BAYESIAN FILTERING WITH INTRACTABLE LIKELIHOOD USING SEQUENTIAL MCMC

François Septier[1], *Gareth W. Peters*[2] *and Ido Nevat*[3]

[1] Institut Mines-Télécom / Télécom Lille1 / LAGIS UMR CNRS 8219, France
[2]Department of Statistical Science, University College London, UK
[3]Wireless & Networking Tech. Lab, CSIRO, Sydney, Australia

ABSTRACT

We develop a sequential estimation methodology for a class of non-linear, non-Gaussian state space models in which the observation process is intractable to express in closed form, but trivial to simulate. In addition we consider models in which the latent state vector and the observation vector are very high dimensional. To overcome these two difficulties we propose the class of Sequential Markov chain Monte Carlo (SMCMC) algorithms in which we incorporate a component of Approximate Bayesian Computation (ABC). In doing so we tackle both the curse of dimensionality via the SMCMC and the intractability of the likelihood via the ABC component. We demonstrate how the proposed algorithm outperforms alternative approaches in two challenging state space model examples.

Index Terms— Bayesian filtering, intractable likelihood, MCMC, approximate Bayesian computation.

1. 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, $f_k(\mathbf{x}_k|\mathbf{x}_{k-1})$, of a hidden Markov process $\{\mathbf{x}_k; k \in \mathbb{N}\}$, $\mathbf{x}_k \in \mathbb{R}^{n_\mathbf{x}}$ and an observation equation describes the likelihood, $g_k(\mathbf{y}_k|\mathbf{x}_k)$, of the observations $\{\mathbf{y}_k; k \in \mathbb{N}\}$, $\mathbf{y}_k \in \mathbb{R}^{n_\mathbf{y}}$. This hidden Markov Model (HMM) has been applied in a wide range of disciplines, e.g. finance [1], digital communication [2], tracking [3] etc. The diversity of the applications clearly shows the flexibility of such a model structure. Within a Bayesian framework, all relevant information about $\mathbf{x}_k$ given observations up to and including time k can be obtained from the filtering distribution $p(\mathbf{x}_k|\mathbf{y}_{0:k})$.

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. Sequential Monte Carlo (SMC) approaches have become a powerful methodology to cope with non-linear and non-Gaussian problems [4].These methods, also known as particle filters (PF), exploit numerical integration techniques for approximating the filtering probability density function of inherently nonlinear non-Gaussian systems. Using these methods, the obtained estimates of sequences of functional w.r.t. to the filtering distribution can be set arbitrarily close to the optimal solution at the expense of computational complexity [5].

However, Monte carlo methods require the density $g_k(\mathbf{y}_k|\mathbf{x}_k)$ to be available point-wise, which may not be the case in some applications. The approximate Bayesian computation (ABC) technique has become a popular scheme to tackle this problem, see discussion in [6] and examples in [7]. Recently, a sequential Monte Carlo method targeting an ABC approximation of the filtering distribution in general non-linear and non-Gaussian HMM has been proposed in [8]. Nevertheless, due to their sampling mechanization, SMC methods tend to be inefficient when applied to high-dimensional problems in which the observation and state vector is high dimensional. Markov chain Monte Carlo (MCMC) methods are generally more effective than PFs in high-dimensional spaces when n_x and n_y are large for a given time k. Their traditional formulation, however, allows sampling from probability distributions in a non-sequential fashion. Recently, sequential MCMC schemes were proposed by [9–14]. These approaches are distinct from the Resample-Move scheme [15] in particle filters where the MCMC algorithm is used to rejuvenate degenerate samples following importance sampling resampling. These methods [9–13] use neither resampling nor importance sampling.

Utilising recent advances in "likelihood-free" inference as well as in Monte-Carlo filtering techniques, we propose a novel filtering approach based on MCMC that sequentially estimates an ABC approximation of the posterior distribution of interest.

In Section 2, the approximate bayesian computation (ABC) method is desribed as well as the recent derivation of this framework for filtering using the sequential Monte-Carlo methodology. Section 3 presents firstly a brief review of the sequential MCMC which represents an alternative to the classical particle filter. Then, the proposed SMCMC-ABC is described. In Section 4, performances of the proposed algorithm are illustrated through numerical simulations with different models. Finally, conclusions are given in Section 5.

