

Particle Filtering Under General Regime Switching

Yousef El-Laham, Liu Yang, Petar M. Djurić, Mónica F. Bugallo

Department of Electrical & Computer Engineering

Stony Brook University, Stony Brook (USA)

{yousef.ellaham, liu.yang.2, petar.djuric, monica.bugallo}@stonybrook.edu

Abstract—In this paper, we consider a new framework for particle filtering under model uncertainty that operates beyond the scope of Markovian switching systems. Specifically, we develop a novel particle filtering algorithm that applies to general regime switching systems, where the model index is augmented as an unknown time-varying parameter in the system. The proposed approach does not require the use of multiple filters and can maintain a diverse set of particles for each considered model through appropriate choice of the particle filtering proposal distribution. The flexibility of the proposed approach allows for long-term dependencies between the models, which enables its use to a wider variety of real-world applications. We validate the method on a synthetic data experiment and show that it outperforms state-of-the-art multiple model particle filtering approaches that require the use of multiple filters.

I. INTRODUCTION

In the past three decades, particle filtering (PF) [1], [2] has emerged as one of the most powerful statistical tools for online state estimation in dynamical systems. PF methods approximate the posterior distribution of an unknown time-varying parameter vector in a state-space model (SSM) using a set of weighted samples. The samples, which are also called particles, are drawn from a probability distribution called the *proposal distribution* and are weighted properly according to the principle of importance sampling [3]. Unlike Kalman filtering and its extensions [4], [5], PF can deal with SSMs that exhibit both nonlinearities and non-Gaussianities. This flexibility has allowed PF methods to thrive in many applications in fields as diverse as signal processing, economics, neuroscience, epidemiology, and ecology [6]–[10].

Model uncertainty introduces an additional layer of complexity to stochastic filtering that is generally difficult to deal with. In this situation, one must determine the model that best represents the system of interest from a set of candidate models, while also jointly estimating the unknown time-varying parameters of the chosen model. The issue of model uncertainty is further complicated if the model can switch from one time instant to the next. In signal processing, the well-known problem of tracking a maneuvering target falls within this class of model selection problems [11]. The trajectory of a maneuvering target is represented via a Markovian switching system (i.e., jump Markov systems) [12], where the model

dynamics change according to the state of a discrete-time, discrete-state Markov chain. More generally, however, systems whose models (or regimes) can change from one time instant to the next are referred to as *regime-switching systems*.

There have been mainly two classes of solutions proposed in the PF literature to deal with the challenge of model uncertainty. In the first class of solutions, a model index that references the candidate models is augmented as an unknown state in the system and is jointly estimated with the unknown time-varying parameters using a single particle filter [12], [13]. While this solution is simple to implement and straightforwardly tackles the joint estimation problem, the disadvantage is that the number of samples assigned to each candidate model cannot be controlled, which can lead to numerical issues and a lack of diversity in the considered models. The second class of solutions employs the use of a bank of particle filters that operate in parallel. Each filter is conditioned on one of the candidate models and state estimates are obtained by weighting the results of each of the filters and then fusing them. In [14], the filters are weighted according to the posterior probability of their respective models, while in [15], they are weighted according to the predictive powers of their respective models. Hybrid solutions which combine this class of approaches with interacting multiple models (IMM) filter have also been proposed to deal with Markovian switching systems [16]. Unfortunately, because a separate filter is required for each model, this class of approaches can be computationally intensive if the number of candidate models is large. We remark that, to the best of our knowledge, non-heuristic implementations of both methods have not been formulated for more general systems, which may exhibit long-term dependencies in the regime switching dynamics beyond those of the Markovian switching systems.

In this work, we propose a novel PF algorithm for general regime switching systems. Similar to the aforementioned solutions, the proposed PF method augments the model index as an unknown in the system that is jointly estimated with the time-varying parameters. Unlike previous approaches in the literature, a diverse set of candidate models can always be considered in the proposed algorithm through appropriate choice of the model index proposal distribution. Furthermore, since our method is not restricted to Markovian switching systems, it can handle more complicated processes that may describe the regime switching dynamics, such as the Pólya urn process [17]. Simulation results for synthetic data experiments validate the performance of the proposed approach.

