MULTIPLE MARGINALIZED POPULATION MONTE CARLO

Bingxin Shen, Mónica F. Bugallo, and Petar M. Djurić

Department of Electrical and Computer Engineering
Stony Brook University, Stony Brook, NY, 11794, USA
phone: + (1) 6316328424, fax: + (1) 6316328494, email: {bxshen, monica, djuric}@ece.sunysb.edu

ABSTRACT

Population Monte Carlo (PMC) algorithms iterate on a set of samples and weights to approximate a stationary target distribution. Their estimation quality and convergence efficiency rely on many factors including the number of samples and the choice of importance function. The computational complexity of the PMC algorithm becomes increasingly challenging as the numbers of the unknowns increases. In this paper, we propose a marginalized PMC algorithm for high-dimensional problems, where the state space of the system is partitioned into several subspaces of lower dimensions and handled by a set of marginalized PMC estimators. Simulation results show the accuracy and feasibility of the method as well as its improvement with respect to other conventional approaches.

1. INTRODUCTION

The Population Monte Carlo (PMC) algorithm is a topic of recent interest in the field of Monte Carlo-based signal processing. The PMC algorithms approximate a stationary target distribution by an iterative importance sampling procedure. An overview of the general PMC algorithm via computation of the products of non-negative sparse matrices is given in [1]. The algorithm has been developed and applied in the fields of quantum physics, polymer science, statistical physics, and statistical sciences.

The PMC algorithms have similarities with Markov Chain Monte Carlo (MCMC) sampling [1]. Both methods are useful tools for the calculation of multi-dimensional integrals. The MCMC algorithms draw samples and move them around the equilibrium distribution in relatively small steps, entailing that it might take a long time to explore the space [2]. PMC employs the resampling/reweighting concept, which updates the weights by learning from previous proposals and target distributions. The advantage of PMC over MCMC algorithms is that they are approximetely unbiased at every iteration and therefore can be stopped at any time. PMC is also more robust than MCMC on initialization parameters [3].

A PMC scheme was applied to missing data problems in [4]. Instead of using a constant importance function or a sequence of importance functions, importance functions that depend on both the iteration and the sample index were proposed. Advantages of this PMC scheme were illustrated for problems with settings of increasing difficulty, where the missing data could not be simulated or approximated through completion devices. A comparison with MCMC for missing data problems was also presented.

In [5], PMC was used to achieve variance reduction, which has always been a critical issue of Monte Carlo methods. A set of importance functions was iteratively optimized to minimize asymptotic variance. PMC methods were applied to restoration of ion channels using a fixed dimension model in [3]. PMC algorithms were also shown to be progressively adapted to a target distribution with a diminishing Kullbak-Leibler divergence in [6].

In many real-world problems, a high dimensional state space makes the PMC implementation very challenging due to the necessity of large number of samples. In some of these problems, some of the unknown parameters are conditionally linear given the remaining parameters. Marginalized PMC (MPMC) was proposed to lower the computational cost by only generating samples of the nonlinear parameters and marginalizing the remaining linear parameters [7]. This approach is based on the well-known Rao-Blackwell theorem.

The computational efficiency of the PMC method can be further improved by the use of a distributed structure. In this paper, we propose a novel method referred to as Multiple PMC (MultiPMC) where the state space of interest is partitioned into several subspaces with lower dimensions and handled by a set of parallel PMC filters. Each PMC filter updates the weights of the samples and the importance functions, if necessary, using information from the other PMC filters. A similar structure used for sequential Monte Carlo methods applied to the problem of target tracking can be found in [8]. A related approach to ours was the one from [9], where the intended application was in speaker recognition. A finite mixture of Gaussians was decomposed into subproblems, which were easier to work with. Then missing data were introduced, and samples were drawn from posteriors. We note, however, that drawing directly from posteriors is often infeasable. In this paper, we employ PMC algorithms to make the generation of samples easy.

The rest of the paper is organized as follows. A brief overview of the current state-of-art is presented in Section 2. The proposed MultiPMC scheme is presented in Section 3 as well as the Multiple MPMC (MultiMPMC). We demonstrate the implementation of the MultiPMC and MultiMPMC by applying it to the problem of frequency estimation of complex sinusoids in Section 4. We conclude with some final thoughts in Section 5.

