Tian-cheng Li, Gabriel Villarrubia, Shu-dong Sun, Juan M. Corchado, Javier Bajo, 2015. Resampling methods for particle filtering: identical distribution, a new method and comparable study. *Frontiers of Information Technology & Electronic Engineering*, **16**(11):969-984. [doi:10.1631/FITEE.1500199]

Resampling methods for particle filtering: identical distribution, a new method, and comparable study

Key words: Particle filter, Resampling, Kullback-Leibler divergence, Kolmogorov-Smirnov statistic

Contact: Tian-cheng Li
E-mail: t.c.li@usal.es; tiancheng.li1985@gmail.com
ORCID: http://orcid.org/0000-0002-0499-5135

Motivation (I)

Resampling is a critical procedure that is of both theoretical and practical significance for efficient implementation of the particle filter. To gain an insight of the resampling process and the filter, this paper contributes in three further respects as a sequel to the tutorial (Li *et al.*, 2015).

1. Identical distribution (ID) is established as a general principle for the resampling design, which requires the distribution of particles before and after resampling to be statistically identical. Three consistent metrics including the (symmetrical) Kullback-Leibler divergence, Kolmogorov-Smirnov statistic, and the sampling variance are introduced for assessment of the ID attribute of resampling, and a corresponding, qualitative ID analysis of representative resampling methods is given.

Motivation (II)

2. A novel resampling scheme that obtains the optimal ID attribute in the sense of minimum sampling variance is proposed.

3. More than a dozen typical resampling methods are compared via simulations in terms of sample size variation, sampling variance, computing speed, and estimation accuracy.

These form a more comprehensive understanding of the algorithm, providing solid guidelines for either selection of existing resampling methods or new implementations.

Reference:
Li, T., Bolic, M., Djurić, P.M., 2015. Resampling methods for particle filtering. *IEEE Signal Process. Mag.*, **32**(3):70-86. [doi:10.1109/MSP.2014.2330626]

Main idea/Method (I)

Three identical distribution metrics

1. The KLD of q from p, denoted as $D_{\mathrm{KL}}(p\|q)$, is a measure of the information lost when q is used to approximate p. For the discrete probability distributions p and q, $D_{\mathrm{KL}}(p\|q)$ is given by

$$D_{\mathrm{KL}}(p\,\|\,q)=\sum_i p(i)\ln\frac{p(i)}{q(i)}.$$

2. The Kolmogorov-Smirnov (K-S) statistic, also referred to as F-discrepancy, is the maximum vertical distance between two empirical distribution functions (EDF):

$$D_{\text{K-S}}=\sup_{-\infty<x<\infty}|F_{1,M}(x)-F_{2,N}(x)|.$$

3. Define the sampling variance (SV) as the mean of the quadratic discrepancies between the numbers of times that the particles are resampled and the expectations:

$$\mathrm{SV}=\frac{1}{M}\sum_{m=1}^{M}(N_t^{(m)}-Nw_t^{(m)})^2,$$

where N is the expected number of particles to resample, $N_t^{(m)}$ is the real number of times that particle m is resampled, $w_t^{(m)}, m=1,\ldots M$ are the weights of original particles.

Main idea/Method (II)

MSV (minimum-sampling-variance) resampling

Step 1: Each particle is first resampled $\text{Floor}(Nw_t^{(m)})$ times, leaving a weight residual $\hat{w}_t^{(m)} = w_t^{(m)} - \text{Floor}(Nw_t^{(m)}) / N$; this step will yield, in total, L particles, where $L = \sum_{m=1}^{M} \text{Floor}(Nw_t^{(m)})$.

Step 2: The particle with relatively large weight residual, top $N-L$, will be further sampled one more time each.

Algorithm 1 MSV resampling

$[\{\tilde{x}_t^{(n)}\}_{n=1}^{N}] = \text{ResampleMSV}[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^{M}, N]$;

$n=0$; $L=0$;

For $m=1:M$

$\quad N_t^{(m)} = \text{Floor}(Nw_t^{(m)})$;

$\quad \hat{w}_t^{(m)} = w_t^{(m)} - N_t^{(m)} / N$;

$\quad L = L + N_t^{(m)}$;

End

For $m=1:M$

$\quad$ **If** $\hat{w}_t^{(m)} \in \text{TopRank}_{N-L}[\{\hat{w}_t^{(m)}\}_{m=1}^{M}]$

$\qquad N_t^{(m)} = N_t^{(m)} + 1$;

$\quad$ **End**

$\quad$ **For** $h = 1 : N_t^{(m)}$

$\qquad n=n+1$;

$\qquad \tilde{x}_t^{(n)} = x_t^{(m)}$;

$\quad$ **End**

End

Main idea/Method (III)

MSV (minimum-sampling-variance) resampling

Theory 1. The MSV resampling approach satisfies that $\forall\, 1 \le m \le M$,

$$\left|N_t^{(m)} - N w_t^{(m)}\right| < 1.$$

Theory 2. The MSV resampling approach minimizes the SV.

