PARTICLE FILTERING OF ARMA PROCESSES OF UNKNOWN ORDER AND PARAMETERS

Iñigo Urteaga and Petar M. Djurić

Department of Electrical & Computer Engineering
Stony Brook University, Stony Brook, NY 11794 USA
{inigo.urteaga, petar.djuric}@stonybrook.edu

ABSTRACT

This paper considers inference on the widely used state-space models described by hidden ARMA state processes of unknown order observed via non-linear functions of the states. We propose a particle filtering method for sequentially inferring the unknown ARMA time-series by Rao-Blackwellization of all the static unknowns. Our method does not rely either on any assumption on the model order or on the static ARMA and state innovation parameters. Consequently, when the ARMA model order is unknown, it can be used without a follow-up model selection procedure. Extensive simulation results validate the proposed method across different ARMA models.

***Index Terms*—** State-space models, time-series, ARMA models, particle filtering, Rao-Blackwellization.

1. INTRODUCTION

This paper considers the processing of observations that are functions of a hidden signal. The objective is to estimate the unknown time-varying signal x_t, given the observations y_t. Models that describe this setting are known as state-space models, and they are often used in many signal processing applications, including speech processing, communications, finance and neuroscience [1]. It is not surprising then that the inference of the hidden states in state-space models has been a widely studied problem. In the case of linear state-space models with additive Gaussian noises, the optimal solution is the celebrated Kalman filter [2]. When the models deviate from the assumptions of linearity and Gaussianity, the processing of the data under such models requires alternative solutions. Sequential Monte Carlo methods, also known as particle filters (PFs), are one of the alternatives. They already have a nice track record in diverse disciplines [3, 4, 5]. Motivated by applications in finance engineering such as representation of asset returns [6, 7] or in neuroscience in the analysis of neural signals [8], we are interested in inference of hidden linear time series observed through non-linear functions.

In this paper, the time-dependency of the hidden signal is characterized by an Autoregressive Moving-Average (ARMA) model, where the current state depends on both the previous values of the signal and a state noise (i.e., innovations). Due to its flexible parameterization, the ARMA model can be fit to any linear time series with high accuracy. An ARMA process is described as ARMA(p, q), where p is the order of the auto-regressive (AR) part, and q, the order of the moving-average (MA) part. Due to the non-linearities induced by the MA part, the sequential estimation of the hidden states modeled as ARMA processes is challenging. Besides, if the observations are also non-linear functions of the states, one has to look for techniques that can handle such non-linearities. Particle filtering methods have the capacity to overcome these challenges. Dealing with unknown state parameters does nothing but complicate the problem further, as PF methods require special care in handling them. Finally, in practice, we usually have no knowledge of the model orders p and q, which confounds the problem additionally. The latter entails that one would need to invoke some model selection procedure as part of the estimation process.

Most of the literature on PFs has focused on AR state processes with non-linear observations [9, 10, 11], where the inference has been performed under the assumption of both known and unknown parameters. When dealing with unknown AR parameters, the PFs resort to joint estimation of the states and the parameters (in other words, particles are generated for all the states and parameters [9, 12]). However, much less work can be found on applying particle filtering to state-space models with full ARMA processes in the state equation. In [13] for example, given that the model order is known, estimation of both the state and the ARMA parameters was proposed, under the assumption of time-varying parameters. It is important to recall that PFs suffer when dealing with fixed model parameters [12] and, consequently, there has been an increasing effort to derive new methods for addressing the problem (e.g., the use of artificial parameter evolution [14], kernel smoothing [12], or density assisted PFs [15]). Some recent work has avoided the limitations of the parameter estimation by resorting to Rao-Blackwellization of the nuisance parameters. The performance of these new PFs is improved because the generated particles are sampled from a space that has a reduced dimensionality. This can readily be done for AR models when the noise in the state equation is Gaussian [16]. The case of ARMA models is more complicated, thus approximations like numerical Rao-Blackwellization have been presented [17]. In all these papers, the model orders were known.

