Evolution Strategies Based Particle Filters for Simultaneous State and Parameter Estimation of Nonlinear Stochastic Models

K. Uosaki and T. Hatanaka

Department of Information and Physical Sciences, Graduate School of Information Science and Technology, Osaka University, Suita, Osaka 565–0871, Japan

(Tel: +81–6–6879–7833; Fax: +81–6–6879–7836; Email:{uosaki, hatanaka}@ist.osaka-u.ac.jp)

Abstract: Recently, particle filters have attracted attentions for nonlinear state estimation. In this approaches, a posterior probability distribution of the state variable is evaluated based on observations in simulation using so-called importance sampling. We proposed a new filter, Evolution Strategies based particle (ESP) filter to circumvent degeneracy phenomena in the importance weights, which deteriorates the filter performance, and apply it to simultaneous state and parameter estimation of nonlinear state space models. Results of numerical simulation studies illustrate the applicability of this approach.

Keywords: Nonlinear filtering, particle filters, Bayesian approach, evolution strategies, importance sampling, selection.

1. Introduction

State estimation of dynamic systems using a sequence of their noisy observations is ubiquitous in control system science. This problem can be solved by a Bayesian approach, that is, inference on the unknown state can be performed according to the posterior probability distribution (pdf), which is obtained by combining a prior pdf for the unknown state with a likelihood function relating them to the observations. When observations come sequentially in time, recursive state estimation, which evaluates the evolving posterior pdf recursively in time, is often interested. However, the posterior pdf only admits an analytical expression for very restricted cases, including linear Gaussian state space models where well-known Kalman filter [1],[17] can be applied. In many realistic problems, state space models include nonlinear and non-Gaussian elements that preclude a closed form of expression for the optimal state estimate and that many approximations have been proposed such as the extended Kalman filter (EKF) and Gaussian sum filter [12],[9]. By the recent progress of computing ability, "particle filtering," a simulation-based method for Bayesian sequential analysis attracts much attentions. In this approach, the integral in Bayes' rule is approximated by a weighted sum based on the discrete grid sequentially chosen by the importance sampling and the estimates are obtained based on corresponding importance weights [5], [2]. A common problem in the particle filter is the degeneracy phenomenon, where almost all importance weights tend to zero after some iteration. Hence, a large computational effort is wasted to updating the particles with negligible weights. In order to resolve this difficulty, several modifications have been proposed such as resampling particle filter (SIR) [10] introducing a selection a resampling steps and Evolution Strategies based Particle Filter (ESP) [19] introducing the concept of Evolution Strategies [16], an Evolutionary Computation approach. In this paper, the ESP filter is applied to simultaneous state and parameter estimation of nonlinear state space models. Numerical simulation studies have been conducted to exemplify the applicability of this approach.

This work is partially supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (C)(2)14550447.

2. Particle Filters

Consider the following nonlinear state space model.

$$x_{k+1} = f(x_k, v_k) \quad (1)$$

$$y_k = g(x_k, w_k) \quad (2)$$

where x_k and y_k are the state variable and observation, respectively, f and g are known possibly nonlinear functions, v_k and w_k are independently identically distributed (i.i.d.) system noise and observation noise sequences, respectively. We assume v_k and w_k are mutually independent. Problem to be considered here is to find the best estimate of the state variable x_k in some sense based on the all available data of observations $y_{1:k} = \{y_1, y_2, \ldots, y_k\}$. We can solve the problem by calculating the posterior pdf of the state variable x_k of time instant k based on all the available data of observation sequence $y_{1:k}$.

The posterior pdf $p(x_k|y_{1:k})$ of x_k based on the observation sequence $y_{1:k}$ satisfies the following recursion:

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1}$$

(Chapman-Kolmogorov equation) (3)

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \quad \text{(Bayes' rule)} \quad (4)$$

with a prior pdf $p(x_0|y_0) \equiv p(x_0)$ of the initial state variable x_0. Here normalizing constant

$$p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k$$

depends on the likelihood $p(y_k|x_k)$, which is determined by the observation equation (2).

