The Monte-Carlo method for filtering with discrete-time observations: Central Limit Theorems

Pierre Del Moral* and Jean Jacod†

Summary

We study an interacting particle system which allows to explicitely compute the filtering measure, via Monte-Carlo techniques, for a discrete-time or continuous-time Markov state process and a discrete-time observation process. This particle system was introduced and proved to converge, as the number of particles goes to infinity, in [3]. Here we are interested in a precise description of the error, in the sense of a central limit theorem for the normalized errors. A particular feature is that we take care of the fact that the transition probabilities for the state process are usually not explicitely known. We describe the normalizing factor as a function of the number of particles, and also as a function of the number of variables which are necessary to simulate (hence of the time actually necessary to perform the procedure).

Keywords: Filtering, Monte-Carlo methods, Diffusion processes, Interacting particle systems

A.M.S. Classification: 60G35, 60F17, 93E11

*LSP, Bât. 1R1, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France (CNRS UMR C5583)

†Laboratoire de Probabilités, Université Paris 6, 4 place Jussieu, 75252 Paris, France (CNRS UMR 7599)

1 Introduction

In [3] we have introduced a Monte-Carlo method for solving a filtering problem where the state follows a continuous-time stochastic differential equation, but the number of observations is finite. More precisely, We have an $I\!R^d$-valued state process $X = (X_t)_{t\geq 0}$ which evolves according to a stochastic differential equation of the form

$$dX_t \;=\; a(X_t)dt + b(X_t)dW_t, \qquad \mathcal{L}(X_0) \;=\; \mu, \tag{1.1}$$

where μ is a known distribution on $I\!R^d$, and a, b are known functions, and W is a d-dimensional Wiener process. We have noisy observations $Y_1, \ldots, Y_N$ at N regularly spaced times, and without loss of generality we will assume that these times are $1, \ldots, N$.

We considered three different cases, as follows:

Case A: The observations take the form

$$Y_i \;=\; h(X_i, \varepsilon_i), \tag{1.2}$$

where the ε_i are i.i.d. q'-dimensional variables, independent of X and with a law having a known density, and h is a known function from $I\!R^d \times I\!R^{q'}$ into $I\!R^q$. For convenience, we take $Y_0 = 0$.

Case B: The observations $Y_1, \ldots, Y_n$ are the values at times $1, \ldots, N$ of an $I\!R^q$-valued process $(Y_t)_{t\geq 0}$ which satisfies

$$dY_t \;=\; h(X_t)dt + \sigma dW'_t, \qquad Y_0 \;=\; 0, \tag{1.3}$$

where h is a known function and W' is a q-dimensional Wiener process, independent of W, and σ is an $q \times q$-matrix.

Case C) The observed values Y_i are the values at times i of an $I\!R^q$-valued process $(Y_t)_{t\geq 0}$ which satisfies

$$dY_t \;=\; a'(X_t, Y_t)dt + b'(X_t, Y_t)dW'_t, \qquad Y_0 = 0, \tag{1.4}$$

where a' and b' are known functions and W' is a q-dimensional Wiener process, independent of W.

In each case, the two sequences $(X_i)_{i\in I\!N}$ and $(X_i, Y_i)_{i\in I\!N}$ are homogeneous Markov chains.

In [3] we described a procedure allowing to approximate the conditional expectations (or filter)

$$\pi_N f \;=\; E(f(X_N)|Y_0 = y_0, \ldots, Y_N = y_N) \tag{1.5}$$

for reasonable functions f on $I\!R^d$, and for a given (fixed) sequence $(y_1, \ldots, y_N)$. This procedure will be recalled below, and it hinges upon the simulation of a system of n interacting particles: we proved that, under suitable assumptions on the coefficients appearing in our equations, the constructed approximate filters $\pi_N^n f$ converge as $n \to \infty$ to $\pi_N f$, and the error normalized with a suitable power of n, say n^α for an α which again depends on the case under consideration, is bounded in all $\mathbb{L}^p$ ($p < \infty$), when f is bounded Borel (and Lipschitz continuous in case (B)). More precisely we got in case A, say, $\alpha = 1/2$ and a bound, for the $\mathbb{L}^1$ estimate, of the form

$$E(\sqrt{n}|\pi_N^n f - \pi_N f|) \ \leq \ F_N \sup_x |f(x)| \tag{1.6}$$

(the expectation above is taken over the simulated variables, while the observed values $(y_1, \ldots, y_N)$ is kept fixed; the estimates are similar in the other two cases). However, we made no attempt to optimize the constant F_N, and in [3] this constant is of the form $F_N = K^{N^2}$, where K depends on the coefficients of the model and on the observation $(y_1, \ldots, y_N)$.

Our aim in this paper is to prove a central limit theorem associated with the above convergence (as $n \to \infty$), for functions f which are bounded, or with polynomial or even exponential growth, according to the case under consideration. A first motivation for this is to get rid of the boundedness assumption on f in (1.6), in order to accomodate such functions as $f(x) = x$. The second and main motivation is to obtain accurate bounds on the normalized error $n^\alpha(\pi_N^n f - \pi_N f)$, especially as N varies, since a bound like F_N above would give rise to the conclusion that our Monte-Carlo procedure is totally out of reach for even moderate sized N. Studying the behaviour of our sharp estimates as $N \to \infty$ in the case where the basic process X is ergodic will hopefully be the object of a forthcoming paper. So here we concentrate on the CLT for fixed N.

In addition to this, we also mention that in Case C we take here other normalizations than in [3], which lead to a faster rate of convergence, although the method is basically the same.

The paper is organized as follows: first we consider a general Monte-Carlo approach, suitable for any Markov chain. The procedure is described in Section 2, the consistency is studied in Section 3 and the CLT is given in Section 4; the assumptions H1 and H2 for these general results are quite technical, and are in fact designed to work for all our three cases. The next three sections are concerned respectively with cases A, B and C: by far, the most difficult case is B. One of the difficulties comes from the fact that we usually do not explicitely know the transition probabilities of the Markov chain $X_0, X_1, \ldots, X_N$, so we might have to resort on an Euler approximation procedure to perform the simulation. So we single out the cases (labeled A-1, B-1 and C-1) where, due to some special feature of the model, we do not need to use an Euler scheme, and A-2, B-2, C-2 where we do need an Euler scheme: we will see that for A-1, B-1 and C-1, and for a suitable real α depending on the case under consideration, the sequence $n^\alpha(\pi_N^n f - \pi_N f)$ converges as $n \to \infty$ to a centered normal variable whose variance is given explicitely, while in cases A-2, B-2 and C-2 we have the same rate, but the limit is biased normal.

Finally, let us observe that the number n of particles is not really an accurate measure of the time necessary to perform the algorithm, especially in cases A-2, B-2 and C-2, since the stepsize of the Euler scheme decreases (hence the simulation time for each simulated path increases) as n increases. A more accurate measure of the time is the number r of single variables one has to simulate. So we devote Section 8 to a version of the CLT in terms of this number r.

2 A Monte-Carlo approach in a general case

1) In the next three sections we let our specific problem aside, and consider a much more general setting in discrete time filtering theory.

The state process $(X_n)_{n\geq 0}$ is a Markov chain on an abstract space $(E, \mathcal{E})$, with transition Q and initial law μ. The observation process $(Y_n)_{n\geq 0}$ takes values in $I\!R^q$, with by convention $Y_0 = 0$, and is such that the pair $(X_n, Y_n)_{n\geq 0}$ is again a Markov chain with transition P which factorizes as

$$P(x, y; dx', dy') \;=\; Q(x, dx')G(x, y; x', y')dy'. \tag{2.1}$$

We are interested in the conditional expectations

$$\pi_{Y,N} f \;=\; E(f(X_N)|Y_0, \ldots, Y_N) \tag{2.2}$$

for a suitable class $\mathcal{C}$ of measurable functions f on E.

We fix the observation sequence $(y_0 = 0, y_1, y_2, \ldots, y_N)$, and instead of $\pi_{Y,N}$ in (2.2) we simply write π_N. Then, with G as in (2.1), we introduce the kernels H_k from $(E, \mathcal{E})$ into itself by

$$H_k f(x) \;=\; \int Q(x, dx')G(x, y_{k-1}; x', y_k)f(x')dx'. \tag{2.3}$$

These kernels are positive and σ-finite, but perhaps unbounded. With this notation, we have a recursive way to write π_N, which goes as follows: First,

$$\pi_0 f \;=\; \mu(f) \;=\; \int f(x)\mu(dx). \tag{2.4}$$

Then π_j is given by the following recursive formula, starting with (2.4):

$$\pi_j f \;=\; \frac{\pi_{j-1} H_j f}{\pi_{j-1} H_j 1}. \tag{2.5}$$

Since the vector $(Y_1, \ldots, Y_N)$ has a density w.r.t. Lebesgue measure on $I\!R^{qN}$, we can asssume that $(y_1, \ldots, y_N)$ is in the set where this density is strictly positive, which amounts to assume that the denominator in (2.5) is strictly positive for all $j = 1, \ldots, N$. Note that $\pi_j f$ is well-defined for all measurable functions f which are either nonnegative or bounded, at least.

For further reference, let us also introduce (with I being the identity kernel) the following kernels $H_{k,j}$ for $0 \leq k \leq j$ and numbers $\beta(j,k)$ for $1 \leq k \leq j$:

$$H_{k,j} \;=\; \begin{cases} I & \text{if } \; k = j, \\ H_{k+1} \ldots H_j & \text{if } \; k < j. \end{cases} \tag{2.6}$$

$$\beta(k,j) \;=\; \prod_{i=k}^{j} \pi_{i-1} H_i 1. \tag{2.7}$$

We deduce from (2.5) that

$$0 \leq k < j \quad \Rightarrow \quad \pi_j f \;=\; \frac{\pi_k H_{k,j} f}{\pi_k H_{k,j} 1}, \tag{2.8}$$

and in particular $\pi_n f = \frac{\mu H_{0,n} f}{\mu H_{0,n} 1}$. Note that this looks like an "explicit" formula, at least when Q and G are explicitely given themselves (which is not the case in the situation we are interested in, at least when the Euler scheme is needed). But even so this explicit formula involves integration on a space of large dimension, which is not feasible in practice. This is why we have to resort on a Monte-Carlo procedure.

3) Now we introduce our Monte-Carlo procedure. The integer n will be fixed, and it counts the number of "particles". We consider an auxiliary measurable space $(F_n, \mathcal{F}_n)$ which may depend on n, and for each $k \geq 1$ we consider:

$$\left.\begin{array}{l} \text{a transition probability } R_{n,k} \text{ from } (E,\mathcal{E}) \text{ into } (E\times F_n, \mathcal{E}\otimes\mathcal{F}_n), \\[2mm] \text{a nonnegative measurable bounded function } \psi_{n,k} \text{ on } (E\times F_n, \mathcal{E}\otimes\mathcal{F}_n). \end{array}\right\} \tag{2.9}$$

The kernels $R_{n,k}$ and functions $\psi_{n,k}$ are chosen somewhat arbitrarily, but they should bear nice relationships with the problem at hand: these relationships will be expressed in the Hypotheses (H1) and (H2) to come later. But a very important property of the kernels $R_{n,k}$ should be that for any x *one can simulate variables according to the law* $R_{n,k}(x,.)$.

