MULTIPLE PARTICLE FILTERING FOR INFERENCE IN THE PRESENCE OF STATE CORRELATION OF UNKNOWN MIXING PARAMETERS

Iñigo Urteaga, Mónica F. Bugallo and Petar M. Djurić

Department of Electrical and Computer Engineering
Stony Brook University, Stony Brook, NY 11794-2350
{inigo.urteaga,monica.bugallo, petar.djuric}@stonybrook.edu

ABSTRACT

We present a novel Rao-Blackwellized multiple particle filtering method for inference of correlated latent states observed via nonlinear functions. We adopt a state-space framework and model the dynamic correlated states using a mixing matrix, embedded in white Gaussian noise. The critical challenges in practice are the lack of knowledge about the mixing parameters and the possibly large dimensionality of the state. We address these issues by implementing Rao-Blackwellization of the unknown parameters and adopting a divide-and-conquer approach. The former strategy amounts to marginalizing out some of the variables; the latter breaks the space of the system in subsystems, and runs a separate particle filter for each of them. The resulting Rao-Blackwellized multiple particle filtering accurately estimates the correlated latent states, as shown by the provided simulation results.

***Index Terms*—** Multiple particle filtering, correlated states, mixing matrix, Rao-Blackwellization, unknown parameters.

1. INTRODUCTION

Studying time-varying data is critical in various fields of science and engineering [1]. To that end, state-space modeling [2] and Bayesian dynamic models [3] are well-established techniques for analysis and prediction of time-series. Examples include neurology [4], urban planning [5], engineering [6] or finance [7]. The state time-series are often multivariate, and they show non-negligible correlation and are not directly observed. Moreover, for the study of correlated states one uses linearly mixing parameters, which result in multivariate Gaussian processes. In the most common setting, the state-space model is assumed to be linear, the noises in the state and observation equations are Gaussian, and the matrices and parameters that appear in these equations are known. In that case, the optimal solution for estimation of the dynamic state is the Kalman filter [8].

In this paper, we do not restrict ourselves to any of these assumptions, as we focus on more general practical settings. On the one hand, we work with Gaussian correlated multivariate states, where the static mixing parameters of the model are not known. On the other, we do not limit ourselves to linear observations. To overcome these challenges and resolve the parameter estimation problem, one needs to resort to suboptimal algorithms and, in this paper, we adopt the sequential Monte Carlo (SMC) or Particle Filtering (PF) method [9, 10]. These methods have been widely applied [11, 12] since the seminal publication of [13], but are known to suffer when dealing with static parameters and high-dimensional states. We address these challenges by applying Rao-Blackwellization [14] and by adopting the principles of Multiple Particle Filtering (MPF) [15].

In this paper, we focus on combining Rao-Blackwellization with MPF for inference in the presence of state correlation with unknown mixing parameters. By marginalizing out the unknown static parameters, we avoid generating particles for them and, thus, we improve the estimation accuracy of the method. With MPF, we focus on a set of marginal densities to mitigate the impact of increasing the state dimensionality. The contribution of this paper is on the efficient combination of both techniques for inference of correlated states with unknown parameters observed via nonlinear functions.

The next section introduces the problem of interest and motivates the need of the PF framework, which is briefly described in Section 3. The drawbacks related to systems with correlated states of unknown mixing parameters are addressed in Section 4, where the new method is proposed. Section 5 presents numerical results and Section 6 provides closing remarks.

2. PROBLEM FORMULATION

We consider state-space models described by a set of hidden processes correlated via an unknown mixing matrix and embedded in white Gaussian noise. These latent states are observed through a set of nonlinear measurements.

Mathematically, let $x_t \in \mathbb{R}^{d_x}$ be the set of latent correlated processes and $y_t \in \mathbb{R}^{d_y}$ the data observed at time t defined by a nonlinear equation, i.e.,

$$x_t = Ax_{t-1} + u_t, \quad \text{state equation,} \tag{1}$$

$$y_t = h(x_t, v_t), \quad \text{observation equation,} \tag{2}$$

where $t = 1, 2, \cdots$; $u_t \in \mathbb{R}^{d_x}$ is a zero mean Gaussian vector with covariance matrix C_u; $A \in \mathbb{R}^{d_x \times d_x}$ is a mixing matrix; $v_t \in \mathbb{R}^{d_v}$ denotes an independent noise process; and $h(x_t, v_t) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_v} \to \mathbb{R}^{d_y}$, is some generic function. We do not restrict the form of $h(\cdot, \cdot)$, and thus allow for data observed via nonlinear functions of the state. The only requirement is that the likelihood, i.e., $f(y_t|x_t)$, is computable up to a proportionality constant.