Since a closed form solution is not admitted except in very restrictive cases such as linear Gaussian state space models, where the well-known Kalman filter [1], [17] can be applied, some approximations should be introduced. The most popular approximation approach is the extended Kalman filter (EKF) [12],[9]:

$$
\begin{aligned}
\hat{x}_{k|k-1} &= f(\hat{x}_{k-1|k-1}) \\
P_{k|k-1} &= \tilde{A}_k P_{k-1|k-1} \tilde{A}_{k-1|k-1}^T + Q \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - g(\hat{x}_{k|k-1})) \\
P_{k|k} &= (I - K_k \tilde{C}_k) P_{k|k-1} \\
K_k &= P_{k|k-1} \tilde{C}_k^T (\tilde{C}_k P_{k|k-1} \tilde{C}_k^T + R)^{-1} \\
\tilde{A}_k &= \left. \frac{df(x)}{dx} \right|_{x=\hat{x}_{k-1|k-1}} \\
\tilde{C}_k &= \left. \frac{dg(x)}{dx} \right|_{x=\hat{x}_{k|k-1}}.
\end{aligned} \tag{5}
$$

This is applicable to nonlinear models with additive Gaussian noise and uses a linearization technique based on a first order Taylor expansions of the nonlinear system and observation equations about the current estimate. However, it approximates the posterior pdf to be Gaussian. If the true density is non-Gaussian, then a Gaussian can never describe it well. In such cases, approximate grid-based filters and particle filters will yield an improvement in performance. They approximate the true posterior pdf with the following weighted empirical distribution of a set of $n \gg 1$ samples $\{x_k^{(i)},\ (i = 1, \ldots, n)\}$ called as particles or discrete grids with associated importance weights $\{w_k^{(i)},\ (i = 1, \ldots, n)\}, w_k^{(i)} > 0,\ \sum_{i=1}^n w_k^{(i)} = 1$,

$$p(x_k|y_{1:k}) \approx \sum_{i=1}^n w_k^{(i)} \delta(x_k - x_k^{(i)}) \tag{6}$$

where $\delta(\cdot)$ is a function such that $\delta(x) = 1$ for $x = 0$ and $\delta(x) = 0$ otherwise.

Here, the particles are generated and associated weights are chosen using the principle of "importance sampling" [6]:

Suppose $p(x) \propto \pi(x)$ is a pdf from which it is difficult to draw samples, but for which $\pi(x)$ can be evaluated (and so $p(x)$). Let $x^{(i)}$ $(i = 1, \ldots, n)$ be samples that are easily generated from a pdf $q(x)$, called an importance density. Then a weighted approximation to the density $p(x)$ is given by

$$p(x) \approx \sum_{i=1}^n w^{(i)} \delta(x - x^{(i)}) \tag{7}$$

with the normalized weight of the i-th particle

$$w^{(i)} \propto \frac{\pi(x^{(i)})}{q(x^{(i)})} \tag{8}$$

So, if the samples $x_k^{(i)}$ in (6) were drawn from an importance density $q(x_k^{(i)}|y_{1:k})$, then the associated normalized weights are defined as by (8) to be

$$w_k^{(i)} \propto \frac{p(x_k^{(i)}|y_{1:k})}{q(x_k^{(i)}|y_{1:k})}. \tag{9}$$

If the importance density $q(x_k|y_{1:k-1})$ is chosen to factorize such that

$$q(x_k|y_{1:k}) = q(x_k|x_{k-1}, y_{1:k}) q(x_{k-1}|y_{1:k-1}). \tag{10}$$

Then we can obtain samples $x_k^{(i)}$ by augmenting each of the existing samples $x_{k-1}^{(i)}$ sampled from the importance density $q(x_{k-1}|y_{1:k-1})$ with the new state sampled from $q(x_k|x_{k-1}, y_{1:k})$. Noting that the posteriori pdf can be rewritten using Bayes' rule as

$$
\begin{aligned}
p(x_k|y_{1:k}) &= \frac{p(y_k|x_k, y_{1:k-1}) p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \\
&= \frac{p(y_k|x_k, y_{1:k-1}) p(x_k|x_{k-1}, y_{1:k-1})}{p(y_k|y_{1:k-1})} p(x_{k-1}|y_{1:k-1}) \\
&\propto p(y_k|x_k) p(x_k|x_{k-1}) p(x_{k-1}|y_{1:k-1})
\end{aligned} \tag{11}
$$

we have

$$
\begin{aligned}
w_k^{(i)} &\propto \frac{p(y_k|x_k^{(i)}) p(x_k^{(i)}|x_{k-1}^{(i)}) p(x_{k-1}^{(i)}|y_{1:k-1})}{q(x_k^{(i)}|x_{k-1}^{(i)}, y_{1:k}) q(x_{k-1}^{(i)}|y_{1:k-1})} \\
&= w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i)}) p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{k-1}^{(i)}, y_{1:k})}.
\end{aligned} \tag{12}
$$

