ESTIMATION OF MULTIMODAL POSTERIOR DISTRIBUTIONS OF CHIRP PARAMETERS WITH POPULATION MONTE CARLO SAMPLING

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

Department of Electrical & Computer Engineering
Stony Brook University, Stony Brook, NY 11794-2350
email: {bxshen, monica, djuric}@ece.sunysb.edu

ABSTRACT

Chirp signals are usually encountered in target tracking problems including radar and sonar systems. The multimodality characterizing the distribution of the chirp signal parameters makes their estimation very challenging. In this paper we apply marginalized population Monte Carlo (MPMC) sampling to the problem of parameter estimation of chirp signals in noise. MPMC reduces the dimension of the vector of unknowns by marginalizing the complex amplitudes, which are conditionally linear on the chirp rates and frequencies. A Gibbs sampling scheme is combined with the MPMC method to further improve the performance. Computer simulations illustrate the validity of the proposed approach.

Index Terms— Population Monte Carlo, marginalization, Gibbs sampling, multimodality.

1. INTRODUCTION

Many signal processing problems involve the estimation of unknown parameters from some given observations. In some of the problems, the multimodality of the posterior distributions of interest of the unknowns makes the estimation very challenging. The parameter estimation of chirp signals is one such problems and has received a great deal of attention in the past decades [1, 2, 3].

The parameter estimation of chirp signals is usually encountered in velocity estimation and target tracking problems in radar, and sonar systems, as well as robotics and embedded sensor systems. Several approaches have been proposed for parameter estimation of a single chirp signal in [1, 2, 3]. These methods have been proven to be effective and achieve the Cramér-Rao lower bound (CRLB) at high signal-to-noise ratios (SNRs).

An iterative maximum likelihood (ML) algorithm was proposed in [4] for parameter estimation of multiple chirp signals. However, this method achieves good results only at high SNRs. The method also involves a great amount of computations due to its iterative nature.

A noniterative ML approach based on the concept of importance sampling was proposed in [5]. The method was designed to estimate the nonlinear parameters, i.e., the chirp rates and frequencies. It was observed that this technique provides good estimates even when the frequency components are closely spaced. However, the approach is still computationally very expensive since it involves grid searches of inverse integrals and calculations of multidimensional integrals.

This work has been supported by the National Science Foundation under Awards CCF-0953316 and CCF-1018323, and the Office of Naval Research under Award N00014-09-1-1154.

All of the above methods produce point estimates. In this paper we propose to apply the population Monte Carlo (PMC) sampling [6, 7] in a Gibbs manner [8] and aim at obtaining an approximation of the posterior distribution of the unknowns of interest. Thus, rather than obtaining point estimates, we aim at estimating the joint posteriors of the unknowns. The posteriors, as is well known, contain all the information about the unknowns, given their priors, the data and the model. The proposed method marginalizes the complex amplitudes, and only particles of the nonlinear parameters, i.e., the chirp rates and frequencies are generated. The contributions of this paper are in the extensions of a previously proposed method for estimating posteriors of sinusoidal parameters [8]. Computer simulations illustrate the feasibility of the method which shows good performance.

The paper is organized as follows. In the next section we provide a general formulation of the problem. In Section 3, we briefly review the PMC method and some recent advances. In Section 4, we propose the Gibbs sampling-inspired PMC and the details of its implementation. We demonstrate the use and performance of the method on the problem of frequency and chirp rate estimation of chirp signals in Section 5. We conclude the paper with Section 6.

2. PROBLEM FORMULATION

We observe a sequence of data y_t which have the following representation

$$y_t = \sum_{k=1}^{K} A_k \exp\left[i(2\pi f_k t + \pi \alpha_k t^2)\right] + v_t, \quad t = 1, 2, ..., d_y \quad (1)$$

where $i = \sqrt{-1}$; A_k is a complex amplitude; f_k denotes frequency; α_k is the chirp rate; and v_t is a complex white Gaussian noise.