This work has been supported by the National Science Foundation under CCF-0515246 and the Office of Naval Research under Award N00014-09-1-1154.

2. PROBLEM FORMULATION

The problem is to estimate an unknown vector of parameters $\mathbf{x}$ based on the vector of observations $\mathbf{y}$. The general model is

$$\mathbf{y} = h(\mathbf{x}) + \mathbf{v}, \tag{1}$$

where the observation $\mathbf{y}$ is a $d_y \times 1$ vector and the unknown parameter $\mathbf{x}$ is a $d_x \times 1$ vector with known prior density $p_0(\mathbf{x})$. In most cases, $h(\cdot)$ is a nonlinear function of unknown parameters. Finally, $\mathbf{v}$ is a $d_y \times 1$ additive noise vector with a known probability distribution $p(\mathbf{v})$.

2.1. PMC algorithm

The underlying principle of PMC is importance sampling [10], [11]. A commonly used point estimator is the minimum mean-square estimator (MMSE), which is defined as

$$\eta_x = \int \mathbf{x} p(\mathbf{x}|\mathbf{y}) d\mathbf{x}, \tag{2}$$

where $p(\mathbf{x}|\mathbf{y})$ is the posterior of $\mathbf{x}$. If we can draw samples from the posterior,

$$\mathbf{x}^{(m)} \sim p(\mathbf{x}|\mathbf{y}), \quad m = 1,2,...,M, \tag{3}$$

where M is the total number of independently drawn samples, then we can compute the integral in equation (2) according to classical Monte Carlo integration by

$$\hat{\eta}_x \simeq \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}^{(m)}. \tag{4}$$

This estimate will converge to the true value by the strong law of large numbers.

However, samples usually cannot be drawn directly from the posterior $p(\mathbf{x}|\mathbf{y})$ in practice. Alternatively, samples can be generated from another probability distribution $q(\mathbf{x})$, called importance function, and the estimate is computed as

$$\begin{aligned} \eta_x &= \int \mathbf{x} p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \\ &= \int \mathbf{x} \frac{p(\mathbf{x}|\mathbf{y})}{q(\mathbf{x})} q(\mathbf{x}) d\mathbf{x} \\ &\simeq \frac{1}{M} \sum_{m=1}^{M} \frac{\mathbf{x}^{(m)} p(\mathbf{x}^{(m)}|\mathbf{y})}{q(\mathbf{x}^{(m)})}. \end{aligned}$$

When $q(\mathbf{x})$ satisfies some conditions, it can be shown that by the strong law of large numbers, this estimate also converges to the true mean of the posterior.

The above approximation can also be written as

$$\eta_x \simeq \sum_{m=1}^{M} w^{(m)} \mathbf{x}^{(m)} \tag{5}$$

where $w^{(m)}$ denotes the weight of sample $\mathbf{x}^{(m)}$, i.e.,

$$w^{(m)} \propto \frac{p(\mathbf{x}^{(m)}|\mathbf{y})}{q(\mathbf{x}^{(m)})}, \tag{6}$$

and

$$\sum_{m=1}^{M} w^{(m)} = 1. \tag{7}$$

PMC employs an iterated and adaptive importance sampling scheme. It also uses resampling as do particle filtering methods, where samples with small weights are most likely removed and ones with large weights are replicated. The method can be summarized as follows. Let j denote the iteration number, $j = 1,2,...$, and let m represent the index of the particle, $m = 1,2,...,M$.

`Step 1.` Choose an importance function $q_j^{(m)}(\mathbf{x})$;

`Step 2.` Draw samples $\mathbf{x}_j^{(m)}$ from $q_j^{(m)}(\mathbf{x})$;

`Step 3.` Compute weights of the samples

$$\tilde{w}_j^{(m)} \propto \frac{p(\mathbf{x}_j^{(m)}|\mathbf{y})}{q_j^{(m)}(\mathbf{x}_j^{(m)})}; \tag{8}$$

`Step 4.` Normalize the weights:

$$w_j^{(m)} = \frac{\tilde{w}_j^{(m)}}{\Sigma_{k=1}^{M} \tilde{w}_j^{(k)}}; \tag{9}$$

`Step 5.` Resample the samples according to their weights;

`Step 6.` If more iterations are needed, set $j = j+1$, and go back to step 1.

