Discrete Filtering Using Branching and Interacting Particle Systems*

D. Crisan [†], P. Del Moral [‡], T. Lyons [§]

23/02/98

Abstract

The stochastic filtering problem deals with the estimation of the current state of a signal process given the information supplied by an associate process, usually called the observation process. We describe a particle algorithm designed for solving numerically discrete filtering problems. The algorithm involves the use of a system of n particles which evolve (mutate) in correlation with each other (interact) according to law of the signal process and, at fixed times, give birth to a number of offsprings depending on the observation process. We present several possible branching mechanisms and prove, in a general context the convergence of the particle systems (as n tends to ∞) to the conditional distribution of the signal given the observation. We then apply the result to the discrete filtering and give several example when the results can be applied.

AMS Subject Classification (1991): 93E11, 60G57, 65U05

Keywords: Filtering, Particle Systems, Branching Algorithms, Interacting Algorithms, Measure Valued Processes, Numerical Solutions

*Supported by the EU TMR Programme Project ERBF MRX CT 960075A

[†] Dept. of Mathematics, Imperial College, Huxley Building, 180 Queen's Gate, London SW7 2BZ, d.crisan@ic.ac.uk

[‡] UMR C55830, CNRS, Bat.1R1, Univ. Paul Sabatier, 31062 Toulouse, France, delmoral@cict.fr

[§] Dept. of Mathematics, Imperial College, Huxley Building, 180 Queen's Gate, London SW7 2BZ, UK, t.lyons@ic.ac.uk

1 Introduction

The stochastic filtering problem consists in effectively estimating the conditional distribution of a process (the signal) given the "noisy" information obtained from a related process (the observation). The basic problem can be identified applications: signal processing, radar control, satellite tracking, weather forecasting, speech recognition are just a few of them (see, for instance, [21] and the reference therein). There are very few cases when the problem admits a solution in closed form and therefore, efficient numerical approximations to the conditional distribution are of great interest.

Several recently suggested approaches are based on the simulation of interacting and branching particle systems. If suitable conditions are imposed on the signal semigroup, the empirical measures of the particle systems can be shown to converge to the solution of the measure valued dynamical system the evolution of the conditional distributions. In this paper we design a particle system approach which allows us to combine the branching and interacting mechanisms introduced in [4, 5, 6], [11, 12, 13, 14, 15] and in [17]. Our particle approximations also cope with discrete time filtering problems in which the signal is a non linear process with transitions that depend on all the data observed in the past. The discrete time and measure valued processes under study also arise in Statistical Physics. In the following we present a brief formulation of the model.

We assume we can model the state space E as a locally compact metric space with the associated Borel σ-field $\mathbf{B}(E)$. We denote by $\mathbf{M}(E)$ the space of all finite non negative Borel measures on E and by $\mathbf{M}_1(E) \subset \mathbf{M}(E)$ the set of probability measures. $\mathbf{M}(E)$ and $\mathbf{M}_1(E)$ are furnished with the weak topology. We recall that the weak topology is metrisable and, under this topology, $\lim_{n\to\infty} \mu_n = \mu$ iff

$$\lim_{n\to\infty} \int_E f(x)\mu_n(dx) = \int_E f(x)\mu(dx), \forall f \in \mathcal{C}_b(E),$$

where $\mathcal{C}_b(E)$ is the space of bounded continuous functions functions on which we consider the norm

$$\|f\| = \sup_{x\in E} |f(x)|.$$

We consider now a set of transitions $\{K_{n,\mu} : \mathbf{M}(E) \to \mathbf{M}(E) \ ; \ n \geq 1, \mu \in \mathbf{M}(E)\}$ which will represent the transitions of the (non linear) signal process. We denote by μK the measure given by $\mu K(A) = \int_E \mu(dx)\, K(x, A)$ where K is any transition on E, $\mu \in \mathbf{M}(E)$ and $A \in \mathbf{B}(E)$, hence

$$\forall f \in \mathcal{C}_b(E) \qquad \mu K(f) = \int \mu(dx)\, K(x, dz)\, f(z). \tag{1}$$

Using (1), the so-called nonlinear filtering equations which represents the dynamics structure of the conditional distributions are decomposed into two separate mechanisms

$$\begin{cases} \widehat{\eta}_n &= \Psi_n(\eta_n) \qquad n \geq 0 \qquad \eta_0 \in \mathbf{M}_1(E) \\ \eta_{n+1} &= \widehat{\eta}_n \, K_{n+1,\widehat{\eta}_n} \end{cases} \tag{2}$$

where $\Psi_n : \mathbf{M}_1(E) \to \mathbf{M}_1(E)$, $n \geq 0$ are applications given by

$$\forall f \in \mathcal{C}_b(E) \qquad \Psi_n(\eta)f = \frac{\eta(g_n f)}{\eta(g_n)}$$

and $g_n : E \to I\!R_+^*$, $n \geq 0$ are bounded positive functions. In nonlinear filtering settings the first mechanism

$$\eta_n \longrightarrow \Psi_n(\eta_n)$$

updates the distribution η_n given the current observation at time n and the second one

$$\widehat{\eta}_n \longrightarrow \widehat{\eta}_n \, K_{n+1,\widehat{\eta}_n}$$

is called the prediction and does not depend on the observation data at time $n+1$, but may depends on all the data observed up to time n.
When the functions $g_n, n \geq 1$ are constant, i.e., $g_n(x) = 1$ for all $x \in E$ the dynamical system (2) describes the time evolution of the density profiles of McKean-Vlasov stochastic processes with mean field drift functions. Such equations also occur in Statistical Physics (see [7],[29] and references therein) and it was proposed by McKean and Vlasov to approximate the corresponding equations by mean field interacting particle systems. A crucial practical advantage of this situation is that the dynamical structure of the non linear stochastic process can be used in the design of an interacting particle system in which the mean field drift is replaced by a natural interaction function. Such models are called in Physics Masters equations and/or weakly interacting particle systems. They are now well understood (see [1], [8], [7],[18], [29], [30] and references therein). Under rather general assumptions, it was shown that the particle density profile (that is the random empirical measures of the particle systems) converges towards the solution of (2) as the number of particles is going to infinity. As a consequence, propagation of chaos occurs.
In contrast to the situation described above the conditional distributions cannot be viewed as the law of a finite dimensional stochastic process which incorporates a mean field drift [3]. We therefore have to find a new strategy to define an interacting particle system which will approximate the desired distributions.
The paper has the following structure:
In section 2 we introduce a **branching and interacting particle system** (BIPS) model and we study the connections between several particular classes of particle approximations. The study the convergence of the empirical measure of the system when the initial number of particles tends to ∞ is performed in section 3. The application of the particle approximations described in section 2 to non linear filtering problems

is explored in section 4. We end this paper with some applications of the former BIPS approximations to some practical problems arising in advanced non linear signal processing.

2 Branching and Interacting particle Systems

The BIPS under study will be a two step Markov chain

$$(N_n, \xi_n) \stackrel{Branching}{\longrightarrow} (\widehat{N}_n, \widehat{\xi}_n) \stackrel{Mutation}{\longrightarrow} (N_{n+1}, \xi_{n+1}) \tag{3}$$

with product state space $\mathcal{E} = \bigcup_{\alpha \in I\!N} (\{\alpha\} \times E^\alpha)$ with the convention $E^\alpha = \emptyset$ if $\alpha = 0$. We will note $\mathcal{F} = \{F_n, \widehat{F}_n \ : \ n \geq 0\}$ the canonical filtration associated to (3) so that

$$F_n \ \subset \ \widehat{F}_n \ \subset F_{n+1}$$

The points of the set E^α, $\alpha \geq 0$ are called particle systems and are mostly denoted by the letters x and z. The parameter $\alpha \in I\!N$ represents the size of the system. The initial number of particle $N_0 \in I\!N$ is a fixed non random number which represents the precision parameter of the BIPS algorithm.

2.1 Description of the model

The evolution in time of the BIPS is defined inductively as follows.