Here, we propose a PF that analytically Rao-Blackwellizes all the unknown ARMA parameters and the state noise. Interestingly, the proposed Rao-Blackwellization does not require knowledge of the model order. Thus, our method has the clear advantage in situations when the model order is unknown. Also, due to the Rao-Blackwellization, it avoids the limitations of particle-based parameter estimation, and provides state estimates with reduced variance. All in all, the suggested PF advances the state of the art because it (1) considers general ARMA(p, q) models, (2) does not require knowledge of model order, (3) does not assume knowledge of neither the ARMA nor the state innovation parameters, and (4) resorts to the joint Rao-Blackwellization of all the unknowns.

The rest of the paper is organized as follows. In Section 2, we formulate the problem. We propose the new particle filtering solu-

This work was supported by NSF under Award CCF-1320626.

tion in Section 3. Section 4 contains the results of extensive computer simulations where we demonstrate the performance of the proposed method. We make our final conclusions in Section 5.

2. PROBLEM FORMULATION

We are interested in a general class of state-space models where the state is modeled by a generic ARMA(p, q) process, with AR parameters a_i, $i = 1, \cdots, p$ and MA parameters b_j, $j = 1, \cdots, q$. The observation is a non-linear function $h(x_t, v_t)$ of the state. The state-space model is represented mathematically as follows:

$$\begin{aligned} x_t &= \textstyle\sum_{i=1}^{p} a_i x_{t-i} + \sum_{j=1}^{q} b_j u_{t-j} + u_t, \quad \text{state eq.} && (1)\\ y_t &= h(x_t, v_t), \qquad\qquad\qquad\qquad \text{observation eq.} && (2) \end{aligned}$$

where the symbols u_t and v_t represent white Gaussian processes independent of each other. We assume that we do not know the order of the ARMA model. Consequently, we also do not know its parameters. Finally, the parameters of the state innovations (i.e., mean and variance), too, are unknown. As for the observation equation, we do not enforce any restrictions except that the function $h(x_t, v_t)$ produces a likelihood $f(y_t|x_t)$ that is computable up to a proportionality constant.

Given the observations $y_{1:t} \equiv \{y_1, y_2, \cdots, y_t\}$, we want to sequentially estimate the posterior density of x_t, $f(x_t|y_{1:t})$. To do so, we resort to particle filtering [18], a well known approach for inference in non-linear/non-Gaussian state-space models. PFs approximate the posterior density of the states given all the available observations by

$$f(x_t|y_{1:t}) \approx \sum_{m=1}^{M} w_t^{(m)} \delta(x_t - x_t^{(m)}), \tag{3}$$

where $x_t^{(m)}$ are particles drawn from a proposal distribution, M is the number of particles, $w_t^{(m)}$ are the weights associated to the particles, and $\delta(\cdot)$ is the Dirac delta function.

The method proceeds sequentially, i.e., $f(x_t|y_{1:t})$ is obtained from $f(x_{t-1}|y_{1:t-1})$ according to

$$\begin{aligned} f(x_t|y_{1:t}) &\propto f(y_t|x_t) \int f(x_t|x_{1:t-1}) f(x_{1:t-1}|y_{1:t-1}) \mathrm{d}x_{1:t-1} \\ &\approx f(y_t|x_t) \sum_{m=1}^{M} w_{t-1}^{(m)} f(x_t|x_{1:t-1}^{(m)}). && (4) \end{aligned}$$

In obtaining $f(x_t|y_{1:t})$, the challenge is the derivation of $f(x_t|x_{1:t-1})$ for an ARMA model when the parameters $(p, q, a_i, b_j, \mu_u, \sigma_u^2)$ are unknown. In the next section we show how this transition density is obtained by Rao-Blackwellization of all the nuisance parameters.

