(X_t)\xi_t\,|\mathcal{Y}] = \tilde{E}[(\psi_t(X_t)\xi_t\,|\mathcal{Y}] = \tilde{E}[(\psi_0(X_0)\,|\mathcal{Y}] = (\pi_0, \psi_0) \tag{5.101}$$

In (5.101), we used the fact that $\psi_t = \varphi$. Also, since U_n converges weakly to π_0 and $\psi_0 \in \mathcal{D}(A) \subset C_b(E)$, $\tilde{P}$-a.s., we have that

$$\lim_{n\to\infty}(U_n(0),\psi_0) = (\pi_0,\psi_0),\ \tilde{P}-a.s. \tag{5.102}$$

Hence, in order to prove the proposition it is enough to show that

$$\lim_{n\to\infty}\tilde{E}[((U_n(t),\psi_t)-(U_n(0),\psi_0))^2] = 0. \tag{5.103}$$

We have the following identity

$$\begin{aligned}(U_n(t),\psi_t)-(U_n(0),\psi_0) &= (U_n(t),\psi_t)-(U_n(\frac{[nt]}{n}),\psi_{\frac{[nt]}{n}})\\ &+\sum_{i=1}^{\frac{[nt]}{n}}(U_n(\frac{i}{n}),\psi_{\frac{i}{n}})-\tilde{E}[(U_n(\frac{i}{n}),\psi_{\frac{i}{n}})\left|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}\right.] \qquad (5.104)\\ &+\sum_{i=1}^{\frac{[nt]}{n}}\tilde{E}[(U_n(\frac{i}{n}),\psi_{\frac{i}{n}})\left|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}\right.]-(U_n(\frac{i-1}{n}),\psi_{\frac{i-1}{n}})\end{aligned}$$

We show that all the terms from the right hand side of (5.104) converge to 0 in $L^2(\Omega)$. For the first term we have the following

$$\begin{aligned}(U_n(t),\psi_t)-(U_n(\frac{[nt]}{n}),\psi_{\frac{[nt]}{n}}) &= \frac{1}{n}\sum_{j=1}^{m_n(t)}\varphi(V_n^j(t))(1-\lambda_n^{[nt]}(V_n^j,t))\\ &+\frac{1}{n}\sum_{j=1}^{m_n(t)}\psi_t(V_n^j(t))\lambda_n^{[nt]}(V_n^j,t)-\psi_{\frac{[nt]}{n}}(V_n^j(\frac{[nt]}{n}))\end{aligned}$$

Using now (5.100) and the fact the φ is bounded, we get the following upper bound for the L_2 norm of the first term

$$\frac{2c(1)(2\left\|\varphi\right\|^2\left\|h\right\|^2 e^{\frac{\|h\|^2}{n}}+C)}{n} \tag{5.105}$$

For the second term we have the following identity

$$\tilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}}(U_n(\frac{i}{n}),\psi_{\frac{i}{n}})-\tilde{E}[(U_n(\frac{i}{n}),\psi_{\frac{i}{n}})\left|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}\right.]\right)^2\right]$$

$$\begin{aligned}
&= \widetilde{E}\left[\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\left((U_n(\frac{i}{n}), \psi_{\frac{i}{n}}) - \widetilde{E}[(U_n(\frac{i}{n}), \psi_{\frac{i}{n}}) \middle| \mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}]\right)^2 \middle| \mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right]\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\sum_{j=1}^{m_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))^2 \nu_n^{i-1}(V_n^j)\right] \\
&\leq \frac{BM(h)}{\sqrt{n}} \qquad (5.106)
\end{aligned}$$

where, for the last inequality we used lemma 6. Lastly, for the third term we have the identity

$$\widetilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}[(U_n(\frac{i}{n}), \psi_{\frac{i}{n}}) \middle| \mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}] - (U_n(\frac{i-1}{n}), \psi_{\frac{i-1}{n}})\right)^2\right]$$

$$\frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\sum_{j=1}^{m_n(\frac{i-1}{n})} \widetilde{E}\left[\left(\psi_{\frac{i}{n}}(V_n^j(\frac{i}{n}))\xi_n^{i-1}(V_n^j, \frac{i}{n}) - \psi_{\frac{i-1}{n}}(V_n^j(\frac{i-1}{n}))\right)^2 \middle| \mathcal{F}_{\frac{i-1}{n}} \vee \mathcal{Y}\right]\right]$$

and one finds the following upper bound for the L_2 norm of the third term

$$\frac{C}{n}. \qquad (5.107)$$

From (5.104), (5.105), (5.106) and (5.107) we get the required limit (5.103) and with it the proof of the proposition.

Proof of Theorem 27

We proved that any subsequence of $U_n(t)$ contains a subsubsequence $U_{n_k}(t)$ such that for $\varphi \in \mathcal{M} \cup \{1\}$ we have $\lim_{k\to\infty}(U_{n_k}(t), \varphi) = \rho_t(\varphi)$, $\widetilde{P}$-a.s. Hence, for any subsequence of $U_n(t)$ we can find a subsubsequence $U_{n_k}(t)$ such that *simultaneously for all* $\varphi \in \mathcal{M} \cup \{1\}$, $\lim_{k\to\infty}(U_{n_k}(t), \varphi) = \rho_t(\varphi)$, $\widetilde{P}$-a.s., i.e., there exists $\Omega' \subset \Omega$, with $\widetilde{P}(\Omega') = 1$ such that for all $\varphi \in \mathcal{M} \cup \{1\}$, we have

$$\lim_{k\to\infty}(U_{n_k}(t), \varphi) = \rho_t(\varphi),\ \forall \omega \in \Omega'$$

We prove now that the limit holds not only for $\varphi \in \mathcal{M} \cup \{1\}$, but for all $\varphi \in C_b(E)$ for all $\omega \in \Omega'$. The first step is to prove that the limit holds for all $\varphi \in C_0(E)$ For this, we fix an arbitrary $\varphi \in C_0(E)$ and $\omega \in \Omega'$. We show that for all $\epsilon > 0$, there exists $k_\epsilon(\omega)$, such that for all $k \geq k_\epsilon(\omega)$, we have

$$|(U_{n_k}(t)(\omega), \varphi) - \rho_t(\varphi)(\omega)| \leq \epsilon, \tag{5.108}$$

To simplify the notation we will omit from now on the dependence on ω. Since $\varphi \in C_0(E)$, there exists $\varphi' \in \mathcal{M}$, such that

$$\|\varphi - \varphi'\| \leq \frac{\epsilon}{2(\rho_t(1) + q)}$$

where $q \overset{\text{def}}{=} \sup_{k>0}(U_{n_k}(t), 1)$ (it is finite since the sequence is convergent to $\rho_t(1)$). Therefore, for all $k > 0$

$$|(U_{n_k}(t), \varphi) - (U_{n_k}(t), \varphi'| + |\rho_t(\varphi) - \rho_t(\varphi')| \leq \frac{\epsilon}{2} \tag{5.109}$$

Since the limit holds for φ', there exists k_ϵ, such that for all $k \geq k_\epsilon$, we have

$$|(U_{n_k}(t), \varphi') - \rho_t(\varphi')| \leq \frac{\epsilon}{2}, \tag{5.110}$$

Now (5.109) and (5.110) give us (5.108).With a further approximation, on can show that

$$\lim_{k \to \infty} (U_{n_k}(t), I_K) = \rho_t(I_K)$$

where I_K is the characteristic function of the arbitrary compact $K \in E$ (with the complement $CK \overset{\text{def}}{=} E \setminus K$) and therefore for any $\delta > 0$ there exists a compact K_δ such that

$$\sup_{k \to \infty} (U_{n_k}(t), I_{CK_\delta}) < \delta \tag{5.111}$$

Let now $\varphi \in C_b(E)$, $\delta = \frac{\epsilon}{3\|\varphi\|}$ and $\varphi'' \in \mathcal{M}$, such that

$$\|\varphi I_{K_\delta} - \varphi''\| \leq \frac{\epsilon}{3(\rho_t(1) + q)}$$

Therefore, for all $k > 0$

$$|(U_{n_k}(t), \varphi I_{K_\delta}) - (U_{n_k}(t), \varphi''| + |\rho_t(\varphi I_{K_\delta}) - \rho_t(\varphi'')| \leq \frac{\epsilon}{3} \tag{5.112}$$

Since the limit holds for φ'', there exists k_ϵ, such that for all $k \geq k_\epsilon$, we have

$$|(U_{n_k}(t), \varphi'') - \rho_t(\varphi'')| \leq \frac{\epsilon}{3}, \tag{5.113}$$

Now (5.111), (5.112) and (5.113) give us (5.108) for any $\varphi \in C_b(E)$, Q.E.D..

Proof of Corollary 4

Since we showed that U_n is tight we only need to show that the limit of any convergent (in distribution) subsequence of U_n is p. Let U_{n_k} be a convergent subsequence of U_n and u its limit. We need to show that the finite dimensional distributions of u coincide with the finite dimensional distribution of p, i.e., that

$$\tilde{E}[f_1((u(t_1), \varphi_1))f_2((u(t_2), \varphi_2))...f_k((u(t_k), \varphi_k))] = \tilde{E}[f_1(\rho_t(\varphi_1))f_2(\rho_t(\varphi_2))...f_k(\rho_t(\varphi_k))] \tag{5.114}$$

for all $f_1, f_2, ..., f_k \in C_b(\mathbb{R})$, $\varphi_1, \varphi_2, ..., \varphi_k \in C_b(E)$, $0 \leq t_1 \leq t_2 \leq ... \leq t_k \leq 1$, $k \in \mathbb{N}$. Obviously U_{n_k} contains a subsequence, which we also denote with U_{n_k}, whose limit at time t_i is ρ_t for all $i = 1, 2, ..., k$. Hence the Dominated Convergence Theorem gives us

$$\lim_{k\to\infty} \tilde{E}[f_1((U_{n_k}(t_1), \varphi_1))f_2((U_{n_k}(t_2), \varphi_2))...f_k((U_{n_k}(t_k), \varphi_k))]$$

$$= \tilde{E}[f_1(\rho_t(\varphi_1))f_2((\rho_t, \varphi_2))...f_k((\rho_t, \varphi_k))] \tag{5.115}$$

But, since U_{n_k} converges in distribution to u, we have that

$$\lim_{k\to\infty} \tilde{E}[f_1((U_{n_k}(t_1), \varphi_1))f_2((U_{n_k}(t_2), \varphi_2))...f_k((U_{n_k}(t_k), \varphi_k))]$$

$$= \tilde{E}[f_1((u(t_1), \varphi_1))f_2((u(t_2), \varphi_2))...f_k((u(t_k), \varphi_k))] \tag{5.116}$$

Finally (5.115) and (5.116) implies (5.114), Q.E.D.

