# MULTIPLE PARTICLE FILTERING WITH IMPROVED EFFICIENCY AND PERFORMANCE

*Petar M. Djurić and Mónica F. Bugallo*

Department of Electrical and Computer Engineering
Stony Brook University, Stony Brook, NY 11794
Email: {petar.djuric,monica.bugallo}@stonybrook.edu

## ABSTRACT

Particle filtering has been widely accepted as an important methodology for processing data represented by state-space models characterized by nonlinearities and/or non-Gaussianities. It is also well documented that particle filtering deteriorates quickly in performance when the dimension of the tracked state becomes large. This limits its application in many science/engineering problems. Previously we have proposed a way of alleviating this deficiency based on the use of multiple particle filtering. According to the approach, a number of particle filters are assigned to track different subsets of the state with time. In this paper, we propose a new method for accurate and efficient implementation of multiple particle filtering. We provide simulation results that demonstrate the performance of the new method.

***Index Terms*—** multiple particle filtering, state-space models, high-dimensional systems

## 1. INTRODUCTION

An important area of signal science is the sequential processing of observations that are represented by state-space models. The objective there is to obtain an estimate of a hidden state (signal) given available observations. The state follows a dynamic model and the observations are functions of the hidden state and some random perturbations that can be interpreted as noise. It is well known that when the state-space model is linear and with additive Gaussian noise, the optimal solution of the tracking is obtained by Kalman filtering [1]. The problem gets much more challenging when the model becomes nonlinear and/or the noises in the system are non-Gaussian. A class of methods that has risen to address these challenges is known as particle filtering. It is based on tracking the relevant densities describing the unknown state by discrete random measures composed of particles and their weights [2].

It is now well understood that particle filtering produces an approximation error that usually increases exponentially with the dimension of the state variable [3]. This problem is known as curse-of-dimensionality. In many science and engineering fields including geophysical sciences [4], evolutionary biology [5], robotics [6], and computational neuroscience [7], models with high-dimensional states are common and the use of particle filtering there may often be prohibited. There are two key issues that cause the curse-of-dimensionality. One is the difficulty of generating good samples in a high-dimensional space. The other is the computation of the likelihoods [8].

This work was supported by NSF under Awards CCF-1320626 and CCF-0953316.

Recently, however, it has been argued that in principle, one may develop local particle filtering algorithms whose approximation error is dimension-free [9]. This is the same argument that we have used in a series of papers where we have studied multiple particle filtering [10, 11, 12, 13, 14, 15]. Namely, most of the processes of interest have an intrinsic sparsity in the sense that any given small subset of states of the system depends only on another very small subset of the states. A similar statement can be made about the measurements. In other words, a given measurement is a function of only a small subset of the states. This sparsity opens up possibilities for particle filtering that can alleviate the curse-of-dimensionality. One of them is to use a number of particle filters, where each particle filter is assigned to track a separate small subset of the unknowns. In doing so, the various particle filters communicate information about the states that they track. Since these individual particle filters operate in much smaller dimensional state-spaces, they are not affected by the overall high-dimensional state-space.

The main contribution of the paper is in the novel way of implementing the information provided by the relevant filters to a given filter. The filters only communicate means and covariances of the states that they track. A given filter needs information from two sets of particle filters. The first set are filters whose states are needed by the given filter for propagation of its particles. The second set is composed of filters whose states are necessary to the given filter to compute the likelihoods of its particles. The novelty in the proposed approach is in the way how the communicated information by the filters is efficiently used.

The paper is organized as follows. In Section 2, we formulate the problem. The novel scheme is described in Section 3. Simulation results are presented in Section 4. Final conclusions are made in Section 5.

