TRACKING A FREQUENCY HOPPED SIGNAL USING PARTICLE FILTERING

N. D. Sidiropoulos[*] *A. Swami* *A. Valyrakis*

Dept. ECE, Tech. Univ. of Crete
73100 Chania - Crete, Greece
nikos@telecom.tuc.gr

Army Research Laboratory
Adelphi, MD, 20783, U.S.A.
a.swami@ieee.org

Dept. ECE, Tech. Univ. of Crete
73100 Chania - Crete, Greece
alevali@telecom.tuc.gr

ABSTRACT

The problem of tracking the frequency and complex amplitude of a frequency-hopped complex sinusoid is considered, using a novel stochastic state-space formulation and particle filtering tools. The problem is of considerable interest for interference mitigation in frequency-hopped wireless networks, and in military communications. The proposed particle filtering approach has a number of desirable features. It affords high-resolution estimates of carrier frequency and hop timing, manageable complexity (linear in the number of processed samples), and flexibility in tracking signals with irregular hopping patterns due to intentional timing jitter. The proposed state-space model is not only parsimonious, but fortuitous as well: it turns out that the associated optimal importance function can be computed in closed form, and thus samples from it can be drawn using rejection techniques. Both prior and optimal importance sampling versions are developed and illustrated in pertinent simulations.

Keywords: Frequency hopping, spectral analysis, estimation of time-varying line spectra, sequential importance sampling, particle filtering

1. INTRODUCTION

Tracking the frequency of a time-varying complex sinusoid is an important problem which arises in numerous applications. In speech processing, for example, one is often interested in tracking formant frequencies. In wireless communications, it arises in the context of frequency hopping, when the receiver has no prior knowledge of the hopping pattern, or is simply out of sync with the transmitter's hopping pattern generator [2, 8, 6, 7].

Both non-parametric time-frequency analysis, and parametric techniques have been developed for the more general problem of tracking a time-varying sinusoid, and can be applied to the problem of tracking a frequency-hopped sinusoid as well. However, existing methods have limitations, especially when used to track a frequency-hopped signal. Non-parametric methods, like the spectrogram, or coarse channelization [2] suffer from limited frequency- and temporal-resolution due to leakage. It is possible to employ time-frequency distributions that are better-adapted to frequency hopping [3], but the results are still not very satisfactory. Parametric methods for frequency hopping explicitly model the frequency as piecewise-constant, assume a "budget" on the number of hops within a given observation interval, and employ dynamic programming to track the sought frequency and complex amplitude parameters [6, 7]. Other than an upper bound on the number of hops, the methods in [6, 7] do not assume anything else about the frequencies or complex amplitudes, which are treated as deterministic unknowns.

A different viewpoint is adopted in this paper. A stochastic non-linear, non-Gaussian state-space formulation is proposed, which captures frequency hopping dynamics in a probabilistic sense. The proposed formulation is naturally well-suited for the application of particle filtering for state estimation. Compared to the prior state-of-art in [6, 7], the new approach has a number of desirable features:

• Computational complexity: The complexity of particle filtering is $O(NT)$, where N is the number of particles and T is the number of temporal samples. The complexity of dynamic programming, on the other hand, is roughly $O(T^4)$. This means that only short segments can be processed by dynamic programming, and then one has to rely on hop periodicity to segment the rest of the data. This has two disadvantages: first, the more samples are processed the better from an estimation performance perspective; second, hop timing is often intentionally randomized as a countermeasure.

• Flexibility: The state-space model in the particle filtering formulation can be easily tailored to match a given scenario (e.g., spread bandwidth and modulation).

The proposed state-space model is simple and fortuitous: the associated optimal importance function can be computed in closed form, and thus samples from it can be drawn using rejection techniques. Both prior and optimal importance sampling versions are developed and compared in pertinent simulations.

