ESTIMATION OF TIME-VARYING AUTOREGRESSIVE SYMMETRIC ALPHA STABLE PROCESSES BY PARTICLE FILTERS*

*Deniz Gençağa[1,**], Ercan E. Kuruoğlu[2], Ayşın Ertüzün[1]*

[1]Electrical and Electronic Engineering Dept., Boğaziçi University, Bebek, 34342, İstanbul, Turkey
[2] ISTI, Consiglio Nazionale delle Ricerche, Pisa, Italy
gencagao@boun.edu.tr, ercan.kuruoglu@isti.cnr.it, ertuz@boun.edu.tr

ABSTRACT

In the last decade alpha-stable distributions have become a standard model for impulsive data. Especially the linear symmetric alpha-stable processes have found applications in various fields. When the process parameters are time-invariant, various techniques are available for estimation. However, time-invariance is an important restriction given that in many communications applications channels are time-varying. For such processes, we propose a relatively new technique, based on particle filters which obtained great success in tracking applications involving non-Gaussian signals and nonlinear systems. Since particle filtering is a sequential method, it enables us to track the time-varying autoregression coefficients of the alpha-stable processes. The method is tested both for abruptly and slowly changing autoregressive parameters of signals, where the driving noises are symmetric-alpha-stable processes and is observed to perform very well. Moreover, the method can easily be extended to skewed alpha-stable distributions.

1. INTRODUCTION

With the availability of increasingly higher computing power, the particle filters have found practical applications in many disciplines, such as communications, astrophysics, biomedicine and finance [1]. In its most general form, particle filters enable us to obtain the optimal Bayesian solution of the systems that can be modelled by non-Gaussian and nonlinear state-space equations [1-2]. For such systems, if the signals are non-stationary, particle filters can still provide us with the optimal Bayesian solution, since the estimation is performed sequentially. However, other Bayesian techniques, such as the Markov Chain Monte Carlo (MCMC) [3], can only be used for stationary signals, since these methods have batch processing nature and discard the time information of the signals.

Generic techniques, that are developed for non-Gaussian signals can be applied to alpha-stable (α-stable) processes, since they possess non-Gaussian distributions too except for α=2, corresponding to the Gaussian case. For stationary cases, MCMC techniques have been applied to estimate the parameters of an α-stable process [4]. However, there is a limited number of studies that have been done for handling the non-stationary cases [5]. In literature, particle filters are utilised to estimate the time-varying autoregressive (AR) coefficients of a process, which is driven by a *Gaussian* noise and embedded in an additive noise, modelled by a symmetric α-stable distribution (*SαS*) [6].

In literature [7-8], it is known that the particle filters can successfully estimate the time-varying autoregressive coefficients of *non-Gaussian* signals, such as the Mixture of Gaussian and Laplacian distributed ones. Motivated by these, a novel method for estimating the time-varying AR coefficients of *α-stable* processes is proposed in this work.

The paper is organized as follows: First, the problem is stated formally and, a brief background information on α-stable processes and particle filters is presented. Then, the proposed method is introduced and the performance analysis is illustrated by the computer simulations.

1.1 α-stable Processes

It is well known that, if we add a large number of random variables of different distributions, the summation variable tends to be more Gaussian distributed as the number of terms goes to infinity. This is known as the Central Limit Theorem (CLT). Moreover, it is necessary that each added random variable is of finite variance. Otherwise, CLT becomes insufficient and Generalized Central Limit Theorem should be used [9]. In this case, the limiting distribution is an α-stable distribution. α-stable distributions are defined in terms of their characteristic functions, since their probability density functions (pdf) cannot be obtained analytically, except for some limited cases (α=2, β=0 Gaussian; α=1, β=0 Cauchy; α=0.5, β=-1 Pearson) [10, p. 14].The characteristic function of α-stable distributions is given as follows:

$$\varphi(t) = \exp\left\{j\delta t - \gamma|t|^{\alpha}\left[1 + j\beta\, sign(t)\,\omega(t,\alpha)\right]\right\} \quad \text{(1a)}$$