## 2. PROBLEM FORMULATION

We study dynamic systems that are modeled according to

$$x_t = f(x_{t-1}, u_t), \tag{1}$$

$$y_t = g(x_t, v_t), \tag{2}$$

where $t = 1, 2, \cdots$ is a time index, $x_t \in \mathbb{R}^{d_x}$ is the state of the system that evolves as a first-order Markov process, $y_t \in \mathbb{R}^{d_y}$ is an observation at time $t$, $f(\cdot)$ and $g(\cdot)$ are known functions which in general are nonlinear, and $u_t \in \mathbb{R}^{d_u}$ and $v_t \in \mathbb{R}^{d_v}$ are random perturbations. The symbol $d_z$ represents the dimension of the variable $z$ in the subscript of $d$. The probability distribution functions of $u_t$ and $v_t$ are assumed known, and $u_t$ and $v_t$ are independent for all $t$. Furthermore, the samples of $u_t$ and $v_t$ are independent from their previous samples.

We are interested in systems with high-dimensional states where there is sparsity in the interdependence among the states, and among the observation and the states. In simple terms, we study systems where, for example, in (1) the first state depends only on itself and a very few of the other states from the previous time instant. Similar relationships exist for the other states. Likewise, an element from $y_t$ depends only on a small subset of the states. To emphasize this, we decompose the system in (1) into $K$ subsystems and write

$$x_{k,t} = f_k(x_{k,t-1}, q_{k,t-1}, u_{k,t}), \quad (3)$$

$$y_{k,t} = g_k(x_{k,t}, z_{k,t}, v_{k,t}), \quad (4)$$

where $k = 1, 2, \cdots K$, and where the subsystems do not share states but may share observations. The function $f_k(x_{k,t-1}, q_{k,t-1}, u_{k,t})$ shows how the state $x_{k,t}$ evolves with time and $q_{k,t-1}$ is a small subset of the remaining states of the system that are needed for the evolution. The symbol $y_{k,t}$ represents the collection of all the observations which are functions of $x_{k,t}$ (and possibly other states). The function $g_k(x_{k,t}, z_{k,t}, v_{k,t})$ represents the signal in the observation, where the signal, in general, depends not only on the desired state $x_{k,t}$ but also on a subset $z_{k,t}$ of the remaining states of the system. We reiterate that the substate vectors $x_{i,t}$ and $x_{j,t}$ with $i \neq j$ do not share elements but the vectors $y_{i,t}$ and $y_{j,t}$ with $i \neq j$ may share them.

The aim is to sequentially estimate the marginal posteriors of $x_{k,t}, p(x_{k,t}|y_{k,1:t})$. Thus, instead of one filter that would process all the data, we want to use $K$ filters, each of them assigned to estimate the marginal posterior of a different state vector $x_{k,t}$. Each of the filters operates in a state-space of dimension $d_{x_k}$, which is much smaller than $d_x$. Since the functions $f_k(\cdot)$ and $g_k(\cdot)$ are in general nonlinear, we want to estimate marginal posteriors with particle filtering. Given that the overall system contains many particle filters, we refer to it as a multiple particle filtering system.

## 3. MULTIPLE PARTICLE FILTERING WITH IMPROVED EFFICIENCY AND PERFORMANCE

We first provide the necessary operations that each filter needs to implement at every time instant $t$. We focus on the $k$th filter. Its objective is to sequentially obtain $p(x_{k,t}|y_{k,1:t})$ from $p(x_{k,t-1}|y_{k,1:t-1})$. We write

$$p(x_{k,t}|y_{k,1:t}) \propto p(y_{k,t}|x_{k,t})p(x_{k,t}|x_{k,t-1}, y_{k,1:t-1})$$
$$= p(y_{k,t}|x_{k,t}) \int p(x_{k,t}|x_{k,t-1})p(x_{k,t-1}|y_{k,1:t-1})\mathrm{d}x_{k,t-1}, \quad (5)$$

where $\propto$ stands for "proportional to." In a particle filtering scheme, the posterior $p(x_{k,t-1}|y_{1:t-1})$ is represented by a discrete random measure given by

$$p_M(x_{k,t-1}|y_{k,1:t-1}) = \sum_{k=1}^{M} w_{k,t-1}^{(m)} \delta\left(x_{k,t-1} - x_{k,t-1}^{(m)}\right), \quad (6)$$

where $\delta(\cdot)$ is the Dirac delta function, $x_{k,t-1}^{(m)}$ are particles of $x_{k,t-1}$, $m$ is the index of the particle, $M$ is the total number of particles, and $w_{k,t-1}^{(m)}$ is the weight corresponding to $x_{k,t-1}^{(m)}$. This makes the implementation of (5) rather easy.

