SEQUENTIAL MONTE CARLO SAMPLING FOR SYSTEMS WITH FRACTIONAL GAUSSIAN PROCESSES

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 the past decades, Sequential Monte Carlo (SMC) sampling has proven to be a method of choice in many applications where the dynamics of the studied system are described by nonlinear equations and/or non-Gaussian noises. In this paper, we study the application of SMC sampling to nonlinear state-space models where the state is a fractional Gaussian process. These processes are characterized by long-memory properties (i.e., long-range dependence) and are observed in many fields including physics, hydrology and econometrics. We propose an SMC method for tracking the dynamic long-memory latent states, accompanied by a model selection procedure when the Hurst parameter is unknown. We demonstrate the performance of the proposed approach on simulated time-series with nonlinear observations.

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

1. INTRODUCTION

The analysis of time-series is critical in a multitude of fields including engineering, science and economics [1, 2, 3]. In all these areas, stochastic processes are widely used to model the behavior of observed and latent data. At times, the process is modeled according to a physical mechanism that generates the data. Other times, the data are described in a purely statistical sense, without providing a meaningful interpretation of the model parameters. Among the relevant features of such models, the "memory" of the data is considered to be one of its most important characteristics.

On the one hand, experts in many scientific areas have studied short-memory processes in the form of AR, MA, ARMA and other Markov processes. The study of ARMA(p, q) time-series was pioneered by Box and Jenkins [4] and later continued by Durbin in state-space form [2]. In simple words, in short-memory processes only few past samples affect the present data values. On the other hand, in long-memory series, the present value is dependent on samples far into the past. These long-memory processes have attracted attention of practitioners in the last decades [5, 6].

The groundwork on long-memory processes was laid by Hurst [7], when he found that the Nile river data manifested long-range dependence. Later, many other hydrological, geophysical, and climatological records were reported to describe similar characteristics. Motivated by the verification that a plethora of real-life time-series manifest such properties [5, 8], we hereby study long-memory processes that are not directly observed.

To do so, we adopt the state-space paradigm, where the evolution of a system is modeled by a hidden long-memory process, associated with a set of sequentially observed data. In particular, we are interested in latent fractional Gaussian processes (described in Section 2) and nonlinear observation models. Due to the implicit nonlinearities of such models for estimation of the hidden process, we use sequential Monte Carlo (SMC) methods, also known as particle filters (PFs) [9, 10, 11]. They have been successfully applied to many fields [12, 13, 14].

In this paper, we present a new SMC method that can track latent fractional Gaussian processes. Known and unknown parameter cases, under nonlinear observation equations, are considered. In Section 2, we provide an overview of the fractional Gaussian process. The problem of interest is formally introduced in Section 3, for which we propose SMC solutions in Section 4. Computer simulation results for performance evaluation of the methods are contained in Section 5. We complete the paper with our conclusions in Section 6.

2. BACKGROUND

The intuitive interpretation of a long-memory process is that the dependence between events that are far apart diminishes slowly with increasing lag τ. In statistical terms, the memory of the process is accurately described by the autocovariance function $\gamma(\tau)$ of the process. Recall that we hereby consider stationary processes and thus, the autocovariance only depends on the time-lag τ.

For ARMA and Markov processes driven by stationary noises with independent samples, the decay of the autocovariance is asymptotically exponential. It has been proven that an upper bound of the form $|\gamma(\tau)| \propto ba^{\tau}$ for some real scalars

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

$0 < b < \infty$ and $0 < a < 1$ exists for short-memory processes. On the contrary, a slower decay of the autocovariance is observed in long-memory data. Specifically, their decay is explicitly modeled by an autocovariance function proportional to $\tau^{-\alpha}$ for some real scalar $\alpha \in (0, 1)$ [5].

A stationary process x_t with slowly decaying autocovariance function is called a stationary long-memory process when the following relation holds:

$$\lim_{\tau \to \infty} \frac{\gamma(\tau)}{c \cdot \tau^{-\alpha}} = 1, \qquad (1)$$

for real scalars c and $\alpha \in (0, 1)$. Note that, because this is an asymptotic definition, it provides us with the ultimate behavior of the autocovariances as the lag tends to infinity. However, it does not specify its magnitude at any fixed lag.

