A NEW IMPORTANCE FUNCTION FOR PARTICLE FILTERING AND ITS APPLICATION TO BLIND DETECTION IN FLAT FADING CHANNELS

Yufei Huang and Petar M. Djurić

Department of Electrical and Computer Engineering
State University of New York at Stony Brook
Stony Brook, NY, 11794-2350
yfhuang@ece.sunysb.edu, djuric@ece.sunysb.edu

ABSTRACT

Particle filtering has drawn much attention in recent years due to its capacity to handle nonlinear and non-Gaussian problems. One crucial issue in particle filtering is the selection of importance function. In this paper, we propose a new type of importance function, which possess advantages over both the posterior and the prior importance functions. In addition, we demonstrate the use of the proposed importance function in blind detection in flat fading channels. Simulation results show its efficiency and performance.

1. INTRODUCTION

Adaptive filtering methods have found wide range of applications in science, engineering, and finance. Among them, a group of new algorithms known as particle filters [1] has drawn much attention in recent years. Based on Monte Carlo sampling strategies, particle filters have strong potential for tackling nonlinear and non-Gaussian problems [1].

One crucial issue in particle filtering is the selection of the importance function. The two choices that are most seen in the literature are the posterior and the prior importance functions. The posterior importance function, although minimizing the variance of the importance weights, is often almost impossible to use due to difficulty in determining the corresponding weights. Therefore, the prior importance function is usually adopted. A major disadvantage of the prior importance function is its ineffectiveness that leads to poor filtering performance. Other algorithms including the auxiliary particle filter have been proposed to improve its effectiveness [2].

In this paper, we look beyond the above two choices and propose a new type of importance function. It is a hybrid of the aforementioned importance functions. This importance function leads to more effective algorithms than the prior importance function because it uses the most recent observations. It is also less restricted than its posterior counterpart, and is thus applicable to more problems.

As an application of our proposed importance function, the problem of adaptive blind detection in flat fading channels is studied. A similar problem has been addressed in [3]. However, here we consider a more realistic situation, i.e., we assume that the noise variance is unknown. Consequently, the scheme proposed in [3] is no longer feasible.

The paper is organized as follows: The particle filter is briefly described in Section 2. The posterior and the prior importance functions are discussed in Section 3. In Section 4, the new hybrid importance function is proposed, and the computation of its weights is outlined. In Section 5, blind detection in flat fading channels by particle filtering with the proposed importance function is studied. The simulation results are shown in Section 6. Concluding remarks are provided in Section 7.

2. THE PARTICLE FILTER

We consider dynamic systems that are described by state-space models. A state-space model can be represented by the equation:

$$\begin{cases} \mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{u}_t) \\ \mathbf{y}_t = g(\mathbf{x}_t, \mathbf{v}_t) \end{cases} \quad (1)$$

where $f(\cdot)$ and $g(\cdot)$ represent the state and observation functions respectively, $\mathbf{x}_t$ is a vector of state parameters at time t, $\mathbf{y}_t$ denotes the observations, and $\mathbf{u}_t$ and $\mathbf{v}_t$ are the two noise vectors all associated with time t. Now let $\mathbf{x}_{0:t} = \{\mathbf{x}_0, \cdots, \mathbf{x}_t\}$ and $\mathbf{y}_{0:t} = \{\mathbf{y}_0, \cdots, \mathbf{y}_t\}$ represent a collection of states and observations from 0 to t, respectively. Our aim is then to estimate $\mathbf{x}_t$ sequentially based on the observations $\mathbf{y}_{0:t}$. Under a Bayesian paradigm, the posterior distribution $p(\mathbf{x}_t|\mathbf{y}_{0:t})$ is the key entity for estimation. Note that when $f(\cdot)$ and $g(\cdot)$ are linear with respect to the states and $\mathbf{u}_t$ and $\mathbf{v}_t$ are Gaussian, $p(\mathbf{x}_t|\mathbf{y}_{0:t})$ can be derived exactly and the well known Kalman filter provides the optimum solution. However in cases of nonlinearity and non-Gaussianity, this distribution is usually difficult to obtain and, furthermore, the subsequent calculation for the Bayesian estimators could be also prohibitively complex. To circumvent these difficulties, the particle filters adopt simulation-based approaches of which the basic technique is the sequential importance sampling [1].

