On MCMC-Based Particle Methods for Bayesian Filtering: Application to Multitarget Tracking
François Septier, Sze Kim Pang, Avishy Carmi, Simon Godsill

To cite this version:
On MCMC-Based Particle Methods for Bayesian Filtering: Application to Multitarget Tracking

François Septier, Sze Kim Pang, Avishy Carmi and Simon Godsill
Signal Processing and Communications Laboratory
Cambridge University, UK

Abstract—Nonlinear non-Gaussian state-space models arise in numerous applications in control and signal processing. In this context, one of the most successful and popular approximation techniques is Sequential Monte Carlo (SMC) methods, also known as particle filters. Nevertheless, these methods tend to be inefficient when applied to high dimensional problems. In this paper, we present an overview of Markov chain Monte Carlo (MCMC) methods for sequential simulation from posterior distributions, which represent efficient alternatives to SMC methods. Then, we describe an implementation of this MCMC-based particle algorithm to perform the sequential inference for multitarget tracking. Numerical simulations illustrate the ability of this algorithm to detect and track multiple targets in a highly cluttered environment.

I. INTRODUCTION

In many applications, we are interested in estimating a signal from a sequence of noisy observations. This problem can generally be stated in a state space form as follows. A transition equation describes the prior distribution of a hidden Markov process \( \{x_k; k \in \mathbb{N}\} \), \( x_k \in \mathbb{R}^n \) and an observation equation describes the likelihood of the observations \( \{z_k; k \in \mathbb{N}\} \), \( z_k \in \mathbb{R}^m \). Within a Bayesian framework, all relevant information about \( x_k \) given observations up to and including time \( k \) can be obtained from the filtering distribution \( p(x_k|z_{0:k}) \).

Except in a few special cases, including linear and Gaussian state space models (Kalman filter) and hidden finite-state space Markov chains, it is impossible to evaluate this filtering distribution analytically. Since the nineties, sequential Monte Carlo (SMC) approaches have become a powerful methodology to cope with non-linear and non-Gaussian problems [1]. These methods, also known as particle filters (PF), exploit numerical representation techniques for approximating the filtering probability density function of inherently nonlinear non-Gaussian systems. Using these methods, the obtained estimates can be set arbitrarily close to the optimal solution (in the Bayesian sense) at the expense of computational complexity. Due to their sampling mechanization, PFs tend to be inefficient when applied to high-dimensional problems such as multi-target tracking.

Markov chain Monte Carlo (MCMC) methods are generally more effective than PFs in high-dimensional spaces. Their traditional formulation, however, allows sampling from probability distributions in a non-sequential fashion. Recently, sequential MCMC schemes were proposed by [2]–[6]. These approaches are distinct from the Resample-Move scheme [7] in particle filter where the MCMC algorithm is used to rejuvenate degenerate samples following importance sampling resampling. These methods [2]–[6] use neither resampling nor importance sampling.

In this paper, we will review existing approaches that use MCMC methods in a sequential setting in order to approximate the filtering distribution. Then, we will apply this method for multitarget tracking. The purpose of multiple object tracking algorithms is to detect, track and identify targets from sequences of noisy observations provided by one or more sensors. The difficulty of this problem has increased as sensor systems in the modern battlefield are required to detect and track objects in very low probability of detection and in hostile environments with heavy clutter.

The paper is organized as follows. In Section II, existing sequential MCMC approaches are reviewed and discussed. Section III describes how such algorithms can be efficiently employed for the challenging multitarget tracking problem. Numerical results are shown in Section IV. Conclusions are given in Section V.

II. MCMC-BASED PARTICLE METHODS

In this section, we will describe the sequential approaches using MCMC proposed in [2]–[6].

A. Targeting the Filtering Distribution

In a Bayesian framework, we are aimed at computing the filtering posterior distribution \( p(x_k|z_{0:k}) \) recursively by

\[
p(x_k|z_{0:k}) \propto \int p(z_k|x_k)p(x_k|x_{k-1})p(x_{k-1}|z_{0:k-1})dx_{k-1}
\]

\( \text{(1)} \)

Unfortunately in many applications, this distribution is analytically intractable. In [3], the authors have proposed to use a MCMC in a sequential setting in order to obtain an approximation of this filtering distribution. This is achieved by using a set of unweighted particles to represent the density \( p(x_{k-1}|z_{0:k-1}) \):

\[
p(x_{k-1}|z_{0:k-1}) \approx \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(x_{k-1} - x_{k-1}^{(j)})
\]

