A Performance Evaluation of Branching Particle Filters: Case Studies

Alexsandro Machado Jacob and Takashi Yoneyama

***Abstract*— This paper presents two case studies for the performance evaluation of branching particle filters with the objective of contributing towards understanding and providing useful insight for practical implementation. A performance study based on the robustness of the estimate in relation to the number of independent simulations of three different modes of implementation of the branching particle filter was also made and the results were compared with those obtained by the extended Kalman filter. In the first case, an one-dimensional nonlinear stable process, all the configurations of the branching particle filter produced slightly better estimates than the extended Kalman one when the process was located at regions where the linear approximation was not good. For the second case, the filters were tested in an one-dimensional unstable system and the branching particle filter presented robust results when compared to the extended Kalman one during the simulation steps.**

***Keywords*— Nonlinear filtering, Zakai equation, Branching algorithm, Monte Carlo approximation, Parallel processing.**

I. INTRODUCTION

The standard continuous-time nonlinear filtering problem consists of estimating a time-homogeneous Markov process $X = \{X_t;\, t \geq 0\}$ with known law, or *signal process*, given $Y = \{Y_t;\, t \geq 0\}$, or *observation process*, defined by

$$Y_t = \int_0^t h(s, X_s)ds + V_t, \;\; 0 \leq t \leq T. \tag{1}$$

The signal process X takes values in $\mathbb{R}^d$, whereas the observation one in $\mathbb{R}^m$, with $d, m \geq 1$. The process $\{V_t; t \geq 0\}$ is a standard m-dimensional Brownian motion, T is a fixed final time, and $h : \mathbb{R}^+ \mathrm{x} \mathbb{R}^d \rightarrow \mathbb{R}^m$ is a bounded continuous function. It is assumed that $Y_0 = 0$ and X_0 is a random variable with law ξ. The classical filtering problem, or the best estimate of X with respect to the minimum square error relative to the true signal, can be summarized as finding the conditional distribution of X_t with respect to the observations Y up to time t, that is,

$$\pi_t(\varphi) = \int_{\mathbb{R}^d} \varphi(x)\pi_t(dx) = \mathbb{E}[\varphi(X_t)|\mathcal{F}_0^t], \; a.s., \tag{2}$$

where $\mathcal{F}_0^t$ is the filtration generated by Y up to time t and $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a bounded continuous function.

Based on the assumptions presented by equation (1), the conditional distribution $\pi_t(\varphi)$ is known to satisfy the measure valued stochastic differential equations called Fujisaki-Kallianpur-Kunita (FKK) [1] and Kushner-Stratonovich [2] equations.

The unnormalised conditional distribution p_t of X_t given the filtration generated by $Y = \{Y_t;\, t \geq 0\}$ has also been studied extensively in the literature. It satisfies the following linear stochastic partial differential equation, or Zakai equation [3],

$$p_t(\varphi) = \pi_0(t) + \int_0^t p_s(A\varphi)ds + \int_0^t p_s(h^*\varphi)dY_s, \tag{3}$$

where A is the infinitesimal generator associated to X. Uniqueness of solution to the measure valued equation presented above has been established in [4].

Bensoussan *et al.* [5] - [6] and Le Gland [7] approximated the solution of the Zakai equation via the splitting-up method for solving stochastic partial differential equations (SPDE). Kushner [8] did the same, but approximated the referred trajectories by using a discrete-time Markov chain, whereas Sun and Glowinski [9] used a pathwise approximation via operator splitting.

Crisan [10] described a system of moving branching particles whose empirical distribution at time t, denoted by $U_n(t)$, converges almost surely to p_t, that is,

$$\lim_{n\rightarrow\infty} (U_n(t), \varphi) \;\; = \;\; p_t(\varphi). \tag{4}$$

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

In fact, Crisan and Lyons [11] early proposed a sequence U_n of similar branching particle systems where the variance of the branching mechanism was given *a priori* and only the (conditional) expectation of that sequence converged to p_t due to an extra degree of randomness introduced. Then, a whole set of copies of the particle system was needed in order to obtain a good approximation to the solution of the Zakai equation. More recently, the new approach presented by Crisan *et al.* [12] converges directly to p_t without any estimate of an average, but the rate of convergence of the mean square error was not deduced correctly. In [10], Crisan deduced the exact rate of convergence of the mean square error and proposed variations of the branching algorithm, or *Branching Particle Filter* (BPF), with better rates of convergence.

The purpose of this paper is to present two case studies for the performance evaluation of BPF with the objective of contributing towards understanding and providing useful

The authors are with the Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, Pça. Mal. Eduardo Gomes, 50 - São José dos Campos, SP - Brazil (e-mail: ajacob@ele.ita.br). This work was supported by The State of São Paulo Research Foundation (FAPESP) under Grant 02/10632-0.

insight for practical implementation. A performance study based on the robustness of the estimate in relation to the number of independent simulations of three different modes of implementation of BPF was also made and the results were compared with those obtained by the extended Kalman filter (EKF). The first case, an one-dimensional nonlinear stable process, all the configurations of BPF produce slightly better estimates than the EKF when the process is located at regions where the linear approximation was not good. For the second case, the filters were tested in an one-dimensional unstable system and the BPF presented robust results when compared to the EKF during the simulation steps. These results are consistent with those obtained by Jacob *et al.* [13] when a Monte Carlo-based nonlinear filter proposed by Davis [14] was submitted to the same tests.

