SEQUENTIAL MONTE CARLO SAMPLING FOR CORRELATED LATENT LONG-MEMORY TIME-SERIES

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

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

ABSTRACT

In this paper, we consider state-space models where the latent processes represent correlated mixtures of fractional Gaussian processes embedded in white Gaussian noises. The observed data are nonlinear functions of the latent states. The fractional Gaussian processes have interesting properties including long-memory, self-similarity and scale-invariance, and thus, are of interest for building models in finance and econometrics. We propose sequential Monte Carlo (SMC) methods for inference of the latent processes where each method is based on different assumptions about the parameters of the state-space model. The methods are extensively evaluated via simulations of the popular stochastic volatility model.

Index Terms— Sequential Monte Carlo, particle filtering, operator fractional Gaussian process, time-series, state-space models.

1. INTRODUCTION

We study the inference of latent correlated time-series that exhibit long-memory and self-similarity properties. We investigate state-space models where the latent states represent correlated mixtures of independent fractional Gaussian processes (fGps) embedded in white Gaussian noise and the observed data are nonlinear functions of the states.

State-space models provide a very flexible framework and therefore have been widely used in many signal processing applications, such as speech processing, communications, finance, and neuroscience [1, 2]. Motivated by applications in finance and econometrics [3], we are particularly interested in inference of time-series observed through nonlinear functions, commonly used in the analysis of asset returns [4, 5].

The paradigm of self-similarity and scale-invariance has recently attracted a lot of attention within the finance community due to the multi-scale nature of econometric data, e.g., data that represent second, minute or daily trading. For capturing these features, both fractal-system analysis [6] and self-similar processes [7] have been popularized. However, econometric data are often multivariate and, thus, the concept of scale invariance needs to be generalized. To that end, a recently suggested approach is based on the use of operator fractional Gaussian processes (OfGps) [8, 9], which are multivariate Gaussian self-similar processes. The approach to modeling multivariate self-similar and scale-invariant data is based on linearly mixing a set of independent fractional Gaussian processes, where each of them may have a different Hurst parameter. The Hurst parameter is used, amongst others, as a measure of the long-term memory of time-series. The stationarity, self-similarity and other properties of operator fractional processes have been studied in [10] and references therein. The fGps [11] have already been widely used to analyze a myriad of signals and systems [12], and in particular, finance data [13, 14, 15, 16].

In this paper, we propose a flexible approach for studying OfGps embedded in noise and observed through nonlinear functions. Our goal is to provide a generic framework for inference of correlated latent long-memory data. We consider settings with different prior knowledge about the model parameters, which lead to non-Gaussian densities. Due to the presence of nonlinearities and non-Gaussianities, we devise Sequential Monte Carlo (SMC) methods, also known as Particle Filters (PFs) [17, 18, 19] for tracking the latent processes. We build upon our previous work [20] on the inference of a latent fGp, by extending it to accommodate the correlation of several long-memory time-series, i.e., by constructing an OfGp. The main contribution of this paper is the novel suite of SMC methods devised for inference of correlated mixtures of latent long-memory processes.

The paper is organized as follows. In the next section, we formulate the problem. The proposed methodology is presented in Section 3. In Section 4, we demonstrate the performance of the methods on data simulated from the stochastic volatility model. The last section has our concluding remarks.

2. PROBLEM FORMULATION

We are interested in making inference of correlated latent processes observed through nonlinear functions. Specifically, we consider state-space models that are described by a hidden OfGp embedded in white Gaussian noise and a nonlinear observation equation.

Let $u_t \in \mathbb{R}^{d_u}$ be a vector of independent fGps, each with a Hurst parameter H_i, variance σ_i^2 and the following autocovariance function [21]:

$$\gamma_{u_i}(\tau) = \frac{\sigma_i^2}{2}\rho_{u_i}(\tau) = \frac{\sigma_i^2}{2}\left[|\tau-1|^{2H_i} - 2|\tau|^{2H_i} + |\tau+1|^{2H_i}\right].$$

For $\frac{1}{2} < H_i < 1, i = 1, \cdots, d_u$, the process has long-range dependence; and, for $H_i = 0.5$, it is uncorrelated.