2.2. MPMC algorithm

MPMC employs a scheme where PMC is only applied to the nonlinear parameters, while the linear parameters are obtained by analytical integrations with prior distributions. In high dimensional problems with some of the unknown parameters being conditionally linear given the remaining parameters, MPMC needs less particles than PMC, and therefore achieves an improved computational efficiency [7, 12, 13].

We assume that the model of the data is

$$\mathbf{y} = h(\mathbf{x}_n) + A(\mathbf{x}_n)\mathbf{x}_l + \mathbf{v}, \tag{10}$$

where the observation $\mathbf{y}$ is a $d_y \times 1$ vector and the unknown parameter $\mathbf{x}$ is a $d_x \times 1$ vector. The vector $\mathbf{x}$ is composed of nonlinear parameters $\mathbf{x}_n$ of dimension d_{x_n} and linear parameters $\mathbf{x}_l$ of dimension d_{x_l}, where $d_x = d_{x_n} + d_{x_l}$, and the prior density of $\mathbf{x}$ is given by $p(\mathbf{x}_n, \mathbf{x}_l)$. As in (1), $h(\cdot)$ is a nonlinear function of the parameters $\mathbf{x}_n$; $A(\mathbf{x}_n)$ is a matrix of functions of the nonlinear parameter $\mathbf{x}_n$ and has dimension $d_y \times d_{x_l}$; and $\mathbf{v}$ is a $d_y \times 1$ noise vector with a known probability distribution $p(\mathbf{v})$.

In the MPMC algorithm, at iteration j, one only generates samples of the nonlinear parameters, $\mathbf{x}_{n,j}^{(m)}$. The corresponding weights to these samples are

$$w_{n,j}^{(m)} \propto \frac{p(\mathbf{x}_{n,j}^{(m)}|\mathbf{y})}{q_{n,j}^{(m)}(\mathbf{x}_{n,j}^{(m)})}, \tag{11}$$

The numerator $p(\mathbf{x}_{n,j}^{(m)}|\mathbf{y})$ is the marginalized posterior of $\mathbf{x}_n$, which has the following property

$$p(\mathbf{x}_{n,j}^{(m)}|\mathbf{y}) \propto \int p(\mathbf{y}|\mathbf{x}_l, \mathbf{x}_{n,j}^{(m)}) p(\mathbf{x}_l, \mathbf{x}_{n,j}^{(m)}) d\mathbf{x}_l. \tag{12}$$

The proposed MPMC algorithm is summarized as follows. Let j denote the iteration number, $j = 1, 2, ...,$ and let m denote the index of the particle, $m = 1, 2, ..., M$.

`Step 1.` Choose an importance function $q_{n,j}^{(m)}(\mathbf{x}_{n,j})$;

`Step 2.` Draw samples $\mathbf{x}_{n,j}^{(m)}$ from $q_{n,j}^{(m)}(\mathbf{x}_{n,j})$;

`Step 3.` Based on $\mathbf{x}_{n,j}^{(m)}$, use the MMSE criterion to estimate the corresponding $\mathbf{x}_l$;

`Step 4.` Compute weights of the samples

$$\tilde{w}_{n,j}^{(m)} = \frac{p(\mathbf{x}_{n,j}^{(m)}|\mathbf{y})}{q_{n,j}^{(m)}(\mathbf{x}_{n,j}^{(m)})}; \tag{13}$$

`Step 5.` Normalize the weights

$$w_{n,j}^{(m)} = \frac{\tilde{w}_{n,j}^{(m)}}{\sum_{k=1}^{M} \tilde{w}_{n,j}^{(k)}}; \tag{14}$$

`Step 6.` Resample the samples according to their weights;

`Step 7.` If more iterations are needed, set $j = j + 1$, and go back to step 1.