The basic concepts about the filtering framework are described in Section 2. Section 3 presents the simple Monte Carlo approach to solve the Zakai equation, whereas Section 4 shows the main results of the BPF. A brief discussion about the BPF implementation occurs in Section 5 and numerical examples are posted in Section 6. The conclusions are in Section 6.

II. The Filtering Framework

Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ be a filtered probability space on which the signal X is described by the following stochastic differential equation (SDE)

$$dX_t = f(t, X_t)\, dt + g(t, X_t)\, dW_t, \tag{5}$$

where $\{W_t; t \geq 0\}$ is a standard d-dimensional Brownian motion independent of V, and

$$dY_t = h(t, X_t)\, dt + dV_t \tag{6}$$

is its corresponding noisy observation, in accordance to the assumptions related to (1) and that

$$||h|| = \max_{1\leq i\leq m} \sup_{x\in\mathbb{R}^m} |h_i(t,x)|, \ for\, all\, t. \tag{7}$$

The functions $f : \mathbb{R}^+ \mathrm{x} \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $g : \mathbb{R}^+ \mathrm{x} \mathbb{R}^d \rightarrow \mathrm{L}(\mathbb{R}^d, \mathbb{R}^d) \simeq \mathbb{R}^{d^2}$ are assumed to be globally Lipschitz and $X_0 \sim \xi$ is a d-dimensional square integrable random vector, $\mathcal{F}_0$-measurable and independent of W and V. These hypothesis satisfy the sufficient conditions for the uniqueness of the solution of equation (5) [15].

A new probability measure $\tilde{P}$ absolutely continuous with respect to P can be defined as

$$\tilde{P}(A) = \mathbb{E}[1_A Z_t], for\, all\, A \in \mathcal{F}_t, t \geq 0, \tag{8}$$

where $Z = \{Z_t, \mathcal{F}_t; t \geq 0\}$ is the exponential martingale

$$Z_t = exp(-\int_0^t h^T(s, X_s) dW_s - \frac{1}{2}\int_0^t ||h(s, X_s)||^2 ds). \tag{9}$$

Under the new measure $\tilde{P}$ and according to Girsanov's theorem [15], the observation process Y becomes a Brownian motion independent of the signal process X. Thus, the conditional law (2) can be calculated by using the Kallianpur-Striebel formula [16]

$$\pi_t(\varphi) = \frac{p_t(\varphi)}{p_t(1)}, \ P - a.s., \tag{10}$$

where

$$p_t(\varphi) = \tilde{\mathbb{E}}[\varphi(X_t)\, exp(\int_0^t h^T(s, X_s) dY_s - \frac{1}{2}\int_0^t ||h(s, X_s)||^2 ds) | \mathcal{F}_t], \tag{11}$$

is a random measure and $\tilde{\mathbb{E}}$ is the expectation with respect to $\tilde{P}$. Further, equation (3) is uniquely satisfied by (11).

Theorem 1 (Kurtz and Ocone [17]): Under the conditions set up above, if U_t is a $\mathcal{F}_t$-adapted, cádlág, measure valued process satisfying

$$\begin{aligned}(U(t), \varphi) &= (\pi_0, \varphi) + \int_0^t (U(s), A\varphi) ds \\ &\quad + \int_0^t (U(s), h^*\varphi) dY_s, \ a.s., \end{aligned} \tag{12}$$

for all t and for a suitable large class of test functions φ, then

$$U(t) = p_t, \ \ a.s., \tag{13}$$

for all t.

Since $U(t)$ is not easy to be obtained, the idea is that it might be possible to approximate the Zakai equation by creating a *sample* from the posterior measure. Next sections show the branching particle approach to solve the filtering problem described here.

III. Monte Carlo Filter

Assume that $V_i(t)$, $i = 1, ..., P$ and $t \in [0, T]$, are independent realizations of the signal X and that they are independent of the observation Y. Let $\delta_{V_i(t)}$ be an impulse measure defined on $\sigma(\mathbb{R}^d)$ for each t, where

$$\Theta_P(t) = \frac{1}{P}\sum_{i=1}^{P} \mu^i_{0:t} \delta_{V_i(t)}, \tag{14}$$

is the empirical measure of P particles with mass determined by

$$\mu^i_{0:t} = exp(\int_0^t h^*(V_i(t)) dY_s - \frac{1}{2}\int_0^t ||h(V_i(t))||_2^2 ds), \tag{15}$$

that is, the likelihood/weight corresponding to $V_i(t)$.