Theory 3. The MSV resampling approach is asymptotically unbiased.

Theories 2 and 3 show that the proposed MSV resampling method minimizes the SV in terms of asymptotical unbiasedness and optimal-weight condition.

Major results (I)

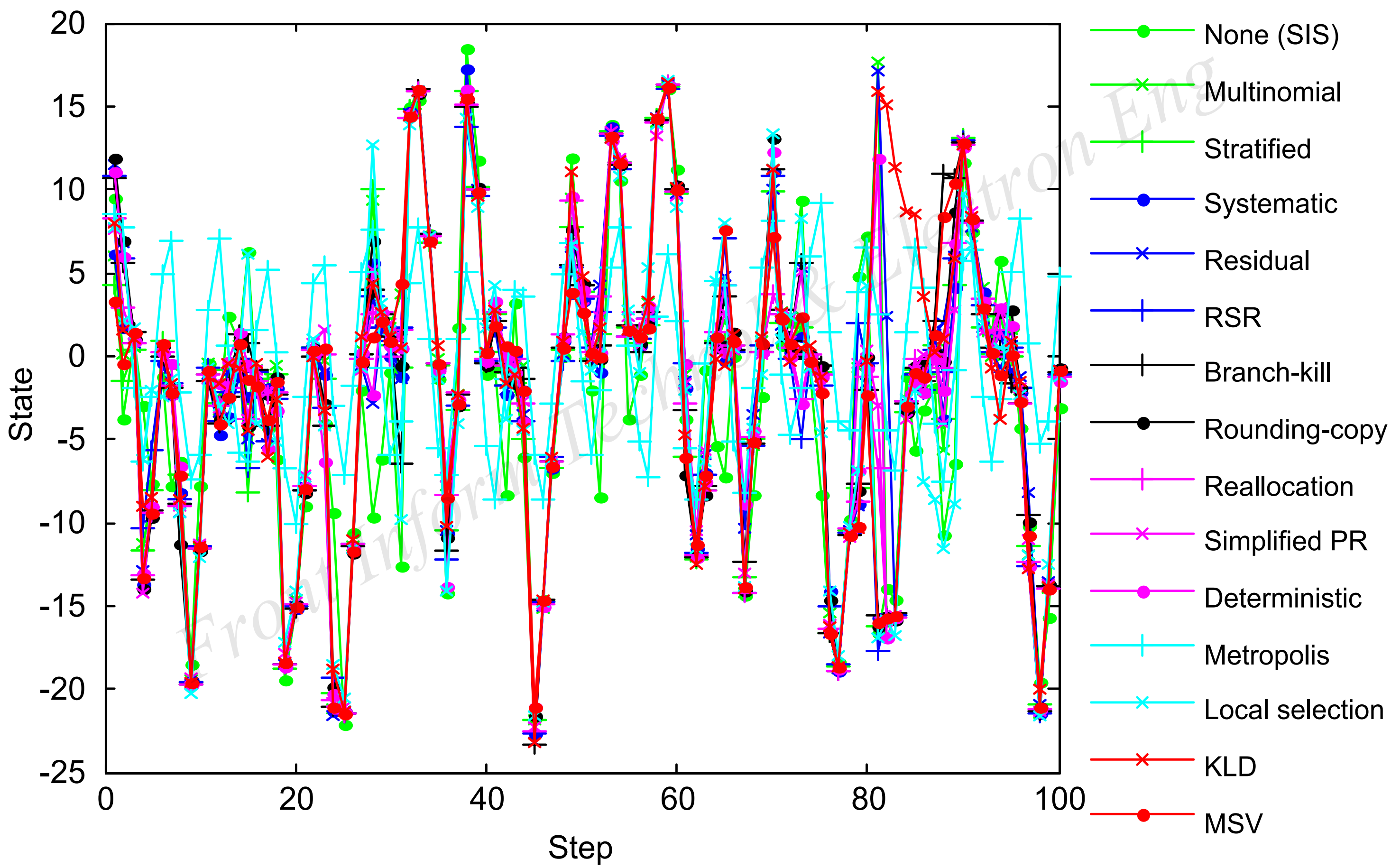

Fig. 1 True state and estimates against time when the starting number of particles is 100