Let $x_t \in \mathbb{R}^{d_x}$ be a set of latent correlated processes (i.e., OfGp) and $y_t \in \mathbb{R}^{d_y}$ the observed vector at time t. We mathematically represent the hierarchical model of interest as follows:

$$\begin{cases} u_{i,t} \sim fGp(H_i, \sigma_i^2), & i = 1, \cdots, d_u, \\ x_t = Au_t + w_t, \\ y_t = h(x_t, v_t), \end{cases} \tag{1}$$

where $t = 1, 2, \cdots$; $w_t \in \mathbb{R}^{d_x}$ is a zero mean Gaussian vector with covariance matrix C_w (we write $w_t \sim \mathcal{N}(0, C_w)$); $A \in \mathbb{R}^{d_x \times d_u}$ is a mixing matrix; $v_t \in \mathbb{R}^{d_v}$ denotes an independent white Gaussian noise; and $h(x_t, v_t) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_v} \to \mathbb{R}^{d_y}$, is some nonlinear function. We point out that we do not restrict the form of the function $h(\cdot, \cdot)$, but we require that the likelihood function $f(y_t|x_t)$ obtained from the function and the density of v_t be known up to a proportionality constant and be computable.

Given a set of observations $y_{1:t} \equiv \{y_1, y_2, \cdots, y_t\}$, we want to sequentially estimate the posterior distribution of x_t, $f(x_t|y_{1:t})$. To do so, we need to proceed sequentially, i.e., for a new observation y_{t+1}, we update $f(x_t|y_{1:t})$ to a new distribution $f(x_{t+1}|y_{1:t+1})$. This density is derived using the Bayes' rule as follows:

$$f(x_{t+1}|y_{1:t+1}) \propto f(y_{t+1}|x_{t+1}) \int f(x_{t+1}|x_{1:t}) f(x_{1:t}|y_{1:t}) \mathrm{d}x_{1:t}. \tag{2}$$

On the one hand, the challenge is to estimate $f(x_{t+1}|x_{1:t})$ for the state-space model as in (1), given its hierarchical nature and the properties of the latent OfGp. On the other, the analytical solution to the integral in (2) is only possible in a very limited number of cases, those with Gaussian noises and linear functions, which is the celebrated Kalman filter [22]. However, we do not restrict ourselves to such settings and thus, resort to SMC methods, widely popular since the seminal publication of [23]. When dealing with nonlinear/non-Gaussian state-space models, these methods approximate the densities of interest by discrete random measures of the form

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

where $x_t^{(m)}$ are particles drawn from a proposal distribution, $w_t^{(m)}$ are the weights associated to the particles, and M is the number of particles. Because sampling from the optimal proposal distribution is intractable for the considered model, we resort to the commonly used transition density of the state.

3. THE PROPOSED METHOD

Here we provide a Bayesian derivation of the state transition densities under different assumptions about the variances of the fGps and the mixing parameters of the OfGp. The derived densities play a central role in the application of the proposed SMC methods.

3.1. Bayesian Analysis

We are interested in the joint filtering density of both latent states in (1), i.e., $f(u_t, x_t|y_{1:t})$ and thus, we need to derive the joint transition density $f(u_{t+1}, x_{t+1}|u_{1:t}, x_{1:t})$. Due to the hierarchical structure of the model, we can factorize it as

$$f(u_{t+1}, x_{t+1}|u_{1:t}, x_{1:t}) = f(u_{t+1}|u_{1:t}) f(x_{t+1}|u_{t+1}). \tag{4}$$