Then the algorithm goes according to the following steps, in which $\Psi_k^n = \Psi_k^n(\omega, dx)$ denotes a random probability measure on $(E,\mathcal{E})$:

• **Step 0:** We set $\Psi_0^n = \mu$ (= the initial distribution).

• **Step** $k \geq 1$**:** Knowing Ψ_{k-1}^n from the previous step, we simulate n i.i.d. variables $(X_{k-1}'^j)_{1\leq j\leq n}$ according to the law Ψ_{k-1}^n; then for each j we simulate a variable (X_k^j, Z_k^j) according to the law $R_{n,k}(X_{k-1}'^j, .)$. Then for every fonction f on E we set

$$W_k^n f \;=\; \frac{1}{n}\sum_{i=1}^n f(X_k^i)\psi_{n,k}(X_k^i, Z_k^i), \tag{2.10}$$

and finally we introduce the random probability measure Ψ_k^n on $(E,\mathcal{E})$ by

$$\Psi_k^n \;=\; \begin{cases} \frac{1}{nW_k^n 1}\sum_{i=1}^n \psi_{n,k}(X_k^i, Z_k^i)\varepsilon_{X_k^i} & \text{if} \quad W_k^n 1 > 0 \\[2mm] \varepsilon_{x_0} & \text{otherwise,} \end{cases} \tag{2.11}$$

where x_0 is an arbitrary point in E (the measure ε_{x_0} above could in fact be replaced by any probability measure on $(E,\mathcal{E})$).

We stop at the end of Step N, and our approximation of $\pi_N f$ is $\Psi_N^n f$ (observe that by construction $\Psi_N^n f$ is well defined for any function f on E, even though $\pi_N f$ could be not well defined).

3 Consistency of the Monte-Carlo procedure

In this section we give conditions under which the approximations $\pi_k^n f$ converge in probability to $\pi_k f$, for all $k \leq N$. The observation $(y_1, \dots, y_N)$ is fixed, so $\pi_k f$ is deterministic, and the randomness comes from the Monte-Carlo procedure only. We are not mainly interested in consistency, though, so our conditions are not the weakest ones possible, but they are designed in view of the CLT to come afterwards.

Let us first introduce bounded positive kernels H_k^n and $H_k'^n$ on $(E, \mathcal{E})$ by (recall that each function $\psi_{n,k}$ is bounded):

$$\left.\begin{array}{lcl} H_k^n f(x) &=& \int R_{n,k}(x; dx', dz) f(x') \psi_{n,k}(x', z), \\ H_k'^n f(x) &=& \int R_{n,k}(x; dx', dz) f(x') \psi_{n,k}(x', z)^2. \end{array}\right\} \tag{3.1}$$

Then the assumption goes as follows:

Hypothesis (H1): There are an algebra $\mathcal{C}$ of measurable functions on $(E, \mathcal{E})$ containing the constants, and two sequences of positive numbers δ_n and δ'_n, and a number $r > 1$, such that (recall that k is in $\{1, \dots, N\}$):

a) $\mathcal{C} \in \mathbb{L}^1(\mu)$.

b) If $||.||$ denotes the supremum norm, then

$$\delta_n \left(\frac{\delta_n ||\psi_{n,k}||}{n} \right)^{r-1} \rightarrow 0. \tag{3.2}$$

c) We have

$$\delta_n \rightarrow \infty, \qquad \frac{\delta_n^2}{n} \rightarrow u \in [0, \infty). \tag{3.3}$$

d) We have $\delta'_n \rightarrow 0$; further, for each $f \in \mathcal{C}$ we have $H_k f \in \mathcal{C}$ and there exists two functions $H'_k f$ and $L_k f$ in $\mathcal{C}$ with $L_k f \geq 0$, such that

$$|H_k^n f - H_k f| \leq \frac{n \delta'_n}{\delta_n^2} L_k f, \qquad |H_k'^n f - \frac{n}{\delta_n^2} H'_k f| \leq \frac{n \delta'_n}{\delta_n^2} L_k f. \tag{3.4}$$

Under (H1-c), as soon as (H1-b) holds for some r it also holds for all $r' \geq r$.

We are interested in the error $\Psi_k^n f - \pi_k f$, normalized by the sequence δ_n above, so we set for any $f \in \mathcal{C}$ (recall that $\Psi_k^n f$ is well defined for all functions f on E, and (H1-a) together with (2.4) and (2.5) imply that $\pi_k f$ is well defined for $f \in \mathcal{C}$):

$$U_k^n f = \delta_n (\Psi_k^n f - \pi_k f). \tag{3.5}$$

We also set for any $f \in \mathcal{C}$:

$$\hat{W}_k^n f = \sum_{i=1}^n \xi_{k,i}^n f, \quad \text{where} \quad \xi_{k,i}^n f = \frac{\delta_n}{n} (f(X_k^i) \psi_{n,k}(X_k^i, Z_k^i) - \Psi_{k-1}^n H_k^n f). \tag{3.6}$$

(let us recall that the variables X_k^i, etc..., depend also on n).

Theorem 3.1 *Assume (H1). Then*

a) We have δ_n-consistency on $\mathcal{C}$, in the sense that the sequence $(U_k^n f)_{n\geq 1}$ is tight for all $k = 0, \ldots, N$ and all $f \in \mathcal{C}$.

b) If $f_1, \ldots, f_m \in \mathcal{C}$, the sequence $(\hat{W}_k^n f_j)_{1\leq k\leq N, 1\leq j\leq m}$ converges in law to a variable $(U_{k,j})_{1\leq k\leq N, 1\leq j\leq m}$ which is centered Gaussian, with covariance $L(k,j;k',j')$ given by

$$L(k,j;k',j') \;=\; \begin{cases} \pi_{k-1}H'_k(f_jf_{j'}) - u(\pi_{k-1}H_kf_j)(\pi_{k-1}H_kf_{j'}) & \text{if} \quad k = k' \\ 0 & \text{if} \quad k \neq k'. \end{cases} \tag{3.7}$$

Remark: It can be deduced from the proof below that for the claim (a) above the assumption (H1-b) is useless, and in (3.3) the sequence δ_n^2/n need only to be bounded, and also in the right side of the last expression in (3.4) it is enough to have $\frac{n}{\delta_n^2}L_kf$. $\square$

Proof. 1) Below, f denotes an arbitrary function in $\mathcal{C}$. We have

$$W_k^n f \;=\; \frac{1}{\delta_n}\hat{W}_k^n f + \Psi_{k-1}^n H_k^n f. \tag{3.8}$$

We also set $g = (f_1, \ldots, f_m)$ and write $\xi_{k,i}^n$ for the random vector with components $(\xi_{k,i}^n f_j)_{0\leq j\leq m}$. Let $\mathcal{H}_0^n$ be the trivial σ-field. For $(k-1)n < i \leq kn$ we set

$$\eta_i^n \;=\; \xi_{k,i-(k-1)n}^n, \qquad \mathcal{H}_i^n \;=\; \sigma(X_\ell^j : (\ell-1)n + j \leq i).$$

The sequence of σ-fields $\mathcal{H}_i^n$ increases with i, and η_i^n is $\mathcal{H}_i^n$-measurable. Below, we take an even integer $r > 2$ satisfying (3.2), and C will denote a constant which may change from line to line, and may depend on r and on g or on the function $f \in \mathcal{C}$ under consideration. Set $\varepsilon_n = \sup_{1\leq k\leq N} \delta_n \left(\frac{\delta_n||\psi_{n,k}||}{n}\right)^{r-1}$.

By construction, conditional on $\mathcal{H}_{(k-1)n}^n$, the variables (X_k^i, Z_k^i) for $i = 1, \ldots, n$ are i.i.d., with law $\Psi_{k-1}^n R_{n,k}(.)$. Hence if $(k-1)n < i \leq kn$ we have

$$E(\eta_i^n | \mathcal{H}_{i-1}^n) \;=\; 0, \tag{3.9}$$

$$E(\eta_i^n \eta_i^{n,*} | \mathcal{H}_{i-1}^n) \;=\; \frac{\delta_n^2}{n^2}\left(\Psi_{k-1}^n H_k'^n(gg^*) - (\Psi_{k-1}^n H_k^n g)(\Psi_{k-1}^n H_k^n g^*)\right). \tag{3.10}$$

Further, the function $g' = |g|^r$ belongs to $\mathcal{C}$, because r is an even integer and $\mathcal{C}$ is an algebra. So if we use the property $E(|Y - E(Y|\mathcal{G})|^r|\mathcal{G}) \leq 2^r E(|Y|^r|\mathcal{G})$ for every m-dimensional variable Y and any σ-field $\mathcal{G}$ we get, again for $(k-1)n < i \leq kn$:

$$\begin{aligned} E(|\eta_i^n|^r | \mathcal{H}_{i-1}^n) \;&\leq\; 2^r \frac{\delta_n^r}{n^r} E(g'(X_k^{i-(k-1)n})\psi_{n,k}(Z_k^{i-(k-1)n})||\psi_{n,k}||^{r-1} \; |\mathcal{H}_{i-1}^n) \\ &\leq\; 2^r \frac{\varepsilon_n}{n} \Psi_{k-1}^n H_k^n g'. \end{aligned} \tag{3.11}$$

2) We will prove the result by induction on N. Let us denote by (C-N) the property (a), and (D-N) the property (b). Since $\Psi_0^n = \mu = \pi_0$ and any f in $\mathcal{C}$ is μ-integrable, we trivially have (C-0).

3) Now, we prove that (D-k)+(C-(k-1))$\Rightarrow$(C-k) if $k \geq 1$. Let $f \in \mathcal{C}$. Using (D-k) and (3.8) and $\delta_n \to \infty$, we see that

$$\text{the sequence } \ (\delta_n(W_k^n f - \Psi_{k-1}^n H_k^n f))_{n\geq 1} \quad \text{is tight.} \tag{3.12}$$

Next (3.4) yields that $\delta_n|\Psi_{k-1}^n H_k^n f - \Psi_{k-1}^n H_k f| \leq \Psi_{k-1}^n L_k f$, and by (C-(k-1)) and $L_k f \in \mathcal{C}$ we have that $\Psi_{k-1}^n L_k f \to^P \pi_{k-1} L_k f$, hence

$$\text{the sequence } \ (\delta_n(\Psi_{k-1}^n H_k^n f - \Psi_{k-1}^n H_k f))_{n\geq 1} \quad \text{is tight.} \tag{3.13}$$

Since $H_k f \in \mathcal{C}$, (C-(k-1)) again yields that the sequence $(\delta_n(\Psi_{k-1}^n H_k f - \pi_{k-1} H_k f))_{n\geq 1}$ is tight. Combining this (3.12) and (3.13) yields

$$\text{the sequence } \ (\delta_n(W_k^n f - \pi_{k-1} H_k f))_{n\geq 1} \quad \text{is tight.} \tag{3.14}$$

Next, combining (2.5) and (2.11), we obtain with the notation $\pi'_k = \pi_{k-1} H_k$:

$$\delta_n(\Psi_k^n f - \pi_k f) \ = \ \frac{\delta_n(W_k^n f - \pi'_k f)\pi'_k 1 - \delta_n(W_1^n - \pi'_k 1)\pi'_k f}{(W_k^n 1)(\pi'_k 1)}$$

on the set $\{W_k^n 1 > 0\}$. But (3.14) implies that $W_k^n 1 \to^P \pi'_k 1$ (recall that $1 \in \mathcal{C}$), and by hypothesis $\pi'_k 1 > 0$, so the above relation and (3.14) applied to f and 1 imply the tightness of the sequence $(\delta_n(W_k^n f - \pi_k f))_{n\geq 1}$. This, being true for all $f \in \mathcal{C}$, gives (C-k).

