Central Limit Theorem for Non Linear Filtering and Interacting Particle Systems

P. Del Moral*, A. Guionnet†

Abstract

Several random particle systems approaches were recently suggested to solve numerically non linear filtering problems. The present analysis is concerned with genetic-type interacting particle systems. Our aim is to study the fluctuations on path space of such particle approximating models.

Keywords : Central Limit, Interacting random processes, Filtering, Stochastic approximation.

code A.M.S : 60F05, 60G35, 93E11, 62L20.

1 Introduction

1.1 Background and motivations

The Non Linear Filtering problem consists in recursively computing the conditional distributions of a non linear signal given its noisy observations. This problem has been extensively studied in the literature and, with the notable exception of the linear-Gaussian situation or wider classes of models (Bènes filters [2]) optimal filters have no finitely recursive solution (Chaleyat-Maurel/Michel [7]).
Although Kalman filtering ([26],[29]) is a popular tool in handling estimation problems its optimality heavily depends on linearity. When used for non linear filtering (Extended Kalman Filter) its performance relies on and is limited

*UMR C55830, CNRS, Bat.1R1, Univ. Paul Sabatier, 31062 Toulouse, delmoral@cict.fr.
†URA 743, CNRS, Bat. 425, Univ. Paris Sud, 91405 Orsay, France, guionnet@stats.matups.fr

by the linearization performed on the concerned model.
It has been recently emphasized that a more efficient way is to use random particle systems to solve numerically the filtering problem. That particle algorithms are gaining popularity is attested by the list of referenced papers. Instead of hand-crafting algorithms often on the basis of ad-hoc criteria, particle systems approaches provide powerful tools for solving a large class of non linear filtering problems.
Several practical problems which have been solved using these methods are given in [5, 6, 12, 13] including Radar/Sonar signal processing and GPS/INS integrations. Other comparisons and examples where the extended Kalman filter fails can be found in [4].
The present paper is concerned with the genetic type interacting particle systems introduced in [15]. We have shown in our earlier papers [15, 16] that, under rather general assumptions, the particle density profiles converge to the desired conditional distributions of the signal when the number of particles is growing. The study of the convergence of the empirical measure on path space and large deviation principles are presented in [19]. In the current work we study the fluctuations on path space of such particle approximations.

1.2 Description of the Model and Statement of Some Results

The basic model for the general non linear filtering problem consists of a time inhomogeneous Markov process $(X_n \; ; \; n \geq 0)$ taking its values in a Polish space $(E, \mathbf{B}(E))$ and, a non linear observation process $(Y_n \; ; \; n \geq 0)$ taking values in $\mathbb{R}^d$ for some $d \geq 1$.
To describe precisely our model, let us introduce some notations. We denote by $\mathbf{M}_1(E)$ the space of all probability measures on E furnished with the weak topology. We recall that the weak topology is generated by the bounded continuous functions. We will denote $\mathcal{C}_b(E)$ the space of these functions.
The classical filtering problem can be summarized as to find the conditional distributions

$$\eta_n(f) \stackrel{\text{def}}{=} E(f(X_n)/Y_1, \ldots, Y_n) \qquad \forall f \in \mathcal{C}_b(E) \qquad n \geq 0 \tag{1}$$

It was proven in a rather general setting by Kunita [27] and Stettner [32] that, given a series of observations $Y = y$, the distributions $(\eta_n \; ; \; n \geq 0)$ are solution of a discrete time measure valued dynamical system of the form

$$\eta_n = \phi(n, \eta_{n-1}) \qquad \forall n \geq 1 \qquad \eta_0 = \eta \tag{2}$$

where η is the law of the initial value of the signal and $\phi(n,.)$ an application on $\mathbf{M}_1(E)$ which depends on the series of observations $(y_n \ ; \ n \geq 1)$ and on the laws of the random perturbations $(V_n \ ; \ n \geq 1)$ (see Lemma 2.1 for its complete description).
The random particle system $(\Omega, F_n, (\xi_n)_{n\geq 0}, P)$ associated to (2) will be a Markov process with product state space E^N, where $N \geq 1$ is the size of the system. The N-tuple of elements of E, i.e. the points of the set E^N, are called particle systems and will be mostly denoted by the letters x, z.
Our dynamical system is then described by

$$\begin{aligned} P(\xi_0 \in dx) &= \prod_{p=1}^{N} \eta_0(dx^p) \\ P(\xi_n \in dx / \, \xi_{n-1} = z) &= \prod_{p=1}^{N} \phi\left(n, \frac{1}{N}\sum_{q=1}^{N} \delta_{z^q}\right)(dx^p) \end{aligned}$$

We use $dx \stackrel{\text{def}}{=} dx^1 \times \ldots \times dx^N$ to denote an infinitesimal neighborhood of $x = (x^1, \ldots, x^N) \in E^N$.
In [19] we develop large deviations principles for the law of the empirical distributions

$$\eta^N(\xi_{[0,T]}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{(\xi_0^i, \ldots, \xi_n^i)} \qquad T > 0$$

on the path space $(\Sigma_T, \mathbf{B}(\Sigma_T))$ with $\Sigma_T = E^{T+1}$ and we prove that it converges exponentially fast to a Dirac measure on the product measure

$$\eta_{[0,T]} \stackrel{\text{def}}{=} \eta_0 \otimes \ldots \otimes \eta_T$$

as the number of particles is growing. Our goal is now to investigate the fluctuations of such particle approximations, that is to show that

$$W_T^N = \sqrt{N}\left(\eta^N(\xi_{[0,T]}) - \eta_{[0,T]}\right)$$

converges in law as $N \to \infty$ to a centered Gaussian field.

One obvious consequence of such a result is that the so-called non linear filtering equation (2) can be regarded as a limiting measure valued dynamical system associated to a system of particles undergoing adaptation in a time varying and random environment. This environment is represented by the observation data and the form of noise source. Several convergence theorems ensuring the convergence of the particle scheme toward the desired distribution were obtained in [15, 16, 18] and [19]. The results presented in this paper makes it possible to estimate the deviations up to order $\sqrt{N}$ between $\eta^N(\xi_{[0,T]})$ and $\eta_{[0,T]}$.

The paper has the following structure:

In section 2 we recall the classical formulation of the filtering problem and the interacting particle system model under study. The main result of this paper is presented in section 3. In a first preliminary subsection 3.1 we present the assumptions needed in the foregoing development. To motivate our work we also compare our interacting particles model with the pure jump type process studied by Shiga and Tanaka in [33]. In subsection 3.2 we prove a Central Limit Theorem for the empirical measures associated to our particle scheme. Our basic tools are the Dynkin-Mandelbaum's Theorem on symmetric statistics and multiple Wiener's integrals as in [33]. Several examples of non linear filtering problems that can be handle in our framework are worked out in section 4.

2 Interacting Particle Systems for Non Linear Filtering

Several random particle system approaches were recently suggested to solve numerically non linear filtering problems (see [16] and [11] for theoretical details and references). These approaches develop new methods for dealing with measure valued dynamical system of the form (1). In a first subsection we present the so-called non linear filtering equations then we present the genetic type interacting particle approximation under study.

2.1 Formulation of the Non Linear Filtering Problem

Let $X = (\Omega_1 = E^{\mathbb{N}}, (\widetilde{F}_n^1)_{n\geq 0}, (X_n)_{n\geq 0}, P_X^0)$ be a time-inhomogeneous discrete time Markov process taking values in E with Feller probability transitions K_n, $n \geq 1$ and the initial distribution η_0.
Let $Y = (\Omega_2 = (\mathbb{R}^d)^{\mathbb{N}}, (\widetilde{F}_n^2)_{n\geq 0}, (Y_n)_{n\geq 0}, P_Y^0)$ be a sequence of independent random variables with measurable positive density g_n with respect to Lebesgue measure. Y is independent of X.