Proposition 1 (Crisan [10]): If $\Theta_P(t)$ is the approximation given by (14), then

$$\tilde{\mathbb{E}}[((\Theta_P(t), \varphi) - p_t(\varphi))^2] = \frac{c_\Theta(t)}{N}, \tag{16}$$

where $c_\Theta(t)$ is a constant independent of P for each t.

A. Monte Carlo Filter Algorithm

Consider a partition with a time discretization $(\mathcal{T}_\delta)$ where

$$0 = \mathcal{T}_0 < \mathcal{T}_1 < \cdots < \mathcal{T} < \cdots < \mathcal{T}_N = T, \tag{17}$$

$$\delta = \frac{T}{N} \tag{18}$$

and

$$\mathcal{T}_{j+1} = \mathcal{T}_j + \delta, \ j = 0, 1, ..., (N-1). \tag{19}$$

Given an initial distribution ξ, the step-by-step algorithm of the so-called *Monte Carlo Filter* (MCF) is described as the following:

At time $t = \mathcal{T}_0$,
Step 0: *Initialization*
- For $i = 1, ..., P$, sample $V_i(t) \sim \xi$;
- Set $t = \mathcal{T}_1$.

While $t \neq \mathcal{T}_N$,
Step 1: *Evolution*
- For $i = 1, ..., P$, evolve $V_i(t)$ in accordance to the model described by (5).

Step 2: *Importance weights evaluation*
- For $i = 1, ..., P$, evaluate the importance weights $\mu^i_{0:t}$ using equation (15).

Step 3: *Conditional law computation*
- Compute the conditional law according to (14);
- Set $j = j + 1$ and go to **Step 1**.

end

IV. Principles of the Branching Particle Filter

Initially, based on the time discretization presented by (17), the inter-branching time interval is defined by

$$\Delta\mathcal{T}_l = \Delta_{l+1} - \Delta_l = n\delta, \;\; l, n \in \mathbb{N} \tag{20}$$

and can be visualized in Figure 1 below.

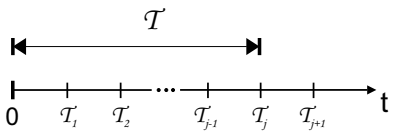

Fig. 1. Discretization of the time line.

Crisan and Lyons [11] constructed a measured value process whose expectation at any time is the conditional distribution of X_t. This has a branching particle system approximation and the particles evolve independently moving with the same law as X and branch according to a mechanism that depends on the trajectory of the particle and Y, but is independent of the events elsewhere in the system. The mechanism is chosen so that it has finite second moment and the mean number of offspring for a particle given the σ-field $\mathcal{F}_{\Delta_{l+1}-} = \sigma(\mathcal{F}_s, s < \Delta_{l+1})$ of events up to time Δ_{l+1} is

$$\mu^i_{\Delta\mathcal{T}_l} = exp(\int_{\Delta_l}^{\Delta_{l+1}} h^T(V_i(t))dY_t - \frac{1}{2}\int_{\Delta_l}^{\Delta_{l+1}} ||h(V_i(t))||^2 dt). \tag{21}$$

The referred variance must be minimal and consistent with the number of offspring being an integer, that is,

$$\nu^i_{\Delta\mathcal{T}_l} = (\mu^i_{\Delta\mathcal{T}_l} - [\mu^i_{\Delta\mathcal{T}_l}])([\mu^i_{\Delta\mathcal{T}_l}] + 1 - \mu^i_{\Delta\mathcal{T}_l}) \leq \frac{1}{4} \tag{22}$$

where $[\mu^i_{\Delta\mathcal{T}_l}]$ is the largest integer smaller than $\mu^i_{\Delta\mathcal{T}_l}$. The result is a cloud of paths where those surviving to the current time provide an estimate for the conditional distribution of X_t in the following manner

$$U_{P(\Delta_l)}(\Delta\mathcal{T}_l) = \frac{1}{P(\Delta_l)} \sum_{i=1}^{P(\Delta_l)} \mu^i_{\Delta_l:\Delta_{l+1}} \delta_{V_i(\Delta_{l+1})}, \tag{23}$$

where $P(\Delta_l)$ is the number of particles corresponding to the instant $t = \Delta_l$.

This approach is feasible in the sense that one carry it out and get a return directly related to the amount of computational effort invested. However, the convergence could still be quite slow. Paths exploring unfruitful directions of exploration are rapidly killed suggesting a model akin to lemmings flowing along and reproducing heavily, but being killed if they drift away form the plausible values of the variables.

Corollary 1 (Crisan [10]): If the length of the inter-branching times is $\frac{1}{P^\alpha}$, where $\alpha \in (\frac{2}{3}, 2)$, then

$$\lim_{P\to\infty} P^{1-\frac{\alpha}{2}} \tilde{\mathbb{E}}[((U_P(t), \varphi) - p_t(\varphi))^2] = c_U(t), \tag{24}$$

where $c_U(t)$ is a constant independent of P.

