*The Annals of Applied Probability*
1999, Vol. 9, No. 2, 275–297

# CENTRAL LIMIT THEOREM FOR NONLINEAR FILTERING AND INTERACTING PARTICLE SYSTEMS[1]

BY P. DEL MORAL AND A. GUIONNET

*LSP-CNRS Toulouse and LMSS-CNRS Orsay*

Several random particle systems approaches were recently suggested to solve nonlinear filtering problems numerically. 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.

## 1. Introduction.

1.1. *Background and motivations.* The nonlinear filtering problem consists in recursively computing the conditional distributions of a nonlinear 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 [7].

Although Kalman filtering [27, 30] is a popular tool in handling estimation problems, its optimality depends heavily on linearity. When used for nonlinear filtering (extended Kalman filter), its performance relies on and is limited by the linearization performed on the model concerned.

It has recently been emphasized that a more efficient way is to use random particle systems to solve the filtering problem numerically. That particle algorithms are gaining popularity is attested to by the list of referenced papers (see for instance [8, 12, 14, 15, 36] and references therein). 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 nonlinear filtering problems.

Several practical problems which have been solved using these methods are given in [5, 6, 18, 19] including radar/sonar signal processing and Global Positioning System/Inertial Navigation System (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 [13]. We have shown in our earlier papers [13, 14] that, under rather general assumptions, the particle density profiles converge to the desired conditional distributions of the signal when the number of particles

Received December 1997; revised March 1998.

[1]Supported by European Community Research Fellowship CEC Contract ERB-FMRX-CT96-0075 and INTAS-RFBR 95-0091.

*AMS* 1991 *subject classifications*. 60F05, 60G35, 93E11, 62L20.

*Key words and phrases*. Central limit, interacting random processes, filtering, stochastic approximation.

is growing. The study of the convergence or the empirical measure on path space and large deviation principles is presented in [17]. 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 nonlinear 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 nonlinear observation process $(Y_n;\ n \geq 0)$ taking values in $\mathbb{R}^d$ for some $d \geq 1$.

To describe our model precisely, 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 by $\mathscr{C}_b(E)$ the space of these functions.

The classical filtering problem can be summarized to find the conditional distributions

$$\eta_n(f) =_{\text{def}} E(f(X_n)/Y_1, \dots, Y_n) \qquad \forall\, f \in \mathscr{C}_b(E),\ n \geq 0. \tag{1}$$

It was proved in a rather general setting by Kunita [28] and Stettner [32] that, given a series of observations $Y = y$, the distributions $(\eta_n;\ n \geq 0)$ are the solution of a discrete-time measure-valued dynamical system of the form

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

where $\eta$ is the law of the initial value of the signal and $\phi(n, \cdot)$ 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 $E$, that is, 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

$$P(\xi_0 \in dx) = \prod_{p=1}^{N} \eta_0(dx^p),$$

$$P(\xi_n \in dx/\, \xi_{n-1} = z) = \prod_{p=1}^{N} \phi\biggl(n, \frac{1}{N}\sum_{q=1}^{N} \delta_{z^q}\biggr)(dx^p).$$

We use $dx =_{\text{def}} dx^1 \times \cdots \times dx^N$ to denote an infinitesimal neighborhood of $x = (x^1, \dots, x^N) \in E^N$.

In [17] we develop large deviations principles for the law of the empirical distributions

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

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

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

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

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

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

One obvious consequence of such a result is that the so-called nonlinear 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 the noise source. Several convergence theorems ensuring the convergence of the particle scheme toward the desired distribution were obtained in [13, 14, 16] and [17]. The results presented in this paper make 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 systems model under study. The main result of this paper is presented in Section 3. In Section 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 Section 3.2 we prove a central limit theorem for the empirical measures associated to our particle scheme. Our basic tools are the Dynkin–Mandelbaum theorem on symmetric statistics and multiple Wiener integrals as in [33]. Several examples of nonlinear filtering problems that can be handled in our framework are worked out in Section 4.

**2. Interacting particle systems for nonlinear filtering.** Several random particle systems approaches were recently suggested to solve nonlinear filtering problems numerically (see [14] and [8] for theoretical details and references). These approaches develop new methods for dealing with measure-valued dynamical system of the form (1). In Section 2.1 we present the so-called nonlinear filtering equations, then we present the genetic-type interacting particle approximation under study.

2.1. *Formulation of the nonlinear 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 $\eta_0$.

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. Here $Y$ 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 $h_n\colon E \to \mathbb{R}^d$, $n \geq 1$, 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 $\nu$.