On the canonical space $(\Omega_0 = \Omega_1 \times \Omega_2, (F_n)_{n\geq 0} = (\widetilde{F}_n^1 \times \widetilde{F}_n^2)_{n\geq 0}, P^0 = P_X^0 \otimes P_Y^0)$ the signal process X and the observation process Y are P_0-independent. Let us set

$$L_n = \prod_{k=1}^{n} \frac{g_k(Y_k - h_k(X_{k-1}))}{g_k(Y_k)} \tag{3}$$

where $h_n : E \to \mathbb{R}^d$, $n \geq 1$, are bounded measurable functions. Note that L is a $(P_0, (\widetilde{F}_n)_{n\geq 0})$-martingale. Then we can define a new probability measure P on $(\Omega_0, (\widetilde{F}_n)_{n\geq 0})$ such that the restrictions P_n^0 and P_n to $\widetilde{F}_n$ satisfy

$$P_n = L_n\, P_n^0 \qquad n \geq 0 \tag{4}$$

One can prove that, under P, X is a time-inhomogeneous Markov process with transition operators K_n, $n \geq 1$ and initial distribution ν.
In addition $V_n = Y_n - h_n(X_{n-1})$, $n \geq 1$, are independent of X and independent random variables with continuous and positive density g_n with respect to Lebesgue measure. The classical filtering problem can be summarized as to estimate the distribution of X_n conditionally to the observations up to time n. Namely,

$$\eta_n(f) = E(f(X_n)/Y_1, \ldots, Y_n) \qquad \forall f \in \mathcal{C}_b(E)$$

The following result is a slight modification of Kunita [27] and Stettner results [32] (see also [16] for details).
If, for any Markov transition K on E and any $\mu \in \mathbf{M}_1(E)$ we denote μM the probability so that for any $f \in \mathcal{C}_b(E)$,

$$\mu M f = \int \mu(dx)(Mf)(x) \qquad \text{with} \qquad Mf(x) = \int f(z) M(x, dz)$$

the dynamical structure of the conditional distribution $(\eta_n \; ; \; n \geq 0)$ is given by the following lemma.

Lemma 2.1 *Given a fixed observation record Y=y, $(\eta_n)_{n\geq 0}$ is solution of the $\mathbf{M}_1(E)$-valued dynamical system*

$$\eta_n = \phi_n(y_n, \eta_{n-1}) \qquad n \geq 1 \qquad \eta_0 = \nu \tag{5}$$

where $y_n \in I\!\!R^d$ is the current observation and ϕ_n is the continuous function given by

$$\begin{aligned} \phi_n(y_n, \eta) &= \psi_n(y_n, \eta) K_n \\ \psi_n(y_n, \eta) f &= \frac{\displaystyle\int f(x)\, g_n(y_n - h_n(x))\, \eta(dx)}{\displaystyle\int g_n(y_n - h_n(z))\, \eta(dz)} \end{aligned}$$

for all $f \in \mathcal{C}_b(E)$, $\eta \in \mathbf{M}_1(E)$ and $y_n \in I\!\!R^d$.

2.2 Interacting Particle Systems Approximations

We have shown in our earlier papers [15, 16, 18, 19] that the non linear filtering equation can be regarded as the limiting measure valued dynamical system associated to a natural interacting particle system scheme. Namely, the N-particle system associated to (5) is defined by

$$\begin{aligned} P_y(\xi_0 \in dx) &= \prod_{j=1}^{N} \eta_0(dx^j) \\ P_y(\xi_n \in dx / \, \xi_{n-1} = z) &= \prod_{j=1}^{N} \phi_n(y_n, \frac{1}{N}\sum_{i=1}^{N} \delta_{z^i})(dx^j) \end{aligned} \tag{6}$$

Let us remark that

$$\phi_n\left(y_n, \frac{1}{N}\sum_{i=1}^N \delta_{z^i}\right) = \left(\sum_{i=1}^N \frac{g_n(y_n - h_n(z^i))}{\sum_{j=1}^N g_n(y_n - h_n(z^i))}\,\delta_{z^i}\right) K_n \tag{7}$$

and therefore

$$P_y(\xi_n \in dx/\,\xi_{n-1} = z) = \prod_{j=1}^N \sum_{i=1}^N \frac{g_n(y_n - h_n(z^i))}{\sum_{j=1}^N g_n(y_n - h_n(z^i))}\, K_n(z^i, dx^j)$$

Using the above observations we see that the particles evolve according two separate mechanisms. They can be modelled as follows :

Initial particle system

$$P_y\left(\xi_0 \in dx\right) = \prod_{p=1}^N \eta_0(dx^p)$$

Selection/Updating

$$P_y\left(\widehat{\xi}_{n-1} \in dx|\xi_{n-1} = z\right) = \prod_{p=1}^N \sum_{i=1}^N \frac{g_n(y_n - h_n(z^i))}{\sum_{j=1}^N g_n(y_n - h_n(z^j))}\delta_{z^i}(dx^p)$$

Mutation/Prediction

$$P_y\left(\xi_n \in dz|\widehat{\xi}_{n-1} = x\right) = \prod_{p=1}^N K_n(x^p, dz^p) \tag{8}$$

Thus, we see that the particles move according the following rules

1. **Updating:** When the observation $Y_n = y_n$ is received, each particle examines the system of particles $\xi_{n-1} = (\xi_{n-1}^1, \ldots, \xi_{n-1}^N)$ and chooses randomly a site ξ_{n-1}^i with probability
$$\frac{g_n(y_n - h_n(\xi_{n-1}^i))}{\sum_{j=1}^N g_n(y_n - h_n(\xi_{n-1}^j))}$$
2. **Prediction:** After the updating mechanism each particle evolves according the transition probability kernel of the signal process.

We see that this particle approximations of the non linear filtering equation belongs to the class of algorithms called genetic algorithms. These algorithms are based on the genetic mechanisms which guide natural evolution: exploration/mutation and updating/selection. They were introduced by J.H. Holland [25] in 1975 to handle global optimization problems on a finite set.

3 Central Limit Theorem for the Empirical Measures on Path Space

3.1 Hypotheses and General notations

In this section we study the fluctuations of the empirical distributions on path space. We will always assume that the signal transition kernels $\{K_n\ ;\ n \geq 1\}$ and the functions $\{g_n, h_n\ ;\ n \geq 0\}$ satisfy the following conditions:

(H0) For any time $n \geq 0$, h_n is bounded continuous and g_n is a positive continuous function.

(H1) K_n is Feller and such that for any time $n \geq 1$ there exists a reference probability measure $\lambda_n \in \mathbf{M}_1(E)$ and a $\mathbf{B}(E)$-measurable function φ_n so that

$$\delta_x K_n \ \sim\ \lambda_n$$

Moreover, there exists a non negative function φ_n such that, for any $p \geq 1$,

$$\left|\log \frac{d\delta_x K_n}{d\lambda_n}(z)\right| \leq \varphi_n(z) \quad \text{ and } \quad \int \exp\left(p\ \varphi_n\right)\, d\lambda_n < \infty \tag{9}$$

We now give some comments on hypotheses (H0) and (H1). First of all we note that very similar assumptions were introduced in [19] to study large deviations principles for the empirical measures on path space. As we will see in section 4 the conditions (H0) and (H1) cover many typical examples of non linear filtering problems.
On the other hand we note that if (H0) is satisfied then it is easily seen that there exists a family of positive functions $\{\alpha_n\ ;\ n \geq 0\}$, such that

$$\alpha_n(y)^{-1} \leq \frac{g_n(y - h_n(x))}{g_n(y)} \leq \alpha_n(y) \qquad \forall (y, x) \in I\!R^d \times E \quad \forall n \geq 0 \tag{10}$$

In such a context, it is not difficult to see that, for any given finite time T, if we denote P_T^N the the law of $(\xi_n)_{0\leq n\leq T}$ on path space, then P_T^N is absolutely continuous with respect to $\eta_{[0,T]}^{\otimes N}$ and

$$\frac{dP_T^N}{d\eta_{[0,T]}^{\otimes N}}(x) = \exp H_T^N(x) \tag{11}$$

with

$$H_T^N(x) = N \sum_{n=1}^{T} \int \log \frac{d\phi(n, m^N(x_{n-1}))}{d\eta_n}\, dm^N(x_n) \tag{12}$$

if

$$m^N(x_n) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_n^i} \qquad 0 \leq n \leq T$$

Therefore, the density of P_T^N only depends on the empirical measure m^N and we find ourselves exactly in the setting of mean field interacting particles with regular Laplace density.

