A NEW APPROACH TO PARTICLE BASED SMOOTHED MARGINAL MAP

S. Saha , P. K. Mandal and A. Bagchi

Dept. of Applied Mathematics, University of Twente
7500 AE, Enschede, The Netherlands
phone: + (31)53-489-3453, fax: + (31)53-489-3800, email: s.saha, p.k.mandal, a.bagchi @ewi.utwente.nl

ABSTRACT

We present here a new method of finding the MAP state estimator from the weighted particles representation of marginal smoother distribution. This is in contrast to the usual practice, where the particle with the highest weight is selected as the MAP, although the latter is not necessarily the most probable state estimate. The method developed here uses only particles with corresponding filtering and smoothing weights. We apply this estimator for finding the unknown initial state of a dynamical system and addressing the parameter estimation problem.

1. INTRODUCTION

The maximum *a posteriori* (MAP) estimate of a stochastic unobserved variable x given the observations y is the value of x that maximizes the posterior density $p(x|y)$. This MAP estimate is specially useful when the posterior has a strong multimodal characteristic. This scenario may often arise in target tracking problems ([2],[7]). For example, the posterior of a target position may be multimodal and in such a situation, the minimum mean square error (MMSE) state estimate may be located in a region between the modes, which has very low probability. For obvious reasons, MAP estimator is therefore more meaningful in such cases. However, in practice, use of MAP is limited in the sense that for a general nonlinear dynamic system, closed form solution for the posterior density is hardly available, whereas analytically approximated model may lead to an inaccurate MAP estimation. In recent times, starting with Gordon's seminal paper ([18]), particle based sequential Monte Carlo method has been getting increasing attention due to its capability of efficiently approximating such difficult posterior distributions. In this method, the posterior is approximated by a cloud of N weighted particles, whose empirical measure closely approximates the true posterior for large N ([4],[1],[12]).

In the previous literature, it has been argued that the MAP estimator in the particle filtering framework can be given by the particle with the highest weight, for example see ([16],[20]). However, the particle with the highest weight does not necessarily represent the most probable state estimate ([5],[19],[6]). Thus, this estimator is not really a fair approximation of the true MAP. In this paper, we present a new method of estimating the marginal smoother MAP. Estimating this MAP essentially involves maximization over the posterior density $p(x_t|y_{1:T})$. Naturally, the crux of the problem lies in constructing this posterior density from the weighted cloud representation of the smoothed distribution. The most straight-forward approach is the kernel based method where a kernel is fitted around each particle to get the approximate continuous density ([15]). This method requires a choice of kernel bandwidth which is not obvious and it is computationally demanding, which restricts its use in many practical applications.

Recently, there has been an interesting development on particle filter MAP estimation ([5],[6]) where the authors estimate the density function from the running particle filters only. This method thus avoids the need of bandwidth selection associated with the kernel based methods. In principle, this new method can provide the probability density function at any support point. We extend here this idea to the smoothing algorithm. Our proposed algorithm is then used to estimate the unknown initial state of a given dynamic system. We also apply the method to the parameter estimation problem.

This work was supported by a research grant from THALES Nederland B.V.

2. PROBLEM STATEMENT

Consider a nonlinear dynamic system given by

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

$$y_t = h(x_t, v_t), \qquad t = 1, 2, \ldots \tag{2}$$

where (x_t) are the unobservable system values (the state) with (known) initial prior density $p(x_0) \equiv p(x_0|x_{-1})$ and (y_t) are the observed values (the measurements). The process noises (w_t) are assumed to be independent of the measurement noises (v_t). The problem here is to estimate the maximum *a posteriori* (MAP) of the unobserved system value x_t from all the observations $y_{1:T} \equiv (y_1, y_2, \ldots, y_T)$, up to time T (where $t < T$) or equivalently, to estimate the value of x_t that maximizes the posterior density (also known as marginal smoothing density) $p(x_t|y_{1:T})$. This can be stated mathematically as