Here, the parameters are defined within the following intervals: $-\infty < \delta < \infty, \quad \gamma > 0, \quad 0 < \alpha \le 2, \quad -1 \le \beta \le 1$.

$$\omega(t,\alpha) = \begin{cases} \tan\frac{\alpha\pi}{2} & if \quad \alpha \neq 1 \\ \frac{2}{\pi}\log|t| & if \quad \alpha = 1 \end{cases} \quad and \quad sign(t) = \begin{cases} 1 & if \quad t > 0 \\ 0 & if \quad t = 0 \\ -1 & if \quad t < 0 \end{cases} \quad \text{(1b)}$$

As shown above, an α-stable distribution is defined by four parameters. Among these, α and β are known as the shape parameters and they determine the thickness of the tails and the symmetry of the distribution, respectively.

*This work is supported by TÜBİTAK-CNR project number: 102E027 and Boğaziçi University Scientific Research Fund project number: 04A201.
** The researcher was supported by NATO-TÜBİTAK A2 fellowship, throughout his research at ISTI-CNR, Italy.

For example, in our work, *SαS* are used. Thus, in our case β parameter is taken to be zero. As α gets smaller, the distributions become more impulsive. δ and γ are known as the measures of the location and the dispersion around it, respectively. In this work, standard (δ = 0, γ = 1) α-stable distributions are considered and they can be generalized easily by using variable transformations. So, for standard *SαS* distributions, (1a) takes the following form:

$$\varphi(t) = \exp\left\{-|t|^{\alpha}\right\} \tag{2}$$

It is well known that AR processes are obtained by filtering a white noise with an all-pole filter. The difference equation, corresponding to such a process, is given as follows:

$$y(t) = \sum_{k=1}^{K} x_k(t)\, y(t-k) + v(t) \tag{3}$$

Above, $x_k(t)$ parameters are known as the autoregressive parameters and v(t) is the driving process. The estimation of the AR coefficients from observation y(t) can be performed by the well known Yule-Walker equations, in case of Gaussian driving processes [11]. In case of α-stable driving processes, methods, such as the Iteratively Reweighted Least Squares [12], Generalized Yule-Walker [10], MCMC based techniques [4] and others [13] have been proposed in the literature. However, in all these methods, the AR coefficients are assumed to be time-invariant. In our work, we consider the case where these parameters are time-varying.

1.2 Particle Filters

Particle filters are used in order to sequentially update *a priori* knowledge about some predetermined state variables by using the observation data. In general, these state variables are the hidden variables in a non-Gaussian and nonlinear state-space modelling system. Such a system can be given by the following equations:

$$\begin{aligned} \mathbf{x}_t &= f_t(\mathbf{x}_{t-1}, \mathbf{v}_t) \\ \mathbf{y}_t &= h_t(\mathbf{x}_t, \mathbf{n}_t) \end{aligned} \tag{4}$$

where $\mathbf{x}_t$ and $\mathbf{y}_t$ represent the hidden state and the observation vectors at current time *t*, respectively. Here, the process and observation noises are denoted by $\mathbf{v}_t$ and $\mathbf{n}_t$, respectively. f_t and h_t are known as the process and observation functions and in their most general case, they are nonlinear. Also, the noise processes in (4) are modelled to be non-Gaussian. Here, the objective is to sequentially obtain the *a posteriori* distribution of the state variables obtained via the observation data gathered up to that time, i.e. $p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})$. If both the process and the observation noises are Gaussianly distributed and the corresponding functions f_t and h_t are linear, then the desired *a posteriori* distribution is also Gaussian and sequentially estimating the mean and variance is sufficient instead of the whole pdf. In this situation, the optimal solution can be obtained by the Kalman filter [14]. For this condition, (4) is expressed as follows:

$$\begin{aligned} \mathbf{x}_t &= \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{v}_t \\ \mathbf{y}_t &= \mathbf{H}_t \mathbf{x}_t + \mathbf{n}_t \end{aligned} \tag{5}$$

where F_t and H_t are linear operators and the noise distributions are Gaussian. For both (4) and (5), the optimal Bayesian solution for the *a posteriori* pdf is given as follows [1-2]:

$$p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{x}_{t-1})}{p(\mathbf{y}_t \mid \mathbf{y}_{t-1})}\, p(\mathbf{x}_{0:t-1} \mid \mathbf{y}_{1:t-1}) \tag{6}$$

In general non-Gaussian situations we may not always have analytical expressions for distributions. Thus, the distributions are expressed in terms of samples, to approximate them. These samples are called as the particles. The expression for the *a posteriori* pdf can be given in terms of particles as follows:

$$p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^i\, \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^i) \tag{7}$$

where w_t^i , $\mathbf{x}_{0:t}^i$, δ(.) denote the weight, i^{th} particle and the Kronecker delta operator, respectively. Then, expectations for function g(.) can be obtained by the following equation:

$$I(f_t) = \int g(\mathbf{x}_{0:t})\, p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})\, d\mathbf{x}_{0:t} \tag{8}$$

where g(.) is a function depending on the estimate [1]. Here, the major problem is to draw samples from an analytically inexpressible non-Gaussian distribution and estimate the integral given by (8) using Monte Carlo integration techniques, shown as follows:

$$\hat{I}_N(f_t) = \sum_{i=1}^{N} g_t(\mathbf{x}_{0:t}^i)\, \tilde{w}_t^i \tag{9}$$

where $\tilde{w}_t^i$ denote the normalized weights given as:

$$\tilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^{N} w_t^i}, \quad i = 1, \ldots, N \tag{10}$$

The particles that take place in equations (7) and (9) are drawn by a method known as the “Importance Sampling” [1-2] and the corresponding “Importance Weight” for each of them is denoted by w_t^i as defined as follows:

$$w_t^i \propto \frac{p(\mathbf{x}_{0:t}^i \mid \mathbf{y}_{1:t})}{q(\mathbf{x}_{0:t}^i \mid \mathbf{y}_{1:t})} \tag{11}$$

where q(.) function is called as the “Importance Function” and drawing samples from this pdf is easier than that of original distribution [1-2]. However, importance sampling shown in (11), can be used in batch processing techniques and should be modified as follows for the sequential applications [1-2]:

$$w_t^i \propto w_{t-1}^i \frac{p(\mathbf{y}_t \mid \mathbf{x}_t^i)\, p(\mathbf{x}_t^i \mid \mathbf{x}_{t-1}^i)}{q(\mathbf{x}_t^i \mid \mathbf{x}_{0:t-1}^i, \mathbf{y}_{1:t})} \tag{12}$$

But, as a consequence of this sequential modification, a phenomenon, known as “Degeneracy”, arises as a problem and causes the importance weight of each particle, but one, to converge to zero as time evolves [1-2]. In order to avoid the degeneracy problem, "Resampling" is performed as an additional step and by this procedure, particles with high importance weights are replicated, while the others are discarded. By doing so, we can approximate the desired pdf in time [1-2].