Given a set of observations $y_{1:t} \equiv \{y_1, y_2, \cdots, y_t\}$, the goal is to sequentially infer the evolution of the correlated states, x_t, by estimating the posterior distribution, i.e., $f(x_t|y_{1:t})$. We do so by updating $f(x_t|y_{1:t})$ to $f(x_{t+1}|y_{1:t+1})$, by using each new observation y_{t+1}. The new density is derived by employing the Bayes' rule

$$f(x_{t+1}|y_{1:t+1}) \propto f(y_{t+1}|x_{t+1}) \int f(x_{t+1}|x_t) f(x_t|y_{1:t}) \mathrm{d}x_t. \tag{3}$$

This work has been supported by the National Science Foundation under Award CCF-1617986 (M. F. Bugallo) and Award CCF-1618999 (P. M. Djurić).

The analytical solution to (3) is attainable in rare cases, e.g., when dealing with Gaussian noises and linear functions, which is the celebrated Kalman filter [8]. Since we do not restrict ourselves to such assumptions, we resort to PF methods, which can overcome these additional challenges.

3. STANDARD PARTICLE FILTERING

PF methods approximate the densities of interest by discrete random measures, i.e.,

$$f(x_t) \approx f^M(x_t) = \sum_{m=1}^{M} w_t^{(m)} \delta(x_t - x_t^{(m)}), \tag{4}$$

where $x_t^{(m)}$ are samples (i.e., particles) drawn from a proposal distribution and $w_t^{(m)}$ are the weights associated to them, which sum up to one. The total number of particles that form the random measure is denoted by M.

With PF, the discrete random measure is recursively updated upon arrival of a new observation. The update consists of generation (sampling) of new particles and computation of their corresponding weights. A final step, known as resampling, is implemented to avoid particle attrition due to the sequential updates over time.

The key for a successful PF method is to propagate particles in regions of the state's space where most of the information is located. The two most common proposal functions are the transition density $f(x_{t+1}|x_{1:t})$ and the optimal importance function $f(x_{t+1}|x_{1:t}, y_{1:t+1})$. The latter minimizes the variance of the resulting weights, conditioned on the available information, but it is often intractable in many problems of interest (as is the case here). Drawing samples from the transition density is easier and leads to particle weight updates given by $w_{t+1}^{(m)} \propto w_t^{(m)} f(y_{t+1}|x_{t+1})$. For the model in (1)-(2), the transition density of the states, when the mixing parameters are known, can be readily derived.

However, there are two main challenges to address in practical situations. On the one hand, it is unrealistic to assume knowledge of the mixing parameters and, on the other, the method suffers as the dimensionality of the state increases. Both problems are critical for PF methods, as the estimation of static parameters is troublesome [16, 17] and the curse of dimensionality undermines their applicability [18, 19].

Here, we deal with these challenges using a two-pronged approach. First, we exploit Rao-Blackwellization [14] to integrate out the static parameters and avoid generating particles for them. This statistical procedure guarantees reduced estimation variance and thus, allows for improved estimation of the state process. The technique has already been successfully applied in conjunction to SMC methods [20, 21, 22]. For further mitigation of the impact of increasing state dimensionality, we build upon the MPF approach [23, 15, 24], where one breaks the high-dimensional distribution of the complete state into smaller dimensional marginal ones.

The novelty of this paper is on the efficient combination of Rao-Blackwellization with MPF methods for inference of correlated latent states with unknown mixing parameters.

4. THE PROPOSED METHOD

We first derive the Rao-Blackwellized density for the problem of interest and then elaborate on the details of the proposed MPF method.

For the model in (1)-(2), given knowledge of the mixing parameters A and C_u, and previous states x_t, the transition density is a multivariate Gaussian

$$f(x_{t+1}|x_t, A, C_u) = \mathcal{N}(x_{t+1}|Ax_t, C_u). \tag{5}$$

In practice, it is unrealistic to assume that the true values of the mixing parameters are known and thus, we marginalize them out. Let us define the following historical data matrices

$$\begin{cases} X_t = [x_1 x_2 \cdots x_t] \in \mathbb{R}^{d_x \times t}, \\ Z_t = [x_0 x_1 \cdots x_{t-1}] \in \mathbb{R}^{d_x \times t}. \end{cases} \tag{6}$$

It can be seen that the estimate of the mixing matrix at time t is $\hat{A}_t = X_t Z_t^\top (Z_t Z_t^\top)^{-1}$ [25]. Instead of resorting to point estimates, we adopt Rao-Blackwellization. That is, we integrate out the unknown mixing parameters A and C_u to derive the transition density.