$$x_{t|T}^{MAP} = \arg\max_{x_t} p(x_t|y_{1:T}). \tag{3}$$

3. MAP ESTIMATOR FOR MARGINAL SMOOTHING DENSITY

In general, no analytical solution is available for this MAP estimator. So we focus our attention here to approximately construct the marginal smoothing density $p(x_t|y_{1:T})$. The marginal smoother can be obtained using forward- backward smoother ([10]) as

$$p(x_t|y_{1:T}) = p(x_t|y_{1:t}) \int \frac{p(x_{t+1}|y_{1:T})\, p(x_{t+1}|x_t)}{p(x_{t+1}|y_{1:t})} dx_{t+1}, \tag{4}$$

where, $p(x_t|y_{1:t})$ and $p(x_{t+1}|y_{1:t})$ are the filtering density and one step ahead predictive density respectively, at time t. The marginal fixed interval smoother $p(x_t|y_{1:T})$ is obtained by backward recursion starting from $p(x_T|y_{1:T})$.

3.1 Particles based forward-backward smoothing

The marginal smoothing distribution can be approximated using Monte Carlo particle based techniques as described in ([3],[13]). The algorithm is derived based on the approximation to equation (4). Here, one starts with the forward filtering pass for computing the filtered distribution at each step using particle filter as

$$\widehat{P}(dx_t|y_{1:t}) = \sum_{j=1}^{N} \omega_t^{(j)} \delta_{x_t^{(j)}}(dx_t). \tag{5}$$

Next, relying on the same set of supports generated by the forward distribution, one performs the backward smoothing pass to determine the smoothing distribution. This smoothing distribution is approximated as

$$\widehat{P}(dx_t|y_{1:T}) = \sum_{i=1}^{N} \omega_{t|T}^{(i)} \delta_{x_t^{(i)}}(dx_t), \tag{6}$$

where the smoothing weights are obtained through the following backward recursion:

$$\omega_{t|T}^{(i)} = \omega_t^{(i)} \sum_{j=1}^{N} [\omega_{t+1|T}^{(j)} \frac{p(x_{t+1}^{(j)}|x_t^{(i)})}{\sum_{k=1}^{N} p(x_{t+1}^{(j)}|x_t^{(k)})\omega_t^{(k)}}] \tag{7}$$

with $\omega_{T|T}^{(i)} = \omega_T^{(i)}$. It is important to note that the forward- backward smoother keeps the same particle support as used in filtering step and re-weights the particles to obtain the approximated particle based smoothed distribution. Thus, success of this method crucially hinges on the filtered distribution having supports where the smoothed distribution is significant.

3.2 Particles based MAP estimator for marginal smoothing density

As mentioned in the introduction earlier, to calculate MAP one needs the posterior density $p(x_t|y_{1:T})$ from the cloud representation. One can get this using kernel based method but with its limitations as stated earlier. This kernel based method can be viewed as a separate post-processor which extracts the density from the weighted particles. Here, we envisage a simple alternative method to compute this density by using the (weighted) particles only. We proceed as follows:

Using Bayes' rule, one can write the one step ahead predictive density in equation (4) as

$$p(x_{t+1}|y_{1:t}) = \frac{p(x_{t+1}|y_{1:t+1})p(y_{t+1}|y_{1:t})}{p(y_{t+1}|x_{t+1})}. \tag{8}$$

Substituting the expression in (8) in equation (4), one obtains, $p(x_t|y_{1:T})$

$$\begin{aligned} &= p(x_t|y_{1:t}) \int \frac{p(x_{t+1}|y_{1:T})p(x_{t+1}|x_t)p(y_{t+1}|x_{t+1})}{p(x_{t+1}|y_{1:t+1})p(y_{t+1}|y_{1:t})} dx_{t+1} \\ &= \frac{p(x_t|y_{1:t})}{p(y_{t+1}|y_{1:t})} \int \left[\frac{p(x_{t+1}|x_t)p(y_{t+1}|x_{t+1})}{p(x_{t+1}|y_{1:t+1})} \right] p(x_{t+1}|y_{1:T}) dx_{t+1} \\ &= \frac{p(x_t|y_{1:t})}{p(y_{t+1}|y_{1:t})} \int \left[\frac{p(x_{t+1}|x_t)p(y_{t+1}|x_{t+1})}{p(x_{t+1}|y_{1:t+1})} \right] \widehat{P}(dx_{t+1}|y_{1:T}). \end{aligned}$$