According to the corollary, the larger the length of the inter-branching times $\Delta\mathcal{T}$ is, the better rate is. However, the order of the length of the inter-branching times cannot be larger than $\frac{1}{P^{\frac{2}{3}}}$ as the last part of evolution of the system is not corrected and hence a bias is introduced. By other hand, if the inter-branching times are of order $\frac{1}{P^2}$, then U_P no longer converges to p_t. By branching so often, the randomness introduced in the system at branching times overpowers the corrective effect and, as a result, just as in the case when the branching variance is fixed, the limiting process is measure valued process whose conditional expectation given the environment Y is p_t [10].

A. Branching Particle Filter Algorithm

Let $\{U_n(t), \mathcal{F}_t; 0 \leq t \leq 1\}$ be a sequence of branching particle systems on $(\Omega, \mathcal{F}, \tilde{P})$ with values in $\mathbb{R}^m$ defined according to (23) and let $P(\mathcal{T}_0)$ be the initial number of particles. The BPF step-by-step algorithm is described as the following:

At time $t = \mathcal{T}_0 = \Delta_0$,
Step 0: *Initialization*
- For $i = 1, ..., P(\mathcal{T}_0)$, sample $V_i(t) \sim \xi$;
- Compute $\Delta\mathcal{T}_0$ according to *a priori* convergence rate and $P(\mathcal{T}_0)$ by using (24);
- Set $t = \Delta_1$.

While $t \leq \Delta_M$, $\Delta_M = \mathcal{T}_N$,
Step 1: *Evolution*
- For $i = 1, ..., P(\Delta_{l-1})$, evolve $V_i(\Delta\mathcal{T}_{l-1})$ in accordance to the model described by (5).

Step 2: *Importance weights evaluation*
- For $i = 1, ..., P(\Delta_{l-1})$, evaluate the importance weights $\mu^i_{\Delta\mathcal{T}_{l-1}}$ using equation (21).

Step 3: *Conditional law computation*
- Compute the conditional law according to (23).

Step 4: *Offspring computation*
- For $i = 1, ..., P(\Delta_{l-1})$,

– Sample $\epsilon \sim \mathcal{U}[0,1]$;
– Compute

$$m^i(\Delta_l) = round(\mu^i_{\Delta\mathcal{T}_l} + \epsilon - 0.5), \quad (25)$$

where $m^i(\Delta_l)$ is the number of particles.

Step 5: *Replacement*

- For $i = 1, ..., P(\Delta_{l-1})$, replace the particles in $V_i(\Delta_l)$ according to $m^i(\Delta_l)$.

Step 6: *Inter-branching interval computation*

- Compute the total number of particles as

$$P(\Delta_l) = \sum_{i=0}^{P(\Delta_{l-1})} m^i(\Delta_l); \quad (26)$$

- Compute $\Delta\mathcal{T}_l$ according to *a priori* convergence rate and $P(\Delta_l)$ by using (24);
- Set $t = \Delta_{l+1}$ and go to **Step 1**.

end

Specifically, in order to keep the number of particles limited, a range containing the minimum and a maximum number of particles for BPF algorithm was established. According to preliminary tests, in this approach a process can be completely killed or the number of particles can be large enough to difficult applications in real time, what is not interesting in many applications. When the number of particles at time Δ_l is larger than an upper limit, then the particles are not replaced; when the number of particles at time Δ_l is smaller than a downer limit, then the particles are replaced for that with wider likelihood.

B. Modified Branching Particle Filter Algorithm

The idea here is to maintain the number of particles fixed during the evolution of the branching system, that is, $P(\Delta_l) = P$ for all l. For this case, it is not necessary to compute a new number of offspring and the replacement procedure must be modified. Based on the resampling ideas presented by Doucet *et al.* [18], the new step-by-step algorithm for the so-called *Modified Branching Particle Filter* (MBPF) is presented as follows:

At time $t = \mathcal{T}_0 = \Delta_0$,

Step 0: *Initialization*

- For $i = 1, ..., P$, sample $V_i(t) \sim \xi$;
- Compute the fixed $\Delta\mathcal{T}$ according to *a priori* convergence rate and P by using (24);
- Set $t = \Delta_1$.

While $t \leq \Delta_M$, $\Delta_M = \mathcal{T}_N$,

Step 1: *Evolution*

- For $i = 1, ..., P$, evolve $V_i(\Delta\mathcal{T}_{l-1})$ in accordance to the model described by (5).

Step 2: *Importance weights evaluation*

- For $i = 1, ..., P$, evaluate the importance weights $\mu^i_{\Delta\mathcal{T}_{l-1}}$ using equation (21).

Step 3: *Conditional law computation*

- Compute the conditional law according to (23).