Additionally, long-memory processes often show self-similarity properties. That is, the whole process possesses same (or similar) shapes as one or more of its parts. This feature was first introduced by Mandelbrot [15] and has been identified in many long-memory processes.

Among the stochastic processes that fulfill such conditions, we focus on fractional Gaussian processes, which represent the increments of the fractional Brownian motion B_t^H. B_t^H is a self-similar process with self-similarity parameter $H \in (\frac{1}{2}, 1)$, known as the Hurst parameter, and stationary Gaussian increments u_t. The increments, $u_t = B_t^H - B_{t-1}^H$ are the fractional Gaussian noise. Because the process u_t is stationary and Gaussian, its sufficient statistics are its mean and autocovariance. For a zero-mean fractional Gaussian process, its autocovariance function is given by

$$\gamma(\tau) = \frac{\sigma^2}{2}\left[|\tau - 1|^{2H} - 2\,|\tau|^{2H} + |\tau + 1|^{2H}\right], \qquad (2)$$

where σ^2 is the variance of the process. For $\frac{1}{2} < H < 1$, the process has long-range dependence; and, for $H = 0.5$, the observations are uncorrelated, as the expression in (2) simplifies to the Kronecker delta function: i.e., $\gamma(\tau) = \sigma^2\delta(\tau)$. We depict the shape of the autocovariance of the fractional Gaussian process for different Hurst parameters in Fig. 1, where the long-term dependence is obvious as $H \to 1$.

If short-memory processes, such as ARMA(p, q) models, were used for modeling non-quickly decaying autocovariance functions, one would need to increase the model orders p and q. Actually, the number of parameters tends to infinity. In practice, an excessive number of parameters is undesirable for many reasons and, especially, because it increases the uncertainty of the statistical inference. Therefore, the parsimonious description provided by the fractional Gaussian process is a more satisfactory model to describe such autocovariance decays.

3. PROBLEM STATEMENT

In this paper, we are interested in the study of latent long-memory processes observed through nonlinear functions.

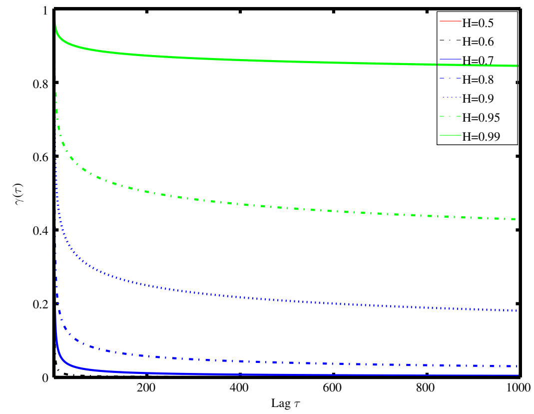

Fig. 1. Autocovariance of the fractional Gaussian noise for different Hurst parameters.

Mathematically, we model such time-series by the following state-space representation:

$$\begin{cases} x_t = u_t, & \text{state equation} \\ y_t = h(x_t, v_t), & \text{observation equation} \end{cases} \qquad (3)$$

where u_t represents a zero-mean fractional Gaussian process with autocovariance function $\gamma_u(\tau)$ as in (2).

Our goal is to sequentially infer the hidden state x_{t+1} as new observations become available. To do so, we are interested on the filtering density $f(x_{1:t+1}|y_{1:t+1})$ which, due to the nonlinearities considered, is approximated by a random measure updated sequentially by SMC sampling. Specifically, the distribution $f(x_{1:t}|y_{1:t})$ is updated to $f(x_{1:t+1}|y_{1:t+1})$ once a new observation y_{t+1} becomes available. We factorize the distribution of interest according to

$$f(x_{1:t+1}|y_{1:t+1}) = f(x_{t+1}|x_{1:t}, y_{1:t+1})f(x_{1:t}|y_{1:t}), \qquad (4)$$

so that the sequential importance sampling methodology can be applied. In a nutshell, SMC sampling consists of sequentially updating a random measure that approximates the distribution of interest by following the recursion in (4). A random measure χ is composed of a set of M particles $x^{(m)}$ and associated weights $w^{(m)}$, i.e., at time instant t, $\chi_t = \{x_{1:t}^{(m)}, w_t^{(m)}\}$, $m = 1, 2, \cdots, M$.