This work was carried out thanks to the support of the National Science Foundation (NSF) under Awards CCF-1617986 and CCF-1618999. The authors thank the Research Computing and Cyberinfrastructure and the Institute for Advanced Computational Science at Stony Brook University for access to the high-performance SeaWulf computing system sponsored by NSF (# 1531492).

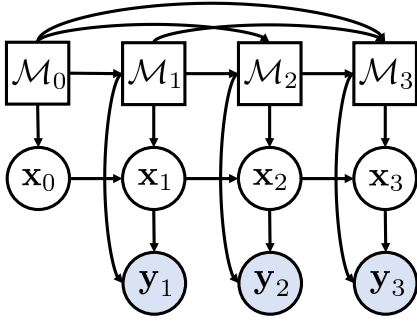


Fig. 1: The considered regime switching SSM formulation for a time horizon $T = 3$. The probability of a model depends on the complete history of models at previous time instants.

II. PROBLEM FORMULATION

Let $\mathbf{x}_t \in \mathbb{R}^{d_x}$ denote a latent state vector, $\mathbf{y}_t \in \mathbb{R}^{d_y}$ denote a measurement vector, and $\mathcal{M}_t \in \{1, \dots, K\}$ denote a model index from a set of K candidate models, where t denotes time index. We consider the generic stochastic filtering problem under model uncertainty over a fixed time horizon T . The generative process is assumed to have the following form:

$$\mathcal{M}_t \sim p(\mathcal{M}_t | \mathcal{M}_{0:t-1}), \quad (1)$$

$$\mathbf{x}_t \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathcal{M}_t), \quad (2)$$

$$\mathbf{y}_t \sim p(\mathbf{y}_t | \mathbf{x}_t, \mathcal{M}_t), \quad (3)$$

for $t = 1, \dots, T$, where the initial model \mathcal{M}_0 is distributed according to $\mathcal{M}_0 \sim p(\mathcal{M}_0)$ and the initial latent state \mathbf{x}_0 is distributed according to $\mathbf{x}_0 \sim p(\mathbf{x}_0 | \mathcal{M}_0)$. A graphical representation of the system is shown in Fig. 1 for $T = 3$.

Our goal is to jointly infer the unknown states $\mathbf{x}_{0:T} \triangleq [\mathbf{x}_0, \dots, \mathbf{x}_T] \in \mathbb{R}^{d_x \times (T+1)}$ and the sequence of unknown models $\mathcal{M}_{0:T} \triangleq (\mathcal{M}_0, \dots, \mathcal{M}_T)$ based on the given observations $\mathbf{y}_{1:T} \triangleq [\mathbf{y}_1, \dots, \mathbf{y}_T] \in \mathbb{R}^{d_y \times T}$ under the Bayesian paradigm. In other words, we would like to approximate the joint posterior distribution $p(\mathbf{x}_{0:T}, \mathcal{M}_{0:T} | \mathbf{y}_{1:T})$ in a recursive manner using a Bayesian filtering solution.

III. REGIME SWITCHING PARTICLE FILTERING

In this section, we derive a generalized regime switching PF (RSPF) algorithm. We begin by first establishing the recursiveness of the joint distribution and then we derive the importance weights of the particles in the novel PF algorithm. We discuss different strategies for sampling models which allow for model diversity. Finally, we elaborate on how one can use the proposed algorithm to obtain the maximum *a posteriori* (MAP) estimate of the model at each time instant. We summarize the proposed approach in Algorithm 1.

A. Deriving the Joint Distribution

At time t , the distribution of interest is $p(\mathbf{x}_{0:t}, \mathcal{M}_{0:t} | \mathbf{y}_{1:t})$. This joint distribution can be decomposed as

$$p(\mathbf{x}_{0:t}, \mathcal{M}_{0:t} | \mathbf{y}_{1:t}) = p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}, \mathcal{M}_{0:t}) p(\mathcal{M}_{0:t} | \mathbf{y}_{1:t}), \quad (4)$$

Algorithm 1 Regime Switching Particle Filtering (RSPF)

- 1: **Initialization:** Draw N samples from the prior of the initial model to determine the model indexes

$$\mathcal{M}_0^{(n)} \sim p(\mathcal{M}_0), \quad n = 1, \dots, N,$$

and draw N samples from the prior of the initial state conditioned on the sampled model indexes

$$\mathbf{x}_0^{(n)} \sim p(\mathbf{x}_0 | \mathcal{M}_0^{(n)}), \quad n = 1, \dots, N.$$

Set the weights as $\tilde{w}_0^{(n)} = \frac{1}{N}$ for $n = 1, \dots, N$.