Chapter 6

A Measure Valued Process Associated with the Kushner-Stratonovitch Equation

6.1 Assumptions

In this chapter we construct a sequence of branching particle systems α_n convergent in distribution to the conditional distribution of the signal given the observation - π_t - , i.e., to the solution of the Kushner-Stratonovitch equation. Again, the algorithm based on this result can be used to solve numerically the filtering problem. The nonlinear character of the equation makes direct numerical work difficult. Using the branching particle systems α_n, we produce a sample approximation of the conditional distribution of X and, based on this, we can construct a numerical algorithm suitable for high dimensional problems (the existing algorithms can only be used in low dimensions). We will construct the particle systems and, implicitly, the measure valued process α, up to a fixed horizon, in fact on the fixed interval $[0,1]$, the construction being identical for any interval $[0,T]$. Then, using an argument based on the Carathèodory extension theorem, the construction can be extended to the whole positive axis.

Let A_0 be the restriction of A to $C_0(E)$. We assume that the domain of A_0 is a *dense algebra* in $C_0(E)$ (in particular that if $f \in \mathcal{D}(A_0)$, then $f^2 \in \mathcal{D}(A_0)$).

Although we consider A and h to be time independent all the results remain valid in the time-dependent case (which can be viewed as a time independent case with the signal process (X_t, t)).

6.2 The Branching Particle Systems α_n

Let $\{\alpha_n(t), \mathcal{F}_t;\ 0 \le t \le 1\}$ be a sequence of branching particle systems on $(\Omega, \mathcal{F}, P)$ with values in $M_F(E)$ defined as follows:

a. The initial state of the systems, $\alpha_n(0)$, is the occupation measure of n particles of mass $\frac{1}{n}$, i.e., $\alpha_n(0) = \frac{1}{n}\sum_{i=1}^n \delta_{x_i^n}$, ($x_i^n \in E$ for all $n > 0$ and $1 \le i \le n$) and is weakly convergent to π_0.

b. At the time $\frac{i}{n}$, $i = 0, 1, \ldots, n-1$, the process consists of the occupation measure of $a_n(\frac{i}{n})$ particles of mass $\frac{1}{n}$ (we will denote the number of particles alive at time t by $a_n(t)$).

c. During the interval $[\frac{i}{n}, \frac{i+1}{n})$ the particles move independently with the same law as the signal X. Let $b_j(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$, $1 \le j \le a_n(\frac{i}{n})$ be the trajectories of the $a_n(\frac{i}{n})$ particles alive in this interval (the number of particles remains constant during the interval).

d. At the end of the interval, the particles branch into a random number of offspring with a mechanism depending on the recent past of the whole system.

e. The mechanism is chosen so that it has finite second moment and the mean number of offspring of the j-th particle given the σ-field $\mathcal{F}_{\frac{i+1}{n}-} \overset{\text{def}}{=} \sigma(\mathcal{F}_s,\ s < \frac{i+1}{n})$ of events up to time $\frac{i+1}{n}$ is

$$\gamma_n^i(b_j) \overset{\text{def}}{=} \frac{a_n(\frac{i}{n})\beta_n^i(b_j)}{\sum_{k=1}^{a_n(\frac{i}{n})} \beta_n^i(b_k)} \tag{6.1}$$

where

$$\beta_n^i(b_j) \overset{\text{def}}{=} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(b_j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(b_j(t))dt\right) \tag{6.2}$$

and the minimal variance $\delta_n^i(b_j)$ consistent with the number of offspring being an integer. The particles branch independently of each other, given $\mathcal{F}_{\frac{i+1}{n}-}$.

Remark 22 *The construction works also for particles which do not branch independenly.*

We observe that $\delta_n^i(b_j) = (\gamma_n^i(b_j) - [\gamma_n^i(b_j)])([\gamma_n^i(b_j)] + 1 - \gamma_n^i(b_j))$ and is always less than $\frac{1}{4}$ ($[x]$ is the largest integer smaller that x). Just before the $(i+1)$-th branching, we will have $a_n^i \stackrel{\text{def}}{=} a_n(\frac{i}{n})$ particles. Let us denote by $\alpha_n(\frac{i+1}{n}-)$ the state of the process just before the $(i+1)$-th branching and by $b_j^{n,k}(s)$, $s \in [\frac{k}{n}, \frac{k+1}{n})$, $j = 1, \dots, a_n^k$ the processes which represents the trajectories of the particles alive during the interval $[\frac{k}{n}, \frac{k+1}{n})$ corresponding to the n-th particle system. As before, we will identify the particles with their path, i.e., the particle $b_j^{n,k}$ is the particle whose path is the process $b_j^{n,k}(s)$, $s \in [\frac{k}{n}, \frac{k+1}{n})$ and vice versa.

Remark 23 *The branching systems α_n will never die and will never explode.*

Proof. Changes in the number of particles take place only at branching times. Due to the fact that the variance of the branching mechanism is the minimal one, the number of offspring for the j-th particle in the interval $[\frac{i}{n}\frac{i+1}{n}]$ is either $[\gamma_n^i(b_j^{n,k})]$, or $[\gamma_n^i(b_j^{n,k})] + 1$. So

$$a_n^{i+1} \leq \sum_{j=1}^{a_n^i}([\gamma_n^i(b_j^{n,k})] + 1) \leq \sum_{j=1}^{a_n^i}(\gamma_n^i(b_j^{n,k}) + 1) \leq 2a_n^i \tag{6.3}$$

Also, since $\sum_{j=1}^{a_n^i} \gamma_n^i(b_j^{n,k}) = a_n^i$, at least one γ_n^i is larger than or equal to 1, hence at least one particle has offspring.

■

Proposition 19 *The process $\{a_n^i, \mathcal{F}_{\frac{i+1}{n}-}; i = 0, \dots, n\}$ is a bounded martingale and $E[(a_n^i)^2] < 2n^2$.*

Proof. From the construction of the particle systems we have that

$$E[a_n^{i+1} | \mathcal{F}_{\frac{i+1}{n}-}] = \sum_{j=1}^{a_n^i} \frac{a_n^i \beta_n^i(b_j^{n,i})}{\sum_{j=1}^{a_n^i} \beta_n^i(b_j^{n,i})} = a_n^i, \quad i = 0, .., n-1$$

and the inequality (6.3) implies that $a_n^i \leq 2^i n$ for $0 \leq i \leq n$. Since the branching mechanisms are independent given $\mathcal{F}_{\frac{i+1}{n}-}$, we obtain

$$E[(a_n^i)^2] = E[< a_n^i >] + n^2 = E[\sum_{k=1}^{i} E\left[(a_n^k - E[a_n^k | \mathcal{F}_{\frac{k}{n}-}])^2 | \mathcal{F}_{\frac{k}{n}-}\right]] + n^2$$

$$= E\left[\sum_{k=1}^{i}\sum_{j=1}^{a_n^k}\delta_n^k(b_j^{n,k})\right] + n^2 \leq \frac{1}{4}\sum_{k=1}^{i} E[a_n^k] + n^2 < 2n^2,$$

which completes the proof of the proposition.

■

Corollary 5 *The mass processes* $(\alpha_n(t), 1)$ *form a tight sequence.*

Proof. We have $(\alpha_n(t), 1) = \frac{1}{n}a_n^{[nt]}$ so

$$P(\sup_{t\leq T}(\alpha_n(t), 1) \geq k) = P(\sup_{i\leq[nT]} a_n^i \geq nk) \leq \frac{1}{nk}E[a_n^{[nT]}]$$

and the last term is actually equal to $\frac{1}{k}$. Hence

$$\lim_{k\to\infty}\sup_{n\geq 0} P(\sup_{t\leq T}(\alpha_n(t), 1) \geq k) = 0$$

which proves our claim.

■

Remark 24 *For any* $t \in [0,1]$ *and* $\varphi \in B(E)$, *the processes* $(\alpha_n(t), \varphi)$ *are bounded and uniformly square integrable.*

Proof. Straightforward from Proposition 19.

■

Proposition 20 *If* $\varphi, \varphi^2 \in \mathcal{D}(A)$, *then the process* $(\alpha_n(t), \varphi)$ *satisfies the following evolution equation*

$$\begin{aligned}(\alpha_n(t), \varphi) \;&=\; (\alpha_n(0), \varphi) + \int_0^t (\alpha_n(s), A\varphi)ds + S_n^{\varphi}(t) + M_n^{\varphi}([nt]) \\ &+ \sum_{i=1}^{[nt]}\frac{1}{n}\sum_{j=1}^{a_n(\frac{i-1}{n})}\varphi(b_j^{n,i-1}(\frac{i}{n}))(\gamma_n^{i-1}(b_j^{n,i-1}) - 1). \qquad (6.4)\end{aligned}$$

where $\{(S_n^{\varphi}(t),\ \mathcal{F}_t),\ t \in [0,1]\}$ *is a bounded martingale with quadratic variation*

$$< S_n^{\varphi} > (t) = \frac{1}{n}\int_0^t (\alpha_n(s), A\varphi^2 - 2\varphi A\varphi)ds \qquad (6.5)$$

and $\{(M_n^{\varphi}(l),\ \mathcal{F}_{\frac{l+1}{n}-}),\ l = 0, 1, ..., n\}$ *is a discrete martingale with quadratic variation*

$$< M_n^{\varphi} > (l) = \frac{1}{n}\sum_{i=1}^{l}(\alpha_n(\frac{i+1}{n}-), \delta_n^i\varphi^2) \tag{6.6}$$

Proof. From the construction of the particle systems we have that

$$E[(\alpha_n(\frac{i+1}{n}), \varphi)|\mathcal{F}_{\frac{i+1}{n}-}] = \frac{1}{n}\sum_{i=1}^{a_n(\frac{i}{n})} \varphi(b_n^j(\frac{i+1}{n}))\gamma_n^i(b_n^j) \tag{6.7}$$

and also

$$E[(\alpha_n(\frac{i+1}{n}), \varphi)^2|\mathcal{F}_{\frac{i+1}{n}-}] - (E[(\alpha_n(\frac{i+1}{n}), \varphi)|\mathcal{F}_{\frac{i+1}{n}-}])^2 = \frac{1}{n}(\alpha_n(\frac{i+1}{n}-), \delta_n^i\varphi^2). \tag{6.8}$$