In SMC sampling, the propagation of the random measure at time instant t is achieved by drawing particles from

$$x_{t+1}^{(m)} \sim \pi(x_{t+1}|x_{1:t}, y_{1:t+1}), \qquad (5)$$

and assigning weights according to

$$w_{t+1}^{(m)} \propto \frac{f\left(y_{t+1}|x_{t+1}^{(m)}\right) f\left(x_{t+1}^{(m)}|x_{1:t}^{(m)}\right)}{\pi\left(x_{t+1}^{(m)}|x_{1:t}^{(m)}, y_{1:t+1}\right)} w_t^{(m)}. \qquad (6)$$

Due to the complexity of drawing samples from the optimal distribution $\pi(x_{t+1}|x_{1:t}, y_{1:t+1})$, we resort here to the simple but frequently used transition density $f(x_{t+1}|x_{1:t})$ that leads to particle weight updates according to $w_{t+1}^{(m)} \propto w_t^{(m)} \cdot f(y_{t+1}|x_{t+1})$. The challenge is on deriving the transition density $f(x_{t+1}|x_{1:t})$ for fractional Gaussian processes.

4. PROPOSED APPROACH

As presented in Section 2, the fractional Gaussian noise is a zero-mean Gaussian stationary process with autocovariance function as in (2). Therefore, at time instant $t+1$, the joint distribution of the whole process follows a zero-mean multivariate Gaussian, i.e., $x_{1:t+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{t+1})$ where

$$\mathbf{C}_{t+1} = \begin{pmatrix} c_{t+1} & \mathbf{c}_{t+1} \\ \mathbf{c}_{t+1}^\top & \mathbf{C}_t \end{pmatrix}, \tag{7}$$

and

$$\begin{cases} c_{t+1} = \gamma_u(0), \\ \mathbf{c}_{t+1} = \left(\gamma_u(1) \; \gamma_u(2) \; \cdots \; \gamma_u(t-1) \; \gamma_u(t) \right), \\ \mathbf{C}_t = \begin{pmatrix} \gamma_u(0) & \gamma_u(1) & \gamma_u(2) & \cdots & \gamma_u(t-2) & \gamma_u(t-1) \\ \gamma_u(1) & \gamma_u(0) & \gamma_u(1) & \cdots & \gamma_u(t-3) & \gamma_u(t-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_u(t-3) & \gamma_u(t-4) & \gamma_u(t-5) & \cdots & \gamma_u(1) & \gamma_u(2) \\ \gamma_u(t-2) & \gamma_u(t-3) & \gamma_u(t-4) & \cdots & \gamma_u(0) & \gamma_u(1) \\ \gamma_u(t-1) & \gamma_u(t-2) & \gamma_u(t-3) & \cdots & \gamma_u(1) & \gamma_u(0) \end{pmatrix}. \end{cases}$$

Note that the covariance matrix of the joint distribution of the fractional Gaussian process at time $t+1$ is determined by computing the Toeplitz matrix of the autocovariance function up to lag $\tau = t$.