The transition density of the fGps given the parameters H_i and σ_i^2, follows a multivariate Gaussian $f(u_{t+1}|u_{1:t}, \sigma_i^2) = \mathcal{N}(\mu_{t+1}, C_{t+1})$ [20], with parameters

$$\begin{cases} \mu_{t+1} = \begin{pmatrix} \mu_{1,t+1} & \cdots & \mu_{d_u,t+1} \end{pmatrix}^\top, \\ C_{t+1} = \begin{pmatrix} \sigma_{1,t+1}^2 & 0 & \cdots \\ 0 & \sigma_{2,t+1}^2 & \cdots \\ \vdots & \vdots & \ddots \\ 0 & 0 & \sigma_{d_u,t+1}^2 \end{pmatrix}, \end{cases} \tag{5}$$

where we can write

$$\begin{cases} \mu_{i,t+1} = \mathbf{c}_{i,t+1} \mathbf{C}_{i,t}^{-1} u_{i,1:t} , \\ \sigma_{i,t+1}^2 = \sigma_i^2 \left(\rho_{u_i}(0) - c_{i,t} C_{i,t}^{-1} c_{i,t}^\top \right) , \end{cases} \tag{6}$$

with

$$\begin{cases} c_{i,t} = \begin{pmatrix} \rho_{u_i}(1) & \rho_{u_i}(2) & \cdots & \rho_{u_i}(t-1) & \rho_{u_i}(t) \end{pmatrix}, \\ C_{i,t} = \begin{pmatrix} \rho_{u_i}(0) & \rho_{u_i}(1) & \rho_{u_i}(2) & \cdots & \rho_{u_i}(t-2) & \rho_{u_i}(t-1) \\ \rho_{u_i}(1) & \rho_{u_i}(0) & \rho_{u_i}(1) & \cdots & \rho_{u_i}(t-3) & \rho_{u_i}(t-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{u_i}(t-3) & \rho_{u_i}(t-4) & \rho_{u_i}(t-5) & \cdots & \rho_{u_i}(1) & \rho_{u_i}(2) \\ \rho_{u_i}(t-2) & \rho_{u_i}(t-3) & \rho_{u_i}(t-4) & \cdots & \rho_{u_i}(0) & \rho_{u_i}(1) \\ \rho_{u_i}(t-1) & \rho_{u_i}(t-2) & \rho_{u_i}(t-3) & \cdots & \rho_{u_i}(1) & \rho_{u_i}(0) \end{pmatrix}, \end{cases} \tag{7}$$

and $\rho_{u_i}(\tau) = \frac{1}{2}\left[|\tau - 1|^{2H_i} - 2|\tau|^{2H_i} + |\tau + 1|^{2H_i}\right]$. Note that the fGp is not Markovian because the whole past history $u_{i,1:t}$ is required for the computation of the sufficient statistics in (6).

If the variances σ_i^2 are unknown, we can marginalize them and derive alternative expressions, given by d_u location-scale Student's t-distributions [24] as follows:

$$f(u_{i,t+1}|u_{i,1:t}) = \mu_{i,t+1} + l_{i,t+1}\mathcal{T}(\nu_{t+1}),$$
$$\text{with} \begin{cases} \nu_{t+1} = \nu_0 + t, \\ \mu_{i,t+1} = c_{i,t} C_{i,t}^{-1} u_{1:t}, \\ \sigma_{i,t}^2 = \frac{\nu_0 \sigma_0 + u_{i,1:t} C_{i,t}^{-1} u_{i,1:t}^\top}{\nu_t}, \\ l_{i,t+1}^2 = \sigma_i^2 \left(\rho_{u_i}(0) - c_{i,t} C_{i,t}^{-1} c_{i,t}^\top \right) . \end{cases} \tag{8}$$

We note that the underlying model contains the unknown Hurst parameters H_i. Most popular methods that estimate the H_is assume direct observation of x_t [25, 26], which is not the case here. We point out that since the H_is are static parameters, if they are unknown, the applied SMC may have problems in dealing with them [27]. Furthermore, the approaches from the literature may (a) break the self-similarity and stationarity properties of the fGp, and (b) hinder the convergence of the SMC method. A straightforward approach to dealing with the unknown H_is is by way of a bank of parallel SMCs with a subsequent model selection scheme [20], which we adopt here.