2. APPROXIMATE BAYESIAN COMPUTATION METHOD

2.1. Introduction

The *Approximate Bayesian Computation* (ABC) method, originally proposed in [16], is a class of algorithmic methods in Bayesian inference using statistical summaries and computer simulations. ABC algorithms have become popular in inverse problems where point-wise evaluation of the likelihood function $g(\mathbf{y}|\mathbf{x})$ is computationally prohibitive or intractable. The resulting approximation of the posterior distribution is

$$p_\epsilon(\mathbf{x}|\mathbf{y}) \propto \int_{\mathbb{R}^{n_\mathbf{y}}} \mathcal{K}_\epsilon(\mathbf{y}-\mathbf{u})g(\mathbf{u}|\mathbf{x})f(\mathbf{x})d\mathbf{u} \qquad (1)$$

where $\mathbf{u}$ is an auxiliary variable on the same space as the observed data $\mathbf{y}$. The function $\mathcal{K}_\epsilon(\cdot)$ is a smoothing kernel with scale parameter ϵ. This approximation improves as ϵ decreases and exactly recovers the target posterior distribution as $\epsilon \to 0$. In practice, the function $\mathcal{K}_\epsilon(\mathbf{y}-\mathbf{u})$ is expressed through low-dimensional vectors of summary statistics, $\Lambda(\cdot)$, such that this kernel weights the intractable posterior through Eq. (1) with high values when $\Lambda(\mathbf{u}) \approx \Lambda(\mathbf{y})$.

2.2. ABC for Filtering

In [8], the authors have proposed an ABC method for filtering in HMMs. They have investigated the theoretical and empirical bias as well as demonstrated the algorithm in real-case study. The authors propose a sequential Monte Carlo (SMC) algorithm to sample from the following target density at time k :

$$\begin{aligned} p_\epsilon(\mathbf{x}_{0:k}, \mathbf{u}_{0:k}|\mathbf{y}_{0:k}) &\propto \prod_{n=1}^{k} \mathcal{K}_{n,\epsilon}(\mathbf{y}_n - \mathbf{u}_n) g_n(\mathbf{u}_n|\mathbf{x}_n) f_n(\mathbf{x}_n|\mathbf{x}_{n-1}) \\ &\propto \mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}_k) g_k(\mathbf{u}_k|\mathbf{x}_k) f_k(\mathbf{x}_k|\mathbf{x}_{k-1}) \\ &\quad \times p_\epsilon(\mathbf{x}_{0:k-1}, \mathbf{u}_{0:k-1}|\mathbf{y}_{0:k-1}) \end{aligned} \tag{2}$$

The ABC approximation of the posterior distribution proposed in this paper is thus given by :

$$p_\epsilon(\mathbf{x}_{0:k}|\mathbf{y}_{0:k}) \propto \prod_{n=1}^{k} \tilde{b}_{n,\epsilon}(\mathbf{y}_n; \mathbf{x}_n) f_n(\mathbf{x}_n|\mathbf{x}_{n-1}) \tag{3}$$

where $\tilde{b}_{n,\epsilon}(\mathbf{y}_n; \mathbf{x}_n) = \int_{\mathbb{R}^{n_\mathbf{y}}} \mathcal{K}_{n,\epsilon}(\mathbf{y}_n - \mathbf{u}_n) g_n(\mathbf{u}_n|\mathbf{x}_n) d\mathbf{u}_n$. We can clearly see with this formulation that the ABC approximation of the posterior in Eq. (3) tends to the true posterior $p_\epsilon(\mathbf{x}_{0:k}|\mathbf{y}_{0:k})$, $\forall k$, as $\epsilon \to 0$. The SMC-ABC algorithm proposed in [8] is summarized in Algorithm 1.

Algorithm 1 SMC-ABC Algorithm [8]

Initialize particle and importance weights $\{\mathbf{x}_{-1}^{(j)}; \widetilde{w}_{-1}^{(j)}\}_{j=1}^{N_p}$
for $k = 1, \ldots, T$ **do**
  for $j = 1, \ldots, N_p$ **do**
    Propose $\{\mathbf{x}_k^{(j)}\} \sim f_k(\mathbf{x}_k|\mathbf{x}_{k-1}^{(j)})$
    Propose $\{\mathbf{u}_k^{(j)}\} \sim g_k(\mathbf{u}_k|\mathbf{x}_k^{(j)})$
    Compute weights $w_k^{(j)} = \mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}_k^{(j)}) \widetilde{w}_{k-1}^{(j)}$
  end for
  Normalization of the weights $\widetilde{w}_k^{(j)} = w_k^{(j)} \left[\sum_{i=1}^{N_p} w_k^{(j)}\right]^{-1}$
  if $N_eff < \eta$ **then** Resample particles
end for

