Copyright © 2002 IFAC
15th Triennial World Congress, Barcelona, Spain

IMPROVEMENT OF THE PARTICLE FILTER BY BETTER CHOICE OF THE PREDICTED SAMPLE SET

J.P. Norton and G.V. Veres

School of Engineering, The University of Birmingham, Edgbaston, B15 2TT, UK
E-mail: j.p.norton@bham.ac.uk

Abstract: An improvement of the standard "particle filter" (PF) Monte Carlo Bayesian estimator is presented and compared with an existing improved reweighted filter in a target tracking example. The PF updates the probability density function (pdf) of the state, represented as the density of state samples (particles). Each particle is time-updated by applying to the state equation a sample from the forcing distribution. At the next observation, the likelihood of each particle is computed by substituting the prediction error into the observation-noise pdf. Any low-likelihood particle has a low probability of appearing in the resampled state set for the next update, so often the sample set collapses. The improved estimator represents the state pdf as weighted samples, and allows free choice of the values at which the posterior pdf is evaluated. This allows enough particles in regions of low probability density and avoids the need for most particles to be in high-density regions.

Keywords: Particle filters, Nonlinear systems, Bearings only tracking, Bayesian estimation, Discrete-time systems.

1. INTRODUCTION

The Bayesian approach to state estimation constructs the probability density function (pdf) of the current state of an evolving system, given the accumulated observation history. For linear Gaussian models, where the pdf can be summarised by means and covariances, the estimation is carried out by the familiar Kalman filter (Jazwinski, 1970). For non-linear non-Gaussian models, there is no simple analytical way to proceed.

Several approximate methods have been proposed, for example the Gaussian sum filter (Alspach and Sorenson, 1972), or numerical integration over a grid in state space (Kitagawa, 1987), (Kramer and Sorenson, 1988); (Bucy and Senne, 1970); (Challa and Bar-Shalom, 2000). It is not easy to decide how many components to use in a Gaussian sum, and the number required is likely to be excessive when the pdf has abrupt changes (e.g. due to bounds on state). In the numerical-integration methods, time updating of the state involves solution of the Chapman-Kolmogorov or Fokker-Planck-Kolmogorov equation. Finite-difference or finite-element methods may be employed, but to keep computing load tolerable the number of grid points must be restricted, which requires tracking of the domain over which the state pdf is non-negligible. The use of Chebyshev's inequality for time-varying truncation of the support of the state pdf has been suggested (Challa and Bar-Shalom, 2000) but even so adequate coverage by a grid may entail very heavy computing unless the state dimension is low.

Particle filters provide a cheaper approximate solution to Bayes estimation of states for non-linear systems. The pdf is represented as the local

density of a scatter of particles, each of which propagates through state space as time elapses (Polson *et al.*, 1992);(West, 1992).The samples are thus concentrated in regions of high probability density. However, in some applications this is a disadvantage, when useful information resides in the tails of the distribution. Moreover, the sample density may prematurely become zero in the tails, preventing later growth in probability density there. This may occur over large regions, because in resampling after the observation update, some sample values have low enough likelihood to be absent from the resampled set (Gordon *et al.*, 1993). The particle set can even collapse to a single point, for example when observability is poor. To prevent this, several methods have been proposed. They modify the sample set, the prediction process or the representation of the state pdf. Among the former, roughening (Gordon *et al.*, 1993) adds jitter as the new samples are drawn, while prior editing rejects samples with low likelihoods. A weighted-particle filter recently proposed (Veres and Norton, 2001) summarises the posterior marginal distributions of the state variables by specified percentile points, computed from the sample set. This grossly simplifies the state pdf by representing it as the Cartesian product of the marginal densities. To mitigate this, a recovery procedure for the state pdf was added, drawing additional samples in regions of high density. Alternatively, several other improved resampling schemes have been developed recently. The basic idea is to eliminate trajectories which have small normalised importance weights and to concentrate upon trajectories with large weights. A suitable measure of degeneracy of the algorithm by the effective sample size N_{eff} was introduced in (Liu, 1996). When $\hat{N}_{eff}$ is below a fixed threshold N_{eff}, the sequential importance resampling (SIR) procedure is used (Rubin, 1988). It is possible to implement the SIR procedure exactly in $O(N)$ operations by using a classical algorithm (Ripley, 1987); (Carpenter *et al.*, 1999); (Doucet *et al.*, 2000); (Pitt and Shephard, 1999). Other resampling procedures which reduce the Monte Carlo variation, such as stratified sampling (Carpenter *et al.*, 1999) and residual sampling (Liu and Chen, 1998) may be applied as an alternative to SIR.