- 2: **for** $t = 1, \dots, T$ **do**

- 3: **Sampling models:** Draw N samples from the model index proposal distribution

$$\mathcal{M}_t^{(n)} \sim q(\mathcal{M}_t | \mathcal{M}_{0:t-1}^{(n)}), \quad n = 1, \dots, N.$$

- 4: **Sampling states:** Draw N samples of the states conditioned on the drawn models

$$\mathbf{x}_t^{(n)} \sim q(\mathbf{x}_t | \mathbf{x}_{t-1}^{(n)}, \mathcal{M}_t^{(n)}, \mathbf{y}_t), \quad n = 1, \dots, N.$$

- 5: **Weighting:** Compute the weights $\{\tilde{w}_t^{(n)}\}_{n=1}^N$ according to (15) and normalize them as

$$w_t^{(n)} = \frac{\tilde{w}_t^{(n)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}}, \quad n = 1, \dots, N.$$

- 6: **Model selection:** Determine the model index estimate $\hat{\mathcal{M}}_t$ by solving the following maximization problem

$$\hat{\mathcal{M}}_t = \arg \max_{k \in \{1, \dots, K\}} p(\mathcal{M}_t = k | \mathbf{y}_{1:t})$$

where $p(\mathcal{M}_t = k | \mathbf{y}_{1:t}) \approx \sum_{n=1}^N w_t^{(n)} \mathbb{1}(\mathcal{M}_t^{(n)} = k)$ for all k .

- 7: **State estimation:** Obtain the state estimate as

$$\hat{\mathbf{x}}_t = \sum_{n=1}^N w_t^{(n)} \mathbf{x}_t^{(n)}.$$

- 8: **Resampling:** If necessary, resample the model indexes and the states using multinomial resampling and set the weights as $\tilde{w}_t^{(n)} = \frac{1}{N}$ for $n = 1, \dots, N$.
 - 9: **end for**
-

where $p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}, \mathcal{M}_{0:t})$ is the posterior of the state trajectory $\mathbf{x}_{0:t}$ conditioned on the model sequence $\mathcal{M}_{0:t}$ and $p(\mathcal{M}_{0:t} | \mathbf{y}_{1:t})$ is the marginal posterior distribution of the model sequence. Following previous analysis in the Bayesian filtering literature, we can readily deduce the conditional posterior of the state trajectory $p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}, \mathcal{M}_{0:t})$ as

$$p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}, \mathcal{M}_{0:t}) = \frac{p(\mathbf{y}_t | \mathbf{x}_t, \mathcal{M}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathcal{M}_t) p(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}, \mathcal{M}_{0:t-1})}{p(\mathbf{y}_t | \mathbf{y}_{1:t-1}, \mathcal{M}_{0:t})}. \quad (5)$$

Remark that we have made the appropriate assumption that $p(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}, \mathcal{M}_{0:t}) = p(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}, \mathcal{M}_{0:t-1})$, since the conditional posterior at time instant $t-1$ does not depend on the model at time t . Applying Bayes' theorem to the marginal

posterior of the model sequence $p(\mathcal{M}_{0:t}|\mathbf{y}_{1:t})$, we also have

$$p(\mathcal{M}_{0:t}|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathcal{M}_{0:t})p(\mathcal{M}_{0:t}|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}. \quad (6)$$

We decompose $p(\mathcal{M}_{0:t}|\mathbf{y}_{1:t-1})$ as

$$p(\mathcal{M}_{0:t}|\mathbf{y}_{1:t-1}) = p(\mathcal{M}_t|\mathcal{M}_{0:t-1}, \mathbf{y}_{1:t-1})p(\mathcal{M}_{0:t-1}|\mathbf{y}_{1:t-1}), \quad (7)$$

where one can show that

$$p(\mathcal{M}_t|\mathcal{M}_{0:t-1}, \mathbf{y}_{1:t-1}) = \frac{p(\mathbf{y}_{1:t-1}|\mathcal{M}_{0:t})p(\mathcal{M}_t|\mathcal{M}_{0:t-1})}{p(\mathbf{y}_{1:t-1}|\mathcal{M}_{0:t-1})} \quad (8)$$