To illustrate the particle filtering algorithm, suppose that at time $t-1$, we have collected N sets of properly weighted samples $\mathbf{x}_{0:t-1}^{(j)} = \{\mathbf{x}_0^{(j)}, \cdots, \mathbf{x}_{t-1}^{(j)}\}$ and their associated weights $w_{t-1}^{(j)}$ for $j = 1, \cdots, N$. In particular, the weighted samples $\{\mathbf{x}_{0:t-1}^{j}, w_t^{(j)}\}_{j=1}^{N}$ are distributed approxi-

This work was supported by the National Science Foundation under Awards No. CCR-9903120 and CCR-0082607.

0-7803-7402-9/02/$17.00 ©2002 IEEE

mately according to $p(\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t-1})$. When the new observations $\mathbf{y}_t$ arrive, the update of the sample sets from $t-1$ to t is carried out as follows:

The Particle filter

- For $j = 1, \cdots, N$
  - Sample $\mathbf{x}_t^{(j)}$ from an importance function $q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t})$ and set $\mathbf{x}_{0:t}^{(j)} = \{\mathbf{x}_{0:t-1}^{(j)}, \mathbf{x}_t^{(j)}\}$.
  - Calculate the weight by

$$\bar{w}_t^{(j)} = w_{t-1}^{(j)} \frac{p(\mathbf{x}_{0:t}{}^{(j)}|\mathbf{y}_{0:t})}{p(\mathbf{x}_{0:t-1}^{(j)}|\mathbf{y}_{0:t-1}) q(\mathbf{x}_t^{(j)}|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t})} \quad (2)$$

- For $j = 1, \cdots, N$, normalize the weights:

$$w_t^{(j)} = \frac{\bar{w}_t^{(j)}}{\sum_{j=1}^{N} \bar{w}_t^{(j)}} \quad (3)$$

The resulting weighted samples $\{\mathbf{x}_t^{(j)}, w_t^{(j)}\}_{j=1}^N$ approximate $p(\mathbf{x}_t|\mathbf{y}_{0:t})$, and the Bayesian estimator of the states $\mathbf{x}_t$ can be easily calculated using the set $\{\mathbf{x}_t^{(j)}, w_t^{(j)}\}_{j=1}^N$. For instance, the minimum mean square estimator (MMSE) of $\mathbf{x}_t$ is computed according to

$$\hat{\mathbf{x}}_{tMMSE} = \sum_{j=1}^{N} \mathbf{x}_t^{(j)} w_t^{(j)}. \quad (4)$$

The choice of the importance function is essential because it determines the efficiency as well as the complexity of the particle filtering algorithm.

3. THE IMPORTANCE FUNCTIONS

3.1. The posterior importance function

The posterior importance function is defined as $q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}) = p(\mathbf{x}_t|\mathbf{x}_{0:t}^{(j)}, \mathbf{y}_{0:t})$. If it is used for generating particles, the importance weights can be obtained from

$$\bar{w}_t^{(j)} = w_{t-1}^{(j)} p(\mathbf{y}_t|\mathbf{x}_{t-1}^{(j)}, \mathbf{y}_{0:t-1}). \quad (5)$$

This choice of the importance function is optimal in the sense of minimizing the variance of the importance weights. As a consequence, more effective samples can be generated and better estimates produced. However, a major difficulty in its use is the calculation of the weights. Note that it requires analytical evaluation of $p(\mathbf{y}_t|\mathbf{x}_{t-1}^{(j)}, \mathbf{y}_{0:t-1})$, which involves complex high dimensional integrations. This difficulty prevents the posterior importance function from being widely used.

