EUSIPCO 2013 1569742537

AR PROCESSES WITH NON-GAUSSIAN ASYMMETRIC INNOVATIONS

Pascal Bondon and Li Song

CNRS UMR 8506, Université Paris-Sud, 91192 Gif-sur-Yvette, France.

ABSTRACT

We consider the problem of modeling non-Gaussian correlated signals by autoregressive models with skew exponential power innovations. Generalized moments and maximum likelihood estimators of the parameters are proposed and large sample properties are established. Finite sample behavior of the estimators is studied via Monte Carlo simulations. An application to real data is considered.

Index Terms**—** Non-Gaussian, asymmetric distribution, autoregressive model, maximum likelihood estimation.

1. INTRODUCTION

Many modern signal processing applications and systems are faced with various and complex noise environments. These noises may be non-Gaussian and have symmetric or asymmetric marginals. The non-Gaussian character of these noise sources can significantly weaken the performance of conventional statistical tools which are designed to operate effectively in Gaussian environments. From the last century, interest has grown in the construction of flexible classes of distributions that, unlike the Gaussian distribution, may be platykurtic or leptokurtic and asymmetric. Among these classes, some families strictly including the Gaussian distribution not only as a limiting case but as a proper member. These families are interesting because they can vary continuously from Gaussianity to non-Gaussianity and from symmetry to asymmetry. For instance, the exponential power distribution introduced by [1] and used by [2, Chapter 3] in robustness analysis, offers departure from Gaussianity. The works by [3] and [4] introduced a class of asymmetric distributions, called skew-normal, which includes the normal distribution as a special case. An alternative method for generating skew distributions from symmetric distributions, based on the introduction of an asymmetry parameter was suggested by [5] which proposed a skew exponential power (SEP) distribution. On the other hand, the epsilon-skew-normal distribution, the origin of which can be traced back to [6] was defined by [7]. A more general class of asymmetric distributions, encompassing the families in [5] and [7] was presented and analyzed by [8]. Here, we shall use the SEP distribution proposed by [8], see also [9]. This SEP distribution family strictly includes the normal and Laplace distributions, accommodates practical values of skewness and kurtosis, and is analytically tractable. Therefore, its use in data modeling, statistical analysis and robustness studies of normal theory methods is attractive.

We are interested in modeling non-Gaussian correlated signals. The dependence of the data is captured by a linear model. In the 1990's, after the publication of the pioneer work by [10], the statistical signal processing community has studied extensively the use of high-order moments to process this type of signals. The moments (or cumulants) approach is attractive since in general it results in estimation techniques that are easy to implement. Nevertheless, high order moments estimators may suffer from their lack of efficiency. Here, we use a parametric approach in which a sufficiently broad family of distributions can be represented. Recent contributions in this domain include [11] which employed an autoregressive moving average (ARMA) model driven by Laplace noise to fit weekly data on sulphate concentration in a Canadian watershed. The generation of ARMA signals with specified marginal distributions was investigated by [12], [13] and [14], among others. ARMA signals driven by infinite variance noises with stable distributions were also studied. Applications of stable distributions in many area of signal processing, including blind channel identification and deconvolution and robust performing were presented by [15]. Recent contributions include radar cross section modeling applied to synthetic aperture radar image processing by [16], and texture information characterization in rotation-invariant image retrieval by [17].