Approximating the above integration by Monte Carlo integration method, one obtains

$$p(x_t|y_{1:T}) \approx \frac{p(x_t|y_{1:t})}{p(y_{t+1}|y_{1:t})} \sum_{j=1}^{N} \left[\frac{p(x_{t+1}^{(j)}|x_t)p(y_{t+1}|x_{t+1}^{(j)})}{p(x_{t+1}^{(j)}|y_{1:t+1})} \right] \omega_{t+1|T}^{(j)}. \tag{9}$$

Furthermore, the filtered density $p(x_{t+1}|y_{1:t+1})$ can be approximated from the running particle filter ([5]) as

$$p(x_{t+1}|y_{1:t+1}) \approx \frac{p(y_{t+1}|x_{t+1}) \sum_k p(x_{t+1}|x_t^{(k)})w_t^{(k)}}{p(y_{t+1}|y_{1:t})}. \tag{10}$$

We can then rewrite equation (9) as

$$p(x_t|y_{1:T}) \approx p(x_t|y_{1:t}) \sum_{j=1}^{N} \left[\frac{p(x_{t+1}^{(j)}|x_t)}{\sum_{k=1}^{N} p(x_{t+1}^{(j)}|x_t^{(k)})\omega_t^{(k)}} \right] \omega_{t+1|T}^{(j)}. \tag{11}$$

The MAP estimate of the marginal smoothing density, $p(x_t|y_{1:T})$ can then be obtained by finding the location of its global maxima. At this point, there are several choices for performing the optimization. In the subsequent section, we describe a method to approximate this MAP with reduced computational budget, which may be practically relevant for many applications.

The particle representation of any distribution may be viewed as an adaptive discrete grid approximation to the true distribution ([8]). Following this representation, we can approximately locate the MAP of $p(x_t|y_{1:T})$ by evaluating this density at the particles $\{x_t^{(i)}\}_{i=1}^{N}$ and finally selecting the particle with the highest density. This leads to the approximate particle based MAP estimate as

$$x_{t|T}^{MAP} \approx \arg\max_{x_t^{(i)}} p(x_t^{(i)}|y_{1:t}) \sum_{j=1}^{N} \left[\frac{p(x_{t+1}^{(j)}|x_t^{(i)})}{\sum_{k=1}^{N} p(x_{t+1}^{(j)}|x_t^{(k)})\omega_t^{(k)}} \right] \omega_{t+1|T}^{(j)}, \tag{12}$$

for $i = 1,..,N$ where N is the number of particles used in cloud representation at each step. The estimator can be further simplified by using equation (7) as

$$x_{t|T}^{MAP} = \arg\max_{x_t^{(i)}} p(x_t^{(i)}|y_{1:t}) \frac{\omega_{t|T}^{(i)}}{\omega_t^{(i)}}, \tag{13}$$

where the filtered density $p(x_t|y_{1:t})$ at the particle cloud $\{x_t^{(i)}\}_{i=1}^{N}$ can be evaluated during the forward filtering step ([5]) as

$$p(x_t^{(i)}|y_{1:t}) \approx \frac{p(y_t|x_t^{(i)}) \sum_j p(x_t^{(i)}|x_{t-1}^{(j)})w_{t-1}^{(j)}}{p(y_t|y_{1:t-1})}. \tag{14}$$

Subsequently, to obtain $x_{t|T}^{MAP}$, one can replace $p(x_t^{(i)}|y_{1:t})$ in equation (13) by the un-normalized filtered density

$$q(x_t^{(i)}|y_{1:t}) = p(y_t|x_t^{(i)}) \sum_j p(x_t^{(i)}|x_{t-1}^{(j)})w_{t-1}^{(j)} \tag{15}$$

because, $p(y_t|y_{1:t-1})$ in equation (14) is independent of $x_t^{(i)}$. We note here that a numerical problem may arise in equation (13) to obtain the MAP if the filtered weights attached to some particles are very small. This may happen when the "particle degeneracy" occurs and the problem can be addressed using a combination of efficient importance proposal ([4],[9],[14]) with resampling steps. The memory requirement of this marginal MAP estimator for each time step is $O(N)$ and the computational complexity is $O(N^2)$. This complexity may possibly be reduced using the method suggested by Klass *et al* ([11]). We do not discuss this any further in this paper.