2. DATA MODEL AND PROBLEM STATEMENT

We propose the following non-linear non-Gaussian stochastic state-space model of a frequency-hopped complex sinusoid. Let $\mathbf{x}_k := [\omega_k, A_k]^T$ denote the state at time k, where $\omega_k \in [-\pi, \pi)$ and $A_k \in \mathbb{C}$ denote instantaneous frequency and complex amplitude. Let $\mathbf{u}_k := [b_k, \tilde{\omega}_k, \tilde{A}_k]^T$ denote an auxiliary sequence of independent and identically distributed (i.i.d.) vectors with independent components and the following marginal statistics: b_k is a binary random variable with $Pr(b_k = 1) = h$; $\tilde{\omega}_k$ is uniformly distributed over $[-\pi, \pi)$, denoted $\mathcal{U}([-\pi, \pi))$; and $\tilde{A}_k$ is $\mathcal{CN}(0, \sigma_A^2)$, i.e., complex circular Gaussian of variance σ_A^2. Then

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) = \begin{cases} \mathbf{x}_{k-1} & , \mathbf{u}_k(1) = 0 \\ [\mathbf{u}_k(2),\ \mathbf{u}_k(3)]^T & , \mathbf{u}_k(1) = 1 \end{cases}$$

[*]Corresponding author. Supported in part by the Army Research Laboratory (ARL) through participation in the ARL Collaborative Technology Alliance (ARL-CTA) for Communications and Networks under Cooperative Agreement DADD19-01-2-0011, and in part by ARO under ERO Contract N62558-03-C-0012.

$$= \begin{cases} \mathbf{x}_{k-1} & , w.p.\ 1-h \\ \left[\mathcal{U}\left([-\pi,\pi)\right),\ \mathcal{CN}(0,\sigma_A^2)\right]^T & , w.p.\ h \end{cases},$$

$$y_k = \mathbf{x}_k(2) e^{j\mathbf{x}_k(1)k} + v_k,$$

where v_k denotes i.i.d. $\mathcal{CN}(0,\sigma_n^2)$ measurement noise, and $u_k(1)$ the hop variable.

The above state-space formulation models frequency hopping in a probabilistic fashion. Hops are random, i.i.d., with hop probability h per sample interval. This is different from traditional models of frequency hopping, which assume that the frequency hops periodically, and is motivated by the following considerations:

- In military communications, intentional jitter is often introduced in the hop timing in order to reduce the probability of detection by unintended receivers and improve immunity to jamming. Timing jitter yields a pseudo-random quasi-periodic, or even seemingly aperiodic hop timing sequence.
- The above probabilistic model captures information about the average hop rate in a "soft" ensemble sense: the expected number of hops over a long observation interval T is hT. While less accurate if the exact hop period is known, probabilistic modeling is more robust with respect to hop period inaccuracies. Finally,
- The proposed probabilistic model is ideally suited for on-line sequential estimation via particle filtering.

It is worth elaborating on some of the implicit assumptions of the proposed state-space model.

1. When the (discrete-time, baseband-equivalent) frequency hops, it hops anywhere within $[-\pi,\pi)$ with a uniform density. This is well-suited for carrier hopping, which is usually discontinuous. Modulation-induced variations can (and should) be neglected when the objective is to estimate carrier frequency, but could also be explicitly modeled using, e.g., a smooth auto-regressive frequency variation model in-between hops, in lieu of the simplified constant model postulated above. This extension is relatively simple.
2. When the frequency hops, the complex amplitude also changes according to an i.i.d. complex Gaussian distribution. This is also well-motivated for carrier hopping, for every time the carrier frequency hops beyond the coherence bandwidth of the channel, a new channel realization is encountered.

The problem, then, can be stated as follows: Given a sequence of observations $\{y_k\}_{k=1}^T$, estimate the sequence of system states $\{\mathbf{x}_k\}_{k=1}^T$ - that is, the unknown carrier frequencies and complex amplitudes.