- At the time $n = 0$:
  The initial particle system $\xi_0 = (\xi_0^1, \ldots, \xi_0^{N_0})$ consists of N_0 independent and identically distributed particles with common law η_0.
- Evolution in time:
  At the time n, the particle system ξ_n consists of N_n particles.
  If $N_n = 0$ the particle system died and we let $\widehat{N}_n = 0$ and $N_{n+1} = 0$.
  Otherwise the branching correction is defined as follows

  1. **Branching Correction:**
     When $N_n > 0$ we associate to $\xi_n = (\xi_n^1, \ldots, \xi_n^{N_n}) \in E^{N_n}$ the weight vector $W_n = (W_n^1, \ldots, W_n^{N_n}) \in I\!R^{N_n}$ given by

     $$\sum_{i=1}^{N_n} W_n^i \, \delta_{\xi_n^i} = \Psi_n(m(\xi_n)) \qquad \text{where} \quad m(\xi_n) = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{\xi_n^i}$$

     Then, each particle ξ_n^i, $1 \leq i \leq N_n$, branches into a random number of offsprings M_n^i, $1 \leq i \leq N_n$ and the mechanism is chosen so that

     $$E(M_n | F_n) = N_n W_n \tag{4}$$

and there exists a finite constant $C < \infty$ so that for any $f \in \mathcal{C}_b(E)$

$$E\left(|\sum_{i=1}^{N_n} M_n^i f(\xi_n^i) - N_n \Psi_n(m(\xi_n)) f|^2 \, |F_n\right) \leq C \, N_n \|f\|^2 \qquad (5)$$

At the end of this stage the particle system $\widehat{\xi}_n$ consists of

$$\widehat{N}_n = \sum_{i=1}^{N_n} M_n^i$$

particles denoted by

$$\widehat{\xi}_n^i = \xi_n^k \qquad 1 \leq k \leq N_n \qquad \sum_{l=1}^{k-1} M_n^l + 1 \leq i \leq \sum_{l=1}^{k-1} M_n^l + M_n^k \qquad (6)$$

2. **Mutation transition:**

   If $\widehat{N}_n = 0$ the particle system dies and $N_{n+1} = 0$.
   Otherwise, each particle moves independently of each other starting off from the parent particle branching site ξ_n^i, $1 \leq i \leq N_n$, with law

$$K_{n+1, m(\widehat{\xi}_n)}(\xi_n^i, dx) \qquad 1 \leq i \leq N_n$$

   where $m(\widehat{\xi}_n)$ is the empirical measure associated to $\widehat{\xi}_n$.
   During this transition the total number of particle doesn't change ($N_{n+1} = \widehat{N}_n$) and the mechanism can be summarized as follows, for any $\alpha \geq 0$ and $z \in E^\alpha$

$$P\left(\xi_{n+1} \in dx | \widehat{\xi}_n = z, \widehat{N}_n = \alpha\right) = \prod_{i=1}^{\alpha} K_{n+1, m(z)}(z^i, dx)$$

   where $dx = dx^1 \times \ldots \times dx^\alpha$ is an infinitesimal neighbourhood of $x \in E^\alpha$ with the conventions $dx = \emptyset$ and $\prod_{i=1}^{\alpha} = 1$ if $\alpha = 0$.

Note that the mutation of each individual offspring $\widehat{\xi}_n^i$, $1 \leq i \leq \widehat{N}_n$ depends on the entire configuration $\widehat{\xi}_n$ of the system. In other words between branching corrections the particle system behaves itself as a interacting particle system.

The above BIPS model enables a unified description of the various particle system approximations presented in [4, 5, 6], [11, 12, 13, 14] and in [17]. We have deliberately left open the question of the choice of the branching correction transition and we will devote a subsection to present several natural choices which can be used in practice. Before that, let us point out some important properties of the BIPS algorithm.

Proposition 2.1 *The process $N = (N_n)_{n\geq 0}$ is a positive integer valued martingale with respect to the filtration $F = (F_n)_{n\geq 0}$ with the following properties*

$$E\left(\left(\frac{N_n}{N_0}-1\right)^2\right) \leq \frac{C\ n}{N_0} \qquad \text{and} \qquad P\left(N_n = 0\right) \leq \frac{C\ n}{N_0} \tag{7}$$

Proof:
¿From the construction of the branching corrections we have

$$\forall n \geq 0 \qquad E\left(N_n | F_{n-1}\right) = N_{n-1}\ I(N_{n-1} > 0) = N_{n-1}$$

It follows that N is an F-martingale. Similarly, (5) implies that

$$E((N_n - N_{n-1})^2 | F_{n-1}) \leq C N_n$$

It is then easy to show that

$$\forall n \geq 0 \qquad E(N_n^2) \leq N_0^2 + C\ n\ N_0$$

or, what amounts to the same thing

$$E\left(\left(\frac{N_n}{N_0}-1\right)^2\right) \leq \frac{C\ n}{N_0}$$

The last assertion is a consequence of this inequality. To be more precise, for any $\epsilon \in]0,1[$ we have

$$P(N_n > 0) \geq P(N_n \geq (1-\epsilon)N_0) \geq P(|N_n - N_0| \leq \epsilon N_0)$$

Then, using Tchebitchev's inequality we find that for any $\epsilon \in]0,1[$

$$P(N_n = 0) \leq \frac{C\ n}{N_0 \epsilon^2}$$

Letting $\epsilon \to 1$ one obtain the desired inequality. ∎

Remark 2.2:
Using Doob's maximal inequality, from Proposition 2.1 we get that

$$E\left(\sup_{k=1,\ldots,n}\left(\frac{N_k}{N_0}-1\right)^2\right) \leq \frac{C\ n}{N_0}\|f\|^2.$$

Remark 2.3:
The last inequality in (7) yields

$$P(N_k > 0, \quad \forall k \in [0,n]) \geq 1 - \frac{Cn}{N_0}.$$

2.2 Branching Corrections

The purpose of this subsection is to present several examples of branching corrections satisfying conditions (4) and (5). We will distinguish two types of branching numbers laws. In the first situation the branching numbers $M = (M_n)_{n \geq 0}$ are chosen independently each other. In contrast to this we present an example of negatively correlated branching numbers.

2.2.1 Independent Branching Numbers

When the branching numbers M_n are independent conditionnally on F_n condition (5) is equivalent to

$$E\left(\sum_{i=1}^{N_n}\left(M_n^i - N_n W_n^i\right)^2 f(\xi_n^i)^2 \,|F_n\right) \leq C N_n$$

where the weight vector $W_n = (W_n^1, \ldots, W_n^{N_n})$ is given by

$$W_n^i = \frac{g_n(\xi_n^i)}{\sum_{j=1}^{N_n} g_n(\xi_n^j)} \qquad 1 \leq i \leq N_n$$

In this situation it is natural to use a branching number law so that

$$E(M_n|F_n) = N_n W_n \qquad V(M_n|F_n) \leq C\, N_n$$

where $V(M_n/F_n)$ denotes the conditional variance of the vector M_n with respect to F_n. Let us now present some classsical examples of independent branching numbers.

Bernouilli branching numbers:

The Bernouilli branching numbers were introduced by two of the authors in [6]. They are defined as a sequence $M_n = (M_n^i,\ 1 \leq i \leq N_n)$ of conditionally independent random numbers with respect to F_n with distribution given for any $1 \leq i \leq N_n$ by

$$P(M_n^i = k|F_n) = \begin{cases} \epsilon(N_n W_n^i) & \text{if} \quad k = [N_n W_n^i] + 1 \\ 1 - \epsilon(N_n W_n^i) & \text{if} \quad k = [N_n W_n^i] \end{cases}$$

where $[a]$ (resp. $\epsilon(a) = a - [a]$) denotes the integer part (resp. the fractional part) of $a \in I\!R$. The required conditions (4) and (5) are derived easily from the fact that

$$\begin{array}{rcl} E(M_n^i|F_n) & = & N_n W_n^i \\ V(M_n^i|F_n) & = & \epsilon(N_n W_n^i)(1 - \epsilon(N_n W_n^i)) \in [0, 1/4] \end{array} \tag{8}$$

for any $1 \leq i \leq N_n$. Using the above formula we see that (5) is satisfied with $C = 1/4$. In addition it can be seen from the relation $\sum_{i=1}^{N_n}(N_n W_n^i) = N_n$ that at least one particle has one offspring (cf. [4] for more details). Therefore using the above branching correction the particle system never dies.

Poisson branching numbers:

The Poisson branching numbers are defined as a sequence $M_n = (M_n^i,\ 1 \leq i \leq N_n)$ of conditionally independent random numbers with respect to F_n with distribution given for any $1 \leq i \leq N_n$ by

$$\forall k \geq 0 \qquad P(M_n^i = k | F_n) = \exp\left(-N_n W_n^i\right)\ \frac{(N_n W_n^i)^k}{k!}$$

In this situation, we have

$$E(M_n^i | F_n) = N_n W_n^i = V(M_n^i | F_n)$$

so that (5) holds with $C = 1$.

Binomial branching numbers:

The binomial branching numbers are defined as a sequence $M_n = (M_n^i,\ 1 \leq i \leq N_n)$ of conditionally independent random numbers with respect to F_n with distribution given for any $1 \leq i \leq N_n$ by

$$\forall 0 \leq k \leq N_n \qquad P(M_n^i = k | F_n) = C_{N_n}^k\ (W_n^i)^k\ (1 - W_n^i)^{N_n - k}$$

In this case (4) and (5) follows from the fact that for any $1 \leq i \leq N_n$

$$\begin{aligned} E(M_n^i | F_n) &= N_n W_n^i \\ V(M_n^i | F_n) &= N_n W_n^i (1 - W_n^i) \end{aligned}$$

for any $1 \leq i \leq N_n$. Moreover, using the above we see that (5) is satisfied with

2.2.2 Branching numbers with negative correlations

We continue the account of the standard branching laws which can be used in the correction step of the algorithm. We have presented sofar some classes of independent branching numbers which give a good rational representation of the current weights. However, for these branching corrections, the total number of particles is not fixed but random. If

$$M_n = \text{Multinomial}\left(N_n, W_n^1, \ldots, W_n^{N_n}\right) \tag{9}$$

then the population size is preserved. In this case we have for any $1 \leq i \neq j \leq N_n$

$$\begin{aligned} E(M_n^i | F_n) &= N_n W_n^i \\ E((M_n^i - N_n W_n^i)^2 | F_n) &= N_n W_n^i (1 - W_n^i) \\ E((M_n^i - N_n W_n^i)(M_n^j - N_n W_n^j) | F_n) &= -N_n W_n^i W_n^j \end{aligned}$$