$$= \frac{p(\mathcal{M}_t|\mathcal{M}_{0:t-1})}{p(\mathbf{y}_{1:t-1}|\mathcal{M}_{0:t-1})} \int p(\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}|\mathcal{M}_{0:t})d\mathbf{x}_{0:t-1} \quad (9)$$

$$= \frac{p(\mathcal{M}_t|\mathcal{M}_{0:t-1})}{p(\mathbf{y}_{1:t-1}|\mathcal{M}_{0:t-1})} p(\mathbf{y}_{1:t-1}|\mathcal{M}_{0:t-1}) \quad (10)$$

$$= p(\mathcal{M}_t|\mathcal{M}_{0:t-1}), \quad (11)$$

since the joint distribution of $\mathbf{x}_{0:t-1}$ and $\mathbf{y}_{1:t-1}$ is conditionally independent from model \mathcal{M}_t given the sequence of models up to time instant $t-1$. We can now establish the recursive solution to the joint posterior $p(\mathbf{x}_{0:t}, \mathcal{M}_{0:t}|\mathbf{y}_{1:t})$ as

$$p(\mathbf{x}_{0:t}, \mathcal{M}_{0:t}|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\mathbf{x}_t, \mathcal{M}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathcal{M}_t) \times p(\mathcal{M}_t|\mathcal{M}_{0:t-1})p(\mathbf{x}_{0:t-1}, \mathcal{M}_{0:t-1}|\mathbf{y}_{1:t-1}). \quad (12)$$

B. Deriving the Particle Filtering Weights

Suppose we draw a set of N samples $\{(\mathbf{x}_{0:t}^{(n)}, \mathcal{M}_{0:t}^{(n)})\}_{n=1}^N$, where each sample $(\mathbf{x}_{0:t}^{(n)}, \mathcal{M}_{0:t}^{(n)})$ is drawn from a proposal distribution $q(\mathbf{x}_{0:t}, \mathcal{M}_{0:t}|\mathbf{y}_{1:t})$ for $n = 1, \dots, N$. Then, the importance weight of each sample is determined according to

$$\tilde{w}_t^{(n)} = \frac{p(\mathbf{x}_{0:t}^{(n)}, \mathcal{M}_{0:t}^{(n)}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{0:t}^{(n)}, \mathcal{M}_{0:t}^{(n)}|\mathbf{y}_{1:t})}, \quad n = 1, \dots, N. \quad (13)$$

Suppose that the proposal distribution can be factored as

$$q(\mathbf{x}_{0:t}, \mathcal{M}_{0:t}|\mathbf{y}_{1:t}) = q(\mathbf{x}_{0:t-1}, \mathcal{M}_{0:t-1}|\mathbf{y}_{1:t-1}) \times q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathcal{M}_t, \mathbf{y}_t)q(\mathcal{M}_t|\mathcal{M}_{0:t-1}). \quad (14)$$

Combining with results from (12), the importance weight of each sample $(\mathbf{x}_{0:t}^{(n)}, \mathcal{M}_{0:t}^{(n)})$ can thus be determined as

$$\tilde{w}_t^{(n)} \propto \tilde{w}_{t-1}^{(n)} \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(n)}, \mathcal{M}_t^{(n)})p(\mathbf{x}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \mathcal{M}_t^{(n)})p(\mathcal{M}_t^{(n)}|\mathcal{M}_{0:t-1}^{(n)})}{q(\mathbf{x}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \mathcal{M}_t^{(n)}, \mathbf{y}_t)q(\mathcal{M}_t^{(n)}|\mathcal{M}_{0:t-1}^{(n)})}. \quad (15)$$