Step 4: *Weights normalization*

- Compute the normalization factor

$$\gamma = \sum_{i=1}^{P} \mu^i_{\Delta\mathcal{T}_{l-1}}; \quad (27)$$

- For $i = 1, ..., P$, normalize the weights

$$\tilde{\mu}^i_{\Delta\mathcal{T}_{l-1}} = \gamma^{-1}\mu^i_{\Delta\mathcal{T}_{l-1}}. \quad (28)$$

Step 5: *Replacement*

- For $i = 1, ..., P$, replace proportionally the particles in $V_i(\Delta_l)$ based on the normalized weights $\tilde{\mu}^i_{\Delta\mathcal{T}_{l-1}}$;
- Set $t = \Delta_{l+1}$ and go to **Step 1**.

end

V. Details about the Implementation

Computer programs were produced to simulate the MCF, BPF and MBPF by using equations (5) and (6).

A. Simulation of Sample Paths

The objective is to use a discretization scheme aiming at the best possible order of accuracy, given the SDE to be solved, from among schemes that use only increments of Brownian path.

Based on the assumptions presented in (5), the kth component of the Milstein scheme [19] has the form

$$X^k_{\mathcal{T}_{l+1}} = X^k_{\mathcal{T}_l} + f^k\delta + g^k\Delta W^k_{\mathcal{T}_l} + \frac{1}{2}(\sum_{l=1}^{d} g^l \frac{\partial g^k}{\partial x^l})\{(\Delta W^k_{\mathcal{T}_l})^2 - \delta\}, \quad (29)$$

where

$$\Delta W_{\mathcal{T}_l} = W_{\mathcal{T}_{l+1}} - W_{\mathcal{T}_l} \quad (30)$$

is the $N(0;\delta)$ increment of the Wiener process W. Particularly, if g does not depend on x, equation (29) becomes the Euler-Maruyama scheme.

When the SDE is one-dimensional or the function $g = g(t,x)$ satisfies the commutativity condition

$$\sum_{k=1}^{n}(\frac{\partial g^j_q}{\partial x^k}g^k_p - \frac{\partial g^j_p}{\partial x^k}g^k_q) = 0 \quad (31)$$

$\forall j = 1, ..., d, \forall p, q = 1, ...r_1$, the best order of accuracy is usually $\mathcal{O}(\delta)$ and it can be obtained using the Milstein scheme. In the general case of an SDE in more than one dimension and not satisfying condition (31), the best obtainable accuracy is $\mathcal{O}(\sqrt{\delta})$, which is obtained using the Euler-Maruyama scheme [19].

B. Integration

To approximate the integrals in (15) and (21), both needed for determining the mean number of offspring for each particle

at the end of each generation, the simplest method is the well-known Euler one

$$\mu^i_{\mathcal{T}_k:\mathcal{T}_{k+m}} = \sum_{j=k}^{m-1} h^T(V_i(\mathcal{T}_j))\Delta Y(\mathcal{T}_j) - \frac{1}{2}\sum_{j=k}^{m-1} h^T h(V_i(\mathcal{T}_j))\delta, \quad (32)$$

where

$$\Delta Y(\mathcal{T}_j) = Y(\mathcal{T}_{j+1}) - Y(\mathcal{T}_j). \quad (33)$$

According to [12], since the integrals involved are stochastic, the use of the trapezoidal rule

$$\mu^i_{\mathcal{T}_k:\mathcal{T}_{k+m}} = \tfrac{1}{2}\sum_{j=k}^{m-1}[h^T(V_i(\mathcal{T}_j)) + h^T(V_i(\mathcal{T}_{j+1}))]\Delta Y(\mathcal{T}_j) - \tfrac{1}{4}\sum_{j=k}^{m-1}[h^T h(V_i(\mathcal{T}_j)) + h^T h(V_i(\mathcal{T}_{j+1}))]\delta \quad (34)$$

does not give a higher-order approximation than the Euler one, however it does improve the leading coefficient of the error significantly. The sums in (34) converge to the same values as those in (32), with no correction term required, due to independence of the processes V_i and Y.

VI. Experimental Results

The main contribution here is to compare the performances of the MCF, BPF, MBPF nonlinear approaches to the traditional EKF. Generally, for the nonlinear filters, just parameters related to the initial number of particles and length of the inter-branching time can be adjusted for improving the rate of convergence; except for the MCF where just the number of particles is important. The initial covariance matrix of the EKF must be adjusted with accuracy in order to minimize the transient in the first steps. The criteria used for the comparison and the numerical examples are presented next.

A. The Absolute Error Criterion

This criterion is based on the performance measure used to verify the estimate accuracy of the filters at fixed $\mathcal{T}$. By fixing R, $j = 1, ..., R$, different initial conditions for the estimated signal $\hat{x}_0^{j,k}$, $k = 1, ..., M$, M different simulated sample paths of the estimated signal $\hat{x}$ are performed and compared to their corresponding real values x. The jth estimated absolute error is defined as

$$\hat{\varepsilon}_j = \frac{1}{M}\sum_{k=1}^{M} |\hat{X}_{\mathcal{T}}^{j,k} - X_{\mathcal{T}}^{j,k}| \quad (35)$$

which is independent and Gaussian for large M.