In this paper, the conventional PF is modified to allow better approximation of $p(\mathbf{x}_t|\mathbf{Y}_{t-1})$ by any given number of particles, i.e. the new approach prevents collapse of the particle set by altering the prediction stage. In the conventional PF, each predicted particle results from applying a single sample of forcing to a single prior particle; there is no integration over the range of forcing and prior state which could have given rise to the specified predicted state.The proposed estimator intead implements approximate convolution of the unforced-state pdf with the forcing pdf, by summing over a number of prior particles, each weighted. The domain over which the predicted state pdf is computed is thereby put at the user's discretion instead of following from the sample properties of the prior particle set and forcing sample set. Collapse of the particle set can therefore be prevented, as can over-confidence, and extra attention can be paid to any especially important region of state space (e.g. the region around any potential target of a hostile vehicle being tracked). Of course, the problem of defining the domain over which to predict the state pdf must be faced, just as in the numerical-integration approaches mentioned above. However, the particles need not be located at regular grid points but can be placed to take account of the latest observation and any bounds such as limits on speed or manoeuvre capability. There is no claim that the proposed method is preferable to all other estimators. Rather, the choice of estimator will be determined by the demands of the application.

2. BAYESIAN DISCRETE-TIME STATE ESTIMATION

Consider the generally non-linear discrete-time state-space model of the form:

$$\mathbf{x}_{t+1} = \mathbf{f}_t(\mathbf{x}_t, \mathbf{w}_t),$$

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t) + \mathbf{v}_t,$$

where functions $\mathbf{f}_t$ and $\mathbf{h}_t$ are known and $\mathbf{x}_t \in \Re^n$ is the sampled state vector at sampling instant t. The observation vector is $\mathbf{y}_t \in \Re^m$ and $\mathbf{w}_t \in \Re^l$ is a random disturbance which can describe both unknown forcing and the results of inaccuracies in model structure. It will be assumed that $\mathbf{w}_t$ has a known pdf $p(\mathbf{w})$. Measurement noise $\mathbf{v}_t$ has a known pdf $p(\mathbf{v})$.

The problem is to estimate at time t_+ the state $\mathbf{x}_t$, by estimating the posterior probability density of $\mathbf{x}_t$ based on past output samples $Y^t \stackrel{def}{=} \{\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_0\}$. The conditional density of $\mathbf{x}_t$ given $\mathbf{x}_{t-1}$ can be derived from the knowledge of $\mathbf{f}_t$ and $p(\mathbf{w}_t)$ and will be denoted by $p(\mathbf{x}_t|\mathbf{x}_{t-1})$. An initial estimate of the pdf $p(\mathbf{x}_0)$ of $\mathbf{x}_0$ is assumed to be known. The probability density of $\mathbf{x}_t$ conditional on Y^{t-1} can then be obtained, in principle, by the time update

$$p(\mathbf{x}_t|Y^{t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1}) p(\mathbf{x}_{t-1}|Y^{t-1}) d\mathbf{x}_{t-1}$$

The observation update then finds the posterior pdf, in theory, by Bayes' rule:

$$p(\mathbf{x}_t|Y^t) = \frac{p(\mathbf{y}_t|\mathbf{x}_t) p(\mathbf{x}_t|Y^{t-1})}{p(\mathbf{y}_t|Y^{t-1})} \qquad (1)$$

where $p(\mathbf{y}_t|\mathbf{x}_t) = p(\mathbf{v}_t = \mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t))$, the observed $\mathbf{y}_t$ is substituted into $p(\mathbf{y}_t|\mathbf{x}_t)$ and

$$p(\mathbf{y}_t|Y^{t-1}) = \int p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|Y^{t-1})d\mathbf{x}_t \quad (2)$$

Analytical solution of this problem is only feasible for a relatively small and restricted selection of system and measurement models.

A recent development (Carpenter *et al.*, 1999) of the conventional PF is summarised next.