For sampling the states, the bootstrap implementation of this method would assume that the proposal distribution of the states is identical to the state transition distribution, i.e., $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathcal{M}_t, \mathbf{y}_t) = p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathcal{M}_t)$, and that the particles are resampled after each time instant. For the bootstrap implementation, the importance weights are given by

$$\tilde{w}_t^{(n)} \propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(n)}, \mathcal{M}_t^{(n)})p(\mathcal{M}_t^{(n)}|\mathcal{M}_{0:t-1}^{(n)})}{q(\mathcal{M}_t^{(n)}|\mathcal{M}_{0:t-1}^{(n)})}, \quad (16)$$

for $n = 1, \dots, N$. The obtained solution is analogous to the weighting function in bootstrap PF, except now, we must taken into account that models can change according to $p(\mathcal{M}_t|\mathcal{M}_{0:t-1})$. We remark that $p(\mathcal{M}_t|\mathcal{M}_{0:t-1})$ determines how the model \mathcal{M}_t is determined from the history of models $\mathcal{M}_{0:t-1}$ and depends on the nature of system being considered.

C. Discussion on Sampling Model Indexes

There are a variety of choices for the model index proposal distribution $q(\mathcal{M}_t|\mathcal{M}_{0:t-1})$. The most obvious choice is the *bootstrap* approach, where we use the model transition function as the proposal, i.e, we set $q(\mathcal{M}_t|\mathcal{M}_{0:t-1}) = p(\mathcal{M}_t|\mathcal{M}_{0:t-1})$. Then, the importance weights simply become the joint likelihood of the sampled model indexes and states. Alternatively, one can use a discrete *uniform* proposal distribution to sample the model indexes, i.e., $q(\mathcal{M}_t = k|\mathcal{M}_{0:t-1}) = \frac{1}{K}$ for all k . Then, each model has an equal chance to be sampled at each step of the algorithm, and thus avoiding the possibility of a model diversity issue. Finally, one can *deterministically* sample an equal number of particles for each model. The weights in this case would be the same as if we had sampled from the discrete uniform distribution.

D. Online Maximum A Posteriori Model Selection

In order to select the most promising model at each time instant from the set of candidate models, we need to solve the following optimization problem:

$$\hat{\mathcal{M}}_t = \arg \max_{k \in \{1, \dots, K\}} p(\mathcal{M}_t = k|\mathbf{y}_{1:t}), \quad (17)$$

where $p(\mathcal{M}_t = k|\mathbf{y}_{1:t})$ denotes the posterior probability of the k th model. This posterior probability of each model can be estimated directly using the set of particles and weights

$$p(\mathcal{M}_t = k|\mathbf{y}_{1:t}) \approx \frac{1}{\sum_{n=1}^N \tilde{w}_t^{(n)}} \sum_{n=1}^N \tilde{w}_t^{(n)} \mathbb{1}(\mathcal{M}_t^{(n)} = k), \quad (18)$$

for $k = 1, \dots, K$, where $\mathbb{1}(\cdot)$ denotes the indicator function. Given the estimated posterior probabilities, one can obtain an approximate solution to the optimization problem in (17).

IV. EXAMPLES OF MODEL SEQUENCE DYNAMICS

Here, we give examples of different regime switching dynamics that can easily be treated using our proposed approach.

A. Independent Regime Dynamics

The simplest case is when the models are generated independently from one another, i.e., the joint distribution of the models can be factored as:

$$p(\mathcal{M}_{0:T}) = \prod_{t=0}^T p(\mathcal{M}_t), \quad (19)$$

where the model independence assumption implies that $p(\mathcal{M}_t|\mathcal{M}_{0:t-1}) = p(\mathcal{M}_t)$. The assumption that the models are independent may be unrealistic for most applications and requires to specify the prior distribution of each model.

B. Markovian Switching Dynamics

We also consider Markovian switching systems, where the model at each time instant only depends on the model at the previous time instant. The joint distribution of the models under this assumption is given by

$$p(\mathcal{M}_{0:T}) = p(\mathcal{M}_0) \prod_{t=1}^T p(\mathcal{M}_t|\mathcal{M}_{t-1}), \quad (20)$$

where we have that $p(\mathcal{M}_t|\mathcal{M}_{0:t-1}) = p(\mathcal{M}_t|\mathcal{M}_{t-1})$. Here, the model transition distribution $p(\mathcal{M}_t|\mathcal{M}_{t-1})$ is represented by a transition probability matrix \mathbf{P}

$$\mathbf{P} = \begin{pmatrix} p_{1,1} & \cdots & p_{1,K} \\ \vdots & \ddots & \vdots \\ p_{K,1} & \cdots & p_{K,K} \end{pmatrix}, \quad (21)$$

where each element $p_{i,j} \triangleq p(\mathcal{M}_t = j|\mathcal{M}_{t-1} = i)$ is defined to be the probability of transitioning from model i to model j and each row of the matrix \mathbf{P} satisfies $\sum_{j=1}^K p_{i,j} = 1$.