3. THE PROPOSED METHOD

In the formulation of the problem, we have introduced state equations described by general ARMA models. For the following derivation, we assume weak-sense stationarity of the ARMA process, thus requiring that the first and second moments exist and are constant with respect to time: i.e., $\mathbb{E}\{x_t\} = \mu$ and $Cov(x_t, x_{t-\tau}) = \gamma(\tau)$. These conditions imply that the mean does not vary with time and that the autocovariance of the process is a function only of the time-difference and not the actual time instants.

For such stationary ARMA models, the joint distribution of the series at time instant t can be written as

$$f(x_{1:t}|\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) = f(\mathbf{x}_t|\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) = \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), \tag{5}$$

$$\text{with} \begin{cases} \mathbf{x}_t = \begin{pmatrix} x_t & x_{t-1} & x_{t-2} & \cdots & x_2 & x_1 \end{pmatrix}^\top \\ \boldsymbol{\mu}_t = \mu \mathbf{1}_t \\ \boldsymbol{\Sigma}_t = \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(t-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(t-2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(t-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma(t-1) & \gamma(t-2) & \gamma(t-3) & \cdots & \gamma(0) \end{pmatrix}, \end{cases}$$

where $\mathbf{1}_t$ is a $t \times 1$ vector of ones and $\mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ represents a multivariate Gaussian distribution with a mean vector $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}_t$. Due to stationarity, the mean $\boldsymbol{\mu}_t$ is equal for all time instants with its value μ being dependent on the state noise mean μ_u and the ARMA parameter set (a_i, b_j, p, q). Similarly, the covariance matrix $\boldsymbol{\Sigma}_t$ is a function of the time-lag $\tau = \{0, 1, \cdots, t-1\}$, the ARMA parameter set and the state noise variance σ_u^2.

Whenever the parameters of the model are unknown, the determination of these sufficient statistics is not possible. Since the goal is to infer the hidden state x_t and we consider the parameters to be of secondary importance (they are nuisance), we resort to Rao-Blackwellization.

Let us assume the following hierarchical parametric model for the vector $\mathbf{x}_t$ of an ARMA(p, q) process at time instant t:

$$\begin{aligned} f(\boldsymbol{\Sigma}_t|\boldsymbol{\Lambda}_t, \nu_t) &= \mathrm{IW}_{\nu_t}(\boldsymbol{\Lambda}_t), && (6)\\ f(\boldsymbol{\mu}_t|\boldsymbol{\eta}_t, \boldsymbol{\Sigma}_t, \kappa_t) &= \mathcal{N}\left(\boldsymbol{\eta}_t, \frac{\boldsymbol{\Sigma}_t}{\kappa_t}\right), && (7)\\ f(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t|\boldsymbol{\eta}_t, \kappa_t, \boldsymbol{\Lambda}_t, \nu_t) &= \mathrm{NIW}(\boldsymbol{\eta}_t, \kappa_t, \boldsymbol{\Lambda}_t, \nu_t), && (8)\\ f(\mathbf{x}_t|\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) &= \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), && (9) \end{aligned}$$

where $\mathrm{IW}_{\nu_t}(\boldsymbol{\Lambda}_t)$ is the inverse Wishart distribution with a scale matrix $\boldsymbol{\Lambda}_t$ and $\nu_t \geq t$ degrees of freedom, and $\mathrm{NIW}(\boldsymbol{\eta}_t, \kappa_t, \boldsymbol{\Lambda}_t, \nu_t)$ is the normal-inverse-Wishart distribution with location $\boldsymbol{\eta}_t$, an inverse scale matrix $\boldsymbol{\Lambda}_t$, and real parameters $\kappa_t > 0$ and ν_t.

