A Particle filter Using SVD Based Sampling Kalman Filter to Obtain the Proposal Distribution

Bin Liu
Graduate University, The Chinese Academy of Sciences
Institute of Acoustics, The Chinese Academy of Sciences
Beijing, China
binliu1981@yahoo.com

Xiao-chuan Ma, Chao-huan Hou
Institute of Acoustics, The Chinese Academy of Sciences
Beijing, China
{maxc,hch}@mail.ioa.ac.cn

Abstract—**In this paper, we propose a novel particle filter (PF), which uses a bank of singular-value-decomposition based sampling Kalman filters (SVDSKF) to obtain the importance proposal distribution. This proposal has two properties. Firstly, it allows the particle filter to incorporate the latest observations into a prior updating routine and, secondly it inherits advantage of having good numerical stability from the singular-value-decomposition (SVD). The convergence results of the numerical simulations we made confirm that the proposed PF method outperforms the standard bootstrap PF as well as other local linearization based PFs.**

Keywords—**Particle Filter, SVD, proposal distribution, SRUKF**

I. INTRODUCTION

To solve the problem of nonlinear problems, the best known filtering algorithms are EKF (extended Kalman filter)[1,2] and UKF (unscented Kalman filter) [1,3~8]. Furthermore, particle filters (PFs) are proposed to solve the problem of non-Gaussian[8~12]. PFs rely on importance sampling techniques which lead to the requirement for the design of proposal distributions in order to approximate the posterior distribution reasonably well. In general, it is hard to design such proposals. The most common strategy is to sample from the probabilistic model of the states evolution (transition prior) which results in the generation of standard bootstrap PF. However, this strategy is easy to fail if the new measurements appear in the tail of the prior or if the likelihood is too peaked in comparison to the prior. These situations do indeed arise in several areas of engineering and finance, where one can encounter sensors that are very accurate (peaked likelihoods) or data that undergoes sudden changes (nonstationarities) [12]. To overcome this problem, researchers proposed to use local linearization techniques, such as EKF or UKF, to generate the proposal distribution of PFs[12]. However, EKF has its inherent drawbacks due to its first-order linearization and UKF and its basis unscented transform (UT) often encounter the ill-conditioned problem of covariance matrix in practice (though it is theoretically positive semi-definite) [8].

The singular-value-decomposition (SVD) concerns the factorization of an arbitrary matrix A into a product UDV' of orthogonal matrices U and V and a "diagonal" matrix D. It's applied frequently in numerical linear algebra and proved to be a robust method for solving ill-conditioned least squares problems[13~15]. Youmin Zhang proposed a SVD based EKF with a good numerical stability in application to aircraft flight state and parameter estimation[2]. Z.Chen presented a SVD sampling Kalman filter (SVDSKF) in [8] which is similar as UKF but can avoid the ill-condition problems that may be encountered when using UKF.

In this paper, we use the SVDSKF to obtain the importance proposal distribution of the PF. Then we get a novel PF method.

II. DYNAMIC STATE SPACE MODEL

Since we are interested in nonlinear, non-Gaussian regression, the state space model can be expressed as follows

$$\begin{aligned} \mathbf{x}_t &= \mathrm{f}\left(\mathbf{x}_{t-1}\right) + \mathbf{v}_{t-1} \\ \mathbf{y}_t &= \mathrm{h}\left(\mathbf{x}_t\right) + \mathrm{n}_t \end{aligned} \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^{n_x}$ denotes the unobserved states (or parameters) of the model, $\mathrm{y}_t \in \mathbb{R}^{n_y}$ the observations, $\mathrm{v}_t \in \mathbb{R}^{n_v}$ the process noise and $\mathrm{n}_t \in \mathbb{R}^{n_n}$ the measurement noise. To complete the specification of the model, the prior distribution (at $t = 0$) is denoted by $p\left(\mathbf{x}_0\right)$. Our goal will be to approximate the posterior distribution $p\left(\mathbf{x}_{0:t} \mid \mathrm{y}_{1:t}\right)$ and one of its marginals, the filtering density $p\left(\mathbf{x}_t \mid \mathrm{y}_{1:t}\right)$, where $\mathrm{y}_{1:t} = \left\{\mathrm{y}_1, \mathrm{y}_2, \cdots, \mathrm{y}_t\right\}$. By computing the filtering density recursively, we do not need to keep track of the complete history of the states.