2.3. Contribution

Application of ABC methods in an SMC context, involves at each iteration replacing evaluation of the intractable likelihood with comparison between the true observation vector $\mathbf{y}_k$ and a model simulated observation vector $\mathbf{u}_k$. Typically, if the dimension of $\mathbf{y}_k$ and $\mathbf{x}_k$ are not too large, then there is a reasonable chance of matching $\mathbf{u}_k$ for a given ABC tolerance ϵ. However, when the dimension of the observation or state vector increases, this probability decreases resulting in particle degeneracy, high variance in the incremental weights and consequently one is forced to either introduce additional bias by increasing the comparative ABC tolerance ϵ or increasing computational complexity by using more particles. Our main contribution involves the design of a Sequential MCMC-ABC algorithm (SMCMC-ABC) that circumvents these problems by allowing greater flexibility in the mutation stage of the particles, e.g. the ability to utilize block-wise Metropolis-Hastings within Gibbs rejection steps to efficiently explore the support of marginal distributions of the high dimensional state vector. Consequently, this proposed SMCMC-ABC can achieve a significant gain compared to this existing SMC-ABC algorithm.

3. SMCMC-ABC ALGORITHM

3.1. Brief review of MCMC based Particle algorithm

As detailed in [17], the sequential MCMC (SMCMC) is a powerful sequential methodology for filtering that targets the general joint posterior distribution of $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$:

$$p(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathbf{y}_{0:k}) \propto g_k(\mathbf{y}_k|\mathbf{x}_k) f_k(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}) \tag{4}$$

A MCMC procedure is used to make inference from this complex distribution, i.e. Eq. (4) is the target distribution. Since we do not have a closed form representation of the posterior distribution $p(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1})$ at time $k-1$, this latter will be approximated with an empirical distribution based on the current particle set

$$p(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}) \approx \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(j)}) \tag{5}$$

where N_p is the number of particles and (j) the particle index. Then, by plugging this particle approximation into Eq. (4),

$$p(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathbf{y}_{0:k}) \approx \frac{1}{N_p} g_k(\mathbf{y}_k|\mathbf{x}_k) \sum_{j=1}^{N_p} f_k(\mathbf{x}_k|\mathbf{x}_{k-1}^{(j)}) \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(j)})$$

Then, having made many joint draws from Eq. (4) using an appropriate MCMC scheme, the converged MCMC output for variable $\mathbf{x}_k$ can be extracted to give an updated marginalized particle approximation to $p(\mathbf{x}_k|\mathbf{y}_{0:k})$. In this way, sequential inference can be achieved. It should be noted that several sampling strategies in the refinement step can be done in order to improve the algorithm. In particular, if the state of interest is multivariate, and n_x is large for each $\mathbf{x}_k$ then $\mathbf{x}_k$ can be divided into P sub-blocks, $\mathbf{x}_k = [\mathbf{x}_{k,\Omega_1}, \ldots, \mathbf{x}_{k,\Omega_P}]$. Then each sub-block can be updated either via a random scan or a deterministic scan using a series of MH-within-Gibbs steps.

More precisely, at the m^{th} MCMC iteration, the procedure, to obtain samples from $p(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathbf{y}_{0:k})$, involves a joint Metropolis-Hastings (MH) proposal step where both $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ are updated jointly, as well as block refinement Metropolis-within-Gibbs steps where the blocks of $\mathbf{x}_k$ are updated individually. This algorithm, summarized in Algorithm 2, will be denoted by SMCMC.

Algorithm 2 SMCMC

Initialize particle set $\{\mathbf{x}_{-1}^{(j)}\}_{j=1}^{N_p}$
for $k = 1, \ldots, T$ **do**
  for $m = 1, \ldots, N_{MCMC}$ **do**
    Joint Draw
    Propose $\{\mathbf{x}_k^*, \mathbf{x}_{k-1}^*\} \sim q_1(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1})$
    Compute the MH acceptance probability $\rho_1 = \min\left(1, \frac{p(\mathbf{x}_k^*, \mathbf{x}_{k-1}^*|\mathbf{y}_{0:k})}{q_1(\mathbf{x}_k^*, \mathbf{x}_{k-1}^*|\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1})} \frac{q_1(\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1}|\mathbf{x}_k^*, \mathbf{x}_{k-1}^*)}{p(\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1}|\mathbf{y}_{0:k})}\right)$
    Accept $\{\mathbf{x}_k^m, \mathbf{x}_{k-1}^m\} = \{\mathbf{x}_k^*, \mathbf{x}_{k-1}^*\}$ with probability ρ_1
    Block Refinement
    Randomly divide $\mathbf{x}_k$ into P blocks $\{\Omega_p\}_{p=1}^{P}$
    for $p = 1, \ldots, P$ **do**
      Propose $\{\mathbf{x}_{k,\Omega_p}^*\} \sim q(\mathbf{x}_{k,\Omega_p}|\mathbf{x}_{k,\setminus\Omega_p}^m, \mathbf{x}_{k-1}^m)$
      Compute the MH acceptance probability $\rho_p = \min \left(1, \frac{p(\mathbf{x}_{k,\Omega_p}^*|\mathbf{x}_{k,\setminus\Omega_p}^m, \mathbf{x}_{k-1}^m, \mathbf{y}_{0:k})}{q(\mathbf{x}_{k,\Omega_p}^*|\mathbf{x}_{k,\setminus\Omega_p}^m, \mathbf{x}_{k-1}^m)} \frac{q(\mathbf{x}_{k,\Omega_p}^m|\mathbf{x}_{k,\setminus\Omega_p}^m, \mathbf{x}_{k-1}^m)}{p(\mathbf{x}_{k,\Omega_p}^m|\mathbf{x}_{k,\setminus\Omega_p}^m, \mathbf{x}_{k-1}^m, \mathbf{y}_{0:k})}\right)$
      Accept $\{\mathbf{x}_k^m\} = \{\mathbf{x}_k^*\}$ with probability ρ_p
    end for
    After a burn in period of N_{burn} , keep every N_{thin} MCMC output $\mathbf{x}_k^{(j)} = \mathbf{x}_k^m$ as the new particle set for approximating $p(\mathbf{x}_k|\mathbf{y}_{0:k})$, i.e. $\widehat{p}(\mathbf{x}_k|\mathbf{y}_{0:k}) = \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(\mathbf{x}_k - \mathbf{x}_k^{(j)})$
  end for
end for