C. Pólya Urn Dynamics

Under a more general formulation, the model at a given time instant t depends on the complete sequence of models $\mathcal{M}_{0:t-1}$. Here, since there are no independence assumptions, the joint distribution of the models is given by

$$p(\mathcal{M}_{0:T}) = p(\mathcal{M}_0) \prod_{t=1}^T p(\mathcal{M}_t|\mathcal{M}_{0:t-1}). \quad (22)$$

If the number of models is finite and *a priori* known, one possibility is to consider a Pólya urn process for the regime dynamics. For the Pólya urn process, the probability of transitioning to a particular model at time instant t depends on how many times that model was chosen in previous time instants. Let $\alpha_{k,t} = \mathbb{1}(\mathcal{M}_t = k)$ be variable indicating if model k was visited at time t for $t = 1, \dots, T$ and let $\beta_k \in \mathbb{N}$ be any positive integer for $k = 1, \dots, K$. Then, the probability of transitioning to model k at time t is given by

$$p(\mathcal{M}_t = k|\mathcal{M}_{0:t-1}) = \frac{\beta_k + \sum_{\tau=0}^{t-1} \alpha_{k,\tau}}{\sum_{j=1}^K (\beta_j + \sum_{\tau=0}^{t-1} \alpha_{j,\tau})}. \quad (23)$$

V. SIMULATIONS

To validate the performance of the proposed RSPF, we generated synthetic measurement sequences of time length $T = 50$ based on eight candidate models, with each model being of the form

$$\mathcal{M}_k : \begin{cases} x_t = a_k x_{t-1} + c_k + u_t \\ y_t = b_k \sqrt{|x_t|} + d_k + v_t \end{cases}, \quad (24)$$

where the parameter settings are $[a_1, \dots, a_8] = [-0.1, -0.3, -0.5, -0.9, 0.1, 0.3, 0.5, 0.9]$, $[c_1, \dots, c_8] = [0, -2, 2, -4, 0, 2, -2, 4]$, $[b_1, \dots, b_8] = [a_1, \dots, a_8]$, and $[d_1, \dots, d_8] = [c_1, \dots, c_8]$. The process noise u_t and observation noise v_t are assumed to be i.i.d. zero-mean Gaussian with equal variances, i.e., $u_t \sim \mathcal{N}(0, \sigma_u^2)$ and $v_t \sim \mathcal{N}(0, \sigma_v^2)$ with $\sigma_u^2 = \sigma_v^2 = 0.1$. The initial state x_0 was generated uniformly from -0.5 to 0.5. We tested the method on two scenarios corresponding to regime switching based on Markovian dynamics and Pólya urn dynamics, respectively.

We first ran our novel algorithm with $N = 2000$ particles per iteration when the model sequence dynamics is Markovian. The transition probability matrix in simulation was

$$\mathbf{P} = \begin{pmatrix} 0.80 & 0.15 & \epsilon & \cdots & \epsilon \\ \epsilon & 0.80 & 0.15 & \cdots & \epsilon \\ \vdots & & \ddots & \ddots & \vdots \\ \epsilon & \cdots & & 0.80 & 0.15 \\ 0.15 & \epsilon & \cdots & \epsilon & 0.80 \end{pmatrix}, \quad (25)$$

where we set $\epsilon = \frac{1}{120}$ so that each row of \mathbf{P} summed to 1.

Three different model index proposal distributions were used (deterministic, uniform, and bootstrap). For comparison, we also ran the multiple model particle filtering (MMPF) algorithm presented in [14], [18], where we drew 250 samples per model. Note that this algorithm considers a forgetting factor parameter $\gamma \in [0, 1]$ that determines how much the observation history influences the model probabilities. The closer γ is to 1, the more the observation history influences the model probabilities. We tested four different settings of this method, each corresponding to a different forgetting factor $\gamma \in \{0, 0.5, 0.9, 1\}$. The results are averaged over 500 Monte Carlo runs and are summarized in Tables I and II. We can see that the novel method, regardless of the choice of the model index proposal distribution, provides a smaller mean squared error (MSE) and more accurate model selection results. For reference, we also plot the average cumulative sum of the MSE in the state estimation in Fig. 2.