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 to estimate the distribution of $X_n$ conditionally to the observations up to time $n$. Namely,

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

The following result is a slight modification of Kunita [28] and Stettner results [32] (see also [14] for details).

If, for any Markov transition $K$ on $E$ and any $\mu \in \mathbf{M}_1(E)$ we denote $\mu M$ the probability so that for any $f \in \mathscr{C}_b(E)$,

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

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 the solution of the $\mathbf{M}_1(E)$-valued dynamical system*

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

*where $y_n \in \mathbb{R}^d$ is the current observation and $\phi_n$ is the continuous function given by*

$$\phi_n(y_n, \eta) = \psi_n(y_n, \eta) K_n,$$

$$\psi_n(y_n, \eta) f = \frac{\int f(x)\, g_n(y_n - h_n(x))\, \eta(dx)}{\int g_n(y_n - h_n(z))\, \eta(dz)}$$

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

2.2. *Interacting particle systems approximations.* We have shown in our earlier papers [13, 14, 16, 17] that the nonlinear 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

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

Let us remark that

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

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 to two separate mechanisms. They can be modelled as follows. The initial particle system is

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

*Selection / updating.*

$$
P_y\bigl(\widehat{\xi}_{n-1} \in dx \mid \xi_{n-1} = z\bigr) = \prod_{p=1}^{N} \sum_{i=1}^{N} \frac{g_n(y_n - h_n(z^i))}{\sum_{j=1}^{N} g_n(y_n - h_n(z^j))}\, \delta_{z^i}(dx^p).
$$

*Mutation / prediction.*

$$
P_y\bigl(\xi_n \in dz \mid \widehat{\xi}_{n-1} = x\bigr) = \prod_{p=1}^{N} K_n(x^p, dz^p). \tag{8}
$$

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 $\xi_{n-1} = (\xi_{n-1}^1, \dots, \xi_{n-1}^N)$ and chooses randomly a site $\xi_{n-1}^i$ 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 to the transition probability kernel of the signal process.

We see that this particle approximation of the nonlinear 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 Holland [26] 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 $\varphi_n$ so that

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

Moreover, there exists a nonnegative function $\varphi_n$ such that, for any $p \geq 1$,

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

We now give some comments on hypotheses (H0) and (H1). First, we note that very similar assumptions were introduced in [17] 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 nonlinear 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

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

In such a context, it is not difficult to see that, for any given finite time $T$, if we denote by $P_T^N$ 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

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

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 the precise Laplace method is now extensively developed (see for instance [29, 35, 3, 1, 21] 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, the precise Laplace's method can be developed (see [1, 29, 21]). In [21], 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 closely related to Shiga and 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 1. This simplifies the analysis considerably. In fact, the proof then mainly relies on a simple formula on multiple Wiener integrals and the Dynkin–Mandelbaum theorem [20] 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 $(\zeta_n)_{n\ge 0}$ with transitions

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

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

($\mathscr{Q}$) There exists constants $c_1, c_2 > 0$ such that

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

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

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

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 $Q(x, x', dz)$ does not depend on the parameter $x'$, the measure kernel $Q$ 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 $(\mathscr{Q})$ is given by:

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

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

A clear disadvantage of condition $(\mathscr{Q}')$ is that it is in general not satisfied when $E$ is not compact and in particular in many nonlinear filtering problems. Our purpose here is to extend the technic of [33] to handle the particle approximation introduced in Section 2.2 and to replace the assumption $(\mathscr{Q}')$ 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 $(y_n;\ 1 \le n \le T)$.

Under (H1) for any $n \ge 1$ 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

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

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

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

with

$$q_{n,y}(x, z) = \frac{g_n(y_n - h_n(z_{n-1}))\ k_n(z_{n-1}, x_n)}{\int_E g_n(y_n - h_n(u))\ k_n(u, x_n)\eta_{n-1}(du)},$$

$$a_{T,y}(x, z) = q_{[0,T],y}(x, z) - \int_{\Sigma_T} q_{[0,T],y}(x', z)\ \eta_{[0,T]}(dx').$$

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

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

is a Hilbert–Schmidt operator on $L^2(\Sigma_T, \eta_{[0,T]})$. 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* $\{W_T^N(\varphi);\ \varphi \in L^2(\eta_{[0,\,T]})\}$ *converges as* $N \to \infty$ *to a centered Gaussian field* $\{W_T(\varphi);\ \varphi \in L^2(\eta_{[0,\,T]})\}$ *satisfying*

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

*for any* $\varphi_1, \varphi_2 \in L^2(\eta_{[0,\,T]})$, *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 $\eta$ instead of $\eta_{[0,\,T]}$. In addition we will write $\phi(n, \cdot)$ instead of $\phi_n(y_n, \cdot)$ and $g_n(\cdot)$ instead of $g_n(y_n - h_n(\cdot))$.