Major results (II)

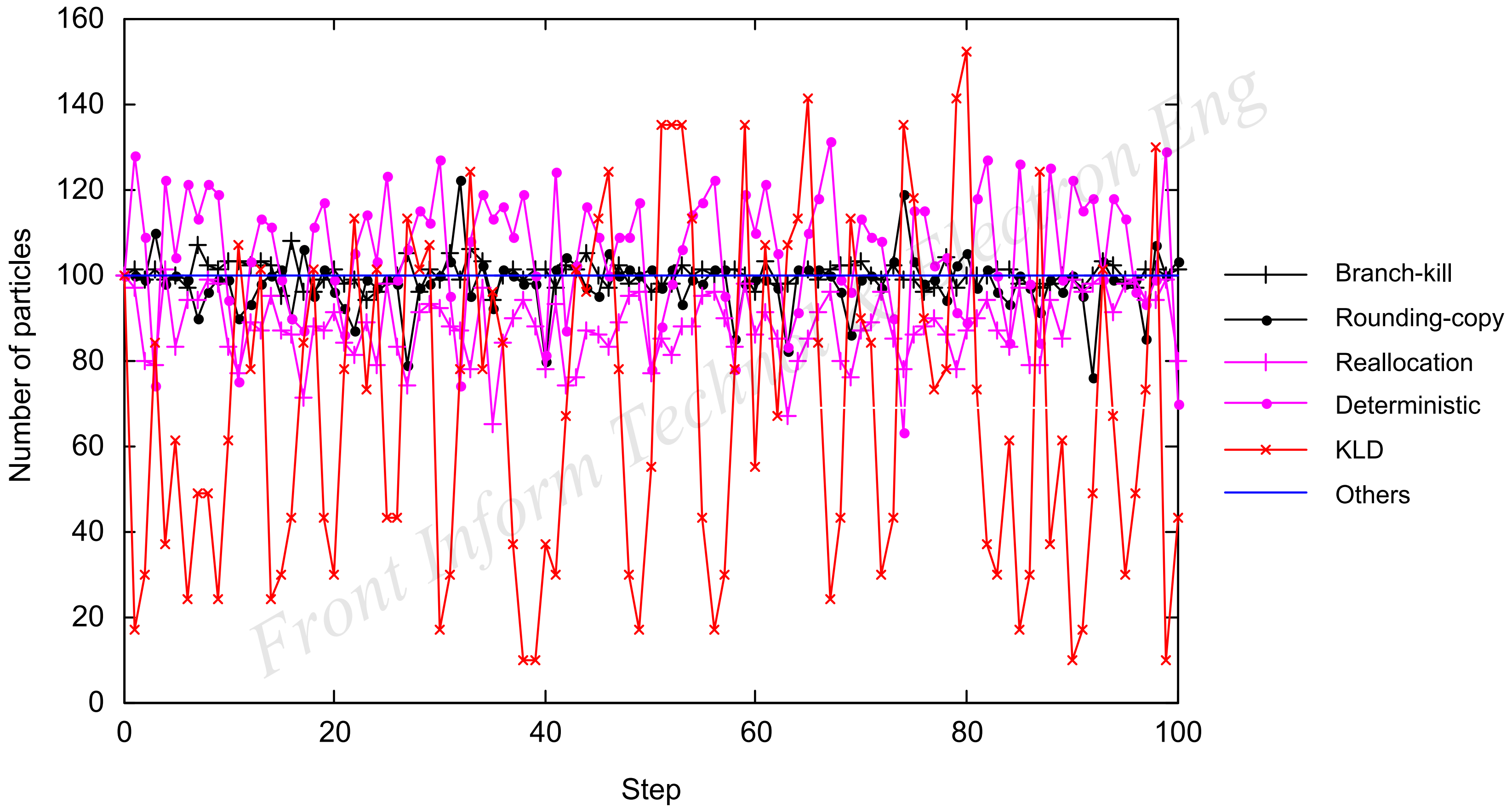

Fig. 2 Fluctuation of the number of particles against time

Major results (III)