From the model in (1), we deduce that the conditional density of x_{t+1} given u_{t+1} is a multivariate Gaussian $f(x_{t+1}|u_{t+1}) = \mathcal{N}(Au_{t+1}, C_w)$, for known A and C_w. Nevertheless, it is unrealistic to know their true values in practice and thus, we marginalize them out. Consider the estimate of the mixing matrix at time t $\hat{A}_t = \mathbf{X}_t \mathbf{U}_t^\top (\mathbf{U}_t \mathbf{U}_t^\top)^{-1}$, where the following historical data matrices have been defined [28]:

$$\begin{cases} \mathbf{X}_t = [x_1 x_2 \cdots x_t] \in \mathbb{R}^{d_x \times t}, \\ \mathbf{U}_t = [u_1 u_2 \cdots u_t] \in \mathbb{R}^{d_u \times t}. \end{cases} \tag{9}$$

Instead of using point estimates, we integrate out the unknown A and C_w to derive the predictive density of x_{t+1}, given u_{t+1}, $\mathbf{X}_t$ and $\mathbf{U}_t$, by following [29]. The resulting density is a multivariate t-distribution

$$f(x_{t+1}|u_{t+1}, \mathbf{X}_t, \mathbf{U}_t) = \mathcal{T}(\nu_{t+1}, \mu_{t+1}, R_{t+1}), \tag{10}$$

with ν_{t+1} degrees of freedom, location parameter $\mu_{t+1} \in \mathbb{R}^{d_x}$ and scale matrix $R_{t+1} \in \mathbb{R}^{d_x \times d_x}$ [24], computed by

$$\begin{cases} \nu_{t+1} = t - d_x - d_u + 1, \\ \mu_{t+1} = \hat{A}_t u_{t+1}, \\ R_{t+1} = \frac{\left(\mathbf{X}_t - \hat{A}_t \mathbf{U}_t\right)\left(\mathbf{X}_t - \hat{A}_t \mathbf{U}_t\right)^\top}{\nu\left(1 - u_{t+1}^\top (\mathbf{U}_{t+1} \mathbf{U}_{t+1}^\top)^{-1} u_{t+1}\right)}. \end{cases} \tag{11}$$

Note that this density does not depend on any of the parameters A and C_w. Now we have the two main transition distributions of the hidden states that are needed for devising our particle filter. This is explained next.

3.2. Particle Filter

We now present SMC methods for joint filtering of the latent states u_t and x_t. Consider that at time instant t the discrete random measure is given by

$$\chi_t = \left\{ \overline{u}_{1:t}^{(m_u)}, \overline{x}_{1:t}^{(m_u)}, w_t^{(m_u)} \right\}, \text{ where } m_u = 1, \cdots, M_u.$$

Upon reception of a new observation at time instant $t+1$, the algorithm proceeds in ways that depend on the available knowledge of the model parameters. The details are as follows:

1. Compute the new correlation values for each latent fGp

$$\rho_{u_i}(\tau) = \frac{1}{2}\left[(\tau+1)^{2H_i} - 2\tau^{2H_i} + (\tau-1)^{2H_i} \right].$$

2. Propagate the latent fGp by conditioning on the available resampled streams $\overline{u}_{1:t}^{(m_u)}$.

   - If σ_i^2 is known,

$$u_{i,t+1}^{(m_u)} \sim f(u_{i,t+1}^{(m_u)} | \overline{u}_{i,1:t}^{(m_u)}) = \mathcal{N}\left(\mu_{i,t+1}^{(m_u)}, \sigma_{i,t+1}^2 \right),$$

$$\text{where} \begin{cases} \mu_{i,t+1}^{(m_u)} = \mathbf{c}_{i,t+1} \mathbf{C}_{i,t}^{-1} \overline{u}_{i,1:t}^{(m_u)}, \\ \sigma_{i,t+1}^2 = \sigma_i^2 \left(\rho_{u_i}(0) - c_{i,t} C_{i,t}^{-1} c_{i,t}^\top \right). \end{cases}$$