Using the above we find that for any $f \in \mathcal{C}_b(E)$,

$$E\left(\left(\sum_{i=1}^{N_n} M_n^i f(\xi_n^i) - N_n \Psi_n(m(\xi_n))f\right)^2 |F_n\right) \leq N_n\ \Psi_n(m(\xi_n))\ (f - \Psi_n(m(\xi_n))f)^2$$

Therefore we see that (4) and (5) are satisfied with $C = 1$. When the transition probability kernels $K_{n,\mu}$ satisfy the assumption

$$\forall (x,z) \in E^2 \qquad K_{n,\mu}(x, \{z\}) = 0 \tag{10}$$

it does follows that for any $1 \leq i \neq j \leq N_n$ $\xi_n^i \neq \xi_n^j$ P-a.s.. In this case the weights M_n may be written as

$$\forall 1 \leq i \leq N_n \qquad M_n^i = \text{Card}\{1 \leq j \leq N_n\ :\ \widehat{\xi}_n^j = \xi_n^i\}$$

where $\widehat{\xi}_n = (\widehat{\xi}_n^1, \ldots, \widehat{\xi}_n^{N_n})$ are conditionally independent random variables with respect to F_n with common law $\Psi_n(m(\xi_n))$. This model of branching numbers was introduced by one of the authors in [11].

Let us look at the special case of a BIPS with multinomial branching corrections. In view of the preceding considerations the size of the systems $(\xi_n, \widehat{\xi}_n\ :\ n \geq 0)$, $n \geq 0$, does not change and is equals to N_0. In addition when (10) holds the dynamics structure of the latter can be written in the simplest form:

- **Initial Particle System**

$$P(\xi_0 \in dx) = \prod_{p=1}^{N_0} \eta_0(dx^p)$$

- **Branching Correction**

$$P(\widehat{\xi}_n \in dx | \xi_n = z) = \prod_{p=1}^{N_0} \Psi_n(\frac{1}{N_0}\sum_{i=1}^{N_0} \delta_{z^i})(dx^p)$$

- **Mutation Transition**

$$P(\xi_{n+1} \in dz | \widehat{\xi}_n = x) = \prod_{p=1}^{N_0} K_{n+1, \frac{1}{N_0}\sum_{i=1}^{N_0} \delta_{z^i}}(x^p, dz^p)$$

We note that $(\xi_n\ :\ n \geq 0)$ is a E^{N_0}-valued Markov chain given by

$$P(\xi_n \in dx | \xi_{n-1} = z) = \prod_{p=1}^{N_0} \Phi(n, \frac{1}{N_0}\sum_{i=1}^{N_0} \delta_{z^i})(dx^p) \tag{11}$$

where

$$\Phi(n, \eta) = \Psi_{n-1}(\eta) K_{n, \Psi_{n-1}(\eta)} \qquad \forall \eta \in \mathbf{M}_1(E)$$

These constructions first appear in [11] and were developed in full details in [12]. Large deviations principles for interacting particle systems of the form (11) and their applications to non linear filtering problems are described in [14].

2.3 Structural Properties

One natural question we address now is the difference between BIPS with conditionally independent branching numbers and BIPS with multinomial branching numbers. Before getting down into the details it may be helpful to make a couple of remarks. In the first place it should be note that the BIPS described above using Multinomial branching and Poisson branching are related to the continuous critical branching superprocesses and the Fleming-Viot processes (see Dawson [9] and references therein) and the same kind of relations exist between the BIPS presented above.

Total mass process

As their continuous time version, the main difference between the two types of BIPS presented above is that the population size of the BIPS with Multinomial branching is constant but the population size of the BIPS with Bernoulli, Poisson or Binomial branching is a martingale with quadratic characteristic

$$< N >_n = \sum_{k=0}^{n-1} V(M_k|F_k) \leq \sum_{k=0}^{n-1} N_k$$

Conditional BIPS

As in continuous time settings the BIPS with multinomial branching arises by conditioning a BIPS with Poisson branching to have constant population size.
To make all this more precise we introduce some additional notations. For any $N_0 \geq 1$ we denote

$$\left(\Omega, (F_n, \widehat{F}_n)_{n\geq 0}, (N_n, \xi_n, \widehat{N}_n, \widehat{\xi}_n)_{n\geq 0}, P^{\mathrm{MB}}_{\mathrm{N}_0}\right)$$

the discrete time Markov model which realizes the BIPS with multinomial branchings and starts with N_0 particles. By construction of the multinomial corrections we have

$$\forall n \geq 0 \qquad N_n = N_0 \qquad P^{\mathrm{MB}}_{N_0} - \text{a.s.}$$

On the other hand we denote

$$\left(\Omega, (F_n, \widehat{F}_n)_{n\geq 0}, (N_n, \xi_n, \widehat{N}_n, \widehat{\xi}_n)_{n\geq 0}, P^{\mathrm{PB}}_{\mathrm{N}_0}\right)$$

the discrete time Markov model which realizes the BIPS with Poisson branchings and starts with N_0 particles.

Proposition 2.4 *For any $A \in \vee_n (F_n \vee \widehat{F}_n)$ we have*

$$P^{\mathrm{PB}}_{N_0}\left(A|N = N_0\right) = P^{\mathrm{MB}}_{N_0}\left(A\right) \qquad P^{\mathrm{PB}}_{N_0} - a.s. \tag{12}$$

Proof:
First we note that conditionally on the event

$$\{N = N_0\} = \bigcap_{n\geq 0} \{N_n = N_0\}$$

we have

$$\forall n \geq 0 \qquad (\xi_n, \widehat{\xi}_n) \in E^{N_0} \times E^{N_0} \qquad P^{\text{PB}}_{\text{N}_0} - \text{a.s.}$$

On the other hand, by construction of the mutation transition we have for any $n \geq 0$, $x, z \in E^{N_0}$

$$P^{\text{PB}}_{\text{N}_0}\left(\xi_n \in dz | N = N_0, \widehat{\xi}_{n-1} = x\right) = P^{\text{MB}}_{\text{N}_0}\left(\xi_n \in dz | \widehat{\xi}_{n-1} = x\right)$$

Using the above observations and the fact that changes in the number of particles only take place at branching corrections, we see that to prove (12) it suffices to check that for any $n \geq 0$, $x, z \in E^{N_0}$

$$P^{\text{PB}}_{\text{N}_0}\left(\widehat{\xi}_n \in dz | N = N_0, \xi_n = x\right) = P^{\text{MB}}_{\text{N}_0}\left(\widehat{\xi}_n \in dz | \xi_n = x\right)$$

Now, by definition of the Poisson branching transitions, for each $n \geq 0$, $x \in E^{N_0}$ and $k \in I\!N^{N_0}$

$$\begin{aligned} P^{\text{PB}}_{\text{N}_0}\left(M_n = k | N = N_0, \xi_n = x\right) &= P^{\text{PB}}_{\text{N}_0}\left(M_n = k | N_n = N_0, \xi_n = x\right) \\ &= \frac{1}{Z(n, N_0)} \prod_{i=1}^{N_0} \exp\left(-N_0 W_n^i\right) \; \frac{(N_0 W_n^i)^{k_i}}{k_i!} \end{aligned}$$

with

$$W_n^i = \frac{g_n(x^i)}{\sum_{j=1}^{N_0} g_n(x^j)} \qquad \forall 1 \leq i \leq N_0$$

and

$$Z(n, N_0) = \sum_{k_1+\ldots+k_{N_0}=N_0} \prod_{i=1}^{N_0} \exp\left(-N_0 W_n^i\right) \; \frac{(N_0 W_n^i)^{k_i}}{k_i!}$$

It is not difficult to see that $Z(n, N_0) = e^{-N_0} \; N_0! \; N_0^{N_0}$ so that

$$P^{\text{PB}}_{\text{N}_0}\left(M_n = k | N = N_0, \xi_n = x\right) = \text{Multinomial}(N_0, W_n^1, \ldots, W_n^{N_0})$$

Or, what amounts to the same thing

$$P^{\text{PB}}_{\text{N}_0}\left(M_n = k | N = N_0, \xi_n = x\right) = P^{\text{MB}}_{\text{N}_0}\left(M_n = k | \xi_n = x\right)$$

This means that the conditionally on the event $\{N = N_0\}$ the Poisson branching corrections become multinomial corrections. This end the proof of the proposition ■

The continuous time version of this result was discovered by Etheridge and March in [16] in their study of the connections between critical branching superprocesses and the Fleming-Viot interacting particle systems.
Before moving on let us remark that each time the multinomial branching numbers described in section 2.2.1 are defined using the population size of the system at last time. Furthermore this transition keeps unchanged the total size of the system.
When using multinomial branching laws one still have the freedom to adapt the size parameter so that to produce a given number of offsprings.
To this end, let $(a_n;\, n \geq 0)$ be the path numbers of offsprings we want to have at each stage of the algorithm (i.e. $N_0 = a_0, N_1 = a_1, \ldots, N_n = a_n, \ldots$).
To do this the corresponding branching laws are defined in replacing at each time n the law (9) by the multinomial distribution

$$M_n = \text{Multinomial}\left(a_{n+1}, W_n^1, \ldots, W_n^{a_n}\right) \tag{13}$$

Let us denote by

$$\left(\Omega, (F_n, \widehat{F}_n)_{n\geq 0}, (N_n, \xi_n, \widehat{N}_n, \widehat{\xi}_n)_{n\geq 0}, P_{\mathrm{N}_0}^{\mathrm{PB}(a)}\right)$$

the discrete time Markov model which realizes the BIPS with the multinomial branchings corrections (13) and starts with N_0 particles. Using the same line of arguments as before one gets

Proposition 2.5 *For any $A \in \vee_n(F_n \vee \widehat{F}_n)$ we have*

$$P_{\mathrm{N}_0}^{\mathrm{PB}}\left(A|N = a\right) = P_{\mathrm{N}_0}^{\mathrm{MB(a)}}\left(A\right) \qquad P_{\mathrm{N}_0}^{\mathrm{PB}} - a.s.$$

The continuous time version of this result was proved by Perkins [25] in his precise study of the structural properties of Dawson-Watanabe and Fleming-Viot processes.