From this joint distribution and, given the last t samples of the process, we can compute the conditional distribution of the next sample x_{t+1}, given $x_{1:t}$. It follows that the resulting transition distribution is the univariate Gaussian

$$\begin{aligned} & f(x_{t+1}|x_{1:t}) = \mathcal{N}(\mu_{x_{t+1}|x_{1:t}}, \sigma^2_{x_{t+1}|x_{1:t}}), \\ & \text{where} \begin{cases} \mu_{x_{t+1}|x_{1:t}} &= \mathbf{c}_{t+1}\mathbf{C}_t^{-1} x_{1:t}, \\ \sigma^2_{x_{t+1}|x_{1:t}} &= \gamma_u(0) - \mathbf{c}_{t+1}\mathbf{C}_t^{-1}\mathbf{c}_{t+1}^\top. \end{cases} \end{aligned} \tag{8}$$

We now present two SMC schemes that leverage the transition density in (8) for fractional Gaussian processes. We first describe in subsection 4.1 the case when the parameter H is known. In subsection 4.2, we present a solution without assuming knowledge of the Hurst parameter.

4.1. Known Hurst parameter

Let us assume that at time instant t, we have the random measure $\chi_t = \left\{ x_{1:t}^{(m)}, w_t^{(m)} \right\}$, where $m = 1, \cdots, M$ and knowledge of the Hurst parameter H. Upon reception of a new data sample y_{t+1}, we proceed as follows:

1. Perform resampling (to avoid sample degeneracy) by drawing from a categorical distribution defined by the random measure χ_t,

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

2. Compute the necessary sufficient statistics in (7) by using the autocovariance function for $\tau = 0, 1, \cdots, t$,

$$\gamma_u(\tau) = \frac{\sigma^2}{2} \left(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H} \right).$$

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

$$x_{t+1}^{(m)} \sim f(x_{t+1}|\overline{x}_{1:t}^{(m)}) = \mathcal{N}(\mu_{x_{t+1}|\overline{x}_{1:t}^{(m)}}, \sigma^2_{x_{t+1}|\overline{x}_{1:t}^{(m)}}),$$

$$\text{where} \begin{cases} \mu_{x_{t+1}|\overline{x}_{1:t}^{(m)}} &= \mathbf{c}_{t+1}\mathbf{C}_t^{-1}\overline{x}_{1:t}^{(m)}, \\ \sigma^2_{x_{t+1}|\overline{x}_{1:t}^{(m)}} &= \gamma_u(0) - \mathbf{c}_{t+1}\mathbf{C}_t^{-1}\mathbf{c}_{t+1}^\top. \end{cases}$$

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

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

and normalize them to obtain a new random measure

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

4.2. Unknown Hurst parameter

In practice, assuming knowledge of the parameter H of the hidden fractional Gaussian process is unrealistic and thus, new alternatives must be considered. It has been extensively reported that SMC sampling techniques suffer when tracking fixed model parameters [16]. To overcome such limitations, various methodologies have been suggested [16, 17]. In these, an approximating model is assumed for the fixed parameter, either by assuming a slowly varying parameter or by approximating the joint distribution of the parameter and the state. However, these approaches raise several concerns in our problem of interest. On the one hand, assuming a varying H would not only break important properties of the process (i.e., stationarity and self-similarity), but would also affect the mixing properties of the states, thus endangering the convergence properties of the SMC algorithm. On the other hand, determining a suitable density for H is challenging, due to its complicated dependency on the state.

Consequently, we resort to an alternative solution that consists of a bank of K parallel SMC filters with different H parameters, $H_k \in [0.5, 1)$, $k = 1, 2, \cdots, K$; followed by a model selection scheme. Each of the SMC filters proceeds as described in subsection 4.1 for a given H_k. For the model selection criteria, we consider the $k = 1, 2, \cdots, K$ predictive likelihoods of the observations

$$f(y_{t+1}|y_{1:t}, H_k) = \int f(y_{t+1}|x_{t+1})f(x_{t+1}|y_{1:t}, H_k)dx_{t+1}. \quad (9)$$

To numerically solve the otherwise analytically non-solvable integral, we propose to take advantage of the random measure provided by each SMC instance. That is, given

$$f(x_{t+1}|y_{1:t}, H_k) \approx \sum_{m=1}^{M} w_t^{(m)} f(x_{t+1}|x_{1:t}^{(m)}, H_k)\delta\left(x_{1:t} - x_{1:t}^{(m)}\right), \quad (10)$$