To construct a confidence interval for the estimated absolute error, the mean of the batch averages is

$$\hat{\varepsilon} = \frac{1}{R}\sum_{j=1}^{R}\hat{\varepsilon_j} = \frac{1}{RM}\sum_{j=1}^{R}\sum_{k=1}^{M}|\hat{X}_{\mathcal{T}}^{j,k} - X_{\mathcal{T}}^{j,k}| \quad (36)$$

and the estimated variance of ε is

$$\hat{\sigma}^2_{\varepsilon} = \frac{1}{R-1}\sum_{j=1}^{R}(\hat{\varepsilon_j} - \hat{\varepsilon})^2. \quad (37)$$

Then, the Student t-distribution with $R-1$ degrees of freedom is used to construct the $(1-\alpha)$% confidence interval for ε according to

$$[\Delta\varepsilon] = (\hat{\varepsilon} - \Delta\hat{\varepsilon},\ \hat{\varepsilon} - \Delta\hat{\varepsilon}), \quad (38)$$

with

$$\Delta\varepsilon = t_{1-\alpha,R-1}\sqrt{\frac{\hat{\sigma^2_{\varepsilon}}}{R}}, \quad (39)$$

where $t_{1-\alpha,R-1}$ is the significance level.

B. Stable System

Consider the filtering problem with the following one-dimensional Ornstein-Uhlenbeck process

$$dx_t = -1.0\,dt + 0.25\,dw_t \quad (40)$$

whose one-dimensional observation process is

$$dy_t = h(x_t)\,dt + dv_t \quad (41)$$

with

$$h(x) = \begin{cases} sin(x) & \text{if } |x| \le \pi/2 \\ +1 & \text{if } x > \pi/2 \\ -1 & \text{if } x < -\pi/2 \end{cases}, \quad (42)$$

where w_t and v_t are independent one-dimensional standard Brownian motions. The filtering was carried out for the range $t \in [0, 5]$ which was divided in $N = 2^{11}$ intervals. All the integrals were approximated by the traditional Euler scheme. The signal process initialization was set to $x(0) \sim N(2.0, 0.2)$ in order to evaluate the performance of the Euler-Maruyama discretization scheme over the nonlinear range of the observation process.

The BPF and the MBPF implementation was made aiming to obtain the best rate of convergence in accordance with the result (24) of the Corollary 1. However, to simplify the algorithm implementation, the inter-branching times computed were multiple of 2^k, $k = 0, 1, 2, ...$

Figure and 2 present a typical realization of the process (40) with the filtering results for the MCF with 100 particles, BPF initialized with 100 particles and range from 1 up to 1000, and EKF.

The performance evaluation was made based on the robustness of the MCF, MBPF and BPF estimates, where the formers were tested with three different configurations: $P = 10, 100, 1000$, and the latter with a given initial number of particles inside a range of working. The EKF robustness estimate was evaluated for 100 runs. To satisfy the *absolute error criterion*, the absolute error ε and its respective 90% confidence interval $[\Delta\varepsilon]$ were computed at instants $t = 1.0s$ and $t = 3.0s$, as presented in Tables I and II, respectively. The performance provided by the nonlinear filters and EKF are similar, except by the fact that here the MCF, MBPF and BPF confidence interval at $t = 1.0s$ are significantly better for $P \ge 100$ due to the signal process be located at the strong nonlinear region of the observation one.

Note that the MCF obtained better results when compared to others nonlinear filters for the same number of initial particles.

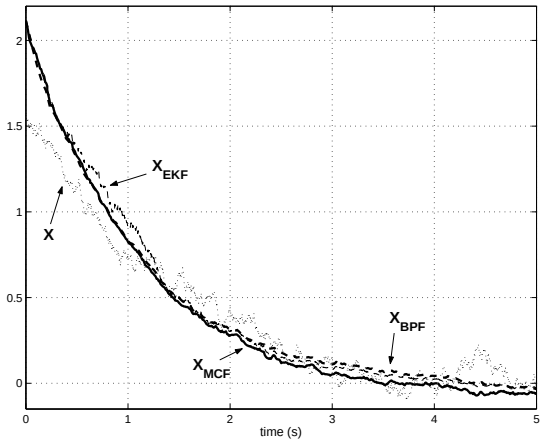

Fig. 2. A realization of (40) and its respective MCF,BPF and EKF estimates.

TABELA I

ESTIMATION AND 90% CONFIDENCE INTERVAL OF $\hat{\varepsilon}$ OBTAINED BY MCF, BPF, MBPF AND EKF AT $t = 1.0s$ FOR THE STABLE SYSTEM.