The problem for implementation of the previous filter is that its transition probability density function (pdf) $p(x_{k,t}|x_{k,t-1})$ and likelihood $p(y_{k,t}|x_{k,t})$ do not depend only on $x_{k,t-1}$ and $x_{k,t}$, respectively. This entails that the different particle filters need to communicate information among themselves. In particular, since the $k$th particle filter uses a transition pdf $p(x_{k,t}|x_{k,t-1}, q_{k,t-1})$, it needs to get information from the particle filters that are in charge of the states that form $q_{k,t-1}$. Similarly, due to the form of $p(y_{k,t}|x_{k,t}, z_{k,t})$, this filter needs information from all the filters whose states are elements of $z_{k,t}$.

In this paper, we investigate particle filters and therefore we have a choice that the necessary exchanged information is provided in terms of particles (and possibly weights) of the states or in terms of moments of the states. As the aim of this paper is to study an efficient multiple particle filtering system, we examine the case when these filters only communicate the mean and the covariance of their state estimates. In a system where there may be dozens or hundreds of particle filters, this choice would prove considerably simpler.

It is well known that a standard particle filtering scheme is composed of three steps, particle propagation, weight computation and resampling. Here we focus on the steps of particle propagation and weight computation. The propagation is critical for good performance, because one needs to generate "good" particles, that is, particles from the space of $x_{k,t}$ where the probability masses are not negligible. A standard procedure, which is rather simple and works remarkably well is to draw samples from the transition density, i.e.,

$$x_{k,t}^{(m)} \sim p(x_{k,t}|x_{k,t-1}^{(m)}). \quad (7)$$

In the case of multiple particle filtering, the problem of this scheme is that we do not have the form of $p(x_{k,t}|x_{k,t-1}^{(m)})$ but instead we have to work with $p(x_{k,t}|x_{k,t-1}^{(m)}, q_{k,t-1})$. A formal way of obtaining $p(x_{k,t}|x_{k,t-1}^{(m)})$ from $p(x_{k,t}|x_{k,t-1}^{(m)}, q_{k,t-1})$ is to use

$$p(x_{k,t}|x_{k,t-1}^{(m)}) = \int p(x_{k,t}|x_{k,t-1}^{(m)}, q_{k,t-1})p(q_{k,t-1})\mathrm{d}q_{k,t-1}, \quad (8)$$

where $p(q_{k,t-1})$ is the distribution of the states $q_{k,t-1}$ at time $t-1$. We assume here that this distribution is Gaussian with mean and covariance provided by the particle filters that feed the $k$th filter with information, which we denote by $\mathcal{N}(\mu_k, \Sigma_k)$. An alternative way of dealing with this problem is to work with (3), and treat both $q_{k,t-1}$ and $u_{k,t}$ as random quantities. Then it may be possible to obtain the required transition pdf $p(x_{k,t}|x_{k,t-1})$ directly. Once obtained, one can generate the particles from (7). We provide a simple example to explain the procedure. Let

$$x_{1,t} = ax_{1,t-1} + bx_{2,t-1} + u_{1,t}, \quad (9)$$

where all the variables are scalars. If we treat $x_{2,t-1}$ as a random variable, we rewrite (9) as

$$x_{1,t} = ax_{1,t-1} + \tilde{u}_{1,t}, \quad (10)$$

where

$$\tilde{u}_{1,t} = bx_{2,t-1} + u_{1,t}, \quad (11)$$

can be viewed as a modified noise variable. We need $p(x_{1,t}|x_{1,t-1}^{(m)})$, and so if we assume that

$$x_{2,t-1} \sim \mathcal{N}(\mu_{2,t-1}, \sigma_{2,t-1}^2), \quad (12)$$

$$u_{1,t} \sim \mathcal{N}(0, \sigma_u^2), \quad (13)$$

we get that the "noise" $\tilde{u}_{1,t}$ is Gaussian characterized as

$$\tilde{u}_{1,t} \sim \mathcal{N}(b\mu_{2,t-1}, b^2\sigma_{2,t-1}^2 + \sigma_u^2). \quad (14)$$

