ON THE IMPLEMENTATION OF PARTICLE FILTERS FOR DOA TRACKING

William Ng, Jean-René Larocque and James P. Reilly

Communications Research Laboratory, McMaster University
Hamilton, Ontario, Canada
larocque@reverb.crl.mcmaster.ca,reillyj@mcmaster.ca

ABSTRACT

This paper addresses practical issues for the implementation of sequential Monte Carlo sampling schemes, also known as particle filtering, for application to tracking problems.

The discussion focusses on ways to improve on previous resampling schemes, resulting in significantly improved performance.

These conclusions are demonstrated and supported by examples of application of the particle filter to a sequential tracking of a known number of directions of arrival.

1. INTRODUCTION

Many problems can be modelled in the state-space paradigm. The hidden parameters of interest evolve in time following the update equation while the observations are functions of these parameters. The objective is to sequentially estimate the hidden parameters, based on the observations. If the system is linear and Gaussian, the optimal filter is the Kalman filter. However, many real problems are neither linear nor Gaussian.

The sequential Monte Carlo sampling methods, also called particle filters, offer promising new approaches to this difficult real-life case. Ground breaking work in the area is presented in [1][2].

In this paper, we offer a modification of the method in [1], by proposing a Metropolis-Hasting resampling [3][4] of the parameters, using an improved proposal distribution. We verify our method by appliyng a particle filter to the sequential tracking of the directions of arrival on targets of interest from an array of sensors.

The paper is organised as follows. Section II presents the state-space model. Section III emphasises the proposed improvements to the algorithm. Results from simulations are presented in Section IV, with conclusions in Section VI.

2. THE STATE-SPACE MODEL

The state-space model for the sequential tracking problem considered is given by:

$$\boldsymbol{x}(t) = \boldsymbol{x}(t-1) + \sigma_v \boldsymbol{v}(t) \quad (1)$$

$$\boldsymbol{y}(t) = \boldsymbol{S}(\boldsymbol{x}(t))\boldsymbol{a}(t) + \sigma_w \boldsymbol{w}(t), \quad (2)$$

where the noise variables $\boldsymbol{v}(t)$ and $\boldsymbol{w}(t)$ are *iid* Gaussian variables with zero mean and unit variance, independant of the parameters. The matrix $\boldsymbol{S}(\boldsymbol{x}) \in \mathcal{C}^{M\times k}$ is the usual steering matrix. In the proposed system of equations, the noise variances σ_v and σ_w are assumed unknown but constant(for a stationary system). The unknown parameter of amplitudes $\boldsymbol{a}(t)$ are allowed to change slowly between snapshots. The directions of arrival, $\boldsymbol{x}(t)$, are to be sequentially estimated based on the observations $\boldsymbol{y}(t)$.

Similar notation to that of [1] is adopted. Let's define the vector of parameters:

$$\boldsymbol{\theta}_{1:t} \triangleq (\{\boldsymbol{a}\}_1^t, \sigma_v, \sigma_w, \boldsymbol{\delta}^2) \quad (3)$$

and the whole parameter space as:

$$\boldsymbol{\alpha}_{0:t} \triangleq (\boldsymbol{x}_{0:t}, \boldsymbol{\theta}_{1:t}). \quad (4)$$

The parameter k, which is the dimension of $\boldsymbol{x}$, is assumed known. This aspect of the algorithm (joint detection of the number of sources) will constitute the focus of another paper.

The joint distribution of all the parameters is:

$$p(\boldsymbol{\alpha}_{0:t}, \boldsymbol{y}_{1:t}) \propto p(\boldsymbol{y}_{1:t}|\boldsymbol{\alpha}_{0:t})p(\boldsymbol{\alpha}_{0:t}) \quad (5)$$

$$\propto p(\boldsymbol{y}_{1:t}|\boldsymbol{x}_{0:t})p(\boldsymbol{x}_{0:t}|\boldsymbol{\theta}_{1:t})p(\boldsymbol{\theta}_{1:t}). \quad (6)$$

It is assumed that the observations, given the states, are *iid* and that the state conditional update likelihood is also *iid*. Therefore, assuming the distribution of the initial states to be uniform, and using the Markov properties of the model, (6) can be written in the form:

$$p(\boldsymbol{\alpha}_{0:t}, \boldsymbol{y}_{0:t}) \propto \prod_{l=1}^{t} p(\boldsymbol{y}(l)|\boldsymbol{x}(l), \boldsymbol{\theta}_{1:t}) p(\boldsymbol{\theta}_{1:t}) \times \prod_{l=1}^{t} p(\boldsymbol{x}(l)|\boldsymbol{x}(l-1), \boldsymbol{\theta}_{1:t}). \tag{7}$$

To complete the model, prior distributions of the parameters are required. These prior distributions are chosen as non-informative as possible. When convenient, the conjugate forms were favoured.