The observation vector is $\mathbf{y} = [y_1, ..., y_{d_y}] \in C^{d_y}$ and the unknown parameter vector is $\mathbf{x} = [|A_1|, \angle A_1, f_1, \alpha_1, ..., |A_K|, \angle A_K, f_K, \alpha_K] \in R^{4K}$, where the symbol $|\cdot|$ denotes magnitude of the argument, and $\angle\cdot$ represents phase of the parameter. The vector $\mathbf{x}$ is composed of nonlinear parameters $\mathbf{x}_n = [f_1, \alpha_1, ..., f_K, \alpha_K] \in \mathcal{X} \subset R^{2K}$ and linear parameters $\mathbf{x}_l = [|A_1|, \angle A_1, ..., |A_K|, \angle A_K] \in \mathcal{A} \subset R^{2K}$. The noise vector $\mathbf{v} = [v_1, ...v_{d_y}]$ has known probability distribution $p(\mathbf{v})$.

We assume that we know the a priori distribution $\pi(\mathbf{x})$, and that given the noise probability distribution, we can write the conditional distribution $p(\mathbf{y}|\mathbf{x})$. Given the observation vector $\mathbf{y}, \pi(\mathbf{x})$, and $p(\mathbf{y}|\mathbf{x})$, we want to compute the posterior distribution $p(\mathbf{x}|\mathbf{y})$, which can be written as

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})\pi(\mathbf{x}),$$

where $\propto$ symbolizes proportionality. We refer to $p(\mathbf{x}|\mathbf{y})$ as our target distribution. We may not be interested in the complete posterior of

978-1-4673-0046-9/12/$26.00 ©2012 IEEE

$\mathbf{x}$, and instead, we only want to obtain the posterior of some subset of $\mathbf{x}$, in this case the frequencies and chirp rates.

3. STATE-OF-THE-ART IN PMC

The PMC methodology was described and summarized in [7]. It produces a set of particles at each iteration that are used for approximation of the posterior distributions of the set of unknown parameters. The iterative structure allows for adaptivity and convergence towards the target distribution [7].

The robustness regarding initialization of the PMC algorithm was shown in [7] with an example of its application to Bayesian modeling of a Gaussian mixture and ion channel models. As the generation of particles with iterations proceeds, the quality of the generated particles improves. This is a very important feature of PMC. For example, it has been shown that PMC can be used for variance reduction [9], where a mixture of generating functions can be iteratively optimized to achieve a minimum variance of a function of interest.

In our previous work, we have addressed the problem of estimation of frequencies of multiple sinusoids in noise with marginalized PMC (MPMC) [10]. There we exploited the principle of Rao-Blackwellization to improve the efficiency of the method by marginalizing the unwanted parameters (all the parameters except the frequencies). In other words, we applied the PMC only on the nonlinear parameters of the model. We have used several PMC algorithms that operate in parallel, each of them producing particles and weights of a subset of the parameters.

We also proposed to use Gibbs sampling in our previous work on PMC [8]. There we generated the particles that represented particles from the joint probability distribution of two or more unknowns in a Gibbs manner [11]. Note that Gibbs sampling belongs to the larger class of Markov-chain Monte Carlo (MCMC) methods and is often used for Bayesian inference [12].

4. THE GIBBS MPMC

In MCMC methods, at iteration j, we generally construct the m-th particle $\mathbf{x}_j^{(m)} = [x_{1,j}^{(m)}, x_{2,j}^{(m)}, \cdots, x_{d_x,j}^{(m)}]^\top$ by drawing each component from a proposal function $q_j(x_k)$ as $x_{k,j}^{(m)} \sim q_j(x_k)$, $k = 1, 2, \cdots, d_x$ [12]. We then accept/reject these proposals individually in parallel MCMC sampling, or we reject the complete $\mathbf{x}_j^{(m)}$ globally.

Gibbs sampling is a special type of MCMC sampling where each $x_{k,j}^{(m)}$ is sampled from a conditional distribution $q_j(x_k|\mathbf{x}_{-k})$. The symbol $\mathbf{x}_{-k}$ is the vector of all the parameters in $\mathbf{x}$ except for x_k at their current values, i.e., one uses the drawn values from the last iteration for the conditioning parameters.