The particle filter with these steps is called "Sequential Importance Sampling Particle Filter" (SIS).

It is known that the SIS filter suffers from the degeneracy phenomenon, where all but one of the normalized importance weights are very close to zero after a few iterations. By this degeneracy, a large computational effort is wasted to updating trajectories whose contribution to the final estimate is almost zero. In order to prevent this phenomenon, several modifications have been introduced. Among them, resampling process is used often. Its idea is to eliminate trajectories whose normalized importance weights are small and to concentrate upon the trajectories with larger weights. It involves generating new grid points $x_k^{*(i)}$ $(i = 1, \ldots, n)$ by resampling from the grid approximation (6) randomly with probability

$$\Pr(x_k^{*(i)} = x_k^{(j)}) = w_k^{(j)} \tag{13}$$

and the weights are reset to $w_k^{*(i)} = 1/n$. The choice of resampling is done by using some criterion such as the effective sample size N_{eff} introduced in [14],

$$N_{eff} = \frac{n}{1 + \text{Cov}_{q(\cdot|y_{1:k})}(w_k(x_k^{(i)}))}, \tag{14}$$

whose estimate is given by

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^n (w_k^{(i)})^2} \tag{15}$$

with the associated normalized weight $w_k^{(i)}$. This indicates how many samples in the particle cloud that actually contribute to the support of the pdf approximation. We can resample if the effective number of samples is less than a predefined threshold $N_{thres} < 1$. Particle filter with this resampling process is called "Sampling Importance Resampling Particle Filter" (SIR) [5], [2].

3. Evolutionary Computation and Evolution Strategies Based Particle Filter

A novel particle filter called Evolution Strategies Based Particle (ESP) filter was proposed to prevent the degeneration in SIS filter [19], by recognizing the similarity and the difference between the importance sampling and resampling processes in SIR filter and evolution processes in Evolution Strategies (ES) originated by Rechenberg and Schwefel [16]. In this section, they are briefly reviewed.

3.1. Evolutionary Computation

Evolutionary computation approach is a computational model of natural evolutionary processes as key elements in the design and implementation of computer-based problem solving systems. A variety of evolutionary computation approaches such as 'Evolutionary Programming' (EP) [8], 'Evolution Strategies' (ES) [16], 'Genetic Algorithm' (GA) [11], and 'Genetic Programming' (GP) [15] have been proposed and studied. Extensive survey and comments are given in [4],[3],[7]. The common conceptual base is simulating the evolution of individuals (candidate solutions) via processes of selection and perturbation. These processes depend on the perceived performance (fitness) of the individuals as defined by the environments.

Evolutionary computation approach maintains a population of structures that evolve according to rules of selection and other operators, such as recombination and mutation. Each individual is evaluated, receiving a measure of its fitness in the environment. *Selection (reproduction)* focuses attention on high-fitness individuals, thus exploiting the available fitness information. *Recombination* (also refer to as *crossover*) and *mutation* perturb those individuals, providing general heuristics for exploration. Here we explain Evolution Strategies (ES) briefly. ES is developed by Rechenberg and Schwefel [16] to solve hydrodynamic problems. It is applied to continuous function optimization in real-valued n-dimensional space. Mutation is applied more often to the solution rather than crossover. The simplest method can be implemented as follows: Let $x^{(k)} = (x_1^{(k)}, \cdots x_n^{(k)}) \in R^n, \ \ (k = 1, \cdots, \mu)$ be each individual in the population.

3.1.1 Generation of initial population

We generate an initial population of parent vectors $\{x^{(k)}, \ (k = 1, \cdots, \mu)\}$ randomly from a feasible range in each dimension.