3.2. The prior importance function

The prior importance function is defined as $q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}) = p(\mathbf{x}_t|\mathbf{x}_{t-1}^{(j)})$. Compared with the posterior importance function, it is attractive due to its simplicity in implementation. First, we can see that sampling from the prior densities is often straightforward. Second, since the weights associated with this importance function are

$$\bar{w}_t^{(j)} = w_{t-1}^{(j)} p(\mathbf{y}_t|\mathbf{x}_t^{(j)}, \mathbf{y}_{0:t-1}) \quad (6)$$

where $p(\mathbf{y}_t|\mathbf{x}_t^{(j)}, \mathbf{y}_{0:t-1})$ is simply the likelihood function at time t, the calculation of the likelihood function is usually easy. Despite the simplicity of applying the prior importance function, the implementation of the particle filter with it is discouraged because it can be very inefficient. Since no information from the observations is used, the generated particles often come from the tails of the posterior distributions, and as a result, the weights have large variations and the estimation results are poor.

4. A HYBRID IMPORTANCE FUNCTION

Consider the situation where the use of the posterior importance function is extremely difficult. In the cases analyzed here, we assume that the state parameters can be divided into two independent parts, e.g., $\mathbf{x}_t = \{\mathbf{x}_{1t}, \mathbf{x}_{2t}\}$ such that generation from $p(\mathbf{x}_{2t}|\mathbf{x}_{2,t-1}^{(j)})$ and $p(\mathbf{x}_{1t}|\mathbf{x}_{2t}^{(j)}, \mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t})$ can be carried out easily. In these cases, we propose to use the following importance function

$$q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}) = p(\mathbf{x}_{1t}|\mathbf{x}_{2t}^{(j)}, \mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}) p(\mathbf{x}_{2t}|\mathbf{x}_{2,t-1}^{(j)}) \quad (7)$$

where $\mathbf{x}_{2t}^{(j)}$ is the proposed sample at t from $p(\mathbf{x}_{2t}|\mathbf{x}_{2,t-1}^{(j)})$. Apparently, (7) is a hybrid between the posterior and the prior importance functions. The weights of the particles generated by the hybrid importance function can be derived from (2). An outline of the derivation is shown in (8). We note from (8) that the distribution $p(\mathbf{y}_t|\mathbf{x}_{2t}^{(j)}, \mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t-1})$ is critical in the computation of the weights, and that therefore, its analytical form should be available.

The advantage of the proposed hybrid importance function over the posterior importance function is in the easy updating of the weights. In addition, since the hybrid importance function includes information from the observations, it generates samples with smaller weight variance than the prior importance function.

The key to the applicability of this importance function is the assumption of knowing the analytical form of $p(\mathbf{x}_{1t}|\mathbf{x}_{2t}^{(j)}, \mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t})$ (including the normalizing constant). Cases that fall within the assumption, for instance, are the ones where the $\mathbf{x}_{1t}$'s are discrete variables, and given $\mathbf{x}_{2t}$, $\mathbf{y}_t$ is linear with $\mathbf{x}_{1t}$. These conditions resemble those of the mixture Kalman filter [4, 3]. However, in the mixture Kalman filter, they provide the possibility to maginalize the nuisance states $\mathbf{x}_{1t}$. On the contrary, in our case $\mathbf{x}_{1t}$ are the states of interests, and the conditions lead to a more efficient and effective importance function. In fact, the mixture Kalman filter can be integrated with the use of the hybrid importance function to further enhance the performance of the method. In addition, other techniques like the smoothing kernel [5] can be incorporated to reduce the variance of the weights.