3. PARTICLE FILTERING SOLUTIONS

Particle filtering has emerged as an important sequential state estimation method for stochastic non-linear and/or non-Gaussian state-space models, for which it provides a powerful alternative to the commonly used extended Kalman filter. See [1, 5] for recent tutorial overviews. In particle filtering, continuous distributions are approximated by discrete random measures, comprising "particles" and associated weights. That is, a certain continuous distribution of interest, say $p(\mathbf{x})$, is approximated as

$$p(\mathbf{x}) \approx \sum_{n=1}^{N} w_n \delta(\mathbf{x} - \mathbf{x}_n),$$

where $\delta(\cdot)$ denotes the Dirac delta functional. A useful simplification stemming from this approximation is that the computation of pertinent expectations and conditional probabilities reduces to summation, as opposed to integration. While this can also be accomplished via direct discretization over a fixed grid, the use of a random measure affords flexibility in adapting the particle locations to better fit the distribution of interest.

3.1. Basics of particle filtering

If we aim for an on-line filtering algorithm, in which the state at time k should be estimated from measurements up to and including time k, the key distribution of interest is the posterior density $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^k\right)$. Given this density, one can estimate the state at time k, e.g., via the associated (posterior) mean, or mode. The basic idea of particle filtering, then, is to begin with a random measure approximation of the initial state distribution, and, as measurements become available, derive updated random measure approximations of $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^k\right)$, $k \in \{1, 2, \cdots\}$. That is, we seek random measure approximations

$$\hat{p}\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^k\right) = \sum_{n=1}^{N} w_{n,k} \delta(\mathbf{x}_k - \mathbf{x}_{n,k})$$

In particle filtering, the updates - the derivation of $\hat{p}\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^k\right)$ from $\hat{p}\left(\mathbf{x}_{k-1} \mid \{y_l\}_{l=1}^{k-1}\right)$ - are based on the Bayes rule [1, 5].

A random measure approximation comprises two components: the particles (locations) and the associated weights. If we could sample from the sought posterior $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^k\right)$, then all particle weights would have been equal. Unfortunately, such direct sampling is not possible in most cases, and thus we resort to sampling from a so-called *importance function* that "resembles" the desired posterior, and from which samples can be drawn with relative ease. The mismatch between the sought density and the importance function is compensated in the calculation of weights, chosen proportional to their ratio evaluated at each particle [1, 5]. The choice of importance function is a very important step in the design of a particle filtering algorithm. Two common choices are discussed next.

3.2. Prior importance function

Perhaps the most intuitive choice of importance function is the *prior importance function* $p(\mathbf{x}_k \mid \mathbf{x}_{n,k-1})$; i.e., the n-th particle is updated by propagating it through the state-evolution part of the system: $\mathbf{x}_{n,k} = f(\mathbf{x}_{n,k-1}, \mathbf{u}_n)$. This is an often-made choice, for simplicity considerations. The drawback is that particles evolve without regard to the latest measurement, which only comes into play in the ensuing weight update. When using the prior importance function, the said weight update at time instant k is given by $w_{n,k} = w_{n,k-1} p(y_k \mid \mathbf{x}_{n,k})$, followed by normalization to enforce $\sum_{n=1}^{N} w_{n,k} = 1$.

Regardless of the particular importance function employed, a common problem in particle filtering is *degeneracy*: the weights of all but a few particles tend to become negligible after a few iterations [1, 5]. Degeneracy can be detected via degeneracy measures, and mitigated via *resampling* techniques [1, 5]. Resampling the discrete measure replicates particles with large weights and removes those with negligible weights. All particle weights become

equal after resampling. There exist several computationally efficient ($O(N)$) resampling schemes that can be used to avoid the quadratic cost of brute-force resampling [1, 5].

3.3. Optimal importance function

From the viewpoint of minimizing the variance of the weights, the optimal importance function is given by [1, 5]

$$p(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) = \frac{p(y_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{n,k-1})}{\int_{\mathbf{x}} p(y_k|\mathbf{x})p(\mathbf{x}|\mathbf{x}_{n,k-1})d\mathbf{x}}.$$

Notice that, in contrast to the prior importance function, the above takes into account the newly available measurement in the particle update itself. While both the prior importance function and the optimal one yield consistent algorithms[1], the optimal one usually works well with much smaller N, and is therefore preferable from a performance point of view. There are, however, two difficulties associated with the use of the optimal importance function. First, it requires multidimensional integration to compute the normalization factor, which is often intractable. Second, sampling from the optimal importance function is more complicated than sampling from the prior. The smaller number of particles needed to attain satisfactory performance with the optimal importance function usually more than offsets the cost of drawing samples from it; the integration problem remains the bottleneck in most cases [1]. Thankfully, for our particular model, it turns out that it is possible to carry out this integration analytically. This is explained next.