The linear model considered in this paper is an AR model. Indeed, AR models are very popular in the signal processing community, see for instance [18] and [19]. They are used for instance for spectral analysis, for modeling speech and audio signals, and for identifying systems in control engineering. Causal AR models with independent innovations have the nice property that optimal (in the mean square sense) nonlinear infinite past predictors reduce to linear finite past predictors. Furthermore, the so-called Yule-Walker estimator of the parameters of an AR model can be easily calculated using the well-known Levinson-Durbin algorithm, leading to fast implementations. The Yule-Walker estimator is strongly consistent and asymptotically efficient when the innovations are Gaussian, see for instance [20, Chap. 8]. AR models with non-Gaussian innovations have been considered in the literature, see e.g. [21] and references therein. For instance, symmetric non-Gaussian innovations were considered by [22], asymmetric Laplace innovations were used by [23], and epsilon-skew-normal innovations were proposed by [24]. In this work, we consider SEP innovations, thus allowing more flexibility in the choice of the conditional distribution of the data. Our model is particularly suited for processes that are skewed and leptokurtic, but which appear to have finite higher order moments. The Yule-Walker estimator of the AR parameters is strongly consistent and asymptotically normal, but it fails to be asymptotically efficient since the innovations are non-Gaussian. Therefore, maximum likelihood estimators (MLE) of the parameters are proposed and their limit distributions are derived.

The rest of this paper is organized as follows. The AR model with SEP innovations is presented in Section 2. Generalized moments estimates and MLE of the parameters are established in Section 3. Numerical simulation results are presented in Section 4, and a real data modeling is considered in Section 5. Concluding remarks can be found in Section 6.

2. MODEL DESCRIPTION

Let $f_\alpha(x) = c\exp(-|x|^\alpha)\,\mathbb{1}_{\mathbb{R}}(x)$ where $\alpha > 0$ is a shape parameter, $c^{-1} = 2\Gamma(1+1/\alpha)$ and Γ is the Gamma function. Function f_α is the probability density function of the exponential power distribution, called generalized Gaussian distribution. It changes gradually from fat-tailed for $0 < \alpha < 2$, to short-tailed for $\alpha \geq 2$, as α increases. With $\alpha = 1$, $\alpha = 2$, and $\alpha \to +\infty$, f_α reduces to the Laplace, Gaussian and uniform densities, respectively. A skewed version of f_α is

$$f_{(\alpha,\epsilon)}(x) = f_\alpha\left[\frac{x}{1+\epsilon}\right]\mathbb{1}_{\{x<0\}} + f_\alpha\left[\frac{x}{1-\epsilon}\right]\mathbb{1}_{\{x\geq 0\}},$$

where $\epsilon \in (-1,1)$ is the skew parameter. We extend the family $f_{(\alpha,\epsilon)}$ to include location and scale parameters. Let Y be a random variable with density $f_{(\alpha,\epsilon)}$. The family of location-scale SEP distributions is defined as the distribution of $Z = \mu + \sigma Y$ for $\mu \in \mathbb{R}$ and $\sigma > 0$. The corresponding density is given by

$$f_\theta(x) = \frac{1}{\sigma}f_\alpha\left[\frac{x-\mu}{(1+\epsilon)\sigma}\right]\mathbb{1}_{\{x<\mu\}} + \frac{1}{\sigma}f_\alpha\left[\frac{x-\mu}{(1-\epsilon)\sigma}\right]\mathbb{1}_{\{x\geq\mu\}},$$

where $\theta = (\alpha,\epsilon,\mu,\sigma)$, and we denote $Z \sim \mathrm{SEP}(\alpha,\epsilon,\mu,\sigma)$. The distribution of Z is unimodal with mode at μ and it has probability mass $(1+\epsilon)/2$ below the mode. A useful stochastic representation of Z as the product of two independent random variables was established by [8], as well as the expressions of the median and the moments of Z. For $r > -1$, let

$$d_r = \int_{-\infty}^{+\infty} |x|^r f_\alpha(x)dx = \Gamma\left(\frac{r+1}{\alpha}\right)\Big/\Gamma\left(\frac{1}{\alpha}\right).$$