   - If σ_i^2 is unknown,

$$u_{i,t+1}^{(m_u)} \sim f(u_{i,t+1}^{(m_u)} | \overline{u}_{i,1:t}^{(m_u)}) = \mu_{i,t+1}^{(m_u)} + l_{i,t+1}^{(m_u)} \mathcal{T}(\nu_{t+1}),$$

$$\text{where} \begin{cases} \nu_{t+1} = \nu_0 + t, \\ \mu_{i,t+1}^{(m_u)} = c_{i,t} C_{i,t}^{-1} \overline{u}_{i,1:t}^{(m_u)}, \\ \sigma_{i,t}^{2(m_u)} = \frac{\nu_0 \sigma_0 + \overline{u}_{i,1:t}^{(m_u)} C_{i,t}^{-1} \left(\overline{u}_{i,1:t}^{(m_u)} \right)^\top}{\nu_t}, \\ l_{i,t+1}^{(m_u)2} = \sigma_{i,t}^{2(m_u)} \left(\rho_{u_i}(0) - c_{i,t} C_{i,t}^{-1} c_{i,t}^\top \right). \end{cases}$$

3. Propagate the latent state by oversampling M_x particles (to improve diversity) from the conditional on the fGp samples.

   - If the mixing parameters are known,

$$x_{t+1}^{(m_u,m_x)} \sim f(x_{t+1} | u_{t+1}^{(m_u)}) = \mathcal{N}(A u_{t+1}^{(m_u)}, C_w).$$

   - If the mixing parameters are unknown,

$$x_{t+1}^{(m_u,m_x)} \sim f(x_{t+1} | u_{t+1}^{(m_u)}) = \mathcal{T}(\nu_{t+1}, \mu_{t+1}^{(m_u)}, R_{t+1}^{(m_u)}),$$

   where

$$\begin{cases} \nu_{t+1} = t - d_x - d_u + 1, \\ \hat{A}_t^{(m_u)} = \mathbf{X}_t^{(m_u)} (\mathbf{U}_t^{(m_u)})^\top (\mathbf{U}_t^{(m_u)} (\mathbf{U}_t^{(m_u)})^\top)^{-1}, \\ \mu_{t+1}^{(m_u)} = \hat{A}_t^{(m_u)} u_{t+1}^{(m_u)}, \\ R_{t+1}^{(m_u)} = \frac{\left(\mathbf{X}_t^{(m_u)} - \hat{A}_t^{(m_u)} \mathbf{U}_t^{(m_u)} \right) \left(\mathbf{X}_t^{(m_u)} - \hat{A}_t^{(m_u)} \mathbf{U}_t^{(m_u)} \right)^\top}{\nu_{t+1} \left(1 - (u_{t+1}^{(m_u)})^\top (\mathbf{U}_{t+1}^{(m_u)} (\mathbf{U}_{t+1}^{(m_u)})^\top)^{-1} u_{t+1}^{(m_u)} \right)}. \end{cases}$$

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

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

   and normalize them to obtain a new random measure

$$\chi_{t+1} = \left\{ u_{1:t+1}^{(m_u)}, x_{1:t+1}^{(m_u,m_x)}, w_{t+1}^{(m_u,m_x)} \right\}.$$

5. Perform downsampling from $M_u \times M_x$ to M_u (this is needed to prevent the growth of the number of samples with time) by drawing the tuple $\left\{ \overline{u}_{1:t+1}^{(m_u)}, \overline{x}_{1:t+1}^{(m_u)} \right\}$ from a categorical distribution defined by the random measure χ_{t+1}

$$\left\{ \overline{u}_{1:t+1}^{(m_u)}, \overline{x}_{1:t+1}^{(m_u)} \right\} \sim \chi_{t+1}, \text{ where } m_u = 1, \cdots, M_u.$$