3.1.2 Evolution operations

1. Crossover

This process allows for mixing of parental information while passing it to their descendants. A typical crossover rule is

$$x'_j = x_{S,j} + \chi \cdot (x_{T,j} - x_{S,j}) \qquad (16)$$

where S and T denote two parent individuals selected at random from the population and $\chi \in [0, 1]$ is a uniform random or deterministic variable. The index j in x'_j indicates j-th component of new individuals. This is a similar operator used in differential evolution [18].

2. Mutation

This process introduces innovation into the population. It is realized by following additive process,

$$\begin{aligned} \sigma'_j &= \sigma_j \exp(\tau' N(0,1) + \tau N_j(0,1) \\ x''_j &= x'_j + \sigma'_j N_j(0,1) \end{aligned} \qquad (17)$$

Here, $N(0, 1)$ denotes a realization of normal random variable with mean and unit variance, $N_j(0, 1)$ denotes random variable sampled anew for counter j of normal random variable with mean and unit variance and σ_j denote the mean step size. The factors τ and τ' are chosen depending the population size μ [4]. In this approach, small variations are much more frequent than larger variations, expressing the state of affairs on the phenotypic level in nature.

3. Selection

This is the completely deterministic process choosing the individuals of higher fitness out of the union of parents and offspring or offspring only to form the next generation in order to evolve towards better search region.

- $(\mu + \lambda)$-selection

This creates λ offspring from μ parents and selected the μ best individuals out of the union of parents and offspring.

- (μ, λ)-selection

This creates λ offspring from μ parents and selected the μ best individuals out of offspring $(\lambda \geq \mu)$.

3.2. Evolution Strategies Based Particle Filter

It can be seen that SIR and ES have similarities; both the importance sampling process in SIR filter and mutation process in ES give perturbation to the parent individuals $x_{k-1}^{(i)}$ with extrapolation by $f(x_{k-1}^{(i)})$, and both resampling process in SIR filter and selection process in ES selects offspring among the perturbed individuals. However, there is a difference between them, i.e., resampling in SIR is carried out randomly and the weights are reset as $1/n$, while the selection in ES is deterministic and the fitness function is never reset. Hence, by replacing the resampling process in SIR by the selection process in ES, we have derived a new particle filter as follows.

Based on the particles $x_{k-1}^{(i)} \ \ (i = 1, \ldots, n)$ sampled from the importance density $q(x_{k-1}|y_{1:k-1})$, we generate ℓ $x_k^{(i,j)}, \ (j = 1, \ldots, \ell)$ sampled from the importance density function $q(x_k|x_{k-1}^{(i)}, y_{1:k})$. Corresponding weights $w_k^{(i,j)}$ are evaluated by

$$w_k^{(i,j)} = w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i,j)}) p(x_{k-1}^{(i,j)}|x_{k-1}^{(i)})}{q(x_k^{(i,j)}|x_{k-1}^{(i)}, y_{1:k})} \qquad i = 1, \ldots, n, \ j = 1, \ldots, \ell \qquad (18)$$

From the set of $n\ell$ particles and weights $\{x_k^{(i,j)}, w_k^{(i,j)}, \ (i = 1, \ldots, n, \ \ j = 1, \ldots, \ell)\}$, we choose n sets with the larger weights, and set as $x_k^{(i)}, w_k^{(i)} (i = 1, \ldots, n)$. This process corresponds to $(n, n\ell)$-selection in ES. Hence, we call this particle filter using $(n, n\ell)$-selection in ES as Evolution Strategies based particle filter Comma (ESP(,)). When we add the particles $x_k^{(i,0)} = f(x_{k-1}^{(i)}), \ (i = 1, \ldots, n)$ in addition to $n\ell$ $x_k^{(i,j)}, \ (i = 1, \ldots, n, \ j = 1, \ldots, \ell)$ sampled from the importance density function $q(x_k|x_{k-1}^{(i)}, y_{1:k})$ as above and evaluate the weights $w_k^{(i,j)}, (i = 1, \ldots, n, j = 0, \ldots, \ell)$ by (18), and then choose n sets of $(x_k^{(i)}, w_k^{(i,j)})$ with larger weights from the ordered set of $n(\ell + 1)$ particles $\{x_k^{(i,j)}, w_k^{(i,j)}, (i = 1, \ldots, n, \ j = 0, \ldots, \ell)\}$, we can obtain another ESP filter. Since this ESP filter uses the selection corresponding to $(n + n\ell)$-selection in ES, we can call this filter as Evolution Strategies based particle filter Plus (ESP(+)). The algorithms are summarized in Fig.1.