Denote $\mathbf{x}_k := [\omega_k, A_k]^T$, where $\omega_k \in [-\pi, \pi)$, and $A_k \in \mathbb{C}$; likewise $\mathbf{x}_{n,k-1} := [\omega_{n,k-1}, A_{n,k-1}]^T$, and a dummy variable $\mathbf{x} := [\omega, A]^T$. Let $D(y_k, \mathbf{x}_{n,k-1}) := \int_{\mathbf{x}} p(y_k|\mathbf{x})p(\mathbf{x}|\mathbf{x}_{n,k-1})d\mathbf{x}$. Then

$$D(y_k, \mathbf{x}_{n,k-1}) = \int_{\omega\in[-\pi,\pi)} \int_{A\in\mathbb{C}} \frac{1}{2\pi\sigma_n^2} e^{-\frac{|y_k - Ae^{j\omega k}|^2}{2\sigma_n^2}} \times$$

$$\left[(1-h)\delta(\omega - \omega_{n,k-1})\delta(A - A_{n,k-1}) + \frac{h}{2\pi}\frac{1}{2\pi\sigma_A^2} e^{-\frac{|A|^2}{2\sigma_A^2}}\right] dAd\omega$$

This integral can be computed by completing the squares, yielding

$$D(y_k, \mathbf{x}_{n,k-1}) = \frac{1}{2\pi}\frac{h}{\sigma_n^2 + \sigma_A^2} e^{-\frac{|y_k|^2}{2(\sigma_n^2+\sigma_A^2)}} +$$

$$\frac{1}{2\pi}\frac{1-h}{\sigma_n^2} e^{-\frac{|y_k - A_{n,k-1}e^{j\omega_{n,k-1}k}|^2}{2\sigma_n^2}}.$$

For the above optimal choice of the importance function, the weight update is given by

$$w_{n,k} \propto w_{n,k-1} p(y_k|\mathbf{x}_{n,k-1}) = w_{n,k-1} D(y_k, \mathbf{x}_{n,k-1}),$$

followed by normalization to 1. What is missing is a way to sample from the optimal importance function. As a first step towards this end, note that $p(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k)$ can be written as a mixture of two pdfs

$$p(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) = (1-\tilde{h})p_0(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) + \tilde{h}p_1(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k),$$

where

$$p_0(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) := \delta(\omega_k - \omega_{n,k-1})\delta(A_k - A_{n,k-1}),$$

$$p_1(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) := \frac{\frac{1}{2\pi}\frac{1}{2\pi\sigma_n^2}\frac{1}{2\pi\sigma_A^2} e^{-\frac{|y_k - A_k e^{j\omega_k k}|^2}{2\sigma_n^2}} e^{-\frac{|A_k|^2}{2\sigma_A^2}}}{\frac{1}{2\pi}\frac{1}{\sigma_n^2+\sigma_A^2} e^{-\frac{|y_k|^2}{2(\sigma_n^2+\sigma_A^2)}}},$$

and

$$\tilde{h} := h\frac{\frac{1}{2\pi}\frac{1}{\sigma_n^2+\sigma_A^2} e^{-\frac{|y_k|^2}{2(\sigma_n^2+\sigma_A^2)}}}{D(y_k, \mathbf{x}_{n,k-1})}.$$

It follows that with probability $1 - \tilde{h}$ we simply copy the previous particle, else we draw a particle from $p_1(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k)$. We will use rejection sampling techniques for this latter step, as explained next.