\( \text{(2)} \)

where \( N_p \) is the number of particles and \( (j) \) the particle index. Then, by plugging this particle approximation into Eq. (1),

\[
p(x_k|z_{0:k}) \approx \frac{1}{N_p} p(z_k|x_k) \sum_{j=1}^{N_p} p(x_k|x_{k-1}^{(j)})
\]

\( \text{(3)} \)
A MCMC procedure is designed using (3) as the target distribution with a proposal distribution of \( q(x_k|x_k^0) \). Then, by using the appropriate acceptance ratio, the desired approximation \( \tilde{p}(x_k|z_{0:k}) \) is obtained by storing every \( N_{\text{thin}} \)th accepted sample (thinning) after the initial burn-in (\( N_{\text{burn}} \)) iterations. Unfortunately, the computational demand of this MCMC-based particle filter by Khan et al. can become excessive as the number of particles increases owing to its direct Monte Carlo computation of the predictive density at each time step (in general \( O(N_p) \) per MCMC iteration). This algorithm, summarized in Algorithm 1, will be denoted below by Marginalized MCMC-Based Particle Algorithm.

Algorithm 1 Marginalized MCMC-Based Particle Algorithm

1: Initialize particle set \( \{x^{(j)}_k\}^{N_p}_{j=1} \)
2: for \( k = 1, \ldots, T \) do
3: for \( m = 1, \ldots, N_{\text{MCMLC}} \) do
4: Propose \( x^*_k \) \( \sim q(x_k|x_{k-1}^m) \)
5: Compute the MH acceptance probability \( p(x_k^m, x_k^*) = \min \left( 1, \frac{p(x_k^m|x_{k-1}^m)p(z_{0:k}|x_k^m)}{p(x_k^*|x_{k-1}^m)p(z_{0:k}|x_k^*)} \right) \)
6: Accept \( x_k^m = x_k^* \) with probability \( p(x_k^m, x_k^*) \)
7: After a burn in period of \( N_{\text{burn}} \), keep every \( N_{\text{thin}} \) MCMC output \( x_k^{(j)} \) as the new particle set for approximating \( p(x_k|z_{0:k}) \), i.e. \( \tilde{p}(x_k|z_{0:k}) = \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(x_k - x_k^{(j)}) \)
8: end for
9: end for

B. Targeting the Joint Posterior Distribution

To avoid numerical integration of the predictive density at every MCMC iteration, an alternative algorithm has been developed in [5]. This algorithm can be considered as a generalisation of the features of some existing sequential MCMC algorithms, which allow for further extensions and insights into the formulation. A special case of it has been used by [2] in an early paper on sequential inference using MCMC. It was applied to a dynamical model with expanding parameter space. It has also been used by [4] which was applied to imputing missing data from nonlinear diffusions.

The MCMC-Based Particle Algorithm in [5] considers the general joint posterior distribution of \( x_k \) and \( x_{k-1} \):

\[
p(x_k, x_{k-1}|z_{0:k}) \propto p(z_k|x_{k})p(x_k|x_{k-1})p(x_{k-1}|z_{0:k-1})
\]

(4)

A MCMC procedure will then be used to make inference from this complex distribution, i.e. Eq. (4) is the target distribution. Clearly, there will not be a closed form representation of the posterior distribution \( p(x_{k-1}|z_{0:k-1}) \) at time \( k-1 \). Instead, like in the previous algorithm, it will be approximated with an empirical distribution based on the current particle set, Eq. (2).

Then, having made many joint draws from Eq. (4) using an appropriate MCMC scheme, the converged MCMC output for variable \( x_k \) can be extracted to give an updated marginalized particle approximation to \( p(x_k|z_{0:k}) \). In this way, sequential inference can be achieved.

More precisely, at the \( m^{th} \) MCMC iteration, the procedure, to obtain samples from \( p(x_k, x_{k-1}|z_{0:k}) \), involves a joint Metropolis-Hastings (MH) proposal step where both \( x_k \) and \( x_{k-1} \) are updated jointly, as well as individual refinement Metropolis-within-Gibbs steps where \( x_k \) and \( x_{k-1} \) are updated individually. This algorithm, summarized in Algorithm 2, will be denoted by MCMC-Based Particle Algorithm.