P	$\hat{\varepsilon}(10^{-1})$	$[\Delta\varepsilon](10^{-1})$
MCF		
10	2.155±0.230	(2.117,2.193)
100	2.043±0.151	(2.018,2.068)
1000	2.023±0.141	(2.000,2.046)
MBPF		
10	2.794±0.458	(2.718,2.870)
100	2.108±0.156	(2.082,2.134)
1000	2.044±0.140	(2.021,2.067)
BPF		
10 [5 15]	2.469±0.316	(2.417,2.521)
10 [1 1000]	2.302±0.281	(2.255,2.349)
100 [50 200]	2.255±0.189	(2.220,2.229)
100 [1 1000]	2.075±0.158	(2.049,2.101)
EKF	2.252±0.158	(2.226,2.278)

TABELA II

ESTIMATION AND 90% CONFIDENCE INTERVAL OF $\hat{\varepsilon}$ OBTAINED BY MCF, BPF, MBPF AND EKF AT $t = 3.0s$ FOR THE STABLE SYSTEM.

P	$\hat{\varepsilon}(10^{-1})$	$[\Delta\varepsilon](10^{-1})$
MCF		
10	1.491±0.114	(1.472,1.510)
100	1.417±0.102	(1.400,1.434)
1000	1.402±0.104	(1.385,1.419)
MBPF		
10	1.567±0.099	(1.551,1.583)
100	1.441±0.110	(1.423,1.459)
1000	1.417±0.109	(1.399,1.435)
BPF		
10 [5 15]	1.563±0.122	(1.543,1.583)
10 [1 1000]	1.495±0.115	(1.476,1.514)
100 [50 200]	1.477±0.113	(1.458,1.496)
100 [1 1000]	1.420±0.106	(1.403,1.437)
EKF	1.419±0.104	(1.402,1.436)

However, according to *Corollary 1*, the convergence rate of the branching particle filters can be better to that presented by the Monte Carlo-based ones for an adequate choice of the inter-branching times and number of particles. The worst results of the performance evaluation for the branching particle system must be investigated better, but some explanations about the practical implementation can be done: 1-) the inter-branching interval computed was multiple of 2^k, and 2-) the results show that the range limitation of the particles affect the performance of the estimate.

C. Unstable System

Now the one-dimensional Ornstein-Uhlenbeck process is defined as

$$dx_t \quad = \quad 0.5\,dt + 0.25\,dw_t, \tag{43}$$

where w_t is a one-dimensional standard Brownian motion and $x(0) \sim N(0.0; 0.20)$. The corresponding observation process is as presented by (41) and (42).

The estimation results of the MCF with 100 particles, BPF initialized with 100 particles and range from 1 up to 1000, and EKF for a typical realization of (43) is shown in Figure 3.

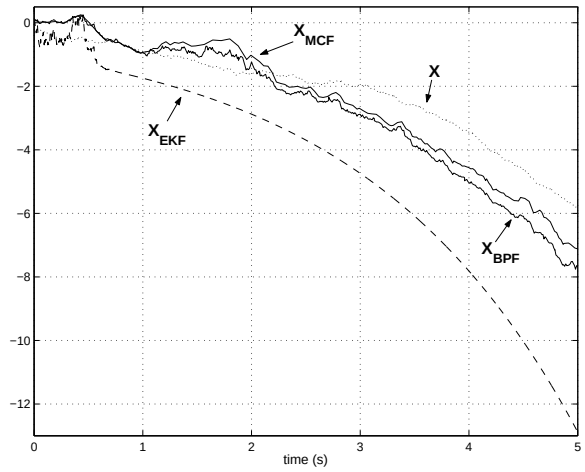

Fig. 3. A realization of (43) and its respective MCF, BPF and EKF estimates.

The performance evaluation here is analogous to to the last example. However, differently from that in the stable case, the EKF failed to provide an acceptable bounded error estimate when compared to the nonlinear filters, as presented by the absolute errors $\hat{\varepsilon}$ in Tables III and IV.

TABELA III

ESTIMATION AND 90% CONFIDENCE INTERVAL OF $\hat{\varepsilon}$ OBTAINED BY MCF, MBPF, BPF AND EKF AT $t = 1.0s$ FOR THE UNSTABLE SYSTEM.

P	$\hat{\varepsilon}(10^{-1})$	$[\Delta\varepsilon](10^{-1})$
MCF		
10	7.027±0.825	(6.890,7.164)
100	6.588±0.566	(6.494,6.682)
1000	6.511±0.529	(6.423,6.599)
MBPF		
10	7.461±0.854	(7.319,7.603)
100	6.902±0.608	(6.801,7.003)
1000	6.570±0.529	(6.482,6.658)
BPF		
10 [5 15]	7.366±0.889	(7.219,7.513)
10 [1 1000]	7.494±1.010	(7.326,7.662)
100 [50 200]	6.623±0.547	(6.532,6.714)
100 [1 1000]	6.651±0.574	(6.556,6.746)
EKF	10.900±0.730	(9.482,12.318)

TABELA IV

ESTIMATION AND 90% CONFIDENCE INTERVAL OF $\hat{\varepsilon}$ OBTAINED BY MCF, MBPF, BPF AND EKF AT $t = 3.0s$ FOR THE UNSTABLE SYSTEM.