3. IMPROVED REWEIGHTED PARTICLE FILTER

The following scheme (Carpenter *et al.*, 1999) aims to prevent collapse of the particle set, while remaining computationally efficient.

Initialisation Start with a random measure with N support points, possibly obtained by stratified sampling, which approximates the pdf $p(\mathbf{x}_0)$, or by drawing from the known initial pdf $p(\mathbf{x}_0)$. By a random measure is understood the particle set (support points) and their weights summing to 1.
Preliminaries (step t) Assume that we have a random measure $(\mathbf{s}_{t-1}, m_{t-1})$, approximating $p(\mathbf{x}_{t-1}|Y^{t-1})$
Prediction: Estimate the density $p(\mathbf{x}_t|Y^t)$, up to a normalising constant K, by

$$p(\mathbf{x}_t|Y^t) = K\sum_{i=1}^{N} m_{t-1}^i p(\mathbf{x}_t|\mathbf{s}_{t-1}^i)p(\mathbf{y}_t|\mathbf{x}_t).$$

Construct an approximating mixture density

$$\hat{p}(\mathbf{x}_t) = \sum_{i=1}^{N} \hat{\beta}_i \hat{p}_i(\mathbf{x}_t).$$

Take a stratified sample from this density using the stratified sampling algorithm (see (Carpenter *et al.*, 1999)), with N_i sample points in the ith category, where N_i has expected value $N\beta_i$
Update: For each i, sample N_i support points $\mathbf{s}_t^j$ from $p_i(\mathbf{x}_t)$, with importance weights given by

$$m_t^j \propto \frac{m_{t-1}^i p(\mathbf{s}_t^j|\mathbf{s}_{t-1}^i)p(\mathbf{y}_t|\mathbf{s}_t^j)}{\hat{\beta}_i \hat{p}_i(\mathbf{s}_t^j)},$$

for $\sum_{l=1}^{i-1} N_l < j \leq \sum_{l=1}^{i} N_l$. The updated random measure is then given by $[(\mathbf{s}_t^j, m_t^j),\ j = 1, \ldots, N]$, where the weights are scaled to sum to 1.

4. IMPROVED PARTICLE FILTER

4.1 *Theory*

Consider the time update

$$p(\mathbf{x}_t|Y^{t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|Y^{t-1})d\mathbf{x}_{t-1} =$$

$$\int p(\mathbf{f}(\mathbf{x}_{t-1}, \mathbf{w}_{t-1}) = \mathbf{x}_t)p(\mathbf{x}_{t-1}|Y^{t-1})d\mathbf{x}_{t-1}. \quad (3)$$

The standard PF only employs one sample of $\mathbf{w}_{t-1}$ per (sample) value of $\mathbf{x}_t$, instead of integrating over $\mathbf{x}_{t-1}$. With $p(\mathbf{x}_{t-1}|Y^{t-1})$ represented by N equal-weight particles as

$$p(\mathbf{x}_{t-1}|Y^{t-1}) = \frac{1}{N}\sum_{j=1}^{N} \delta(\mathbf{x}_{t-1}^j) \quad (4)$$

(3) gives

$$p(\mathbf{x}_t|Y^{t-1}) = \frac{1}{N}\sum_{j=1}^{N} \int \delta(\mathbf{x}_{t-1}^j) \times \quad (5)$$

$$p(\mathbf{f}(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1}) = \mathbf{x}_t)d\mathbf{x}_{t-1} = \frac{1}{N}\sum_{j=1}^{N} p(\mathbf{w}_{t-1}^j),$$

where $\mathbf{w}_{t-1}^j$ is the value (assumed unique for the moment) which satisfies $\mathbf{f}(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1}^j) = \mathbf{x}_t$.

Using (5), Bayes' rule (1) has as its numerator at any specified value of $\mathbf{x}_t$

$$\frac{p(\mathbf{v}_t = \mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t))}{N}\sum_{j=1}^{N} p(\mathbf{w}_{t-1}^j). \quad (6)$$

Using (6), we can select any sample values of $\mathbf{x}_t$ we like and compute their probability densities, representing $p(\mathbf{x}|Y)$ as the sum of modulated (weighted) delta functions. In the next updating cycle, the delta functions in (4) and thence the terms in the sum in (5) carry these weights. The particles can be located wherever necessary to characterise the state pdf adequately, and are no longer wasted in regions of high but near-uniform density nor necessarily absent from regions of low density which may, nevertheless, lead to high density later. Common sense suggests curvature as a criterion for their location: interpolation between particles is less accurate the higher the local curvature, so density of particles should (all other things being equal) be highest where the curvature is highest. An additional criterion is that coverage should be adequate over the whole support region where probability density is above some some selected threshold, to prevent premature collapse into the highest-likelihood region.