4. ALGORITHM

- Given observation $y_{1:T}$, For $i = 1,..,N$, where N is the number of particles

<u>Forward Filtering step</u>

- Assume $p(x_0)$, draw $x_0^{(i)}$ from $p(x_0)$, set $\omega_0^{(i)} = \frac{1}{N}$.
- Run Particle Filter to generate and store $x_t^{(i)}, \omega_t^{(i)}$ for $t = 0,...,T$
- Evaluate (un-normalized) filtered pdf for $t = 1,...,T$, at cloud points i

$$q(x_t^{(i)}|y_{1:t}) = p(y_t|x_t^{(i)}) \sum_j p(x_t^{(i)}|x_{t-1}^{(j)})\omega_{t-1}^{(j)}$$

starting with $q(x_0^{(i)}) = p(x_0^{(i)})$ and store

Backward Smoothing step

- Set $\omega_{T|T}^{(i)} = \omega_T^{(i)}$
- For $t = T-1,...,0$ evaluate the smoother importance weights as

$$\omega_{t|T}^{(i)} = \omega_t^{(i)} \sum_{j=1}^{N} [\omega_{t+1|T}^{(j)} \frac{p(x_{t+1}^{(j)}|x_t^{(i)})}{\sum_{k=1}^{N} p(x_{t+1}^{(j)}|x_t^{(k)})\omega_t^{(k)}}]$$

- Evaluate the approximate smoother MAP as

$$x_{t|T}^{MAP} = \arg\max_{x_t^{(i)}} q(x_t^{(i)}|y_{1:t}) \frac{\omega_{t|T}^{(i)}}{\omega_t^{(i)}}$$

5. NUMERICAL EXAMPLES

Since for a linear-Gaussian model, the marginal smoothed MAP can be obtained analytically using Kalman smoother, we have first validated the estimate of the particle based marginal smoothed MAP against it. The result is satisfactory. As it does not give any further insight, the result is not included here. After this successful initial validation step, we have applied this marginal smoother MAP to estimate the unknown initial condition of the state. Subsequently, using same approach we have addressed parameter estimation problems by considering parameter as additional state.

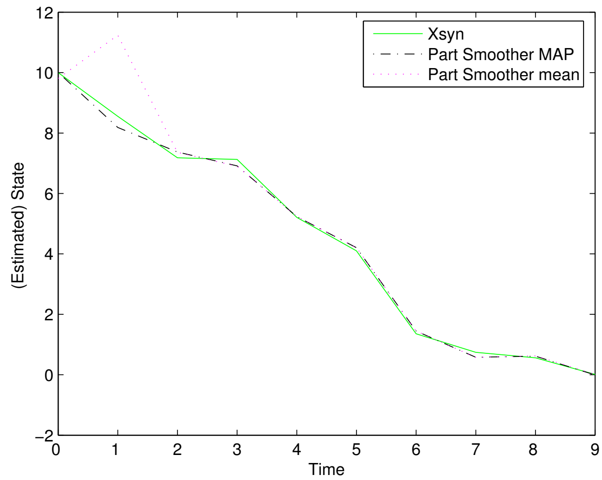

Figure 1: MAP and mean of the marginal smoothing posterior for the first 10 time steps