Let $Y \sim \mathrm{SEP}(\alpha,\epsilon,0,1)$ and $\mu_r = \mathrm{E}(Y^r)$. Then

$$\mu_r = \frac{(1-\epsilon)^{r+1} + (-1)^r(1+\epsilon)^{r+1}}{2}d_r.$$

Let $Z \sim \mathrm{SEP}(\alpha,\epsilon,\mu,\sigma)$. We have $Z = \mu + \sigma Y$, and then $Z - \mathrm{E}(Z) = \sigma(Y - \mu_1)$. The mean, variance, skewness and kurtosis of Z are respectively,

$$\begin{aligned} \mathrm{E}(Z) &= \mu - 2\epsilon\sigma d_1, \\ \mathrm{Var}(Z) &= \sigma^2[(1+3\epsilon^2)d_2 - 4(\epsilon d_1)^2], \\ \mathrm{E}(Z-\mathrm{E}(Z))^3 &= \sigma^3[\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3], \\ \mathrm{E}(Z-\mathrm{E}(Z))^4 &= \sigma^4[\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4]. \end{aligned} \tag{1}$$

For the wide ranges of skewness and kurtosis that can be obtained when the shape and skew parameters vary, see [8].

An $\mathrm{AR}(p)$ model with SEP innovations is defined by the difference equation

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t, \tag{2}$$

where (Z_t) is a sequence of independent and identically distributed (iid) random variables with $Z_t \sim \mathrm{SEP}(\alpha,\epsilon,\mu,\sigma)$, and polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ with real coefficients has no zeros in the closed unit disk $\{z \in \mathbb{C} : |z| \leq 1\}$. The (unique) stationary solution (X_t) of (2) has the MA representation $X_t = \sum_{i=0}^\infty \psi_i Z_{t-i}$, where (ψ_i) are the coefficients in the Taylor series expansion of $1/\phi(z)$ for $|z| \leq 1$. We have $\sum|\psi_i| < +\infty$, and then finiteness of $\mathrm{E}(|Z_t|^k)$ imply finiteness of $\mathrm{E}(|X_t|^k)$ for all $k \geq 1$.

3. PARAMETER ESTIMATION

Fitting model (2) to some data consists in choosing p and estimating the parameter vector $\eta = (\phi',\alpha,\epsilon,\mu,\sigma)'$ where $\phi = (\phi_1,\ldots,\phi_p)'$ and ϕ' denotes the transpose of ϕ. In the following, C is the interior of the domain of vectors ϕ such that $\phi(z)$ has no zeros in the closed unit disk. We shall first propose generalized moments estimates which will be used as initial values in a quasi-Newton method to find MLE.

The standard Yule-Walker equations for model (2) are

$$M_2\phi = m_2,$$

where M_2 is the invertible covariance matrix $[m_{2,i-j}]_{i,j=1}^p$, $m_2 = (m_{2,1},\ldots,m_{2,p})'$ and $m_{2,k} = \mathrm{E}(X_0X_k) - \mathrm{E}(X_0)^2$. The Yule-Walker estimator $\widehat{\phi}_n$ of parameter vector ϕ based on observations $(X_t)_{t=1}^n$ is $\widehat{\phi}_n = \widehat{M}_2^{-1}\widehat{m}_2$, where $\widehat{M}_2$ is the sample covariance matrix $[\widehat{m}_{2,|i-j|}]_{i,j=1}^p$, $\widehat{m}_2 = (\widehat{m}_{2,1},\ldots,\widehat{m}_{2,p})'$ with $\widehat{m}_{2,k} = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \overline{X})(X_{t+k} - \overline{X})$ and $\overline{X} = \frac{1}{n}\sum_{t=1}^n X_t$. According to [20, Theorem 8.1.1], $\widehat{\phi}_n \xrightarrow{a.s.} \phi$ and $n^{1/2}(\widehat{\phi}_n - \phi) \xrightarrow{d} \mathcal{N}(0, \mathrm{Var}(Z)M_2^{-1})$ as $n \to \infty$. Moreover, the covariance matrix $\mathrm{Var}(Z)M_2^{-1}$ depends only on the parameter vector ϕ.