We now proceed with deriving the Rao-Blackwellized distribution of the joint state (the integration details can be found in [19] and [20]). We write

$$\begin{aligned} f(\mathbf{x}_t) &= \int_{\boldsymbol{\Sigma}_t} \int_{\boldsymbol{\mu}_t} f(\mathbf{x}_t|\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) f(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t|\boldsymbol{\eta}_t, \kappa_t, \boldsymbol{\Lambda}_t, \nu_t) \mathrm{d}\boldsymbol{\mu}_t \mathrm{d}\boldsymbol{\Sigma}_t \\ &= \mathcal{T}_{\nu_t - t + 1}\left(\boldsymbol{\eta}_t, \frac{\boldsymbol{\Lambda}_t(1+\kappa_t)}{\kappa_t(\nu_t - t + 1)}\right), \end{aligned} \tag{10}$$

where $\mathcal{T}_{\nu_t - t + 1}\left(\boldsymbol{\eta}_t, \frac{\boldsymbol{\Lambda}_t(1+\kappa_t)}{\kappa_t(\nu_t - t + 1)}\right)$ stands for a multivariate t-distribution where $\boldsymbol{\eta}_t$ is its location vector, $\frac{\boldsymbol{\Lambda}_t(1+\kappa_t)}{\kappa_t(\nu_t - t + 1)}$ is the scale matrix, and $\nu_t - t + 1$ represents the degrees of freedom.

Let us further rewrite the obtained joint multivariate t-distribution as

$$\begin{bmatrix} x_t \\ \mathbf{x}_{t-1} \end{bmatrix} \sim \mathcal{T}_{\nu_t - t + 1}\left(\boldsymbol{\eta}_t, \frac{\boldsymbol{\Lambda}_t(1+\kappa_t)}{\kappa_t(\nu_t - t + 1)}\right), \tag{11}$$

$$\text{with parameters} \begin{cases} \boldsymbol{\eta}_t = \begin{bmatrix} \eta_t \\ \boldsymbol{\eta}_{t-1} \end{bmatrix} \\ \boldsymbol{\Lambda}_t = \begin{bmatrix} \lambda_t & \mathbf{l}_t^\top \\ \mathbf{l}_t & \mathbf{L}_t \end{bmatrix}. \end{cases}$$

Next, we write the conditional density $f(x_t|x_{1:t-1})$ and obtain [19]

$$f\left(x_t|x_{1:t-1}\right) = \mathcal{T}_{\nu_t - t+2}\left(\mu_{t|1:t-1}, \sigma^2_{t|1:t-1}\right), \tag{12}$$

with

$$\begin{cases} \mu_{t|1:t-1} = \eta + \mathbf{l}_t^\top \mathbf{L}_t^{-1}\left(\mathbf{x}_{t-1} - \boldsymbol{\eta}_{t-1}\right) \\ \sigma^2_{t|1:t-1} = \frac{h_{t|1:t-1}(1+\kappa_t)}{\kappa_t(\nu_t - t+1)}\left(\lambda_t - \mathbf{l}_t^\top \mathbf{L}_t^{-1}\mathbf{l}_t\right) \\ h_{t|1:t-1} = \frac{\nu_t - t+1}{\nu_t}\left[1 + \frac{\kappa_t\left(\mathbf{x}_{t-1} - \boldsymbol{\eta}_{t-1}\right)^\top}{(1+\kappa_t)}\mathbf{L}_t^{-1}\left(\mathbf{x}_{t-1} - \boldsymbol{\eta}_{t-1}\right)\right]. \end{cases}$$

We have thus shown that the transitive density of the Rao-Blackwellized ARMA time series is a Student's t-distribution, dependent on a set of hyper-parameters $\boldsymbol{\eta}_t$, $\boldsymbol{\Lambda}_t$, κ_t and ν_t. To provide meaningful values for these parameters, we resort to the Empirical Bayes Method [21]. Namely, we estimate the hyperparameters from the available data. Furthermore, we leverage the stationarity of the ARMA model to yield accurate estimation of the hyperparameters. Specifically, we compute the prior mean and covariance as (a) the empirical stationary mean $\hat{x_t}$ and, (b) the Toeplitz matrix formed by the empirical autocovariance function $\hat{\gamma}(\tau)$, $\tau = \{0, 1, \cdots, t-1\}$, respectively. More precisely, we use