J predictive samples are drawn from the transition density, i.e., $x_{t+1}^{(m,j)} \sim f(x_{t+1}|x_{1:t}^{(m)}, H_k)$, $j = 1, 2, \cdots, J$. Thus, we approximate the predictive likelihood of the next observation with

$$f(y_{t+1}|y_{1:t}, H_k) \approx \sum_{m=1}^{M} \frac{w_t^{(m)}}{J} \sum_{j=1}^{J} f(y_{t+1}|x_{t+1}^{(m,j)}). \quad (11)$$

Based on this metric, the best model parameter H^* that describes the observed sequence $y_{1:t+1}$ is obtained from

$$H^* = \underset{H_k}{\operatorname{argmax}} \sum_{i=1}^{t} \log f(y_{i+1}|y_{1:i}, H_k). \quad (12)$$

5. SIMULATION RESULTS

We evaluate the suggested SMC sampling methods under the stochastic volatility model, where the log-volatility of the observed time-series is a fractional Gaussian process. Volatility models with long-memory characteristics have been widely studied in finance [8]. The goal is to estimate the evolving volatility of an observed series of stock prices. Mathematically, the corresponding state-space model is written as

$$\begin{cases} x_t = u_t, \\ y_t = e^{x_t/2} v_t, \end{cases} \quad (13)$$

where u_t is a zero-mean fractional Gaussian process with autocovariance function as in (2) and $v_t \sim \mathcal{N}(0, \sigma_v^2)$. For the following simulation results,[1] we assume $\sigma^2 = 1$ and $\sigma_v^2 = 1$. We evaluate the performance of the methods for different Hurst parameter values.

The state filtering accuracy of the proposed SMC for the known Hurst parameter case is presented in Fig. 2. The results illustrate the implicit benefits of tracking long-memory processes. As $H \to 1$, more information is provided by past samples in estimating the present value of the series and thus, the performance of the SMC sampling technique improves. That is, it is possible to more accurately estimate the time-evolving log-variance of the observed time-series, because of the slow decay of the autocovariance function in such long-memory processes.

[1]The presented results are performance metrics averaged over $R = 100$ realizations of 200 time-instant series with $M = 1000$ particles.

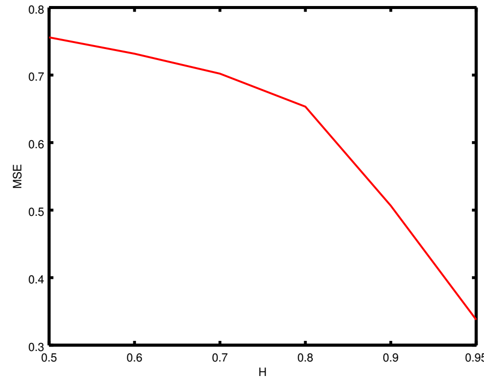

Fig. 2. MSE of the state for different Hurst parameters.

SMC bank K=6	True H					
	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$	$H = 0.95$
$H_1 = 0.5$	0.75585	0.7384	0.7414	0.76453	0.7468	0.70245
$H_2 = 0.6$	0.76235	0.73161	0.71785	0.71312	0.66156	0.59087
$H_3 = 0.7$	0.78312	0.73957	0.70206	0.67159	0.57899	0.47794
$H_4 = 0.8$	0.82449	0.76833	0.70837	0.65323	0.52737	0.39924
$H_5 = 0.9$	0.9034	0.83263	0.74516	0.67107	0.50654	0.35076
$H_6 = 0.95$	0.98512	0.90572	0.80342	0.71992	0.51913	0.33772

Table 1. MSEs of the state for a bank of SMC filters.

As described in subsection 4.2, we resort to a bank of parallel SMC filters when the Hurst parameter is unknown. The results of this experiment are contained in Table 1, where it is clear that the most accurate inference is attained when the correct parameter value is assumed. Furthermore, for rapidly decaying autocovariance functions (i.e., $0.5 \leq H < 0.75$), there is a minimal performance difference among the SMC filters that assume H values in this interval. On the contrary, as the long-memory effect becomes more evident ($H \geq 0.75$), only the filters with values close to the true one provide good accuracy. In conclusion, the relevance of long-memory dependence is highlighted.