4) In order to complete the induction, it remains to prove that (C-(k-1))$\Rightarrow$(D-k) for $k \geq 1$. So we assume (C-(k-1)). Using the fact that $\mathcal{C}$ is an algebra containing the constants (hence if $f \in \mathcal{C}$ there is $f' \in \mathcal{C}$ with $|f| \leq f'$) and (H1-d) (which in particular implies that $\sup_n \delta'_n < \infty$) and (3.3), one sees that there are functions $h_j \geq 1$ (hence $\Psi_{j-1}^n h_j \geq 1$) in $\mathcal{C}$ such that, for $(j-1)n < i \leq jn$:

$$\left| E(\eta_i^n \eta_i^{n,*} | \mathcal{H}_{i-1}^n) - \frac{1}{n}\left(\Psi_{j-1}^n H'_j(gg^*) - \frac{\delta_n^2}{n}(\Psi_{j-1}^n H_j g)(\Psi_{j-1}^n H_j g^*)\right)\right| \leq \ \frac{\delta'_n}{n}(\Psi_{j-1}^n h_j)^2.$$

Now, (C-(k-1)) yields that $\Psi_{j-1}^n f \to^P \pi_{j-1} f$ for $f \in \mathcal{C}$ and $j \leq k$, so we deduce that for all $t \leq k$ (using also (3.3) and (3.10)):

$$\sum_{i=1}^{[nt]} E(\eta_i^n \eta_i^{n,*} | \mathcal{H}_{i-1}^n) \ \to^P \ A_t \ , \tag{3.15}$$

where

$$A_t \ = \ \sum_{j=1}^{[t]} (\pi_{j-1} H'_j(gg^*) - u(\pi_{j-1} H_j g)(\pi_{j-1} H_j g^*))$$
$$+\{t\} \left(\pi_{[t]} H'_{1+[t]}(gg^*) - u(\pi_{[t]} H_{1+[t]} g)(\pi_{[t]} H_{1+[t]} g^*)\right),$$

where $[x]$ and $\{x\}$ are the integer and fractional parts of x.

Note also that $\Psi_{j-1}^n H_j^n f \to^P \pi_{j-1} H_j f$ for $f \in \mathcal{C}$ and $j \leq k$ by (C-(k-1)) and (3.4), hence by (3.11):

$$\sum_{i=1}^{[nt]} E(|\eta_i^n|^r | \mathcal{H}_{i-1}^n) \ \to^P \ 0, \tag{3.16}$$

It remains to apply a well known theorem on the convergence of triangular arrays of martingale increments (use (3.9)) satisfying the Lindeberg condition by (3.16): if $Z_t^n = \sum_{i=1}^{[nt]} \eta_i^n$, we deduce from (3.15) that the processes Z^n (indexed by the time interval $[0,k]$) converge in law to a continuous Gaussian martingale Z with $Z_0 = 0$ and bracket A. If we observe that $\hat{W}_k^n g = Z_{kn}^n - Z_{(k-1)n}^n$, Property (D-k) immediately follows. $\square$

4 A general Central Limit Theorem

In this section we give a central limit theorem for the errors given by (3.5). For this we need an assumption stronger than (H1), and which goes as follows:

Hypothesis (H2): We have (H1) with some class $\mathcal{C}$ and some sequence δ_n. Furthermore there exists a subspace $\mathcal{C}'$ of $\mathcal{C}$ containing the constants and an increasing sequence of functions $g_p \in \mathcal{C}$ with $0 \leq g_p \leq 1$ and $\lim_p g_p = 1$, with the following properties: for each $f \in \mathcal{C}'$ we have $H_k f \in \mathcal{C}'$ and there exist a function $K_k f \in \mathcal{C}$, a nonnegative function $L_k' f \in \mathcal{C}$, and positive numbers $\varepsilon_{n,p}(f)$ with $\lim_n \varepsilon_{n,p}(f) = 0$ for each p, such that

$$\rho_k^n f(x) := |\delta_n H_k^n f(x) - \delta_n H_k f(x) - K_k f(x)| \leq \begin{cases} L_k' f(x) & \text{for all } x \\ \varepsilon_{n,p}(f) L_k' f(x) & \text{if } \ g_p(x) \geq \frac{1}{2}. \end{cases} \tag{4.1}$$

Theorem 4.1 *Assume (H2). Let $f_1, \ldots, f_m$ be functions in $\mathcal{C}'$. Then the sequence of variables $(U_k^n f_i)_{1 \leq k \leq N, 1 \leq i \leq m}$ converges in law to a variable $(U_{k,i})_{1 \leq k \leq N, 1 \leq j \leq m}$ which is Gaussian with mean $M(k,i)$ and covariance $\Gamma(k,i;k',i')$ given by*

$$M(k,i) \ = \ \sum_{s=1}^{k} \frac{1}{\beta(s,k)} \pi_{s-1} K_s H_{s,k}(f_i - \pi_k f_i), \tag{4.2}$$

$$\Gamma(k,i;k',i') \ = \ \sum_{s=1}^{k \wedge k'} \frac{1}{\beta(s,k)\beta(s,k')} \pi_{s-1} H_s' \{(H_{s,k}(f_i - \pi_k f_i))(H_{s,k'}(f_{i'} - \pi_{k'} f_{i'}))\}. \tag{4.3}$$

Observe that (H2) implies that $H_{s,k} f \in \mathcal{C}'$ if $f \in \mathcal{C}'$, hence the right side of (4.2) is well defined.

Proof. 1) In a first step, we prove that if $f \in \mathcal{C}'$ and $1 \leq k \leq N$, then

$$\Psi_{k-1}^n \rho_k^n f \ \to^P \ 0. \tag{4.4}$$

To see this, we first observe that, by (4.1),

$$\Psi^n_{k-1}\rho^n_k f \;\leq\; Y^n_p \;:=\; \varepsilon_{n,p}(f)\Psi^n_{k-1}L'_k f + 2\Psi^n_{k-1}((1-g_p)L'_k f)$$

for every p. Moreover $(1-g_p)L'_k k \in \mathcal{C}$, hence by Theorem 3.1 we have that $2\Psi^n_{k-1}((1-g_p)L'_k f) \to^P \delta_{k,p} := 2\pi_{k-1}((1-g_p)L'_k f)$, and also $\Psi^n_{k-1}L'_k f \to^P \pi_{k-1}L'_k f$ as $n \to \infty$. Therefore $Y^n_p \to^P \delta_{k,p}$ because $\lim_n \varepsilon_{n,p}(f) = 0$ for every p. Therefore if $\varepsilon > 0$, we get

$$\limsup_n P(\Psi^n_{k-1}\rho^n_k f > \varepsilon) \;\leq\; \limsup_n P(Y^n_p > \varepsilon) \;\leq\; P(\delta_{k,p} > \varepsilon). \tag{4.5}$$

But since g_p increases to 1 and $L'_k f$ is π_{k-1}-integrable, the dominated convergence theorem yields $\lim_p \delta_{k,p} = 0$ everywhere. Hence the left side of (4.5) is 0, and we have (4.4).

2) In our second step, we take an arbitrary function $f \in \mathcal{C}'$. We set

$$a_k f \;=\; \pi_{k-1}H_k f, \qquad a'_k f \;=\; \pi_{k-1}K_k f.$$

We have (3.8), hence (2.5) and (2.11) yield that on the set $\{W^n_k 1 > 0\}$, one has

$$\Psi^n_k f - \pi_k f \;=\; \frac{\delta_n \Psi^n_{k-1}H^n_k f + \hat{W}^n_k f}{\delta_n \Psi^n_{k-1}H^n_k 1 + \hat{W}^n_k 1} - \frac{a_k f}{a_k 1}.$$

Then

$$U^n_k f \;=\; \frac{\alpha^n_k f + \beta^n_k f}{\gamma^n_k} \qquad \text{on the set } \{W^n_k 1 > 0\}, \tag{4.6}$$

where

$$\begin{aligned}
\alpha^n_k f \;&=\; (a_k 1)\hat{W}^n_k f - (a_k f)\hat{W}^n_k 1,\\
\beta^n_k f \;&=\; \delta_n\left((\Psi^n_{k-1}H^n_k f)(a_k 1) - (\Psi^n_{k-1}H^n_k 1)(a_k f)\right),\\
\gamma^n_k \;&=\; (a_k 1)\left(\Psi^n_{k-1}H^n_k 1 + \frac{1}{\delta_n}\hat{W}^n_k 1\right).
\end{aligned}$$

As in the proof of Theorem 3.1 (see after (3.12)) we get that $\Psi^n_{k-1}H^n_k g \to^P a_k g$ for every $g \in \mathcal{C}$, while the sequence $(\hat{W}^n_k g)$ is tight by the same theorem. Hence we obtain

$$\gamma^n_k \;\to^P\; (a_k 1)^2. \tag{4.7}$$

Next, we set

$$\begin{aligned}
\beta'^n_k f \;&=\; \delta_n\left((\Psi^n_{k-1}H_k f)(a_k 1) - (\Psi^n_{k-1}H_k 1)(a_k f)\right),\\
\beta''^n_k f \;&=\; (\Psi^n_{k-1}K_k f)(a_k 1) - (\Psi^n_{k-1}K_k 1)(a_k f)).
\end{aligned}$$

It readily follows from (4.4) (recall that the constants are in $\mathcal{C}'$) that

$$\beta^n_k f - \beta'^n_k f - \beta''^n_k f \;\to^P\; 0. \tag{4.8}$$

Finally, $W^n_k 1 \to^P a_k 1 > 0$ and $\beta''^n_k f \to^P (a'_k f)(a_k 1) - (a'_k 1)(a_k f) = (a_k 1)(a'_k(f - \pi_k f))$. Then by (4.6), (4.7) and (4.8) and by the fact that $P(W^n_k 1 > 0) \to 1$, we get

$$U^n_k f - \frac{\alpha^n_k f + \beta'^n_k f}{(a_k 1)^2} \;\to^P\; \frac{a'_k(f - \pi_k f)}{a_k 1}.$$

3) Now we come to our claim. Using now a vector notation with g having the components $f_1,\dots,f_r$ (as in the proof of Theorem 3.1), we can rewrite what precedes as such:

$$U_k^n g-\frac{\hat W_k^n(ga_k1-a_kg)+U_{k-1}^n((a_k1)(H_kg)-(a_kg)(H_k1))}{(a_k1)^2}\ \to^P\ \frac{a'_k(g-\pi_kg)}{a_k1},$$

which is also

$$(a_k1)U_k^n g-\hat W_k^n(g-\pi_kg)-U_{k-1}^n(H_kg-(\pi_kg)(H_k1))\ \to^P\ a'_k(g-\pi_kg). \tag{4.9}$$