In between two branches the particles move according to the prescribed law, hence for t in the interval $[\frac{i}{n}, \frac{i+1}{n})$

$$(\alpha_n(t), \varphi) = (\alpha_n(\frac{i}{n}), \varphi) + \int_{\frac{i}{n}}^{t}(\alpha_n(s), A\varphi)ds + S_n^{\varphi,i}(t), \tag{6.9}$$

where $\{(S_n^{\varphi,i}(t), \mathcal{F}_t),\ t \in [\frac{i}{n}, \frac{i+1}{n}]\}$ is a bounded martingale (we use proposition 19) with the quadratic variation

$$< S_n^{\varphi,i} > (t) = \frac{1}{n}\int_{\frac{i}{n}}^{t}(\alpha_n(s), A\varphi^2 - 2\varphi A\varphi)ds. \tag{6.10}$$

In order to compute the evolution equation of $(\alpha_n(t), \varphi)$, we need to add all the parts coming from the particles motion and all the parts coming from the particle branching, which gives

$$(\alpha_n(t), \varphi) = (\alpha_n(0), \varphi) + \int_0^t(\alpha_n(s), A\varphi)ds + S_n^{\varphi}(t) + \sum_{i=1}^{[nt]}\Big((\alpha_n(\frac{i}{n}), \varphi) - (\alpha_n(\frac{i}{n}-), \varphi)\Big)$$

where $\{(S_n^{\varphi}(t),\ \mathcal{F}_t),\ t \in [0, 1]\}$ is the bounded martingale

$$S_n^{\varphi}(t) \triangleq S_n^{\varphi,[nt]}(t) + \sum_{i=0}^{[nt]-1} S_n^{\varphi,i}(\frac{i+1}{n})$$

with the quadratic variation presented in (6.5). We then split the term coming from the branching into a martingale part and a bounded variation part and obtain

$$\begin{aligned}(\alpha_n(t),\varphi) &= (\alpha_n(0),\varphi) + \int_0^t (\alpha_n(s), A\varphi)ds + S_n^\varphi(t) + M_n^\varphi([nt]) \\ &\quad + \sum_{i=1}^{[nt]} \left(\bar{E}[(\alpha_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] - (\alpha_n(\frac{i}{n}-),\varphi) \right), \end{aligned} \tag{6.11}$$

and $\{(M_n^\varphi(l), \mathcal{F}_{\frac{l+1}{n}-}),\ l = 0,1,...,n\}$ is the square integrable martingale

$$M_n^\varphi(0) \triangleq 0,$$

$$M_n^\varphi(l) \triangleq \sum_{i=1}^{l} \left((\alpha_n(\frac{i}{n}),\varphi) - \bar{E}[(\alpha_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] \right).$$

with quadratic variation

$$< M_n^\varphi > (l) = \sum_{i=1}^{l} \bar{E}\left[\left((\alpha_n(\frac{i}{n}),\varphi) - E[(\alpha_n(\frac{i}{n}),\varphi)|\mathcal{F}_{\frac{i}{n}-}] \right)^2 |\mathcal{F}_{\frac{i}{n}-} \right] \tag{6.12}$$

The proposition follows now easily from (6.7), (5.72), (6.11) and (6.12).

■

We introduce also the notation

$$l_n^i(b_j^{n,i},r) \triangleq \exp\left(\int_{\frac{i}{n}}^r h^*(b_j^{n,i}(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^r h^*h(b_j^{n,i}(t))dt \right), \quad r \in [\frac{i}{n}, \frac{i+1}{n}]$$

$$\mu_n^i(b_j^{n,i},r) \triangleq \frac{l_n^i(b_j^{n,i},r)}{\sum_{k=1}^{a_n^i} l_n^i(b_k^{n,i},r)}, \quad r \in [\frac{i}{n}, \frac{i+1}{n}]$$

Using Itô calculus, we get that

$$\begin{aligned}\gamma_n^i(b_j^{n,i}) &= 1 + \int_{\frac{i}{n}}^{\frac{i+1}{n}} a_n^i \mu_n^i(b_j^{n,i},r) \left(h^*(b_j^{n,i}(r)) - (\sum_{k=1}^{a_n^i} \mu_n^i(b_k^{n,i},r) h^*(b_k^{n,i}(r))) \right) \\ &\quad \times \left(dY_r - (\sum_{k=1}^{a_n^i} \mu_n^i(b_k^{n,i},r) h(b_k^{n,i}(r)))dr \right) \end{aligned} \tag{6.13}$$

6.3 The Tightness of the Sequence α_n

To prove the tightness we need to show that the sequence α_n satisfies the conditions of Theorem 21, i.e., to prove that the mass processes satisfy the compact containment condition (which we did in Corollary 5) and that we can control the oscillations of the processes $\{\alpha_n\}_{n>0}$. For this, we prove that, for a dense sequence $\{\varphi_i\}_{i\geq 0} \subset C_0(E)$ (we will take them to be in the domain of A_0), any arbitrary sequence of stopping times $\{\tau_n\}_{n\geq 0}$ and any real positive sequence $\{\rho_n\}_{n\geq 0}$ with $\lim_{n\to\infty} \rho_n = 0$, we have

$$\lim_{n\to\infty} P(|(\alpha_n(\tau_n+\rho_n),\varphi_i) - (\alpha_n(\tau_n),\varphi_i)| \geq \epsilon) = 0 \tag{6.14}$$

As before, we have the following result.

Lemma 7 *For all $i = 0, \ldots, n-1$, we have*

$$\tilde{E}\left[\frac{1}{n}\sum_{j=1}^{a_n^i} \delta_n^i(V_n^j)\right] \leq \frac{M(h)}{\sqrt{n}} \tag{6.15}$$

where $M(h)$ is a constant depending only on h.

Proof. Similar to Lemma 6 from the (6.13) and the fact that if we have an integer random variable with mean μ and minimal variance ν, then $\nu \leq |\mu - 1|$.

■

Corollary 6 *For all $n > 0$ and $t \in [0,1]$, we have*

$$\tilde{E}\left[((\alpha_n(t),1)-1)^2\right] \leq \frac{M(h)}{\sqrt{n}} \tag{6.16}$$

Proof. Straightforward from (6.15)

■

Corollary 7 *For all $n > 0$ we have*

$$\tilde{E}\left[\left(\sup_{t\in[0,1]} (\alpha_n(t),1)\right)^2\right] \leq 4 + \frac{4M(h)}{\sqrt{n}} \tag{6.17}$$

Proof. Straightforward from (6.16) and Doob's maximal inequality.

■

The limit (6.14) follows from the next proposition.

Proposition 21 *For any arbitrary sequence of stopping times* $\{\tau_n\}_{n\geq 0}$*, any real positive sequence* $\{\rho_n\}_{n\geq 0}$ *with* $\lim_{n\to\infty}\rho_n = 0$ *and* $\varphi \in D(A_0)$*, we have*

$$\lim_{n\to\infty} E[|(\alpha_n(\tau_n+\rho_n),\varphi)-(\alpha_n(\tau_n),\varphi)|]=0. \tag{6.18}$$

Proof. Using (6.4), we get

$$\begin{aligned} E[|(\alpha_n(\tau_n+\rho_n),\varphi)-(\alpha_n(\tau_n),\varphi)|] &\leq E[|\int_{\tau_n}^{\tau_n+\rho_n}(\alpha_n(s),A\varphi)ds|] \\ &+E[|S_n^{\varphi}(\tau_n+\rho_n)-S_n^{\varphi}(\tau_n)|]+E[|M_n^{\varphi}([n(\tau_n+\rho_n)])-M_n^{\varphi}([n\tau_n])|] \\ &+E\left[\left|\sum_{i=[n\tau_n]}^{[n(\tau_n+\rho_n)]}\frac{1}{n}\sum_{j=1}^{a_n^{i-1}}\varphi(b_j^{n,i-1}(\frac{i}{n}))(\gamma_n^{i-1}(b_j^{n,i-1})-1)\right|\right] \end{aligned} \tag{6.19}$$

Since the first three terms of the right hand side of (6.19) can be easily shown to converge to 0, we only prove this for the last one. Based on the identity (6.13), we have the following

$$\begin{aligned} &\sum_{i=[n\tau_n]}^{[n(\tau_n+\rho_n)]}\frac{1}{n}\sum_{j=1}^{a_n^{i-1}}\varphi(b_j^{n,i-1}(\frac{i}{n}))(\gamma_n^{i-1}(b_j^{n,i-1})-1) \\ &= \int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\rho_n)]}{n}}\frac{1}{n}\sum_{j=1}^{a_n^{[nr]}}\varphi(b_j^{n,[nr]}(\frac{[rn]}{n}))a_n^{[nr]}\mu_n^{[nr]}(b_j^{n,[nr]},r) \\ &\quad\times(h^*(b_j^{n,[nr]}(r)-(\sum_{k=1}^{a_n^{[nr]}}\mu_n^{[nr]}(b_k^{n,[nr]},r)h^*(b_k^{n,[nr]}(r)))) \\ &\quad\times(dY_r-(\sum_{k=1}^{a_n^{[nr]}}\mu_n^{[nr]}(b_k^{n,[nr]},r)h(b_k^{n,[nr]}(r)))dr) \end{aligned} \tag{6.20}$$

By taking the modulus then the expectation and using the fact that

$$\sum_{k=1}^{a_n^{[nr]}}\mu_n^{[nr]}(b_k^{n,[nr]},r)=1,$$

we obtain from (6.20)

$$E\left[\left|\sum_{i=[n\tau_n]}^{[n(\tau_n+\rho_n)]}\frac{1}{n}\sum_{j=1}^{a_n(\frac{i-1}{n})}\varphi(b_j^{n,i-1}(\frac{i}{n}))(\gamma_n^{i-1}(b_j^{n,i-1})-1)\right|\right]$$

$$\begin{aligned}
&\leq 2\|\varphi\|\|h\|\sqrt{\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\rho_n)]}{n}}\frac{E[(a_n^{[nr]})^2]}{n^2}dr}+2\|\varphi\|\|h\|^2\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\rho_n)]}{n}}\frac{E[a_n^{[nr]}]}{n}dr \\
&\leq 2\|\varphi\|\|h\|\left(\sqrt{\left(1+\frac{M(h)}{\sqrt{n}}\right)\frac{[n(\tau_n+\rho_n)]-[n\tau_n]}{n}}+\|h\|\frac{[n(\tau_n+\rho_n)]-[n\tau_n]}{n}\right)
\end{aligned}$$

which proves that also the last term in (6.19) converges to zero (it is of order $\sqrt{\max(\rho_n, 1/n)}$).