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

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

Furthermore, by definition of $a$ and the fact that $\eta$ is a product measure, it is easily checked that

$$\forall\ n \geq 2, \qquad \operatorname{Trace} A^n = 0.$$

Standard spectral theory (see [31] for instance) then shows that $\det_2(I - A)$ is equal to 1 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

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

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

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

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 $E_{\eta^{\otimes N}}(\cdot)$ [resp. $E_{P^N}(\cdot)$] to denote expectations with respect to the measure $\eta^{\otimes N}$ (resp. $P^N$) on $\Sigma_T^N$ and, unless otherwise stated, the sequence $\{x^i;\ i \geq 1\}$ is regarded as a sequence of $\Sigma_T$-valued and independent random variables with common law $\eta$.

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, page 189 in [34] or Scheffé's Lemma 5.10, page 55 in [39]) which says that, if a sequence of nonnegative random variables $\{X_N;\ N \geq 1\}$ converges almost surely towards some random variable $X$ as $N \to \infty$, then we have

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

The equivalence still holds if $X_N$ only converges in distribution by Skorohod's theorem (see, for instance, Theorem 1, page 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 $\varphi$ 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

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

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

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

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 $\phi \in L^2_{\text{sym}}(\eta^{\otimes m})$ is then defined by a completion argument.

We are going to prove that, if $=_{\text{law}}$ denotes the equality in law, we have the following lemma.

LEMMA 3.2.

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

*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 $\varphi \in L^2(\eta)$,*

$$\lim_{N\to\infty} (H^N(x) + iW_T^N(\varphi)) =_{\text{law}} \tfrac{1}{2} I_2(f) + iI_1(\varphi) - \tfrac{1}{2}\operatorname{Trace} AA^*.$$

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

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

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

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

The proof of Theorem 3.1 is thus complete. □

PROOF OF LEMMA 3.2. Since $\eta_n = \phi(n, \eta_{n-1})$ and

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

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

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

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

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

with

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

where

$$\overline{q}_n(x^j) = \frac{g_n(x^j_{n-1})}{\eta_{n-1}(g_n)}.$$

Following (17), we also have

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

By the representation

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

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

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

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

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

with

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

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

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 $H^N(x^1, \dots, x^N)$. To this end, we shall rely on $L^2$ technics and, more precisely, the Dynkin–Mandelbaum construction of multiple Wiener integrals as a limit of symmetric statistics. For completeness we present this result.

Let $\{\zeta^i;\ i \ge 1\}$ be a sequence of independent and identically distributed random variables with values in an arbitrary measurable space $(\mathscr{X}, \mathscr{B})$. To every symmetric function $h(z_1, \dots, z_m)$ there corresponds a statistic,

$$\sigma_m^N(h) = \sum_{1 \le i_1 < \cdots < i_m \le N} h(\zeta^{i_1}, \dots, \zeta^{i_m})$$

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

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

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

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

where $\mu$ is the probability distribution of $\zeta^1$.

We call such functions $\{h_m;\ m \geq 0\}$ canonical. Finally we denote by $\mathscr{H}$ the set of all sequences $h = (h_0, h_1(z_1), \dots, h_m(z_1, \dots, z_m), \dots)$ where $h_m$ are canonical and

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

As in [33] we will use the following theorem repeatedly.

THEOREM 3.3 (Dynkin–Mandelbaum [20]). *For $h \in \mathscr{H}$ the sequence of random variables $Z_N(h) = \sum_{m \geq 0}(1/N^{m/2})\sigma_m^N(h_m)$ 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 $x = (x^1, \dots, x^N)$, which is a sequence of $\Sigma_T$-valued and independent random variables with common law $\eta$, the above theorem is applicable. The technical trick to identify the limit of this functional is to decompose the symmetric statistic $H^N(x^1, \dots, x^N)$ 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

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

For a second time, we discuss the limit of $J_2^N$: we introduce the decomposition

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

As 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,

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

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

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

with

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

and where $b^{\sim} \in L^3_{\mathrm{sym}}(\eta^{\otimes 3})$, $\int b^{\sim}(x, y, z)\, \eta(dz) = 0$ 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 $N^{-1/2}$ and therefore vanishes. Furthermore, Theorem 3.3 also applies to the first term and leads to

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

Combining (24) and (26), one gets

$$\lim_{N\to\infty} J^N_2 = -\tfrac{1}{2} \int \sum_{n=1}^{T} (q_n(x, x) - 1)^2\, \eta(dx) - \tfrac{1}{2} I_2(\overline{b}). \tag{27}$$