To that end, we proceed as in [26] and obtain the density of x_{t+1}, given the past data $x_{1:t}$, as a multivariate t-distribution

$$f(x_{t+1}|x_{1:t}) = \mathcal{T}_{\nu_{t+1}}\left(x_{t+1}|\mu_{t+1}, R_{t+1}\right), \tag{7}$$

where ν_{t+1} denotes degrees of freedom, $\mu_{t+1} \in \mathbb{R}^{d_x}$ is the location parameter, and $R_{t+1} \in \mathbb{R}^{d_x \times d_x}$ represents the scale matrix [27]. These distribution parameters are computed by

$$\begin{cases} \nu_{t+1} = t - d_x - d_x + 1, \\ \mu_{t+1} = \hat{A}_t z_{t+1}, \\ R_{t+1} = \frac{\left(X_t - \hat{A}_t Z_t\right)\left(X_t - \hat{A}_t Z_t\right)^\top}{\nu_{t+1}\left(1 - x_t^\top (Z_{t+1} Z_{t+1}^\top)^{-1} x_t\right)}. \end{cases} \tag{8}$$

This transition density (which does not depend on the mixing parameters A and C_u) allows for implementation of a standard PF method for inference of correlated processes when the mixing parameters are unknown.

In order to apply the previous scheme to systems of higher dimensions, we are interested on an MPF-type solution. In MPF, one decomposes the space of the latent state into subspaces and assigns a particle filter to each of them. Consequently, instead of obtaining the full joint filtering density, one focuses on a set of marginal filtering densities of a lower dimensional subspaces. Mathematically, the space of the state $x_t \in \mathbb{R}^{d_x}$ is partitioned into N disjoint subspaces $x_{n,t} \in \mathbb{R}^{d_{x_n}}$, $n = 1, 2, \cdots, N$, where $\prod_{n=1}^{N} \mathbb{R}^{d_{x_n}} = \mathbb{R}^{d_x}$. For each of the subpaces $x_{n,t}$, a separate PF is run to sequentially obtain $f(x_{n,t+1}|x_t, y_{t+1})$.

However, the correlation between states as in (1)-(2) is a hurdle for the MPF, as the propagation and weighting steps of each PF depend on all the other states. To overcome such difficulties, one can share information across the individual PFs [24]. We note that, the approach in [24] relies on knowledge of the mixing parameters, which is not the case here. Therefore, the critical issue is how to combine the random measure of each PF with the information shared by other PFs when the mixing parameters A and C_u are unknown. We overcome these challenges by leveraging the Rao-Blackwellized density in (7) within a MPF method.

Let us denote each subspace vector as $x_{n,t} \in \mathbb{R}^{d_{x_n}}$ and the overall state information as $z_t = \left(x_{1,t}\ x_{2,t}\ \cdots\ x_{N,t}\right)^\top \in \mathbb{R}^{d_x}$, and define

$$\begin{cases} X_{n,t} = [x_{n,1}\ x_{n,2}\ \cdots\ x_{n,t}] \in \mathbb{R}^{d_{x_n} \times t}, \\ Z_t = [x_0\ x_1\ \cdots\ x_{t-1}] \in \mathbb{R}^{d_x \times t}. \end{cases} \tag{9}$$

For each subspace, the transition density when no knowledge of the mixing parameters is assumed, is given by

$$f(x_{n,t+1}|x_{1:t}) = \mathcal{T}_{\nu_{n,t+1}}\left(x_{n,t+1}|\mu_{n,t+1}, R_{n,t+1}\right),$$
$$\text{with} \begin{cases} \nu_{n,t+1} = t - d_{x_n} - d_x + 1, \\ \hat{A}_{n,t} = X_{n,t} Z_t^\top (Z_t Z_t^\top)^{-1} \\ \mu_{n,t+1} = \hat{A}_{n,t} z_{t+1}, \\ R_{n,t+1} = \frac{\left(X_{n,t} - \hat{A}_{n,t} Z_t\right)\left(X_{n,t} - \hat{A}_{n,t} Z_t\right)^\top}{\nu_{n,t+1}\left(1 - z_{t+1}^\top (Z_{t+1} Z_{t+1}^\top)^{-1} z_{t+1}\right)}. \end{cases} \tag{10}$$

The above density illustrates how, for each subspace $x_{n,t}$, the local information $X_{n,t}$ is fused with the overall knowledge via z_{t+1} and Z_t.