3. MULTIPLE PMC

Besides marginalizing the linear parameters, one can avoid generation of too many samples for accurate estimation in high-dimensional problems by partitioning the problem into subproblems and use "independent" PMC algorithms for each subproblem [8, 9]. The partitioning often depends on the problem [14]. By decomposing the original problem, one can considerably reduce the computational complexity.

3.1. Multiple PMC

We will further assume that the model in equation (1) can be partitioned into K subproblems as follows:

$$\mathbf{y} = \sum_{k=1}^{K} h_k(\mathbf{x}_k) + \mathbf{v}, \tag{15}$$

where $[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K]$ forms the unknown vector $\mathbf{x}$ in the general model described by (1).[1]

We assign each unknown vector $\mathbf{x}_k$ a PMC filter with the target distribution $p(\mathbf{x}_k|\mathbf{y})$. The sample generation/propagation and resampling step of each PMC estimator can be implemented as the algorithm stated in Section 2.1. The key question is the weight updating of the samples for $\mathbf{x}_k$ based on the other PMC filters. Theoretically the weight update $\tilde{w}_{k,j}^{(m)}$ should be carried out by

$$\tilde{w}_{k,j}^{(m)} = \frac{p(\mathbf{x}_{k,j}^{(m)}|\mathbf{x}_{-k}, \mathbf{y})}{q_{k,j}^{(m)}(\mathbf{x}_j^{(m)})}, \tag{16}$$

where $\mathbf{x}_{-k}$ contains the true values of all unknowns except $\mathbf{x}_k$. This form of update requires the knowledge of $\mathbf{x}_{-k}$, which is not available. Here we propose to implement the updates as

$$\tilde{w}_{k,j}^{(m)} = \frac{p(\mathbf{x}_{k,j}^{(m)}|\tilde{\mathbf{x}}_{-k}, \mathbf{y})}{q_{k,j}^{(m)}(\mathbf{x}_{k,j}^{(m)})}, \tag{17}$$

where $\tilde{\mathbf{x}}_{-k}$ are the most recent estimated values of all the unknowns except $\mathbf{x}_k$

$$\tilde{\mathbf{x}}_{-k} = \tilde{\mathbf{x}} \setminus \tilde{\mathbf{x}}_k = [\tilde{\mathbf{x}}_1^\top, \tilde{\mathbf{x}}_2^\top, ..., \tilde{\mathbf{x}}_{k-1}^\top, \tilde{\mathbf{x}}_{k+1}^\top, ..., \tilde{\mathbf{x}}_K^\top]^\top,$$

and

$$\tilde{\mathbf{x}}_k = \sum_{m=1}^{M} w_{k,j}^{(m)} \mathbf{x}_{k,j}^{(m)}, \tag{18}$$

where $w_{k,j}^{(m)}$ is the normalized weight.

In each iteration, the PMC estimators use the exchanged estimates to compute their weights in an alternating way. For good performance of the method, we propose that in each iteration the implementation order of the PMC filters is selected randomly.

3.2. Multiple MPMC

The distributed structure of multiple estimators can also be applied to MPMC methods. If we modify (10) for a model of type (15), we can write

$$\mathbf{y} = \sum_{k=1}^{K} (h_k(\mathbf{x}_{k,n}) + A_k(\mathbf{x}_{k,n})\mathbf{x}_{k,l}) + \mathbf{v}, \tag{19}$$

where $[\mathbf{x}_{1,n}, \mathbf{x}_{1,l}, \mathbf{x}_{2,n}, \mathbf{x}_{2,l}, ..., \mathbf{x}_{K,n}, \mathbf{x}_{K,l}]$ form the unknown vector $\mathbf{x}$ in the general model described in (10).