We combine the Gibbs sampling with PMC and refer to it as Gibbs PMC [8]. Instead of drawing a particle from $p(\mathbf{x})$, which is usually infeasible, we obtain a particle $\mathbf{x}^{(m)}$ from $q(\mathbf{x})$ and assign it an importance weight given by

$$\tilde{w}^{(m)} = \frac{p(\mathbf{x}^{(m)})}{q(\mathbf{x}^{(m)})}.$$

The weights and the particles form a random measure, $\chi = \{\mathbf{x}^{(m)}, w^{(m)}\}_{m=1}^M$, where $w^{(m)}$ denotes the normalized weight corresponding to particle $\mathbf{x}^{(m)}$ and M is the total number of particles. Here we reiterate that in PMC, we implement the generation of particles in iterations. For example, in iteration one, we get the random measure χ_1, in iteration two, the random measure χ_2 and so on. The objective is that, as we proceed with iterations, we improve the accuracy of the approximation. To that end, for obtaining better generating functions, one uses the approximations from the previous iterations. One way of exploiting the previous iteration is to employ resampling, which is another operation that is common in particle filtering [13]. That is, we construct new generating functions by using particles from the previous iteration that are selected based on their weights.

In this paper we propose an approach for constructing generating functions for the chirp problem. Instead of drawing particles of particular scalar unknowns from their conditional distributions, we draw the particles from a conditional distribution for a set of frequency and chirp rate, where the conditioning is on the remaining unknowns. We basically mimic the Gibbs sampling idea, where as explained, we replicate the same steps except that our conditionals are not obtained from the target distribution. Note that in PMC, we assume that we *cannot* generate from the conditionals of the target distribution, and therefore we work with a different joint distribution, but one that allows for easy drawing of particles.

We now describe the specific steps of the proposed scheme. At iteration $j = 0$, we initialize the particle streams by drawing them from the prior $\pi(\mathbf{x})$. We draw M particles, and to each of them we assign weights according to

$$\tilde{w}_0^{(m)} = p(\mathbf{y}|\mathbf{x}_0^{(m)}).$$

The Gibbs MPMC method as applied to our problem can be summarized as follows:

`Step 1.` Randomly choose the order of generation of the parameters related to different frequency components $\mathbf{x}_j^{(m)} = [f_{1,j}^{(m)}, \alpha_{1,j}^{(m)}, f_{2,j}^{(m)}, \alpha_{2,j}^{(m)}, \cdots, f_{K,j}^{(m)}, \alpha_{K,j}^{(m)}]^\top$. Let the order be $l_1, l_2, \cdots, l_K$.

`Step 2.` For $m = 1, 2, \cdots, M$, choose a particle for conditioning based on the normalized weights of the particles. Let the selected particle be with index λ_m. Then generate new particles according to

$$f_{l_1,j}^{(m)}, \alpha_{l_1,j}^{(m)} \sim q_{l_1,j}\left(f_{l_1}, \alpha_{l_1}|f_{l_2,j}^{(\lambda_m)}, \alpha_{l_2,j}^{(\lambda_m)}, \cdots, f_{l_K,j}^{(\lambda_m)}, \alpha_{l_K,j}^{(\lambda_m)}\right),$$

for $n = 2, 3, \cdots, K-1$,

$$f_{l_n,j}^{(m)}, \alpha_{l_n,j}^{(m)} \sim q_{l_n,j}\left(f_{l_n}, \alpha_{l_n}|f_{l_1,j}^{(m)}, \alpha_{l_1,j}^{(m)}, \cdots f_{l_{n-1},j}^{(m)}, \alpha_{l_{n-1},j}^{(m)}, f_{l_{n+1},j-1}^{(\lambda_m)}, \alpha_{l_{n+1},j-1}^{(\lambda_m)}, \cdots, f_{l_K,j-1}^{(\lambda_m)}, \alpha_{l_K,j-1}^{(\lambda_m)}\right),$$

and

$$f_{l_K,j}^{(m)}, \alpha_{l_K,j}^{(m)} \sim q_{l_K,j}\left(f_{l_K}, \alpha_{l_K}|f_{1,j}^{(m)}, \alpha_{1,j}^{(m)}, \cdots, f_{K-1,j}^{(m)}, \alpha_{K-1,j}^{(m)}\right).$$