The study of the fluctuations for mean field interacting particle systems via precise Laplace method is now extensively developed (see for instance [28] [35],[3],[1], [20] and references therein).
Various methods are based on the fact that the law of mean field interacting processes can be viewed as a mean field Gibbs measure on path space (see (11)). In such a setting, precise Laplace's method can be developed (see [1], [28], [20]...). In [20], the study of the fluctuations for mean field Gibbs measures was extended to analytic potentials which probably includes our setting.
However the present analysis is more related to Shiga/Tanaka's paper [33]. In this article, the authors restrict themselves to dynamics with independent initial data so that the partition function of the corresponding Gibbs measure is constant and equal to one. This simplifies considerably the analysis. In fact, the proof then mainly relies on a simple formula on multiple Wiener integrals and Dynkin-Mandelbaum theorem [30] on symmetric statistics. Also the pure jump McKean-Vlasov process studied in [33] is rather close to our model.
To motivate our work and illustrate the connections of our model with [33] we recall that the discrete time version of the interacting N-particle system studied in [33] is an E^N-valued Markov chain $(\zeta_n)_{n\geq 0}$ with transitions

$$P(\zeta_n \in dz/\zeta_{n-1} = x) = \prod_{p=1}^{N} \frac{1}{N} \sum_{j=1}^{N} Q(x^i, x^j, dz^p) \tag{13}$$

where $Q(x, x', dz)$ is a suitable measure kernel satisfying the following condition

(Q) There exists constants $c_1, c_2 > 0$ such that

$$c_1\, Q(x, x', dz) \leq Q(x, x'', dz) \leq c_2\, Q(x, x', dz) \qquad \forall x, x', x'' \in E$$

Our setting is therefore rather similar since, following (6), we have the same type of transitions

$$P_y(\xi_n \in dx / \xi_{n-1} = z) \;=\; \prod_{p=1}^{N} \sum_{i=1}^{N} \frac{g_n(y_n - h_n(z^i))}{\sum_{j=1}^{N} g_n(y_n - h_n(z^j))} K_n(z_i, dx^p)$$

except that the interaction does not appear linearly as in (13) which simplifies the analysis in [33].

It is also interesting to note that in the case of McKean's model of Boltzmann's Equations (see [33]) or more generally if the kernel $Q(x, x', dz)$ doesn't depend on the parameter x' the measure kernels Q has the form

$$Q(x, x', dz) = K(x, dz)$$

where K is a Markov transition on E. This also corresponds to our setting when $h_n \equiv 0$ since in this case the observation process is independent of the signal process.
In this situation an equivalent condition of (Q) is given by

(Q$'$) There exists of a reference probability measure $\lambda \in \mathbf{M}_1(E)$ and a constant $c > 0$ such that

$$\forall \mu \in \mathbf{M}_1(E) \quad \forall (x,z) \in E^2 \qquad \delta_x K \sim \lambda \quad \text{and} \quad c^{-1} \leq \frac{d\delta_x K}{d\lambda}(z) \leq c$$

A clear disadvantage of condition (Q$'$) is that it is in general not satisfied when E is not compact and in particular in many non linear filtering problems. Our purpose here is to extend the technic of [33] to handle the particle approximation introduce in section 2.2 and to replace the assumption (Q$'$) by the exponential moment condition (H1).

3.2 Main Result

Now we introduce some additional notations and the Hilbert-Schmidt integral operator which governs the fluctuations of the empirical measures of our particle algorithm. ¿From now on we fix a time $T > 0$ and a series of observations $(y_n\ ;\ 1 \leq n \leq T)$.
Under (H1) for any $n \geq 1$ there exists a reference probability measure $\lambda_n \in \mathbf{M}_1(E)$ such that $\delta_x K_n \sim \lambda_n$. In this case we shall use the notation

$$\forall (x,z) \in E^2 \qquad k_n(x,z) \stackrel{\text{def}}{=} \frac{d\delta_x K_n}{d\lambda_n}(z)$$

For any $x = (x_0, \ldots, x_T)$ and $z = (z_0, \ldots, z_T) \in \Sigma_T$ set

$$q_{[0,\mathrm{T}],y}(x,z) = \sum_{n=1}^{T} q_{n,y}(x,z)$$

with

$$\begin{aligned} q_{n,y}(x,z) &= \frac{g_n(y_n - h_n(z_{n-1}))\ k_n(z_{n-1}, x_n)}{\int_E g_n(y_n - h_n(u))\ k_n(u, x_n) \eta_{n-1}(du)} \\ a_{T,y}(x,z) &= q_{[0,\mathrm{T}],y}(x,z) - \int_{\Sigma_T} q_{[0,\mathrm{T}],y}(x',z)\ \eta_{[0,\mathrm{T}]}(dx') \end{aligned}$$

Now we remark that from the exponential moment condition (H1) we have that $q_{[0,\mathrm{T}],y} \in L^2(\eta \otimes \eta)$. It follows that $a_{T,y} \in L^2(\eta \otimes \eta)$ and therefore the integral operator $A_{T,y}$ given by

$$A_{T,y}\varphi(x) = \int a_{T,y}(z,x)\ \varphi(z)\ \eta_{[0,\mathrm{T}]}(dz) \qquad \forall \varphi \in L^2(\Sigma_T, \eta_{[0,\mathrm{T}]})$$

is an Hilbert-Schmidt operator on $L^2(\Sigma_T, \eta_{[0,\mathrm{T}]})$. We can now formulate our main result

Theorem 3.1 *Assume that conditions (H0) and (H1) are satisfied. For any observation record $Y = y$ and $T > 0$ the integral operator $I - A_{T,y}$ is invertible and the random field $\{W_T^N(\varphi)\ ;\ \varphi \in L^2(\eta_{[0,\mathrm{T}]})\}$ converges as $N \to \infty$ to a centered Gaussian field $\{W_T(\varphi)\ ;\ \varphi \in L^2(\eta_{[0,\mathrm{T}]})\}$ satisfying*

$$\begin{aligned} & E_y\left(W_T(\varphi_1) W_T(\varphi_2)\right) \\ &= \left((I - A_{T,y})^{-1}(\varphi_1 - \eta(\varphi_1)), (I - A_{T,y})^{-1}(\varphi_2 - \eta(\varphi_2))\right)_{L^2(\eta_{[0,\mathrm{T}]})} \end{aligned}$$

for any $\varphi_1, \varphi_2 \in L^2(\eta_{[0,\mathrm{T}]})$, in the sense of convergence of finite dimensional distributions.

We begin with the main line of the proof and present some basic facts on symmetric statistics and multiple Wiener integrals which are needed in the sequel.

To clarify the presentation, we simplify the notations suppressing the time parameter $T > 0$ and the observations parameter y in our notations. For instance we will write a instead of $a_{T,y}$, $q(x,z)$ instead of $q_{[0,T],y}(x,z)$ and η instead of $\eta_{[0,T]}$. In addition we will write $\phi(n,\centerdot)$ instead of $\phi_n(y_n,\centerdot)$ and $g_n(\centerdot)$ instead of $g_n(y_n - h_n(\centerdot))$.

Let us first recall how one can see that $I - A$ is invertible. This is in fact classical now (see [1] and [33] for instance). First one notices that, under our assumptions, A^n, $n \geq 2$ and $A\,A^*$ are trace class operators with

$$\begin{aligned} \mathrm{Trace} A^n &= \int \ldots \int a(x_1,x_2)\ldots a(x_n,x_1)\;\eta(dx_1)\ldots\eta(dx_n) \\ \mathrm{Trace} A A^* &= \int_{\Sigma_T^2} a(x,z)^2\;\eta(dx)\;\eta(dz) = \|a\|^2_{L^2(\eta\times\eta)} \end{aligned}$$

Furthermore by definition of a and the fact that η is a product measure it is easily checked that

$$\forall n \geq 2 \qquad \mathrm{Trace} A^n = 0$$

Standard spectral theory (see [31] for instance) then shows that $det_2(I - A)$ is equal to one and therefore that $I - A$ is invertible.