2.4 Complexity and Efficiency

The multinomial branching numbers ensure that the population size is constant and prevent extinction or explosion of the algorithm. The price to pay is that the multinomial branching correction is time consuming:
For instance the sampling of a measure concentrated at N_0 points using the standard inversion formula is performed by an algorithm which uses N_0 tests. It follows that N_0 independent sampling of the same measure will use N_0^2 test operations. In contrast to this the Bernouilli distribution is extremely time saving. Using the same inversion formula we just have to use one test operation. So that N_0 independent sampling of a Bernouilli distribution will use no more than N_0 test operations.
In terms of complexity it seems then logical to choose Bernouilli branching instead of multinomial. Nevertheless the precision of the particle approximation can be altered by

the random fluctuations of the population size associated to a given choice of branching law.

The BIPS aproach introduced at the beginning of this section leads to a variety of random particle algorithms. In section 2.2 we proposed several examples of branching corrections that can be used in practice. There are actually no techniques for determining the "optimal choice" of branching correction. Nevertheless if we want to estimate the integrals $\widehat{\eta}_n f$ at each time $n \geq 0$, for some test function $f \in \mathcal{C}_b(E)$, the key condition (5) page 5 is pivotal. Roughtly speaking it ensures that the dynamics structure of the BIPS during each branching corrections is not far from the updating transition of the original system (2). Before discussing the meaning of conditions (4) and (5) let us recall the conditional expectation of a $\mathbf{M}(E)$-valued random measure relative to a σ-field (cf. H.Kunita [20]). Let $\mu(\omega)$ be an $\mathbf{M}(E)$-valued random variable defined on a probability space $(\Omega, \mathcal{F}, P)$. The conditional expectation of μ relative to a sub-σ-field $G \subset F$ is defined as a $\mathbf{M}(E)$-valued random variable $E(\mu/G)$ such that

$$F\left(E(\mu/G)\right) = E(F(\mu)/G)$$

holds for all continuous affine functions $F : \mathbf{M}(E) \longrightarrow \mathbb{R}$ ($F \in \mathcal{C}_b(\mathbf{M}(E))$ is affine if there exists a real constant c and a function $f \in \mathcal{C}_b(E)$ such that for every $\nu \in \mathbf{M}(E)$ $F(\nu) = c + \nu(f)$).
The first condition (4) ensures that the estimator is unbiased, i.e.,

$$E(\sum_{i=1}^{N_n} M_n^i \; \delta_{\xi_n^i} | F_n) = N_n \; \Psi_n(m(\xi_n))$$

This property of being unbiased is natural and expresses the fact that the averaged estimate coincide with the desired result. The second condition (5) is related to the deviation of the estimate. Thus, it is, again, natural to suppose that an unbiased branching correction is better, if its deviation is smaller.
Let us examine the deviation

$$E\left(|\sum_{i=1}^{N_n} M_n^i f(\xi_n^i) - N_n \Psi_n(m(\xi_n)) f|^2 \, |F_n \right) \tag{14}$$

associated to the branching number laws presented in section 2.2.

$$\sum_{i=1}^{N_n} V(M_n^i | F_n) \; f(\xi_n^i)^2 \qquad \text{(Cond. Independent branching numbers)}$$

$$N_n \; \Psi_n(m(\xi_n))(f - \Psi_n(m(\xi_n))f)^2 \qquad \text{(Multinomial branching numbers)}$$

Note that the last deviation can be written

$$\frac{N_n}{2} \sum_{i \neq j} W_n^i \; W_n^j \; (f(\xi_n^i) - f(\xi_n^j))^2$$

We quote a result of two of the authors [4, 5, 6] who showed that the Bernouilli branching numbers realize the minimum deviation (14) with respect to all conditionally independent branching numbers.
The "optimal choice" of the branching correction at each time $n \geq 0$ strongly depends on the function $f \in \mathcal{C}_b(E)$ and on the weight vector $W_n = (W_n^1, \ldots, W_n^{N_n})$:
Suppose we want to estimate the integrals $\Psi_n(\eta_n)f$, $n \geq 0$, where f is constant function on some Borel subset $B \subset E$.
For any weight vector W_n, the deviation (14) corresponding to a multinomial branching correction is null as soon as the particles ξ_n^i belong to B for any $1 \leq i \leq N_n$.
On the other hand, if the weight vector W_n is given by

$$W_n = \frac{p_n}{N_n}$$

with

$$p_n = (p_n^1, \ldots, p_n^{N_n}) \in I\!N^{N_n} \qquad \text{and} \qquad \sum_{i=1}^{N_n} p_n^i = N_n$$

the multinomial branching correction is null for any test function f.

3 Convergence Theorems

In the previous section we introduced a general model of particle systems which move, die and produce offsprings in accordance with the dynamics structure of the measure valued process (2). In this section we prove that the so-called particle density profile, i.e., the random empirical measure of the system, converges weakly to the solution of (2) as the initial number of particle tends to infinity.
The main difference from the interacting particle models developped in [11] and [12] is that the total number of particles is not constant but forms an integer valued martingale. For this reason the basic state space for the study of the convergence is now $\mathbf{M}(E)$ instead of $\mathbf{M}_1(E)$.
Let us introduce the random measures which will be used in the sequel. We denote by $(\eta_n^N, \widehat{\eta}_n^N \; ; \; n \geq 0)$ the random measures given by

$$\eta_n^N = \frac{1}{N_0} \sum_{i=1}^{N_n} \delta_{\xi_n^i} \qquad \text{and} \qquad \widehat{\eta}_n^N = \frac{1}{N_0} \sum_{i=1}^{\widehat{N}_n} \delta_{\widehat{\xi}_n^i} \tag{15}$$

with the convention $\sum_{\emptyset} = 0$ the null measure on E.
It is also convenient to introduce the normalised measures, $(m(\xi_n), m(\widehat{\xi}_n) \; ; \; n \geq 0)$

$$\forall f \in \mathcal{C}_b(E) \qquad m(\xi_n)f = \frac{\eta_n^N(f)}{\eta_n^N(1)} \qquad \text{and} \qquad m(\widehat{\xi}_n)f = \frac{\widehat{\eta}_n^N(f)}{\widehat{\eta}_n^N(1)} \tag{16}$$

with the convention $m(\xi_n)f = 0$ (resp. $m(\widehat{\xi}_n)f = 0$) if η_n^N (resp. $\widehat{\eta}_n^N$) is the null measure on E.
When the particle systems ξ_n (resp. $\widehat{\xi}_n$) is not dead $m(\xi_n)$ (resp. $m(\widehat{\xi}_n)$) is the empirical measure associated to ξ_n (resp. $\widehat{\xi}_n$).