2. THE PROPOSED METHOD

In this work, we propose a new method, which enables us to sequentially track the time-varying AR parameters of an α-stable process from the observation data. The corresponding AR model is given in (3). Here, these AR coefficients are expressed in terms of particles and form the state vector, which is given by the following the state-space equations:

$$\begin{aligned} \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{v}_t \\ y_t &= \mathbf{y}_{t-1}^T \mathbf{x}_t + n_t \end{aligned} \tag{13}$$

where vectors are defined as $\mathbf{y}_{t-1}=(y_{t-1},..........,y_{t-K})^T$ and $\mathbf{x}_t=(x_1(t),......,x_K(t))^T$. Moreover, in order to be able to model the system correctly, the statistical properties of the process noise, given in the first equation of (13), should be known. However, when we do not have any *a priori* information regarding to the states, as in this case, an additional estimation technique should be used to model the process equation accurately. Since there is no *a priori* information, the state transition matrix is taken to be the identity matrix. When there is no information about the process noise, the method proposed in [7-8] can be used in order to sequentially model the covariance matrix of the zero mean process noise from the past data. That is, the process noise is modelled by a Gaussian distribution. Here, in case of a scalar state variable, the variance of the process noise can be estimated from the variances of the particles regarding to the previous AR coefficients as follows:

$$\sigma_{v(t)}^2 = \sigma_{x(t-1)}^2 \left(\frac{1}{\lambda} - 1\right) \tag{14}$$

where λ is called as the "Forgetting Factor" and takes values between zero and one. In case of vector variables, the covariance matrix of the process noise vector is estimated as follows:

$$\mathbf{\Sigma}_{v(t)} = \mathbf{\Sigma}_k \left(\frac{1}{\lambda} - 1\right) \tag{15}$$

where Σ_k is a diagonal matrix, whose elements are variances of the particles, corresponding to the related AR coefficient at time (t-1). After forming the state-space equations, the next issue is the choice of the importance function. Here, we choose the *a priori* transition pdf, which is given by $q(\mathbf{x}_t^i \mid \mathbf{x}_{0:t-1}^i, \mathbf{y}_{1:t}) = p(\mathbf{x}_t^i \mid \mathbf{x}_{t-1}^i)$. As a result of this selection, the importance weight calculation (12) of each particle becomes as follows [1-2, 7-8]:

$$w_t^i \propto w_{t-1}^i p(y_t \mid \mathbf{x}_t^i) \tag{16}$$

where $p(y_t \mid \mathbf{x}_t^i)$ is the likelihood term. Since the observation noise n_t of (13) has an α -stable distribution, this term cannot be estimated analytically, except for a very limited number of cases, which are mentioned at the introduction. Thus, in order to estimate these importance weights, we take the inverse Fourier transform of the related standard characteristic function numerically at each time instant and evaluate its value for the observation data as follows:

$$p(y_t \mid \mathbf{x}_t^i) = \frac{1}{\pi}\int_0^\infty \exp(-t^\alpha)\cos\left[\left(y_t - \mathbf{y}_t^T \mathbf{x}_t^i\right)t\right]dt \tag{17}$$

3. EXPERIMENTS

In this section, the theory given in the previous sections is justified by several computer simulations. In these simulations, a synthetically generated first order AR process is used, which can be given in the following form:

$$y(t) = x(t)y(t-1) + n(t) \tag{18}$$

where the AR coefficients are time-varying and represented by x(t). The driving process n(t) is generated from various α-stable distributions, as explained below. In all cases the distributions are symmetric (β = 0) and standard (γ = 1, δ = 0). For each experiment, 20 ensembles are used in order to estimate the Normalized Mean Square Errors (NMSE), which can be given as follows:

$$NMSE(t) = \frac{\sum_{i=1}^{20} (\hat{x}_i(t) - x(t))^2}{\sum_{i=1}^{20} \left|\mathbf{x}\right|^2} = \frac{\sum_{i=1}^{20} (\hat{x}_i(t) - x(t))^2}{\sum_{i=1}^{20}\sum_{t=1}^{1000} x^2(t)} \tag{19}$$

where i denotes the related ensemble and $\hat{x}_i(t), x(t)$ denote the Minimum Mean Square Estimate (MMSE) and the original AR coefficients, respectively. In all experiments, residual resampling [15], 100 particles and 1000 time samples are used.