Rewriting $J^N_3$ in the following way

$$J^N_3 = \frac{1}{N} \sum_{i<j} c(x^i, x^j) + \frac{1}{2N} \sum_{i=1}^{N} c(x^i, x^i)$$

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

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

a clear application of Theorem 3.3 yields

$$\lim_{N\to\infty} J^N_3 =_{\mathrm{law}} \tfrac{1}{2} I_2(c) + \tfrac{1}{2} \int \sum_{n=1}^{T} (\overline{q}_n(x) - 1)^2 \eta(dx). \tag{28}$$

By virtue of Theorem 3.3, we deduce from (23), (27) and (28) that the limit of the sum of symmetric statistics $J^N_1 + J^N_2 + J^N_3$ is given by

$$\begin{aligned} \lim_{N\to\infty} J^N_1 + J^N_2 + J^N_3 &=_{\mathrm{law}} \tfrac{1}{2}(I_2(a + a^*) - I_2(\overline{b}) + I_2(c)) \\ &\quad + \tfrac{1}{2} \int \sum_{n=1}^{T} \big((\overline{q}_n(x) - 1)^2 - (q_n(x, x) - 1)^2\big) d\eta(x). \end{aligned} \tag{29}$$

It is now easily seen that

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

and

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

Therefore we obtain from (29) that

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

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

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

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

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

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

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

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

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

From the above it is also not difficult to check that

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

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 follows directly from Theorem 3.3 since it shows that $W_T^N(\varphi)$ converges in law toward $I_1(\varphi)$.

**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 nonlinear filtering problems.

We now turn to some examples of nonlinear 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 a nonlinear filtering problem, assume the functions $h_n\colon E \to \mathbb{R}^d$, $n \ge 1$, are bounded continuous and the densities $g_n$ given by

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

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

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

where $(V_n)_{n\ge 1}$ 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

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

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

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

Notice that

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

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

$$g_n(v) = \tfrac{1}{2}\,\alpha_n \exp(-\alpha_n|v|), \qquad \alpha_n > 0.$$

In this case,

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

Observe that

$$-\|h_n\| \le |y|-|y-h_n(x)| \le \|h_n\| \qquad \forall (y,x)\in\mathbb{R}\times E.$$

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

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

EXAMPLE 4. Suppose that $E=\mathbb{R}^m$, $m\ge 1$ and $K_n$, $n\ge 1$ 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 nonnegative matrix and $b_n\colon \mathbb{R}^m \to \mathbb{R}^m$ is a bounded continuous function. This corresponds to the situation where the signal process is given by

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

where $(W_n)_{n\ge 1}$ 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

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

Indeed, we then find out that

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

which insures the exponential moment condition (H1) with

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

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

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

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 as

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

with

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

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

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

EXAMPLE 6. Let us suppose that $E = \mathbb{R}$ and

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

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

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

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

$$\forall\, x, z \in \mathbb{R} \qquad \exp(-\varphi_n(z))\,p_n(z) \leq \sqrt{\varepsilon_n(x)}\exp\left(-\tfrac{1}{2}\varepsilon_n(x)\,z^2\right) \leq \exp(\varphi_n(z))\,p_n(z).$$

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

The interacting particle system model (3) is not only designed to solve the nonlinear filtering equation. It can also model systems which arise in biology and physics. This class of particle scheme is also used to solve numerical function optimization problems, image processing and others (see, for instance, [22, 23, 24, 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 nonlinear 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 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

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

with

$$\varepsilon_n =_{\mathrm{def}} \min(p_n, 1 - p_n).$$

CONCLUDING REMARKS. In the current work we have presented an approximation of the two-step transitions of the system (2),

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

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 nonlinear filtering settings, $\eta_n$ is usually called the one-step predictor and $\widehat{\eta}_n$ is the optimal filter. In [14, 16] 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 because the distribution of the particle systems

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

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 [11, 10, 9, 8]. 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 suggestions and comments that improved the presentation of this work.

## REFERENCES

[1] BEN AROUS, G. and 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* **5** 65–92.

[3] BOLTHAUSEN, E. (1986). Laplace approximation for sums of independant random vectors I. *Probab. Theory Related Fields* **72** 305–318.

[4] BUCY, R. S. (1994). Lectures on discrete time filtering. In *Signal Processing and Digital Filtering.* Springer, New York.

[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. and SALUT, G. (1997). Optimal non-linear filtering in GPS/INS integration. *IEEE Trans. Aerospace Electron. Systems* **33** 835–850.

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

[8] CRISAN, D., DEL MORAL, P. and LYONS, T. J. (1997). Non linear filtering using branching and interacting particle systems. Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier **01-98**.