3.2. Derivation of the SMCMC-ABC

It is straightforward to show that the ABC approximation of the posterior given in Eq. (2) admits the following marginal :

$$p_\epsilon(\mathbf{x}_k, \mathbf{x}_{k-1}, \mathbf{u}_k|\mathbf{y}_{0:k}) \quad \propto \quad \mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}_k) g_k(\mathbf{u}_k|\mathbf{x}_k) f_k(\mathbf{x}_k|\mathbf{x}_{k-1}) \times p_\epsilon(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}) \tag{8}$$

As in the sequential MCMC described in Section 3.1, we propose to use a set of unweighted particles to represent the density $p_\epsilon(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1})$:

$$p_\epsilon(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}) \approx \frac{1}{N_p} \sum_{j=1}^{N_p} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(j)}) \tag{9}$$

where N_p is the number of particles and (j) the particle index. Then, by plugging this particle approximation into Eq. (8), a MCMC procedure can be employed to draw samples from :

$$p_\epsilon(\mathbf{x}_k, \mathbf{x}_{k-1}, \mathbf{u}_k|\mathbf{y}_{0:k}) \quad \propto \quad \mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}_k) g_k(\mathbf{u}_k|\mathbf{x}_k) \times \sum_{j=1}^{N_p} f_k(\mathbf{x}_k|\mathbf{x}_{k-1}^{(j)}) \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(j)}) \tag{10}$$

A MCMC procedure can be designed to sequentially target this ABC approximation of posterior distribution $p_\epsilon(\mathbf{x}_k, \mathbf{x}_{k-1}, \mathbf{u}_k|\mathbf{y}_{0:k})$. In this SMCMC-ABC, the joint draw will involve a Metropolis-Hastings proposal step where $\mathbf{x}_k$, $\mathbf{x}_{k-1}$ and $\mathbf{u}_k$ are jointly updated. The refinement step will be discussed in the next section.

3.3. Discussion on the Block Refinement Step

The use of proposals for sub-blocks of the state vector in the original SMCMC algorithm was shown to provide a significant impact on the accuracy of the resulting particle estimate of the filtering distribution [18]. The idea involves the partitioning of the state vector $\mathbf{x}_k$ into P blocks and then the successive sampling of updates for each block via a Metropolis-Hasting within Gibbs rejection mechanism. When one further incorporates the model component of auxiliary variables associated to synthetic observations, due to the ABC structure, the ability to incorporate blockwise updates shows even greater performance gains compared to standard SMC-ABC solutions.

More precisely, in the ABC framework the auxiliary random vector $\mathbf{u}_k$ has to be sampled from the likelihood $g_k(\cdot|\mathbf{x}_k)$ in order to avoid the computation of the likelihood which is intractable. Consequently, if the proposal used in the block refinement step of the SMCMC-ABC only moves a sub-block of the state vector, denoted $\mathbf{x}_{k,\Omega_p}$, then the auxiliary variable $\mathbf{u}_k$ must also be sampled. However, in some cases, it will not be necessary to sample all the elements of this auxiliary variable $\mathbf{u}_k$. Reducing the dimension of the state to sample in this refinement step shows marked improvement in the final estimation accuracy. The likelihood density involved in the target distribution, Eq. (11), can be written as :

$$g_k(\mathbf{u}_k|\mathbf{x}_k) \quad = \quad g_k(\mathbf{u}_{k,\Gamma_p}|\mathbf{u}_{k,\setminus\Gamma_p}, \mathbf{x}_{k,\Omega_p}, \mathbf{x}_{k,\setminus\Omega_p}) \times g_k(\mathbf{u}_{k,\setminus\Gamma_p}|\mathbf{x}_{k,\Omega_p}, \mathbf{x}_{k,\setminus\Omega_p}) \tag{11}$$