Two experiments are conducted:

a) AR coefficient is taken to be changing sinusoidally with time. The effect of various α parameters is examined and the estimated AR trajectories and their instantaneous NMSE are plotted as a function of time (Fig. 1).

b) AR coefficient is taken to be 0.99 until the 500th sample, where it changes abruptly to 0.95. This is examined for 4 different α parameters (Fig. 2).

4. DISCUSSION AND CONCLUSIONS

In this work, a new method is proposed in order to estimate the time-varying AR processes, which are driven by symmetric-α-stable processes. The performance of the method is tested for several values of the α parameter and it is observed to perform very well and the quality of the MMSE of the AR coefficients increases as the value of the α decreases, that is, as the process becomes more heavy-tailed. This is illustrated in Fig. 1. It is also noted that the NMSE value at the peaks of the AR waveform decreases significantly. This is due to the slow variation of the AR coefficients throughout these regions. Tracking performance in case of an abruptly changing AR coefficient is shown in Fig. 2. When Figs. 1 and 2 are compared, it is seen that the quality of the estimates increase as the time variation of the AR coefficients decrease. As a final remark, the proposed method can easily be extended for the skewed ($\beta \neq 0$) α-stable distributions by including the β parameter in (2) and (17). As a future work, the skewed case will be examined and parameters of the α-stable distribution will also be estimated beside the AR coefficients.

REFERENCES

[1] Doucet A., de Freitas N., Gordon N., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[2] Doucet A., Godsill S., Andrieu C., "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering", *Statistics and Computing*, (2000), 10, 197-208.

[3] Robert C., Casella G., *Monte Carlo Statistical Methods*, Springer, 1999.

[4] Godsill S. , Kuruoğlu E. E., "Bayesian Inference for Time Series with Heavy-Tailed Symmetric α-Stable Noise Processes", *Heavy Tails'99, Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, June. 3-5, 1999, American University, Washington D.C., USA.

[5] Thavaneswaran A., Peiris S., "Estimation for Regression with Infinite Variance Errors", *Mathematical and Computer Modelling*, (29), 1999, pp. 177-180.

[6] Lombardi M., Godsill S., "Monte Carlo Bayesian Filtering and Smoothing for TVAR Signals in Symmetric α-stable noise", *EUSIPCO 2004*, pp. 865-872, Vienna, Austria, 2004.

[7] Djuric P. M., Kotecha J. H., Esteve F., Perret E., "Sequential Parameter Estimation of Time-Varying Non-Gaussian Autoregressive Processes", *EURASIP Journal on Applied Signal Processing*, 2002:8, 865-875.

[8] Djuric P. M., Kotecha J. H., Tourneret J. Y., Lesage S., "Adaptive Signal Processing by Particle Filters and Discounting of Old Measurements", *Proc. IEEE Conf. Acous., Speech, Signal Proc.* , Salt Lake City, USA, 2001.

[9] Samorodnitsky G., Taqqu M. S., *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*, Chapman & Hall, 1994.

[10] Nikias C. L., Shao M., *Signal Processing with Alpha-Stable Distributions and Applications*, Prentice-Hall, 1995.

[11] Haykin S., *Adaptive Filter Theory*, Prentice-Hall, 1996.

[12] Kuruoğlu E. E., Rayner P., Fitzgerald W., "Least L_p Norm Estimation of Autoregressive Model Coefficients of Symmetric α-Stable Processes", *IEEE Signal Proc. Letters*, (4), no. 7, July 1997.

[13] Davis R.A., Knight K., Liu J., "M-estimation for autoregressions with infinite variance", *Stochastic Proc. and their Applications,* Vol. 40, Issue 1, Feb 1992, pp. 145-180.

[14] Anderson B., Moore J., *Optimal Filtering*, Prentice-Hall, 1979.

[15] Liu J., Chen R., "Sequential Monte Carlo Methods for Dynamic Systems", *J. American Stat. Assoc.*, 93(443): 1032-1044,1998.