5.1 Estimation of (unknown) initial condition

5.1.1 Linear State Space

We have considered the following linear Gaussian model:

$$x_k = 0.8x_{k-1} + w_k \tag{16}$$

$$y_k = x_k + v_k \tag{17}$$

where $w_k \sim N(0,1)$ and $v_k \sim N(0,0.1)$. In this model, the initial state x_0 is assumed to be unknown (constant). The synthetic data $\{x_k, y_k\}_{k=0:500}$ is generated starting with $x_0^* = 10$. To estimate the unknown initial state x_0, we start with initial prior $p(x_0) \sim U[0,\ 20]$ where $U[a,\ b]$ denotes uniform probability density function with lower bound a and upper bound b respectively. We use ”efficint proposal” as given in ([4]) in the forward filtering step with particle sample size $N = 500$. The estimate of the initial unknown state is given by the particle based MAP of $p(x_0|y_{0:T})$. We repeat this MAP state estimate for 30 Monte Carlo runs. The mean and variance of the estimator are shown in Table 1. The result shows that the smoothed initial density peaks around the true initial state, even though we have started with a pretty wide uniform initial prior. We also plot for a particular realization, the (backward) evolution of the marginal smoother estimates (i.e. mean and the MAP) for the first 10 time steps and the un-normalized filtered and smoothed probability density functions (pdfs) of x_0 in figure 1 and figure 2 respectively. As expected, the mean and MAP are almost similar and the smoothed density is more concentrated than the filtered density around the true value 10.

$Mean(x_{0\|500}^{MAP})$	$Var(x_{0\|500}^{MAP})$
9.9726	0.0915

Table 1: Mean and Variance of estimated initial state

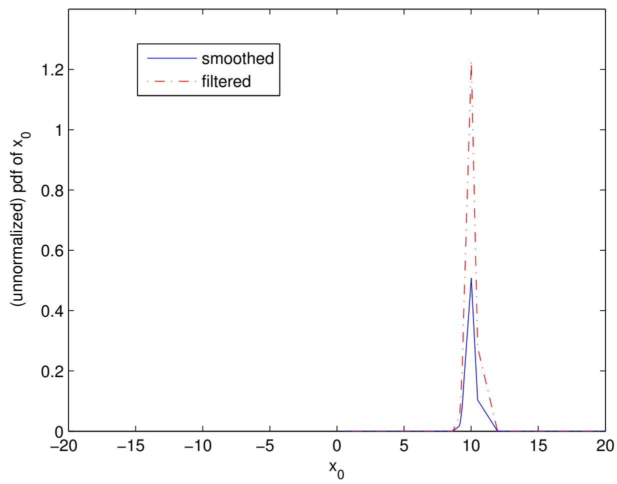

Figure 2: Filtered and smoothed probability density functions for the initial state x_0

5.1.2 Nonlinear State Space

Next, we consider the nonlinear time series model.

$$x_k = \frac{x_{k-1}}{2} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + w_k, \tag{18}$$

$$y_k = \frac{x_k^2}{20} + v_k, \quad k = 1,2,\ldots \tag{19}$$

where $w_k \sim N(0,10)$ and $v_k \sim N(0,1)$. The synthetic data $\{x_k, y_k\}_{k=0:500}$ is generated starting with $x_0^* = 10$. Like previous case, we start with initial prior $p(x_0) \sim U[0,\ 20]$. For this nonlinear problem, we use the ”Exact Moment matching (EMM) proposal” as given in ([17]) during forward filtering step with particle sample size $N = 500$. The estimate of the initial unknown state is given by the particle based MAP of $p(x_0|y_{0:T})$. We repeat this MAP state estimate for 30 Monte Carlo runs. The mean and variance of the estimator are shown in Table 2. The result in Table 2 is really remarkable as we can see by comparing with Table 1. Even for highly nonlinear model as considered above and with wide uniform initial prior, the result is almost as good as in linear case. Of course the variance is somewhat larger, but that is to be expected given the highly nonlinear nature of the problem. It is also interesting to