Procedure *ESP*

```
For   k = 0
      i = 1,...,n, sample x_0^(i) ~ q(x_0|y_0);
      i = 1,...,n, evaluate the weight
        w_0^(i) = p(y_0|x_0^(i)) p(x_0^(i)) / q(x_0^(i)|y_0).
For   k ≥ 1
      i = 1,...,n
        set x_k^(i,0) = f(x_{k-1}^(i))
      j = 1,...,ℓ
        sample x̃_k^(i,j) ~ q(x_k|x_{k-1}^(i), y_{1:k});
      i = 1,...,n and j = 0,1,...,ℓ,
        evaluate the weight
        w_k^(i,j) = w_{k-1}^(i) (p(y_k|x̃_k^(i,j)) p(x̃_k^(i,j)|x_{k-1}^(i)) / q(x̃_k^(i)|x̃_k^(i), y_{1:k})).
      Sort the set of pairs {x̃_k^(i,j), w_k^(i,j) (i = 1,
      ...,n, j = 0,1,...,ℓ)}
      by the size of w_k^(i,j) in descending
      order.
      Take the first n x_k^(i) from the ordered
      set {x̃_k^(i), w̃_k^(i)}.
      i = 1,...,n, normalize the weight
        w_k^(i) = w_k^(i) / Σ_{i=1}^n w_k^(i).
      Let p(x_k|y_{1:k}) ≈ Σ_{i=1}^n w̃_k^(i) δ(x_k − x_k^(i))
```

Fig. 1. Algorithm for ESP filters. ESP(+): with the underlined part; ESP(,): without the underlined part

4. Simultaneous State and Parameter Estimation by Evolution Strategies Based Particle Filter

The ESP filter is applied here to simultaneous state and parameter estimation of nonlinear systems. Consider the nonlinear state space model (1) with unknown parameter θ and (2), where a posteriori pdf $p(x_k, \theta|y_{1:k})$ should be approximated to estimate state and parameter simultaneously, Application of Bayes' rule (4) provides

$$p(x_{k+1}, \theta|y_{1:k+1}) \quad \propto \quad p(y_{k+1}|x_{k+1}, \theta)p(x_{k+1}|\theta, y_{1:k+1}) \times p(\theta|y_{1:k+1})$$

Since the form of the theoretical pdf $p(\theta|y_{1:k})$ is not known for unknown parameter case, we replace θ by θ_k at time k, and simply include θ_k in an augmented state vector $\mathit{x}_k = (x_k, \theta_k)^T$, where θ_k evolves as

$$\theta_{k+1} = \theta_k + \eta_k \tag{19}$$

and η_k is a normal random disturbance with zero-mean and very small variance. Then approximation of the true posteriori pdf is given by

$$p(\mathit{x}_k|y_{1:k}) \approx \sum_{i=1}^{n} w_k^{(i)} \delta(\mathit{x}_k - \mathit{x}_k^{(i)}) \tag{20}$$

If particles $\mathit{x}_k^{(i)}$ in (20) were drawn from an importance density

$$q(\mathit{x}_k^{(i)}|\mathit{x}_{k-1}^{(i)}, y_{1:k}) \quad = \quad q_x(x_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \times q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \tag{21}$$

with importance densities for x_k and θ_k, $q_x(x_k^{(i)}|x_{k-1}^{(i)},\ \theta_{k-1}^{(i)}, y_{1:k})$ and $q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})$, and the associated normalized weights are evaluated by

$$w_k^{(i)} \quad \propto \quad w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i)}, \theta_k^{(i)})}{q_x(x_k^{(i)}, \theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})} \times \frac{p(x_k^{(i)}, \theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)})}{q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})}. \tag{22}$$

Then, the SIS, SIR and ESP filters are defined as above.