As a consequence, if we want to propose a move for the p-th block $\mathbf{x}_{k,\Omega_p}$, the best strategy is to jointly sample $\mathbf{u}_{k,\Gamma_p}$ in which Γ_p denotes the smallest subset of $\mathbf{u}_k$ such that the conditions 1 and either the condition 2 or 3 are satisfied :

Cond. 1: $\mathbf{u}_{k,\Gamma_p}$ *can be sampled from* $g_k(\mathbf{u}_{k,\Gamma_p}|\mathbf{u}_{k,\setminus\Gamma_p}, \mathbf{x}_{k,\Omega_p}, \mathbf{x}_{k,\setminus\Omega_p})$
Cond. 2: The marginal likelihood distribution $g_k(\mathbf{u}_{k,\setminus\Gamma_p}|\mathbf{x}_{k,\Omega_p}, \mathbf{x}_{k,\setminus\Omega_p})$ *can be evaluated point-wise*
Cond. 3: This marginal likelihood could be reduced as

$$g_k(\mathbf{u}_{k,\setminus\Gamma_p}|\mathbf{x}_{k,\Omega_p}, \mathbf{x}_{k,\setminus\Omega_p}) = g_k(\mathbf{u}_{k,\setminus\Gamma_p}|\mathbf{x}_{k,\setminus\Omega_p}) \tag{12}$$

where $\mathbf{x}_{k,\setminus\Omega_p}$ consists of all the elements in $\mathbf{x}_k$ which are not elements of the block $\mathbf{x}_{k,\Omega_p}$. In conclusion, in this refinement step, if $\mathbf{x}^*_{k,\Omega_p}$ is sampled from the proposal distribution $q(\mathbf{x}_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})$ and $\mathbf{u}^*_{k,\Gamma_p}$ is sampled from $g_k(\mathbf{u}_{k,\Gamma_p}|\mathbf{u}^m_{k,\setminus\Gamma_p}, \mathbf{x}^*_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p})$, the acceptance ratio will be either Eq. (6) if Condition 2 is satisfied or simplifies to Eq. (7) if Condition 3 is satisfied. The SMCMC-ABC is summarized in Algorithm 3.

Algorithm 3 SMCMC-ABC

Initialize particle set $\{\mathbf{x}_{-1}^{(j)}\}_{j=1}^{N_p}$
for $k = 1, \ldots, T$ **do**
  for $m = 1, \ldots, N_{MCMC}$ **do**
    Joint Draw
    Propose $\{\mathbf{x}^*_k, \mathbf{x}^*_{k-1}\} \sim q_1(\mathbf{x}_k, \mathbf{x}_{k-1}|\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1})$
    Propose $\mathbf{u}^*_k \sim g_k(\mathbf{u}_k|\mathbf{x}^*_k)$
    Compute the MH acceptance probability $\rho_1 = \min\left(1, \frac{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}^*_k) f_k(\mathbf{x}^*_k|\mathbf{x}^*_{k-1})}{q_1(\mathbf{x}^*_k, \mathbf{x}^*_{k-1}|\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1})} \frac{q_1(\mathbf{x}_k^{m-1}, \mathbf{x}_{k-1}^{m-1}|\mathbf{x}^*_k, \mathbf{x}^*_{k-1})}{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - \mathbf{u}_k^{m-1}) f_k(\mathbf{x}_k^{m-1}|\mathbf{x}_{k-1}^{m-1})}\right)$
    Accept $\{\mathbf{u}^m_k, \mathbf{x}^m_k, \mathbf{x}^m_{k-1}\} = \{\mathbf{u}^*_k, \mathbf{x}^*_k, \mathbf{x}^*_{k-1}\}$ with probability ρ_1
    Block Refinement
    Randomly divide $\mathbf{x}_k$ into P blocks $\{\Omega_p\}_{p=1}^P$
    Find associated partition of the observations $\{\Gamma_p\}_{p=1}^P$
    for $p = 1, \ldots, P$ **do**
      Propose $\{\mathbf{x}^*_{k,\Omega_p}\} \sim q(\mathbf{x}_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})$
      Propose $\mathbf{u}^*_{k,\Gamma_p} \sim g_k(\mathbf{u}_{k,\Gamma_p}|\mathbf{u}^m_{k,\setminus\Gamma_p}, \mathbf{x}^*_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p})$
      Compute the MH acceptance probability ρ_{p,c_2} or ρ_{p,c_3}
      Accept $\{\mathbf{u}^m_{k,\Gamma_p}, \mathbf{x}^m_k\} = \{\mathbf{u}^*_{k,\Gamma_p}, \mathbf{x}^*_k\}$ with probability ρ_{p,c_2} or ρ_{p,c_3}
    end for
    After a burn in period of N_{burn} , keep every N_{thin} MCMC output $\mathbf{x}_k^{(j)} = \mathbf{x}_k^m$ as the new particle set
  end for
end for