$Mean(x_{0\|500}^{MAP})$	$Var(x_{0\|500}^{MAP})$
9.7165	0.9236

Table 2: Mean and Variance of estimated initial state

study the behaviour of the smoother where the initial distribution is of larger interval. Starting with $p(x_0) \sim U[-40,\ 40]$, the (backward) evolution of the marginal smoother estimates (i.e. mean and the MAP) for the first 10 time steps for a particular realization are shown in figure 3 while the corresponding un-normalized filtered and smoothed pdfs for x_0 are shown in figure 4. It is interesting to note that the smoothed pdf of the initial state is bimodal (the smaller peak is near -10). Although the dominant mode is very close to the true initial state, $x_0^* = 10$, the contribution from the weaker mode, shifts the smoothed mean away from x_0^* (as seen from figure 3, the smoothed mean is near 8 here). This further strengthens the justification of using MAP in such scenario.

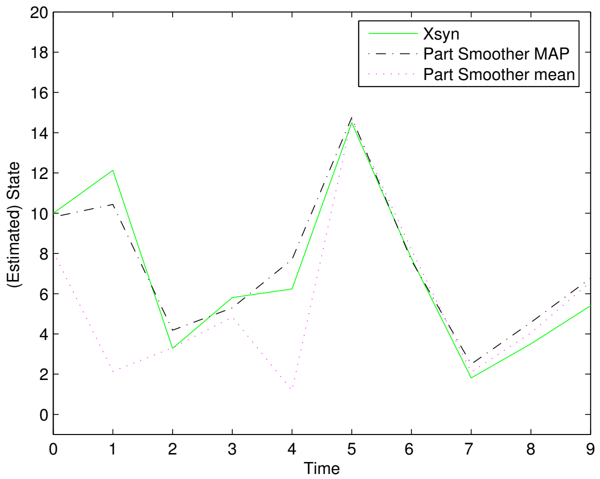

Figure 3: MAP and mean of the marginal smoothed posterior for the first 10 time steps

5.2 Parameter estimation

One of the common approaches of estimating a parameter in a state-space model is to augment the parameter as an extra state with a small artificial dynamics and then take the filtered estimate as the estimate of the parameter. The artificial evolution, however, in effect, renders the fixed parameter into a slowly varying one. As a result, the variance of the filtered estimate of the parameter goes on increasing with time ([21]) which limits the precision of the resulting estimate. Looking from another perspective in this augmented framework, one may observe that only the initial augmented state is not corrupted by artificial noise. Hence in our approach, we consider the marginal smoother of the initial augmented state to be the estimate of the true (fixed) parameter. It is expected that as more and more observations are available, the smoothed estimate would converge to the true parameter value. We proceed here with the following dynamic system:

$$x_{k+1} = f(x_k, w_{k+1}; \theta), \quad (20)$$

$$y_k = h(x_k, v_k), \qquad k = 0, 1, \ldots \quad (21)$$

where θ is a fixed unknown parameter, (x_k) are the unobservable state with (known) initial prior density $p(x_0)$ and (y_k) are the observation. The process noises (w_k) are assumed to be independent of

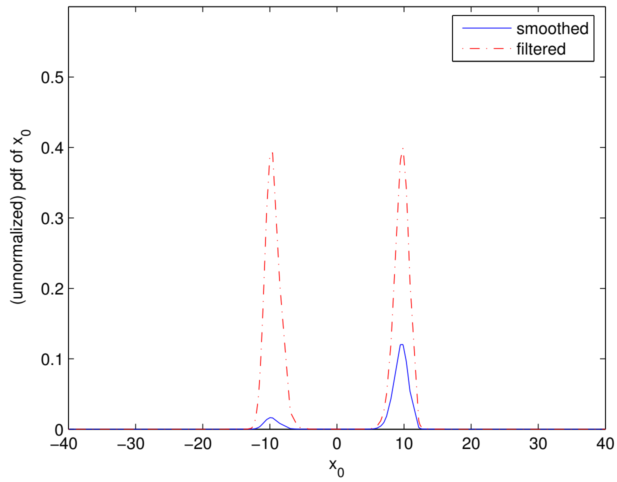

Figure 4: Filtered and smoothed probability density functions for the initial state x_0