Within MPF methods, several alternatives can be considered on how to share information across PFs. In this paper, we simply resort to the state estimates provided by each PF, i.e., $z_t = \hat{x}_t = \left(\hat{x}_{1,t}\ \hat{x}_{2,t} \cdots \hat{x}_{N,t}\right)^\top$. Sharing these estimates provides relevant information while incurring minimal communication cost.

In summary, at time instant t, the Rao-Blackwellized MPF method for inference of correlated states with unknown mixing matrices considers a random measure $f^M(x_{n,t})$ per subspace $x_{n,t}$

$$f^M(x_{n,t}) = \sum_{m=1}^{M} w_t^{(m)} \delta\left(x_{n,t} - x_{n,t}^{(m)}\right), n = 1, \cdots, N,$$

and all state estimates up to time instant t, included in Z_t and z_{t+1}. Upon the reception of a new observation, at time instant $t+1$, the algorithm proceeds as follows:

1. Propagate the particles for each subspace by sampling from the transition density

$$x_{n,t+1} \sim f(x_{n,t+1}|x_{1:t}^{(m)}) = \mathcal{T}_{\nu_{n,t+1}}\left(x_{n,t+1}|\mu_{n,t+1}^{(m)}, R_{n,t+1}^{(m)}\right),$$
$$\text{with} \begin{cases} \nu_{n,t+1} = t - d_{x_n} - d_x + 1, \\ \hat{A}_{n,t}^{(m)} = X_{n,t}^{(m)} Z_t^\top (Z_t Z_t^\top)^{-1}, \\ \mu_{n,t+1}^{(m)} = \hat{A}_{n,t}^{(m)} z_{t+1}, \\ R_{n,t+1}^{(m)} = \frac{\left(X_{n,t}^{(m)} - \hat{A}_{n,t}^{(m)} Z_t\right)\left(X_{n,t}^{(m)} - \hat{A}_{n,t}^{(m)} Z_t\right)^\top}{\nu_{n,t+1}\left(1 - z_{t+1}^\top (Z_{t+1} Z_{t+1}^\top)^{-1} z_{t+1}\right)}. \end{cases}$$

   Note that the sufficient statistics depend on each previous stream of (resampled) particles $\overline{x}_{n,t}^{(m)}$.

2. Compute the non-normalized weights for the drawn particles according to

$$\widetilde{w}_{n,t+1}^{(m)} \propto f(y_{t+1}|x_{n,t+1}^{(m)}),$$

   and normalize them to obtain a new random measure per subspace

$$f^M(x_{n,t+1}) = \sum_{m=1}^{M} w_{n,t+1}^{(m)} \delta\left(x_{n,t+1} - x_{n,t+1}^{(m)}\right).$$

3. Compute the state estimates $z_t = \left(\hat{x}_{1,t}\ \hat{x}_{2,t}\ \cdots\ \hat{x}_{N,t}\right)^\top$, from each subspace random measure

$$\hat{x}_{n,t} = \sum_{m=1}^{M} x_{n,t+1}^{(m)} w_{n,t+1}^{(m)}. \tag{11}$$

   This information is shared amongst all N PFs for subsequent time instants.

4. Perform resampling of each state subspace (to avoid sample degeneracy) by drawing from a categorical distribution defined by the random measure

$$\overline{x}_{n,t}^{(m)} \sim f^M(x_{n,t}), \text{ where } m = 1, \cdots, M.$$

   We note that this step does not have to be performed at every time instant.