$$
\begin{aligned}
\bar{w}_t^{(j)} &= w_{t-1}^{(j)} \frac{p(\mathbf{x}_{0:t}{}^{(j)}|\mathbf{y}_{0:t})}{p(\mathbf{x}_{0:t-1}^{(j)}|\mathbf{y}_{0:t-1})p(\%bfx_{1t}|\mathbf{x}_{2t}^{(j)},\mathbf{x}_{0:t-1}^{(j)},\mathbf{y}_{0:t})p(\mathbf{x}_{2t}^{(j)}|\mathbf{x}_{2,t-1}^{(j)})} \\
&\propto w_{t-1}^{(j)} \frac{p(\mathbf{y}_t|\mathbf{x}_{1t}^{(j)},\mathbf{x}_{2t}^{(j)})p(\mathbf{x}_{1t}^{(j)},\mathbf{x}_{2t}^{(j)}|\mathbf{x}_{1,t-1}^{(j)},\mathbf{x}_{2,t-1}^{(j)})p(\mathbf{x}_{0:t-1}^{(j)}|\mathbf{y}_{0:t-1})}{p(\mathbf{x}_{0:t-1}^{(j)}|\mathbf{y}_{0:t-1})p(\mathbf{x}_{1t}^{(j)}|\mathbf{x}_{2t}^{(j)},\mathbf{x}_{0:t-1}^{(j)},\mathbf{y}_{0:t})p(\mathbf{x}_{2t}^{(j)}|\mathbf{x}_{2,t-1}^{(j)})} \\
&= w_{t-1}^{(j)} \frac{p(\mathbf{y}_t|\mathbf{x}_{1t}^{(j)},\mathbf{x}_{2t}^{(j)})p(\mathbf{x}_{1t}^{(j)}|\mathbf{x}_{1,t-1}^{(j)})p(\mathbf{y}_t|\mathbf{x}_{2t}^{(j)},\mathbf{x}_{0:t-1}^{(j)},\mathbf{y}_{0:t-1})}{p(\mathbf{y}_t|\mathbf{x}_{1t}^{(j)},\mathbf{x}_{2t}^{(j)})p(\mathbf{x}_{1t}^{(j)}|\mathbf{x}_{1,t-1}^{(j)})} \\
&= w_{t-1}^{(j)} p(\mathbf{y}_t|\mathbf{x}_{2t}^{(j)},\mathbf{x}_{0:t-1}^{(j)},\mathbf{y}_{0:t-1})
\end{aligned} \tag{8}
$$

5. ADAPTIVE BLIND DETECTION IN FLAT FADING CHANNELS WITH UNKNOWN NOISE VARIANCE

5.1. Problem formulation

We consider detection of digital signals in flat fading channels. In the baseband, at time t, the received signal y_t is obtained from the M-ary transmitted signal s_t as

$$y_t = h_t s_t + e_t \tag{9}$$

where h_t and e_t are the complex fading coefficients and additive ambient noise. The noise e_t is assumed to be complex Gaussian with zero mean and unknown variance σ^2. The flat fading channel is further assumed to be a Rayleigh process, and thus h_t is a complex Gaussian process. To model the temporal correlation of the fading channel, a second order autoregressive (AR) process is used such that

$$h_t = -a_1 h_{t-1} - a_2 h_{t-2} + v_t \tag{10}$$

where a_1 and a_2 are the known model coefficients, and $v_t \sim \mathcal{CN}(0, 1)$. The coefficients are closely related to the physical characteristics of the underlying fading process and can be determined by fitting the autocorrelation function of the true fading process [6].

From a particle filtering perspective, we need to formulate the problem in a state space representation. It can be expressed as

$$\begin{cases} \sigma_t^2 = \sigma_{t-1}^2 \\ \mathbf{h}_t = \mathbf{D}\mathbf{h}_{t-1} + \mathbf{g}v_t \\ y_t = \mathbf{g}^\top \mathbf{h}_t s_t + e_t \end{cases} \tag{11}$$

where $\mathbf{h}_t = [h_t \; h_{t-1}]^\top$, $\mathbf{g} = [1 \; 0]^\top$, and

$$\mathbf{D} = \begin{bmatrix} -a_1 & -a_2 \\ 1 & 0 \end{bmatrix}.$$

At any instant of time t, the unknowns of the problem are s_t, h_t and σ_t^2, and our main objective is to detect the transmitted signal s_t sequentially without sending pilot signals.

5.2. A particle filtering solution

In the implementation of particle filtering, the first rule of thumb is to marginalize out as many nuisance parameters as possible. Here, we observe that given σ_t^2 and s_t, (11) is linear regarding h_t. Therefore, the mixture Kalman filter can be used to marginalize out h_t. Next, we are ready to apply particle filtering on s_t and σ_t^2. In choosing the importance function, we notice that the posterior importance function is intractable due to the presence of σ_t^2. Consequently, one would usually resort to using the prior importance function. However, since s_t is a discrete variable, we can instead adopt the hybrid importance function which is taken as