It has so far been assumed that the state equation has a unique solution $\mathbf{w}_{t-1}^j(\mathbf{x}_{t-1}^j, \mathbf{x}_t)$ for use in (5). This is usually so if $\dim(\mathbf{w}_{t-1}) = \dim(\mathbf{x}_t)$, but not in the commoner case $\dim(\mathbf{w}_{t-1}) < \dim(\mathbf{x}_t)$, as in target tracking. The latter may be handled by taking a subset $\mathbf{x}_t'$ of the elements of $\mathbf{x}_t$ for which the state equations $\mathbf{x}_t' = \mathbf{f}'(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1}^j)$

have a unique solution $\mathbf{w}_{t-1}^j(\mathbf{x}_t')$, then using $\mathbf{x}_t''^j = \mathbf{f}''(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1}^j)$ to find the rest $\mathbf{x}_t''^j$ of $\mathbf{x}_t^j$. This gives

$$p(\mathbf{x}_t^j|Y^{t-1}) \equiv p(\mathbf{x}_t', \mathbf{x}_t''^j) = p(\mathbf{w}_{t-1}^j(\mathbf{x}_t'))p(\mathbf{x}_{t-1}^j)$$

(allowing for the particles at time $t-1$ having weights $p(\mathbf{x}_{t-1}^j|Y_{t-1})$, instead of all having the same weight as in the usual PF).

Now we have a free choice of the values $\mathbf{x}_t'$. The chosen values need only cover the subspace of $\mathbf{x}_t'$ adequately, so N' can be much smaller than N. Good coverage of the state space just depends on choosing the values of $\mathbf{x}_t'$ intelligently; in particular, we should avoid wasting effort on values giving very big innovations. It is reasonable to ignore $\mathbf{x}_t'$ giving innovations $\nu_t^j = \mathbf{y}_t - \mathbf{h}'(\mathbf{x}_t'^j)$ with, say, $|\nu_{ti}^j| \geq 3r_{ii}^{1/2}$, $1 \leq i \leq m$, where the observation noise covariance matrix $\mathbf{R} = diag(r_{ii})$. We can economize further by dropping values $\mathbf{x}_t'$ which give very small $p(\mathbf{x}_t', \mathbf{x}_t''^j|Y_{t-1})$ as soon as $\mathbf{w}_{t-1}^j(\mathbf{x}_t')$ has been computed. This process of distributing the $\mathbf{x}_t'$ values according to the behaviour of $p(\mathbf{x}_t'|Y_t)$ suggests adaptive placing, deciding where to put successive values according to the results of the observation update at time t for values chosen earlier in the same update cycle.

For the next update cycle, we have to generate N samples of $\mathbf{x}_t$ from $N'N''$ particles $(\mathbf{x}_t', \mathbf{x}_t'')$. Some possible schemes are:

a) Make $N'N'' = N$ and use the same particles;
b) With $N'N'' > N$, take the N samples with the largest weights (risking losing tails);
c) With $N'N'' > N$, take the N samples from the weighted particles, with equal probabilities. If a particle is sampled more than once, throw the second one away and carry on until there are N particles in all;
d) With $N'N'' > N$, take $N_1, N_2, \ldots, N_m$ samples from the particle sets with weights in specified percentile ranges $1, 2, \ldots, m$.

N equal-weight samples are taken with replacement from the new particle set with probabilities defined by their weights.