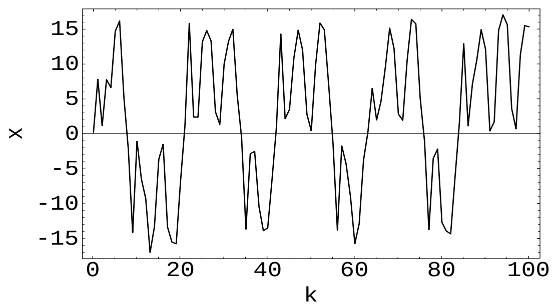
4.1. Numerical Examples

Numerical simulation are carried out to exemplify the applicability of the proposed ESP filter. First, we consider the following nonlinear state space model

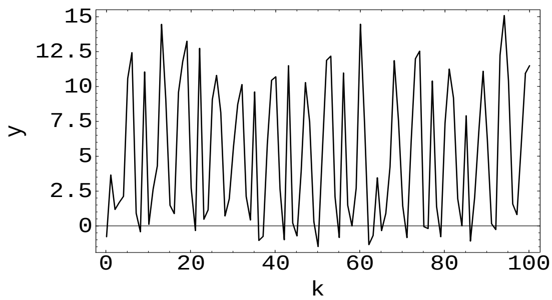
$$x_k = \frac{x_{k-1}}{2} + \frac{\theta x_{k-1}}{1 + x_{k-1}^2} + 8\cos(1.2k) + v_k$$
$$= f(x_{k-1}, \theta) + v_k \tag{23}$$

$$y_k = \frac{x_k^2}{20} + w_k \tag{24}$$

where v_k and w_k are i.i.d. zero-mean normal random variables with variance 10 and 1, respectively, and value of the parameter θ is known to be 25. The normal distribution with mean $f(x_{k-1}^{(i)})$ and variance 10 is chosen as the importance density $q(x_k|x_{k-1}^{(i)}, y_{1:k})$. A sample behavior of the true state and corresponding observation processes is shown in Fig.2. Sample paths of the estimates by the



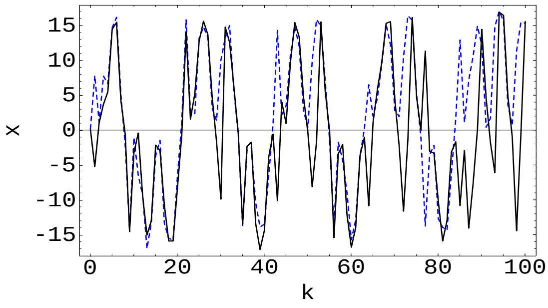
(a) True state



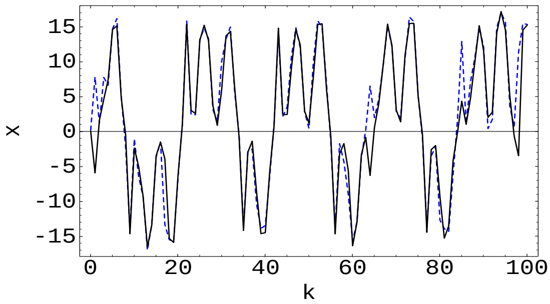
(b) Observation

Fig. 2. Sample behavior of state and observation processes

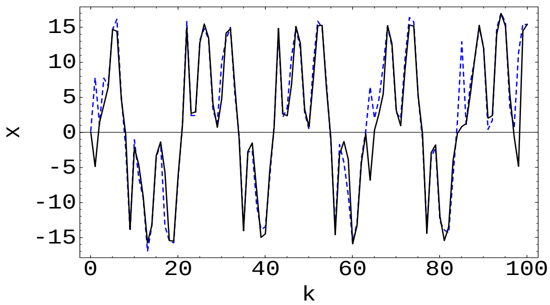
particle filters (SIS ($n = 200$), SIR ($n = 100, N_{eff} = 50$), and the proposed ESP(,) ($n = 100, \ell = 2$)) and ESP(+) ($n = 100, \ell = 1$)) are given in Fig.3, and that of EKF as well for comparison. Particle