Now set $\theta_{k,k}g=g-\pi_kg$, and $\theta_{s,k}g=H_{s,k}\theta_{k,k}g$ for $1\le s<k$. By (2.8) we have $\pi_s\theta_{s,k}g=0$, and by (2.6) we have

$$\theta_{s,k}g\ =\ H_s\theta_{s,k}g. \tag{4.10}$$

Notice that (H2) implies that the components of $\theta_{s,k}g$ are in $\mathcal{C}'$. Then (4.10) and (4.9) applied to $\theta_{s,k}g$ instead of g and s instead of k, yield for $1\le s\le k$:

$$(a_s1)U_s^n\theta_{s,k}g-\hat W_s^n(\theta_{s,k}g)-U_{s-1}^n(\theta_{s-1,k}g)\ \to^P\ a'_s(\theta_{s,k}g).$$

This, plus the fact that $U_0^nh=0$ for all h, shows that (recalling (2.7) and denoting by M_k the vector with components $(M(k,i):1\le i\le r)$ in (4.2)):

$$U_k^n g-\sum_{s=1}^k\frac{1}{\beta(s,k)}\hat W_s^n(\theta_{s,k}g)\ \to^P\ M_k.$$

In other words, if $V_k^ng=U_k^ng-M_k$, the sequence $(V_k^nf_i)_{1\le k\le N,1\le i\le m}$ has the same asymptotic behaviour than the sequence $(\sum_{s=1}^k\frac{1}{\beta(s,k)}\hat W_s^n(\theta_{s,k}f_i))_{1\le k\le N,1\le i\le m}$. It remains to apply Theorem 3.1: the above sequence converges in law to a centered Gaussian vector $(W(k,i):1\le k\le N,1\le i\le m)$, whose covariance is given by

$$E(W(k,i)W(k',i'))\ =\ \sum_{s=1}^{k\wedge k'}\frac{1}{\beta(s,k)\beta(s,k')}L'_s(k,i;k',i'),$$

where L'_s above is $L(k,i;k',i')$ in (3.7), with $\theta_{s,k}f_i$ and $\theta_{s,k'}f_{i'}$ instead of f_i and $f_{i'}$. Since $\theta_{s,k}g=H_{s,k}(g-\pi_kg)$ and $\pi_{s-1}H_s\theta_{s,k}g=0$, we deduce that the covariance is indeed $\Gamma(k,i;k',i')$ as given by (4.3), and the proof is finished. $\Box$

5 Case A

1) Recall that in case A the state process is the solution of Equation (1.1), and the observations have the form (1.2). For convenience, we take $Y_0=0$. We will assume either (A0) and (A2), or (A1) and (A2) below:

(A0) In (1.1) the functions a and b are uniformly Lipschitz.

(A1) In (1.1) the functions a and b are 2 times differentiable with bounded partial derivatives of orders 1 and 2, and the matrix $(bb^*)(x)$ is uniformly non-degenerate.

(A2) In (1.2), for any $x \in I\!R^d$ the variable $Y = h(x, \varepsilon_1)$ admits a density $y \mapsto g(x, y)$ and the function g is bounded and *explicitly known*.

We are in the setting of Section 2 with Q being the transition kernel of the Markov chain $(X_n)_{n\geq 0}$ and with G in (2.1) given by

$$G(x, y; x', y') \;=\; g(x', y'). \tag{5.1}$$

Let us also recall that, if necessary, we can simulate the processes X via an Euler approximation scheme. If the stepsize is $1/m$, we obtain the processes X^m. This process is not homogeneous Markov, but the chain $(X^m_n)_{n\in I\!N}$ is homogeneous Markov with a transition kernel which we denote by $Q^{(m)}$. (A1) implies that the kernels Q and $Q^{(m)}$ admit densities, denoted by q and $q^{(m)}$. According to Bally and Talay [2], we have with constants C, C' depending only on the coefficients a and b:

$$q(x, x') + q^{(m)}(x, x') \;\leq\; Ce^{-C'|x-x'|^2}, \tag{5.2}$$

$$\left.\begin{array}{rcl} q^{(m)}(x, x') - q(x, x') &=& \frac{\nu(x,x')}{m} + \frac{R_m(x,x')}{m^2}, \\ \text{with} \quad |\nu(x, x')| + |R_m(x, x')| &\leq& Ce^{-C'|x-x'|^2}, \end{array}\right\} \tag{5.3}$$

2) Next, we specialize the general Monte-Carlo procedure described above to our case. The reader will check that it is exactly the procedure described in [3]. The auxiliary set F_n will be $F_n = \{0\}$, and this amounts to not introducing the auxiliary variables Z^i_k at all, and to consider that $R_{n,k}$ is a transition probability from $(E, \mathcal{E})$ into itself. Then we subdivise the case into two subcases:

Subcase A-1: We can simulate exactly a random variable with law $Q(x, .)$, for any x. Then we take $R_{n,k} = Q$ for all n, k.

Subcase A-2: We have to resort to Euler scheme for simulation. Then we take m_n to be the smallest integer $\geq \sqrt{n}$, and we take $R_{n,k} = Q^{(m_n)}$ for all n, k.

It remains to specify the function $\psi_{n,k}$ in (2.9), and for this we take (recall that $F_n = \{0\}$ here, and $y_1, \ldots, y_k$ denote the observations):

$$\psi_{n,k}(x, 0) \;=\; g(x, y_k). \tag{5.4}$$

3) Next, we show that Hypotheses (H1) and (H2) are satisfied for suitable classes $\mathcal{C}$ and $\mathcal{C}'$. We introduce the following classes of Borel functions from $I\!R^d$ into $I\!R$ (α denotes a nonnegative real):

$$\left.\begin{array}{lll}\mathcal{D} & = & \{f \text{ Borel and bounded}\} \\ \mathcal{D}_{\text{pol}} & = & \{f \text{ Borel}, \ |f(x)| \leq C(1+|x|^\alpha) \ \text{ for some constants } C, \alpha\} \\ \mathcal{D}_{\text{exp}} & = & \{f \text{ Borel}, \ |f(x)| \leq Ce^{\alpha|x|} \ \text{ for some constants } C, \alpha\}.\end{array}\right\} \tag{5.5}$$

Proposition 5.1 *In Subcase A-1 we assume (A0) and (A2) and we let $\mathcal{C}$ be either $\mathcal{D}$ or $\mathcal{D}_{pol}$. In Subcase A-2 we assume (A1) and (A2) and we let $\mathcal{C}$ be any one of the classes $\mathcal{D}$, $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. As soon as the initial probability μ integrates all functions in $\mathcal{C}$* (which is always the case when $\mathcal{C}=\mathcal{D}$), *we have (H1) and (H2) with $\mathcal{C}$ and $\mathcal{C}'=\mathcal{C}$ and $\delta_n=\sqrt{n}$, and with H'_kf and K_kf given by*

$$H'_kf(x) \ = \ \int Q(x,dx')f(x')g(x',y_k)^2. \tag{5.6}$$

$$K_kf(x) \ = \ \begin{cases} 0 & \text{in Subcase A-1} \\ \int \nu(x,x')g(x',y_k)f(x')dx' & \text{in Subcase A-2.}\end{cases} \tag{5.7}$$

Observe that, in view of (5.2) and (5.3), the functions H'_kf and K_kf are defined for all $f\in\mathcal{C}$.

Proof. That $\mathcal{C}$ is an algebra containing the constants is obvious. (H1-a) holds by our hypothesis on μ. Taking $\delta_n=\sqrt{n}$, (H1-c) is obvious, and (H1-b) holds with any $r>2$ because g is bounded (recall (5.4)).

Consider first Subcase A-1. Recalling (2.3), (3.1), (5.6), (4.1) and (5.7), we have $H^n_k=H_k$ and $H'^n_k=H'_k$ and $\rho^n_kf=0$. Further, (A0) and the boundedness of g in (A2) and well known results on stochastic differential equation with Lipschitz coefficients yield that H_kf and H'_kf belong to $\mathcal{C}$ as soon as $f\in\mathcal{C}$: so (H1-d) holds with $L_kf=0$ and $\delta'_n=0$, and (H2) holds with $g_p=1$ and $\varepsilon_{n,p}(f)=0$ and $L'_kf=0$, hence the result.

Consider now Subcase A-2. Using (5.2) and (5.3), one readily gets that for any $f\in\mathcal{C}$ the functions H_kf, H'_kf and K_kf are in $\mathcal{C}$ as well. We also have

$$H^n_kf(x)-H_kf(x) \ = \ \int (q^{(m_n)}(x,x')-q(x,x')f(x')g(x',y_k)dx'.$$

Then using the fact that $|g|\leq C$ for some constant C, that m_n is equivalent to $\sqrt{n}$, and (5.3), we readily get the first part of (3.4) with $\delta'_n=\frac{1}{\sqrt{n}}$, and the second part is proved similarly, as well as (4.1) with $g_p=1$ and $\varepsilon_{n,p}(f)=\frac{1}{\sqrt{n}}$, hence (H1) and (H2) hold. $\square$

4) We can now apply Theorems 3.1 and 4.1. In particular in Subcase A-1, the mean value $M(k,i)$ in (4.2) is 0. Then we get the following corollary (in which H'_k and K_k are given by (5.6) and (5.7), and $H_kf(x)=\int q(x,x')f(x')g(x',y_k)dx'$):

Corollary 5.2 *In Subcase A-1 we assume (A0) and (A2) and we let $\mathcal{C}$ be either $\mathcal{D}$ or $\mathcal{D}_{pol}$. In Subcase A-2 we assume (A1) and (A2) and we let $\mathcal{C}$ be any one of the classes $\mathcal{D}$, $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. Assume also that the initial probability μ Then for any $f \in \mathcal{C}$ the variables $U_N^n f = \sqrt{n}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable, characterized as such:*

(i) In Subcase A-1 it is centered with variance

$$\Gamma_N f \;=\; \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H'_s\{(H_{s,N}(f - \pi_N f))^2\}. \tag{5.8}$$

(ii) In Subcase A-2 it has the same variance as above, and its mean is

$$M_N f \;=\; \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)). \tag{5.9}$$

6 Case B

1) In Case B, the state process X still satisfies (1.1), and the observations are given by (1.3). The results are more difficult to prove than in the other two cases, and they depend on the set of assumptions. We give below different sets of assumptions:

(B0): The functions a, b, h are bounded.

(B1): The functions a, b, h are Lipschitz, and the matrix σ is invertible.

(B2): The functions a, b, h are differentiable with bounded and Lipschitz derivatives, and the matrix σ is invertible.

We will assume at least (B1), which implies in particular that the two sequences $(X_i)_{i\in\mathbb{N}}$ and $(X_i, Y_i)_{i\in\mathbb{N}}$ are Markov chains and that (2.1) holds with some function G, which we presently compute.