4.2 *Algorithm*

1. Draw N samples $\mathbf{x}_0 = \mathbf{x}_0^i : i = 1, \ldots, N$ from the known prior $p(\mathbf{x}_0)$. Divide the state variables into subsets $\mathbf{x}_t'$ and $\mathbf{x}_t''$.
2. At each sampling time $t \geq 1$:
   (a) Define bounds for subsets $\mathbf{x}_{t|t-1}'$ and $\mathbf{x}_{t|t-1}''$.
   (b) Draw N' samples of $\mathbf{x}_{t|t-1}'$ within the defined bounds.
   (c) For each sample $\mathbf{x}_{t|t-1}'^i$, $i = 1, \ldots, N'$
      (1) Solve state equation $\mathbf{x}_{t|t-1}'^i = \mathbf{f}'(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1})$ for each $j = 1, \ldots, N$ to get $\mathbf{w}_{t-1}^j(\mathbf{x}_{t|t-1}'^i)$. Choose only those $\mathbf{w}_{t-1}^j(\mathbf{x}_{t|t-1}'^i)$ within the defined bounds and calculate their probabilities.
      (2) Use $\mathbf{x}_{t|t-1}''^j = \mathbf{f}''(\mathbf{x}_{t-1}^j, \mathbf{w}_{t-1}^j)$ to find the rest $\mathbf{x}_{t|t-1}''^j$ of $\mathbf{x}_t^j$. Choose only those $\mathbf{x}_{t|t-1}''^j$ within the bounds
      (3) Form set of $N'N''$ samples $\mathbf{X}_{t|t-1} = (\mathbf{x}_{t|t-1}', \mathbf{x}_{t|t-1}'')$ and set of corresponding probabilities of the process noise.
   (d) If set $\mathbf{X}_{t|t-1}$ is empty, loosen the bounds and return to step (b).
   (e) Apply Bayes' rule (1),(6) to the $N'N''$ samples
   (f) Generate N samples of $\mathbf{x}_t$ from the $N'N''$ particles $(\mathbf{x}_{t|t-1}', \mathbf{x}_{t|t-1}'')$

5. SIMULATION

Here the performance of the new scheme is compared with the improved reweighted PF (IRPF) (Carpenter *et al.*, 1999) in a bearings-only tracking problem (Gordon *et al.*, 1993).

A target moves in the $x-y$ plane (with x, y Cartesian coordinates) according to the second-order motion model

$$\mathbf{x}_{t+1} = F\mathbf{x}_t + \Gamma\mathbf{w}_t, \tag{7}$$

where $\mathbf{x}_t = (x, \dot{x}, y, \dot{y})_t^T$, $\mathbf{w}_t = (w_x, w_y)_t^T$,

$$F = \begin{pmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Gamma = \begin{pmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{pmatrix}$$

where the sampling interval is $T = 1$. The process noise is zero-mean Gaussian white noise with covariance matrix $\mathbf{Q}$: $E[\mathbf{w}_k\mathbf{w}_j^T] = \mathbf{Q}\delta_{jk}$, where $\mathbf{Q} = q\mathbf{I}_2$ and $\mathbf{I}_2$ is the 2×2 identity matrix.

A fixed observer at the origin takes noisy measurements z_t of the target bearing

$$z_t = tan^{-1}(y_t/x_t) + v_t. \tag{8}$$

The measurement noise is a zero-mean Gaussian white noise with variance r: $E[v_k v_j] = r\delta_{kj}$; for the example $\sqrt{r} = 0.002$.

The initial actual target state is

$$\mathbf{x}(0) = [5000 \quad -76 \quad 3000 \quad -184]^T.$$

For the first $10s$ the target moves with constant velocity and then it begins a right turn with cross-track acceleration of $4g$. The turn lasts for $10s$ then the target moves with constant velocity. After a further $7s$ the target turns left with cross-track acceleration of $5g$ and continues turning for

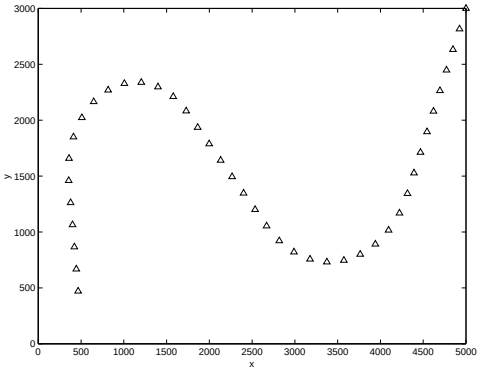

Fig. 1. True target track in the $x-y$ plane: $\triangle$ – target position (start: top right)