`Step 3.` Compute the weights of the particles by

$$\tilde{w}_j^{(m)} = \frac{p\left(\mathbf{y}|\mathbf{x}_j^{(m)}\right)}{\prod_{n=1}^K q_{l_n,j}(f_{l_n,j}^{(m)}, \alpha_{l_n,j}^{(m)})}.$$

Step 4. Normalize the weights according to

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

Step 5. Resample the particles according to their weights.

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

Step 7. At the end, obtain an approximation of the posterior of the unknowns. Note that all available weighted particles (from all samplers and all iterations) are used to approximate the posterior $p(\mathbf{x}|\mathbf{y})$.

The computed weights are stored as they were obtained by the last expression. They are normalized at the end of the algorithm with the weights of the particles from all the iterations in order to get the best possible approximation of the distribution of interest.

5. PARAMETER ESTIMATION OF MULTIPLE CHIRP SIGNALS

We demonstrate the proposed method on the problem of frequency estimation of complex chirp signals in noise [5]. The data were modeled as in (1). The frequencies were normalized and $0 < f_1 < f_2 < ... < f_K < 1$. For the chirp rates we had $0 \leq \alpha_k \leq 2$. The measurement noise v_t was white complex Gaussian of the form

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

with real and imaginary components that were independent and had distribution $\mathcal{N}(0, \frac{\sigma_v^2}{2})$. We reiterate that the vector of unknowns was

$$\mathbf{x} = [|A_1|, \angle A_1, f_1, \alpha_1, ..., |A_K|, \angle A_K, f_K, \alpha_K, \sigma_v^2]$$

and therefore the space of unknowns had dimension $4K + 1$.

We were primarily interested in the chirp rates and frequencies, so we worked with the MPMC method as described in [14]. Thus, the parameter space of interest was $\mathbf{x} = [f_1, \alpha_1, ..., f_K, \alpha_K]$. The posterior $p(\mathbf{x}|\mathbf{y})$ could be obtained in a closed analytical form [14], but particles could not be obtained from it. Here, we applied a scheme where each of the conditionals was a truncated Gaussian centered at the selected particle from the previous iteration. The conditioning parameters were used for deciding the truncation points of the Gaussian. For example, if the chirp rate $\alpha_{k,j}$ needed to be generated, we used as a generating function the truncated Gaussian centered at $\alpha_{k,j-1}^{(m)}$ with cutoff points determined by the most recent particles representing the closest smaller and larger chirp rates from the remaining components.[1] The joint sampling of the parameters could be easily carried out by initially setting a set of bivariate Gaussians with covariances $[\sigma_1^2\mathbf{I}, \sigma_2^2\mathbf{I}, \cdots, \sigma_L^2\mathbf{I}]$, where L represents the number of Gaussians, all used for particle generation. At the first iteration, using the particles and the weights the correlation coefficient ρ of f and α is estimated. Let this estimate be $\hat{\rho}_1$. Then one constructs $\mathbf{C}_1$ by

$$\mathbf{C}_1 = \begin{bmatrix} 1 & \hat{\rho}_1 \\ \hat{\rho}_1 & 1 \end{bmatrix}$$

and forms the next set of Gaussians using $[\sigma_1^2\mathbf{C}_1, \sigma_2^2\mathbf{C}_1 \cdots, \sigma_L^2\mathbf{C}_1]$. The method proceeds in an obvious way. In the next iteration, it again estimates ρ, constructs $\mathbf{C}$, and updates the importance functions.