As a comparison, Berzunzi et al. [2] made use of the individual refinement step to move the current state \( x_k \) as well as the particle representation \( x_{k-1} \). This can potentially lead to poor mixing in high-dimensional problems due to the highly-disjoint predictive density of the particle representation. On the other hand, Gollighot and Wilkinson [4] made use of only the joint draw to move the MCMC chain. This can potentially reduce the effectiveness of the MCMC as refinement moves are not employed to explore the structured probabilistic space.

In [6], Septier et al. have proposed to incorporate, into the framework of this MCMC-based particle scheme, several attractive features of genetic algorithms and simulated annealing. We refer the reader to that paper for details.

Algorithm 2 MCMC-Based Particle Algorithm

1: Initialize particle set \( \{x^{(j)}_{k-1}\}^{N_p}_{j=1} \)
2: for \( k = 1, \ldots, T \) do
3: for \( m = 1, \ldots, N_{\text{MCMLC}} \) do
4: Joint Draw
5: Propose \( \{x^*_k, x^*_{k-1}\} \sim q_k(x_k, x_{k-1}|x_{k-1}^{m-1}, x_{k-1}^{m-1}) \)
6: Compute the MH acceptance probability \( \rho_1 = \min \left( 1, \frac{q_k(x_k^*, x_{k-1}^*|x_{k-1}^{m-1}, x_{k-1}^{m-1})}{q_k(x_k^{m-1}, x_{k-1}^{m-1}|x_{k-1}^*, x_{k-1}^*)} \right) \)
7: Accept \( \{x^*_k, x^*_{k-1}\} = \{x^*_k, x^*_{k-1}\} \) with probability \( \rho_1 \)
8: Refinement
9: Propose \( \{x^*_{k-1}\} \sim q_{k-1}(x_{k-1}^{m-1}|x_{k-1}^{m-1}) \)
10: Compute the MH acceptance probability \( \rho_2 = \min \left( 1, \frac{q_{k-1}(x_{k-1}^{m-1}|x_{k-1}^{m-1})}{q_{k-1}(x_{k-1}^{m-1}|x_{k-1}^{m-1})} \right) \)
11: Accept \( \{x^*_{k-1}\} = \{x^*_{k-1}\} \) with probability \( \rho_2 \)
12: Propose \( \{x^*_k\} \sim q_k(x_k|x_{k-1}^{m-1}) \)
13: Compute the MH acceptance probability \( \rho_3 = \min \left( 1, \frac{q_k(x_k^{m-1}|x_k^{m-1})}{q_k(x_k^{m-1}|x_k^{m-1})} \right) \)
14: Accept \( \{x^*_k\} = \{x^*_k\} \) with probability \( \rho_3 \)
15: After a burn in period of \( N_{\text{burn}} \), keep every \( N_{\text{thin}} \) MCMC output \( x_k^{(j)} \) as the new particle set for approximating \( p(x_k|z_{0:k}) \), i.e. \( \tilde{p}(x_k|z_{0:k}) = \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(x_k - x_k^{(j)}) \)
16: end for
17: end for

III. APPLICATION TO MULTITARGET TRACKING

A. Problem Formulation

In this study, we consider a time-varying number of targets. As a consequence, the target states are variable dimension quantities since targets can appear or disappear from the scene randomly over time. In order to model this birth and death process, we choose to use a set of existence variables \( e_k \) with elements \( e_{k,i} \in \{0,1\} \) model the birth and death process for each individual target. In this formulation, the targets’ kinematic vector is thus regarded as fixed dimensional quantity with \( N_{\text{max}} \) targets, each of which being active or inactive according to its existence variable \( e_{k,i} \).
In this application, the aim is thus to compute, at time $t_k$, the filtering posterior distribution $p(x_k, e_k | z_{0:k})$ where $x_k = [x_{k,1} \cdots x_{k,N_{max}}, y_{k,1} \cdots y_{k,N_{max}}, \dot{x}_{k,1} \cdots \dot{x}_{k,N_{max}}, y_{k,1} \cdots y_{k,N_{max}}]^T$, $e_k = [e_{k,1} \cdots e_{k,N_{max}}]^T$, and $z_{0:k}$ are respectively the targets’ kinematics vector, the existence vector and the observation set from time $t_0$ to $t_k$.