P	$\hat{\varepsilon}$	$[\Delta\varepsilon]$
MCF		
10	1.576±0.224	(1.539,1.613)
100	1.440±0.117	(1.421,1.460)
1000	1.419±0.104	(1.402,1.436)
MBPF		
10	2.001±0.260	(1.917,2.085)
100	1.562±0.130	(1.540,1.584)
1000	1.426±0.109	(1.408,1.444)
BPF		
10 [5 15]	2.049±0.357	(1.990,2.108)
10 [1 1000]	1.843±0.300	(1.793,1.893)
100 [50 200]	1.738±0.175	(1.709,1.767)
100 [1 1000]	1.475±0.118	(1.455,1.494)
EKF	2.807±0.206	(2.732,2.882)

VII. CONCLUSION

Two case studies for the performance evaluation of branching particle filters were used with the objective of contributing towards understanding and providing useful insight for practical implementation. A performance study based on the robustness of the estimate in relation to the number of independent simulations of three different modes of implementation of the branching particle filter was also made and the results were compared with those obtained by the extended Kalman filter. Under favorable conditions, the performance provided by the modified nonlinear filter and the extended Kalman filter were similar, but here the former obtained better results on regions where the linear approximation of the observation process was not good. Additionally, based on the example of an unstable system, it was shown that the nonlinear filter, for an adequate number of independent realizations, is much better than the extended Kalman filter.

ACKNOWLEDGEMENT

The first author would like to thank The State of São Paulo Research Foundation (FAPESP) under Grant 02/10632-0.

REFERÊNCIAS

[1] M. Fujisaki, G. Kallianpur, and H. Kunita, "Stochastic differential equations for the nonlinear filtering problem," *Osaka J. Math.*, vol. 9, pp. 19–40, 1972.

[2] H. J. Kushner, "Dynamical equations for optimal nonlinear filtering," *Journal of Differential Equations*, vol. 3, pp. 179–190, 1967.

[3] M. Zakai, "On the optimal filtering for diffusion process," *Z. Wahrsch. Verw.*, vol. 11, pp. 230–243, 1969.

[4] A. G. Bhatt, G. Kallianpur, and R. L. Karandikar, "Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering," *Ann. Probab.*, vol. 23, no. 4, pp. 1895–1938, 1995.

[5] A. Bensoussan, R. Glowinski, and A. Rascanu, "Approximation of the zakai equation by splitting up method," *SIAM J. Control Optim.*, vol. 28, pp. 1420–1431, 1990.

[6] ——, "Approximation of some stochastic differential equations by the splitting up method," *Appl. Math. Optim.*, vol. 25, pp. 81–106, 1992.

[7] F. Le Gland, *Stochastic partial differential equations and their applications*, ser. Lecture Notes in Control and Inform. Sci. 176. New York: Springer-Verlag, 1992, ch. Splitting-up approximation for the SPDEs and SDEs with application to nonlinear filtering.

[8] H. J. Kushner, *Probability methods for approximations in stochastic control and elliptic equations*, ser. Mathematics in Science and Engineering 129. New York: Academic Press, 1977.

[9] M. Sun and R. Glowinski, "Pathwise approximation and simulation for the zakai filtering equation through operator splitting," *Calcolo*, vol. 30, pp. 257–297, 1993.

[10] D. Crisan, "Exact rates of convergence for a branching particle approximation to the solution of the zakai equation," *The Annals of Probability*, vol. 31, no. 2, pp. 693–718, 2003.

[11] D. Crisan and T. Lyons, "Nonlinear filtering and measure-valued processes," *Probability Theory and Related Fields*, vol. 109, pp. 217–244, 1997.

[12] D. Crisan, J. Gaines, and T. Lyons, "Convergence of a branching particle method to the solution of the zakai equation," *SIAM Journal of Applied Probability*, vol. 58, no. 5, pp. 1568–1590, 1998.

[13] A. M. Jacob, E. M. Hemerly, and T. Yoneyama, "Improvements on davis' monte carlo-based nonlinear filtering," in *Congresso Brasileiro de Automática (to appear)*, Gramado, 2004.

[14] M. H. A. Davis, "New approach to filtering for nonlinear systems," *IEE Proc. D, Control Theory and Appl.*, vol. 128, no. 5, pp. 166–172, 1981.

[15] R. S. Lipster and A. N. Shirayaev, *Statistics of randon processes I: general theory*. New York: Springer-Verlag, 1977.

[16] G. Kallianpur and H. Striebel, "Estimations of stochastic processes: Arbitrary system process with additive white noise observation errors," *Ann. Math. Statist.*, vol. 39, pp. 785–801, 1968.

[17] T. G. Kurtz and D. L. Ocone, "Unique characterization of conditional distributions in nonlinear filtering," *Ann. Probab.*, vol. 16, pp. 80–107, 1988.

[18] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo methods in practice*, ser. Statistics for Engineering and Information Science. New York: Springer-Verlag, 2001.

[19] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, 3rd ed. Berlin: Springer, 1999.