5. SIMULATION RESULTS

We evaluate the proposed method by simulating an illustrative practical scenario, where a set of observed time-series are driven by several latent correlated processes. Specifically, we model a 6-dimensional state process ($d_x = 6$) with standard multivariate Gaussian independent and identically distributed innovations u_t, i.e.,

$$\begin{aligned} x_t &= Ax_{t-1} + u_t, \\ \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \\ x_{5,t} \\ x_{6,t} \end{pmatrix} &= \begin{pmatrix} 0.7 & 0.12 & 0.08 & 0 & 0 & 0 \\ 0 & 0.7 & 0.12 & 0.08 & 0 & 0 \\ 0.15 & 0 & 0.7 & 0.12 & 0.08 & 0 \\ 0 & 0.15 & 0 & 0.70 & 0.12 & 0.08 \\ 0 & 0 & 0.15 & 0 & 0.70 & 0.12 \\ 0 & 0 & 0 & 0.15 & 0 & 0.7 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \\ x_{4,t-1} \\ x_{5,t-1} \\ x_{6,t-1} \end{pmatrix} + u_t. \end{aligned} \tag{12}$$

The latent process x_t drives the mean and volatility of three observed time-series ($d_y = 3$)

$$\begin{cases} y_{1,t} = x_{1,t} + e^{x_{2,t}/2} v_{1,t}, \\ y_{2,t} = x_{3,t} + e^{x_{4,t}/2} v_{2,t}, \\ y_{3,t} = x_{5,t} + e^{x_{6,t}/2} v_{3,t}, \end{cases} \tag{13}$$

where $v_{1,t}$, $v_{2,t}$ and $v_{3,t}$ are independent standard Gaussian noises.

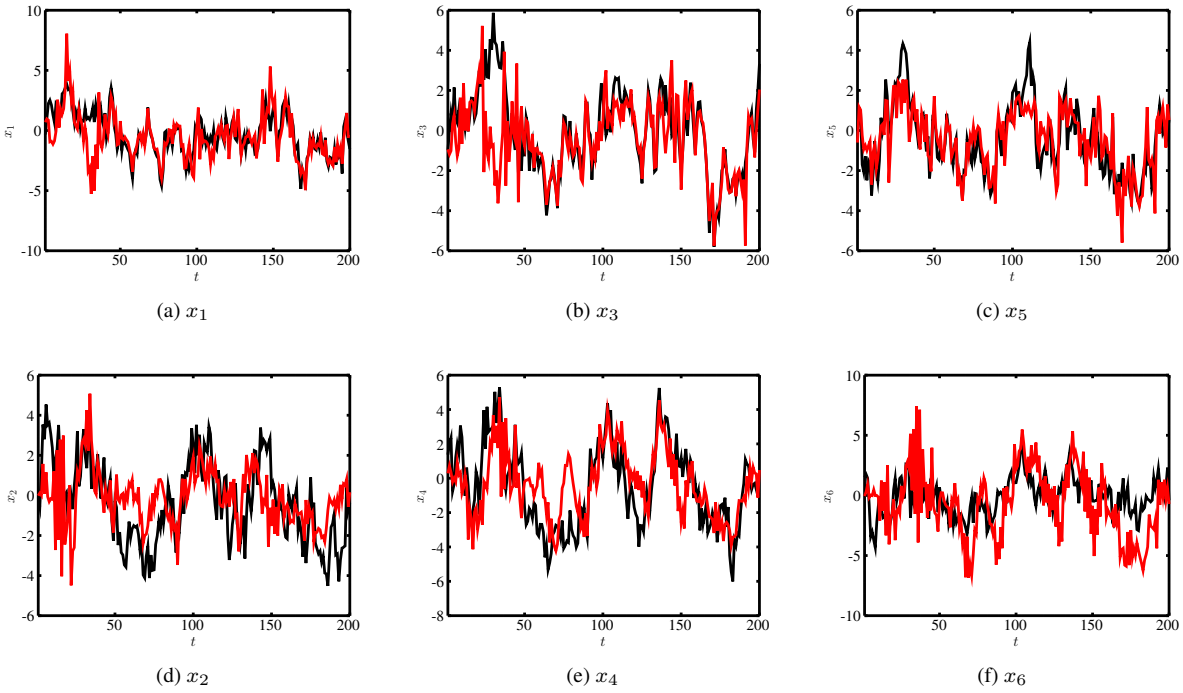
A practical motivation for this problem setup can be found in finance, where each observation describes the price evolution of an asset, whose dynamics are captured by a trend and a log-volatility [28]. Each mean and log-volatility process not only follows its own dynamics, but is also correlated via the mixing matrix A. In practice, how the properties of the assets are correlated amongst each other is hard to know beforehand and thus, the relevance of this problem for the evaluation of the proposed method.