■

6.4 The Convergence of the Sequence to the Solution of K-S Equation

This section follows the steps of the analogoes one for the sequence U_n. Here are results:

Proposition 22 *Every subsequence of $\alpha_n(t)$ contains a subsubsequence $\alpha_{n_k}(t)$ such that for $\varphi \in \mathcal{M} \cup \{1\}$ we have*

$$\lim_{k\to\infty}(\alpha_{n_k}(t), \varphi) = \pi_t(\varphi),\ P-a.s., \tag{6.21}$$

where $\mathcal{M}$ is a countable and uniformly dense set in $C_K(E)$ (the space of continuous functions with compact support).

Theorem 28 *Every subsequence of $\alpha_n(t)$ contains a subsubsequence $\alpha_{n_k}(t)$ convergent P-a.s. to the conditional distribution of the signal π_t*

Corollary 8 *The sequence $\alpha_n(t)$ is convergent in measure to π_t, i.e., to the law of $X(t)$ given the observation σ-field $\mathcal{Y}_t$.*

Corollary 9 *The sequence α_n is convergent in distribution to the measure valued process that represents the conditional law of X given the observation Y.*

Theorem 28, Corollary 8 and Corollary 9 have similar proofs to their analogous in the previous chapter, so we will prove only Proposition 22.

Proof of Proposition 22

The idea of the proof is to 'unnormalise' $\alpha_n(t)$ by multiplying it by a suitably chosen 'mass' process $\xi_n(t)$. We then prove that convenient subsequences of $\xi_n\alpha_n(t)$ converge to the unnormalised conditional distribution of the signal ρ_t and simultaneously the mass of $\alpha_n(t)$ converges to 1. Hence $\alpha_n(t)$ converges to ρ_t normalised, i.e., to π_t.

Again we use the solution of the backward stochastic partial differential equation (5.94) and assume that, as in the previous Chapter, the condition **U** holds.

U. For all $t \in [0,1]$, there exists a countable set $\mathcal{M}$, uniformly dense in $C_K(E)$, such that for all $\varphi \in \mathcal{M}$ the SPDE (5.94) has a solution $\psi_s \in \mathcal{D}(A)$ .which satisfies

$$B \overset{\text{def}}{=} \tilde{E}\left[\sup_{s\in[0,t]} \|\psi_s\|^2\right] < \infty. \tag{6.22}$$

and

$$C \overset{\text{def}}{=} \tilde{E}\left[\sup_{s\in[0,t]} \|A\psi_s^2 - 2\psi_s A\psi_s\|\right] < \infty. \tag{6.23}$$

Obviously, the set $\mathcal{M}$ from above will be the set for which we will prove the limit (6.21). From 6.22 and 5.100, we have that

$$\widetilde{E}\left[\left(\psi_{\frac{i+1}{n}}\left(b_j^{n,i}\left(\frac{i+1}{n}\right)\right)\beta_n^i(b_j^{n,i}) - \psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i}{n}\right)\right)\right)^2 \middle| \mathcal{F}_{\frac{i}{n}}\right] \leq \frac{C}{n} \tag{6.24}$$

for $j = 1, ..., a_n^i$ and $i = 0, ..., n-1$ and

$$\widetilde{E}\left[\left(\psi_t\left(b_j^{n,i}(t)\right) - \psi_{\frac{[nt]}{n}}\left(b_j^{n,[nt]}\left(\frac{[nt]}{n}\right)\right)\right)^2 \middle| \mathcal{F}_{\frac{[nt]}{n}}\right] \leq \frac{BM'(h)+C}{n} \tag{6.25}$$

for $j = 1, ..., a_n^{[nt]}$ and $t \in [0,1]$. Let now $\xi_n(t) = \prod_{i=0}^{[nt]} \sum_{k=1}^{a_n^i} \frac{\beta_n^i(b_k)}{a_n^i}$. The process $\{\xi_n(t), \mathcal{F}_t; t \in [0,1]\}$ is a positive martingale with mean 1. We have

$$\sum_{k=1}^{a_n^i} \frac{l_n^i(b_j^{n,i}, r)}{a_n^i} = 1 + \int_{\frac{i}{n}}^{r} \sum_{k=1}^{a_n^i} \frac{l_n^i(b_j^{n,i}, r)}{a_n^i} h^*(b_j^{n,i}(r))dY_r \tag{6.26}$$

hence for $p \geq 2$, we get from (6.26), using Burkholder's inequality and Gronwall's inequality

$$\tilde{E}\left[\left(\sum_{k=1}^{a_n^i}\frac{l_n^i(b_j^{n,i},r)}{a_n^i}\right)^p\right] \leq e^{K_p\|h\|^p(r-\frac{i}{n})}$$

hence

$$\tilde{E}\left[(\xi_n(t))^p\right] = \prod_{i=0}^{[nt]}\tilde{E}\left[\left(\sum_{k=1}^{a_n^i}\frac{l_n^i\left(b_j^{n,i},\frac{i+1}{n}\right)}{a_n^i}\right)^p\right] \leq e^{K_p\|h\|^p}\ \forall t\in[0,1] \tag{6.27}$$

We want to prove now that, at least for a subsequence of α_n

$$\lim_{n\to\infty}(\xi_n(t)\alpha_n(t),\varphi) = \rho_t(\varphi), \quad \tilde{P}-a.s. \tag{6.28}$$

As in the proof of Proposition 18, we show that

$$\lim_{n\to\infty}\widetilde{E}[((\xi_n\alpha_n(t),\psi_t)-(\xi_n\alpha_n(0),\psi_0))^2] = 0 \tag{6.29}$$

Since ξ_n is changing only at branching times and is constant otherwise, the inequality (6.25) implies that

$$\lim_{n\to\infty}\widetilde{E}[((\xi_n\alpha_n(t),\psi_t)-(\xi_n\alpha_n(\frac{[nt]}{n}),\psi_{\frac{[nt]}{n}}))^2] = 0 \tag{6.30}$$

As in (5.104), we have

$$\begin{aligned}(\xi_n\alpha_n(\frac{[nt]}{n}),\psi_{\frac{[nt]}{n}})-(\xi_n\alpha_n(0),\psi_0) \;=\; & \sum_{i=1}^{\frac{[nt]}{n}}\Bigg((\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}}) \\ & -\widetilde{E}[(\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}})\Big|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}]\Bigg) \\ & +\sum_{i=1}^{\frac{[nt]}{n}}\Bigg(\widetilde{E}[(\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}})\Big|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}] \\ & -(U_n(\frac{i-1}{n}),\psi_{\frac{i-1}{n}})\Bigg)\end{aligned} \tag{6.31}$$

Now

$$\tilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}}(\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}})-\widetilde{E}[(\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}})\Big|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}]\right)^2\right]$$

$$\begin{aligned}
&= \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}}\tilde{E}\left[\left(\xi_n\left(\frac{i}{n}\right)\right)^2\sum_{j=1}^{a_n^i}\left(\psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i}{n}\right)\right)\right)^2\delta_n^i(b_j^{n,i})\right]\\
&\leq \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}}\tilde{E}\left[\sup_{x\in E}|\psi_{\frac{i}{n}}(x)|^2\left(\xi_n\left(\frac{i}{n}\right)\right)^2\sum_{j=1}^{a_n^i}\delta_n^i(b_j^{n,i})\right]\\
&= \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}}\tilde{E}\left[\sup_{x\in E}|\psi_{\frac{i}{n}}(x)|^2\right]\tilde{E}\left[\left(\xi_n\left(\frac{i}{n}\right)\right)^2\sum_{j=1}^{a_n^i}\delta_n^i(b_j^{n,i})\right]\\
&\leq \frac{B}{n^2}\sum_{i=1}^{\frac{[nt]}{n}}\sqrt{\tilde{E}\left[\left(\xi_n\left(\frac{i}{n}\right)\right)^4\right]\tilde{E}\left[a_n^i\sum_{j=1}^{a_n^i}\left(\delta_n^i(b_j^{n,i})\right)^2\right]}\\
&\leq \frac{BM''(h)}{\sqrt{n}}
\end{aligned}$$

hence the first term from the right hand side of (6.31) is of order $\frac{1}{\sqrt{n}}$. Also we have

$$\tilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}}\widetilde{E}[(\xi_n\alpha_n(\frac{i}{n}),\psi_{\frac{i}{n}})\Big|\mathcal{F}_{\frac{i}{n}-}\vee\mathcal{Y}]-(U_n(\frac{i-1}{n}),\psi_{\frac{i-1}{n}})\right)^2\right]$$

$$\begin{aligned}
&= \frac{1}{n^2}\tilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}}\sum_{j=1}^{a_n^i}\xi_n\left(\frac{i}{n}\right)\psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i}{n}\right)\right)\gamma_n^i(b_j^{n,i})-\xi_n\left(\frac{i-1}{n}\right)\psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i-1}{n}\right)\right)\right)^2\right]\\
&= \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}}\tilde{E}\left[\left(\xi_n\left(\frac{i-1}{n}\right)\right)^2\sum_{j=1}^{a_n^i}\left(\psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i}{n}\right)\right)\beta_n^i(b_j^{n,i})-\psi_{\frac{i}{n}}\left(b_j^{n,i}\left(\frac{i-1}{n}\right)\right)\right)^2\right]
\end{aligned}$$

and using (6.24) we get that the second term from the right hand side of (6.31) is of order $\frac{1}{n}$, hence

$$\lim_{n\to\infty}\widetilde{E}[((\xi_n\alpha_n(\frac{[nt]}{n}),\psi_{\frac{[nt]}{n}})-(\xi_n\alpha_n(0),\psi_0))^2]=0 \qquad (6.32)$$

Finally (6.30) and (6.32) imply (6.29). We now know that out of any subsequence of $\xi_n \alpha_n(t)$ we can obtain a subsubsequence such that (6.28) holds. Furthermore, since we know that $\lim_{n\to\infty}(\alpha_n(t), 1) = 1$ in L_2 sense, we can find a subsequence such that simultaneously (6.28) and $\lim_{n\to\infty}(\alpha_n(t), 1) = 1$, $\tilde{P} - a.s.$ hold. For this subsequence we have that

$$\lim_{n\to\infty} \xi_n(t) = \lim_{n\to\infty} \xi_n(t)(\alpha_n(t), 1) = \rho_t(1) \quad \tilde{P} - a.s.$$

and hence

$$\lim_{n\to\infty} \alpha_n(t) = \frac{\rho_t}{\rho_t(1)} = \pi_t \tag{6.33}$$

and the limit holds $\tilde{P}$ -a.s. and hence also P-a.s., since the two measures are absolutely continuous with respect to each other.