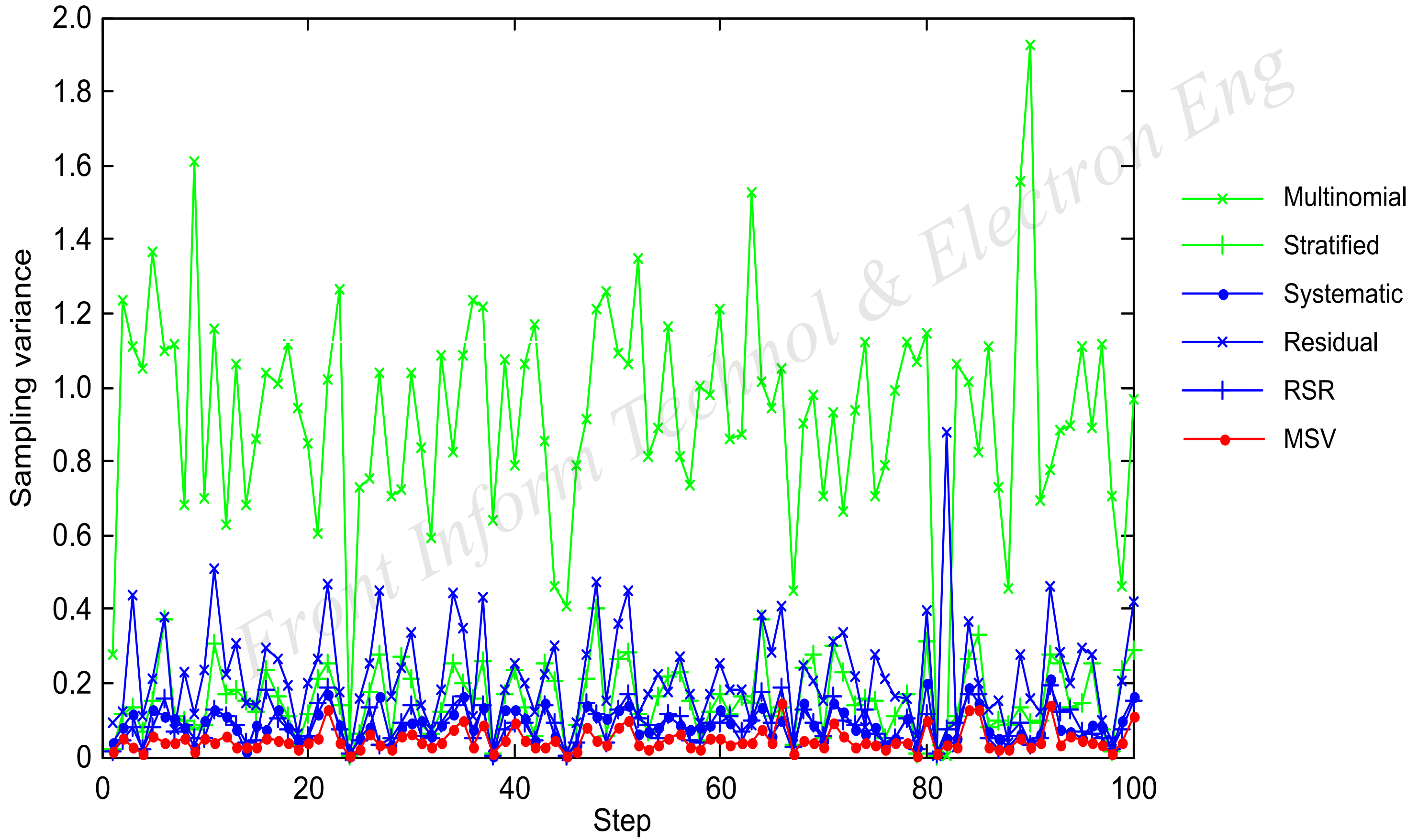

Fig. 3 Sampling variance of different resampling methods that satisfy the optimal weighting condition

Major results (IV)