Thus, the pdf $p(x_{1,t}|x_{1,t-1}^{(m)})$ is Gaussian, or more precisely,

$$p(x_{1,t}|x_{1,t-1}^{(m)}) = \mathcal{N}(ax_{t-1}^{(m)} + b\mu_{2,t-1}, b^2\sigma_{2,t-1}^2 + \sigma_u^2). \quad (15)$$

Thus, one can proceed with direct sampling of $x_{1,t}^{(m)}$ from the pdf in (15).

Often, however, it will be difficult to obtain a closed form of $p(x_{k,t}|x_{k,t-1}^{(m)})$ in this manner. In that case, the procedure is a bit more computationally intensive in that one first generates samples $q_{k,t-1}^{(m)}$ from a Gaussian whose parameters are obtained from the other particle filters and then draws $x_{k,t}^{(m)}$ from $p(x_{k,t}|x_{k,t-1}^{(m)}, q_{k,t-1}^{(m)})$. We point out that we could have applied the same procedure in the above example, with the difference being that it is computationally less efficient. In its implementation we draw more particles.

The situation is more critical in the computation of the weights of the particles. Under the assumption that the particles were propagated according to $p(x_{k,t}|x_{k,t-1})$ and that resampling was performed at $t-1$, the computation of the weights amounts to evaluating

$$w_{k,t}^{(m)} \quad \propto \quad p(y_{k,t}|x_{k,t}^{(m)}). \quad (16)$$

Again we have a problem because we do not have the form of the likelihood $p(y_{k,t}|x_{k,t}^{(m)})$ but instead $p(y_{k,t}|x_{k,t}^{(m)}, z_{k,t})$. A straightforward way of handling the problem is to obtain the required likelihood via

$$p(y_{k,t}|x_{k,t}^{(m)}) \quad = \quad \int p(y_{k,t}|x_{k,t}^{(m)}, z_{k,t})p(z_{k,t})\mathrm{d}z_{k,t}. \quad (17)$$

An equivalent approach is to work with (4) and treat $z_{k,t}$ as a random variable. The information about $z_{k,t}$ is obtained from the relevant particle filters and it represents the moments of the propagated particles of the reporting filters. The procedure to obtain the form of $p(y_{k,t}|x_{k,t}^{(m)})$ is the same as the one described by the above example. Basically, we deal with transformation of random variables where the objective is to obtain the distribution of the new noise random variable.

Very often, the function in (4) has the form

$$y_{k,t} \quad = \quad g_k(x_{k,t}, z_{k,t}) + v_{k,t}, \quad (18)$$

where $v_{k,t} \sim \mathcal{N}(0, \Sigma_k)$. Note that we have to compute from (18) the likelihood of $x_{k,t}^{(m)}$. In the case when it is not possible to straightforwardly obtain a closed form of the likelihood, one may resort to an approximation. If we assume that the likelihood is Gaussian, we only need to obtain $\mathbb{E}(y_{k,t}|x_{k,t}^{(m)})$ and $\mathrm{Cov}(y_{k,t}|x_{k,t}^{(m)})$. We recall that our interest is in systems where the observations are nonlinear functions of the states and therefore it will not be always possible to obtain the exact likelihood. If we resort to approximated likelihood, we need to find the approximated mean and covariance. There are various ways of obtaining them, and one standard approach is to employ a Taylor expansion of the nonlinear function $g_k(x_{k,t}^{(m)}, z_{k,t})$ around the mean of $z_{k,t}$. Again, for simplicity, let $y_{k,t}$ and $z_{k,t}$ be scalars and let $\mathbb{E}(z_{k,t}) = \eta_{k,t}$ and $\mathrm{Var}(z_{k,t}) = \sigma_{k,t}^2$. Then we can find the needed mean and variance readily by [16]