Let us now sketch the proof of Theorem 3.1. First, let us denote by P^N the distribution induced by $(\xi_n)_{0\leq n\leq T}$ on the path space $(\Sigma_T^N, \mathbf{B}(\Sigma_T)^N)$ where

$$\Sigma_T^N = \Sigma_T \times \ldots \times \Sigma_T \qquad \text{and} \qquad \mathbf{B}(\Sigma_T)^N = \mathbf{B}(\Sigma_T) \times \ldots \times \mathbf{B}(\Sigma_T).$$

As we noticed in the introduction (see (11)), P^N is absolutely continuous with respect to $\eta^{\otimes N}$ and

$$\frac{dP^N}{d\eta^{\otimes N}}(x) = \exp H^N(x) \qquad \text{for} \quad x = (x^1,\ldots,x^N) \in \Sigma_T^N$$

with

$$\begin{aligned} H^N(x) &= \sum_{n=1}^{T}\sum_{i=1}^{N} \log \frac{d\phi(n, m^N(x_{n-1}))}{d\eta_n}(x_n^i) \\ &= N \sum_{n=1}^{T} \int \log \frac{d\phi(n, m^N(x_{n-1}))}{d\eta_n}\; dm^N(x_n) \end{aligned}$$

In what follows we use $E_{\eta^{\otimes N}}(\boldsymbol{.})$ (resp. $E_{P^N}(\boldsymbol{.})$) to denote expectations with respect to the measure $\eta^{\otimes N}$ (resp. P^N) on Σ_T^N and, unless otherwise stated, the sequence $\{x^i\ ;\ i \geq 1\}$ is regarded as a sequence of Σ_T-valued and independent random variables with common law η.

To prove theorem 3.1, it is enough to study the limit of $\{E_{P^N}(\exp(iW_T^N(\varphi)))\ ;\ N \geq 1\}$ for functions $\varphi \in L^2(\eta)$.

Writing

$$E_{P^N}(\exp(iW_T^N(\varphi))) = E_{\eta^{\otimes N}}(\exp(iW_T^N(\varphi) + H^N(x))),$$

one finds that the convergence of $\{E_{P^N}(\exp(iW_T^N(\varphi)))\ ;\ N \geq 1\}$ follows from the convergence in law and the uniform integrability of $\exp(iW_T^N(\varphi)+H^N(x))$ under the product law $E_{\eta^{\otimes N}}$. The last point is clearly equivalent to the uniform integrability of $\exp H^N(x)$ under $E_{\eta^{\otimes N}}$.

The proof of the uniform integrability of $\exp H^N(x)$ then relies on a classical result (see for instance theorem 5 pp. 189 in [34] or Scheffé's Lemma 5.10 p.55 in [39]) which says that, if a sequence of non negative random variables $\{X_N\ ;\ N \geq 1\}$ converges almost surely towards some random variable X as $N \to \infty$ then we have

$$\lim_{N\to\infty} E(X_N) = E(X) < \infty \Longleftrightarrow \{X_N\ ;\ N \geq 1\} \text{ is uniformly integrable}$$

The equivalence still holds if X_N only converges in distribution by Skorohod's Theorem (see for instance Theorem 1 pp. 355 in [34]).
Since $E_{\eta^{\otimes N}}(\exp H^N(x)) = 1$ it is clear that the uniform integrability of $\{\exp H^N(x)\ ;\ N \geq 1\}$ follows from the convergence in distribution of $H^N(x)$ towards a random variable H such that $E(\exp H) = 1$.

Thus, it suffices to study the convergence in distribution of $\{iW_T^N(\varphi) + H^N(x)\ ;\ N \geq 1\}$ for $L^2(\eta)$ functions φ to conclude.

To state such a result, we first introduce Wiener integrals.

Let $\{I_1(\varphi)\ ;\ \varphi \in L^2(\eta)\}$ be a centered Gaussian field satisfying

$$E\left(I_1(\varphi_1)I_1(\varphi_2)\right) = (\varphi_1, \varphi_2)_{L^2(\eta)}$$

If we set, for each $\varphi \in L^2(\eta)$ and $m \geq 1$

$$h_0^\varphi = 1 \qquad h_m^\varphi(z_1, \ldots, z_m) = \varphi(z_1)\ldots\varphi(z_m),$$

the multiple Wiener integrals $\{I_m(h_m^\varphi)\ ;\ \varphi \in L^2(\eta)\}$ with $m \geq 1$, are defined by the relation

$$\sum_{m\geq 0} \frac{t^m}{m!}\, I_m(h_m^\varphi) = \exp\left(tI_1(\varphi) - \frac{t^2}{2}\|\varphi\|_{L^2(\eta)}\right).$$

The multiple Wiener integral $I_m(\phi)$ for $\phi \in L^2_{\mathrm{sym}}(\eta^{\otimes m})$ is then defined by a completion argument.

We are going to prove that, if $=_{\mathrm{law}}$ denotes the equality in law,

Lemma 3.2

$$\lim_{N\to\infty} H^N(x) =_{\mathrm{law}} \frac{1}{2} I_2(f) - \frac{1}{2}\, \mathit{Trace} AA^* \tag{14}$$

where f is given by

$$f(y,z) = a(y,z) + a(z,y) - \int_{\Sigma_T} a(u,y)\ a(u,z)\ \eta(du). \tag{15}$$

In addition, for any $\varphi \in L^2(\eta)$,

$$\lim_{N\to\infty} (H^N(x) + iW_T^N(\varphi)) =_{\mathrm{law}} \frac{1}{2} I_2(f) + iI_1(\varphi) - \frac{1}{2}\, \mathit{Trace} AA^*$$

Following the above observations, we get for any $\varphi \in L^2(\eta)$,

$$\begin{aligned}
\lim_{N\to\infty} E_{P^N}\left(\exp iW_T^N(\varphi)\right) &= \lim_{N\to\infty} E_{\eta^{\otimes N}}\left(\exp\left(iW_T^N(\varphi) + H^N(x)\right)\right) \\
&= E\left(\exp\left(iI_1(\varphi) + \frac{1}{2} I_2(f) - \frac{1}{2}\mathrm{Trace} AA^*\right)\right)
\end{aligned}$$

Moreover, Shiga-Tanaka's formula of Lemma 1.3 in [33] shows that for any $\varphi \in L^2_{\mathrm{sym}}(\eta)$,

$$E\left(\exp\left(iI_1(\varphi) + \frac{1}{2} I_2(f) - \frac{1}{2}\mathrm{Trace} AA^*\right)\right) = \exp\left(-\frac{1}{2}\|(I - A)^{-1}\varphi\|^2_{L^2(\eta)}\right) \tag{16}$$

The proof of theorem 3.1 is thus complete.

Proof of Lemma 3.2

Since $\eta_n = \phi(n, \eta_{n-1})$ and

$$\eta_n(du) = \frac{\eta_{n-1}\left(g_n(\centerdot) k_n(\centerdot, u)\right)}{\eta_{n-1}\left(g_n\right)}\ \ \lambda_n(du) \tag{17}$$

we deduce that for any $\mu \in \mathbf{M}_1(E)$ and $u \in E$

$$\begin{aligned} \frac{d\phi(n,\mu)}{d\eta_n}(u) &= \frac{d\phi(n,\mu)}{d\lambda_n}(u) \quad \frac{d\lambda_n}{d\phi(n,\eta_{n-1})}(u) \\ &= \frac{\mu\left(g_n(\centerdot)k_n(\centerdot,u)\right)}{\eta_{n-1}\left(g_n(\centerdot)k_n(\centerdot,u)\right)} \Big/ \frac{\mu\left(g_n\right)}{\eta_{n-1}\left(g_n\right)} \end{aligned}$$