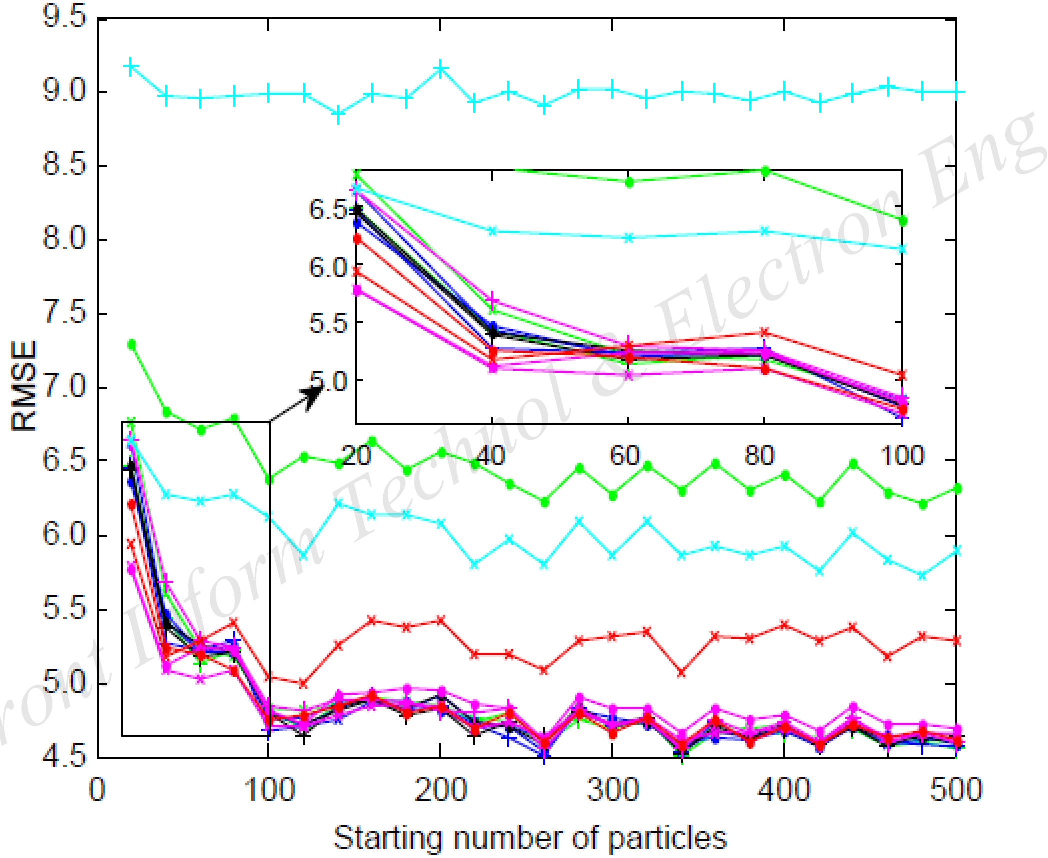

Fig. 4 Average RMSEs of PFs against different starting numbers of particles (legend is the same as that of Fig. 1)

Major results (V)

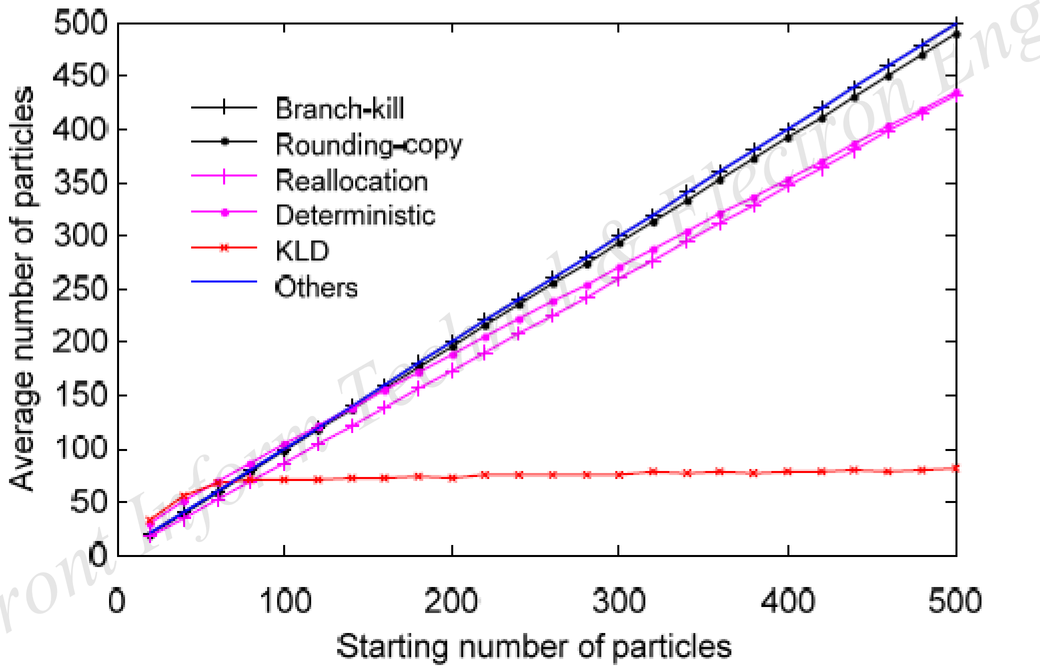

Fig. 5 Average numbers of particles of 100 steps of different particle filters (after resampling) against different starting numbers of particles