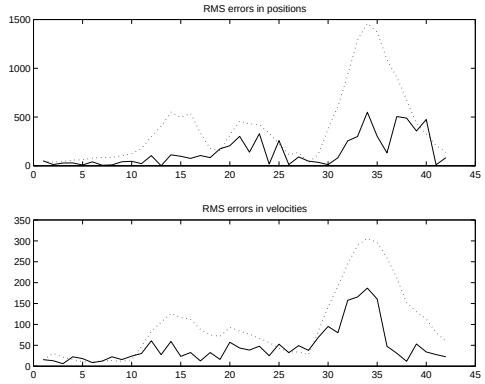

Fig. 2. Solid line - IPF, dotted line - IRPF

$10s$. When the turn is finished the target moves with constant velocity for another $5s$. Fig. 1 gives the actual target path in the $x-y$ plane.

The prior information specified is

$$\hat{\mathbf{x}}(0) = [5040 \quad -66 \quad 2970 \quad -172]^T.$$

and the initial state covariance matrix is diagonal with elements $p_{11}(0) = 1600$, $p_{22}(0) = 100$, $p_{33}(0) = 900$, $p_{44}(0) = 144$.

The process noise standard deviation was $50m/s^2$ for IPF and $25m/s^2$ for IRPF. Different values of the process noise standard deviation were chosen to obtain the best possible performance for each filter.

The initial particles are assumed to be normally distributed $\mathcal{N}(\hat{\mathbf{x}}(0), P)$. The number of particles for the IRPF was $N = 4000$ and for IPF was $N = 100$. Construction of the approximating mixture density for bearings-only problems can be found in (Carpenter *et al.*, 1999).

This implementation of the proposed approach for this example consists of:

1) Prediction of the new states from the previous time instant

$$\hat{\mathbf{x}}^j_{t|t-1} = \mathbf{f}(\hat{\mathbf{x}}^j_{t-1}, \mathbf{w}^j_{t-1}).\ j = 1, \ldots, N$$

2) Definition of the bounded set from which to draw updated positions.

   Since the measurements are in polar coordinates and state vector in Cartesian coordinates, position samples are initially drawn in polar coordinates, subject to polar bounds, then converted to Cartesian coordinates. The bound for azimuth is straightforward; after the observation is obtained the samples are selected so that azimuth is in the interval $[z_t - 3\sqrt{r}, z_t + 3\sqrt{r}]$. This ensures that the particles will not have negligible likelihoods. To improve the quality of the estimation we need at least a rough estimate of range. Samples which satisfy the bounds on predicted azimuth are chosen, and the minimum and maximum of predicted range from the chosen samples are selected as bounds for range.
3) Draw $N1 > N$ ($N1 = 300$ in this example) samples in polar coordinates within the bounds defined above. Convert polar coordinates to Cartesian coordinates. Note that in this realisation we do not aim to obtain a feasible solution from every updated position.
4) For each obtained position
   (i) solve $\mathbf{x}'_{t|t-1} = \mathbf{f}'(\mathbf{x}^j_{t-1}, \mathbf{w}_{t-1})$ to obtain a unique solution $\mathbf{w}^j_{t-1}(\mathbf{x}'_{t|t-1})$, then use $\mathbf{x}''^j_{t|t-1} = \mathbf{f}''(\mathbf{x}^j_{t-1}, \mathbf{w}^j_{t-1})$ to find the rest $\mathbf{x}''^j_{t|t-1}$ of $\mathbf{x}^j_{t|t-1}$;
   (ii) apply the bounds on $p(\mathbf{w}_{t-1})$ and target's speed, retaining only samples which lie within the bounds; in this implementation the bounds applied were $\mathbf{w}^j_{t-1} \in [-3\sqrt{q}, 3\sqrt{q}]$ for forcing and $[150, 250]$ for speed, where $= [1\ 1]^T$.
   (iii) If there is a feasible solution, combine the pair of position coordinates with the feasible pair of velocity coordinates. Otherwise loosen the bounds. Repeat steps 3) and 4) until at least N samples are feasible.
5) For each particle apply Bayes' rule (1), (6).
6) Sample in two stages from the obtained sample set. First choose samples with likelihoods higher than specified threshold. If this produces fewer than N samples, reduce the threshold. Then choose the N samples with highest $p(\mathbf{w}_{t-1})$.