Therefore the function H^N can be written in the form

$$H^N(x) = H_1^N(x) + \overline{H}_1^N(x)$$

with

$$\begin{aligned} H_1^N(x) &= \sum_{n=1}^{T}\sum_{i=1}^{N} \log\left(\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j)\right) \\ \overline{H}_1^N(x) &= -N\sum_{n=1}^{T} \log\left(\frac{1}{N}\sum_{j=1}^{N} \overline{q}_n(x^j)\right) \end{aligned}$$

where

$$\overline{q}_n(x^j) = g_n(x_{n-1}^j)/\eta_{n-1}(g_n)$$

Following (17), we also have

$$\overline{q}_n(z) = \int_{\Sigma_T} q_n(y,z)\ \eta(dy)$$

By the representation

$$\log z = (z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3(\epsilon z + (1-\epsilon))^3}$$

which is valid for all $z > 0$ with $\epsilon = \epsilon(z)$ such that $\epsilon(z) \in [0,1]$ we obtain the decomposition

$$\begin{aligned} H^N(x) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} a(x^i,x^j) - \frac{1}{2}\sum_{n=1}^{T}\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1\right)^2 \\ &\qquad + \frac{N}{2}\sum_{n=1}^{T}\left(\frac{1}{N}\sum_{j=1}^{N} \overline{q}_n(x^j) - 1\right)^2 + R^N \\ &= J_1^N + J_2^N + J_3^N + R^N \end{aligned} \tag{18}$$

where the remainder term satisfies $R^N = R_1^N + R_2^N$ with

$$|R_1^N| \leq \frac{1}{3}\sum_{n=1}^{T}\sum_{i=1}^{N}|\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1|^3 \; \theta_n^i(x) \tag{19}$$

$$\text{with} \quad \theta_n^i(x)^{-1} = \min\,(\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j), 1)^3 \tag{20}$$

$$|R_2^N| \leq \frac{NC_T}{3}\sum_{n=1}^{T}|\frac{1}{N}\sum_{j=1}^{N}\overline{q}_n(x^j) - 1|^3$$

for some finite constant $C_T < \infty$.
In the last inequality we have used the fact that, by (H1), the positive functions $\overline{q}_n$ are bounded from above and from below (see (10)).

Our aim is now to discuss a limit of a functional $H^N(x^1,\ldots,x^N)$. To this end, we shall rely on L^2-technics and more precisely Dynkin-Mandelbaum construction of multiple Wiener integrals as a limit of symmetric statistics. For completeness we present this result.
Let $\{\zeta^i \; ; \; i \geq 1\}$ be a sequence of independent and identically distributed random variables with values in an arbitrary measurable space $(\mathcal{X},\mathcal{B})$. To every symmetric function $h(z_1,\ldots,z_m)$ there corresponds a statistic

$$\sigma_m^N(h) = \sum_{1\leq i_1<\ldots<i_m\leq N} h(\zeta^{i_1},\ldots,\zeta^{i_m})$$

with the convention $\sigma_m^N = 0$ for $m > N$. Every integrable symmetric statistic $S(\zeta^1,\ldots,\zeta^N)$ has a unique representation of the form

$$S(\zeta^1,\ldots,\zeta^N) = \sum_{m\geq 0} \sigma_m^N(h_m) \tag{21}$$

where $h_m(z_1,\ldots,z_m)$ are symmetric functions subject to the condition

$$\int h_m(z_1,\ldots,z_{m-1},u)\; \mu(du) = 0 \tag{22}$$

where μ is the probability distribution of ζ^1.
We call such functions $\{h_m \; ; \; m \geq 0\}$ canonical. Finally we denote by $\mathcal{H}$ the set of all sequences $h = (h_0, h_1(z_1),\ldots,h_m(z_1,\ldots,z_m),\ldots)$ where h_m are canonical and

$$\sum_{m\geq 0}\frac{1}{m!}E(h_m^2(\zeta^1,\ldots,\zeta^m)) < \infty$$

As in [33] we will use repeatedly the following

Theorem 3.3 (Dynkin-Mandelbaum [30])
For $h \in \mathcal{H}$ the sequence of random variables $Z_N(h) = \sum_{m\geq 0} \frac{1}{N^{m/2}}\sigma_m^N(h_m)$ converges in law, as $N \to \infty$, to

$$W(h) = \sum_{m\geq 0} \frac{I_m(h_m)}{m!}$$

Since $H^N(x)$ can be regarded as a symmetric statistic of $x = (x^1, \ldots, x^N)$ which is a sequence of Σ_T-valued and independent random variables with common law η the above theorem is applicable. The technical trick to identify the limit of this functional is to decompose the symmetric statistic $H^N(x^1, \ldots, x^N)$ as in (21). In view of (18) it clearly suffices to prove that the remainder terms $(R_i^N)_{i=1,2}$ cancel as $N \to \infty$ and to apply Theorem 3.3 to each symmetric statistic J_1^N, J_2^N, J_3^N separately.
Let us now decompose the functions $(J_i^N)_{1\leq i\leq 3}$. First we note that

$$J_1^N = \frac{1}{N}\sum_{i=1}^{N} a(x^i, x^i) + \frac{1}{N}\sum_{i<j}(a + a^*)(x^i, x^j)$$

where we recall that $a \in L^2(\eta \otimes \eta)$ and satisfies

$$\int_{\Sigma_T} a(z, z)\ \eta(dz) = \int_{\Sigma_T} a(x, z)\ \eta(dz) = \int_{\Sigma_T} a(z, x)\ \eta(dz) = 0 \qquad \forall x \in \Sigma_T$$

∎

Therefore, a clear application of theorem 3.3 yields

$$\lim_{N\to\infty} J_1^N =_{\text{law}} \frac{1}{2} I_2(a + a^*). \tag{23}$$

In a second time, we discuss the limit of J_2^N.
For that aim we introduce the decomposition

$$\begin{aligned}
J_2^N &= -\frac{1}{2N^2}\sum_{i=1}^{N}\sum_{n=1}^{T}(q_n(x^i, x^i) - 1)^2 - \frac{1}{N^2}\sum_{(i,j)\neq}\sum_{n=1}^{T}(q_n(x^i, x^i) - 1)(q_n(x^i, x^j) - 1) \\
&\quad -\frac{1}{2N^2}\sum_{(i,j)\neq}\sum_{n=1}^{T}(q_n(x^i, x^j) - 1)^2 - \frac{1}{2N^2}\sum_{(i,j,k)\neq}\sum_{n=1}^{T}(q_n(x^i, x^j) - 1)(q_n(x^i, x^k) - 1) \\
&= J_{2,1}^N + J_{2,2}^N + J_{2,3}^N + J_{2,4}^N
\end{aligned}$$

As it is easily seen $J_{2,1}^N$ is of order $(1/N)$ since q_n belongs to $L^2(\eta \otimes \eta)$. Also, by the law of large numbers under $\eta^{\otimes N}$, we have almost surely

$$\lim_{N\to\infty} J_{2,2}^N = \int \sum_{n=1}^{T} (q_n(x,x)-1)(q_n(x,y)-1)\eta(dx)\eta(dy) = 0$$

Similarly, $\eta^{\otimes N}$-almost surely,

$$\lim_{N\to\infty} J_{2,3}^N = -\frac{1}{2}\int \sum_{n=1}^{T} (q_n(x,x)-1)^2 \; \eta(dx). \tag{24}$$

Finally let us decompose $J_{2,4}^N$

$$J_{2,4}^N = -\frac{1}{2N^2} \sum_{(i,j,k)\neq} b(x^i, x^j, x^k)$$

with

$$b(x,y,z) = \sum_{n=1}^{T} (q_n(x,y)-1)(q_n(x,z)-1)$$

After some manipulations one gets the decomposition

$$\begin{aligned} J_{2,4}^N &= -\frac{1}{2N^2} \sum_{i<j<k} b^{\sim}(x^i,x^j,x^k) - \frac{N-2}{2N^2} \sum_{(j,k)\neq} \overline{b}(x^j,x^k) \\ &= -\frac{1}{2N^2} \sum_{i<j<k} b^{\sim}(x^i,x^j,x^k) - \frac{N-2}{N^2} \sum_{i<j} \overline{b}(x^i,x^j) \end{aligned} \tag{25}$$

with

$$\begin{aligned} \overline{b}(x,y) &= \int b(u,x,y)\;\eta(du) \\ &= \int \sum_{n=1}^{T} (q_n(u,x)-1)(q_n(u,y)-1)\;\eta(du) \end{aligned}$$

and where $b^{\sim} \in L^3_{\text{sym}}(\eta^{\otimes 3})$, $\int b^{\sim}(x,y,z)\;\eta(dz) = 0$ and $\int \overline{b}(x,y)\,\eta(dy) = 0$. Theorem 3.3 shows that the first term in the r.h.s. of (25) is of order $N^{-\frac{1}{2}}$ and therefore vanishes. Furthermore, Theorem 3.3 also applies to the first term and leads to

$$\lim_{N\to\infty} J_{2,4}^N = -\frac{1}{2} I_2(\overline{b}). \tag{26}$$

If we combine (24) and (26) one gets

$$\lim_{N\to\infty} J_2^N = -\frac{1}{2}\int \sum_{n=1}^{T} (q_n(x,x)-1)^2\ \eta(dx) - \frac{1}{2}I_2(\overline{b}) \tag{27}$$