We denote in fact by $X(x)$ the (unique strong) solution of (1.1) starting at x. We also set

$$Z(x) \;=\; \int_0^1 h(X(x)_t)dt. \tag{6.1}$$

Then if instead of $Y_0 = 0$ we start with $Y_0 = y$, conditionally on $(X_s)_{s\in[0,1]}$ the law of $Y_1 - y - Z(x)$ is $\mathcal{N}(0, \sigma\sigma^*)$. So if ϕ denotes the density of $\mathcal{N}(0, \sigma\sigma^*)$, the law of Y_1 conditionally on $(X_s)_{s\in[0,1]}$ has the density $y' \mapsto \phi(y' - y - Z(x))$. Therefore if $K(x, x'; dx'')$ denotes a regular version of the conditional distribution of $Z(x)$, knowing $X(x)_1 = x'$, we get

$$G(x, y; x', y') \;=\; \int K(x, x'; dx'')\phi(y' - y - x'').$$

It follows in particular that (remember (2.3))

$$H_k f(x) \;=\; E(f(X(x)_1)\phi(y_k - y_{k-1} - Z(x))). \tag{6.2}$$

The transition semi-group of the process X is denoted by (Q_t), and we set $S_n = Q_{1/n}$ for integers $n \geq 1$: so with the general notation of this paper, we have $Q = S_n^n$. We can also simulate the processes X via an Euler approximation scheme. If the stepsize is $1/n$, we obtain processes $X(x)^n$, where x denotes the initial value, and the discrete-time process $(X(x)^n_{i/n})_{i\in I\!N}$ is Markov, with a transition probability denoted by $S_n^{(n)}$.

2) Next, we specialize our Monte-Carlo procedure to the present case. Again it will be the same as in [3]. For a fixed n, the auxiliary set F_n will be $F_n = (I\!R^d)^n$ with its Borel σ-field $\mathcal{F}$. Here again we subdivise the case into two subcases:

Subcase B-1: We can simulate exactly a random variable with law $S_n(x,.)$. Then we take

$$R_{n,k}(x; dx', dx_1, \ldots, dx_n) = S_n(x, dx_1) \ldots S_n(x_{n-1}, dx_n)\varepsilon_{x_n}(dx'). \tag{6.3}$$

Subcase B-2: We have to resort to Euler scheme for simulation. Then we take

$$R_{n,k}(x; dx', dx_1, \ldots, dx_n) = S_n^{(n)}(x, dx_1) \ldots S_n^{(n)}(x_{n-1}, dx_n)\varepsilon_{x_n}(dx'). \tag{6.4}$$

It remains to specify the function $\psi_{n,k}$ in (2.9). For this, we denote by ϕ the density of the normal law $\mathcal{N}(0, \sigma\sigma^*)$, and we set (recall that $y_0 = 0, y_1, \ldots, y_k$ denote the observations):

$$\psi_{n,k}(x; x_1, \ldots, x_n) \;=\; \phi\left(y_k - y_{k-1} - \frac{1}{n}\sum_{i=1}^n h(x_i)\right). \tag{6.5}$$

3) Here we give a number of more or less well known results about Equation (1.1) under, at least, (B1). We use matrix notation, and ∇ denotes the gradient. Recall that $X(x)$ is the solution of (1.1) started at x. Let $\theta_n(t) = i/n$ if $i/n < t \leq (i+1)/n$. The Euler approximation with stepsize $1/n$ is the solution of the equation

$$dX(x)^n_t \;=\; a(X(x)^n_{\theta_n(t)})dt + b(X(x)^n_{\theta_n(t)})dW_t, \qquad X(x)^n_0 = x, \tag{6.6}$$

and the error process is

$$V(x)^n \;=\; \sqrt{n}(X(x)^n - X(x)). \tag{6.7}$$

We also set

$$Z(x)^n \;=\; \frac{1}{n}\sum_{i=1}^n h(X(x)_{\frac{i}{n}}, \qquad Y(x)^n \;=\; \frac{1}{n}\sum_{i=1}^n h(X(x)^n_{\frac{i}{n}}. \tag{6.8}$$

Recall that (B1) holds. We set $c = 0$ if (B0) also holds, and $c = 1$ otherwise. Then well known estimates are (the constants $K_.$ below depend on a, b, and the other parameters are explicitely stated in the expression of the constant; p denotes a strictly positive integer; as usual, for any process Y we set $Y_t^* = \sup_{s\leq t}|Y_s|$):

$$E(X(x)_t^{*,p}) + E((X(x)^n)_t^{*,p}) \;\leq\; K_{p,t}(1 + |x|^p), \tag{6.9}$$

$$t \leq 1 \quad \Rightarrow \quad E(|X(x) - x|_t^p) + E(|X(x)^n - x|_t^p) \;\leq\; K_p t^{p/2}(1 + c|x|^p), \tag{6.10}$$

$$E((X(x)-X(y))_t^{*,p})+E((X(x)^n-X(y)^n)_t^{*,p}) \;\leq\; K_{p,t}|x-y|^p, \tag{6.11}$$
$$E((V(x)^n)_t^{*,p}) \;\leq\; K_{p,t}(1+c|x|^p). \tag{6.12}$$

We readily deduce that:

$$E(|Z(x)-Z(y)|^p)+E(|Y(x)^n-Y(y)^n|^p) \;\leq\; K_p|x-y|^p, \tag{6.13}$$

$$E(|Y(x)^n-Z(x)^n|^p)+E(|Z(x)^n-Z(x)|^p)+E(|Y(x)^n-Z(x)|^p) \leq K_p\frac{1+c|x|^p}{n^{p/2}}. \tag{6.14}$$

Now we assume (B2) with the same convention as above for c. We can differentiate $x \mapsto X(x)$ in the $\mathbb{L}^p$ sense, and the differential $\hat{X}(x)$ is a $d\times d$-dimensional process, solution to the linear equation

$$d\hat{X}(x)_t \;=\; \nabla a(X(x)_t)\hat{X}_t dt+\nabla b(X(x)_t)\hat{X}_t dW_t, \qquad \hat{X}(x)_0=I, \tag{6.15}$$

whose Euler approximation is

$$d\hat{X}(x)_t^n \;=\; \nabla a(X(x)^n_{\theta_n(t)})\hat{X}^n_{\theta_n(t)}dt+\nabla b(X(x)^n_{\theta_n(t)})\hat{X}^n_{\theta_n(t)}dW_t, \qquad \hat{X}(x)^n_0=I. \tag{6.16}$$

Then we have the estimates:

$$E(\hat{X}(x)_t^{*,p})+E((\hat{X}(x)^n)_t^{*,p}) \;\leq\; K_{p,t}, \tag{6.17}$$
$$E((\hat{X}(x)-\hat{X}(y))_t^{*,p})+E((\hat{X}(x)^n-\hat{X}(y)^n)_t^{*,p}) \;\leq\; K_{p,t}|x-y|^p, \tag{6.18}$$
$$E((\hat{X}^n(x)-\hat{X}(x))_t^{*,p}) \;\leq\; K_{p,t}\frac{1+c|x|^p}{n^{p/2}}. \tag{6.19}$$

Finally, under (B2) again, the sequence $V(x)^n$ converges stably in law to a limiting process $V(x)$ (see Kurtz and Protter [4]), which is the solution of the linear equation

$$\left.\begin{array}{l} dV(x)_t=\nabla a(X(x)_t)V(x)_t dt+\nabla b(X(x)_t)V(x)_t dW_t-\frac{1}{\sqrt{2}}\nabla b(X(x)_t)b(X(x)_t)d\tilde{W}_t, \\ V(x)_0=0, \end{array}\right\} \tag{6.20}$$

where $\tilde{W}$ is a $d\times d$-dimensional Wiener process, independent of W. Then we have

$$E(V(x)_t^{*,p}) \;\leq\; K_{p,t}(1+c|x|^p), \tag{6.21}$$
$$E((V(x)-V(y))_t^{*,p}) \;\leq\; K_{p,t}|x-y|^p(1+c|x|^p). \tag{6.22}$$

4) Now we can show (H1) for suitable classes $\mathcal{C}$. We denote by L_b the class of all bounded Lipschitz functions on $\mathbb{R}^d$, and by L_{pol} the class of all functions f satisfying $|f(x)-f(y)|\leq C|x-y|(1+|x|^\alpha+|y|^\alpha)$ for some constants C, α.

Proposition 6.1 *(H1) is satisfied with $\delta_n=\sqrt{n}$ in the following two cases:*

a) We have (B0) and (B1), and with $\mathcal{C}=L_b$.

b) We have (B1) and $\int|x|^p\mu(dx)<\infty$ for all $p\in(0,\infty)$, and with $\mathcal{C}=L_{\text{pol}}$.

In these two cases, H'_kf is defined as such for $f\in\mathcal{C}$:

$$H'_kf(x) \;=\; E(f(X(x)_1)\phi(y_k-y_{k-1}-Z(x))^2). \tag{6.23}$$

Proof. In both cases $\mathcal{C}$ is an algebra containing the constants, and (H1-a) holds. Since $||\psi_{n,k}||$ does not depend on (n,k) we have (H1-b) for all $r>2$, and (H1-c) is obvious. In view of (6.2) and (6.23), we deduce from (6.11) and (6.13), and also (6.9) in Case (b), that $H_k f\in\mathcal{C}$ and $H'_k f\in\mathcal{C}$ if $f\in\mathcal{C}$ in both cases.

It remains to prove (3.4). To unify the presentation, we write $X(x)^n=X(x)$ in Subcase B-1, where we do not need Euler approximation (so (6.9) and (6.12) hold trivially with this convention). With this convention, and with $\phi_k(y)=\phi(y_k-y_{k-1}-y)$, we have

$$H_k^n f(x) \;=\; E(f(X(x)_1^n)\phi_k(Y(x)^n)). \tag{6.24}$$

Therefore we can write

$$\begin{aligned} H_k^n f(x)-H_kf(x) \;&=\; E\left((f(X(x)_1^n)-f(X(x)_1))\phi_k(Y(x)^n)\right)\\ &\quad +E\left(f(X(x)_1)\nabla\phi_k(\tilde Y(x)^n)(Y(x)^n-Z(x))\right), \end{aligned}$$

where $\tilde Y(x)^n$ lies in between $Y(x)^n$ and $Z(x)$. Now, $\nabla\phi_k$ is bounded, so in case (a) we readily deduce from (6.12) and (6.14) that $|H_k^n f-H_kf|\le\frac{C}{\sqrt n}$ for some constant C (recall that $c=0$ in this case), and the first half of (3.4) holds with $L_kf=C$; the second half is proved similarly (just replace ϕ_k by ϕ_k^2).

Now suppose that we are in Case (b). Suppose also that $f\in\mathcal{C}$ satisfies for some constants C and r: $|f(x)-f(y)|\le C|x-y|(1+|x|^r+|y|^r)$. Now we have $c=1$, and (6.9), (6.12) and (6.14) and Cauchy-Schwarz inequality yield

$$|H_k^n f(x)-H_kf(x)| \;\le\; K\frac{1+|x|^{r+2}}{\sqrt n}$$

for some constant K: hence the first half of (3.4) holds with $L_kf(x)=K(1+|x|^{r+2})$, which belongs to $\mathcal{C}$. The proof of the second half is similar. $\square$

It follows from this that Theorem 3.1 holds in case B, under the assumptions of the previous proposition, and with $\mathcal{C}$ described above. Note however that if (B1) holds, it does not hold for the class $\mathcal{C}=L_b$ if μ does not integrates the functions $x\mapsto|x|^p$ for all $p<\infty$.