Let

$$\widehat{Z}_t = X_t - \widehat{\phi}_{n,1}X_{t-1} - \cdots - \widehat{\phi}_{n,p}X_{t-p}.$$

The moments estimates of $(\alpha, \epsilon, \mu, \sigma)$ are obtained by solving (1) where the left-hand sides are replaced by the corresponding sample moments obtained from $(\widehat{Z}_t)$. More precisely, let

$$\widehat{c}_{k,Z} = \frac{1}{n}\sum_{t=p+1}^{n}\left(\widehat{Z}_t - \overline{\widehat{Z}}\right)^k \quad \text{where} \quad \overline{\widehat{Z}} = \frac{1}{n}\sum_{t=p+1}^{n}\widehat{Z}_t.$$

Observe that d_r and μ_r only depends on α and (α, ϵ), respectively. Then the standardized skewness and kurtosis defined respectively by $\mathrm{E}(Z-\mathrm{E}(Z))^3/\operatorname{Var}(Z)^{3/2}$ and $\mathrm{E}(Z-\mathrm{E}(Z))^4/\operatorname{Var}(Z)^2$ only depend on (α, ϵ). Generalized moments estimates $(\widehat{\alpha}_n, \widehat{\epsilon}_n)$ of (α, ϵ) are obtained by minimizing the sum of squares

$$w_1\left(\frac{\widehat{c}_{3,Z}}{\widehat{c}_{2,Z}^{3/2}} - \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{[(1+3\epsilon^2)d_2 - 4(\epsilon d_1)^2]^{3/2}}\right)^2 + w_2\left(\frac{\widehat{c}_{4,Z}}{\widehat{c}_{2,Z}^{2}} - \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{[(1+3\epsilon^2)d_2 - 4(\epsilon d_1)^2]^{2}}\right)^2,$$

where w_1 and w_2 are positive weights. Then estimates $(\widehat{\mu}_n, \widehat{\sigma}_n)$ of (μ, σ) are given by

$$\widehat{\mu}_n = \overline{\widehat{Z}} + 2\widehat{\epsilon}_n\widehat{\sigma}_n d_1(\widehat{\alpha}_n),$$
$$\widehat{\sigma}_n^2 = \widehat{c}_{2,Z}[(1+3\widehat{\epsilon}_n^2)d_2(\widehat{\alpha}_n) - 4(\widehat{\epsilon}_n d_1(\widehat{\alpha}_n))^2]^{-1}.$$

Since (Z_t) is non-Gaussian, $\widehat{\phi}_n$ is not asymptotically efficient in general. We now present the MLE of η, and in the following we suppose that η_0 is the true value of η. We consider the likelihood estimator based on maximization of the conditional likelihood of $(X_1, \ldots, X_n)$ conditionally to $(X_1, \ldots, X_p)$. According to (2), the logarithm of the conditional likelihood is

$$L_n(\eta) = \textstyle\sum_{t=p+1}^{n} \ln f_\theta(X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}).$$

Our main result is the following.

Theorem 1. Let (X_t) be defined by (2) where $\phi = \phi_0$ and (Z_t) are iid random variables with a $\mathrm{SEP}(\alpha_0, \epsilon_0, \mu_0, \sigma_0)$ distribution, and let $\eta_0 = (\phi_0', \alpha_0, \epsilon_0, \mu_0, \sigma_0)' \in C \times (1, +\infty) \times (-1, 1) \times (-\infty, +\infty) \times (0, +\infty)$. Then, there exists a sequence of estimators $(\widetilde{\eta}_n)$ such that, for any $\epsilon > 0$, there exists an event E with $\mathsf{P}(E) > 1 - \epsilon$ and an n_0 such that on E, for $n > n_0$, $\frac{\partial L_n}{\partial \eta}(\widetilde{\eta}_n) = 0$ and L_n attains a relative maximum at $\widetilde{\eta}_n$. Furthermore, as $n \to \infty$,