We tested the method by conducting simulations as follows. We simulated $d_y = 50$ observations of $K = 2$ closely spaced chirp signals with complex amplitudes $A_1 = 1$ and $A_2 = 1$, frequencies $f_1 = 0.3$ and $f_2 = 0.32$, chirp rates $\alpha_1 = 0.001$ and $\alpha_2 = 0.002$, respectively. The value of the noise power was defined by using the SNR

$$\text{SNR}_k = 10 \log_{10} \frac{|A_k|^2}{\sigma_v^2}$$

measured in dB.

Figure 1 shows the multimodality of the periodogram of a set of observations at $SNR = 5dB$. The lower figure shows the details of the area of interest, where we can see several local maxima as red ellipses. We applied the Gibbs MPMC to this set of data with $M = 1000$ particles per iteration.

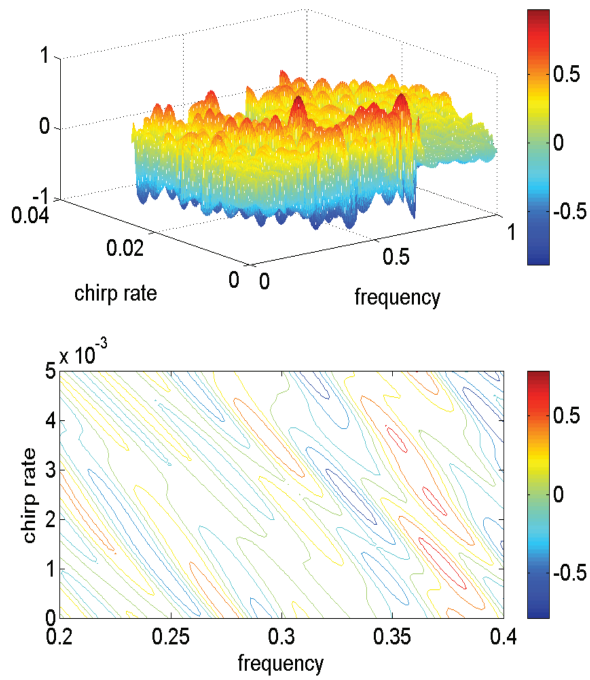

Fig. 1. Periodogram of a set of observations (top) and contour plot of the periodogram in part of the frequency-chirp rate plane (bottom).

All available weighted particles (from all samplers and 10 iterations) were used to approximate the marginalized posteriors $p(f_1, f_2|\mathbf{y})$ and $p(\alpha_1, \alpha_2|\mathbf{y})$, shown in Figure 2. The figure confirms that the marginalized posteriors of the frequencies and chirp rates have most of the probability masses around the true values.

The performance of the method was quantified in terms of the mean square error (MSE) given by

$$\text{MSE}_{x_d} = \frac{1}{R} \sum_{r=1}^{R} (\widehat{x}_d(r) - x_d)^2, \tag{2}$$

where R represents the number of realizations, $\widehat{x}_d(r)$ denotes the estimate of the d-th unknown obtained in the r-th run, and x_d is its

[1] For the lowest value of chirp rate α, the lower cutoff point is 0, and for the highest value of chirp rate the upper cutoff point is 2; for the frequency f when $k = 1$, the lower cutoff point is 0, and when $k = K$ the upper cutoff point is 1.

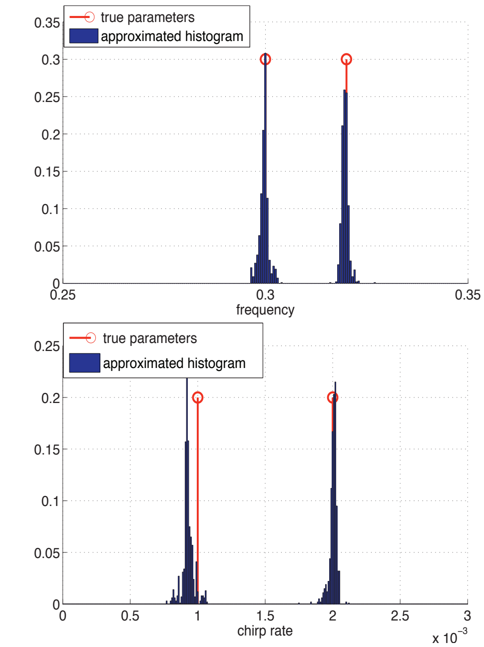

Fig. 2. The approximated posteriors $p(f_1, f_2|\mathbf{y})$ (top) and $p(\alpha_1, \alpha_2|\mathbf{y})$ (bottom).