$$\mathbb{E}(y_{k,t}|x_{k,t}^{(m)}) \quad \approx \quad g_k(x_{k,t}^{(m)}, \eta_{k,t}) + \frac{1}{2}g_k''(\eta_{k,t})\sigma_{k,t}^2, \quad (19)$$

$$\mathrm{Var}(y_{k,t}|x_{k,t}^{(m)}) \quad \approx \quad \left|g_k'(\eta_{k,t})\right|^2 \sigma_{k,t}^2 + \sigma_v^2, \quad (20)$$

where $g_k'(\cdot)$ and $g_k''(\cdot)$ are the first and second derivatives of $g_k(x_{k,t}^{(m)}, z_{k,t})$ with respect to $z_{k,t}$, respectively.

## 4. SIMULATION RESULTS

We present simulation results that show the validity of the new approach. In particular, we discuss two scenarios. In the first one, a linear system coupled in both state and observation equations is considered. The justification for that example is not only to exhibit the advantage of the new method over the standard particle filter but also to show its close performance to the bound imposed by the Kalman filter, which is the optimal solution for that case. The second example is a system also coupled in both state and observation equations, with nonlinearities in the observations. The obtained results clearly show a better performance of the proposed method over the standard particle filter. In both examples, $K = d_x$.

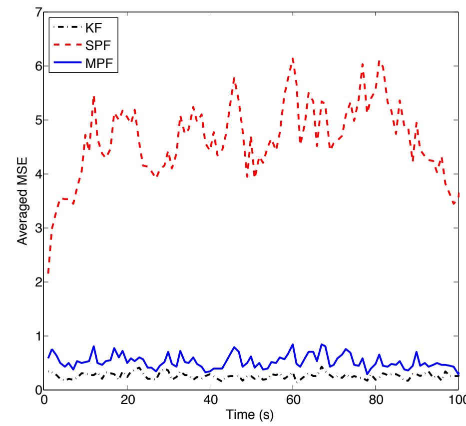


**Fig. 1**. Averaged MSE comparison for the linear example.

### 4.1. Linear example

We considered a system with the following state equations:

$$\begin{aligned} x_{1,t} &= .7x_{1,t-1} - .3x_{d_x,t-1} + u_{1,t} \\ x_{2,t} &= .7x_{2,t-1} - .3x_{1,t-1} + u_{2,t} \\ &\vdots \\ x_{d_x,t} &= .7x_{d_x,t-1} - .3x_{d_x-1,t-1} + u_{d_x,t}, \end{aligned} \quad (21)$$

where $u_{i,t}, i = 1, \cdots, d_x$ were independent and identically distributed zero-mean Gaussian perturbations with variance $\sigma_{u_i}^2 = 1$. The observations were also linear and given by

$$\begin{aligned} y_{1,t} &= 2x_{1,t} + x_{2,t} + v_{1,t} \\ y_{2,t} &= 2x_{2,t} + x_{3,t} + v_{2,t} \\ &\vdots \\ y_{d_x,t} &= 2x_{d_x,t} + x_{1,t} + v_{d_x,t}, \end{aligned} \quad (22)$$

with $v_{i,t}$ being independent zero-mean Gaussian random variables of variance $\sigma_{v_i}^2 = 1$. We note that at each time instant there were as many observations as parameters in the state and that each component of the state vector participated in two observations, i.e., at time instant $t$, the $i$th component ($i > 1$) of the state, $x_{i,t}$, contributed to $y_{i,t}$ and $y_{i-1,t}$, whereas $x_{1,t} = 1$ was present in the measurements $y_{1,t}$ and $y_{d_x,t}$.