- The amplitudes are chosen *iid* with different variances (to accomodate for different signal amplitudes and different SNR):

$$p(\boldsymbol{a}_{1:t}|\sigma_w^2, \boldsymbol{\delta}^2) \sim \prod_{1}^{t} \mathcal{N}(\mathbf{0}, \sigma_w^2 \Delta) \tag{8}$$

  where $\Delta = \operatorname{diag}(\delta_1^2, \ldots, \delta_k^2)$ and the δ_i^2 are hyperparameters to be determined.

- The prior distribution on the noise variances are all assumed to follow the inverse Gamma distribution, which is the conjugate distribution:

$$p(\sigma_v^2) \sim \mathcal{IG}(\frac{v_0}{2}, \frac{\gamma_0}{2}) \tag{9}$$

$$p(\sigma_w^2) \sim \mathcal{IG}(v_1, \gamma_1) \tag{10}$$

$$p(\boldsymbol{\delta}^2) \sim \prod_{i=1}^{k} \mathcal{IG}(n_0, g_0) \tag{11}$$

With these considerations, the generic algorithms described in [5][6][7] can be used almost as described. The sequential importance sampling step and the selection step in these cases work well. Here, we focus attention on the next step, i.e. the MCMC resampling step, to recreate diversity amongst the particles.

It is well understood that only a hand-full of particles at best will have meaningful associated weights. Therefore, any estimate based on these very few particles would show a large variance. The easiest remedy is to multiply/suppress the particles according to their importance weights. However, this does not add diversity. Some papers simply suggest to add a perturbation to the child particles. But a more clever approach, proposed by Andrieu and Doucet[1][7], uses an MCMC step on each particle.

3. THE MCMC DIVERSITY STEP

The selection procedure based on the importance weights described earlier greatly diminishes the diversity of the particles. A strategy to rejuvenate the diversity relies on the use of a MCMC step. This approach is very interesting in the potential for the application of the reversible jump algorithm [4]. Using an RJMCMC step would allow the algorithm to jointly estimate the model order k, but that is out of the scope of this discussion.

The MCMC step, as described in [1] has an invariant distribution $\prod_{i=1}^{N} p(\boldsymbol{\alpha}_{0:t}|\boldsymbol{y}_{1:t})$, which is applied to each of the N particles, one at the time. In the original paper, all the parameters are re-sampled using a Gibbs sampler. To re-sample the state parameter $\boldsymbol{x}^{(i)}(t)$, the proposal distribution first suggested is the importance function used at the sampling step. This choice slows down the acquisition of the tracks because it relies too much on the current estimates of all the other parameters. Early in the tracking, when the parameters are not estimated accurately, this proposal distribution does not offer the flexibility to explore the whole parameter space and therefore slows down considerably the acquisition of the track. Also, this situation creates a vicious circle in which the algorithm has great difficulty at breaking: if the particles have small variances in $\boldsymbol{x}$, the estimated σ_v will be small, which in turn will prevent the particles, at the next iteration, to adequately explore the parameter space.

This paper proposes the use of a Metropolis-Hasting one-at-the-time scheme with a less informative candidate distribution that will show more variance in its candidates. This allows the algorithm to pick up on any fast variation of the states, or to acquire a track faster, by allowing potentially larger jumps. The proposal distribution is:

$$\boldsymbol{x}_c^{(i)}(t) \sim \mathcal{N}(\boldsymbol{x}^{(i)}(t-1), \sigma_v^{(i)}(t)). \tag{12}$$

The candidate $\boldsymbol{x}_c^{(i)}(t)$ is accepted with probability α_x, defined as $\alpha_x = \min(r_x, 1)$, with:

$$r_x = \frac{e^{(\boldsymbol{y}(t)-\boldsymbol{S}(\boldsymbol{x}^{(i)}(t))\boldsymbol{a}^{(i)}(t))^H \boldsymbol{D}_w^{-1}(\boldsymbol{y}(t)-\boldsymbol{S}(\boldsymbol{x}^{(i)}(t))\boldsymbol{a}^{(i)}(t)))}}{e^{(\boldsymbol{y}(t)-\boldsymbol{S}(\boldsymbol{x}_c^{(i)}(t))\boldsymbol{a}^{(i)}(t))^H \boldsymbol{D}_w^{-1}(\boldsymbol{y}(t)-\boldsymbol{S}(\boldsymbol{x}_c^{(i)}(t))\boldsymbol{a}^{(i)}(t)))}},$$

with the covariance matrix $\boldsymbol{D}_w \triangleq \sigma_w^{2^{(i)}}(t)\boldsymbol{I}_k$. Even though this proposal distribution is suboptimal by being independent of the observations, it has proved to be the most effective approach in practice.