the measurement noises (v_k). We start with the usual procedure of augmenting the state space by treating the parameter as additional state. Note that the dimension of the state increases by the numbers of parameters augmented. Now the augmented state space can be written as

$$x_{k+1} = f(x_k, \theta_k, w_{k+1}) \quad (22)$$

$$\theta_{k+1} = \theta_k + \eta_{k+1} \quad (23)$$

$$y_k = h(x_k, v_k), \qquad k = 0, 1, \ldots \quad (24)$$

with $\theta_0 = \theta$, which is unknown here. Now, using notation $X_{k+1} = [x_{k+1}\ \ \theta_{k+1}]^t$ and $W_{k+1} = [w_{k+1}\ \ \eta_{k+1}]^t$, the above model can be rewritten as

$$X_{k+1} = g(X_k, W_{k+1})$$

$$y_k = h(X_k, v_k)$$

Then we estimate the initial state vector X_0 using marginal MAP smoother. The corresponding estimation for the augmented state θ_0 is taken as the estimated parameter. We consider the following two numerical examples for this parameter estimation approach. We begin with a linear example:

$$x_k = \theta x_{k-1} + w_k \quad (25)$$

$$y_k = x_k + v_k \quad (26)$$

with $w_k \sim N(0,1)$ and $v_k \sim N(0, 0.1)$ and (unknown) true parameter $\theta = \theta^* = 0.5$. We take $\eta_k \sim N(0, 0.0025)$. Note that θ_0 is independent of x_0. With $p(x_0) \sim N(0,5)$, we started with $p(\theta_0) \sim U[-5,\ 5]$. We use $N = 1000$ particles and state transition density as our proposal during forward filtering step. The mean and variance of the estimator of θ over 30 Monte Carlo runs is shown in Table 3 below. In this case as well, we see the same type of results as in the previous subsection. Although the assumption of uniform initial prior is radically different from the knowledge of exact initial condition (parameter), we see the parameter estimate to be quite good. Next we consider the following nonlinear example:

$Mean(\theta_{0\|500}^{MAP})$	$Var(\theta_{0\|500}^{MAP})$
0.4220	0.0700

Table 3: Mean and Variance of estimated parameter

$$x_k = \frac{x_{k-1}}{2} + \frac{\theta x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + w_k, \quad (27)$$

$$y_k = \frac{x_k^2}{20} + v_k, \quad (28)$$

where $w_k \sim N(0,10)$ and $v_k \sim N(0,1)$. The true parameter is $\theta = \theta^* = 25$. With known $p(x_0) \sim N(0,5)$, we started with $p(\theta_0) \sim U[-50, 50]$. We use $N = 1000$ particles and state transition density as proposal during forward filtering step. We set $\eta_k \sim N(0,5)$. The estimate of θ for 30 Monte Carlo runs is shown in Table 4. As remarked after Table 3, we see the same pattern in a nonlinear problem as well.

$Mean(\theta_{0\|500}^{MAP})$	$Var(\theta_{0\|500}^{MAP})$
27.2595	1.5410

Table 4: Mean and Variance of estimated parameter

6. CONCLUSION

We have presented a new method for the MAP state estimate from the weighted particles representation of the smoother distribution. We applied it to estimate the unknown initial state of a dynamic system and used this approach to the parameter estimation problem. We observed that this estimation procedure works quite well even in nonlinear cases. Furthermore, as observed from our numerical examples, the smoothing density may be multimodal, which accentuates the need of such MAP estimators. There are several possibilities to extend this work. We are currently looking into the issues of estimating multiple parameters as well as simultaneous estimation of initial state and parameters. As stated earlier, the smoothing distribution here relies on the supports generated during the filtering operation. One may look into the aspect of generating different supports for smoothing in the context of smoother MAP estimation. We note that the MAP estimator in equation (13) is based on the discrete particle approximation of the continuous state space and thus limited to selecting one among those particles. This may lead to a coarse estimate. However, one may refine the estimate by using equation (11) with continuous optimization techniques. Finally, the computational load is a major concern and we plan to look into this in more details in the future.