III. GENERIC PARTICLE FILTER

For completeness, we present a generic PF algorithm here, which involves the following steps as in table I. PFs allow us to approximate the posterior distribution $p\left(\mathbf{x}_{0:t} \mid \mathrm{y}_{1:t}\right)$ using a set of N weighted samples (particles) $\left\{\mathrm{x}_{0:t}^{(i)}; i = 1, \cdots, N\right\}$,which are drawn from an importance proposal distribution $q\left(\mathbf{x}_{0:t} \mid \mathrm{y}_{1:t}\right)$. In framework of the PF algorithm, we can restrict ourselves to importance functions of the form .

$$q\left(\mathbf{x}_{0:t} \mid \mathrm{y}_{1:t}\right) = q\left(\mathbf{x}_0\right) \prod_{k=1}^{t} q\left(\mathbf{x}_k \mid \mathrm{y}_{1:k}, x_{1:k-1}\right) \tag{2}$$

to obtain a recursive formula to evaluate the importance weights[9]

$$w_t \propto \frac{p\left(\mathrm{y}_t \mid \mathrm{y}_{1:t-1}, \mathrm{x}_{0:t}\right) p\left(\mathrm{x}_t \mid \mathrm{x}_{t-1}\right)}{q\left(\mathrm{x}_t \mid \mathrm{y}_{1:t}, \mathrm{x}_{1:t\text{-}1}\right)} \tag{3}$$

TABLE I . GENERIC PF

1.Sequential importance sampling step
● For $i=1,\cdots,N$, sample $\tilde{x}_t^{(i)} \sim q\left(x_t \mid x_{0:t-1}^{(i)}, y_{1:t}\right)$ and update the trajectories $\tilde{x}_{0:t}^{(i)} \triangleq \left(\tilde{x}_t^{(i)}, \tilde{x}_{0:t-1}^{(i)}\right)$
● For $i=1,\cdots,N$, evaluate the importance weights up to a normalizing constant: $w_t^{(i)} = \dfrac{p\left(\tilde{x}_{0:t}^{(i)} \mid y_{1:t}\right)}{q\left(\tilde{x}_t^{(i)} \mid \tilde{x}_{0:t-1}^{(i)}, y_{1:t}\right) p\left(\tilde{x}_{0:t-1}^{(i)} \mid y_{1:t-1}\right)}$
● For $i=1,\cdots,N$, normalize the weights: $\tilde{w}_t^{(i)} = w_t^{(i)}\left[\sum_{j=1}^{N} w_t^{(j)}\right]^{-1}$
2.Selection step
● Multiply/suppress samples $\left(\tilde{x}_{0:t}^{(i)}\right)$ with high/low importance weights $\tilde{w}_t^{(i)}$, respectively, to obtain N random samples $\left(x_{0:t}^{(i)}\right)$ approximately distributed according to $p\left(x_{0:t}^{(i)} \mid y_{1:t}\right)$
● For $i=1,\cdots,N$, set $w_t^{(i)} = 1/N$
3. Output:
The output of the algorithm is a set of samples that can be used to approximate the posterior distribution as follows
$p\left(x_{0:t}^{(i)} \mid y_{1:t}\right) \approx \hat{p}\left(x_{0:t}^{(i)} \mid y_{1:t}\right) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\left(x_{0:t}^{(i)}\right)}\left(dx_{0:t}^{(i)}\right)$
One obtains straightforwardly the following estimate of $E\left(g_t\left(x_{0:t}\right)\right)$, $E\left(g_t\left(x_{0:t}\right)\right) \approx \frac{1}{N}\sum_{i=1}^{N} g_t\left(x_{0:t}\right)$

There are infinitely many possible choices for $q\left(\mathbf{x}_{0:t} \mid y_{1:t}\right)$, the only condition being that its support must include that of $p\left(\mathbf{x}_{0:t} \mid y_{1:t}\right)$. The simplest choice is to just sample from the prior, $p\left(\mathbf{x}_t \mid \mathbf{x}_{t-1}\right)$, in which case the importance weights is equal to the likelihood, $p\left(y_t \mid y_{1:t-1}, x_{0:t}\right)$. This is the most widely used distribution and it leads to the most general bootstrap PF[9]. This proposal is simple to compute but also can be inefficient since it ignores the most recent evidence, y_t .The selection (resampling) step in Table I is used to eliminate the particles having low importance weights and to multiply particles having high importance weights[9,12].

IV. SRUKF BASED PF

Compared with the original formulation of the UKF, the SRUKF has added benefit of numerical stability[7]. A PF using SRUKF as its proposal distributions has been proposed in [16]. The SRUKF based PF is a candidate algorithm for performance comparison with the proposed SVDSKF based PF.