Rewriting J_3^N in the following way

$$J_3^N = \frac{1}{N}\sum_{i<j} c(x^i,x^j) + \frac{1}{2N}\sum_{i=1}^{N} c(x^i,x^i)$$

with $c(x,y) = \sum_{n=1}^{T} (\overline{q}_n(x)-1)(\overline{q}_n(y)-1)$ and noting that

$$c \in L^2_{\mathrm{sym}}(\eta\otimes\eta) \qquad \text{and} \quad \int c(x,z)\ \eta(dz) = 0$$

a clear application of theorem 3.3 yields

$$\lim_{N\to\infty} J_3^N =_{\mathrm{law}} \frac{1}{2}I_2(c) + \frac{1}{2}\int \sum_{n=1}^{T} (\overline{q}_n(x)-1)^2\eta(dx). \tag{28}$$

By virtue of Theorem 3.3, we deduce from (23), (27) and (28) that the limit of the sum of symmetric statistics $J_1^N + J_2^N + J_3^N$ is given by

$$\begin{aligned}\lim_{N\to\infty} J_1^N + J_2^N + J_3^N =_{law} &\ \frac{1}{2}\Big(I_2(a+a^*) - I_2(\overline{b}) + I_2(c)\Big)\\ &+\frac{1}{2}\int\sum_{n=1}^{T}\Big((\overline{q}_n(x)-1)^2 - (q_n(x,x)-1)^2\Big)\,d\eta(x)\end{aligned} \tag{29}$$

It is now easily seen that

$$\begin{aligned}\overline{b}(x,y) - c(x,y) &= \sum_{n=1}^{T}\int (q_n(u,x)-\overline{q}_n(x))(q_n(u,y)-\overline{q}_n(y))\ \eta_n(du)\\ &= \int a(u,x)\ a(u,y)\ \eta(du)\end{aligned}$$

and

$$\begin{aligned}&\sum_{n=1}^{T}\int((q_n(x,y)-1)^2 - (\overline{q}_n(x)-1)^2)\eta(dy)\eta(dx)\\ &= \sum_{n=1}^{T}\int (q_n(x,y)-\overline{q}_n(y))^2\eta(dy)\eta(dx)\\ &= \int a(x,y)^2\eta(dx)\eta(dy) = \mathrm{Trace} AA^*\end{aligned}$$

Therefore we obtain from (29) that

$$\lim_{N\to\infty}(J_1^N + J_2^N + J_3^N) =_{law} \frac{1}{2} I_2(f) - \frac{1}{2}\mathrm{Trace} AA^*$$

with

$$f(x,z) = a(x,z) + a(z,x) - \int_{\Sigma_T} a(u,x)\ a(u,z)\ \eta(du).$$

To complete the proof of the theorem the only point is to check that the remainder terms R_1^N and R_2^N cancel as $N \to \infty$.
We begin by noting that the law of the iterated logarithm clearly implies that R_2^N vanishes as N is going to infinity.
On the other hand in accordance with (18) and (20) we have

$$|R_1^N| \leq \frac{2}{3}\ |J_2^N|\ \sum_{n=1}^{T}\ \max_{1\leq i\leq N} |\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1|\ \max_{1\leq i\leq N} \theta_n^i(x)^3$$

To take the final step we notice that for any $n \geq 1$ and $u \in E$

$$E_{\eta_{n-1}^{\otimes N}}\left((\frac{1}{N}\sum_{j=1}^{N} q_n(u,x^j) - 1)^6 \right) \leq \frac{C}{N^3}\int_E\ (q_n(u,v)-1)^6\ \ \eta_{n-1}(dv)$$

for some finite constant $C < \infty$. Therefore for any $\epsilon > 0$ and $n \geq 1$

$$\begin{aligned} \eta^{\otimes N}\left(\max_{1\leq i\leq N} |\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1| > \epsilon\right) &\leq \sum_{i=1}^{N} \eta^{\otimes N}\left(|\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1| > \epsilon\right) \\ &\leq \frac{C}{\epsilon^6\ N^2}\int_E\ (q_n(u,v)-1)^6\ \ \eta_n(du)\eta_{n-1}(dv) \end{aligned} \tag{30}$$

Under the exponential moment condition (H1) the last term in the right hand side of (30) is bounded so that Borel-Cantelli's lemma gives

$$\lim_{N\to\infty}\max_{1\leq i\leq N} |\frac{1}{N}\sum_{j=1}^{N} q_n(x^i,x^j) - 1| = 0 \qquad \text{a.e.}$$

¿From the above it is also not difficult to check that

$$\lim_{N\to\infty}\max_{1\leq i\leq N} \theta_n^i(x) = 1 \qquad \text{a.e.}$$

It follows from the above limits that $\lim_{N\to\infty} R_1^N = 0$. This ends the proof of the first point of lemma 3.2.

The proof of the second point directly follows from theorem 3.3 since it shows that $W_T^N(\varphi)$ converges in law towards $I_1(\varphi)$.

■

4 Applications

At first sight the hypothesis (H0) and (H1) may seem very restrictive since we assume the existence of a reference probability measure under which the transition probabilities of the signal are absolutely continuous together with an additional exponential moment condition. The aim of this section is to show that these conditions cover many typical examples of non linear filtering problems.

We now turn to some examples of non linear filtering problems where conditions (H0) and (H1) hold. We begin with some examples of observation noise sources satisfying (H0)

Example 1 *As a typical example of non linear filtering problem assume the functions $h_n : E \to \mathbb{R}^d$, $n \geq 1$, are bounded continuous and the densities g_n given by*

$$g_n(v) = \frac{1}{((2\pi)^d |R_n|)^{1/2}} \ \exp\left(-\frac{1}{2} v' \ R_n^{-1} \ v\right)$$

where R_n is a $d\times d$ symmetric positive matrix. This correspond to the situation where the observations are given by

$$Y_n = h_n(X_{n-1}) + V_n \qquad \forall n \geq 1 \tag{31}$$

where $(V_n)_{n\geq 1}$ is a sequence of $\mathbb{R}^d$-valued and independent random variables with Gaussian densities.

After some easy manipulations one gets the bounds (10) with

$$\log \alpha_n(y) = \frac{1}{2} \ \|R_n^{-1}\| \ \|h_n\|^2 + \|R_n^{-1}\| \ \|h_n\| \ \|y\|$$

where $\|h_n\| = \sup_{x\in E} |h_n(x)|$ and $\|R_n^{-1}\|$ is the spectral radius of R_n^{-1}.