5) Now we turn to (H2). We have the

Proposition 6.2 *(H2) is satisfied with $\delta_n=\sqrt n$ in the following two cases:*

a) We have (B0) and (B2), and with $\mathcal{C}=L_b$ and $\mathcal{C}'$ being the set of all $f\in\mathcal{C}$ which are differentiable, with all its partial derivatives in $\mathcal{C}$.

b) We have (B2) and $\int|x|^p\mu(dx)<\infty$ for all $p\in(0,\infty)$, and with $\mathcal{C}=L_{\rm pol}$ and $\mathcal{C}'$ being the set of all $f\in\mathcal{C}$ which are differentiable, with all its partial derivatives in $\mathcal{C}$.

In these two cases, K_kf is defined as such for $f\in\mathcal{C}$:

$$K_kf(x)=\begin{cases} 0 & \text{in Subcase B-1}\\ E(\nabla f(X(x)_1)V(x)_1\phi(y_k-y_{k-1}-Z(x)) & \\ +f(X(x)_1)\nabla\phi(y_k-y_{k-1}-Z(x))\int_0^1\nabla h(X(x)_t)V(x)_tdt) & \text{in Subcase B-2.}\end{cases} \tag{6.25}$$

We begin with a lemma:

Lemma 6.3 *Under the assumptions of Proposition 6.2, for any $f \in \mathcal{C}'$ we have $H_k f \in \mathcal{C}'$ and $K_k f \in \mathcal{C}$, and furthermore the function $\rho_k^n f$ of (4.1) converges pointwise to* 0.

Proof. 1) Let $f \in \mathcal{C}'$ (here, $\mathcal{C}$ and $\mathcal{C}'$ are defined as in our claim). Since we have (B2), $\hat{X}(x)_t$ is the derivative of $x \mapsto X(x)_t$ in all $I\!\!L^p$, we have that $H_k f$ is differentiable, with gradient given by (recall that $\phi_k(y) = \phi(y_k - y_{k-1} - y)$):

$$\nabla H_k f(x) \;=\; E(\nabla f(X(x)_1)\hat{X}(x)_1\phi_k(Z(x)) + f(X(x)_1)\nabla\phi_k(Z(x))\int_0^1 \nabla h(X(x)_t)\hat{X}(x)_t dt).$$

Then if we use (6.11), (6.13), (6.17), (6.18) and $c = 0$ in case (a) and (6.9) in case (b), plus the fact that ϕ_k and its first derivatives are bounded, we easily get that $H_k f \in \mathcal{C}'$. Similarly, using in addition (6.21) and (6.22) in Subcase B-2, we get $K_k f \in \mathcal{C}$ (this is of course trivial in Subcase B-1).

2) Now we prove that $\rho_k^n f \to 0$ pointwise, which amounts to prove that $\sqrt{n}(H_k^n f(x) - H_k f(x)) \to K_k f(x)$ for all x. As in the proof of Proposition 6.1, in Subcase B-1 we set $X(x)^n = X(x)$, hence $V(x)^n = 0$. In view of (6.2) and (6.24), we have $\sqrt{n}(H_k^n f(x) - H_k f(x)) = E(A_n + B_n)$, where

$$A_n \;=\; \nabla f(\tilde{X}(x)_1^n)V(x)_1^n\phi_k(Y(x)^n) + f(X(x)_1)\nabla\phi_k(\tilde{Y}(x)^n)\frac{1}{n}\sum_{i=1}^n \nabla h(\tilde{X}(x)^n_{\frac{i}{n}})V(x)^n_{\frac{i}{n}},$$

$$B_n \;=\; f(X(x)_1)\nabla\phi_k(\tilde{Y}(x)^n)B'_n, \quad B'_n \;=\; \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n h(X(x)^n_{\frac{i}{n}}) - \int_0^1 h(X(x)_t)dt\right),$$

where $\tilde{X}(x)_t^n$ is in between $X(x)_t$ and $X(x)_t^n$ and $\tilde{Y}(x)^n$ is in between $Z(x)$ and $Y(x)^n$.

In view of (6.9), (6.12), (B-2) and of the properties of f and the boundedness of ϕ_k and $\nabla\phi_k$, it is clear that the sequences (A_n) and (B_n) are bounded in $I\!\!L^2$. Hence it suffices to prove that $B_n \to^P 0$ and $A_n \to^{\mathcal{L}} A$, where

$$A \;=\; \nabla f(X(x)_1)V(x)_1\phi_k(Z(x)) + f(X(x)_1)\nabla\phi_k(Z(x))\int_0^1 \nabla h(X(x)_t)V(x)_t dt$$

in Subcase B-2, and $A = 0$ in Subcase B-1 (above, $V(x)$ is defined by (6.20) on an extension of the original space).

3) In Subcase B-1, we have $A_n = 0$ and $A = 0$. In Subcase B-2, recall that the sequence $(V(x)^n)$ converges stably in law to $V(x)$, as processes. Furthermore $X(x)^n \to^P X(x)$ uniformly on $[0,1]$, and $Y(x)^n \to^P Z(x)$ by (6.12) and (6.14), while

$$\frac{1}{n}\sum_{i=1}^n \nabla h(X(x)_{\frac{i}{n}}V(x)^n_{\frac{i}{n}} - \int_0^1 \nabla h(X(x)_t^n)V(x)_t^n dt \;\to^P\; 0$$

by (6.10) and its analogue for the processes $V(x)^n$, and since ∇h is Lipschitz. Hence it becomes clear that $A_n \to^{\mathcal{L}} A$.

4) It remains to prove that $B_n \to^P 0$, and for this it is enough to have $B'_n \to^P 0$. We have $B'_n = \sum_{i=1}^n \beta_i^n$, where

$$\beta_i^n \;=\; \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} (h(X(x)_{\frac{i}{n}}) - h(X(x)_t))dt.$$

We can write $\beta_i^n = \beta_i'^n + \beta_i''^n$, where

$$\beta_i'^n \;=\; \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} \nabla h(X(x)_t)(X(x)_{\frac{i}{n}} - X(x)_t)dt,$$

$$\beta_i''^n \;=\; \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} (\nabla h(\tilde{X}_t^n) - \nabla h(X(x)_t))(X(x)_{\frac{i}{n}} - X(x)_t)dt,$$

and where $\tilde{X}_t^n$ is in between $X(x)_t$ and $X(x)_{i/n}$ when $\frac{i-1}{n} < t \leq \frac{i}{n}$.

First we have by repeated use of Cauchy-Schwarz inequality and (6.10):

$$E(|\beta_i''^n|^2) \;\leq\; \frac{C_x}{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} \left(E(|\nabla h(\tilde{X}_t^n) - \nabla h(X(x)_t)|^4)\right)^{1/2} dt$$

for some constant C_x, possibly depending on x. Hence

$$E(|\sum_{i=1}^n \beta_i''^n|^2) \;\leq\; n\sum_{i=1}^n E(|\beta_i''^n|^2) \;\leq\; C_x\int_0^1 \left(E(|\nabla h(\tilde{X}_t^n) - \nabla h(X(x)_t)|^4)\right)^{1/2} dt,$$

and since $\tilde{X}_t^n \to^P X(x)_t$ for all t and ∇h is bounded and continuous, we readily deduce that the above tends to 0: hence $\sum_{i=1}^n \beta_i''^n \to^P 0$.

Second, by using (1.1) we may write $\beta_i'^n = \gamma_i^n + \delta_i^n$, where

$$\gamma_i^n \;=\; \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} \nabla h(X(x)_t)dt\int_t^{\frac{i}{n}} a(X(x)_s)ds,$$

$$\delta_i^n = \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} \nabla h(X(x)_t)dt\int_t^{\frac{i}{n}} b(X(x)_s)dW_s = \sqrt{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\left(\int_{\frac{i-1}{n}}^{s} \nabla h(X(x)_t)dt\right)b(X(x)_s)dW_s.$$

Now we have $|\gamma_i^n| \leq C(1+X(x)_1^*)n^{-3/2}$, hence $\sum_{i=1}^n \gamma_i^n \to^P 0$ (use (6.9)). Finally, we have

$$E(|\delta_i^n|^2) \;\leq\; nE\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} |b(X(x)_s)|^2\left(\int_{i-1}^{s} |\nabla h(X(x)_t)|dt\right)^2 ds\right) \;\leq\; \frac{C_x}{n^2}$$

for another constant C_x (use again (6.9)). Since further $E(\delta_i^n|\mathcal{F}_{(i-1)/n}) = 0$, it follows that

$$E(|\sum_{i=1}^n \delta_i^n|^2) \;=\; \sum_{i=1}^n E(|\delta_i^n|^2) \;\leq\; \frac{C_x}{n} \;\to\; 0.$$

Putting all these together, we deduce that $B'_n \to^P 0$, and we are finished. $\square$

Proof of Proposition 6.2. In view of Proposition 6.1 and Lemma 6.3, it only remains to prove that (4.1) holds with suitable functions g_p, $L'_k f$ and constants $\varepsilon_{n,k}(f)$.

1) Take $f \in \mathcal{C}'$ and $k \in \{1, \ldots, N\}$, and set

$$M_k^n f(x,y) \;=\; \delta_n(H_k^n f(x) - H_k f(x) - H_k^n f(y) + H_k f(y)).$$

Recall also our convention that $X(x)^n = X(x)$ in Subcase B-1. In view of (6.2) and (6.24), we have $M_k^n f(x,y) = E(A_n + B_n + D_n + E_n)$, where

$$\begin{aligned}
A_n &= \sqrt{n}\phi_k(Y(x)^n)A'_n, \qquad A'_n = f(X(x)_1^n) - f(X(x)_1) - f(X(y)_1^n) + f(X(y)_1),\\
B_n &= \sqrt{n}f(X(x)_1)B'_n, \qquad B'_n = \phi_k(Y(x)^n) - \phi_k(Z(x)) - \phi_k(Y(y)^n) + \phi_k(Z(y)),\\
D_n &= \sqrt{n}(f(X(x)_1^n) - f(X(y)_1^n)(\phi_k(Y(x)^n) - \phi_k(Y(y)^n)),\\
E_n &= \sqrt{n}(f(X(x)_1) - f(X(y)_1)(\phi_k(Y(y)^n) - \phi_k(Z(y))).
\end{aligned}$$

By our definition of $\mathcal{C}'$, we have

$$|f(x) - f(y)| + |\nabla f(x) - \nabla f(y)| \;\leq\; C|x-y|(1 + |x|^\alpha + |y|^\alpha)$$

for some constants C, α, with furthermore $\alpha = 0$ in case (a). Observe that this implies also $|f(x)| + |\nabla f(x)| \leq C(1 + |x|^{\alpha+c})$ for another constant C (below, the constant C changes from line to line; recall that $c = 0$ in case (a), and $c = 1$ in case (b)).