We assign each unknown vector $\mathbf{x}_{k,n}$ an MPMC filter, and use MMSE to estimate the corresponding marginalized linear unknowns $\mathbf{x}_{k,l}$. The sample generation/propagation and resampling step of each MPMC filter can be implemented in the usual way. The proposed weight update is implemented by

$$\tilde{w}_{k,n,j}^{(m)} = \frac{p(\mathbf{x}_{k,n,j}^{(m)}|\tilde{\mathbf{x}}_{-k,n}, \mathbf{y})}{q_{k,n,j}^{(m)}(\mathbf{x}_{k,n,j}^{(m)})}, \tag{20}$$

where $\tilde{\mathbf{x}}_{-k,n}$ represents the most recent estimated values of all nonlinear unknowns except $\mathbf{x}_{k,n}$, i.e.,

$$\tilde{\mathbf{x}}_{-k,n} = \tilde{\mathbf{x}}_n \setminus \tilde{\mathbf{x}}_{k,n} = [\tilde{\mathbf{x}}_{1,n}^\top, \tilde{\mathbf{x}}_{2,n}^\top, ..., \tilde{\mathbf{x}}_{k-1,n}^\top, \tilde{\mathbf{x}}_{k+1,n}^\top, ..., \tilde{\mathbf{x}}_{K,n}^\top]^\top,$$

and

$$\tilde{\mathbf{x}}_{k,n} = \sum_{m=1}^{M} w_{k,n,j}^{(m)} \mathbf{x}_{k,n,j}^{(m)}. \tag{21}$$

The implementation order of each MPMC estimator is randomized at each iteration as before.

[1] Equation (15) represents only one particular case where MultiPMC can be applied.

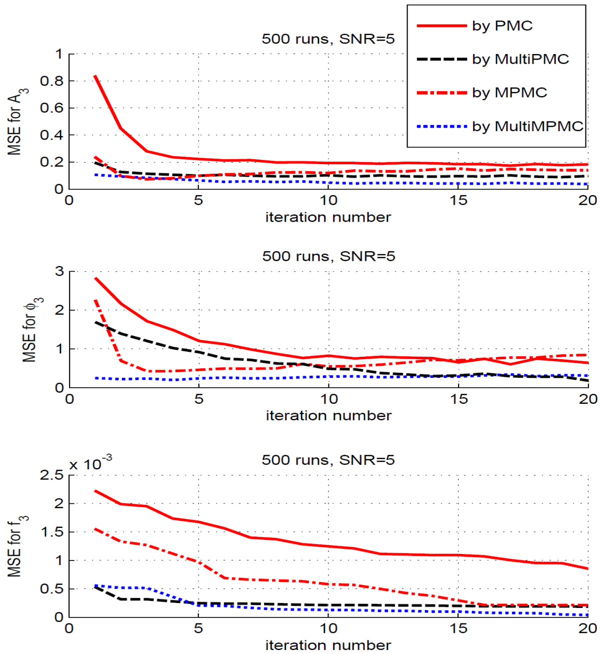

Fig. 1. Estimates for f_3 vs iterations using the PMC, MultiPMC, MPMC, and MultiMPMC algorithms.

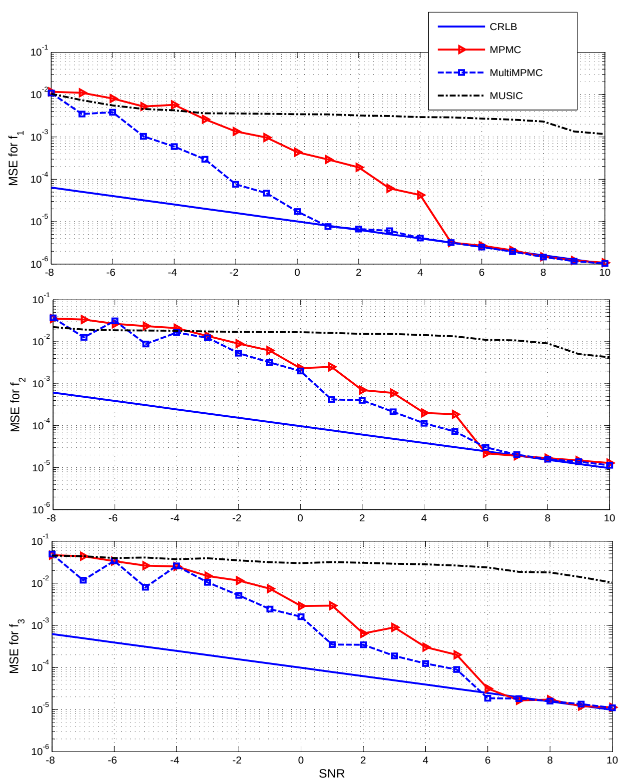

Fig. 2. Estimates vs SNR using MPMC, MultiMPMC and MUSIC algorithms.