Example 2 *Suppose that $d = 1$ and g_n is a Cauchy density*

$$g_n(v) = \frac{\theta_n}{\pi \ (v^2 + \theta_n^2)} \qquad \theta_n > 0$$

In this situation we have that

$$\frac{g_n(y-h_n(x))}{g_n(y)} = \frac{y^2+\theta_n^2}{(y-h_n(x))^2+\theta_n^2} \qquad \forall (y,x) \in I\!R \times E$$

Notice that

$$\frac{y^2+\theta_n^2}{y^2+\theta_n^2+\|h_n\|^2+2|y|\,\|h_n\|} \leq \frac{g_n(y-h_n(x))}{g_n(y_n)} \leq 1+\left(\frac{y}{\theta_n}\right)^2$$

It follows that (10) holds with

$$\alpha_n(y) = 1+\left(\left(\frac{y}{\theta_n}\right)^2 \vee \frac{(|y|+\|h_n\|)^2}{y^2+\theta_n^2}\right)$$

Example 3 *Suppose $d=1$ and g_n is a bilateral exponential density*

$$g_n(v) = \frac{1}{2}\,\alpha_n \exp\left(-\alpha_n|v|\right) \qquad \alpha_n > 0$$

In this case

$$\frac{g_n(y-h_n(x))}{g_n(y_n)} = \exp\left(\alpha_n\left(|y|-|y-h_n(x)|\right)\right)$$

Observe that

$$-\|h_n\| \leq |y|-|y-h_n(x)| \leq \|h_n\| \qquad \forall (y,x) \in I\!R \times E$$

One concludes that (10) is satisfied with $\alpha_n(y) = \exp\left(\alpha_n\|h_n\|\right)$.

Let us now investigate condition (H1) through some examples of signals that can be handled in our framework.

Example 4 *Suppose that $E = I\!R^m$, $m \geq 1$ and K_n, $n \geq 1$ are given by*

$$K_n(x,dz) = \frac{1}{((2\pi)^m|Q_n|)^{1/2}}\ \exp\left(-\frac{1}{2}(z-b_n(x))'\,Q_n^{-1}\,(z-b_n(x))\right)$$

where Q is a $m \times m$ symmetric non negative matrix and $b_n : I\!R^m \to I\!R^m$ is a bounded continuous function. This correspond to the situation where the signal process is given by

$$X_n = b_n(X_n) + W_n \qquad \forall n \geq 1 \tag{32}$$

where $(W_n)_{n\geq 1}$ is a sequence of $\mathbb{R}^m$-valued and independent random variables with Gaussian densities.
It is not difficult to check that (H1) is satisfied with

$$\lambda_n(dz) = \frac{1}{((2\pi)^m |Q_n|)^{1/2}} \; \exp\,(-\frac{1}{2} z' \, Q_n^{-1} \; z) \; dz.$$

Indeed, we then find out that

$$\log \frac{d\delta_x K_n}{d\lambda_n} = const. - b_n(x)' Q_n^{-1} z$$

which insures the exponential moment condition (H1) with

$$\varphi_n(z) = \frac{1}{2} \|b_n\|^2 \|Q_n^{-1}\| + \|Q_n^{-1}\| \; \|b_n\| \; |z| \qquad \forall z \in \mathbb{R}.$$

Example 5 *Suppose $E = \mathbb{R}$ and K_n, $n \geq 1$, are given by*

$$K_n(x, dz) = \frac{1}{2}\, \alpha \exp\,(-\alpha |z - b(x)|) dz \qquad \alpha > 0 \quad b \in \mathcal{C}_b(\mathbb{R})$$

This corresponds to the situation where the signal process X is given by

$$X_n = b(X_{n-1}) + W_n \qquad n \geq 1$$

where $(W_n)_{n\geq 1}$ is a sequence of $\mathbb{R}^m$-valued and independent random variables with bilateral exponential densities. Note that K_n may be written

$$K_n(x, dz) = \frac{1}{2}\, \alpha \exp\,(\alpha(|z| - |z - b(x)|)) \;\; \lambda_n(dz)$$

with

$$\lambda_n(dz) = \frac{1}{2}\, \alpha \exp\,(-\alpha |z|) \; dz$$

It follows that (H1) holds since $\log \frac{\delta_x K_n}{d\lambda_n}(z) \leq \alpha \|b\|$

Finally for completeness we examine a situation where (H1) is not satisfied

Example 6 *Let us suppose that $E = \mathbb{R}$ and*

$$K_n(x, dz) = \sqrt{\frac{\epsilon_n(x)}{2\pi}} \; \exp\left(-\frac{1}{2}\epsilon_n(x) \; z^2\right) \, dz$$

where $\epsilon_n : \mathbb{R} \to \mathbb{R}$ is a continuous function such that

$$\forall x \in \mathbb{R} \qquad \epsilon_n(x) > 0 \qquad and \qquad \lim_{|x|\to\infty} \epsilon_n(x) = 0.$$

Let us assume that K_n satisfies (H1) for some function φ_n. Since $\delta_x K_n$ is absolutely continuous with respect to Lebesgue measure for any $x \in E$, the probability measure λ_n described in (H1) is absolutely continuous with respect to Lebesgue measure. Therefore, there exist a probability density p_n such that

$$\forall x, z \in I\!R \qquad e^{-\varphi_n(z)}\, p_n(z) \leq \sqrt{\epsilon_n(x)}\, \exp\left(-\frac{1}{2}\epsilon_n(x)\, z^2\right) \leq e^{\varphi_n(z)}\, p_n(z).$$

Letting $|x| \to \infty$ one gets $e^{-\varphi_n(z)} p_n(z) = 0$ for any $z \in I\!R$ which is absurd since we also assumed $\int e^{\varphi_n(z)} p_n(z) dz < \infty$.

The interacting particle system model (3) is not only design to solve the non linear filtering equation. It can also model systems which arise in Biology and in Physics. This class of particle scheme are also used to solve numerical function optimization problems, image processing,...(see for instance [21, 22, 23, 37, 38] and references therein). The settings are the same as before except that the functions $\{g_n(y_n - h_n(\cdot))\ ;\ n \geq 1\}$ are replaced by a sequence of deterministic and bounded positive functions still denoted by $\{g_n(\cdot)\ ;\ n \geq 1\}$.
In this framework the limiting measure valued system (2) is used to predict the evolution in time of the finite population model (3). In contrast to the non linear filtering settings a crucial practical advantage of this situation is that the state space E is finite so that the analysis of the limiting process is much simpler. In view of the preceding development the conclusions of theorem 3.1 remain valid if we replace the hypothesis (H0) by the condition (H'0) given by

(H'0) For any time $n \geq 1$ there exists some $\alpha_n > 0$ such that

$$\alpha_n^{-1} \leq g_n(x) \leq \alpha_n \qquad \forall x \in E$$

To our knowledge the Central Limit Theorem 3.1 is the first result of this kind in the theory of genetic-type algorithms.
These models inspired by natural evolution usually encode a potential solution to a specific problem on a simple chromosome-like information. In this description of the genetic algorithm each chromosome is modelled by a binary string of fixed length. The resulting genetic algorithm is a Markov chain with state space E^N where N is the size of the system and the state space of each particle is $E = \{0, 1\}$. In this situation the mutation transition at time $n \geq 1$ is obtained by flipping each particle of each chromosome of the population with the probability $p_n > 0$ so that the transition probability kernels K_n have

the form

$$K_n(x,z) = p_n\ \delta_{x+1}(z) + (1-p_n)\ \delta_x(z)$$

To see that (H1) is satisfied it suffices to note that

$$0 < \epsilon_n \leq K_n(x,z) \leq 1 \qquad \forall (x,z) \in E \times E$$

with

$$\epsilon_n \stackrel{\text{def}}{=} \min(p_n, 1-p_n)$$

Concluding Remarks

In the current work we have presented an approximation of the two steps transitions of the system (2)

$$\eta_{n-1} \xrightarrow{\text{Updating}} \widehat{\eta}_{n-1} \stackrel{\text{def}}{=} \psi_n(Y_n, \eta_{n-1}) \xrightarrow{\text{Prediction}} \eta_n = \widehat{\eta}_{n-1} K_n$$

by a two steps Markov chain taking values in the set of empirical distributions. Namely,

$$\eta_{n-1}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n-1}^i} \xrightarrow{\text{Selection}} \widehat{\eta}_{n-1}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{\xi}_{n-1}^i} \xrightarrow{\text{Mutation}} \eta_n = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}$$

In non linear filtering settings η_n is usually called the one step predictor and $\widehat{\eta}_n$ is the optimal filter. In [16, 18] we prove that the particle density profiles $\{\widehat{\eta}_n^N\ ;\ N \geq 1\}$ weakly converge to the optimal filter $\widehat{\eta}_n$ as $N \to \infty$. One open question is to study the Central Limit Theorem for the empirical distributions

$$\frac{1}{N}\sum_{i=1}^N \delta_{(\xi_0^i, \widehat{\xi}_0^i, \ldots, \xi_T^i, \widehat{\xi}_T^i)}$$

In contrast to the situation examined before the main difficulty here comes from the fact that the distribution of the particle systems

$$\{\xi_n^i, \widehat{\xi}_n^i\ ;\ 0 \leq n \leq T\}$$

is usually not absolutely continuous with respect to some product measure.