In a first subsection we summarize the key concepts and the technical tools necessary to carry out the proof of limit theorems. For further information the reader is referred to Parthasarathy [24] and or Billingsley [2]

3.1 Measure Valued Random Variables

Recall that $\mathbf{M}(E)$ with the topology of weak convergence is a complete separable metrisable space with metric ρ defined as follows

$$\rho(\mu,\nu) = \sum_{m\geq 0} 2^{-m}\ |\mu f_m - \nu f_m| \wedge 1 \quad \forall \mu,\nu \in \mathbf{M}(E) \tag{17}$$

where $(f_m)_{m\geq 0}$ is a suitable sequence of uniformly continuous functions such that $\|f_m\| \leq 1$ for all $m \geq 1$ and $f_0 \equiv 1$. In this paper we study sequences of $\mathbf{M}(E)$-valued random variables. The basic state space for the study of the weak convergence is the set of all probability measures on $\mathbf{M}(E)$ denoted by $\mathbf{M}_1(\mathbf{M}(E))$. By $\mathcal{C}_b(\mathbf{M}(E))$ we denote the space of all bounded continuous functions on $\mathbf{M}(E)$ furnished with the uniform norm

$$\|F\| = \sup_{\mu\in\mathbf{M}(E)} |F(\mu)|$$

For an $F \in \mathcal{C}_b(\mathbf{M}(E))$ and $\Phi \in \mathbf{M}_1(\mathbf{M}(E))$ we write

$$\Phi F = \int F(\mu)\ \Phi(d\mu)$$

We say that a sequence $(\Phi_{N_0})_{N_0\geq 0}$, $\Phi_{N_0} \in \mathbf{M}_1(\mathbf{M}(E))$, converges weakly to a measure $\Phi \in \mathbf{M}_1(\mathbf{M}(E))$ if

$$\forall F \in \mathcal{C}_b(\mathbf{M}(E)) \qquad \lim_{N_0\to+\infty} \Phi_{N_0} F = \Phi F$$

Now we introduce the Kantorovitch-Rubinstein or Vasershtein metric on the set $\mathbf{M}_1(\mathbf{M}(E))$ defined by

$$\begin{aligned} D(\Phi,\Psi) &= \inf\Big\{\int \rho(\mu,\nu)\,\Theta(d(\mu,\nu))\ :\ \Theta \in \mathbf{M}_1(\mathbf{M}(E)\times\mathbf{M}(E)) \\ &\qquad p_1\circ\Theta = \Phi \text{ and } p_2\circ\Theta = \Psi\} \end{aligned} \tag{18}$$

(see for instance [30] and references therein). The metric D gives the topology of weak convergence on $\mathbf{M}_1(\mathbf{M}(E))$.

Let $(\mu,\mu_{N_0})_{N_0\geq 1}$ be a sequence of measure valued random variables on some probability space such that μ_{N_0} have distributions $\Phi_{N_0} \in \mathbf{M}_1(\mathbf{M}(E))$, $N_0 \geq 1$ and μ is a measure valued random variable with distribution $\Phi \in \mathbf{M}_1(\mathbf{M}(E))$. We can apply the monotone convergence theorem to prove that

$$D(\Phi_{N_0}, \Phi) \leq \sum_{m \geq 0} 2^{-m} \; E[|\mu_{N_0} f_m - \mu f_m| \wedge 1]$$

so by the dominated convergence theorem

$$\forall f \in \mathcal{C}_b(E) \quad \lim_{N_0 \to +\infty} E[|\mu_{N_0} f - \mu f|] = 0 \Longrightarrow \lim_{N_0 \to +\infty} D(\Phi_{N_0}, \Phi) = 0 \tag{19}$$

In addition, if μ is a fixed probability distribution the functions

$$F_\mu(\nu) = |\nu f - \mu f| \wedge 1, \qquad f \in \mathcal{C}_b(E)$$

are continuous for the weak convergence topology in $\mathbf{M}(E)$ and therefore

$$\forall f \in \mathcal{C}_b(E) \quad \lim_{N_0 \to +\infty} E[|\mu_{N_0} f - \mu f| \wedge 1] = 0 \Longleftrightarrow \lim_{N_0 \to +\infty} D(\Phi_{N_0}, \Phi) = 0. \tag{20}$$

3.2 Convergence Results

The aim of this section is to prove that the random measures introduced in (15) as well as their normalizations (16) converge weakly to the solution of the system (2). We start with the following lemma

Lemma 3.1 *Let us suppose that, for all $f \in \mathcal{C}_b(E)$, we have*

$$\lim_{N_0 \to \infty} E\left[\left(\eta_n^N(f) - \eta_n(f)\right)^2\right] = 0, \tag{21}$$

then, for all $f \in \mathcal{C}_b(E)$

$$\lim_{N_0 \to \infty} E\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f)\right)^2\right] = 0. \tag{22}$$

Moreover, let us suppose that there exists a constant c_n such that for all $f \in \mathcal{C}_b(E)$, we have

$$E\left[\left(\eta_n^N(f) - \eta_n(f)\right)^2\right] \leq \frac{c_n \, \|f\|^2}{N_0} \tag{23}$$

then, as well, there exists a constant $\hat{c}_n$ such that for all $f \in \mathcal{C}_b(E)$

$$E\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f)\right)^2\right] \leq \frac{\hat{c}_n \, \|f\|^2}{N_0} \tag{24}$$

Proof:
If $N_n > 0$, then we have the following consecutive relations

$$\begin{aligned}
\Psi_n(\eta_n^N)(f) - \Psi_n(\eta_n)(f) &= \frac{\eta_n^N(g_n f)}{\eta_n^N(g_n)} - \frac{\eta_n(g_n f)}{\eta_n(g_n)} \\
&= \frac{\eta_n^N(g_n f)}{\eta_n^N(g_n)} - \frac{\eta_n^N(g_n f)}{\eta_n(g_n)} + \frac{\eta_n^N(g_n f)}{\eta_n(g_n)} - \frac{\eta_n(g_n f)}{\eta_n(g_n)} \\
&\leq \frac{1}{\eta_n(g_n)} \left| \eta_n^N(g_n f) - \eta_n(g_n f) \right| \\
&\quad + \frac{||f||}{\eta_n(g_n)} \left| \eta_n^N(g_n) - \eta_n(g_n) \right| \qquad (25)
\end{aligned}$$

Condition (5) yields

$$E\left[\left(\hat{\eta}_n^N f - \Psi_{n-1}(\eta_n^N)(f) \right)^2 I(N_n > 0) \right] \leq \frac{C}{N_0} ||f||^2 \qquad (26)$$

and Proposition 2.1 tells us that

$$P(N_n = 0) \leq \frac{Cn}{N_0} \qquad (27)$$

Putting together (21), (25), (26) , (27) and the fact that $\hat{\eta}_n(f) = \Psi_n(\eta_n)(f)$, we obtain

$$\begin{aligned}
E&\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f) \right)^2 \right] \\
&= E\left[(\hat{\eta}_n(f))^2 I(N_n = 0) \right] + E\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f) \right)^2 I(N_n > 0) \right] \\
&\leq \frac{Cn\,||f||^2}{N_0} + \frac{2C}{N_0} ||f||^2 + \frac{4}{(\eta_n(g_n))^2} E\left[\left(\eta_n^N(g_n f) - \eta_n(g_n f) \right)^2 \right] \\
&\quad + \frac{4}{(\eta_n(g_n))^2} ||f||^2\, E\left[\left(\eta_n^N(g_n) - \eta_n(g_n) \right)^2 \right] \qquad (28)
\end{aligned}$$

hence (22) holds true. Also from (23) and (28), we get (24). ■

Theorem 3.2 *Let us suppose that the mappings $\mu \to K_{n,\mu} f$, $n \geq 0$, defined on $\mathbf{M}(E)$ with values in $\mathcal{C}_b(E)$ are continuous (pointwise). Then for all $n > 0$ and for all $f \in \mathcal{C}_b(E)$*

$$\lim_{N_0 \to \infty} E\left[\left(\eta_n^N(f) - \eta_n(f) \right)^2 \right] = 0, \qquad (29)$$

$$\lim_{N_0 \to \infty} E\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f) \right)^2 \right] = 0. \qquad (30)$$

Moreover if for every $f \in \mathcal{C}_b(E)$, $\nu \in \mathbf{M}(E)$ and $n \geq 1$ there exist some constant $C_n(\nu)$ and a finite set of bounded functions $\mathcal{H}_n(\nu)$ such that

$$\forall \mu \in \mathbf{M}(E) \qquad ||K_{n,\nu} f - K_{n,\mu} f|| \leq C_n(\nu)\, ||f|| \sum_{h \in \mathcal{H}_n(\nu)} |\nu h - \mu h| \tag{31}$$

then for all $n > 0$ there exist constant c_n, $\hat{c}_n$, such that for all $f \in \mathcal{C}_b(E)$

$$E\left[\left(\eta_n^N(f) - \eta_n(f)\right)^2\right] \leq \frac{c_n\, ||f||^2}{N}. \tag{32}$$

$$E\left[\left(\hat{\eta}_n^N(f) - \hat{\eta}_n(f)\right)^2\right] \leq \frac{\hat{c}_n\, ||f||^2}{N}. \tag{33}$$

Proof:
The previous lemma tells us that (30) follows from (29) and (33) follows from (32), thus we only need to prove (29) and (32) which we do by induction. The initial step is satisfied by hypothesis, since, using the independence of the initial distribution of the particles, we have

$$E\left[\left(\eta_0^N(f) - \eta_0(f)\right)^2\right] = \frac{1}{N^2} \sum_{i=1}^{N} E\left[\left(f(\xi_0^j) - \eta_0(f)\right)^2\right] \leq \frac{||f||^2}{N}$$

We show now that

$$\lim_{N_0 \to \infty} E\left[\left(\eta_n^N(f) - \eta_n(f)\right)^2\right] = 0$$

implies

$$\lim_{N_0 \to \infty} E\left[\left(\eta_{n+1}^N(f) - \eta_{n+1}(f)\right)^2\right] = 0.$$