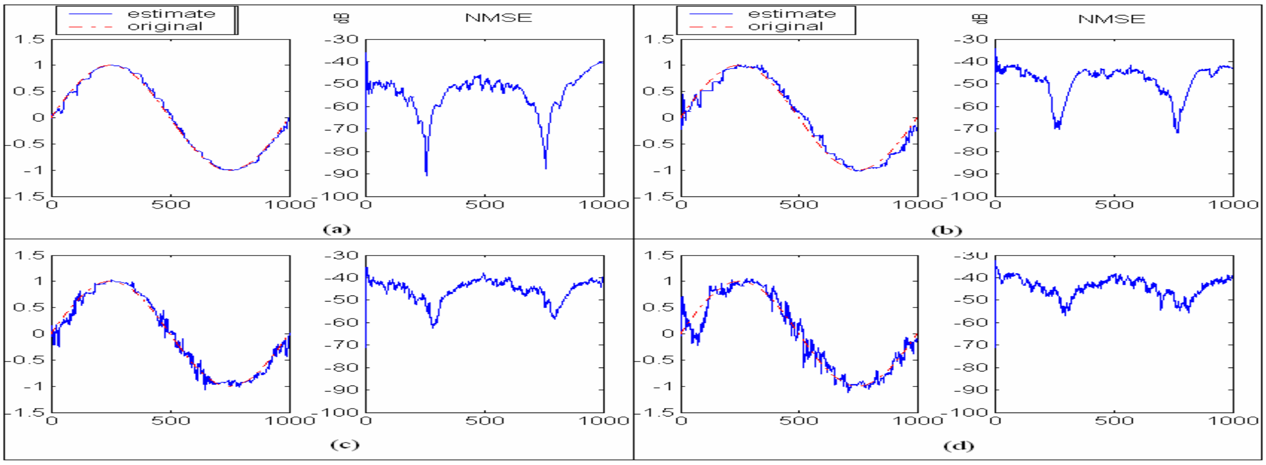

Fig. 1 Estimation of the sinusoidally varying AR parameter for different α values: a) α = 0.5, b) α = 1, c) α = 1.5, d) α = 2

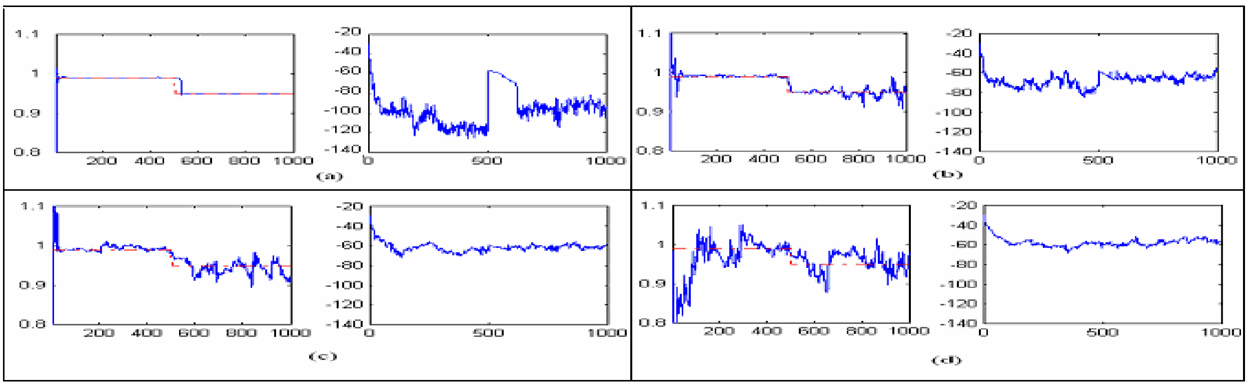

Fig. 2 Estimation of the abruptly varying AR parameter for different α values: a) α = 0.5, b) α = 1, c) α = 1.5, d) α = 2