Major results (VI)

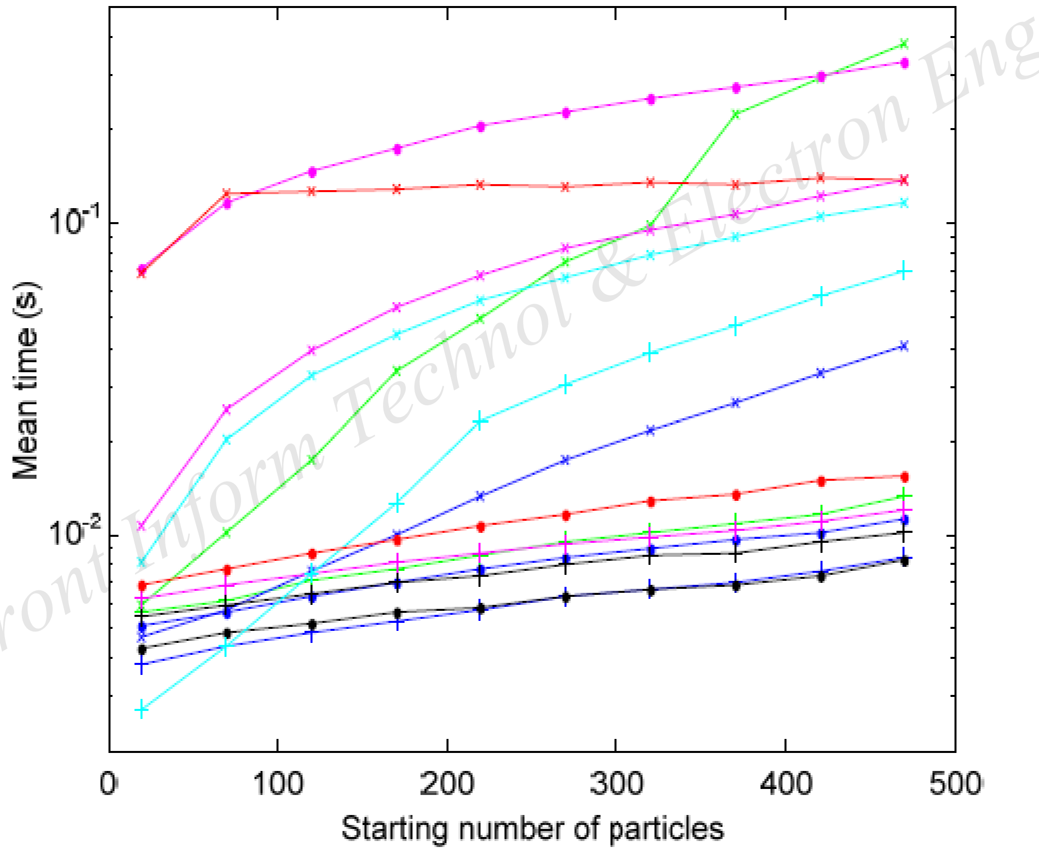

Fig. 6 Average processing time of 100 runs of different resampling methods against different starting numbers of particles (legend is the same as that of Fig. 1).

Conclusions

- This paper has contributed in three more aspects.
  - ID has been established as a fundamental principle for resampling, which can be measured by the KL divergence, K-S statistic, and the sampling variance. This affords a useful perspective to compare and assess existing resampling methods or to design new methods.
  - Following the ID principle, a simple resampling method, called MSV resampling, is proposed that obtains the optimal ID attribute in terms of SV.
  - A comprehensive comparable study of more than a dozen representative resampling methods, including the proposed MSV resampling based on a classical state space model, is given. The results show that, most unbiased resampling methods do not exhibit much difference in terms of estimation accuracy (despite significantly biased resampling methods performing very badly), but they show significant differences in terms of SV and computing time and may provide special benefits in specific problems