$$\begin{cases} \boldsymbol{\eta}_t = \hat{x}_t \mathbf{1}_t, \\ \qquad \text{where } \hat{x_t} = \frac{1}{t}\sum_{i=1}^t x_i, \\ \boldsymbol{\Lambda}_t = \begin{pmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \hat{\gamma}(2) & \cdots & \hat{\gamma}(t-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(t-2) \\ \hat{\gamma}(2) & \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(t-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(t-1) & \hat{\gamma}(t-2) & \hat{\gamma}(t-3) & \cdots & \hat{\gamma}(0) \end{pmatrix} \\ \qquad \text{where } \hat{\gamma}(\tau) = \frac{1}{t-\tau}\sum_{i=1}^{t-\tau}\left(x_i - \hat{x}\right)\left(x_{i+\tau} - \hat{x}\right). \end{cases}$$

A lower bound for the hyperparameters κ_t and ν_t is provided by the dimensionality of the estimation problem. We must have $\kappa_t \geq 1$ and $\nu_t \geq t$ (as we are estimating t autocovariance values).

Finally, we note that the size of the covariance matrix increases with time (i.e., the dimensionality of the matrix is time dependent). Since this will impose a computational burden for long time series, an alternative is to truncate the covariance matrix to a predefined maximum lag τ_{max}. This truncation is justified by the short memory nature of ARMA processes, as most of the information is contained in few of the last samples. Even if the exact form of the covariance matrix for the ARMA(p, q) is in general intractable ([22], [23]), we can assert the following for the autocovariance function's form [1]:

- For AR processes, it decays exponentially.
- For MA processes, it is zero after the first q lags.
- For the general ARMA(p, q), it decays exponentially for lags bigger than $m = max(p, q)$.

We can therefore safely restrict our analysis of the covariance to a predetermined dimensionality $\tau_{max} >> m$.

3.1. Particle Filter

Based on the above derivation of the Rao-Blackwellized transition density, we now present how to implement the PF. Let us assume that at time instant t, we have the random measure $\chi_t = \left\{x_t^{(m)}, w_t^{(m)}\right\}$, where $m = 1, \cdots, M$. Then we proceed as follows:

1. Estimate the stationary sufficient statistics at time t:

$$\begin{cases} \hat{x_t}^{(m)} = \frac{1}{t}\sum_i x_i^{(m)} \\ \hat{\gamma}(\tau)^{(m)} = \frac{1}{(t-\tau)}\sum_{i=1}^{t-\tau}\left(x_i^{(m)} - \hat{x_t}^{(m)}\right)\left(x_{i+\tau}^{(m)} - \hat{x_t}^{(m)}\right) \\ \qquad \text{where } \tau = \{0, 1, \cdots, t-1\}. \end{cases}$$

2. Perform resampling of the state (to avoid sample degeneracy) by drawing from a categorical distribution defined by the random measure χ_t:

$$\overline{x}_t^{(m)} \sim \chi_t, \text{ where } m = 1, \cdots, M.$$

3. Propagate the particles by sampling from the conditional density, given the resampled streams:

$$x_{t+1}^{(m)} \sim f(x_{t+1}|\overline{x}_{1:t}^{(m)}) = \mathcal{T}_{\nu_{t+1}-t+1}\left(\mu_{t+1|1:t}^{(m)}, \left(\sigma_{t+1|1:t}^{(m)}\right)^2\right),$$