The process associated with the Kushner-Stratonovitch equation can be viewed as a genetic process with infinitely-many-alleles. Its space of "types" is E and the two phenomena occurring are random genetic drift (governed by the infinitesimal generator A) and reproduction. There is no mutation or selection involved. The individuals in this process in generation k will branch with a mean depending on their path in the "type"-space and the observed path and with the minimal possible variance.

Chapter 7

Some Numerical Simulations

7.1 General Remarks

Let us first review what we have achieved in the last two chapters. We constructed a measure valued process $\{\mathcal{X}(t), t \in [0,1]\}$, whose conditional expectation given the observation path is exactly ρ, the unnormalised conditional distribution of the signal. We proved that the conditional mean of $\mathcal{X}_n(t)$ - the particle systems used to construct $\mathcal{X}(t)$ - converges to ρ_t. We then constructed a sequence of branching particle systems $\{U_n(t), t \in [0,1]\}$ almost surely convergent directly (without the need of taking the conditional expectation) to the unnormalised conditional distribution of the signal. Lastly, we constructed $\{\alpha_n(t), t \in [0,1]\}$ a sequence of branching particle systems almost surely convergent to $\{\pi_t; t \in [0,1]\}$ - the (normalised) conditional distribution of the signal within the time frame $[0,1]$.

Each of the three constructions can be used to solve numerically the filtering problem. Jessica Gaines from the University of Edinburgh produced computer programs for each of the three. However, as one can expect, the algorithm based on the first construction performs far less well than the other two, since as pointed out in Chapter 5, one needs to take a larger number particles initially, therefore increasing the number of computation needed . In the first algorithm (construction) we fixed the variance of the branching mechanism beforehand, in the next two we took the minimal one which allowed the existence of an integer valued branching mechanism. This way we introduced extra randomness in the system, which leads to poorer performance.

The difference between the second and the third construction is that, while in the second construction we normalise only at the end (all the time, we approximate the unnormalised conditional distribution of the signal), in the third one we normalise (almost) as we go (all the time, we approximate the normalised conditional distribution of the signal). Therefore, in the third algorithm the number of particles is kept under control. If one uses the second algorithm, the population might explode or die out. Apart from this, there is no difference from a numerical point of view between the two algorithms.

In the following we present two examples solved with the second algorithm. The metric space will be $E \equiv \mathbb{R}^d$ with $d = 1$ in the first case and $d = 6$ in the second case. The signal/observation process will satisfy the following system of stochastic differential equations

$$\begin{aligned} dX_t &= f(X_t)dt + g(X_t)dW_1(t) \\ dY_t &= h(X_t)dt + dW_2(t) \end{aligned}$$

where $\{(W_1(t), W_2(t)); t \geq 0\}$ is a standard Brownian motion.

We run the algorithm up to a final time T. As in the description of the algorithm we divide the time interval into n equal intervals $\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]$, $k = 0, 1, ..., n-1$. At time $t = 0$, we generate n particles randomly with distribution $\pi(0)$. Then during an arbitrary interval the particles undergo two stages. In the first stage, from time $\frac{kT}{n}$ up to the final time $\frac{(k+1)T}{n}$, each particle moves along a trajectory, determined by numerical solution of (5.2), using independent simulations of a Brownian path for each particle. In the second stage, at time $\frac{(k+1)T}{n}$, each particle is replaced by a number of offspring, with the mean number of offspring being determined by the trajectory of the particle during stage one, as given by equation (5.4) and minimal variance.

7.2 The Numerical Examples

In this section we present two numerical examples done by Jessica Gaines (Department of Mathematics, University of Edinburgh). The first example consists of a one-dimensional signal, $x(t)$, and a one-dimensional observation, $y(t)$, given by

$$\begin{aligned} dx_t &= -\alpha x \; dt + \sigma dw_1(t) \\ dy_t &= \arctan(x)dt + dw_2(t) \end{aligned}$$

where $w_1(t)$ and $w_2(t)$ are independent one-dimensional standard Brownian motions. The parameter values used for the figures below are $\alpha = 1$ and $\sigma = 0.25$. The distribution of $x(0)$ was taken as normal with mean 1 and variance 0.25 and filtering was carried out from $t = 0$ until $t = 5$.

In Figure 1 below, we show the historical process for a simulation with 20 particles at the initial time. The past is shown only for particles alive at the final time. The signal is shown in Figure 2, along with the quartiles of the distribution of particles, for a simulation starting with 160 particles. In both pictures the simulation time has been divided into 160 generations and the time step used to calculate the trajectories of the particles is $h = 2^{-8}$.

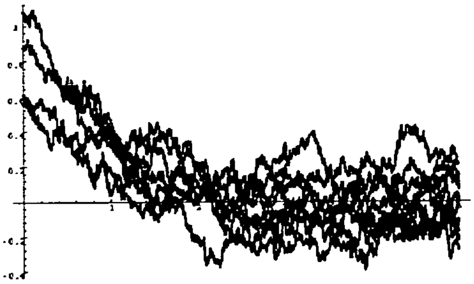

Figure 1: The historical process

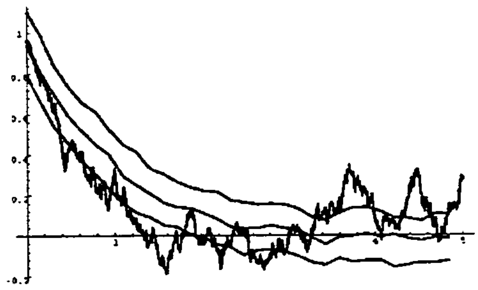

Figure 2: The signal and the quartiles

Figure 3 compares the expected mean of the signal calculated by numerical solution of the Zakai equation on the one hand and by the branching particle system on the other. The curve corresponding to the Zakai equation is lower at the final time than the other. The conditional densities of the signal at various times as calculated by solving the Zakai equation are shown in Figure 4 . The graphs progress with time from the right to the left of the picture.

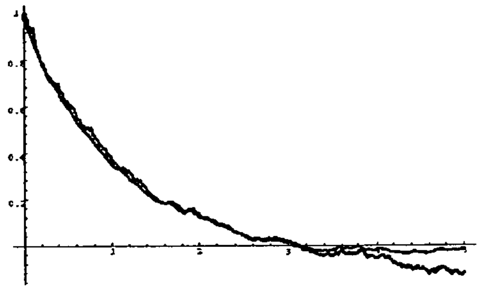

Figure 3: The conditional mean

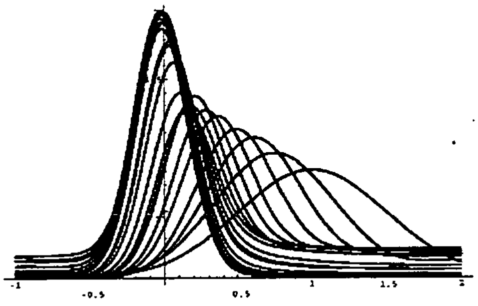

Figure 4: The conditional density

In the second example, we have a six-dimensional signal, representing the position and velocity of an object moving in $\mathbb{R}^3$ and a four-dimensional observation, consisting of angles measured by observers at two different positions. We suppose that the object starts at time $t_0 = 0$ from an initial position and with an initial velocity both drawn from normal distributions. The observers do not start measuring until time $t_1 > t_0$. The velocity of the object is subject to white noise. The observations also include noise. The signal $(x(t), v(t))$, where $x, v \in \mathbb{R}^3$ are the position and velocity of the object, is therefore given by

$$\begin{aligned} dx(t) &= v\, dt \\ dv(t) &= -Av\; dt + B\; dw(t) \end{aligned}$$

where $w(t)$ is standard Brownian motion in $\mathbb{R}^3$. A and B are diagonal matrixes with constant entries and $A(1,1) = A(2,2) = \epsilon$, $A(3,3) = g + \epsilon$, where g is the gravitational constant. The observation vector, $y(t) \in \mathbb{R}^4$ is defined by

$$\begin{aligned} dy_1(t) &= \alpha_1 \left(\arctan \frac{x_3 - p_3}{x_1 - p_1} \right) dt + d\bar{w}_1(t) \\ dy_2(t) &= \alpha_2 \left(\arctan \frac{x_3 - p_3}{x_2 - p_2} \right) dt + d\bar{w}_2(t) \end{aligned}$$

$$\begin{aligned} dy_3(t) &= \alpha_3 \left(\arctan \frac{x_3 - q_3}{x_1 - q_1} \right) dt + d\bar{w}_3(t) \\ dy_4(t) &= \alpha_4 \left(\arctan \frac{x_3 - q_3}{x_2 - q_2} \right) dt + d\bar{w}_4(t) \end{aligned}$$

where $p, q \in \mathbb{R}^3$ are the positions of the two observers and the four standard Brownian motions $\bar{w}_i$, $i = 1, \ldots, 4$ are all independent.

For the figures below, the parameter values chosen are $\epsilon = 0.01$, $B(i,i) = 3$, $\alpha_i = 1$, $i = 1, \ldots, 3$. We started observations at $t_1 = 0.5$ and simulated until $T = 3$. The mean and standard deviation of the initial values are:

	Mean	S.D.
x_1	100	10
x_2	100	10
x_3	0	0
v_1	-30	5
v_2	0	5
v_3	25	0

The number of generations per unit time and the time step used for simulating the trajectories of the particles are the same as for the first example. The pictures in Figures 4–8 show the cloud of particles at four different points in time (only the first three coordinates, i.e., the position of the particles and without their velocity coordinates). The signal has also been plotted in each picture (again, only the first three coordinates).