3.4. Sampling from the optimal importance function: Rejection

The basic idea of rejection-based sampling can be summarized as follows [4, pp. 40-42]. Suppose we wish to draw samples from a density $\phi(\mathbf{x})$, for which there exists a *dominating density* $g(\mathbf{x})$ and a known constant c such that $\phi(\mathbf{x}) \leq cg(\mathbf{x}), \forall \mathbf{x}$. In practice, we choose $g(\mathbf{x})$ to be easy to sample from, and such that c is as small as possible. The rejection method then works as follows. We i) draw a sample $\mathbf{x}$ from $g(\cdot)$ and an independent sample $U \sim \mathcal{U}([0, 1])$; ii) set $\tau := c\frac{g(\mathbf{x})}{\phi(\mathbf{x})}$; iii) test whether $U\tau \leq 1$; if so, we accept the sample $\mathbf{x}$; else we reject it and repeat the process.

It can be shown that the above rejection method generates samples from the desired density $\phi(.)$, and the mean number of iterations until a sample is accepted is c (thus the desire to keep $c \geq 1$ as small as possible). Furthermore, the distribution of the number of trials is geometic with parameter $1 - \frac{1}{c}$, which means that the probabilities of longer trials decay exponentially [4, p. 42].

In our present context, we wish to sample from the density $p_1(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k)$. Define

$$\mu := \frac{|y_k|\sigma_A^2}{\sigma_n^2 + \sigma_A^2}, \quad \sigma^2 := \frac{\sigma_n^2\sigma_A^2}{\sigma_n^2 + \sigma_A^2}.$$

Using the triangle inequality, it can be shown that the following is a suitable dominating density:

$$g(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) = \frac{e^{-\frac{(|A_k| - \mu)^2}{2\sigma^2}}}{(2\pi)^{5/2}Q_0\sigma},$$

for which it holds that $p_1(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) \leq cg(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k)$, with

$$c := \sqrt{2\pi}Q_0/\sigma \geq 1,$$

$$Q_0 := \int_{r=0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}} dr = \frac{1}{2}\text{erfc}(-\frac{|y_k|\sigma_A}{\sigma_n\sqrt{2(\sigma_n^2 + \sigma_A^2)}}).$$

[1] In the sense that the pertinent discrete measure approximations converge to the sought continuous distributions as $N \to \infty$, see [1] and references therein.

Through experimentation, we have found that even better results can be attained using an *outer rejection loop*, which declines candidates $\mathbf{x}_{n,k}$ generated through rejection when the following metric exceeds a certain small value (set to 3×10^{-3} in our experiments):

$$\tilde{h}(y_k, \mathbf{x}_{n,k}) := h \frac{\frac{1}{2\pi} \frac{1}{\sigma_n^2 + \sigma_A^2} e^{-\frac{|y_k|^2}{2(\sigma_n^2 + \sigma_A^2)}}}{D(y_k, \mathbf{x}_{n,k})},$$

where $D(\cdot, \cdot, \cdot)$ was defined in Sec. 3.3. This outer rejection loop selects particles that are consistent with the new measurement (cf. the functional form of the denominator) and, at the same time, have large weight after the associated update. We do not have a full explanation at this point, yet this version of the algorithm appears to yield the best results - in particular, better than the one based on the optimal importance function. Note that the latter is optimal with respect to minimizing the variance of the weights after the update (and typically works better than the one based on the prior importance function), but it is not necessarily optimal in terms of the performance - complexity trade-off.

4. SIMULATIONS

We now present simulation results for the three algorithms: the basic one using the prior importance function (denoted P), the one using the optimal importance function (O), and the one using the outer rejection loop as above (V). Fig. 1 shows a plot of a typical simulation run, using the posterior mean to form instantaneous frequency estimates and multinomial resampling for all three algorithms. Monte-Carlo (MC) simulation results are presented in Fig. 2. The Root Mean Square Error (RMSE) frequency estimation performance of the three algorithms is assessed using the following parameters: $h = 0.01$, $T = 100$, $\sigma_A^2 = 1$, $\sigma_n^2 = 0.2$, and the number of MC trials is 300. The execution time for P is $O(NT)$, whereas for O and V the execution time is also an increasing function of h. As a result, O and/or V can be faster than P, even for the same number of particles. For our simulation setup above, P, O, and V, each with 1K particles, have about the same average execution time, yet V does much better in terms of RMSE as shown in Fig. 2. It takes 3K particles for O and 5K particles for P to reach the performance of V with 1K particles.