4. NUMERICAL STUDY

We compare the performance of the proposed SMCMC-ABC approach with the SMC-ABC [8] and an extension of it, denoted SMC-RM-ABC, that includes a Resample-Move stage after the resampling step, under two challenging models. To perform fair comparison between the SMC-RM and the SMCMC, the resample move stage adopted in the SMC-RM utilizes an identical block refinement stage with an equivalent Markov kernel of the SMCMC for moving each particle just after the resampling step (Line 9 - Algo 1). We note that the results are produced as average over 50 realizations from the model with time series of length $T = 50$. In all studies the kernel $\mathcal{K}_{k,\epsilon}$ is selected to be a multivariate Gaussian with scale ϵ.

$$\rho_{p,c2} = \min\left(1, \frac{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - [\mathbf{u}^*_{k,\Gamma_p}, \mathbf{u}^m_{k,\setminus\Gamma_p}]) g_k(\mathbf{u}^*_{k,\setminus\Gamma_p}|\mathbf{x}^*_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}) f_k(\mathbf{x}^*_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}|\mathbf{x}^*_{k-1})}{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - [\mathbf{u}^m_{k,\Gamma_p}, \mathbf{u}^m_{k,\setminus\Gamma_p}]) g_k(\mathbf{u}^m_{k,\setminus\Gamma_p}|\mathbf{x}^m_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}) f_k(\mathbf{x}^m_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}|\mathbf{x}^m_{k-1})} \frac{q(\mathbf{x}^m_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})}{q(\mathbf{x}^*_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})}\right) \tag{6}$$

$$\rho_{p,c3} = \min\left(1, \frac{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - [\mathbf{u}^*_{k,\Gamma_p}, \mathbf{u}^m_{k,\setminus\Gamma_p}]) f_k(\mathbf{x}^*_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}|\mathbf{x}^*_{k-1})}{\mathcal{K}_{k,\epsilon}(\mathbf{y}_k - [\mathbf{u}^m_{k,\Gamma_p}, \mathbf{u}^m_{k,\setminus\Gamma_p}]) f_k(\mathbf{x}^m_{k,\Omega_p}, \mathbf{x}^m_{k,\setminus\Omega_p}|\mathbf{x}^m_{k-1})} \frac{q(\mathbf{x}^m_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})}{q(\mathbf{x}^*_{k,\Omega_p}|\mathbf{x}^m_{k,\setminus\Omega_p}, \mathbf{x}^m_{k-1})}\right) \tag{7}$$

4.1. Linear and Gaussian Model

Let us firstly consider the dynamic linear and Gaussian model

$$\begin{cases} \mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{b}_k \\ \mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \tag{13}$$

with $n_x = n_y$, $\mathbf{A}_k = 0.99\mathbf{I}_{n_x}$ and $\mathbf{H}_k = \mathbf{I}_{n_x}$. The state and observation noise are zero-mean multivariate Gaussian random variables with covariance matrix $\mathbf{\Sigma}_b = 2\mathbf{I}_{n_x}$ and $\mathbf{\Sigma}_v = \mathbf{I}_{n_x}$.

This example is used as a comparative illustration of the accuracy of the ABC algorithms in which the likelihood is tractable and additionally the optimal solution is known in closed form via the Kalman filter. It is always optimal from a statistical perspective to utilize the exact likelihood in the algorithms described and therefore the SMCMC, SMC-RM and SMC algorithms act as direct comparisons to their ABC counterparts. Furthermore, we study the degradation in performance as the ABC tolerance is increased. Finally, the impact of sampling only a sub-vector of the auxiliary observation at each iteration of the block refinement step is analyzed. In the (SMCMC or SMC-RM)-ABC-FAS, all the auxiliary observation is sampled ($\Gamma_p = \{1, \ldots, n_y\}$) while in the ABC-BAS versions, only a sub-vector is sampled (in this example $\Gamma_p = \Omega_p$ are of size 2).