with parameters

$$\begin{cases} \mu_{t+1|1:t} = \hat{x_t}^{(m)} + \mathbf{l}_{t+1}^\top \mathbf{L}_{t+1}^{-1}\left(\overline{\mathbf{x}}_t - \boldsymbol{\eta}_t\right) \\ \sigma^2_{t+1|1:t} = \frac{h_{t+1|1:t}(1+\kappa_{t+1})}{\kappa_{t+1}(\nu_{t+1} - t)}\left(\lambda_{t+1} - \mathbf{l}_{t+1}^\top \mathbf{L}_{t+1}^{-1}\mathbf{l}_{t+1}\right) \\ h_{t+1|1:t} = \frac{\nu_{t+1} - t}{\nu_{t+1}}\left[1 + \frac{\kappa_{t+1}(\overline{\mathbf{x}}_t - \boldsymbol{\eta}_t)^\top}{(1+\kappa_{t+1})}\mathbf{L}_{t+1}^{-1}\left(\overline{\mathbf{x}}_t - \boldsymbol{\eta}_t\right)\right], \end{cases}$$

where

$$\begin{cases} \boldsymbol{\eta}_t = \hat{x}_t \mathbf{1}_t \\ \lambda_{t+1} = \hat{\gamma}(0)^{(m)} \\ \mathbf{l}_{t+1}^\top = \begin{pmatrix} \hat{\gamma}(1)^{(m)} & \hat{\gamma}(2)^{(m)} & \cdots & \hat{\gamma}(t-1)^{(m)} & 0 \end{pmatrix} \\ \mathbf{l}_{t+1} = \begin{pmatrix} \hat{\gamma}(1)^{(m)} & \hat{\gamma}(2)^{(m)} & \cdots & \hat{\gamma}(t-1)^{(m)} & 0 \end{pmatrix}^\top \\ \mathbf{L}_{t+1} = \begin{pmatrix} \hat{\gamma}(0)^{(m)} & \hat{\gamma}(1)^{(m)} & \cdots & \hat{\gamma}(t-2)^{(m)} & \hat{\gamma}(t-1) \\ \hat{\gamma}(1)^{(m)} & \hat{\gamma}(0)^{(m)} & \cdots & \hat{\gamma}(t-3)^{(m)} & \hat{\gamma}(t-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\gamma}(t-2)^{(m)} & \hat{\gamma}(t-3)^{(m)} & \cdots & \hat{\gamma}(0)^{(m)} & \hat{\gamma}(1)^{(m)} \\ \hat{\gamma}(t-1)^{(m)} & \hat{\gamma}(t-2)^{(m)} & \cdots & \hat{\gamma}(1)^{(m)} & \hat{\gamma}(0)^{(m)} \end{pmatrix}. \end{cases}$$

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

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

and normalize them to obtain a new random measure

$$\chi_{t+1} = \left\{x_{t+1}^{(m)}, w_{t+1}^{(m)}\right\}.$$

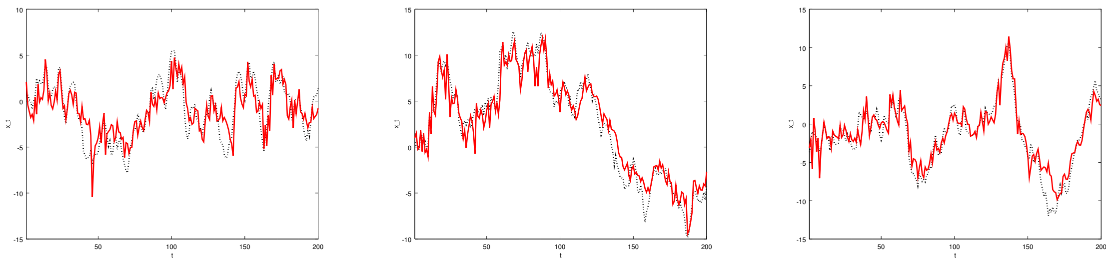
4. SIMULATION RESULTS

We evaluate the proposed method on the following stochastic volatility model:

$$x_t = \sum_{i=1}^p a_i x_{t-i} + \sum_{j=1}^q b_j u_{t-j} + u_t, \tag{13}$$

$$y_t = e^{(x_t/2)} v_t, \tag{14}$$

where the log-volatility x_t is an ARMA(p, q) process and the driving noises are independent and identically distributed standard Gaussian processes. This is a challenging problem not only because of the hidden unknown ARMA process, but also because of the nonlinearity in the observation equation: the goal is to estimate the log-volatility of a random process. Nevertheless, we show that the proposed PF successfully tracks x_t for different ARMA(p, q) models. Figure 1 shows the tracking results for a particular run of the following processes, each evolving for 200 time units:



(a) ARMA(1,1) (b) ARMA(2,2) (c) ARMA(4,2)

Fig. 1: Estimation (solid-red) of the hidden state (dotted-black)

- ARMA(1,1):
$$x_t = 0.9x_{t-1} + u_t + 0.3u_{t-1}.$$
- ARMA(2,2):
$$x_t = 0.8x_{t-1} + 0.15x_{t-2} + u_t + 0.5u_{t-1} + 0.3u_{t-1}.$$
- ARMA(4,2):
$$x_t = 0.6x_{t-1} + 0.2x_{t-2} + 0.05x_{t-3} + 0.1x_{t-4} + u_t + 0.5u_{t-1} + 0.3u_{t-1}.$$

Regarding the covariance truncation approach suggested in Section 3, the results shown in Fig. 2 endorse our proposal to truncate the autocovariance to a maximum lag (e.g., $\tau_{max} \approx 15$). The figure shows the MSE of the estimated process (for different ARMA model orders) as a function of τ_{max}. The results were obtained from 50 realizations with 500 particles. Not only the computational burden is dramatically reduced (time savings of more than an order of magnitude observed in our simulations), but (a) the autocovariance estimation becomes more reliable and (b) in predicting the next sample, the information gain provided by samples further in the past proves to be negligible.

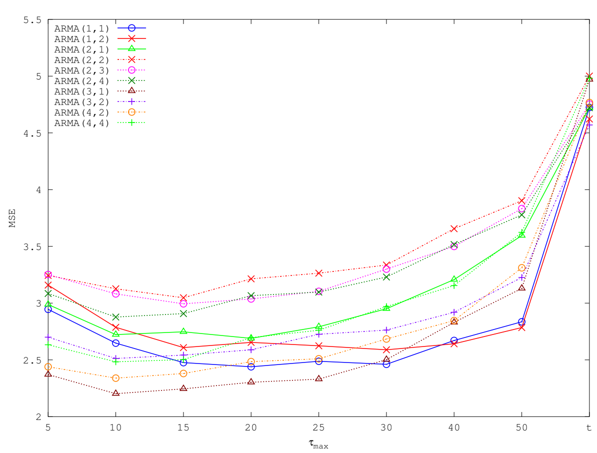


Fig. 2: Covariance truncation influence on ARMA MSE

Figure 2 also allows us to conclude that the performance of our suggested PF is consistent across different ARMA(p, q) models. Because no model order is assumed, the suggested PF provides good estimation accuracy independent of the model order. The explanation is that the information about x_t is solely extracted from the estimated covariance and past states. In doing so, we avoid exploring the parameter space with particles. As a result, the proposed method is much less sensitive to the curse of dimensionality.

In order to further illustrate the benefit of a model order independent PF, we provide a performance comparison between the proposed method and a Gaussian PF where the parameters and the state are jointly estimated, given the ARMA model order (`Gauss PF`). In this example, we assume that the hidden process has a zero mean. Thus, in our general method $\hat{x}_t = 0$ and $\kappa_t \to \infty$. As shown in Table 1, the performance of our PF pays the price of not knowing the model order, as the `Gauss PF` approach provides a slightly better performance. However, the proposed method is computationally an order of magnitude less demanding than the `Gauss PF`. Besides, when the model order is unknown, one would need to apply a number of `Gauss PF`s (with different assumptions of model orders), followed by a model selection procedure. Because no model order knowledge is required, this inconvenience is avoided by our proposed PF.