Next, we conducted the proposed method on the Pólya urn process. The initial counts of the eight models were a random permutation of the integers from 1 to 8. The parameter settings for RSPF and MMPF were the same as above. Table III and IV show the results, which are averaged over 500 Monte Carlo simulations. Again, we can see that the novel method outperforms MMPF with the settings $\gamma = 0.5, 0.9$ and 1 by far, and slightly outperforms the MMPF with a forgetting factor of 0 in terms of model selection accuracy, and in terms of state

	Average	Best	Worst
Novel (Deterministic)	0.2443	0.0566	4.8527
Novel (Uniform)	0.2446	0.0573	4.7388
Novel (Bootstrap)	0.2462	0.0546	4.9030
MMPF ($\gamma = 0$)	0.5986	0.0792	9.0508
MMPF ($\gamma = 0.5$)	9.9912	0.2635	279.7545
MMPF ($\gamma = 0.9$)	51.4122	1.5156	900.7603
MMPF ($\gamma = 1$)	63.5191	1.5136	1002.7609

TABLE I: State estimation MSE (Markovian dynamics).

	Average	Best	Worst
Novel (Deterministic)	0.9407	1	0.5000
Novel (Uniform)	0.9402	1	0.5000
Novel (Bootstrap)	0.9419	1	0.5000
MMPF ($\gamma = 0$)	0.8437	1	0.5200
MMPF ($\gamma = 0.5$)	0.5089	0.9200	0.1200
MMPF ($\gamma = 0.9$)	0.2348	0.9200	0
MMPF ($\gamma = 1$)	0.2180	0.9200	0

TABLE II: Model selection accuracy (Markovian dynamics).

	Average	Best	Worst
Novel (Deterministic)	0.4112	0.0663	2.3021
Novel (Uniform)	0.4111	0.0672	2.2512
Novel (Bootstrap)	0.4116	0.0644	2.3921
MMPF ($\gamma = 0$)	0.4995	0.0734	2.6993
MMPF ($\gamma = 0.5$)	4.4573	0.3854	28.0283
MMPF ($\gamma = 0.9$)	8.6751	1.9418	32.9856
MMPF ($\gamma = 1$)	11.1040	2.0316	43.6506

TABLE III: State estimation MSE (Pólya urn dynamics).

	Average	Best	Worst
Novel (Deterministic)	0.9003	1	0.6800
Novel (Uniform)	0.8996	1	0.6600
Novel (Bootstrap)	0.8996	1	0.6800
MMPF ($\gamma = 0$)	0.8526	1	0.6200
MMPF ($\gamma = 0.5$)	0.2867	0.5600	0.0600
MMPF ($\gamma = 0.9$)	0.1690	0.4200	0
MMPF ($\gamma = 1$)	0.1571	0.4800	0

TABLE IV: Model selection accuracy (Pólya urn dynamics).

estimation MSE. The average cumulative sum of the MSE for the state estimation is shown in Fig. 3.

VI. CONCLUSIONS

In this paper, we introduced a novel particle filtering algorithm for regime switching systems. The proposed method allows for the treatment of stochastic filtering problems under model uncertainty, where the model can change from one time instant to the next. Moreover, our algorithm does not have any restrictions on the regime switching dynamics and can work for systems that are not Markovian switching systems. We validated our method on two synthetic data experiments, where in the first experiment we considered a Markovian switching system and in the second experiment we considered a system where regimes changed according to a Pólya urn process.