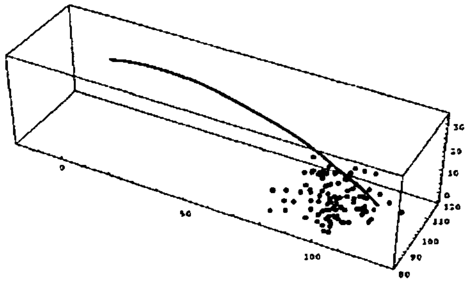

Figure 5: The particle system: time 0

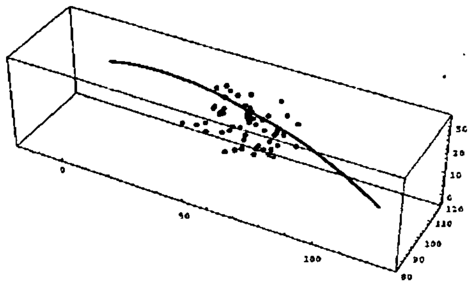

Figure 6: The particle system: time 1

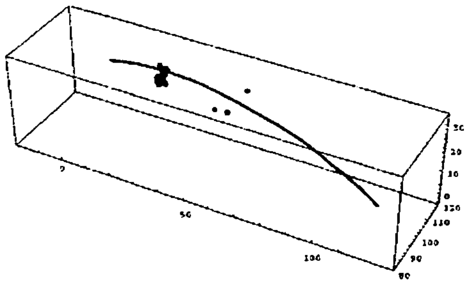

Figure 7: The particle system: time 2

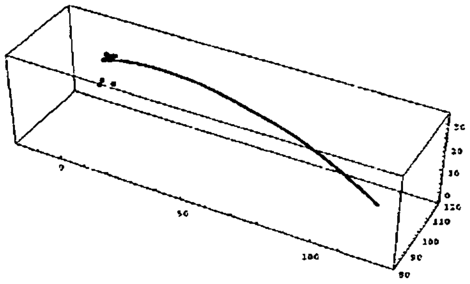

Figure 8: The particle system: time 3

Chapter 8

Appendix

8.1 Uniform upper bounds for the solution of Zakai equation

In this section we present a generic example for which the assumption **U** holds. Let W be a standard m-dimensional Brownian motion defined on probability space $(\Omega, \mathcal{F}, P)$ and $A : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ be the second order differential operator with smooth coefficients

$$A\zeta(x) = \sum_{i,j=1}^{d} a_{i,j}(x)\frac{\partial^2\zeta(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} f_i(x)\frac{\partial\zeta(x)}{\partial x_i}, \ \zeta \in \mathcal{D}(A) \subset C_b(\mathbb{R}^d)$$

It is well known that, provided the matrix a is symmetric and positive definite, for a suitable choice of Riemannian metric on $\mathbb{R}^d$, A takes the form

$$A\zeta = \Delta\zeta + f \cdot \nabla\zeta \tag{8.1}$$

where Δ is the Laplace-Beltrami operator, $\nabla\zeta = (\nabla\zeta^l)_{l=1,..,d}$ is the gradient of ζ, f is a smooth vector field and '$\_ \cdot \_$' is the usual scalar product. In the following we will work with A as having the form described in (8.1). Let also $h : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^m$ be a smooth function. For $t \in [0,\infty)$, we will denote by $h(t)$, the function $x \to h(t,x)$.

We consider the following stochastic partial differential equation

$$\begin{aligned} d\psi_t &= A\psi_t dt + h^*(t)\psi_t dW_t, \quad t \geq 0 \qquad (8.2)\\ \psi_0 &= \varphi \end{aligned}$$

where φ is a smooth bounded function. We want to prove that, under certain boundedness conditions on h and f and φ, we have that

$$E\left[\sup_{t\in[0,T]} {\psi_t}^2\right] < \infty \tag{8.3}$$

$$E\left[\sup_{t\in[0,T]} |\nabla\psi_t|^2\right] < \infty \tag{8.4}$$

where E is the expectation with respect to the probability P (in (8.4) $|\nabla\psi_t|^2 \overset{\text{def}}{=} \nabla\psi_t \cdot \nabla\psi_t$). We introduce the following quantities ($\|\cdot\|$ is the supremum norm)

$$\begin{aligned}
\|h\| &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \|h_i(t)\|, & \|Ah\| &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \|Ah_i(t)\| \\
\|\nabla h\|^2 &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \left\||\nabla h_i(t)|^2\right\|, & \left\|\frac{\partial h}{\partial t}\right\| &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \left\|\frac{\partial h_i(t)}{\partial t}\right\| \\
\|\nabla\nabla h\|^2 &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \left\|\sum_{l=1}^{d} |\nabla(\nabla h_i(t))^l|^2\right\|, & \|\nabla Ah\|^2 &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \left\||\nabla Ah_i|^2\right\| \\
\left\|\frac{\partial\nabla h}{\partial t}\right\|^2 &= \sup_{t\in[0,T]}\max_{i=1,\dots,m} \left\|\left|\nabla\frac{\partial h_i(t)}{\partial t}\right|^2\right\|, & \|\nabla f\|^2 &= \sup_{t\in[0,T]}\max_{i=1,\dots,d} \left\||\nabla f_i(t)|^2\right\| \\
\|\nabla\varphi\|^2 &= \left\||\nabla\varphi|^2\right\|
\end{aligned}$$

Theorem 29 *If $\|h\|$, $\|\nabla h\|$, $\|Ah\|$, $\|\frac{\partial h}{\partial t}\|$ are finite then ψ_t is uniformly bounded on $[0,T]$, P-a.s. and*

$$E\left[\sup_{t\in[0,T]} {\psi_t}^2\right] \le \exp 2m(\|h\| + \|Ah\| + \left\|\frac{\partial h}{\partial t}\right\|)^2 + 2m\,\|\nabla h\|^2 \tag{8.5}$$

We also prove a similar theorem for the gradient of ψ_t

Theorem 30 *If $\|h\|$, $\|\nabla h\|$, $\|Ah\|$, $\|\frac{\partial h}{\partial t}\|$, $\|\nabla\nabla h\|$, $\|\nabla Ah\|$, $\|\frac{\partial\nabla h}{\partial t}\|$, $\|\nabla f\|$ are finite then $\nabla\psi_t$ is uniformly bounded on $[0,T]$, P-a.s. and $E\left[\sup_{t\in[0,T]}|\nabla\psi_t|^2\right]$ is finite.*

Remark 25 *One can also find an explicit upper bound for $E\left[\sup_{t\in[0,T]}|\nabla\psi_t|^2\right]$.*

The proofs follow the following route. One first obtains the robust form of (8.2) which is a deterministic PDE with random coefficients. Using the maximum principle for the operator A, one obtains an upper bound for the solution of the deterministic

PDE, which in turn gives an uniform upper bound for ψ_t. This is then integrated by dividing the interval into small intervals and iterating the procedure for each interval. The explicit upper bound is obtained by taking the limit when the length of the small intervals goes to 0.

The Proofs

Let $\theta_t \overset{\text{def}}{=} \exp(-h^*(t)W_t)$, $t \geq 0$, and $\overline{\psi}_t \overset{\text{def}}{=} \psi_t\theta_t$, then, by Itô's rule, we have

$$d\theta_t = \theta_t\left(-\frac{\partial h^*(t)W_t}{\partial t}dt - h^*(t)dW_t + \frac{1}{2}|h(t)|^2dt\right)$$

thus

$$\begin{aligned} d\overline{\psi}_t &= \theta_t d\psi_t + \psi_t d\theta_t - \overline{\psi}_t|h(t)|^2dt \\ \frac{d\overline{\psi}_t}{dt} &= \theta_t A(t)\psi_t - \overline{\psi}_t\left(\frac{\partial h^*(t)W_t}{\partial t} + \frac{1}{2}|h(t)|^2\right) \qquad (8.6) \end{aligned}$$

We want to transform (8.6) into an equation in $\overline{\psi}_t$. We have

$$\begin{aligned} \frac{\partial\overline{\psi}_t}{\partial x_i} &= \theta_t\frac{\partial\psi_t}{\partial x_i} - \overline{\psi}_t\frac{\partial h^*(t)W_t}{\partial x_i} \\ \frac{\partial^2\overline{\psi}_t}{\partial x_i\partial x_j} &= \theta_t\frac{\partial^2\psi_t}{\partial x_i\partial x_j} - \theta_t\frac{\partial\psi_t}{\partial x_i}\frac{\partial h^*(t)W_t}{\partial x_j} - \frac{\partial\overline{\psi}_t}{\partial x_j}\frac{\partial h^*(t)W_t}{\partial x_i} - \overline{\psi}_t\frac{\partial^2 h^*(t)W_t}{\partial x_i\partial x_j} \\ \frac{\partial^2\overline{\psi}_t}{\partial x_i\partial x_j} &= \theta_t\frac{\partial^2\psi_t}{\partial x_i\partial x_j} - \frac{\partial\overline{\psi}_t}{\partial x_i}\frac{\partial h^*(t)W_t}{\partial x_j} - \overline{\psi}_t\frac{\partial h^*(t)W_t}{\partial x_i}\frac{\partial h^*(t)W_t}{\partial x_j} \\ &\quad - \frac{\partial\overline{\psi}_t}{\partial x_j}\frac{\partial h^*(t)W_t}{\partial x_i} - \overline{\psi}_t\frac{\partial^2 h^*(t)W_t}{\partial x_i\partial x_j} \end{aligned}$$

Hence

$$\begin{aligned} \theta_t A\psi_t &= \theta_t\Delta\psi_t + f\cdot\theta_t\nabla\psi_t \\ &= A\overline{\psi}_t + 2\nabla(h^*(t)W_t)\nabla\overline{\psi}_t + \overline{\psi}_t\left(Ah^*(t)W_t + |\nabla(h^*(t)W_t)|^2\right) \end{aligned}$$

and, if we define the quantity

$$e_t(h,z) \overset{\text{def}}{=} Ah^*(t)W_t + |\nabla(h^*(t)W_t)|^2 - \frac{\partial h^*(t)W_t}{\partial t} - \frac{1}{2}|h(t)|^2,$$

and introduce the operator where

$$\widehat{A}\zeta = A\zeta + 2\nabla(h^*(t)W_t)\nabla\zeta$$

then $\overline{\psi}_t$ satisfies the equation

$$\begin{aligned} \frac{d\overline{\psi}_t}{dt} &= \widehat{A}\overline{\psi}_t + \overline{\psi}_t e_t(h,z) \\ \overline{\psi}_0 &= \varphi \end{aligned} \tag{8.7}$$

and assume that all of them are finite. We have the following inequality

$$\frac{d\overline{\psi}_s}{dt} - \widehat{A}\overline{\psi}_s \leq \overline{\psi}_t \left(\left(\sum_{i=1}^{m} \max_{s\in[0,t]} |W^i_s| \right) \left(\|Ah\| + \left\| \frac{\partial h}{\partial t} \right\| \right) + \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W^i_s| \right)^2 \|\nabla h\|^2 \right)$$

where $s \in [0,t]$. Using then the maximum principle we obtain that

$$\overline{\psi}_s \leq \varphi \exp \left(\left(\sum_{i=1}^{m} \max_{s\in[0,t]} |W^i_s| \right) \left(\|Ah\| + \left\| \frac{\partial h}{\partial t} \right\| \right) + \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W^i_s| \right)^2 \|\nabla h\|^2 \right) \tag{8.8}$$

and hence

$$\psi_s \leq \varphi \exp \left(\left(\sum_{i=1}^{m} \max_{s\in[0,t]} |W^i_s| \right) \left(\|h\| + \|Ah\| + \left\| \frac{\partial h}{\partial t} \right\| \right) + \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W^i_s| \right)^2 \|\nabla h\|^2 \right)$$

for $s \in [0,t]$. We observe that the exponential of $(\max_{s\in[0,t]} |W^i_s|)^2$ is integrable only for t small, so what we do is the following. We divide the interval $[0,1]$ in n small intervals $\left[\frac{k}{n}, \frac{k+1}{n}\right]$, then we proceed as before for the first interval $\left[0, \frac{1}{n}\right]$ and then for the interval $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ we repeat the argument with $\exp(-h^*(t)(W_{\frac{k}{n}+t} - W_t))$, instead of $\exp(-h^*(t)W_t)$. After doing this for all the intervals we obtain

$$\begin{aligned} \psi_s \leq \varphi \prod_{k=1,\dots,n} \exp & \left(\sum_{i=1}^{m} \max_{s\in[0,\frac{1}{n}]} |W^i_{s+\frac{k}{n}} - W^i_{\frac{k}{n}}| \left(\|h\| + \|Ah\| + \left\| \frac{\partial h}{\partial t} \right\| \right) \right. \\ & \left. + \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W^i_{s+\frac{k}{n}} - W^i_{\frac{k}{n}}| \right)^2 \|\nabla h\|^2 \right) \end{aligned} \tag{8.9}$$