We simulated the above system and implemented the proposed method. We divided the state into subspaces, each with a mean and log-volatility pair associated with a particular asset. This results in $N = 3$ PFs: $y_{n,t}$ with $x_{2n-1,t}$ and x_{2n} for $n = 1, 2, 3$, each with $M = 1,000$ particles.

We show in Fig. 1 the estimation performance of the method described in Section 4. We observe that the Rao-Blackwellized MPF is able to successfully track the latent state. Note how the estimation is more accurate for the mean states ($x_{1,t}$, $x_{3,t}$ and $x_{5,t}$) than the log-volatility processes ($x_{2,t}$, $x_{4,t}$ and $x_{6,t}$).

We point out that the accuracy of the method improves as time evolves (see estimates after $t = 100$ in Fig. 1). The reasoning behind this is that, with time, we observe more data and thus, both the state estimates, and the computed sufficient statistics in (10) become more precise. That is, the more data that are observed, the better the state estimates are, and consequently, the more precise the posterior parameters become (which in turn improves the state estimation via more accurate sufficient statistics).

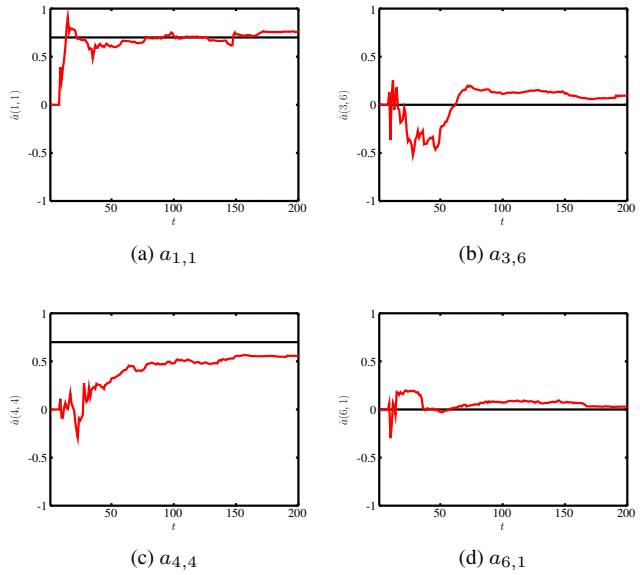
This improvement on the sufficient statistics is plotted, for some components of the mixing matrix A, in Fig. 2. There, we show how the mean estimates improve over time and become quite accurate after an unstable initial period.



(a) x_1 (b) x_3 (c) x_5

(d) x_2 (e) x_4 (f) x_6

Fig. 1: True state process (in black) and estimated (in red).



(a) $a_{1,1}$ (b) $a_{3,6}$

(c) $a_{4,4}$ (d) $a_{6,1}$

Fig. 2: Parameter true value (in black) and estimated (in red).

We further evaluated the proposed Rao-Blackwellized MPF method and compared it to other state-of-the-art PF methods. In Table 1 we provide the state estimation mean squared error (MSE) averaged over 50 realizations of 100-long time-series. There, we compare the performance of the proposed method to the standard PF (SPF) algorithm and the MPF method, when the mixing parameters are known (which serve as benchmarks of the proposed method).

	State estimation error (MSE)		
	SPF, known A, C_u	MPF, known A, C_u	MPF, unknown A, C_u
x_1	0.717278492	0.695521012	1.44317724
x_2	1.71300497	1.58838685	4.07465372
x_3	0.730489575	0.727382485	1.49508191
x_4	1.97312377	1.85175375	5.56804227
x_5	0.801862576	0.775937176	1.4138755
x_6	1.86307229	1.70220069	4.62055982

Table 1: MSE performance of different PF alternatives.

It is apparent that the MPF method provides estimates that are more accurate than the SPF approach for the known parameter case. When dealing with unknown mixing parameters, as expected, the estimation accuracy deteriorates. However, due to the implemented Rao-Blackellization procedure, the proposed method is still able to track the latent correlated states. We emphasize that in practical scenarios there will be no knowledge of the mixing parameters, but the proposed method is still able to track the latent states accurately.

6. CONCLUSIONS

In this paper, a new Rao-Blackwellized multiple particle filtering scheme is proposed for systems with hidden states correlated through unknown mixing parameters. The proposed solution relies on the multiple particle filtering concept, where each of the filter works on a subspace of the overall state. In addition, the mixing parameters are marginalized per subspace, and therefore, constitute a reduced burden in the performance. Computer simulations reveal a good estimation accuracy and reduced computational complexity of the proposed method.