REFERENCES

- [1] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings F (radar and signal processing)*, vol. 140, no. 2. IET, 1993, pp. 107–113.
- [2] 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.
- [3] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer Science & Business Media, 2013.
- [4] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.
- [5] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [6] R. Karlsson and N. Bergman, "Auxiliary particle filters for tracking a maneuvering target," in *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No. 00CH37187)*, vol. 4. IEEE, 2000, pp. 3891–3895.
- [7] D. Creal, "A survey of sequential Monte Carlo methods for economics and finance," *Econometric reviews*, vol. 31, no. 3, pp. 245–296, 2012.
- [8] A. E. Brockwell, A. L. Rojas, and R. E. Kass, "Recursive Bayesian decoding of motor cortical signals by particle filtering," *Journal of Neurophysiology*, vol. 91, no. 4, pp. 1899–1907, 2004.
- [9] C. Jégat, F. Carrat, C. Lajaudie, and H. Wackernagel, "Early detection and assessment of epidemics by particle filtering," in *geoENV VI—Geostatistics for Environmental Applications*. Springer, 2008, pp. 23–35.
- [10] M. Dowd, "A sequential Monte Carlo approach for marine ecological prediction," *Environmetrics: The official journal of the International Environmetrics Society*, vol. 17, no. 5, pp. 435–455, 2006.
- [11] M. F. Bugallo, S. Xu, and P. M. Djurić, "Performance comparison of ekf and particle filtering methods for maneuvering targets," *Digital Signal Processing*, vol. 17, no. 4, pp. 774–786, 2007.
- [12] L. Liang-Qun, X. Wei-Xin, H. Jing-Xiong, and H. Jian-Jun, "Multiple model Rao-Blackwellized particle filter for manoeuvring target tracking," *Defence Science Journal*, vol. 59, no. 3, p. 197, 2009.
- [13] S. McGinnity and G. W. Irwin, "Multiple model bootstrap filter for maneuvering target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 3, pp. 1006–1012, 2000.
- [14] B. Liu, "Instantaneous frequency tracking under model uncertainty via dynamic model averaging and particle filtering," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1810–1819, 2011.
- [15] I. Urteaga, M. F. Bugallo, and P. M. Djurić, "Sequential monte carlo methods under model uncertainty," in *2016 IEEE Statistical Signal Processing Workshop (SSP)*. IEEE, 2016, pp. 1–5.
- [16] Y. Boers and J. N. Driessen, "Interacting multiple model particle filter," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 344–349, 2003.
- [17] R. Gouet, "Strong convergence of proportions in a multicolor Pólya urn," *Journal of Applied Probability*, pp. 426–435, 1997.
- [18] B. Liu, "Robust particle filter by dynamic averaging of multiple noise models," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 4034–4038.

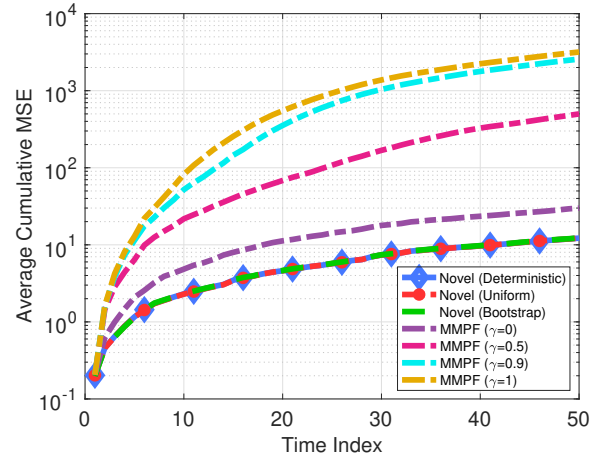


Fig. 2: Cumulative sum of MSE (Markovian dynamics).

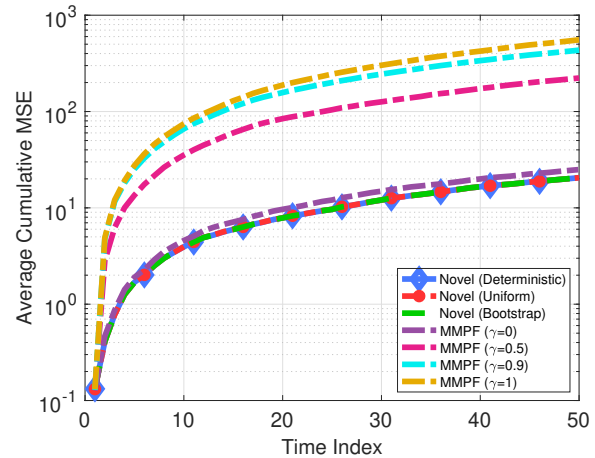


Fig. 3: Cumulative sum of MSE (Pólya urn dynamics).