5. CONCLUSIONS

We have developed three new particle filtering algorithms for tracking a frequency-hopped complex sinusoid, based on a novel stochastic state-space formulation. The algorithms range from a plain-vanilla version that uses the prior importance function (P), to a more advanced version that employs the optimal importance function (O), and, finally, an improvement of the latter using a problem-specific outer rejection loop (V). The two latter algorithms afford considerably better performance - complexity trade-offs.

6. REFERENCES

[1] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[2] L. Aydin and A. Polydoros, "Hop-timing estimation for FH signals using a coarsely channelized receiver," *IEEE Trans. Communcations*, vol. 44, no. 4, pp. 516–526, Apr. 1996.

[3] S. Barbarossa and A. Scaglione, "Parameter estimation of spread spectrum frequency-hopping signals using time-frequency distributions," in *Proc. Signal Proc. Advances in Wireless Communications*, pp. 213–216, Apr. 1997.

[4] L. Devroye, *Non-uniform random variate generation*, Springer-Verlag, New York, 1986. Available on-line at http://cgm.cs.mcgill.ca/∼luc/rnbookindex.html

[5] P. Djuric, J.H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, J. Miguez, "Particle Filtering,", *IEEE Signal Processing Magazine*, pp. 19–38, Sep. 2003.

[6] X. Liu, N. D. Sidiropoulos, and A. Swami, "Blind high resolution localization and tracking of multiple frequency hopped signals," *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 889–901, Apr. 2002.

[7] X. Liu, N. D. Sidiropoulos, and A. Swami, "Joint Hop Timing and Frequency Estimation for Collision Resolution in Frequency Hopped Networks," *IEEE Trans. Wireless Communications*, to appear, Nov. 2005.

[8] M. K. Simon, U. Cheng, L. Aydin, A. Polydoros, and B. K. Levitt, "Hop timing estimation for noncoherent frequency-hopped M-FSK intercept receivers," *IEEE Trans. Communications*, vol. 43, no. 2/3/4, pp. 1144–1154, Feb./Mar./Apr. 1995.

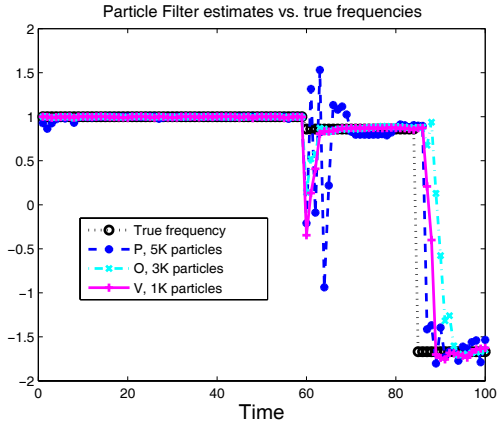

Fig. 1. Typical sample run of the three algorithms using different number of particles for each.

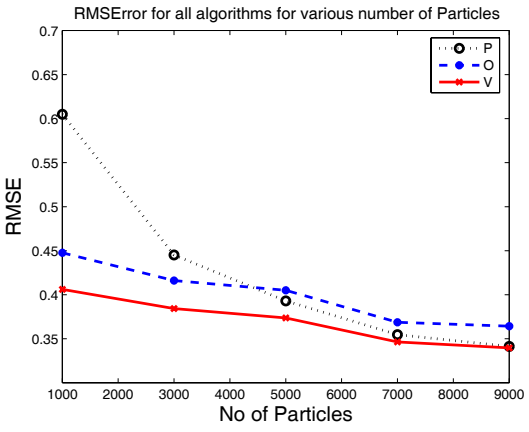

Fig. 2. MC simulation results for the three algorithms.