true value. Figure 3 shows the MSE of the method as a function of SNR with $J = 10$ iterations and $M = 1000$ particles per iteration. At each run, the estimates were obtained from all available particles. Each point in the figure was averaged over $R = 100$ runs. In the figure, the performance of the proposed scheme achieves the CRLB at SNR$>$ 2dB, which is as good as the results in [5]. However, the method in [5] requires a large number of integral calculations and grid searches of inverse integrals in each run to achieve the CRLB. More importantly, the method in [5] does not provide estimates of the posterior as does the proposed method.

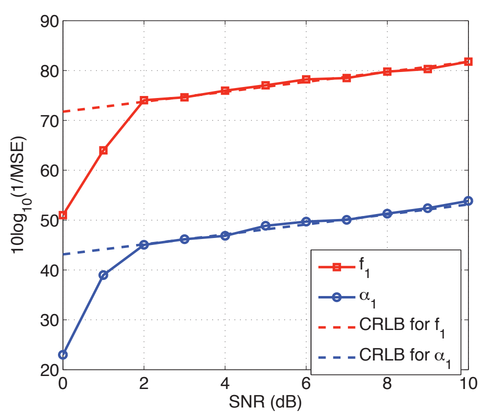

Fig. 3. MSE and CRB as functions of SNR.

6. CONCLUSION

The most critical issue for estimation of the posterior distribution of chirp signals is the multimodality. In this paper, we propose the Gibbs MPMC algorithm to estimate the parameters of multiple chirp signals. The marginalization of linear parameters lowers the computational cost by only generating particles of the nonlinear parameters, namely the frequencies and the chirp rates. We also propose that the generating functions of the particles be alternating conditionals, as in the case of Gibbs sampling. Thus, the overall proposal function is a product of conditionals, where the sampling from each conditional is easy. The method is tested with simulations and the obtained results show good performance.

7. REFERENCES

[1] T. J. Abatzoglou, "Fast maximum likelihood joint estimation of frequency and frequency rate," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-13, pp. 708–715, 1986.

[2] P. M. Djurić and S. M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. 38, pp. 2118–2126, 1990.

[3] O. Besson, M. Ghogho, and A. Swami, "Parameter estimation for random amplitude chirp signals," *IEEE Trans. on Signal Processing*, vol. 47, pp. 3208–3219, 1999.

[4] K. S. Arun R. M. Liang, "Parameter estimation for superimposed chirp signals," in *the Proceedings of ICASSP*, San Francisco, CA, USA, 1992, vol. 5, pp. 273–276.

[5] S. Saha and S. Kay, "Maximum likelihood parameter estimation of superimposed chirp signals using Monte Carlo importance sampling," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 224–230, 2002.

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

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

[8] P. M. Djurić, B. Shen, and M. F. Bugallo, "Population Monte Carlo methodology a la Gibbs sampling," in *the Proceedings of EUSIPCO*, Barcelona, Spain, 2011.

[9] 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. 420–448, 2007.

[10] B. Shen, M. F. Bugallo, and P. M. Djurić, "Multiple marginalized population Monte Carlo," in *the Proceedings of EUSIPCO*, Aalborg, Denmark, 2010.

[11] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[12] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, New York, 1995.

[13] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, 2003.

[14] M. Bugallo, M. Hong, and P. M. Djurić, "Marginalized population Monte Carlo," in *the Proceedings of ICASSP*, Taipei, Taiwan, 2009.