[9] CRISAN, D., GAINES, J. and LYONS, T. J. (1998). A particle approximation of the solution of the Kushner–Stratonovitch equation. *SIAM J. Appl. Math.* To appear.

[10] CRISAN, D. and LYONS, T. J. (1996). Convergence of a branching particle method to the solution of the Zakai equation. Imperial College, London. Preprint.

[11] CRISAN, D. and LYONS, T. J. (1997). Nonlinear filtering and measure valued processes. *Probab. Theory Related Fields* **109** 217–244.

[12] DEL MORAL, P. (1996). Non linear filtering using random particles. *Theory Probab. Appl.* **40** 690–701.

[13] DEL MORAL, P. (1996). Non-linear filtering: interacting particle resolution. *Markov Processes Related Fields* **2** 555–580.

[14] DEL MORAL, P. (1996). Measure valued processes and interacting particle systems. Application to non linear filtering problems. Unpublished manuscript.

[15] 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* 14-96. *J. Appl. Probab.* To appear.

[16] DEL MORAL, P. (1997). Filtrage non linéaire par systèmes de particules en intéraction. *C.R. Acad. Sci. Paris Sér. I Math.* **325** 653–658.

[17] DEL MORAL, P. and GUIONNET, A. (1997). Large deviations for interacting particle systems. Applications to non linear filtering problems. *Stochastic Process. Appl.* To appear.

[18] DEL MORAL, P., NOYER, J. C. and SALUT, G. (1995). Résolution particulaire et traitement nonlinéaire du signal: application radar/sonar. *Traitement du signal* **12** 287–301.

[19] DEL MORAL P., RIGAL, G., NOYER, J. C. and SALUT, G. (1993). Traitement non-linéaire du signal par reseau particulaire: application radar. In *14th colloque GRETSI sur le Traitement du Signal et des Images* 399–402.

[20] DYNKIN, E. B. and MANDELBAUM, A. (1983). Symmetric statistics, Poisson processes and multiple Wiener integrals. *Ann. Statist.* **11** 739–745.

[21] GUIONNET, A. (1997). About precise Laplace's method; Applications to fluctuations for mean field interacting particles. Preprint.

[22] GOLDBERG, D. E. (1985), Genetic algorithms and rule learning in dynamic control systems. In *Proceedings of the First International Conference on Genetic Algorithms* 8–15. Erlbaum, Hillsdale, NJ.

[23] GOLDBERG, D. E. (1987). Simple genetic algorithms and the minimal deceptive problem. In *Genetic Algorithms and Simulated Annealing* (Lawrence Davis, ed.) Pitman, New York.

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

[25] GORDON, N. J., SALMON, D. J. and SMITH, A. F. M. (1993). Novel Approach to non-linear/non-Gaussian Bayesian state estimation. *IEE Proceedings* **140** 107–133.

[26] HOLLAND, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Univ. Michigan Press.

[27] JAZWINSKI, A. H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

[28] KUNITA, H. (1971). Asymptotic behavior of nonlinear filtering errors of Markov processes. *J. Multivariate Anal.* **1** 365–393.

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

[30] LIANG, D. and MCMILLAN, J. (1989). Development of a marine integrated navigation system. *Kalman Filter Integration of Modern Guidance and Navigation Systems, AGARD-LS-166, OTAN.*

[31] SIMON, B. (1977). *Trace Ideals and Their Applications.* Cambridge Univ. Press.

[32] STETTNER, L. (1989). On invariant measures of filtering processes. *Stochastic Differential Systems. Lecture Notes in Control and Inform. Sci.* **126**. Springer, New York.

[33] SHIGA, T. and TANAKA, H. (1985). Central limit theorem for a system of markovian paricles with mean field interaction. *Z. Wahrsch. Verw. Gebiete* **69** 439–459.

[34] SHIRYAEV A. N. (1996). *Probability*, 2nd ed. Springer, New York.

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

[36] VAN DOOTINGH, M., VIEL, F., RAKOTOPARA, D. and GAUTHIER, J. P. (1991). Coupling of nonlinear 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** 453–472.
[39] WILLIAMS, D. (1992). *Probability with Martingales*. Cambridge Univ. Press.

LSP-CNRS UMR C55830, BAT.1R1
UNIVERSITÉ PAUL SABATIER
31062 TOULOUSE
FRANCE
E-MAIL: delmoral@cict.fr

LMSS-CNRS URA 743, BAT. 425
UNIVERSITÉ PARIS SUD,
91405 ORSAY
FRANCE
E-MAIL: guionnet@stats.matups.fr