RMS errors in position and velocity after 1 run are presented in Fig. 2.

It can be seen from Fig. 2 that the performance of IRPF is worse than IPF for this example. Especially after $30T$, the performance of the IRPF gets much worse in both position and velocity, because of ambiguity in which direction to go. The IRPF and the new PF treat this situation differently.The IRPF tries to spread its particles as wide as possible in the direction of the ambiguity. As a result, taking the sample mean would give little useful

information about the target position. The new IPF keeps the spread of the particles within the defined bounds, preventing increasing spread of the samples. It is worth noticing that in this case we cannot guarantee that the actual target state will be in the 95% confidence region in 95% of cases, but the right choice of bounds and rules for handling them will keep the cloud of particles close to the target.

6. CONCLUSION

A IRPF with radically modified treatment of the prediction stage has been proposed. It is potentially useful in cases where there is high uncertainty in some of the state components and the sample set thus has a wide spread. The mean of the IRPF's particle states then gives large estimation errors due to the wide spread. It has been shown that IPF performance is clearly better that of IRPF for a bearings-only tracking example. However, the IPF has its own limitations; for example, if the innovation likelihoods show two or more peaks with similar heights, the ambiguity may induce poor choice of predicted-state samples.

7. ACKNOWLEDGEMENT

This work was carried out with the support of EPSRC grant GR/M04570, the Defence Evaluation and Research Agency, Farnborough, and BAE Systems, Filton.

8. REFERENCES

Alspach, D. L. and H. H. Sorenson (1972). Non-linear Bayesian estimation using Gaussian sum approximation. *IEEE Trans. on Automatic Control* **AC − 17**, 439–447.

Bucy, R. S. and K. D. Senne (1970). Realization of optimum discrete-time nonlinear estimators. *Proc. 1st Symp. on Nonlinear Estimation Theory and Its Applications*, *San Diego*, *Ca* pp. 6–17.

Carpenter, J., P. Clifford and P. Fearnhead (1999). An improved particle filter for nonlinear problems. *IEE Proceedings on Radar*, *Sonar and Navigation* **146**, 2–7.

Challa, S. and Y. Bar-Shalom (2000). Nonlinear filter design using Fokker-Planck-Kolmogorov probability density evolutions. *IEEE Transactions on Aerospace and Electronic Systems* **36(1)**, pp 309–315.

Doucet, A., S.J. Godsill and C.Andrieu (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing* **10**, 197–208.

Gordon, N.J., D.J. Salmond and A. F. M. Smith (1993). Novel approach to nonlinear/ non-Gaussian Bayesian state estimation. *IEE Proceedings on Radar*, *Sonar and Navigation* **140**, 107–113.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.

Kitagawa, G. (1987). Non-Gaussian state-space modelling of non-stationary time series. *Journal of the American Statistical Association* **82**, 1032–1063.

Kramer, S. C. and H. W. Sorenson (1988). Recursive Bayesian estimation using piecewise constant approximation. *Automatica* **24**, 789–801.

Liu, J.S. (1996). Metropolized independent sampling with comparison to rejection sampling and importance sampling. *Statistics and Computing* **6**, 113–119.

Liu, J.S. and R. Chen (1998). Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association* **93**, 1032–1044.

Pitt, M.K. and N. Shephard (1999). Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association* **94**, 590–599.

Polson, N.G., B.P. Carlin and D. S. Stoffer (1992). A Monte-Carlo approach to non-normal and nonlinear state-space modelling. *Journal of the American Statistical Association* **87**, 493–500.

Ripley, B.D. (1987). *Stochastic Simulation*. Wiley. New York.

Rubin, D.B. (1988). Using the SIR Algorithm to Simulate Posterior Distributions. In *Bayesian Statistics*, **vol.3** (J.M. Bernando *et al.* (Ed)). pp. 395–402. Oxford University Press.

Veres, G. V. and J. P. Norton (2001). Weighted particle filter for state estimation. *Proceedings of the IASTED International Conference : Control and Applications*, *Banff*, *Canada* pp. 115–120.

West, M. (1992). Modelling with mixtures. In *Bayesian Statistics*, **vol.4** (J.M. Bernardo *et al.* (Ed)), pp. 503-524. Oxford University Press.