By Cauchy-Schwarz inequality and (6.12) and (6.13), plus (6.9) in case (b), we easily get

$$E(|D_n|) \;\leq\; C|x-y|(1 + |y|^{\alpha+c}). \tag{6.26}$$

Similarly, in view of (6.14), we get

$$E(|E_n|) \;\leq\; C|x-y|(1 + |x|^{\alpha+2c} + |y|^{\alpha+2c}). \tag{6.27}$$

By Taylor's formula we have

$$\begin{aligned}
A'_n \;=\; \int_0^1 \Big[&\nabla f(X(y+u(x-y))_1^n)\hat{X}(y+u(x-y))_1^n\\
&-\nabla f(X(y+u(x-y))_1)\hat{X}(y+u(x-y))_1\Big]\,(x-y)du.
\end{aligned}$$

Then (6.12), (6.17) and (6.18), plus (6.9) in case (b), yield

$$E(|A_n|) \;\leq\; C|x-y|(1 + |x|^{\alpha+2c} + |y|^{\alpha+2c}). \tag{6.28}$$

Similarly, we can write

$$\begin{aligned}
B'_n \;=\; \int_0^1 \Bigg[&\nabla\phi_k(Y(y+u(x-y))_1^n)\frac{1}{n}\sum_{i=1}^n \nabla h(X(y+u(x-y))_{i/n}^n)\hat{X}(y+u(x-y))_{i/n}^n\\
&-\nabla\phi_k(Z(y+u(x-y)))\int_0^1 \nabla h(X(y+u(x-y))_t\hat{X}(y+u(x-y))_t dt\Bigg]\,(x-y)du.
\end{aligned}$$

Then (6.12), (6.14), (6.18), (6.19), and an estimate similar to (6.14) with h replaced by ∇h in (6.8), plus (6.9) in case (a), allow to deduce that

$$E(|B_n|) \;\leq\; C|x-y|(1 + |x|^{\alpha+2c} + |y|^{\alpha+2c}). \tag{6.29}$$

Putting (6.26), (6.27), (6.28) and (6.29) together yields

$$|M_k^n f(x,y)| \ \leq \ C|x-y|(1+|x|^{\alpha+2c}+|y|^{\alpha+2c}). \tag{6.30}$$

2) Now we check (4.1). First, let B_p be the closed ball of center 0 and radius p in $\mathbb{R}^d$. We can find a C^∞ function $g_p : \mathbb{R}^d \to [0,1]$ which equals 1 on B_p and 0 on B_{p+1}^c. The sequence (g_p) is increasing, has limit 1, and $g_p \in \mathcal{C}$. Next, let $f \in \mathcal{C}'$. We know by Proposition 6.1 that (3.4) holds with some $L_k f \in \mathcal{C}$, and $K_k f \in \mathcal{C}$ by Lemma 6.3, hence $L'_k f = 1 + L_k f + (K_k f)^2$ also belongs to $\mathcal{C}$ (which is an algebra) and the first part of (4.1) is met with this function $L'_k f$.

Finally, (6.30) and $K_k f \in \mathcal{C}$ again imply that the sequence of functions $(\rho_k^n f)_{n\geq 1}$ is uniformly equicontinuous on each compact set B_p. Since $\rho_k^n f \to 0$ pointwise, we even have that $\varepsilon_{n,p}(f) = \sup_{1\leq k\leq N} \sup_{x\in R_{p+1}} \rho_k^n f(x)$ converges to 0, while $\{g_p \geq 1/2\} \subset B_{p+1}$: so the second part of (4.1) is met with these sequences $\varepsilon_{n,p}$ and the same $L'_k f$ as above, and we are finished. $\square$

6) We can now apply Theorem 4.1. In particular in Subcase B-1, the mean value $M(k,i)$ in (4.2) is 0, and we have the following corollary (recall that H_k, H'_k and K_k are given by (6.2), (6.23) and (6.25) repectively):

Corollary 6.4 *Assume (B0) (resp. $\int |x|^p \mu(dx) < \infty$ for all $p < \infty$), and (B2). Let f be a differentiable function, belonging to L_b (resp. L_{pol}) as well as all its first derivatives. Then the variables $U_N^n f = \sqrt{n}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable, characterized as such:*

(i) In Subcase B-1 it is centered with variance

$$\Gamma_N f \ = \ \sum_{s=1}^N \frac{1}{\beta(s,N)^2} \pi_{s-1} H'_s\{(H_{s,N}(f-\pi_N f))^2\}. \tag{6.31}$$

(ii) In Subcase B-2 it has the same variance as above, and its mean is

$$M_N f \ = \ \sum_{s=1}^N \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f-\pi_N f)). \tag{6.32}$$

7 Case C

1) The state process X still satisfies (1.1), and the observations are given by (1.4). The assumptions are as follows:

(C1) In (1.1) and (1.4) the functions a, b, a' and b' are 3 times differentiable with bounded partial derivatives of orders 1, 2 and 3, and the matrices $(bb^*)(x)$ and $(b'b'^*)(x,y)$ are uniformly non-degenerate.

The two sequences $(X_i)_{i\in I\!N}$ and $(X_i, Y_i)_{i\in I\!N}$ are Markov chains, and the transitions Q and P admit densities denoted by q and p, and in (2.11) the function G is

$$G(x,y;x',y') \;=\; \frac{p(x,y;x',y')}{q(x,x')}. \tag{7.1}$$

Let us also recall that, if necessary, we can simulate the processes (X, Y) via an Euler approximation scheme. If the stepsize is $1/m$, we obtain processes $(X(x,m), Y(x,y,m))$ (where x and y denote the initial values), whose transition kernels at time 1 are $P^{(m)}$. (C1) implies that the kernel $Q^{(m)}$ admits a density denoted by $p^{(m)}$. According again to Bally and Talay [2], we have with constants C, C' depending only on the coefficients a, a', b and b':

$$|\nabla^j_{y'} p(x,y;x',y')| + |\nabla^j_{y'} p^{(m)}(x,y;x',y')| \;\le\; Ce^{-C'(|x-x'|^2+|y-y'|^2)}, \tag{7.2}$$

for $j = 0, 1, 2, 3$ (where $\nabla^0_{y'} p = p$ and $\nabla^j_{y'} p$ denotes the vector of all partial derivatives of p, of order j, w.r.t. the components of y') and

$$\left.\begin{array}{l} p^{(m)}(x,y;x',y') - p(x,y;x',y') \;=\; \frac{\nu(x,y;x',y')}{m} + \frac{R_m(x,y;x',y')}{m^2}, \\ \text{with} \quad |\nu(x,y;x',y')| + |R_m(x,y;x',y')| \;\le\; Ce^{-C'(|x-x'|^2+|y-y'|^2)}, \end{array}\right\} \tag{7.3}$$

and further the function ν is continuous.

2) Next, we specialize our general Monte-Carlo procedure to the present case, according again to [3]. However, we will take here a normalization which differs from the one in [3], allowing for a better rate of convergence (to the, mild, expense of having three derivatives in (C1) instead of two). The auxiliary set F_n will be $F = I\!R^q$ with its Borel σ-field $\mathcal{F}$. Then here again we subdivise the case into two subcases (recall that $y_0 = 0, y_1, \ldots, y_k$ denote the observations):

Subcase C-1: We can simulate exactly a random variable with law $P(x,y;.)$, for any (x,y). Then we take $R_{n,k}(x;dx',dz) = P(x,y_{k-1};dx',dz)$ for all n, k.

Subcase C-2: We have to resort to Euler scheme for simulation. Then we take m_n to be the smallest integer $\ge n^{2/(4+q)}$, and we take $R_{n,k}(x;dx',dz) = P^{(m_n)}(x,y_{k-1};dx',dz)$ for all n, k.

It remains to specify the function $\psi_{n,k}$ in (2.9), and for this we choose a Borel bounded function ψ from $I\!R^q$ into $[0,\infty)$ such that

$$\int \psi(y)dy \;=\; 1, \qquad \int y\psi(y)dy \;=\; 0, \qquad \int |y|^3\psi(y)dy \;<\; \infty. \tag{7.4}$$

(This is slightly stronger than (5.2) in [*]). Then we set

$$\psi_n(y) \;=\; n^{q/(4+q)}\psi(yn^{1/(4+q)}) \tag{7.5}$$

and

$$\psi_{n,k}(x,z) \;=\; \psi_n(z-y_k). \tag{7.6}$$

3) Next, we show that Hypotheses (H1) and (H2) are satisfied for suitable classes $\mathcal{C}$ and $\mathcal{C}'$. First we observe that here,

$$H_k f(x) \;=\; \int p(x, y_{k-1}; x', y_k) f(x') dx', \tag{7.7}$$

which is in $\mathcal{C}$ as soon as $f \in \mathcal{C}$ by (7.2), when $\mathcal{C}$ is any one of the classes $\mathcal{D}$, $\mathcal{D}_{\text{pol}}$ or $\mathcal{D}_{\text{exp}}$ (recall (5.5).

Proposition 7.1 *Assume (C1). Let $\mathcal{C}$ be any one of the classes $\mathcal{D}$, $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. As soon as the initial probability μ integrates all functions in $\mathcal{C}$* (which is always the case when $\mathcal{C} = \mathcal{D}$)*, we have (H1) and (H2) with $\mathcal{C}$ and $\mathcal{C}' = \mathcal{C}$ and $\delta_n = n^{\frac{2}{4+q}}$, and with $H'_k f$ and $K_k f$ are given by*

$$H'_k f \;=\; \delta H_k f, \qquad \textit{where } \delta \;=\; \int \psi(y)^2 dy, \tag{7.8}$$

$$K_k f(x) = \begin{cases} K'_k f(x) & \text{in Subcase C-1} \\ K'_k(f) + \int \nu(x, y_{k-1}; x', y_k) f(x') dx' & \text{in Subcase C-2.} \end{cases} \tag{7.9}$$

where

$$K'_k f(x) \;=\; \frac{1}{2} \sum_{j,\ell=1}^{q} \int z^j z^\ell \psi(z) dz \int \frac{\partial^2 p}{\partial y'^j \partial y'^\ell}(x, y_{k-1}; x', y_k) f(x') dx', \tag{7.10}$$

Observe that, in view of (7.2) and (7.3), the functions $H'_k f$ and $K_k f$ are defined for all $f \in \mathcal{C}$.

Proof. a) that $\mathcal{C}$ is an algebra containing the constants is obvious. (H1-a) holds by our hypothesis on μ. Take $\delta_n = n^{\frac{2}{4+q}}$. Then (H1-c) is obvious with $u = 0$. We have $||\Psi_{n,k}|| = C n^{\frac{q}{4+q}}$, hence (H1-b) holds for any $r > 2$. Using (7.2) and (7.3), one also readily gets that for any $f \in \mathcal{C}$ the functions $H_k f$, $H'_k f$ and $K_k f$ are in $\mathcal{C}$ as well.

b) Let us now consider Subcase C-1. We have by (3.1) and (7.6):

$$H^n_k f(x) \;=\; \int p(x, y_{k-1}; x', y) f(x') \psi_n(y - y_k) dx' dy.$$

Hence by (7.4) and (7.5):

$$\begin{aligned} H^n_k f(x) - H_k f(x) \;&=\; \int (p(x, y_{k-1}; x', y) - p(x, y_{k-1}; x', y_k)) f(x') \psi_n(y - y_k) dx' dy \\ &=\; \int (p(x, y_{k-1}; x', y_k + \frac{z}{n^{\frac{1}{4+q}}}) - p(x, y_{k-1}; x', y_k)) f(x') \psi(u) dx' dz. \end{aligned}$$

Then we use (7.2) and $\int y \psi(y) dy = 0$ and (7.4) and the boundedness of ψ, together with a third order Taylor expansion for the function $y' \mapsto p(x, y_{k-1}; x', y')$ around $y' = y_k$, to get that

$$|H^n_k f - H_k f - \frac{1}{n^{\frac{2}{4+q}}} K'_k f(x)| \;\leq\; \frac{1}{n^{\frac{3}{4+q}}} L'_k f$$

for some function $L'_k f \in \mathcal{C}$: this yields the first part of (3.4) with $L_k f = |K'_k f| + L'_k$ and $\delta'_n = 1/n^{\frac{2+q}{4+q}}$, and also (4.1) with $K_k f = K'_k f$ and $\varepsilon_{n,p}(f) = 1/n^{\frac{1}{4+q}}$ and $g_p = 1$. Hence it remains only to prove the second half of (3.4).