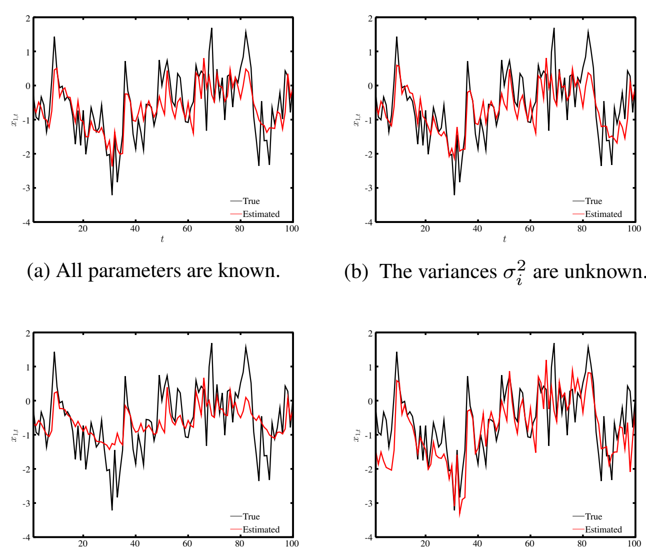
4. SIMULATION RESULTS

We evaluate the proposed SMC method on the stochastic volatility model, which is commonly used in finance [30]. The model is given by

$$\begin{cases} u_{1,t} \sim fGp(H_1, \sigma_1^2), \\ u_{2,t} \sim fGp(H_2, \sigma_2^2), \\ \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} + \mathcal{N}(0, \sigma_w^2 I), \\ y_t = x_{1,t} + e^{x_{2,t}/2} v_t, \end{cases} \tag{12}$$

where v_t represents a standard Gaussian variable.

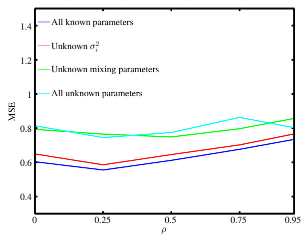
Note that ρ is the idiosyncratic correlation between the trend x_1 and log-volatility x_2 of a return y_t observed over time, while σ_w^2 is the variance of the additive noise w_t in (1).



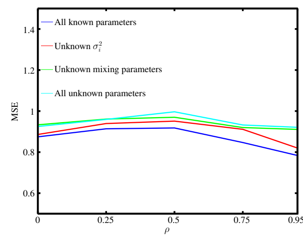
(a) All parameters are known. (b) The variances σ_i^2 are unknown.

(c) The mixing parameters are unknown. (d) All parameters are unknown.

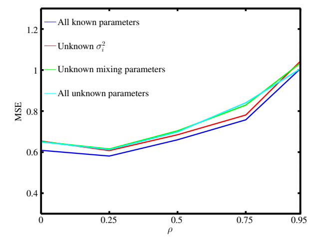
Fig. 1: Estimated (red) and true (black) states x_1.



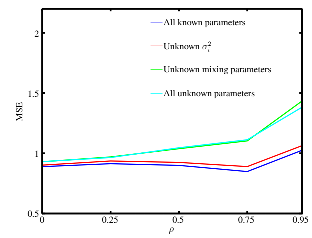
(a) u_1 for $H_1 = 0.5$, $H_2 = 0.5$.



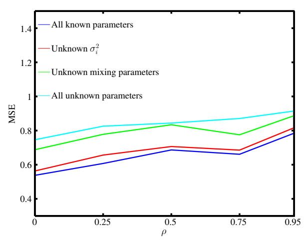
(b) u_2 for $H_1 = 0.5$, $H_2 = 0.5$.



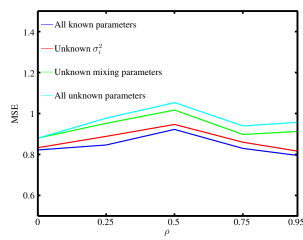
(c) x_1 for $H_1 = 0.5$, $H_2 = 0.5$.



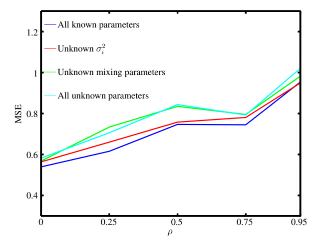
(d) x_2 for $H_1 = 0.5$, $H_2 = 0.5$.