Figure 1 shows the log-relative MSE error between the Monte-Carlo algorithms and the Kalman filter which corresponds in this model to the optimal algorithm and Table 1 shows the MSE for different numbers of particles for each algorithm. These results demonstrate two clearly evident features, firstly that the SMCMC-ABC algorithm performs close to optimal for small ABC tolerances and outperforms both SMC-ABC and the modified version SMC-RM-ABC for this high-dimensional problems in terms of MSE. In addition, the incorporation of the block-refinement stage with sub-sampling of the auxiliary ABC observation vector (ABC-BAS) significantly improves the mixing of the MCMC stage in all cases. We quantify this through the average acceptance probabilities for the SMCMC-ABC algorithm for a range of dimensions as shown in Figure 2. Let us finally remark that the value of ϵ should not be too small in order to prevent the collapse of the Monte-carlo approximation of the ABC posterior and not too high in order to have small approximation error between the ABC posterior and the true posterior.

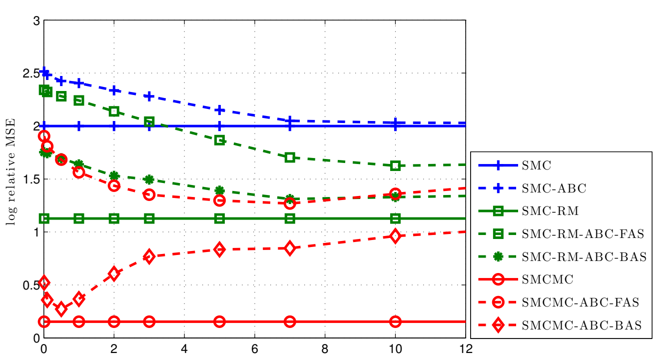

Fig. 1. Evolution of the log-relative time average MSE error as a function of the ABC tolerance ϵ with $n_x = 30$ and $N_p = 500$ particles. All results present log MSE relative to the optimal Kalman filter so zero represents the Kalman reference.

4.2. Heavy-tailed Observation Noise

In this example, we consider the following model:

$$\begin{cases} \mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{b}_k \\ y_{k,i} = x_{k,i} + v_{k,i} \qquad \forall i = \{1, \ldots, n_x\} \end{cases} \tag{14}$$

N_p	*Algorithm*	*Standard*	*ABC-FAS*	*ABC-BAS*
500	SMC	68.85	69.01	
	SMC-RM	39.09	49.3	40.23
	SMCMC	16.26	37.4	20.62
1000	SMC	58.62	58.71	
	SMC-RM	31.03	49.3	32.57
	SMCMC	15.69	32.11	19.51
Kalman filter		14.45		

Table 1. Time average of the MSE obtained for the linear and Gaussian model with $n_x = n_y = 20$. The MSE of the ABC version of the Monte-Carlo algorithms corresponds to the best one obtained on different values for ϵ.

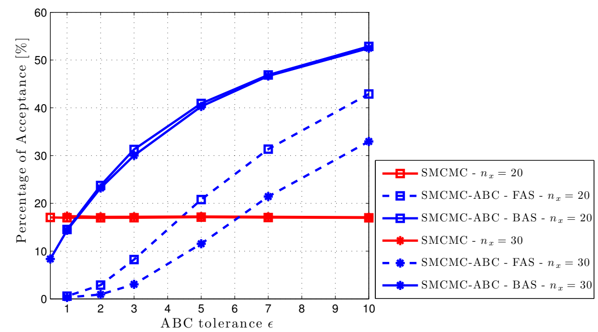

Fig. 2. Evolution of the acceptance rate of the different variants of the SMCMC-ABC as a function of the ABC tolerance ϵ.

where $n_x = n_y = 20$ and the observation noise samples are distributed according to the heavy tailed stable mixture model:

$$\{v_{k,i}\}_{i=1}^{n_x} \overset{\text{i.i.d.}}{\sim} \lambda\mathcal{S}_\alpha(v_{k,i}|\beta,\gamma,\mu) + (1-\lambda)\mathcal{S}_\alpha(-v_{k,i}|\beta,\gamma,\mu) \tag{15}$$

$\mathcal{S}_\alpha(\cdot|\beta,\gamma,\mu)$ denotes the α-stable distribution with characteristic exponent $0 < \alpha < 2$, dispersion parameter $\gamma > 0$, location parameter μ and skewness parameter $\beta \in [-1; 1]$. The likelihood under this model is in general intractable to write in closed form, however we consider for each mixture component the special case of Levy distribution ($\alpha = 0.5$, $\beta = 1$) in order to make a comparison with the non-ABC versions. Figure 3 shows the filtering result in terms of the log time average MSE of the algorithms as a function of the ABC tolerance ϵ. We observe again the out performance of the SMCMC-ABC algorithm versus its competitors in all simulations.