V. SVDSKF BASED PF

The SVDSKF is close in spirit to UKF[8]. Similarly to the implementations of PFs based on EKF or UKF[12], we propose the SVDSKF based PF as follows in table II.

TABLE II . SVDSKF-BASED PF

1. Initialization: $t=0$
● For $i=1,\cdots,N$, draw the states (particles) $x_0^{(i)}$ from the prior $p\left(x_0\right)$ and set,
$\bar{x}_0^{(i)} = E\left[x_0^{(i)}\right]$
$\mathbf{P}_0^{(i)} = E\left[\left(x_0^{(i)} - \bar{x}_0^{(i)}\right)\left(x_0^{(i)} - \bar{x}_0^{(i)}\right)^T\right]$
$\bar{x}_0^{(i)a} = E\left[x^{(i)a}\right] = \left[\left(\bar{x}_0^{(i)}\right)^T \; 0 \; 0\right]^T$
$\mathbf{P}_0^{(i)a} = E\left[\left(x_0^{(i)a} - \bar{x}_0^{(i)a}\right)\left(x_0^{(i)a} - \bar{x}_0^{(i)a}\right)^T\right] = \begin{bmatrix} \mathbf{P}_0^{(i)} & 0 & 0 \\ 0 & \mathbf{Q} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix}$
where $\mathbf{Q}$ =process noise cov, $\mathbf{R}$ =measurement noise cov
2. For $t=1,2,\cdots$
(a) Importance sampling step
• For $i=1,\cdots,N$:
- Update the particles with the SVD-based KF:
* Compute the SVD and eigen-point covariance matrix
$\mathbf{P}_{t-1}^{(i)a} = \mathbf{USV}^T$
$\chi_{0,t-1}^{(i)a} = \bar{x}_{t-1}^{(i)}$
$\chi_{j,t-1}^{(i)a} = \bar{x}_{t-1}^{(i)} + \rho \mathbf{U}_j \sqrt{s_j}$, $j=1,\cdots,n_a$
$\chi_{j,t-1}^{(i)a} = \bar{x}_{t-1}^{(i)} - \rho \mathbf{U}_j \sqrt{s_j}$, $j=n_a+1,\cdots,2n_a$
where s_j is the jth diagonal element of $\mathbf{S}$.
* Propagate particles into future (time update)
$\chi_{t\|t-1}^{(i)x} = f\left(\chi_{t-1}^{(i)x}, \chi_{t-1}^{(i)v}\right)$ $\bar{x}_{t\|t-1}^{(i)} = \sum_{j=0}^{2n_a} w_j^{(m)} \chi_{j,t\|t-1}^{(i)x}$
$\mathbf{P}_{t\|t-1}^{(i)} = \sum_{j=0}^{2n_a} w_j^{(c)} \left[\chi_{j,t\|t-1}^{(i)x} - \bar{x}_{t\|t-1}^{(i)}\right]\left[\chi_{j,t\|t-1}^{(i)x} - \bar{x}_{t\|t-1}^{(i)}\right]^T$
$y_{t\|t-1}^{(i)} = h\left(\chi_{t\|t-1}^{(i)x}, \chi_{t-1}^{(i)n}\right)$ $\bar{y}_{t\|t-1}^{(i)} = \sum_{j=0}^{2n_a} w_j^{(m)} y_{j,t\|t-1}^{(i)}$
TO BE CONTINUED

* Incorporate new observation (measurement update)

$$\mathbf{P}_{x_t y_t} = \sum_{j=0}^{2n_a} w_j^{(c)} \left[\chi_{j,t|t-1}^{(i)x} - \overline{x}_{t|t-1}^{(i)} \right] \left[y_{j,t|t-1}^{(i)} - \overline{y}_{t|t-1}^{(i)} \right]^T$$

$$K_t = \mathbf{P}_{x_t y_t} \mathbf{P}_{\bar{y}_t \bar{y}_t}^{-1} \quad \overline{x}_t^{(i)} = \overline{x}_{t|t-1}^{(i)} + K_t \left(y_t - \overline{y}_{t|t-1}^{(i)} \right)$$

$$\hat{\mathbf{P}}_t^{(i)} = \mathbf{P}_{t|t-1}^{(i)} - K_t \mathbf{P}_{\bar{y}_t \bar{y}_t} K_t^T$$

-Sample $\hat{x}_t^{(i)} \sim q\left(x_t^{(i)} \mid x_{0:t-1}^{(i)}, y_{1:t} \right) = N\left(\overline{x}_t^{(i)}, \hat{\mathbf{P}}_t^{(i)} \right)$