7. ACKNOWLEDGEMENT

We are grateful to Dr. J. N. Driessen and Dr. Y. Boers of THALES Nederland B.V. for explaining to us their particle filter MAP algorithm which led to the development of this work.

REFERENCES

[1] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A Tutorial on Particle Filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE transaction on signal processing*, vol. 50(2), pp. 174–188, Feb. 2002.

[2] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Academic Press,New York, 1995.

[3] M. Briers, A. Doucet and S. R. Maskell, "Smoothing algorithm for state-space models," *Tech. Report CUED/F-INFENG/TR.498*, Cambridge University Engineering Department , Aug. 2004.

[4] A. Doucet, S. Godsill and C. Andrieu , "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[5] J. N. Driessen and Y. Boers, "Particle filter MAP estimation in dynamical systems, " in *The IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, Birmingham, UK, April. 2008.

[6] J. N. Driessen and Y. Boers, "MAP Estimation in Nonlinear Dynamic Systems" *submitted to IEEE transactions on signal processing* , 2008.

[7] H. A. P. Blom, E. A. Bloem, Y. Boers and J. N. Driessen, "Tracking closely spaced targets:Bayes outperformed by an approximation?, "in *Proc. of the 11th International Conference on Information Fusion* , Cologne, Germany, 2008, accepted.

[8] S. Godsill, A. Doucet and M. West, "Maximum a posteriori sequence estimation using Monte Carlo particle filters," *Annals of the Institute of Statistical Mathematics* , vol. 53, pp. 82–96, 2001.

[9] D. Guo, X. Wang and R. Chen, "New sequential Monte Carlo methods for nonlinear dynamic systems," *Statistics and Computing*, vol. 15(2), pp. 135–147, Apr. 2005.

[10] G. Kitagawa, "Non-Gaussian state-space modeling of nonstationary time series," *Journal of the American Statistical Association*, vol. 82(400), pp. 1032–1063, 1987.

[11] M. Klaas, M. Briers, N. de Freitas, S. Maskell and D. Lang, "Fast Particle smoothing: if I had a million particles," in *Proc. of the 23rd International Conference on Machine Learning* , Pittsburgh, Pennsylvania, 2006, pp. 481–488.

[12] J. S. Liu and R. Chen, "Sequential Monte Carlo Methods for Dynamic Systems," *Journal of the American Statistical Association*, vol. 93(443), pp. 1032–1044, 1998.

[13] M. Hurzeler and H. R. Kunsch,"Monte Carlo Approximations for General State-Space Models, " *Journal of Computational and Graphical Statistics*, vol. 7(2), pp. 175–193, June. 1998.

[14] S. Saha, P. K. Mandal, Y. Boers, H. Driessen and A. Bagchi, "Gaussian proposal density using moment matching in SMC methods," *Memorandum 1832, Department of Applied Mathematics, University of Twente*, ISSN 1874-4850, pp. 11–25, 2007.

[15] B. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC, 1986.

[16] S. K. Zhou, R. Chellappa and B. Moghaddam, "Visual Tracking and Recognition using appearance -adaptive Models in particle filters," *IEEE transaction on image processing*, vol. 13(11), pp. 1491–1506, Nov. 2004.

[17] S. Saha, P. K. Mandal, Y. Boers and H. Driessen, "Exact moment matching for efficient importance functions in SMC methods," in *Proc. of NSSPW 2006* , Cambridge,UK, 2006.

[18] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F - Radar and Signal Processing*, vol. 140(2), pp. 107–113, 1993.

[19] O. Cappe, S. J. Godsill and E. Moulines,"An overview of existing methods and recent advances in sequential Monte Carlo," *IEEE Proceedings*, vol. 95(5), pp. 899–924, 2007.

[20] J. V. Candy, "Bootstrap particle filtering," *IEEE Signal Processing Magazine*, vol. 24(4), pp. 73–85, July. 2007.

[21] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," *Sequential Monte Carlo Methods in Practice (A. Doucet, N. D. Freitas and N. Gordon Eds.)* Springer, 2001.