Table 1: PF comparison

ARMA model	State estimation error (MSE)	
	`Gauss PF`	`Suggested PF`
ARMA(1,1)	1.542	2.0693
ARMA(1,2)	2.2353	2.1961
ARMA(2,1)	1.8744	2.1588
ARMA(2,2)	1.6844	1.930
ARMA(2,4)	1.4482	1.8029
ARMA(3,1)	1.5831	1.9982
ARMA(3,2)	1.456	2.0118
ARMA(3,3)	1.0376	1.3216
ARMA(4,2)	1.6928	2.1136

5. CONCLUSIONS

We have proposed a new particle filter that tracks a hidden ARMA process of unknown order and parameters in the presence of nonlinear observations. The method is based on the Rao-Blackwellization of the static ARMA and state innovation parameters. Furthermore, the method does not require knowledge of the ARMA model order. The presented simulation results show the validity of the proposed method and suggest good estimation accuracy across different ARMA(p, q) models. Future work includes detailed study of the method when the state is modeled by multivariate ARMA processes.

6. REFERENCES

[1] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications: With R Examples (Springer Texts in Statistics)*, Springer, 2010.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, 1979.

[3] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, pp. 1032–1044, 1998.

[4] P. M. Djurić and M. F. Bugallo, *Particle Filtering*, chapter 5 in "Adaptive Signal Processing: Next Generation Solutions", pp. 271–331, John Wiley & Sons, Inc., 2010.

[5] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of Nonlinear Filtering*, vol. 12, pp. 656–704, 2009.

[6] S. T. Rachev, J. S. J. Hsu, B. S. Bagasheva, and F. J. Fabozzi, *Bayesian Methods in Finance*, John Wiley & Sons, 2008.

[7] D. E. Johnston, I. Urteaga, and P. M. Djurić, "Replication and optimization of hedge fund risk factor exposures," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 8712–8716.

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

[9] C. M. Carvalho and H. F. Lopes, "Simulation-based sequential analysis of Markov switching stochastic volatility models," *Computational Statistics & Data Analysis*, vol. 51, no. 9, pp. 4526 – 4542, 2007.

[10] H. F. Lopes and R. S. Tsay, "Particle filters and Bayesian inference in financial econometrics," *Journal of Forecasting*, vol. 30, no. 1, pp. 168–209, 2011.

[11] A. Doucet and V. B. Tadic, "Parameter estimation in general state-space models using particle methods," *Annals of the Institute of Statistical Mathematics*, vol. 55, pp. 409–422, 2003.

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

[13] R. Huang, H. Zheng, and E. E. Kuruoglu, "Time-varying ARMA stable process estimation using sequential Monte Carlo," *Signal, Image and Video Processing*, vol. 7, pp. 951–958, 2013.

[14] N. 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, 1993.

[15] P. M. Djurić, M. F. Bugallo, and J. Míguez, "Density assisted particle filters for state and parameter estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2004, vol. 2, pp. ii – 701–704.

[16] P. M. Djurić, M. Khan, and D. E. Johnston, "Particle filtering of stochastic volatility modeled with leverage," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 327–336, 2012.

[17] I. Urteaga and P. M. Djurić, "Estimation of ARMA state processes by particle filtering," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 8033–8037.

[18] 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, September 2003.

[19] K. P. Murphy, "Conjugate Bayesian analysis of the Gaussian distribution," Tech. Rep., University of British Columbia, 2007.

[20] A. Gelman, J. B. Carlin, H. S. Stern, D. D. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, third edition, 2013.

[21] B. P. Carlin and T. A. Louis, *Bayes and Empricial Bayes Methods for Data Analysis*, Chapman and Hall/CRC, 2000.

[22] P. Shaman, "Approximations for stationary covariance matrices and their inverses with application to ARMA models," *The Annals of Statistics*, vol. 4, no. 2, pp. 292–301, 1976.

[23] P. Shaman, "An approximate inverse for the covariance matrix of moving average and autoregressive processes," *The Annals of Statistics*, vol. 3, no. 2, pp. 532–538, 1975.