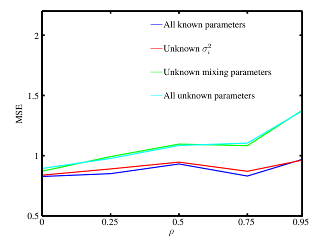
(e) u_1 for $H_1 = 0.7$, $H_2 = 0.7$.



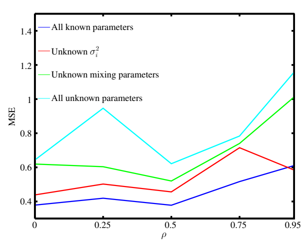
(f) u_2 for $H_1 = 0.7$, $H_2 = 0.7$.



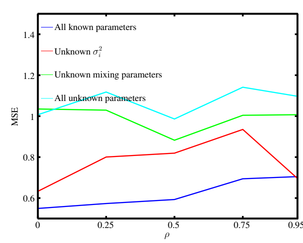
(g) x_1 for $H_1 = 0.7$, $H_2 = 0.7$.



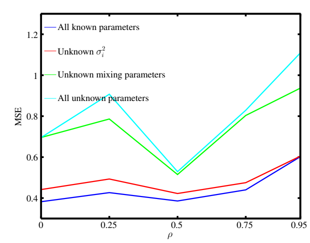
(h) x_2 for $H_1 = 0.7$, $H_2 = 0.7$.



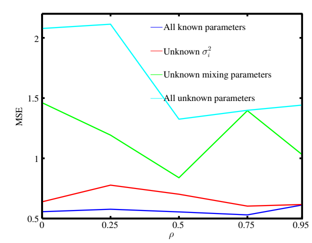
(i) u_1 for $H_1 = 0.9$, $H_2 = 0.9$.



(j) u_2 for $H_1 = 0.9$, $H_2 = 0.9$.



(k) x_1 for $H_1 = 0.9$, $H_2 = 0.9$.



(l) x_2 for $H_1 = 0.9$, $H_2 = 0.9$.

Fig. 2: Filtering MSE for latent fGn and state

The latent OfGp allows for modeling of clustering in the return and volatility [4], i.e., the tendency of asset returns to show large magnitudes in periods of high volatility and calmness in periods of low volatility. Furthermore, we relate the expected asset return to its risk or volatility. We consider an idiosyncratic correlation between the return and its volatility through the mixing matrix A and, at the same time, allow for some random perturbations. The model is used to illustrate the performance of the suggested method, but we do not claim that it fits to any particular instance of real data.

In Fig. 1, we show on a particular realization the capability of the proposed SMC method to track the latent OfGp under different circumstances (we present the estimates of x_1, but similar results are obtained for other latent variables). The plots show a good tracking accuracy and suggest that the impact of not knowing σ_i^2 is less severe than not knowing the mixing parameters. This intuition is backed up by results presented in Fig. 2, where the suggested method is evaluated for different combinations of Hurst parameters and ρ values. All the results have been averaged over 25 realizations and obtained with known H_i values and $M_u = 500$, $M_x = 20$, $\sigma_w^2 = 0.01$.

We note that with the increase of memory of the latent process (i.e., when $H_i \to 1$), the MSE of the latent states decreases. This effect is more evident for the variables x_1 and x_2, but it is also observed for the fGps u_1 and u_2. We also have worse estimation accuracy for the log-volatility when compared to the trend due to its implicit nonlinearity. That is, the method is able to track the trend (i.e., u_1 and x_1) much better than the log-volatility (i.e., u_2 and x_2).

Finally, the results in Fig. 2 reveal that the performance of the SMC methods is consistent for different values of ρ. The justification comes from the way the proposed approach deals with the mixing parameters (by integrating them out). We reiterate that the marginalization is generic, as we do not assume any particular value or structure for A and C_w.