4. SIMULATIONS

In this section, we consider the problem of estimating the frequencies of complex sinusoids. The model of the observations is given by

$$y_t = \sum_{k=1}^{K} A_k e^{i(2\pi f_k t+\phi_k)} + v_t, \quad t = 1,2,...,d_y, \tag{22}$$

where $i=\sqrt{-1}$, $0 < f_1 < f_2 < ... < f_K < 1$; A_k and ϕ_k are the amplitude and phase, respectively, of the k-th frequency component; and v_t is white complex Gaussian noise. The parameters to be estimated are $\mathbf{x} = [A_1, \phi_1, f_1, ..., A_K, \phi_K, f_K]$, and therefore the space of unknowns has dimension $3K$.

The complex noise was drawn from the distribution

$$v_t \sim \mathcal{CN}(0, \sigma_v^2),$$

or more specifically,

$$real(v_t) \sim \mathcal{N}(0, \tfrac{\sigma_v^2}{2}), \text{ and } imag(v_t) \sim \mathcal{N}(0, \tfrac{\sigma_v^2}{2}),$$

and the value of the variance was defined by using the signal-to-noise ratio (SNR)

$$SNR = \frac{A_k^2}{{\sigma_v}^2},$$

where SNR was measured in dB, A_k was the amplitude of the signal, and σ_v^2 was the noise power.

The model described by (22) can be rewritten as

$$y_t = \textstyle\sum_{k=1}^{K} \tilde{A}_k e^{i2\pi f_k t} + v_t, \quad t = 1,2,...,d_y,$$

where f_k are the nonlinear parameters, and $\tilde{A}_k = A_k e^{i\phi_k}$ are the linear parameters.

We had $d_y = 25$ observations with $K = 3$ sinusoids generated according to the model. Observations were generated using three frequencies $f_1 = 0.2$, $f_2 = 0.5$ and $f_3 = 0.52$, with amplitudes $A_1 = 1$, $A_2 = 1$ and $A_3 = 1$, and phases $\phi_1 = 0$, $\phi_2 = 0$ and $\phi_3 = \pi/4$. PMC, MultiPMC, MPMC and MultiMPMC were applied to estimate the parameters in this 9-dimensional problem. For the prior of $\tilde{A}_k$ in MPMC and MultiMPMC we used

$$\tilde{A}_k \sim \mathcal{CN}(0, 5).$$

The initial importance functions for the frequencies had preselected means at the estimates obtained by the Yule-Walker method [15], and a predetermined variance vector [16] given by $\mathbf{v} = {\sigma_0}^2 \times [1^2, 0.1^2, 0.01^2, 0.001^2, 0.0001^2]^\top$ with ${\sigma_0}^2 = 0.1^2$. At the initial step, for the frequencies, a variance from the available five variances was assigned randomly to each particle with probability of $\frac{1}{5}$. After each iteration, the weights of available variances for each parameter were updated separately according to the performance of the samples. Updated importance functions had means located at the previous samples after resampling, and had variance coming from the predetermined variance vector with updated weights. In order to keep every variance valid after each iteration, rescaling was employed to ensure that the minimum weight for each available variance was 0.05.

The performances of the methods for estimation of the parameters of the third sinusoid are shown in Figure 1. The performances of the algorithms to this problem were quantified based on the MSE of the parameters to be estimated, given by

$$MSE = \frac{1}{R}\sum_{r=1}^{R}(\hat{x}^r - x)^2, \tag{23}$$

where r represented the r-th run of the algorithm, $\hat{x}^r$ denoted the estimates obtained in the r-th run, and x was the true value of the parameter. All the points on the plot were averaged over $R = 500$ runs for iteration number of 1 to 20 with $SNR = 5\ dB$. In each run, an amount of $M = 600$ samples were generated from an initial importance function for PMC and MPMC, and $M_k = 200$ for each filter, which sums up to $M = 600$ total samples, when implementing the MultiPMC and MultiMPMC algorithms. It can be concluded from the plots that MultiPMC and MultiMPMC perform accurately and converge much faster.