For big enough n the right hand side of (8.9) is integrable. By integrating its square and taking into account that the terms under the product are independent and identically distributed, we obtain that

$$E \left[\sup_{t\in[0,1]} |\psi_t|^2 \right] \leq \|\varphi\|^2 \left(\int_0^\infty \exp \left(2z \left(\|h\| + \|Ah\| + \left\| \frac{\partial h}{\partial t} \right\| \right) + 2z^2 \|\nabla h\|^2 \right) \alpha_{\frac{1}{n}}(z) dz \right)^{nm}$$

where

$$\alpha_{\frac{1}{n}}(z) = \frac{\sqrt{2n}}{\sqrt{\pi}} \exp\left(-\frac{nz^2}{2}\right)$$

is the density of the running maximum of Brownian motion in the interval $\left[0, \frac{1}{n}\right]$. We compute first the integral in the equation above. To simplify the formulae, we make the notations $d = 2(\|h\| + \|Ah\| + \|\frac{\partial h}{\partial t}\|)$ and $c = 2\,\|\nabla h\|^2$. Hence

$$\frac{\sqrt{2n}}{\sqrt{\pi}} \int_0^\infty \exp(dz + cz^2) \exp\left(-\frac{nz^2}{2}\right) dz$$

$$\begin{aligned}
&= \frac{\sqrt{2n}}{\sqrt{\pi}} \exp\left(\frac{d^2}{2(n-2c)}\right) \int_0^\infty \exp - \frac{\left(z - \frac{d}{n-2c}\right)^2}{2\left(\frac{1}{n-2c}\right)} dz \\
&= \frac{\sqrt{2n}}{\sqrt{\pi}} \exp\left(\frac{d^2}{2(n-2c)}\right) \left(\int_0^\infty \exp - \frac{z^2}{2\left(\frac{1}{n-2c}\right)} dz + \int_{-\frac{d}{n-2c}}^0 \exp - \frac{z^2}{2\left(\frac{1}{n-2c}\right)} dz \right) \\
&\leq \frac{\sqrt{2n}}{\sqrt{\pi}} \exp\left(\frac{d^2}{2(n-2c)}\right) \left(\frac{\sqrt{2\pi\left(\frac{1}{n-2c}\right)}}{2} + \frac{d}{n-2c} \right) \\
&\leq \left(\sqrt{\frac{n}{n-2c}} + \frac{d\sqrt{2}}{\sqrt{n\pi}} \right) \exp \frac{d^2}{2(n-2c)}
\end{aligned}$$

Hence

$$\begin{aligned}
E\left[\sup_{t\in[0,1]} |\psi_t|^2\right] &\leq \lim_{n\to\infty} \left(\sqrt{\frac{n}{n-2c}} + \frac{d\sqrt{2}}{\sqrt{n\pi}} \right)^{nm} \exp \frac{nmd^2}{2(n-2c)} \\
E\left[\sup_{t\in[0,1]} |\psi_t|^2\right] &\leq \exp \frac{md^2}{2} + mc
\end{aligned}$$

and finally we find the required upper bound (8.5).

We also want the equation satisfied by $\widehat{\psi}_t \overset{\text{def}}{=} \nabla\overline{\psi}_t \cdot \nabla\overline{\psi}_t$. By differentiating (8.7), we obtain

$$\begin{aligned}
\frac{d\nabla\overline{\psi}_t}{dt} &= \widehat{A}\nabla\overline{\psi}_t + (\nabla f + 2\nabla\nabla(h^*(t)W_t))\nabla\overline{\psi}_t + \nabla\overline{\psi}_t e_t(h,z) + \overline{\psi}_t \nabla e_t(h,z) \\
\nabla\overline{\psi}_0 &= \nabla\varphi
\end{aligned}$$

Also

$$A\widehat{\psi}_t = 2\nabla\overline{\psi}_t \cdot A\nabla\overline{\psi}_t + 2|\nabla\nabla\overline{\psi}_t|^2,$$

Hence

$$\begin{aligned}
\frac{d\widehat{\psi}_t}{dt} &= 2\nabla\overline{\psi}_t \cdot \frac{d\nabla\overline{\psi}_t}{dt} \\
&= A\widehat{\psi}_t - 2|\nabla\nabla\overline{\psi}_t|^2 + 2\nabla\overline{\psi}_t \cdot (\nabla f + 2\nabla\nabla(h^*(t)W_t))\nabla\overline{\psi}_t \\
&\quad + \widehat{\psi}_t e_t(h,z) + 2\nabla\overline{\psi}_t \cdot \nabla e_t(h,z)\overline{\psi}_t \\
\widehat{\psi}_0 &= \nabla\varphi \cdot \nabla\varphi,
\end{aligned}$$

which, in turn, implies

$$\frac{d\widehat{\psi}_t}{dt} - \widehat{A\psi}_t \leq \widehat{\psi}_t \left(a_1 \sum_{i=1}^{m} \max_{s\in[0,t]} |W_s^i| + a_2 \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W_s^i| \right)^2 + a_3 \right) + \overline{\psi}_t^2 \qquad (8.10)$$

where we used the following further notation

$$\begin{aligned}
a_1 &= 2m\|\nabla f\| + 4m\|\nabla\nabla h\| + \|Ah\| + \left\|\frac{\partial h}{\partial t}\right\| + \|\nabla Ah\| + \left\|\frac{\partial \nabla h}{\partial t}\right\| \\
a_2 &= \|\nabla h\|^2 + \|\nabla\nabla h\|^2 \\
a_3 &= \|h\|^2 \|\nabla h\|^2
\end{aligned}$$

Using once again the maximum principle in (8.10) and (8.8) one shows that

$$\begin{aligned}
\widehat{\psi}_t \leq& \left(\|\nabla\varphi\|^2 + t\|\varphi\|^2 \exp\left(a_4 \sum_{i=1}^{m} \max_{s\in[0,t]} |W_s^i| + a_5 \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W_s^i| \right)^2 \right) \right) \\
& \times \exp\left(a_1 \sum_{i=1}^{m} \max_{s\in[0,t]} |W_s^i| + a_2 \sum_{i=1}^{m} \left(\max_{s\in[0,t]} |W_s^i| \right)^2 + a_3 \right) \qquad (8.11)
\end{aligned}$$

where

$$\begin{aligned}
a_4 &= \|Ah\| + \left\|\frac{\partial h}{\partial t}\right\| \\
a_5 &= \|\nabla h\|^2
\end{aligned}$$

We also have

$$|\nabla\psi_t|^2 \le 2\widehat{\psi}_t \exp\left(2\|h\|\sum_{i=1}^{m}\max_{s\in[0,t]}|W_s^i|\right) + 2\psi_t^2\|\nabla h\|^2\sum_{i=1}^{m}\left(\max_{s\in[0,t]}|W_s^i|\right)^2 \qquad (8.12)$$

From (8.12) and (8.11), we obtain that (8.4) in the same way as for (8.3).

8.2 The construction of a positive integer valued random variable with given mean and variance

Let $\alpha > 0$ be the given mean and $v > 0$ be the given variance.

Proposition 23 *There exist positive integer valued random variables with mean α and variance v if and only if the following inequality holds*

$$\alpha^2 - \alpha + v \ge 0.$$

Proof. Let us suppose that we can find Λ, a positive integer valued random variable with mean α and variance v. Then

$$|\Lambda - \frac{1}{2}| \ge \frac{1}{2}$$

hence

$$E[(\Lambda - \frac{1}{2})^2] = E[(\Lambda)^2] - E[(\Lambda)] + \frac{1}{4} = \alpha^2 - \alpha + v + \frac{1}{4} \ge \frac{1}{4}$$

which proves half of the claim (E is the expectation over the probability space that the random variable is defined on). To prove the other half, we only need to construct Λ, a positive integer valued random variable with mean $\alpha \in (0,1]$ and variance v because any other case can be reduce to this one by adding a constant. Let $\delta \ge 0$ be the following quantity

$$\delta = \alpha^2 - \alpha + v \ge 0$$

and $\beta > 1$ be a positive integer such that $\beta \ge 1 + \frac{\delta}{\alpha}$. Then the random variable $\Lambda : \Omega \to \{0, 1, \beta\}$ such that

$$\tilde{P}[\Lambda = 0] = 1 - \alpha + \frac{\delta}{\beta}$$

$$\tilde{P}[\Lambda = 1] = \alpha - \frac{\delta}{\beta - 1}$$
$$\tilde{P}[\Lambda = \beta] = \frac{\delta}{\beta^2 - \beta}$$

satisfies our requirements.

■

Remark 26 *If $v \geq \frac{1}{4}$, then*

$$\alpha^2 - \alpha + v = (\alpha - \frac{1}{2})^2 + v - \frac{1}{4} \geq 0$$

and hence we can construct a positive integer valued random variable with any mean $\alpha > 0$.

8.3 Miscellaneous

In this section we state several classical results used in the thesis.

Proposition 24 (Gronwall's inequality) *Suppose that the continuous function $g(t)$ satisfies*

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s)ds; \; 0 \leq t < T$$

with $\beta \geq 0$ and $\alpha : [0, T] \to \mathbb{R}$ integrable. Then

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s) e^{\beta(t-s)} ds; \; 0 \leq t < T.$$

Let $(X, \mathcal{S}, \mu)$ be a measure space.