5. CONCLUSIONS

We studied the estimation of correlated latent stochastic processes with long-memory properties. The used models are hierarchical, where the latent states are composed of a set of independent fractional Gaussian processes and an operator fractional Gaussian process embedded in white Gaussian noise. The operator fractional Gaussian process is formed by correlated mixtures of the fractional Gaussian processes. The inference of these processes is made from noisy nonlinear observations of the states. We adopted a Bayesian methodology and proposed flexible Sequential Monte Carlo methods for estimating all the latent states. The methods are able to track both the independent fractional Gaussian processes and the operator fractional Gaussian process. The simulation results validate the accuracy of the proposed approach.

6. REFERENCES

[1] J. Durbin and S. J. Koopman, *Time Series Analysis by State-Space Methods*, Oxford Statistical Science Series. Oxford University Press, 2001.

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

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

[4] S. T. Rachev, J. S. J. Hsu, B. S. Bagasheva, and F. J. Fabozzi, *Bayesian Methods in Finance*, Frank J. Fabozzi Series. Wiley, 2008.

[5] 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.

[6] L. E. Calvet and A. J. Fisher, Eds., *Multifractal Volatility: Theory, Forecasting, and Pricing*, Academic Press, 2008.

[7] D. J. Scansaroli, *Stochastic Modeling with Temporally Dependent Gaussian Processes: Applications to Financial Engineering, Pricing and Risk Management*, Ph.D. thesis, Lehigh University Lehigh Preserve, 2012.

[8] G. Didier and V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," *Bernoulli*, vol. 17, no. 1, pp. 1–33, 2011.

[9] P.-O. Amblard and J.-F. Coeurjolly, "Identification of the Multivariate Fractional Brownian Motion," *Signal Processing, IEEE Transactions on*, vol. 59, no. 11, pp. 5152–5168, Nov 2011.

[10] G. Didier, H. Helgason, and P. Abry, "Demixing multivariate-operator self-similar processes," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, April 2015, pp. 3671–3675.

[11] B. B. Mandelbrot and J. W. V. Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.

[12] J. Beran, *Statistics for Long-Memory Processes*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1994.

[13] R. T. Baillie, "Long memory processes and fractional integration in econometrics," *Journal of Econometrics*, vol. 73, no. 1, pp. 5 – 59, 1996.

[14] E. Bayraktar, H. V. Poor, and K. R. Sircar, "Efficient estimation of the Hurst parameter in high frequency financial data with seasonalities using wavelets," in *Computational Intelligence for Financial Engineering, 2003. Proceedings. 2003 IEEE International Conference on*, March 2003, pp. 309–316.

[15] P. Cheridito, "Arbitrage in fractional Brownian motion models," *Finance and Stochastics*, vol. 7, no. 4, pp. 533–553, 2003.

[16] R. Cont, "Long range dependence in financial markets," in *Fractals in Engineering*, J. Lvy-Vhel and E. Lutton, Eds., pp. 159–179. Springer London, 2005.

[17] A. Doucet, N. D. Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

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

[19] 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(5), pp. 19–38, September 2003.

[20] I. Urteaga, M. F. Bugallo, and P. M. Djurić, "Sequential Monte Carlo sampling for systems with fractional Gaussian processes," in *Signal Processing Conference (EUSIPCO), 2015 Proceedings of the 23th European*, 2015.

[21] W. Palma, *Long-Memory Time Series: Theory and Methods*, Wiley Series in Probability and Statistics. Wiley, 2007.

[22] R. E. 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.

[23] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *Radar and Signal Processing, IEEE Proceedings*, vol. 140, no. 2, pp. 107 –113, apr 1993.

[24] J. M. Bernardo and A. F. Smith, *Bayesian Theory*, Wiley Series in Probability and Statistics. Wiley, 2009.

[25] M. S. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for long-range dependence: An empirical study," *Fractals*, vol. 3, pp. 785–798, 1995.

[26] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *Information Theory, IEEE Transactions on*, vol. 45, no. 3, pp. 878–897, Apr 1999.

[27] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics. Springer, 2001.

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

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

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