For the unknown Hurst parameter case, we run the bank of SMC schemes in parallel and thus, there is a need for a model selection scheme that identifies the correct H value. When evaluating the likelihood of the observed time-series for different models, only those with assumed H_k parameter close to the correct one are able to provide the best results (see Fig. 3). The results presented here are obtained for $J = 100$, which provides an improved predictive performance at an acceptable computational overhead. We illustrate the validity and accuracy of the proposed model selection procedure with Table 2 and confirm that it improves as more instances of the time-series are observed.

Finally, we emphasize once more the difficulty of distinguishing the processes with rapidly decaying autocovariance functions (i.e., $0.5 \leq H < 0.75$). By contrast, for long-memory processes, the long-range dependence induced as $H \to 1$ helps the SMC sampling to identify the correct model. In other words, only when capturing the influence of the sam-

SMC bank K=6	True H					
	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$	$H = 0.95$
$H_1 = 0.5$	40	53	44	36	28	21
$H_2 = 0.6$	13	15	5	4	10	3
$H_3 = 0.7$	5	6	6	13	8	8
$H_4 = 0.8$	9	5	11	9	5	8
$H_5 = 0.9$	9	4	5	5	9	11
$H_6 = 0.95$	24	17	29	33	40	49

SMC bank K=6	True H					
	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$	$H = 0.95$
$H_1 = 0.5$	71	42	16	2	1	3
$H_2 = 0.6$	11	29	17	11	2	1
$H_3 = 0.7$	9	19	34	29	1	2
$H_4 = 0.8$	2	6	21	39	21	5
$H_5 = 0.9$	3	3	10	12	38	17
$H_6 = 0.95$	4	1	2	7	37	72

Table 2. Model selection confusion matrix at $t = 10$ (left) and $t = 200$ (right).

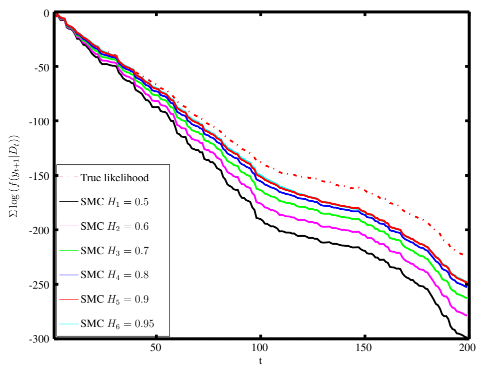

Fig. 3. Cumulative log-likelihood of a time-series ($H = 0.9$).

ples deep in the past on the present values, the SMC method is able to accurately predict the next observations.

6. CONCLUSIONS

In this paper, inference of long-memory processes under non-linear state-space models has been considered. We have studied the characteristics of the fractional Gaussian noise and have derived its transition density. We have proposed Sequential Monte Carlo methods that successfully estimate latent fractional Gaussian processes, when the Hurst parameter is either known or unknown. Comprehensive simulation results demonstrate the validity of the suggested SMC methods and the model selection procedure.

REFERENCES

[1] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer Series in Statistics. Springer, 2nd edition, 1991.

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

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

[4] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day series in time series analysis and digital processing. Holden-Day, 1976.

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

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

[7] H. E. Hurst, "Long-term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–808, 1951.

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

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

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

[11] 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.

[12] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2004.

[13] P. J. van Leeuwen, "Particle filtering in geophysical systems," *Monthly Weather Review*, vol. 12, no. 137, pp. 4089–4114., 2009.

[14] D. Creal, "A survey of Sequential Monte Carlo Methods for Economics and Finance," *Econometric Reviews*, vol. 31, no. 3, pp. 245–296, 2012.

[15] 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.

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

[17] 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, 2004 (ICASSP '04)*, May 2004, vol. 2, pp. ii – 701–704.