The performances of the proposed methods in terms of the MSE of the estimated frequencies for various values of SNR are shown in Figure 2. All the points on the plot are averaged over $R = 500$ runs with sample size of $M = 600$ and iteration number of $J = 20$. The methods are also compared with the MUltiple SIgnal Classification (MUSIC) algorithm, which estimates the pseudospectrum of the observations using Schmidt's eigenspace analysis method [17]. The conventional MUSIC algorithm performs similarly with the proposed methods at low SNRs, but does not improve with SNR as do the PMC methods. The poor performance of MUSIC was caused by the small difference between f_2 and f_3. The proposed methods perform well, and their MSEs converge to the Crámer-Rao lower bound (CRBL) as the SNR increases.

5. CONCLUSION

In this paper, we propose new PMC algorithms for high-dimensional nonlinear problems. The algorithms have distributed structures, and we refer to them as MultiPMC and MultiMPMC. With the approach, a high-dimensional problem is partitioned into several subproblems with lower dimensions and handled by a set of PMC or MPMC filters. Simulation results have shown the accuracy of the estimates and the feasibility of the methods.

6. REFERENCES

[1] Y. Iba, " Population Monte Carlo algorithms," *Transactions of the Japanese Society for Artificial Intelligence*, vol. 16, no. 2, pp. 279–286, 2001.

[2] C. P. Robert and G. Casella, *Monte Carlo statistical methods*. New York: Springer, 2004.

[3] O. Cappé, A. Guillin, J. M. Marin, and C. P. Robert, "Population Monte Carlo," *Journal of Computational and Graphical Statistics*, vol. 13, pp. 907–929, 2004.

[4] G. Celeux, J. M. Marin, and C. P. Robert, "Iterated importance sampling in missing data problems," *Computational Statistics & Data Analysis*, vol. 50, no. 12, pp. 3386–3404, August 2006.

[5] R. Douc, A. Guillin, J. M. Marin, and C. P. Robert, "Minimum variance importance sampling via Population Monte Carlo," *ESAIM: Probability and Statistics*, vol. 11, pp. 427–447, 2007.

[6] ——, "Convergence of adaptive mixtures of importance sampling schemes," *Annals of statistics*, vol. 35, pp. 420–448, 2007.

[7] M. F. Bugallo, M. Hong, and P. M. Djurić, "Marginalized Population Monte Carlo," *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009 (ICASSP 2009)*, pp. 2925–2928, 2009.

[8] P. M. Djuric, T. Lu, and M. Bugallo, "Multiple particle filtering," *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007 (ICASSP 2007)*, pp. 1181–1184, 2007.

[9] A. Doucet, S. Sénécal, and T. Matsui, "Space alternating data augmentation: Application to finite mixture of gaussians and speaker recognition," *IEEE International Conference on Acoustics, Speech and Signal Processing, 2005 (ICASSP 2005)*, pp. 713–716, 2005.

[10] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Method in Practice*. New York: Springer, 2001.

[11] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 1968.

[12] C. Andrieu and A. Doucet, "Joint bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Processing*, vol. 47, no. 10, pp. 2667–2676, October 1999.

[13] C. Andrieu, A. Doucet, , and C. P. Robert, "Computational advances for and from bayesian analysis," *Statistical Science*, vol. 19, no. 1, pp. 118–127, 2004.

[14] D. G. Lainiotis, *International Symposium on Systems Optimization and Analysis*. Springer Berlin / Heidelberg, 1979, ch. Partitioning: The multi-model framework for estimation and control, I: Estimation, pp. 252–290.

[15] P. Stoica, R. L. Moses, B. Friedlander, and T. Soderstrom, "Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements," *IEEE Trans. Acoustics. Speech, and Signal Processing*, vol. 37, no. 3, March 1989.

[16] O. Cappé, R. Douc, A. Guillin, J. M. Marin, and C. P. Robert, "Adaptive importance sampling in general mixture classes," *Statistics and Computing*, vol. 18, no. 4, pp. 447–459, December 2008.

[17] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. AP-34, pp. 276–280, March 1986.