Secondly, it cannot be assumed that the amplitudes $\boldsymbol{a}(t)$ have the same variances. Therefore, forcing Δ in (8) to have constant diagonal elements, as in previous work, leads to poorer results if the signals have different SNRs: the tracks with high SNRs show very tight tracking, while the others exhibit more variance in the error. It seems more appropriate to attribute a different prior variance for each amplitude as in (8).

In doing so, the prior distribution on the amplitude remains uninformative, in the sense that it does not assume the same variance. In summary, the MCMC step for regenerating the diversity of the particles includes the following proposal distribution for the M-H algorithms.

- The parameters of interest:

$$\boldsymbol{x}_c^{(i)}(t) \sim \mathcal{N}(\boldsymbol{x}^{(i)}(t-1), \sigma_v^{2^{(i)}}(t)) \tag{13}$$

- The parameter of amplitudes:

$$\boldsymbol{a}_c^{(i)}(t) \sim \mathcal{N}(\boldsymbol{m}_a^{(i)}, \sigma_w^{2^{(i)}} \Sigma_a^{(i)}) \tag{14}$$

with $\Sigma_a^{-1^{(i)}} = \boldsymbol{S}^H(\boldsymbol{x}^{(i)}(t))\boldsymbol{S}(\boldsymbol{x}^{(i)}(t)) + \Delta^{(i)^{-1}}$ and $\boldsymbol{m}_a^{(i)} = \Sigma_a^{(i)} \boldsymbol{S}^H(\boldsymbol{x}^{(i)}(t))\boldsymbol{y}(t)$.

- The update noise variance:

$$\sigma_{v_c}^{2^{(i)}} \sim \mathcal{IG}(\frac{v_o + kt}{2}, \frac{\gamma_0}{2} + \beta_v^{(i)}) \tag{15}$$

with $\beta_v^{(i)} = \frac{\sum_{l=1}^{t}(\boldsymbol{x}^{(i)}(l) - \boldsymbol{x}^{(i)}(l-1))'(\boldsymbol{x}^{(i)}(l) - \boldsymbol{x}^{(i)}(l-1))}{2}$

- The observation noise variance:

$$\sigma_{w_c}^{2^{(i)}} \sim \mathcal{IG}(v_1 + (k+M)t, \gamma_1 + \beta_w^{(i)}) \tag{16}$$

with

$$\beta_w^{(i)} = \sum_{l=1}^{t}(\boldsymbol{y}(t) - \boldsymbol{S}(\boldsymbol{x}_t^{(i)})\boldsymbol{a}_t^{(i)})'(\boldsymbol{y}(t) - \boldsymbol{S}(\boldsymbol{x}_t^{(i)})\boldsymbol{a}_t^{(i)})$$

- The hyper-parameters δ_k:

$$\delta_{k_c}^{2^{(i)}} \sim \mathcal{IG}(k + n_0, g_0 + \frac{\boldsymbol{a}^{H^{(i)}}(t)\boldsymbol{a}^{(i)}(t)}{\sigma_w^{2^{(i)}}}) \tag{17}$$

4. SIMULATIONS RESULTS

The proposed algorithm is now applied to simulation data, generated for $k_o = 2$ sources with the parameters described in table 4.1. The received array is composed of 8 elements. The amplitudes follow an 4^{th} order AR

Parameter	σ_v^2	σ_w^2	$\boldsymbol{x}(0)$	σ_a^2
Value	1 deg.	0.3	$[50^\circ, 100^\circ]$	0.3

Table 4.1: Parameters of the state-space model for simulated data

process with coefficients $[0.1, 0.2, -0.3, 0.5]$, and noise variance σ_a^2, with an initial SNR of 20dB. The directions of arrival follow a random walk, with noise variance σ_v^2.

The particle filter uses 300 particles only (as opposed to one hundred times that number as suggested in previous works) and is run for 500 observations. Figure 1, 2 and 3 show the results.