We let the system evolve for $T = 100$ time units and set the size of the state to $d_x = 50$. We compared the Kalman filter (labeled as KF), the standard particle filter (labeled as SPF), which generated particles of dimension $d_x$ at each step, and the new multiple particle filter (labeled as MPF), which used $d_x$ filters of dimension 1, i.e., one filter per dimension of the state. The standard particle filter and the multiple particle filter generated $M = 100$ particles per dimension of the state. To deal with the coupling of the states given in equation (21), the filters exchanged the means and variances of their particles at the previous time instant. Therefore, all the particles of a given filter were propagated using the same information (mean and variance) from the coupled state. An identical approach was used to deal with the coupling of the observations while computing the likelihoods of the particles. However, in this case the filters exchanged the means and variances of their particles calculated once the propagation step was completed but before the computation of the particle weights.

Figure 1 shows the mean square error (MSE) of the state calculated from 1000 realizations of the system and averaged over all the states. It is obvious that the traditional particle filtering suffers from the large dimension of the state. We note the large difference in performance between the standard particle filter and the new multiple particle filter. The latter achieves a performance very close to the bound given by the Kalman filter.

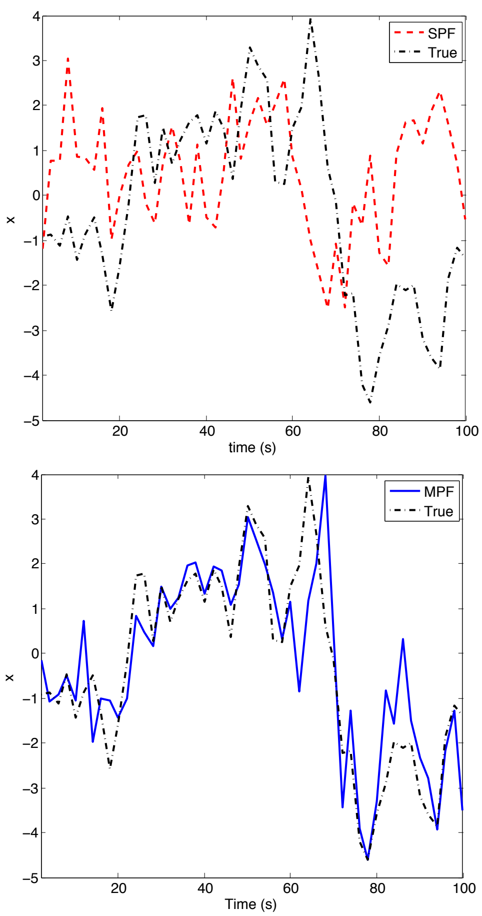


**Fig. 2**. Tracking one randomly chosen state variable in one of the realizations from the nonlinear example. Top: SPF vs. True trajectory. Bottom: MPF vs. True trajectory.

### 4.2. Nonlinear example

We considered the same state equation as in the previous example given by expression (21), and a set of nonlinear observations formulated as

$$
\begin{aligned}
y_{1,t} &= x_{2,t} + e^{\frac{x_{1,t}}{2}} v_{1,t} \\
y_{2,t} &= x_{3,t} + e^{\frac{x_{2,t}}{2}} v_{2,t} \\
&\vdots \\
y_{d_x,t} &= x_{1,t} + e^{\frac{x_{d_x,t}}{2}} v_{d_x,t},
\end{aligned} \tag{23}
$$

with $v_{i,t}$ denoting independent zero-mean Gaussian random variables of variance $\sigma_{v_i}^2 = 1$. Again, at each time instant there were as many observations as states and each component of the state vector participated in two observations.

We set $T = 100$ time units, $d_x = 50$ as the state dimension, and compared the standard and multiple particle filters that used $M = 100$. Figure 2 shows the tracking of one randomly chosen state variable in one particular realization by both algorithms and compares their results with the true trajectory. It is easy to note that the multiple particle filter tracks the trajectory more accurately than the standard one. Finally, Fig. 3 displays the MSE calculated from 1000 realizations and averaged over all the states. The results confirm that the multiple particle filter clearly outperforms the standard one.