7. REFERENCES

[1] Brian D. O. Anderson and John B. Moore, *Optimal Filtering*, Dover Books on Electrical Engineering. Dover Publications, 2012.

[2] James Durbin and Siem Jan Koopman, *Time Series Analysis by State-Space Methods: Second Edition*, Oxford Statistical Science Series. Oxford University Press, 2 edition, 2012.

[3] Mike West and Jeff Harrison, *Bayesian Forecasting and Dynamic Models*, Springer Series in Statistics. Springer New York, 1999.

[4] Tohru Ozaki, *Time Series Modeling of Neuroscience data*, Chapman & Hall/CRC Interdisciplinary Statistics, 2012.

[5] Anthony Stathopoulos and Matthew G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 121 – 135, 2003.

[6] Robert H. Shumway and David S. Stoffer, *Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics)*, Springer, 3rd edition, 2010.

[7] Olivier Basdevant, "On applications of state-space modelling in macroeconomics," *SSRN Electronic Journal*, 2003.

[8] Rudolph Emil Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[9] Arnaud Doucet, Nando De Freitas, and Neil Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[10] Petar M. Djurić and Mónica F. Bugallo, *Particle Filtering*, chapter 5, pp. 271–331, John Wiley & Sons, Inc., 2010.

[11] Branko Ristic, Sanjeev Arulampalam, and Neil Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2004.

[12] Drew Creal, "A survey of Sequential Monte Carlo methods for economics and finance," *Econometric Reviews*, vol. 31, no. 3, pp. 245–296, 2012.

[13] Neil J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Radar and Signal Processing, IEEE Proceedings*, vol. 140, no. 2, pp. 107 –113, 4 1993.

[14] George Casella and Christian P. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[15] Petar M. Djurić and Mónica F. Bugallo, "Particle filtering for high-dimensional systems," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on*, Dec 2013, pp. 352–355.

[16] Jane Liu and Mike West, *Combined Parameter and State Estimation in Simulation Based Filtering*, chapter 10 in "Sequential Monte Carlo Methods in Practice", pp. 197–224, Springer, 2001.

[17] Nikolas Kantas, Arnaud Doucet, Sumeetpal S. Singh, Jan Maciejowski, and Nicolas Chopin, "On particle methods for parameter estimation in state-space models," *Statistical Science*, vol. 30, no. 3, pp. 328–351, 8 2015.

[18] Fred Daum and Jim Huang, "Curse of dimensionality and particle filters," in *2003 IEEE Aerospace Conference Proceedings*, 3 2003, vol. 4, pp. 1979–1993.

[19] Thomas Bengtsson, Peter Bickel, and Bo Li, *Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems*, vol. 2 of *Collections*, pp. 316–334, Institute of Mathematical Statistics, 2008.

[20] Arnaud Doucet, Neil J. Gordon, and Vikram Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *Signal Processing, IEEE Transactions on*, vol. 49, no. 3, pp. 613–624, 3 2001.

[21] Thomas Schön, Fredrik Gustafsson, and Per-Johan Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *Signal Processing, IEEE Transactions on*, vol. 53, no. 7, pp. 2279–2289, 7 2005.

[22] Iñigo Urteaga, Mónica F. Bugallo, and Petar M. Djurić, "Sequential monte carlo sampling for correlated latent long-memory time-series," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 6580–6584.

[23] Pau Closas and Mónica F. Bugallo, "Improving accuracy by iterated multiple particle filtering," *IEEE Signal Processing Letters*, vol. 19, no. 8, pp. 531–534, 8 2012.

[24] Petar M. Djurić and Mónica F. Bugallo, "Multiple particle filtering with improved efficiency and performance," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4 2015, pp. 4110–4114.

[25] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*, vol. 1 of *Fundamentals of Statistical Signal Processing*, Prentice-Hall PTR, 1993.

[26] Seymour Geisser, "Bayesian-estimation in multivariate analysis," *Annals of American Statistics*, vol. 36, no. 1, pp. 150–159, 1965.

[27] José M. Bernardo and Adrian F.M. Smith, *Bayesian Theory*, Wiley Series in Probability and Statistics. Wiley, 2009.

[28] Neil Shephard, Ed., *Stochastic Volatility: Selected Readings*, Advanced texts in econometrics. Oxford University Press, 2005.