Since $P(N_{n+1} = 0) \leq \frac{C(n+1)}{N}$ (Proposition 2.1), we have

$$\begin{aligned} E\left[\left(\eta_{n+1}^N(f) - \eta_{n+1}(f)\right)^2 I(N_{n+1} = 0)\right] &= E\left[(\eta_{n+1}(f))^2\, I(N_{n+1} = 0)\right] \\ &\leq \frac{C(n+1)\, ||f||^2}{N}, \end{aligned}$$

hence we only need to concentrate on the set $\{N_{n+1} > 0\}$. For $N_{n+1} > 0$, we have that

$$\begin{aligned} \eta_{n+1}^N(f) - \eta_{n+1}(f) &= \hat{\eta}_n^N K_{n,\hat{\eta}_n^N}(f) - \hat{\eta}_n K_{n,\hat{\eta}_n}(f) \\ &= \hat{\eta}_n^N K_{n,\hat{\eta}_n^N}(f) - \hat{\eta}_n K_{n,\hat{\eta}_n^N}(f) \\ &\quad + \hat{\eta}_n K_{n,\hat{\eta}_n^N}(f) - \hat{\eta}_n K_{n,\hat{\eta}_n}(f) \end{aligned} \tag{34}$$

Using the induction hypothesis we have that

$$\lim_{N_0 \to \infty} E\left[\left(\hat{\eta}_n^N K_{n,\hat{\eta}_n^N}(f) - \hat{\eta}_n K_{n,\hat{\eta}_n^N}(f)\right)^2\right] = 0$$

and using the Lebesgue Dominated Convergence theorem and the continuity of the mapping $\mu \to K_{n,\mu} f$ we get

$$\lim_{N_0 \to \infty} E\left[\left(\hat{\eta}_n K_{n,\hat{\eta}_n^N}(f) - \hat{\eta}_n K_{n,\hat{\eta}_n}(f)\right)^2\right] = 0$$

hence our claim holds true. For the second part of the claim we proceed similarly by induction and use the bound (31) instead of the Dominated Convergence theorem and the fact that for all $n > 0$ and $\mu \in \mathbf{M}(E)$ we have $||K_{n,\nu} f|| \leq ||f||$. ■

Theorem 3.2 and the form of the metric ρ introduced in (17) provide the following straighforward corollary

Corollary 3.3 *If, as in Theorem 3.2, $\mu \to K_{n,\mu} f$, $n \geq 0$, defined on $\mathbf{M}(E)$ with values in $\mathcal{C}_b(E)$ are continuous (pointwise) then, for all $n > 0$, $\lim_{N_0 \to \infty} E\left[\rho(\eta_n^N, \eta_n)\right] = 0$ and $\lim_{N_0 \to \infty} E\left[\rho(\hat{\eta}_n^N, \hat{\eta}_n)\right] = 0$. Moreover, if for every $f \in \mathcal{C}_b(E)$, $\nu \in \mathbf{M}(E)$ and $n \geq 1$ there exist some constant $C_n(\nu)$ and a finite set of bounded functions $\mathcal{H}_n(\nu)$ such that condition (31) is satisfied, then for all $n > 0$ there exist constant c_n and $\hat{c}_n$ such that $E\left[\rho(\eta_n^N, \eta_n)\right] \leq \frac{c_n}{\sqrt{N}}$ and $E\left[\rho(\hat{\eta}_n^N, \hat{\eta}_n)\right] \leq \frac{\hat{c}_n}{\sqrt{N}}$, where ρ is the metric introduced in (17). In both cases, we obtain that η_n^N, respectively $\hat{\eta}_n^N$, converges to η_n, respectively $\hat{\eta}_n$, in probability. Under the same conditions, similar properties are valid for the normalised measures $(m(\xi_n), m(\widehat{\xi}_n)\ ;\ n \geq 0)$ introduced in (16).*

The following are some examples for which condition (31) holds true.

Example 1 *If the transition kernels $K_{n,\mu}$, $n \geq 1$, $\mu \in \mathbf{M}(E)$ does not depend on the measure μ condition (31) is trivially satisfied with*

$$C_n(\nu, f) = 1 \qquad \text{and} \qquad \mathcal{H}_n(\nu, f) = \{f\}$$

Example 2 *If $E = I\!R$ and the transition kernels $K_{n,\mu}$, $n \geq 1$, $\mu \in \mathbf{M}(E)$ are given by*

$$K_{n,\mu}(x, dz) = \frac{1}{\sqrt{2\pi}} \exp -\frac{1}{2}\left(z - \int a_n(x, u)\, \mu(du)\right)^2$$

where $a_n \in \mathcal{C}_b(I\!R^2)$. We begin by noting that for any $\mu, \nu \in \mathbf{M}(E)$ $f \in \mathcal{C}_b(I\!R)$ and $x \in I\!R$ we have

$$K_{n,\mu} f(x) - K_{n,\nu} f(x) = \int f(z)\, \left(e^{I_{n,\mu}(x,z)} - e^{I_{n,\nu}(x,z)}\right)\, \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\, dz$$

where the functions $I_{n,\gamma}(u, v)$ are given by

$$I_{n,\gamma}(u, v) = -\frac{1}{2}\left(\int a_n(u, w)\, \gamma(dw)\right)^2 + v \int a_n(u, w)\, \gamma(dw)$$

for any $(u,v)\in I\!R^2$*,* $n\geq 1$ *and* $\gamma\in \mathbf{M}(E)$*. Using the fact that*

$$\forall(\alpha,\beta)\in I\!R^2 \qquad |e^{\alpha}-e^{\beta}|\leq|\alpha-\beta|\;(e^{\alpha}+e^{\beta})$$

we can find a constant $C<\infty$ *such that*

$$\left|e^{I_{n,\mu}(x,z)}-e^{I_{n,\nu}(x,z)}\right|\leq C\;(1+|z|)e^{\|a_n\||z|}\left|\int a_n(x,u)\;\mu(du)-\int a_n(x,v)\;\nu(dv)\right|$$

This yields

$$\|K_{n,\mu}f-K_{n,\nu}f\|\leq\sup_{x\in I\!R}\left|\int a_n(x,u)\;\mu(du)-\int a_n(x,v)\;\nu(dv)\right|$$

When functions $(a_n\;;\;n\geq 1)$ *have the form*

$$a_n(x,z)=\sum_{r=1}^{R}\;b_{n,r}(x)\;c_{n,r}(z)$$

with $b_r,c_r\in\mathcal{C}_b(I\!R)$*,* $1\leq r\leq R$*, we see that condition (31) holds and*

$$C_n(\nu,f)=1+\max_{1\leq r\leq R}\|b_r\| \qquad \mathcal{H}_n(\nu,f)=\{K_{n,\nu}f,\;c_r\;;1\leq r\leq R\}$$

The BIPS algorithm with multinomial branching corrections have been extensively studied in [11] and [12]. Unless otherwise stated, from now on we assume that the Markov chain $(N_n,\xi_n,\widehat{N}_n,\widehat{\xi}_n\;;\;n\geq 0)$ is assumed to be defined using the Binomial branching corrections introduced in section 2.2. Our aim is to prove the almost sure convergence of the sequences $\eta_n^N f$ to $\eta_n f$ as $N_0\to\infty$, for any time $n\geq 0$ and any test function $f\in\mathcal{C}_b(E)$. We do this via a Borel-Cantelli argument. Since the branching mechanism is of Bernoulli type, $P(N_n>0)=1$ for all $n\geq 0$ and

$$\begin{aligned}E\left[\left(\hat{\eta}_n^N f-\Psi_{n-1}(\eta_n^N)(f)\right)^4\right] &\leq \left(\frac{3}{16N_0^2}+\frac{1}{48N_0^3}+\frac{3n}{64N_0^3}\right)\|f\|^4\\ &\leq \left(\frac{n}{N^2}\right)\|f\|^4 \end{aligned}\tag{35}$$

Theorem 3.4 *Again let us suppose that, for every* $f\in\mathcal{C}_b(E)$*,* $\nu\in\mathbf{M}(E)$ *and* $n\geq 1$ *there exist some constant* $C_n(\nu)$ *and a finite set of bounded functions* $\mathcal{H}_n(\nu)$ *such that (31) holds then for all* $n>0$ *there exist constants* c_n *and* $\hat{c}_n$ *such that for all* $f\in\mathcal{C}_b(E)$

$$\left(E\left[\left(\eta_n^N(f)-\eta_n(f)\right)^4\right]\right)^{\frac{1}{4}}\leq\frac{c_n\,\|f\|}{\sqrt{N}}\tag{36}$$

$$\left(E\left[\left(\hat{\eta}_n^N(f)-\hat{\eta}_n(f)\right)^4\right]\right)^{\frac{1}{4}}\leq\frac{\hat{c}_n\,\|f\|}{\sqrt{N}}\tag{37}$$

Proof:
As before we prove the theorem by induction. The initial step is satisfied by hypothesis, since, using the independence of the initial distribution of the particles, we have

$$\begin{aligned} E\left[\left(\eta_0^N(f)-\eta_0(f)\right)^4\right] &= \frac{1}{N^4}\sum_{i=1}^{N} E\left[\left(f(\xi_0^i)-\eta_0(f)\right)^4\right] \\ &\quad +\frac{6}{N^4}\sum_{1\le i<j\le n} E\left[\left(f(\xi_0^i)-\eta_0(f)\right)^2\right] E\left[\left(f(\xi_0^j)-\eta_0(f)\right)^2\right] \\ &\le \frac{4\,||f||^4}{N^2}. \end{aligned} \tag{38}$$