Lemma 8 (Fatou) *If f_n is a sequence of integrable functions such that $f_n \geq 0$, a.e., and*

$$\liminf \int f_n d\mu < \infty,$$

then there exists an integrable function f such that $f = \liminf f_n$, a.e., and one has

$$\int f d\mu \leq \liminf \int f_n d\mu.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{W_t, \mathcal{F}_t; t \geq 0\}$ be a d-dimensional standard Brownian motion defined on it. Let also $\{X_t, \mathcal{F}_t; t \geq 0\}$ be a d-dimensional adapted process such that

$$P\left[\int_0^T (X_t^i)^2 dt < \infty\right] = 1;\ 1 \leq i \leq d,\ 0 \leq T < \infty.$$

We set

$$Z_t(X) = \exp\left[\sum_{i=1}^{d} \int_0^t X_s^i dW_s^i - \frac{1}{2}\int_0^t \|X_s\|^2\, ds\right],\ 0 \leq t < \infty.$$

Assume that $Z_t(X)$ is a martingale and define, for each $0 \leq T < \infty$, a probability measure $\tilde{P}_T$ on $\mathcal{F}_T$ by

$$\tilde{P}_T(A) = E[1_A Z_T(X)];\ A \in \mathcal{F}_t.$$

The martingale property shows that the family of probability measures $\{\tilde{P}_T; T \geq 0\}$ satisfies the consistency condition

$$\tilde{P}_T(A) = \tilde{P}_t(A),\ A \in \mathcal{F}_t,\ 0 \leq t < T.$$

Theorem 31 (Girsanov, Cameron, Martin) *Assume that $Z_t(X)$ is a martingale and define $\{\tilde{W}_t, \mathcal{F}_t; t \geq 0\}$ by*

$$\tilde{W}_t^i = W_t^i - \int_0^t X_s^i ds;\ 1 \leq i \leq d,\ 0 \leq t < \infty.$$

Then, for each fixed $T \in [0, \infty)$, the process $\{\tilde{W}_t, \mathcal{F}_t; t \in [0, T]\}$ is a d-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, P)$

For a proof of Girsanov's theorem see [25], pp 191.

Theorem 32 (Doob's maximal inequality) *Let $\{X_t, \mathcal{F}_t; t \geq 0\}$ be a right-continuous submartingale and $[\sigma, \tau]$ a subinterval of $[0, \infty)$. Then*

$$E\left[\left(\sup_{\sigma < t \leq \tau} X_t\right)^p\right] \leq \left(\frac{p}{p-1}\right) E\left[(X_\tau)^p\right],\ p > 1.$$

provided $X_t \geq 0$ a.s. P for every $t \leq 0$, and $E\left[(X_\tau)^p\right] < \infty$.

For a proof of Doob's maximal inequality see [25], pp 13-14.

Theorem 33 (Burkholder-Davis-Gundy) *Let M be a continuous square integrable local martingale. For every $m > 0$ there exist universal positive constants k_m, K_m (depending only on m), such that*

$$k_m E[< M >_T^m] \leq E[(\max_{0 \leq s \leq T} |M_s|)^{2m}] \leq K_m E[< M >_T^m]$$

holds for every stopping time T.

For a proof of the Burkholder-Davis-Gundy inequalities, see [25], pp 166.

Bibliography

[1] D. ALDOUS, *Stopping Times and Tightness*, Ann. Prob., 6, (1978).

[2] V. E. BENEŠ, *Exact Finite-Dimensional Filters for Certain Diffusions with Nonlinear Drift*, Stochastics, 1981, Vol. 5, pp 65-92.

[3] A. BENSOUSSAN, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, 1992.

[4] A. BENSOUSSAN, J. L. LIONS, *Applications des Inéquations Variationnelles en Contrôle Stochastique*, Méthodes Mathématiques de l'Informatique, Dunod, Bordas, Paris 1978.

[5] P. BERNARD, D. TALAY, L. TUBARO, *Rate of Convergence of a Stochastic Particle Method for the Kolmogorov Equation with Variable Coefficients*, Mathematics of Computation, 1994, Vol. 4 No 4. pp. 247-251.

[6] Z. BRZEZNIAK, M. CAPINSKI, F. FLANDOLI, *Approximation for Diffusion in Random Fields*, Stochastic Analysis and Applications, 8(3), 293-313, 1990.

[7] B. CHAUVIN, A. ROUAULT, *KPP Equation and Supercritical Branching Brownian Motion in the Subcritical Speed Area. Application to Spatial Trees*, Probability Theory and Related Fields 80 (1988), 299-314.

[8] B. CHAUVIN, A. ROUAULT, *A stochastic Simulation for solving Scalar Reaction-Diffusion Equation*, Adv. Appl. Probability 12 (1990), 88-100.

[9] D. CRISAN, *A Direct Computation of the Beneš Filter Conditional Density*, admitted for publication in Stochastics and Stochastic Reports, anonymous FTP at ftp.maths.ed.ac.uk/pub/danc.

[10] D. CRISAN, T. LYONS, *Nonlinear Filtering and Measure Valued Processes*, submitted to Probability Theory and Related Fields, anonymous FTP at ftp.maths.ed.ac.uk/pub/danc.

[11] Y. DAVIDOV, *Genetic Algorithms and Robotics* London., World Scientific, 1990.

[12] D. A. DAWSON, *Stochastic Evolution Equation and Related Measure Processes*, Journal of Multivariate Analysis

[13] D. A. DIVER, *Applications of genetic algorithms to the solutions of ordinary differential equations*, Journal of Physics A:Math. Gen. 26 (1993) 3503-3513.

[14] K. D. ELWORTHY, H. Z. ZHAO, J. G. GAINES, *The Propagation of Traveling Waves for Stochastic Generalized KPP Equations*, Preprint, 1994

[15] S. N. ETHIER, T. G. KURTZ, *Markov Processes, Characterization and Convergence*, Wiley Series in Probability and Statistics, 1986.

[16] F. FLANDOLI, *Some Remarks on Regularity Theory and Stochastic Flows for Parabolic SPDE's*, Pisa Preprints, 135/92.

[17] F. FLANDOLI, K. -U. SCHAUMLOFELL, *Stochastic Parabolic Equations in Bounded Domains: Random Evolution Operator and Lyapunov Exponents*, Stochastics and Stochastic Reports, Vol 29, 461-485.

[18] A. E. GELFAND, A. F. M. SMITH, *Sampling-Based Approaches to Calculating Marginal Densities*, J. Amer. Statist. Assoc., 85, 398-409 (1990).

[19] A. E. GELFAND, S. E. HILLS. A. RACINE-POON, A. F. M. SMITH *An Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling*, J. Amer. Statist. Assoc., 85, 972-985 (1990).

[20] D. GEMAN, S. GEMAN, *Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images* IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, 721-741 (1984).

[21] D. E. GOLDBERG, *Genetic Algorithms in Search, Optimisation and Machine Learning*, New-York, Addison Wesley, 1989.

[22] I. ISCOE, *A weighted occupation time for a class of measure valued branching processes*, Probability Theory and Related Fields, 1986, Vol. 71, No.1, pp. 85-116.

[23] R. E. KALMAN, *A new Approach to Linear Filtering and Prediction Problems*, Journal of Basic Eng., 35-45, (1980).

[24] R. E. KALMAN, R. S. BUCY, *New Results in Linear Filtering and Prediction Theory*, Journal of Basic Eng., 95-108, (1981).

[25] I. KARATZAS, S. E. SHREVE, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer-Verlag, 1988.

[26] H. KIELHÖFER, *Global Solutions of Semilinear Evolutions Equations Satisfying an Energy Inequality*, Journal of Differential Equations 36, 188-222 (1980).

[27] A. KOLEN, E. PESCH , *Genetic local search in combinatorial optimisation*, Discrete Applied Mathematics, 48, 1994, 273-284.

[28] K. KAVAZU, S. WATANABE, *Branching processes with immigration and related limit theorems*, Theory Prob. Applications 16 (1971).

[29] T. G. KURTZ, *Approximation of Population Processes*, Soc. for Ind. and Appt. Math., 1981.

[30] T. G. KURTZ, D. L. OCONE, *Unique Characterisation of Conditional Distribution in Nonlinear Filtering*, The Annals of Probability, 1988, Vol 16, No. 1, 80-107.

[31] T. G. KURTZ, P. PROTTER, *Weak Limit Theorems for Stochastic Integrals and Stochastic Differential Equations*, The Annals of Probability, 1991, Vol 19, No. 3, 1035-1070.

[32] K. ITO AND H. P. MCKEAN, *Diffusion Processes and their Sample Paths*, Academic Press, New York, 1965.

[33] R. S. LIPTSER, A. N. SHIRYAYEV, *Statistics of Random Processes I*, Applications of Mathematics, Springer-Verlag, 1977.

[34] T. M. MACROBERT, *Functions of a Complex Variable*, St Martin's Press, New York, 1954.

[35] D. L. OCONE, *Topics in Nonlinear Filtering Theory*, Ph. D. Thesis, MIT, Cambridge 1980.

[36] E. PARDOUX, *Filtrage Non Linéaire et Equations aux Dérivées Partielle Stochastiques Associées*, Ecole d'Eté de Probabilités de Saint-Flour XIX - 1989, Lecture Notes in Mathematics, 1464, Springer-Verlag. 1991.

[37] E. PARDOUX, *Stochastic Partial Differential Equations and Filtering of Diffusion Processes*, Stochastics, Vol. 3, pp. 127-167, 1979.

[38] P. PRIOURET, *Processus de diffusion et equations differentielles stochastiques*, Ecole d'Eté de Probabilités de Saint-Flour XIX - 1974, Lecture Notes in Mathematics, pp. 37-113, Springer-Verlag, 1974.

[39] S. ROELLY-COPPOLETTA, *A Criterion of Convergence of Measure-Valued Processes: Application to Measure Valued Branching Processes*, Stochastics, 1986, Vol 17, pp. 43-65.

[40] A. S. SHERMAN, C. S. PESKIN, *A Monte Carlo Method for Scalar Reaction Diffusion Equations*, SIAM J. Sci. Statistics and Computing 7. (1986), 1360-1372.

[41] T. SHIGA, S. WATANABE, *Bessel diffusions as a one-parameter family of diffusion processes*, Zeitschrift fur Wahrsch. verw. Gebiete 27 (1973).

[42] S. WATANABE, *A Limit Theorem of Branching Processes and Continuous State Branching Processes*, H. Math. Kyoto Univ. 8 (1968).

[43] D. WILLIAMS, *Probability with Martingales*, Cambridge University Press, 1991.

[44] M. YOR, *Some Aspects of Brownian Motion. Pt. 1. Some Special Functionals*, Lectures in Mathematics, ETH, Zürich, Birkhäuser Verlag. 1992.