We may write

$$\begin{aligned} H'^n_k f(x) - \frac{n}{\delta_n^2} H_k f(x) = & \int (p(x, y_{k-1}; x', y) - p(x, y_{k-1}; x', y_k)) f(x') \psi_{n,k}(y)^2 dx' dy \\ = & \frac{n}{\delta_n^2} \int (p(x, y_{k-1}; x', y_k + \frac{z}{n^{\frac{1}{4+q}}}) - p(x, y_{k-1}; x', y_k)) f(x') \psi(z)^2 dx' dz, \end{aligned}$$

and the same arguments as above (it is in fact slightly simpler) lead to the second part of (3.4), with $\delta'_n = 1/n^{\frac{1}{4+q}}$.

c) Finally we consider the Subcase C-2. Now we have

$$H^n_k f(x) = \int p^{(m_n)}(x, y_{k-1}; x', y) f(x') \psi_n(y - y_k) dx' dy.$$

Then $H^n f - H_k f = \tilde{H}^n_k f + \hat{H}^n_k f$, where

$$\begin{aligned} \tilde{H}^n_k f(x) &= \int (p^{(m_n)}(x, y_{k-1}; x', y) - p(x, y_{k-1}; x', y_))) f(x') \psi_{n,k}(y) dx' dy, \\ \hat{H}^n_k f(x) &= \int (p(x, y_{k-1}; x', y) - p(x, y_{k-1}; x', y_k)) f(x') \psi_n(y - y_k) dx' dy. \end{aligned}$$

Since $m_n \sim \delta_n$, (7.3) and the continuity of ν readily give that

$$|\delta_n \tilde{H}^n_k f(x) - \int \nu(x, y_{k-1}; x', y_k) f(x') dx'| \leq \frac{1}{\delta_n} L''_k f(x)$$

for some $L''_k f \in \mathcal{C}$. Since $\hat{H}^n_k f$ is nothing else than the value of $H^n_k f - H_k f$ for Subcase C-1, we deduce readily from the above and from the study of Subcase C-1 that we have the first part of (3.4), as well as (4.1) with $K_k f$ given by (7.9).

Hence it remains only to prove the second half of (3.4). We may write

$$\begin{aligned} & H'^n_k f(x) - \frac{n}{\delta_n^2} \delta H_k f(x) \\ &\quad = \int (p^{(m_n)}(x, y_{k-1}; x', y) - p^{(m_n)}(x, y_{k-1}; x', y_k)) f(x') \psi_{n,k}(y)^2 dx' dy \\ &\qquad + \frac{n}{\delta_n^2} \delta \int (p^{(m_n)}(x, y_{k-1}; x', y_k) - p(x, y_{k-1}; x', y_k)) f(x') dx' \\ &= \frac{n}{\delta_n^2} \int (p^{(m_n)}(x, y_{k-1}; x', y_k + \frac{z}{n^{\frac{1}{4+q}}}) - p^{(m_n)}(x, y_{k-1}; x', y_k)) f(x') \psi(z)^2 dx' dz \\ &\qquad + \frac{n}{\delta_n^2} \delta \int (p^{(m_n)}(x, y_{k-1}; x', y_k) - p(x, y_{k-1}; x', y_k)) f(x') dx', \end{aligned}$$

and the same arguments as above lead to the second part of (3.4). $\square$

4) We can now apply Theorems 3.1 and 4.1, and we have the following corollary (where H_k and K_k are defined in (7.7) and (7.9)):

Corollary 7.2 *Assume (C1) and that μ integrates all functions in the class $\mathcal{C}$ which is either $\mathcal{D}$ or $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. Then for any $f \in \mathcal{C}$ the variables $U_N^n f = n^{\frac{2}{4+q}}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable with variance*

$$\Gamma_N f \;=\; \delta \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H_s\{(H_{s,N}(f - \pi_N f))^2\}. \tag{7.11}$$

and with mean

$$M_N f \;=\; \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)). \tag{7.12}$$

8 CLT and the number of simulated variables

As said in the introduction, the above results are not quite satisfactory, because the number n of particles in the procedure does not really measure the time necessary to perform. A much better measurement is the number $r = r_n$ of "single" variables which have to be simulated. By a "single" simulated variable we mean a random variable which may take values in $\mathbb{R}^d$ or $\mathbb{R}^q$ and which can be simulated according to well-established procedures: in case an Euler scheme is needed, only Gaussian variables have to be simulated. In the other case, we must simulate variables according to the distribution $Q(x,.)$ for example, and the difficulty depends on the actual form of Q.

So in all cases, the computational time is proportional to the number r_n described above, with a factor which depends only on d and q and (in cases A-1, B-1 and C-1) on the transitions Q and P.

It is quite easy to compute r_n in all our cases, for any fixed N:

Case A-1: we have $r_n = 2Nn$,

Case A-2: we have $r_n = (1 + m_n)nN \sim Nn^{3/2}$,

Case B: we have $r_n = (1 + n)Nn \sim Nn^2$ in both subcases,

Case C-1: we have $r_n = 3Nn$,

Case C-2: we have $r_n = (1 + 2Nm_n)n \sim 2Nn^{\frac{6+q}{4+q}}$.

Then we readily deduce from Corollaries 5.2, 6.4 and 7.2 that:

Corollary 8.1 *In Subcase A-1, assume (A0) and (A2) and that μ integrates all functions in the class $\mathcal{C}$ which is either $\mathcal{D}$ or $\mathcal{D}_{pol}$. Then for any $f \in \mathcal{C}$ the variables $\sqrt{r_n}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a centered Gaussian variable with variance*

$$\sqrt{2N} \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H'_s\{(H_{s,N}(f - \pi_N f))^2\}, \tag{8.1}$$

where H'_k is given by (5.6) and $H_k f(x) = \int q(x,x')f(x')g(x',y_k)dx'$.

Corollary 8.2 *In Subcase A-2, assume (A1) and (A2) and that μ integrates all functions in the class $\mathcal{C}$ which is either $\mathcal{D}$ or $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. Then for any $f \in \mathcal{C}$ the variables $r_n^{\frac{1}{3}}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable with mean and variance given respectively by*

$$N^{\frac{1}{3}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)). \tag{8.2}$$

$$N^{\frac{2}{3}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H'_s\{(H_{s,N}(f - \pi_N f))^2\}, \tag{8.3}$$

where H_k and H'_k as in (8.1) and K_k as in (5.7).

Corollary 8.3 *In Case B, assume (B0) (resp. $\int |x|^p \mu(dx) < \infty$ for all $p < \infty$), and (B2). Let f be a differentiable function, belonging to L_b (resp. L_{pol}) as well as all its first derivatives. Then the variables $r_n^{\frac{1}{4}}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable, characterized as such (with H_k, H'_k and K_k as in (6.2), (6.23) and (6.25) repectively):*

(i) In Subcase B-1 it is centered with variance

$$\sqrt{N} \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H'_s\{(H_{s,N}(f - \pi_N f))^2\}. \tag{8.4}$$

(ii) In Subcase B-2 it has the same variance as above, and its mean is

$$N^{\frac{1}{4}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)). \tag{8.5}$$

Corollary 8.4 *In Subcase C-1, assume (C1) and that μ integrates all functions in the class $\mathcal{C}$ which is either $\mathcal{D}$ or $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. Then for any $f \in \mathcal{C}$ be the variables $r_n^{\frac{2}{4+q}}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable with mean and variance given (with H_k and K_k as in (7.7) and (7.9)) repectively by:*

$$(3N)^{\frac{2}{4+q}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)), \tag{8.6}$$

$$\delta(3N)^{\frac{4}{4+q}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H_s\{(H_{s,N}(f - \pi_N f))^2\}. \tag{8.7}$$

Corollary 8.5 *In Subcase C-2, assume (C1) and that μ integrates all functions in the class $\mathcal{C}$ which is either $\mathcal{D}$ or $\mathcal{D}_{pol}$ or $\mathcal{D}_{exp}$. Then for any $f \in \mathcal{C}$ the variables $r_n^{\frac{2}{6+q}}(\Psi_N^n f - \pi_N f)$ converge in law, as $n \to \infty$, to a Gaussian variable with mean and variance given (with H_k and K_k as in (7.7) and (7.9)) repectively by:*

$$(2N)^{\frac{2}{6+q}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)} \pi_{s-1} K_s H_{s,N}(f - \pi_N f)), \tag{8.8}$$

$$\delta(2N)^{\frac{4}{6+q}} \sum_{s=1}^{N} \frac{1}{\beta(s,N)^2} \pi_{s-1} H_s\{(H_{s,N}(f - \pi_N f))^2\}. \tag{8.9}$$

Remarks: 1) Equation (1.3) is a particular case of Equation (1.4). Hence if the coefficients of Equations (1.1) and (1.3) satisfy the assumption (C1) we can treat Case B with the method of Case C. Then, comparing the previous Corollaries 8.3 and 8.4, in the case where we do not need an Euler scheme for our simulations, we see that (as far as rates are concerned) we should use the method of Case C if $q \leq 3$, and of Case B if $q \geq 5$, while for $q = 4$ the two methods give the same rate. If we need an Euler scheme, then comparing Corollaries 8.3 and 8.5, we see that we should use the method of Case C if $q = 1$ and the method of Case B if $q \geq 3$, while for $q = 2$ the two methods give the same rate.

2) In fact in case C-1 the assumptions on the coefficients are somewhat too strong. The reader will check by himself that for C-1 we only need the coefficients a, b, a', b' to be Lipshitz and bb^* and $b'b'^*$ be non-degenerate at each point, as long as the set of test functions is either $\mathcal{D}$ or $\mathcal{D}_{\text{pol}}$: this is enough to imply that the transition P has a density (a fact needed in (7.7)).

References

[1] V. Bally, D. Talay: The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution fuction. Probab. Th. Related Fields, **104**, 43-60, (1996).

[2] V. Bally, D. Talay: The law of the Euler scheme for stochastic differential equations: II. Approximation of the density, Monte Carlo Methods and Applications, **2**, 93-128, (1996).

[3] P. Del Moral, J. Jacod, P. Protter: The Monte-Carlo method for filtering with discrete-time observations. Preprint (1998).

[4] T.G. Kurtz, P. Protter: Wong Zakai corrections, random evolutions, and simulation schemes for SDE's, *Stochastic Analysis*, ed. Mayer-Wolf, Merzbach and Schwarz, AcademicPress: New York, 1991, 331-146.

[5] M. Ledoux, M. Talagrand: *Probability in Banach Spaces, Isoperimetry and Processes*, Springer-Verlag, 23, (1991)