Several variants of the genetic-type particle scheme studied in this paper have been recently suggested in [8, 9, 10, 11]. In these variants the size of the system is not fixed but random. It is obvious that the situation becomes considerably more involved when the size of the particle systems is random and the study of the fluctuations of such schemes requires a more delicate analysis.

All the results presented in this study and the referenced papers on the subject provide some information about the speed of convergence of a class of particle algorithms until a finite given time. The study of their long time behavior is a rather different subject which we will investigate in another paper.

Acknowledgments :

The authors wish to thank Professor Dominique Bakry and Professor Michel Ledoux for fruitful discussions. We also wish to thank the anonymous referee for his suggestions and comments that improve the presentation of this work. We acknowledge the European Community for the research fellowships CEC Contract No ERB-FMRX-CT96-0075 and INTAS-RFBR 95-0091.

References

[1] BEN AROUS, G., BRUNAUD, M. (1990). Methode de Laplace : Etude variationnelle des fluctuations de diffusions de type “champ moyen”. *Stochastics* **31-32**, 79-144

[2] BENES B.E. (1981). Exact Finite-Dimensional Filters for Certain Diffusions with Nonlinear Drift. *Stochastics,* **Vol.5**, *pp 65-92.*

[3] BOLTHAUSEN, E.(1986) Laplace approximation for sums of independant random vectors I.*Prob. Th. Rel. Fields* **72**, 305-318

[4] R.S. BUCY, Lectures on discrete time filtering, Springer Verlag, Signal Processing and Digital Filtering, 1994.

[5] CARVALHO H. (1995), Filtrage optimal non linéaire du signal GPS NAVSTAR en racalage de centrales de navigation, *Thèse de L'Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse.*

[6] CARVALHO H., DEL MORAL P., MONIN A., SALUT G. (1997). Optimal non-linear filtering in GPS/INS Integration. *IEEE Trans. on Aerospace and Electronic Systems*, July 1997, Volume 33, Number 3, pp. 835-850.

[7] CHALEYAT-MAUREL M., MICHEL D. (1983), Des résultats de non existence de filtres de dimension finie. *C.R.Acad.Sci. Paris, Serie I,* **t. 296**.

[8] CRISAN D., LYONS T.J. (1997). Nonlinear Filtering And Measure Valued Processes. Proba. Theory and Related Fields, 109, 217-244.

[9] CRISAN D., LYONS T.J. (1996). Convergence of a Branching Particle Method to the Solution of the Zakai Equation.*Imperial College, London, preprint.*

[10] CRISAN D., J. GAINES, LYONS T.J. (1998). A Particle Approximation of the Solution of the Kushner-Stratonovitch Equation. To appear in SIAM Journal of Applied Math., 1998.

[11] CRISAN D., DEL MORAL P., LYONS T.J. (1997), Non Linear Filtering Using Branching and Interacting Particle Systems, *Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier,* **No 01-98**.

[12] DEL MORAL P., RIGAL G., NOYER J.C., SALUT G. (1993), Traitement non-linéaire du signal par reseau particulaire : Application Radar. *quatorzième colloque GRETSI, Juan les Pins, 13-16 Septembre.*

[13] DEL MORAL P., NOYER J.C., SALUT G. (1995). Résolution particulaire et traitement non-linéaire du signal : Application Radar/Sonar. *Traitement du signal, Septembre.*

[14] DEL MORAL P. (1996). Non linear filtering using random particles. *Theor. Prob. Appl.,* **Vol. 40**, *(4).*

[15] DEL MORAL P. (1996). Non-linear Filtering: Interacting particle resolution. *Markov Processes and Related Fields,* **Vol.2**

[16] DEL MORAL P. (1996). Measure Valued Processes and Interacting Particle Systems. Application to Non Linear Filtering Problems. *Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier,* **No 15-96**. *To appear in Annals of Applied Proba.*

[17] DEL MORAL P. (1996). A Uniform Theorem for the Numerical Solving of Non Linear Filtering Problems. *Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier, No 14-96, to appear in Journal of Applied Probability* .

[18] DEL MORAL P (1997), Filtrage non linéaire par systèmes de particules en intéraction. *C.R. Acad. Sci. Paris, t. 325, Série I, p. 653-658.*

[19] DEL MORAL P., GUIONNET A. (1997). Large Deviations for Interacting Particle Systems. Applications to Non Linear Filtering Problems. *Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier, No 05-97, submitted.*

[20] GUIONNET A. (1997). About Precise Laplace's Method; Applications to Fluctuations for Mean Field Interacting Particles, *preprint*

[21] GOLDBERG D.E. (1985), Genetic Algorithms and Rule Learning in Dynamic Control Systems, Proceedings of the First International Conference on Genetic Algorithms, Lawrence Erlbaum Associates Hillsdale, NJ, pp. 8-15.

[22] GOLDBERG D.E. (1987), Simple genetic algorithms and the minimal deceptive problem, in Lawrence Davis (Eds.), Genetic Algorithms and simulated annealing, Pitman.

[23] GOLDBERG D.E. (1989), Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA.

[24] GORDON N.J., SALMON D.J., SMITH A.F.M. (1993). Novel Approach to non-linear/non-Gaussian Bayesian state estimation.,*IEE*

[25] HOLLAND J.H. (1975) Adaptation in Natural and Artificial Systems.*University of Michigan Press, Ann Arbor*

[26] JAZWINSKI A.H. (1970). Stochastic Processes and Filtering Theory. *Academic Press.*

[27] KUNITA H. (1971). Asymptotic Behavior of Nonlinear Filtering Errors of Markov Processes. *Journal of Multivariate Analysis,* **Vol. 1***, No 4, p. 365-393*

[28] KUSUOKA, S.; TAMURA, Y.; Gibbs measures for mean field potentials. J.Fac.Sci.Univ. Tokyo ; Sect. IA, Math. 31(1984),

[29] LIANG D., McMILLAN J.. Development of a marine integrated navigation system. *Kalman Filter Integration of Modern Guidance and Navigation Systems, AGARD-LS-166, OTAN.*

[30] DYNKIN E.B., MANDELBAUM A. (1983), Symmetric Statisyics, Poisson Processes and Multiple Wiener Integrals, *Annals Statist.* **11***, 739-745*

[31] SIMON, B. (1977) Trace ideals and their applications. *London Mathematical Society Lecture Notes series* **35** Cambridge University press.

[32] STETTNER L. (1989). On Invariant Measures of Filtering Processes. *Stochastic Differential Systems, Lecture Notes in Control and Information Sciences,* **126***, Springer-Verlag.*

[33] SHIGA T., TANAKA, H. (1985) ; Central limit theorem for a system of markovian paricles with mean field interaction. Z.f. Wahrscheinlichkeitstheorie verw. Geb., 69, 439-459.

[34] SHIRYAEV A.N., Probability, Graduate Texts in Maths, 95, Second edition, Springer-Verlag, 1996.

[35] TANAKA, H. (1984) Limit theorems for certain diffusion processes. Proceedings of the Taniguchi Symp, Katata 1982, 469-488, Tokyo, Kinokuniya.

[36] VAN DOOTINGH M., VIEL F., RAKOTOPARA D., GAUTHIER J.P. (1991), Coupling of non-linear control with a stochastic filter for state estimation : Application on a free radical polimerization reactor,*I.F.A.C. International Symposium ADCHEM'91, Toulouse, France, October 14-15*

[37] VOSE M.D. (1993), Modelling simple genetic algorithms, in Foundations of Genetic Algorithms, Morgan Kaufmann.

[38] VOSE M.D. (1995), Modelling simple genetic algorithms, Elementary Computations, 3 (4), 453-472..

[39] WILLIAMS D. (1992) Probability with Martingales,Cambridge Mathematical textbooks.