In this work, we consider that the targets evolve independently of one another and the existence variable and the targets’ kinematics are independent, so the transition probability distribution can be expanded as follows:

$$p(x_k, e_k | x_{k-1}, e_{k-1}) = \prod_{n=1}^{N_{max}} p(x_{k,n} | x_{k-1,n}, e_{k,n}, e_{k-1,n})p(e_{k,n} | e_{k-1,n})$$

(5)

We now describe the various densities involved in the computation of the filtering posterior distribution.

1) The prior distribution of the existence variables: Each target’s existence variable will be modeled as a discrete Markov chain [8] which is independent of all other states. In this paper, the birth process is modeled as a Bernoulli like the death process, i.e.:  

$$p(e_{k,n} | e_{k-1,n}) = \delta(e_{k,n} - 1)[1 - P_D]\delta(e_{k-1,n} - 1) + P_B \delta(e_{k,n} - 1) + \delta(e_{k,n})[(1 - P_B)\delta(e_{k-1,n} - 1)$$

(6)

where $P_B$ and $P_D$ are probability values for a target to become respectively active (“alive”) or inactive (“dead”).

2) The transition probability of the targets: The transition probability of the $n^{th}$ target can be partitioned according to $e_{k,n}$ and $e_{k-1,n}$ as follows:

$$p(x_{k,n} | x_{k-1,n}, e_{k-1,n}) = \begin{cases} p_b(x_{k,n}) & \text{if } \{e_{k,n}, e_{k-1,n}\} = \{1, 0\} \\ p_d(x_{k,n}) & \text{if } e_{k,n} = 0 \\ p_a(x_{k,n} | x_{k-1,n}) & \text{if } \{e_{k,n}, e_{k-1,n}\} = \{1, 1\} \end{cases}$$

(7)

Target Birth - The target can appear anywhere uniformly in the surveillance area of $L_x$ by $L_y$. The speed is also uniformly distributed between $-V_{max}$ and $V_{max}$ i.e.

$$p_b(x_{k,n}) = U(x_{k,n} | 0, L_x)U(\dot{x}_{k,n} | -V_{max}, V_{max})$$

$$U(y_{k,n} | 0, L_y)U(\dot{y}_{k,n} | -V_{max}, V_{max})$$

(8)

where $U(\cdot | a, b)$ is the continuous uniform distribution for the interval $[a,b]$.

Target Death - For an inactive target, i.e. $e_{k,i} = 0$, we will keep the target state at some $x_{\text{death}}$, which is the state where an inactive target is represented.

$$p_d(x_{k,n}) = \delta(x_{\text{death}})$$

(9)

Target Update - This case corresponds to active targets that will be updated according to the near constant velocity model

$$p_u(x_{k,n} | x_{k-1,n}) = N(x_{k,n} | A_{k,n}x_{k-1,n}, Q_{k,n})$$

(10)

where the matrices $A_{k,n}$ and $Q_{k,n}$ are defined as follows:

$$A_{k,n} = \begin{bmatrix} I_2 & \tau_k I_2 \\ 0_2 & I_2 \end{bmatrix}, Q_{k,n} = \sigma^2_{x,n} \begin{bmatrix} (\tau_k^2 / 3)I_2 & (\tau_k^2 / 2)I_2 \\ (\tau_k^2 / 2)I_2 & \tau_k I_2 \end{bmatrix}$$

(11)

with $\tau_k = t_k - t_{k-1}$.

3) Association Free Observation Model: An association free observation is considered in this study. At each time step $t_k$, a set or frame of sensor measurements $z_k = \{z_k^1, ..., z_k^{M_k}\}$ is received from a sensor scanning within an observation space, where $M_k$ is the number of measurements (both target and clutter) returned by the sensor. Since any element of $z_k$ may originate from a true target or from clutter, a Poisson process is adopted to model the observations falling in a specific region. The number of the $n^{th}$ target measurements is randomly generated from a Poisson distribution having mean $\Lambda_n$ whereas the number of clutter measurements have a mean number $\Lambda_C$ [9]. Accordingly, the likelihood function of the observations can be expressed as

$$p(z_k | x_k) = \frac{e^{-\mu_k} \Lambda_k}{M_k!} \prod_{m=1}^{M_k} \lambda(\tilde{z}_k^m)$$

(12)

where $\mu_k = \Lambda_C + \sum_{n=1}^{N_{T,k}} \Lambda_n$ is the expected total number of measurements received at time $t_k$ and

$$\lambda(\tilde{z}_k^m) = \sum_{n=1}^{N_{T,k}} \Lambda_n p_z(\tilde{z}_k^m | x_k) + \Lambda_C p_C(\tilde{z}_k^m)$$