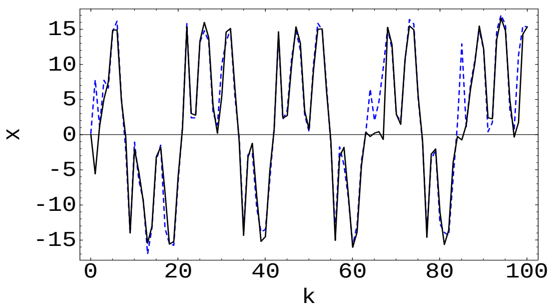
(a) Estimate by SIS



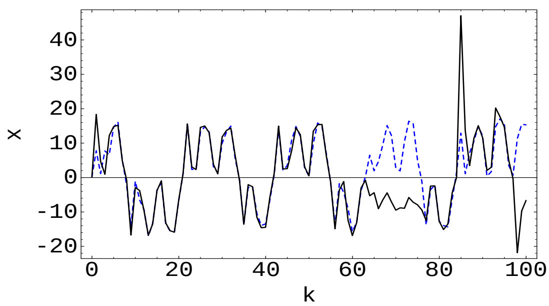
(b) Estimate by SIR



(c) Estimate by ESP(,)



(d) Estimate by ESP(+)



(e) Estimate by EKF

Fig. 3. Sample paths of state estimates (solid line: estimate, dashed line: true state)

filters, especially SIR and proposed ESP filters, show well behaviors in nonlinear state estimation, while the estimate by EKF cannot follow the true state.

Figure 4 shows the 2-dimensional plots of squared errors at $k = 1000$ and processing time [s] until $k = 1000$. ESP filters show

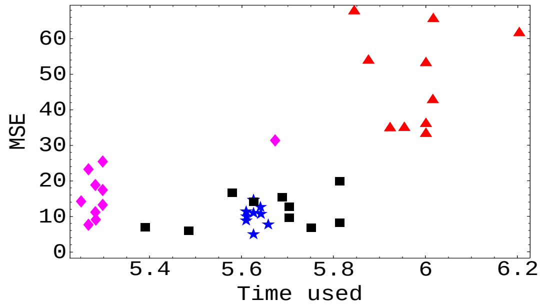


Fig. 4. Squared estimation errors and processing time (triangle: SIS, box: SIR, star: ESP(,), diamond: ESP(+))

similar performance as SIR both in squared estimation errors and processing time, and their fluctuations are smaller than SIR. It implies that ESP filters are more stable than SIR.

Next, we consider the unknown parameter case where the true value of $\theta = 25$ in (23) is not known. Here, only the results by ESP(,) with the importance densities $q_x(x_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \sim \mathcal{N}(f(x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, 10)$ and $q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \sim \mathcal{N}(\theta_{k-1}^{(i)}, 0.01)$ are shown in Fig.5 since the EKF does not work as before. Though the estimate approach to the true ones, the convergence speed is slow and the filter leaves much for improvement. For examples, better choice of design parameters n, N_{eff} and ℓ and choice of evolution operations should be pursued since the estimation performance, of course, depends on the choice of them.

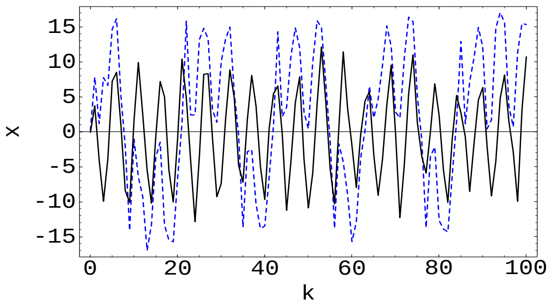
5. Conclusions

The novel particle filter, which was developed by recognizing the similarity and the difference between the importance sampling and resampling processes in the SIR filter and mutation and selection processes in ES and substituting (μ, λ)-selection in ES into resampling process in SIR, is applied to simultaneous state and parameter estimation of nonlinear state space models. It works stably and provides small mean square errors compared to EKF filter. Application of other evolution operations such as crossover and modification of mutation will have the potential to create much higher performance particle filters.

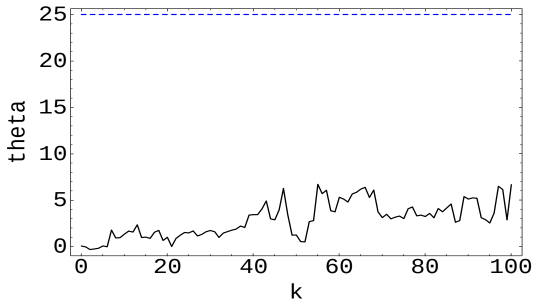
References

[1] B. D. O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, NJ, 1979.