-Set $\hat{x}_{0:t}^{(i)} \triangleq \left(x_{0:t-1}^{(i)}, \hat{x}_t^{(i)} \right)$ and $\hat{\mathbf{P}}_{0:t}^{(i)} \triangleq \left(\mathbf{P}_{0:t-1}^{(i)}, \hat{\mathbf{P}}_t^{(i)} \right)$

- For $i = 1, \cdots, N$, evaluate the importance weights up to a normalizing constant.

$$W_t^{(i)} \propto \frac{p\left(y_t \mid \hat{x}_t^{(i)} \right) p\left(\hat{x}_t^{(i)} \mid x_{t-1}^{(i)} \right)}{q\left(\hat{x}_t^{(i)} \mid x_{0:t-1}^{(i)}, y_{1:t} \right)}$$

- For $i = 1, \cdots, N$, normalize the importance weights.

(b) Selection step

Multiply/Suppress particles $\left(\hat{x}_{0:t}^{(i)}, \hat{\mathbf{P}}_{0:t}^{(i)} \right)$ with high/low importance weights $\tilde{W}_t^{(i)}$,respectively, to obtain N random particles $\left(\tilde{x}_{0:t}^{(i)}, \tilde{\mathbf{P}}_{0:t}^{(i)} \right)$.

(c) Output:

The output is generated in the same manner as for the generic particle filter.

Parameter and weights of the SVD-based KF module:

$$w_0^{(m)} = \frac{\kappa}{n_a + \kappa}, w_j^{(m)} = \frac{1}{2(n_a + \kappa)} (j \neq 0)$$

$$w_0^{(c)} = \frac{\kappa \cdot n_a}{1.96 \times (n_a + \kappa)}, w_j^{(c)} = \frac{n_a}{1.96 \times 2 \times (n_a + \kappa)} (j \neq 0)$$

where κ is a small tuning parameter. ρ is a scaling parameter (a good choice is $1 \leq \rho \leq 1.414$) for controlling the extent of covariance.

VI. SIMULATION

For this experiment, a time-series is generated by the following process model $x_{t+1} = 1 + \sin(w\pi t) + \phi_1 x_t + v_t$, where v_t is a $Gamma(3,2)$ random variable modeling the process noise, and $w = 4e-2$ and $\phi_1 = 0.5$ are scalar parameters. A non-stationary observation model,

$$y_t = \begin{cases} \phi_2 x_t^2 + n_t & t \leq 30 \\ \phi_3 x_t - 2 + n_t & t > 30 \end{cases}$$

is used, with $\phi_2 = 0.2$ and $\phi_3 = 0.5$ .The observation noise, n_t ,is draw from a Gaussian distribution $N\left(0, 1e-5\right)$ .Given only the noise observations, y_t , different filters are used to estimate the underlying clean state sequence x_t for $t = 1 \ldots 60$. The experiment is repeated 100 times with random re-initialization for each run. Each of the PFs used in the experiment, that is the generic bootstrap PF, EKF based PF, SRUKF based PF and the SVDSKF based PF, uses 200 particles and residual re-sampling strategy[12]. The SRUKF module parameters are set to $\alpha = 1, \beta = 0$ and $\kappa = 2$ which are optimal for the scalar case. We run each of the PFs 100 times respectively to get the means and variances of the mean-square-error (MSE) of the state estimates shown in Table III. An example of the estimation results is shown in Figure 1, and a comparison of computing burdens of each PF is illustrated in Figure 2.

TABLE III . The mean and variance of the MSE calculated over 100 independent runs

Algorithm	MSE	
	mean	*variance*
Bootstrap PF	0.357	0.055
EKF based PF	0.315	0.018
SRUKF based PF	0.320	0.069
SVDSKF based PF (proposed)	0.167	0.015

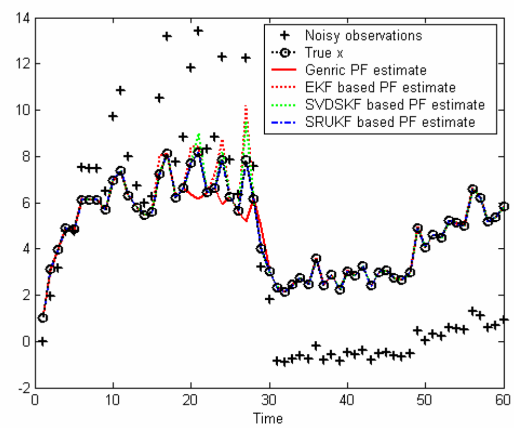

Figure 1. Filter estimates (posterior means) vs. True state

VII. CONCLUSIONS

This paper proposes a new PF algorithm. It incorporates a robust numerical method, i.e. SVD, into the framework of PF. Meanwhile it incorporates the latest observations into a prior updating routine of PF. Both of the two characters guarantee this method's superiority to the generic bootstrap PF algorithm. Furthermore, simulation results show that the proposed algorithm also has better estimation performance than other existing similar PF algorithms that are based on local linearization techniques such as the EKF and SRUKF. However, the proposed method needs some more computing costs than the others mentioned above.