(13)

with $p_z(\cdot)$ and $p_C(\cdot)$ being the likelihood functions of target and clutter measurements and $N_{T,k}$ the number of targets at time $t_k$. For the measurements issuing from the $n^{th}$ target, $\tilde{z}_k^m = [z_{k,x}^m, z_{k,y}^m]^T$ is considered to be drawn from $p_z(\tilde{z}_k^m | x_k) = N(\tilde{z}_k^m | x_k, \Sigma_x)$. The clutter measurements are drawn uniformly in the surveillance region, i.e. $p_C(\tilde{z}_k^m) = U(z_{k,x}^m | 0, L_x)U(z_{k,y}^m | 0, L_y)$.

B. Application of MCMC-Based Particle Algorithm

In this section, we will describe one possible implementation of the MCMC-based Particle approach (Algorithm 1) for this problem. For the joint draw of $\{x_k, e_k, x_{k-1}, e_{k-1}\}$, the following proposal distribution is used:

$$q_1(x_k, e_k, x_{k-1}, e_{k-1} | x_{k-1}^m, e_{k-1}^m, x_k^m, e_k^m) \propto \sum_{j=1}^{N_{T,k}} p(x_k | x_{k-1}^j, e_k)p(e_k | e_{k-1}^j)\delta(x_{k}^{(j)})\delta(e_{k}^{(j)})$$

(14)

Then, in refinement step, we sample successively each of the individual targets and their associated existence variables $\{x_{k,n}, e_{k,n}\}_{n=1}^{N_{max}}$ by using this proposal distribution:

$$q_3(x_{k,n}, e_{k,n} | x_{k-1}^m, e_{k-1}^m, x_{k-1}^m, e_{k-1}^m) = p(x_{k,n} | x_{k-1}^m, e_{k,n} | e_{k-1}^m)$$

(15)

In this implementation, $\{x_{k-1}, e_{k-1}\}$ are not sampled in the refinement step, so the proposal distribution $q_2(\cdot)$ in Algorithm 1 is not defined.

IV. NUMERICAL SIMULATIONS

Consider a time-varying number of targets (max of 5) moving using the near-constant velocity model defined in Eq. (10) as shown in Fig. 1. Note that targets 1-3 are born at time $k = 1$, targets 4-5 are born at time $k = 25$, and target 1 dies at time $k = 50$. The starting and stopping positions for each track are labelled with a circle and triangle respectively.
Individual target motions follow the near constant velocity motion model with a sampling period of $\tau_k = 3s$ and process noise variance $\sigma^2_{x,n} = 0.5$. The observation scene is a square of $5000m \times 5000m$. The parameters involved in the observation model are $\Lambda_0^n = 1$, $\Lambda_C = 20$, and $\Sigma_x = 100 \times I_2$. The MCMC-Based particle scheme is implemented using $N_p = 4000$ particles, a burn-in period of $N_{burn} = 1000$ and a thinning value of 6.

The estimated tracks for a single run are shown in Fig. 2. The MCMC-Based particle algorithm has successfully detected and tracked the multiple targets in a hostile environment with heavy clutter. Fig. 3 shows the mean and standard deviation of the estimated cardinality distribution. From this figure, we can see that the proposed algorithm is capable of adequately tracking the varying number of targets in the observation scene.

![Fig. 1. True tracks in the $xy$ plane. Targets move with near-constant velocity along the paths shown. Start/Stop positions are shown with $\bigcirc/\triangle$.](image1)

![Fig. 2. Tracking results for a single run of the algorithm](image2)

Fig. 3. True cardinality (solid line) shown versus estimated mean cardinality and corresponding standard deviation for the MCMC-Based particle Algorithm over 50 Monte Carlo runs

**V. CONCLUSION**

In this paper, we have reviewed the various MCMC-Based particle algorithms that have been developed over the past years for approximating the filtering distribution in a general state-space model. These methods clearly represent interesting alternatives to SMC methods, especially for high-dimensional problems. This approach is then efficiently applied to detect, track and identify multiple targets in very low probability of detection and in hostile environments with heavy clutter. Simulations show that this approach exhibits a good tracking performance.

**ACKNOWLEDGMENT**

The authors would like to thank the Statistical and Applied Mathematical Sciences Institute - SMC program for providing a collaborative research environment that assisted with the development of our ideas.

**REFERENCES**