$$
\begin{aligned}
q(s_t, \sigma_t^2 | s_{t-1}^{(j)}, \sigma_{t-1}^{2(j)}, \mathbf{y}_{0:t}) &= p(s_t|\sigma_t^{2(j)}, \mathbf{y}_{0:t}) p(\sigma_t^2|\sigma_{t-1}^{2(j)}) \\
&= p(s_t|\sigma_t^{2(j)}, \mathbf{y}_{0:t}) \delta(\sigma_{t-1}^{2(j)})
\end{aligned} \tag{12}
$$

where $\sigma_t^{2(j)} = \sigma_{t-1}^{2(j)}$, $\delta(\cdot)$ is the Dirac delta function, and the last equality is obtained based on the state equation $\sigma_t^2 = \sigma_{t-1}^2$. The corresponding weight is obtained from (8) as

$$
\begin{aligned}
w_t^{(j)} &= w_{t-1}^{(j)} p(y_t|\sigma_t^{2(j)}, \mathbf{y}_{0:t-1}) \\
&= w_{t-1}^{(j)} \sum_{s_t \in \mathcal{A}} p(y_t|s_t, \sigma_t^{2(j)}, \mathbf{y}_{0:t-1})
\end{aligned} \tag{13}
$$

where $\mathcal{A} = \{a_1, \cdots, a_M\}$ is the alphabet space of s_t.

Next, we discuss the sampling of s_t and σ_t^2 from (12) and the calculation of the weight (13). First of all, we notice that no sampling for σ_t^2 is required. Although it simplifies the sampling process, the absence of sampling introduces lack of diversity on σ_t^2. To address this problem, smoothing kernel techniques can be used during the resampling procedure. As for s_t, since it is discrete, the sampling of it only requires the evaluation of the importance function on $\mathcal{A}$. In particular, we have

$$p(s_t|\sigma_t^{2(j)}, \mathbf{y}_{0:t}) \propto p(y_t|s_t, \sigma_{t-1}^{2(j)}, \mathbf{y}_{0:t-1}) \tag{14}$$

Now, considering also (13), we find that both the sampling of s_t and the calculation of the weight are achieved by computing $p(y_t|s_t, \sigma_t^{2(j)}, \mathbf{y}_{0:t-1})$. This distribution is the likelihood function after marginalizing out h_t and can be obtained from the predictive procedure of the Kalman filter as

$$p(y_t|s_t, \sigma_t^{2(j)}, \mathbf{y}_{0:t-1}) = \mathcal{N}(m_t^{(j)}, c_t^{(j)}) \tag{15}$$

where $m_t^{(j)} = \mathbf{g}^\top \mathbf{D} \mu_{t-1}^{(j)}$ and $c_t^{(j)} = \mathbf{g}^\top \mathbf{\Sigma}_t^{(j)} \mathbf{g} + \sigma_t^{2(j)}$ with $\mathbf{\Sigma}_t = \mathbf{D}\mathbf{P}_{t-1}^{(j)}\mathbf{D}^\mathsf{H} + \mathbf{g}^\top\mathbf{g}$ and H denoting the Hermitian transpose. Moreover, $\mu_{t-1}^{(j)}$ and $\mathbf{P}_{t-1}^{(j)}$ are computed from the update steps of the Kalman filter that are expressed, in the time index t, as $\mu_t^{(j)} = \mathbf{D}\mu_{t-1}^{(j)} + \mathbf{K}_t^{(j)}(y_t - m_t^{(j)})$ and $\mathbf{P}_t^{(j)} = (\mathbf{I} - \mathbf{K}_t^{(j)}\mathbf{g}s_t^{(j)})\mathbf{\Sigma}_t^{(j)}$ where $\mathbf{K}_t^{(j)} = \mathbf{\Sigma}_t^{(j)}\mathbf{g}c_t^{(j)-1}s_t^{(j)}$.

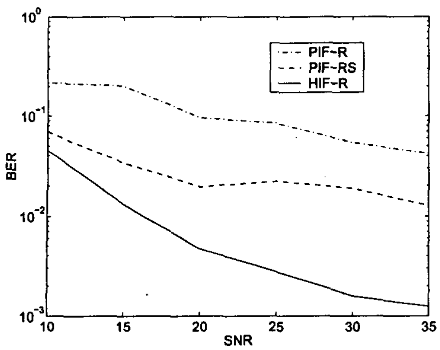

Figure 1: Comparison plot of the BERs between the particle filtering detectors using the hybrid and the prior importance function at various SNRs.