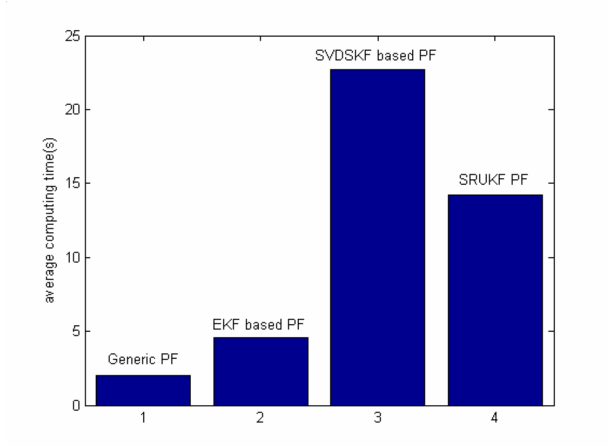

Figure 2. Comparison of computing burdens indexed by average computing time in a simulation including 100 independent runs of each PF algorithm respectively

REFERENCES

[1] Julier S J, and Uhlmann J K, "Unscented filtering and nonlinear estimation," Proc of the IEEE, vol.92, No.3, pp.401-422, March 2004.

[2] Youmin Zhang, Guanzhong Dai and Hongcai Zhang, "A SVD-based extended Kalman filter and applications to aircraft flight state and parameter estimation," Proc of the American Control Conference, Baltimore, Maryland, pp. 1809-1813, June 1994.

[3] Julier S J, Uhlmann J K, and Durrant-Whyte H F, "A new approach for the nonlinear transformation of means and covariances in filters and estimators," IEEE Trans on Automatic Control, vol.45, No.3, pp. 477-482, 2000.

[4] Julier S J, and Uhlmann J K, "Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations," Proc of the American Control Conference, Jefferson City, pp.887-892, 2002.

[5] Julier S J, "The scaled unscented transformation," Proc of American Control Conf, Jefferson City, pp.4555-4559, 2002.

[6] Julier S J, and Uhlmann J K, "A consistent, debiased method for converting between polar and Cartesian coordinate systems," The Proc of AeroSense: The 11th Int Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, pp.110-121, 1997.

[7] Van der Merwe.R., and Wan, E.A., "The square-root unscented kalman filter for state and parameter-estimation," Proc of the International Conference on Acoustic, Speech, and Signal Processing(ICASSP), Salt Lake City, Utah, vol. 6, pp. 3461-3464, May 2001.

[8] Z.Chen, S. Haykin, "Bayesian Filtering: from Kalman filters to particle filters, and beyond," Technical report, Adaptive Systems Lab, McMaster University, Hamilton, ON, Canada, 2003.

[9] Arulampalam S, Maskell S, and Gordon N, etc, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking," IEEE Trans on Signal Processing, vol.50, No.2, pp.174-188, 2002.

[10] Gordon N J, Salmond D J, and Smith A F M, "Novel approach to nonlinear/non-Gaussian bayesian state estimation," IEE Proceedings-F, vol.140, No.2, pp.107-113, 1993.

[11] Doucet A, De Freitas A, Gordon N, eds., Sequential Monte Carlo methods in practice, New York, Springer-Verlag, 2001.

[12] Van der Merwe, R.,Doucet, and A., de Freitas, "The unscented particle filter," Technical Report CUED/F-INFENG/TR 380, Cambridge University Engineering Department, 2000.

[13] Charles G.Cullen, An Introduction to Numerical Linear Algebra. PWS, Boston, MA,1994.

[14] G.Golub and C.Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1983.

[15] Lloyd N.Trefethen and David Bau, III Numerical Linear Algebra, SIAM, 1997.

[16] Jiaxiang Yu, Deyun Xiao and Xiuting Yang, "Square root unscented particle filter with application to angle-only tracking," Proc of the 6th World Congress on Intelligent Control and Automation, Dalian,China, pp. 1548-1552, June 2006.