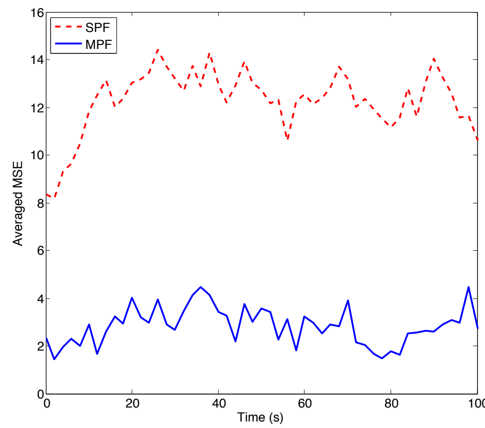


**Fig. 3**. Averaged MSE comparison for the nonlinear example.

## 5. CONCLUSION

In this paper we proposed an approach for implementing multiple particle filtering that is both efficient and accurate. The states of the system are tracked by a number of particle filters, where each filter tracks a subset of the complete state. The particle filters convey information to relevant particle filters about their tracking in the form of mean and covariance of their subset of states. This information by the receiving particle filter is then readily used by them to propose particles in a computationally efficient way and more importantly, to compute accurately the likelihoods of their particles. The proposed methodology is demonstrated by computer simulations.

## 6. REFERENCES


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

[2] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001.

[3] T. Bengtsson, P. Bickel, and B. Li, "Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems," *Probability and Statistics: Essays in honor of David A. Freedman*, vol. 2, pp. 316–334, 2008.

[4] P. J. Van Leeuwen, "Particle filtering in geophysical systems," *Monthly Weather Review*, vol. 137, no. 12, pp. 4089–4114, 2009.

[5] E. L. Ionides, C. Bretó, and A. A. King, "Inference for nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18438–18443, 2006.

[6] G. G. Rigatos, "Extended Kalman and particle filtering for sensor fusion in motion control of mobile robots," *Mathematics and Computers in Simulation*, vol. 81, no. 3, pp. 590–607, 2010.

[7] K. Friston, "The free-energy principle: a rough guide to the brain?," *Trends in Cognitive Sciences*, vol. 13, no. 7, pp. 293–301, 2009.

[8] P. Bickel, B. Li, and T. Bengtsson, "Dynamic conditional independence models and Markov chain Monte Carlo methods," in *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, vol. 92, pp. 1403–1412. 1997.

[9] P. Rebeschini and R. Van Handel, "Can local particle filters beat the curse of dimensionality?," *arXiv preprint arXiv:1301.6585*, 2013.

[10] P. M. Djurić, T. Lu, and M. F. Bugallo, "Multiple particle filtering," in *the Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007, pp. III–1181.

[11] M. F. Bugallo, T. Lu, and P. M. Djurić, "Target tracking by multiple particle filtering," in *the Proceedings of the Aerospace Conference*, 2007.

[12] M. F. Bugallo and P. M. Djurić, "Target tracking by symbiotic particle filtering," in *the Proceedings of the IEEE Aerospace Conference*, 2010, pp. 1–7.

[13] P. M. Djurić and M. F. Bugallo, "Particle filtering for high-dimensional systems," in *the Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, 2013.

[14] M. F. Bugallo and P. M. Djurić, "Particle filtering in high-dimensional systems with Gaussian approximations," in *the Proceedings of the IEEE Acoustics, Speech and Signal Processing*, 2014, pp. 8013–8017.

[15] M. F. Bugallo and P. M. Djurić, "Gaussian particle filtering in high-dimensional systems," in *the Proceedings of the IEEE Workshop on Statistical Signal Processing (SSP)*, 2014, pp. 129–132.

[16] A. Papoulis and S. U. Pillai, *Proability, Random Variables, and Stochastic Processes*, Mc-Graw Hill, New York, Fourth edition, 2001.