Hence (36) is satisfied for $n=0$ and $c_0=4$. We assume now that (36) is satisfied for an arbitrary time n and prove that this implies that (37) is satisfied for n and (36) is satisfied for $n+1$. From (25), (35) and the induction hypothesis we get that

$$\begin{aligned} \left(E\left[\left(\hat{\eta}_n^N(f)-\hat{\eta}_n(f)\right)^4\right]\right)^{\frac14} &\le n^{\frac14}\frac{||f||}{\sqrt{N}}+\frac{2\,||g_n||}{\eta_n(g_n)}\frac{c_n\,||f||}{\sqrt{N}} \\ &\le \left(n^{\frac14}+\frac{2}{\eta_n(\widetilde{g}_n)}\right)\frac{c_n\,||f||}{\sqrt{N}} \end{aligned} \tag{39}$$

where $\widetilde{g}_n=\frac{g_n}{||g_n||}$. To get (39) we assumed that $c_n\ge 1$ (otherwise we take $c_n=1$). Hence

$$\hat{c}_n=\left(n^{\frac14}+\frac{2}{\eta_n(\widetilde{g}_n)}\right)c_n$$

Finally from (31), (34) and (39) we have,

$$\left(E\left[\left(\eta_{n+1}^N(f)-\eta_{n+1}(f)\right)^4\right]\right)^{\frac14}\le\frac{c_{n+1}\,||f||}{\sqrt{N}}$$

where

$$c_{n+1}=\left(n^{\frac14}+\frac{2}{\eta_n(\widetilde{g}_n)}\right)\left(1+C_n(\eta_n)\sum_{h\in\mathcal{H}_n(\eta_n)}||h||\right)c_n \tag{40}$$

Corollary 3.5 *If the transition kernels $K_{n,\mu}$, $n>0$ are independent of the measure μ i.e., $K_{n,\mu}\equiv K_n$, then*

$$\left(E\left[\left(\eta_n^N(f)-\eta_n(f)\right)^4\right]\right)^{\frac14}\le 4\prod_{i=1}^{n-1}\left(i^{\frac14}+\frac{2}{\eta_i(\widetilde{g}_i)}\right)\frac{||f||}{\sqrt{N}} \tag{41}$$

$$\left(E\left[\left(\hat{\eta}_n^N(f)-\hat{\eta}_n(f)\right)^4\right]\right)^{\frac14}\le 4\prod_{i=1}^{n}\left(i^{\frac14}+\frac{2}{\eta_i(\widetilde{g}_i)}\right)\frac{||f||}{\sqrt{N}} \tag{42}$$

Corollary 3.6 *Under the same conditions as in Theorem 3.4 we have, almost surely,* $\lim_{N_0\to\infty} \eta_n^N = \eta_n$ *and* $\lim_{N_0\to\infty} \hat{\eta}_n^N = \hat{\eta}_n$.

Proof:
Using a Borel-Cantelli argument, from (36) and (37) we obtain that, almost surely,

$$\lim_{N_0\to\infty} |\eta_n^N(f_m) - \eta_n(f_m)| = 0$$

$$\lim_{N_0\to\infty} |\hat{\eta}_n^N(f_m) - \hat{\eta}_n(f_m)| = 0$$

simultaneously for all the bounded continuous functions f_m which appear in the definition (17) of the metric ρ (including the constant function $f_0 \equiv 1$) which imply $\lim_{N_0\to\infty} \rho(\eta_n^N, \eta_n) = 0$, respectively, $\lim_{N_0\to\infty} \rho(\hat{\eta}_n^N, \hat{\eta}_n) = 0$. ■ *Remark* 3.7:

The results of Theorem 3.4 and Corrolary 3.6 also hold for the empirical measures $(m(\xi_n)$ and $m(\widehat{\xi_n})$. This is straightforward from (36), (37) and the relations

$$\begin{aligned}
|m(\xi_n)f - \eta_n f| &= |\eta_n^N f - \eta_n f + (1 - N_n/N_0)\ m(\xi_n)f| \\
&\leq |\eta_n^N f - \eta_n f| + |\eta_n^N 1 - \eta_n 1| \|f\| \\
|m(\widehat{\xi_n})f - \widehat{\eta}_n f| &= |\widehat{\eta}_n^N f - \widehat{\eta}_n f + (1 - N_n/N_0)\ m(\widehat{\xi}_n)f| \\
&\leq |\widehat{\eta}_n^N f - \widehat{\eta}_n f| + |\widehat{\eta}_n^N 1 - \widehat{\eta}_n 1| \|f\|
\end{aligned}$$

4 Application to the Nonlinear Filtering Problem

The nonlinear filtering problem consists in computing the conditional distributions of internal states in dynamical systems when partial observations are made and random perturbations are present in the dynamics as well as in the sensor. The object of this section is to apply the results obtained in the previous section to this problem. For a detailed discussion of the filtering problem the reader is referred to the pioneering paper of Stratonovich [28] and to the more rigorous studies of Shiryaev [27] and Kallianpur-Striebel [19]. More recent developments can be found in Ocone [22] and Pardoux [23]. We don't present here the standard change of reference probability approach since in the model under study the signal transition will also dependent on the data observed in the past and on the last value of the optimal filter but follow closely the approach of Stettner [26] and Kunita [20].

The basic model for the general filtering problem consists of a 'signal' process $X = (X_n\ ;\ n \geq 0)$ taking values in a locally compact separable metric space E and an 'observation' process $Y = (Y_n\ ;\ n \geq 0)$ taking values in $\mathbb{R}^d$ for some $d \geq 1$. The classical filtering problem is to find conditional distribution of the signal given the observation process $\widehat{\eta}_n$, where

$$\widehat{\eta}_n f = E(f(X_n)/Y_0, \ldots, Y_{n-1}, Y_n) \qquad \forall f \in \mathcal{C}_b(E),\ \ n \geq 0$$

with the associated one step predictor conditional probability η_n, where

$$\eta_n f = E(f(X_n)/Y_0, \ldots, Y_{n-1}) \qquad \forall f \in \mathcal{C}_b(E), \;\; n \geq 0$$

We assume that the initial value X_0 of the signal is an E-valued random variable with law $\eta_0 \in \mathbf{M}_1(E)$ and the one step transition of X at time $n \geq 1$, denoted by $K_{y_{n-1},\widehat{\eta}_{n-1}}$, (possibly) depends on the condional distribution $\widehat{\eta}_{n-1}$ and on the observation data Y_{n-1}. The corresponding transition functions are connected to a given family of transitions

$$\{K_{y,\mu} \; ; \; \mu \in \mathbf{M}_1(E), \; y \in \mathbb{R}^d\}$$

satisfying the following condition

(H1) For any $y \in \mathbb{R}^d$, $f \in \mathcal{C}_b(E)$ the mapping $\mu \in \mathbf{M}(E) \rightarrow K_{y,\mu} f \in \mathcal{C}_b(E)$ is continuous (pointwise).

We also assume that the observation process has the form

$$Y_n = h_n(X_n) + V_n \qquad n \geq 0$$

where $h_n : E \rightarrow \mathbb{R}^d$ are continuous and $(V_n \; ; \; n \geq 0)$ are independent random variables with density $(g_n \; ; \; n \geq 0)$ with respect to Lebesgue measure on $\mathbb{R}^d$ and that the observation noise $(V_n \; ; \; n \geq 0)$ and the signal $(X_n \; ; \; n \geq 0)$ are independent. The function h_n and g_n satisfy the following condition:

(H2) For any time $n \geq 0$, h_n is bounded continuous and g_n is a positive continuous function.

The problem of estimating the conditional distributions of the signal with respect to the observations is of course related to that of recursively computing the conditional distributions $(\eta_n, \widehat{\eta}_n \; ; \; n \geq 0)$. Kunita and Stettner showed that η_n and $\widehat{\eta}_n$ satisfy the following recurrence relations:

Proposition 4.1 (Kunita [20],Stettner [26]) *Given the observations $Y = y$ the conditional distributions $(\eta_n \widehat{\eta}_n \; ; \; n \geq 0)$ are solution of the $\mathbf{M}_1(E)$-valued dynamical system given by*

$$\left\{ \begin{array}{rcl} \widehat{\eta}_n & = & \Psi_n(y_n, \eta_n) \qquad\qquad n \geq 0 \qquad \eta_0 \in \mathbf{M}_1(E) \\ \eta_{n+1} & = & \widehat{\eta}_n \, K_{y_n,\widehat{\eta}_n} \end{array} \right. \tag{43}$$

where

- *y_n is the given current observation at time n.*
- *$K_{y_n,\widehat{\eta}_n}$ is the transition function of the signal at time n.*

- $\Psi_n(y_n, \centerdot) : \mathbf{M}_1(E) \to \mathbf{M}_1(E)$ *is the continuous function given by*

$$\forall \eta \in \mathbf{M}_1(E) \quad \forall f \in \mathcal{C}_b(E) \qquad \Psi_n(y_n,\eta)f = \frac{\int f(x)\, g_n(y_n - h_n(x))\, \eta(dx)}{\int g_n(y_n - h_n(z))\, \eta(dz)}$$

Equation (43) is usually called the non linear filtering equation. It involves two separate mechanisms. Namely the first one

$$\eta_n \longrightarrow \Psi_n(y_n, \eta_n)$$

updates the predictor conditional distribution η_n given the current observation $Y_n = y_n$. This first mechanism is called the correction or the updating transition. The second transition

$$\widehat{\eta}_n \longrightarrow \widehat{\eta}_n\, K_{y_n, \widehat{\eta}_n}$$

does not depend on the observation at time $n+1$ and is usually called the prediction step.