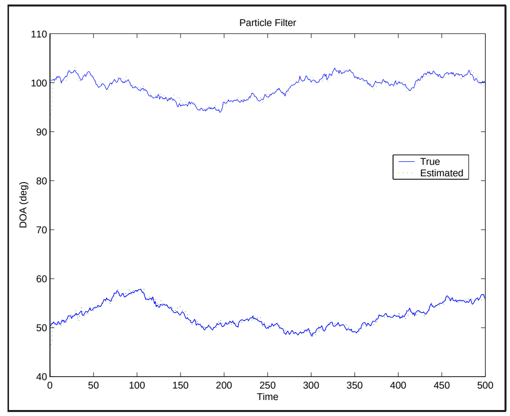

Figure 1: Sequential estimates of the directions of arrival

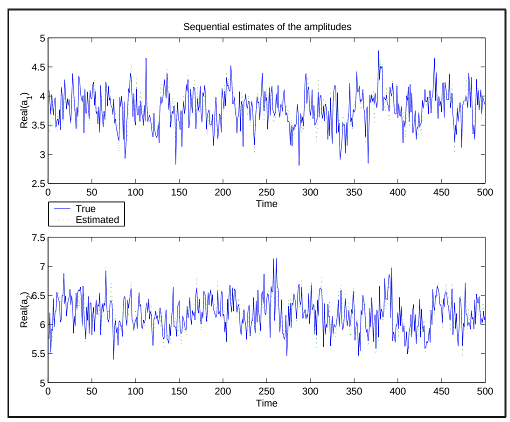

Figure 2: Sequential estimates of amplitudes

Figure 1 clearly shows that the directions of arrival are well traced by their estimates throughout the entire tracking process. Similar findings can be seen in Figure 2, which shows that the trajectories of the real part of the amplitude paramter. The imaginary part is not shown, but is also correctly estimated. At the same time, the noise variances converge slowly to the correct values.

These results use the suggested proposal function and k_0 different distributions when sampling the vector $\boldsymbol{\delta}_t^{2^{(i)}}$, each distribution of which depends on $\boldsymbol{a}_{k,t}^{(i)}$. This

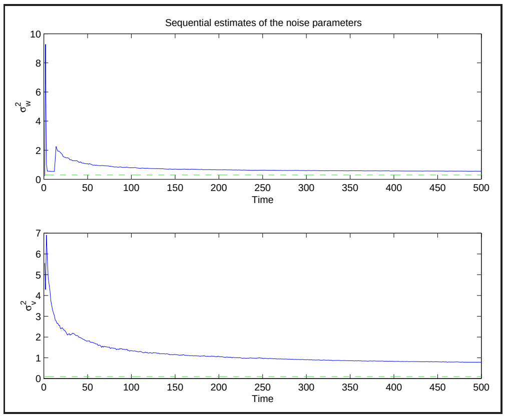

Figure 3: Sequential estimates of variances of the noises (dashed line is the true value)

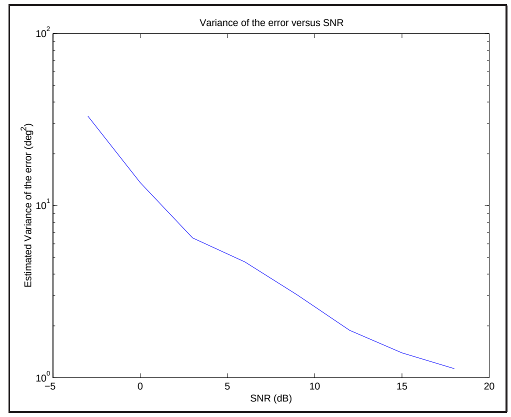

Figure 4: Variance of the estimates vs SNR

approach provides better candidates to $\boldsymbol{\delta}_{k,t}^{2^{(i)}}$ and hence to all the interdependent parameters, especially when the amplitudes differ a lot.

4.1. Performance of the tracking

In this subsection, the proposed algorithm was applied to 50 different scenarios of 50 observations, for different values of SNR, in order to estimate the variance of the estimate as a function of the SNR. For each run, the first 25 sequential estimates (considered in the acquisition mode) were discarded and the following 25 estimates were used to get one sample of the variance. Table 4.2 summurizes the parameters of the simulations and figure 4 shows the results.

Parameter	σ_v^2	σ_a^2	$\boldsymbol{x}(0)$	$\boldsymbol{a}(1)$	SNR
Value	5 deg.	0.005	[90°]	[4]	0-18dB

Table 4.2: Parameters of the state-space model for Monte Carlo estimation

5. CONCLUSION

In this paper, suggestions were made to improve a previously proposed method for recreating the diversity of the particles with an MCMC step. The new filter requires fewer particles to achieve the same performance and converges faster. It is also less sensitive the to value of the other nuisance parameters.

Such improvements are explained by the use of a less informative candidate function in the MCMC move for diversity, which allow the particles to explore the space more freely.

6. ACKNOWLEDGEMENTS

The authors are greatful to Dr. Christophe Andrieu for the helpful discussions. The authors wish to acknowledge and thank the following institutions for their support of this project: Mitel Corporation, Kanata, Ontario, Canada; The Centre for Information Technology Ontario (CITO); and the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] C. Andrieu, N. D. Freitas, and A. Doucet, "Sequential MCMC for bayesian model selection," *Proceedings of HOS Conference*, 1999.

[2] P. Djuric, "Sequential estimation of random parameters under model uncertainty," *Proceedings ICASSP*, 2000.

[3] W. Hasting, "Monte Carlo sampling methods using Markov Chains and their applications," *Biometrika*, vol. 57, no. 1, p. 97, 1970.

[4] P. Green, "Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[5] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer, 1999.

[6] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in practice*. Chapman and Hall, 1998.

[7] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.