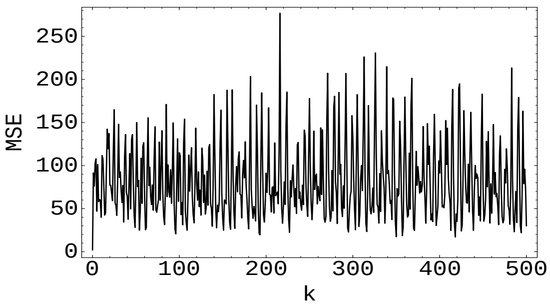
[2] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking," *IEEE Trans. on Signal Processing*, vol. SP-50, pp. 174–188, 2002.



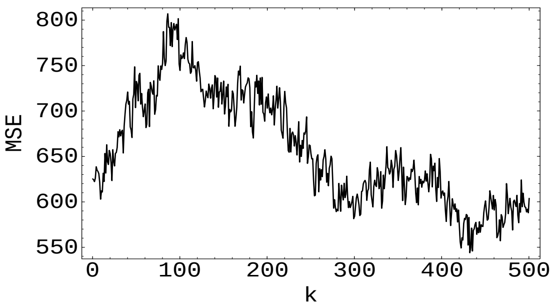
(a) A sample path of the state estimate



(b) A sample path of the parameter estimate



(c) MSE for the state estimate (10 simulation runs)



(d) MSE for the parameter estimate (10 simulation runs)

Fig. 5. Simulation results in simultaneous state and parameter estimation by ESP(,)

[3] T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press, London, 1996.

[4] T. Bäck and H.-P. Schwefel, “An overview of evolutionary algorithms for parameter optimization,” *Evolutionary Computation*, vol/ 1, pp. 1–23, 1993.

[5] A. Doucet, “On sequential simulation-based methods for Bayesian filtering,” *Technical Report CUED/F-INFENG/TR 310*. Department of Engineering, Cambridge University, 1998.

[6] A. Doucet, N. de Freitas and N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, NY, 2001.

[7] D. B. Fogel, *Evolutionary Computation*. IEEE Press, NJ, 1995.

[8] L. J. Fogel, A. J. Owens and M. J. Walsh, “Artificial intelligence through a simulation of evolution,” *Proc. of the 2nd Cybernetics Science Symp.*, pp. 131–155, 1965, reprinted in D. B. Fogel (eds.), *Evolutionary Computation, The Fossil Record*, pp. 230–254, IEEE Press, NY.

[9] G. C. Goodwin and J. C. Agüero, “State parameter estimation for linear and nonlinear systems,” *Proc. 7th International Conf. on Control, Automation, Robotics and Vision*, 2002.

[10] N. Gordon, D. Salmond and A. Smith, “A novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEE Proc. on Rader and Signal Processing*, vol. 140, pp. 107-113, 1993.

[11] J. H. Holland, *Adaptation in Natural and Artificial Systems*. MIT Press, 1992.

[12] A. H. Jazwinski, *Stochastic Process and Filtering Theory*. Academic Press, NY, 1970.

[13] Kalman, R. and R. Bucy (1961). New results in linear filtering and prediction theory. *J. of Basic Engineering, Trans. ASME Series D*, vol. 83, pp. 95–108.

[14] A. Kong, J. S. Liu and W.H. Wong, “Sequential imputations and Bayesian missing data problems,” *J. of American Statistical Association*, vol. 89, 1994, pp. 278–288.

[15] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.

[16] H.-P. Schwefel, *Evolution and Optimum Seeking*, J. Wiley, NY, 1995.

[17] H. Sorenson, *Kalman Filtering: Theory and Application*, IEEE Press, NJ, 1985.

[18] R. Storn and K. Price, “Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces,” *Technical Report TR-95-012*. ICSI, 1995.

[19] K. Uosaki, Y. Kimura and T. Hatanaka, “State estimation of nonlinear stochastic systems by evolution strategies based particle filter,” *Proc. 9th Int. Symposium on Artificial Life and Robotics*, pp. 406–409, 2004.