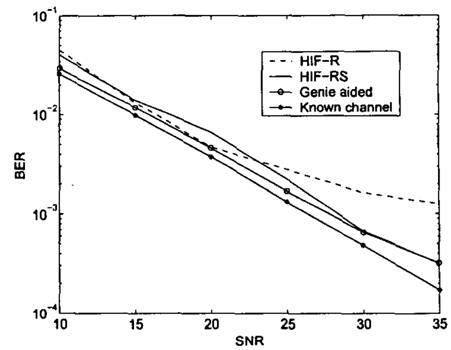

Figure 2: Plot of the BERs at various SNRs of the two HIF solutions and of the known channel and genie aided detectors.

6. SIMULATION

The performance of the proposed particle filtering scheme using the hybrid importance function is studied in this section. The coefficients of the fading channel model (10) were $a_1 = -1.99348$ and $a_2 = 0.996$. They reflect a physical scenario of a Doppler spread of 113 Hz and data rate of 10 Kbps. This AR process is normalized to have a unit power, and thus the signal to noise ratio (SNR) is obtained as $10\log(1/\sigma^2)$. The transmitted signal is BPSK modulated with differential coding. In the following examples, four different particle filtering schemes were examined. Two of them used the hybrid importance function: one with a resampling procedure at every 10 transmissions (HIF-R) and the other, with a smoothing kernel for σ^2 applied at every resampling step (HIF-RS). There were also two schemes that used the prior importance function. One of them employed resampling at every 10 transmissions (PIF-R) and the other, used both resampling and a smoothing kernel (PIF-RS). Furthermore, for the ones using the hybrid importance function, 150 particles were drawn at every filtering step. However, 2000 particles were generated for the ones using the prior importance function. Finally, the MMSE estimator was used to estimate s_t.

In the first example, we provide the bit error rates (BERs) of several particle filters under various SNRs. To compute the BER at a tested SNR, a symbol stream was transmitted continuously until 300 errors were collected (to allow the algorithms to reach the stable state, the errors among the first 100 symbols were ignored). We plotted the simulation results in Figure 1. We find that, although the PIF-RS solution provided improvement over the results of PIF-R, the HIF-R clearly achieved the best performance. In addition, it only used about 1/13 of the number of particles used by the two PIF filters. These facts clearly demonstrate the better efficiency and effectiveness of the hybrid importance function.

In the second example, we compared the performance of the HIF filters with the performance of the known channel and genie aided detectors, where the noise variance was known. These two detectors served as benchmarks for lower bounds. We plotted the BERs vs SNRs in Figure 2. We notice that by applying the smooth kernel to σ^2, big improvements were achieved by HIF-RS. Especially, at high SNRs, the performance of the HIF-RS filter approaches that of the genie aided detector with known noise variance.

7. CONCLUSION

In this paper, we proposed a hybrid importance function which encompasses advantages of both the posterior and the prior importance functions. We have shown its application to the blind detection in flat fading channels. Simulation results showed much improved performance over the ones that use the prior importance function.

8. REFERENCES

[1] A. Doucet, J. de Freitas, and N. Gordon, Eds., *Sequential Monte-Carlo Methods in Practice*, Springer-Verlag, 2000.

[2] M. Pitt and N. Shephard, "Filtering via simulation: auxiliary particle filter," *Journal of the American Statistical Association*, vol. 94, pp. 590–599, 1999.

[3] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2079–2094, Sept. 2000.

[4] R. Chen and J. Liu, "Mixture Kalman filters," *Journal of Royal Statist. Soc. B*, vol. 62, pp. 493–508, 2000.

[5] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. De Freitas, and N. J. Gordon, Eds., New York, 2000, Springer-Verlag.

[6] M. Sternad, L. Lindbom, and A.Ahlén, "Tracking of time-varying mobile radio channels with WLMS algorithms: A case study on D-AMPS 1900 channels," in *Proceeding of IEEE VTC 2000, Tokyo, Japan*, May 2000, pp. 2507–2511.