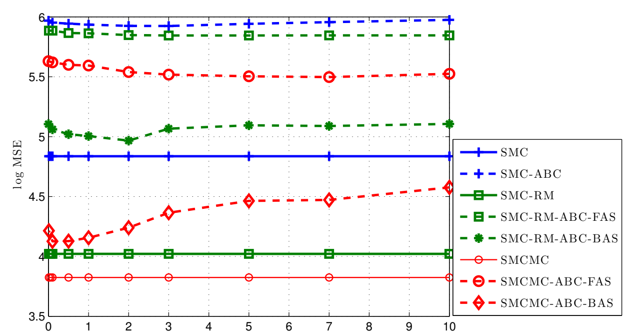

Fig. 3. Evolution of the logarithm of the time average MSE of the algorithms as a function of the ABC tolerance ϵ.

5. CONCLUSION

In this work, we address the challenging problem of filtering in the context of high-dimensional state and observation vectors with an intractable likelihood. In order to solve this problem, a novel sequential MCMC based on an ABC approximation of the filtering distribution is derived. Numerical simulations clearly show that this proposed SMCMC-ABC approach outperforms existing ABC methods. Future work will be dedicated to the development of an automatic learning strategy for the optimal ABC tolerance ϵ.

6. REFERENCES

[1] M. Pitt and N. Shephard, "Filtering Via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–599, 1999.

[2] P. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle Filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, Sept. 2003.

[3] N. Gordon, D. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F, Radar Signal Process.*, vol. 140, pp. 107–113, 1993.

[4] 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.

[5] P. Del Moral, "Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications," Springer, New York, 2004.

[6] G. W. Peters, Y. Fan, and S. A. Sisson, "On sequential Monte Carlo, partial rejection control and approximate Bayesian computation," *Statistics and Computing*, vol. 22, no. 6, pp. 1209–1222, Nov. 2012.

[7] G. W. Peters, S. A. Sisson, and Y. Fan, "Likelihood-free Bayesian inference for α-stable models," *Computational Statistics and Data Analysis*, pp. 1–14, 2010.

[8] A. Jasra, S. S. Singh, J. S. Martin, and E. McCoy, "Filtering via approximate Bayesian computation," *Statistics and Computing*, May 2010.

[9] C. Berzuini, N. G. Best, W. R. Gilks, and C. Larizza, "Dynamic Conditional Independence Models and Markov Chain Monte Carlo Methods," *Journal of the American Statistical Association*, vol. 92, no. 440, pp. 1403–1412, 1997.

[10] Z. Khan, T. Balch, and F. Dellaert, "MCMC-Based Particle Filtering for Tracking a Variable Number of Interacting Targets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1805–1819, Nov. 2005.

[11] A. Golightly and D. Wilkinson, "Bayesian sequential inference for nonlinear multivariate diffusions," *Statistics and Computing*, vol. 16, no. 4, pp. 323–338, 2006.

[12] S. Pang, J. Li, and S. Godsill, "Models and Algorithms for Detection and Tracking of Coordinated Groups," *IEEE Aerospace Conference*, pp. 1–17, 2008.

[13] F. Septier, A. Carmi, S. Pang, and S. Godsill, "Multiple Object Tracking Using Evolutionary and Hybrid MCMC-Based Particle Algorithms," in *15th IFAC Symposium on System Identification, (SYSID 2009)*, Saint Malo, France, July 2009.

[14] A. Carmi, F. Septier, and S. Godsill, "The Gaussian MCMC particle algorithm for dynamic cluster tracking," *Automatica*, vol. 48, no. 10, pp. 2454–2467, 2012.

[15] W. R. Gilks and C. Berzuini, "Following a Moving Target-Monte Carlo Inference for Dynamic Bayesian Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, pp. 127–146, 2001.

[16] J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and A. Feldman, "Population Growth of Human Y Chromosomes: A Study of Y Chromosome Microsatellites," *Mol. Biol. Evol.*, vol. 16, no. 12, pp. 1791–1798, 1999.

[17] F. Septier, S. Pang, A. Carmi, and S. Godsill, "On MCMC-Based Particle Methods for Bayesian Filtering : Application to Multitarget Tracking," in *International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2009)*, Aruba, Dutch Antilles, Dec. 2009.

[18] F. Septier, J. Cornebise, S. Godsill, and Y. Delignon, "A Comparative Study of Monte-Carlo Methods for Multitarget Tracking," in *IEEE International Workshop on Statistical Signal Processing (SSP)*, Nice, France, June 2011.