In this formulation the conditional distributions $(\eta_n, \widehat{\eta}_n\ ;\ n \geq 0)$ are parametrized by a given observation record $(y_n\ :\ n \geq 0)$ and they are solution of the measure valued dynamical system given by (43) so that the BIPS approaches introduced in section 2 can be applied. In the following, we treat only the BIPS algorithm constructed using the Bernoulli branching corrections introduced in section 2.2. The algorithm with multinomial corrections is described in all details in[11] and [12].

For the moment, we assume that the BIPS $((N_n^y, \xi_n^y), (\widehat{N}_n^y, \widehat{\xi}_n^y)\ ; n \geq 0)$ depend on the arbitrary, but *fixed*, observation record $(y_n\ ;\ n \geq 0)$. Then the approximation of the desired conditional distributions $(\eta_n^y, \widehat{\eta}_n^y\ ;\ n \geq 0)$ is guaranteed by the theorems 3.2, 3.4 and their corollaries. Indeed, using Corollary 3.3 Corollary 3.5 and Corollary 3.6, we find:

Proposition 4.2 *If the conditions (H1) and (H2) hold then, for any time $n \geq 0$, we have $\lim_{N_0 \to \infty} E[\rho(\eta_n^{N,y}, \eta_n^y)] = 0$ and $\lim_{N_0 \to \infty} E[\rho(\hat{\eta}_n^{N,y}, \eta_n^y)] = 0$ Furthermore, if we assume that the transition $K_{y,\mu}$ does not depend on the parameter μ then $\lim_{N_0 \to \infty} \eta_n^{N,y} = \eta_n^y$ and $\lim_{N_0 \to \infty} \hat{\eta}_n^{N,y} = \hat{\eta}_n^y$. and*

$$E[(\eta_n^{N,y}(f) - \eta_n^y(f))^4] \leq 2^8 \prod_{i=1}^{n-1} \left(i^{\frac{1}{4}} + \frac{2}{\eta_i^y(\tilde{g}_i^y)} \right)^4 \frac{\|f\|^4}{N_0^2} \tag{44}$$

$$E[(\hat{\eta}_n^{N,y}(f) - \hat{\eta}_n^y(f))^4] \leq 2^8 \prod_{i=1}^{n} \left(i^{\frac{1}{4}} + \frac{2}{\eta_i^y(\tilde{g}_i^y)} \right)^4 \frac{\|f\|^4}{N_0^2} \tag{45}$$

We remove now the assumption of having a fixed observation record $(y_n\ ;\ n \geq 0)$ and impose another condition of the filtering system:

(H3) For any time $n \geq 0$ we have $m := E\left[\prod_{i=1}^{n}\left(1+\frac{2}{\eta_i(\widetilde{g}_i)}\right)\right] < \infty$

For any $N_0 \geq 1$ and $n \geq 0$ we denote by Φ_n^N ($\widehat{\Phi}_n^N$) the distribution of the random measure η_n^N (ref. $\widehat{\eta}_n^N$) and we denote by Φ_n ($\widehat{\Phi}_n$) the distribution of the random measure η_n (ref. $\widehat{\eta}_n$).

Proposition 4.3 *Assume that the conditions (H2) and (H3) hold true and that the transition $K_{y,\mu}$ does not depend on the parameter μ. Then, for any time $n \geq 0$, we have*

$$\forall n \geq 0 \qquad \lim_{N_0\to\infty} D(\Phi_n^N, \Phi_n) = 0 \quad \text{and} \quad \lim_{N_0\to\infty} D(\widehat{\Phi}_n^N, \widehat{\Phi}_n) = 0 \tag{46}$$

where D is the Vasershtein metric introduced in (18).

Proof:
Straightforward from Proposition 4.2 and (19). ■

Corollary 4.4 *If the transition kernels $K_{y,\mu}$, are independent of the measure μ the conditions (H2) and (H3) are satisfied, then*

$$E[(\eta_n^N(f) - \eta_n(f))^4] \leq \frac{(4m)^4\|f\|^4}{N_0^2} \tag{47}$$

Finally we present two examples for which (H3) is satisfied.

Example 3 *As a typical example of a non-linear filtering problem, assume the functions $h_n : E \to I\!R^d$, $n \geq 1$, 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

$$Y_n = h_n(X_n) + V_n \qquad \forall n \geq 1 \tag{48}$$

where $(V_n)_{n\geq 1}$ is a sequence of $I\!R^d$-valued and independent random variables with Gaussian densities. it is easy to show that the exists a constant M_h^n which depends only on h_n and R_n^{-1} so that for all $i \geq 0$

$$\begin{aligned} \eta_i(\widetilde{g}_i)) &\geq M_h^n \exp(-\|R_n^{-1}\|\ \|h_n\|\ \|Y_n\|) \\ &\geq M_h^n \exp(-\|R_n^{-1}\|\ \|h_n\|^2 - \|R_n^{-1}\|\ \|\ \|h_n\|\ \|V_n\|) \end{aligned}$$

where $\|R_n^{-1}\|$ is the spectral radius of R_n^{-1}. This implies, using the independence of the random variables V_n and the existence of their exponential moments, that condition (H3) holds.

Example 4 *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 one can easily check that condition (H3) is satisfied.

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] BILLINGSLEY P., (1968), Convergence of probability measures ,*Wiley, New York.*

[3] 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**.

[4] CRISAN D., LYONS T.J. (1996). Nonlinear Filtering And Measure Valued Processes. *Imperial College London, preprint.*

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

[6] CRISAN D., GAINES J., LYONS T.J. (1996). A Particle Approximation of the Solution of the Kushner-Stratonovitch Equation. *Imperial College, London, preprint.*

[7] DAWSON, D.A., GARTNER, J. (1987) Large déviations from the Mc Kean-Vlasov limit for weakly interacting diffusions. *Stochastics*, **20**, 247-308

[8] DAWSON D.A. .Stochastic Evolution Equation and Related Measure Processes. *Journal of Multivariate Analysis,* **5**, *1-52.*

[9] DAWSON D.A. . *Ecoledété de Saint Flour*

[10] DEL MORAL P. (1996). Non linear filtering using random particles. *Theory of Probability and its Applications, Vol. 40, (4), pp. 690-701.*

[11] DEL MORAL P. (1996). Non-linear Filtering: Interacting particle solution. *Markov Processes and Related Fields, Vol. 2, pp. 555-580*

[12] 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 Probability.

[13] 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.*

[14] 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.*

[15] DEL MORAL P., GUIONNET A. (1997). A Central Limit Theorem for Non Linear Filtering using Interacting Particle Systems. *Publications du Laboratoire de Statistiques et Probabilités, Université Paul Sabatier, No 11-97*

[16] ETHERIDGE A., MARCH P. (1991), A note on superprocesses, Prob. Th. Rel. Fields 89, 141-147.

[17] GORDON N.J., SALMON D.J., SMITH A.F.M. (1993). Novel Approach to non-linear/non-Gaussian Bayesian state estimation., *IEE Proceedings, Part F, Vol. 140, (2), 107-113.*

[18] HITSUDA, M., TANAKA, H.(1981) Central limit theorem for a simple diffusion model for interacting particles. *Hiroshima Mathematical Journal*,**11**, 415-423

[19] KALLIANPUR G., STRIEBEL C. (1967), Stochastic differential equations occuring in the estimation of continuous parameter stochastic processes, *Tech. Rep. No 103, Department of statistics, Univ. of Minnesota.*

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

[21] NIEMANN H., (1990) Pattern Analysis and Understanding, Springer-Verlag.

[22] OCONE D.L. (1980), Topics in Nonlinear Filtering Theory, *Ph.D. Thesis, MIT, Cambridge.*

[23] PARDOUX E. (1991), Filtrage Non Linéaire et Equations aux Dérivés Partielles Stochastiques Associées, *Ecole d'été de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Mathematics, 1464, Springer-Verlag*

[24] PARTHASARATHY K.R. (1968), Probability measures on Metric Spaces, *Academic Press Inc., New York.*

[25] PERKINS E.A., (1991), Conditional Dawson-Watanabe processes and Fleming-Viot processes, Seminar in Stochastic Processes, 1991, Birkhauser, 142-155.

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

[27] SHIRYAEV A.N. (1966), On stochastic equations in the theory of conditional Markov processes, *Theor. Prob. Appl. (11) (179-184).*

[28] STRATONOVITCH R.L.,(1960), Conditional Markov processes, *Theor. Prob. Appl. (5) (156-178).*

[29] SZNITMAN A.S. (1984). Non linear reflecting diffusion process and the propagation of chaos and fluctuations associated. *J. Functional Anal.* **56** *(3) 311-336.*

[30] SZNITMAN A.S. (1989). Topics in Propagation of Chaos. *Ecole d'été de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Mathematics,* **No 1464.**