(i) $\widetilde{\eta}_n \xrightarrow{a.s.} \eta_0$,

(ii) $n^{1/2}(\widetilde{\eta}_n - \eta_0) \xrightarrow{d} \mathcal{N}(0, \Sigma)$, where

$$\Sigma = \sigma_0^2(1-\epsilon_0^2)\frac{\Gamma(1+1/\alpha_0)}{\alpha_0\Gamma(2-1/\alpha_0)}\left(\begin{array}{c|c} M_2^{-1} & \cdots \\ \hline \vdots & \cdots \end{array}\right).$$

(iii) The covariance matrix Σ can be estimated strongly consistently by replacing η_0 by $\widetilde{\eta}_n$ in its expression. One may also replace M_2 by the estimated covariance matrix $[\widehat{m}_{2,|i-j|}]_{i,j=1}^{p}$.

Remark 1. The MLE $\widetilde{\eta}_n$ is asymptotically efficient, i.e., Σ is the inverse of the Fisher information matrix of η evaluated at η_0.

4. MONTE CARLO SIMULATIONS

In this section, we illustrate the finite sample behavior of MLE by Monte Carlo simulations. All the experiments are based on 1000 replications, and the numbers of data considered are $n = 300$ and $n = 1000$. The data generating process is a causal AR(1) model with $Z_t \sim \mathrm{SEP}(\alpha, \epsilon, \mu, \sigma)$. We fix $(\mu, \sigma) = (10, 1)$ and we study the influence of ϕ, α and ϵ. We take $\phi = 0.3$ and $\phi = 0.8$ (simulation results show that there are no differences between $\phi > 0$ and $\phi < 0$), $\alpha = 1$, $\alpha = 2$ and $\alpha = 3$, and $\epsilon = -0.95, -0.9, -0.8, \ldots, 0.8, 0.9, 0.95$. The mean square errors (MSE) of the parameter estimates measure the performances of the estimators.

Figure 1 displays the MSE of the five parameters in $\widetilde{\eta}_n$ when $n = 1000$. The MSE of $\widetilde{\alpha}_n$ and $\widetilde{\epsilon}_n$ decrease as α decreases. Only the MSE of $\widetilde{\phi}_n$ and $\widetilde{\mu}_n$ seem to be notably influenced by ϕ. The larger ϕ is, the smaller the MSE of $\widetilde{\phi}_n$ is and the larger the MSE of $\widetilde{\mu}_n$ is. Parameter ϵ does not affect the MSE of $\widetilde{\alpha}_n$ and $\widetilde{\sigma}_n$. As the absolute value of ϵ decreases, the MSE of $\widetilde{\phi}_n$, $\widetilde{\epsilon}_n$ and $\widetilde{\mu}_n$ increase.

Figure 2 presents the MSE of the five parameters in $\widetilde{\eta}_n$ when $n = 300$ and $n = 1000$, and $\phi = 0.8$, $\alpha = 2$. To plot on a same figure the different MSE, we use the relative MSE which is the MSE divided by the square of the true value of the parameter (when $\epsilon = 0$, the relative MSE of $\widetilde{\epsilon}_n$ is not defined). Not surprisingly, the relative MSE increases as n decreases. We see that the relative MSE of $\widetilde{\epsilon}_n$ is the largest and the relative MSE of $\widetilde{\phi}_n$ and $\widetilde{\sigma}_n$ are the smallest.

5. REAL DATA EXAMPLE

We consider the Dow-Jones Utilities index between September 20, 1967 and November 26, 1968. This series is nonstationary, but the differenced series shows no obvious deviations from stationarity, see [25, Example 5.1.1]. Figure 3(c) and (d), as well as Akaike information criterion suggest to fit an AR(2) model to the differenced series. Figure 3(b) indicates that the differenced series is asymmetric, and this is confirmed by the rejection of the null hypothesis at the 99% significance level in the symmetry test of [26].

We fit an AR(2) model with SEP innovations to the differenced series and we compare with an AR(2) model with Gaussian innovations. The results are given in table 1. The 95% confidence intervals for α and ϵ deduced from table 1 are $(1.05, 1.21)$ and $(-0.26, -0.02)$, respectively. Therefore,

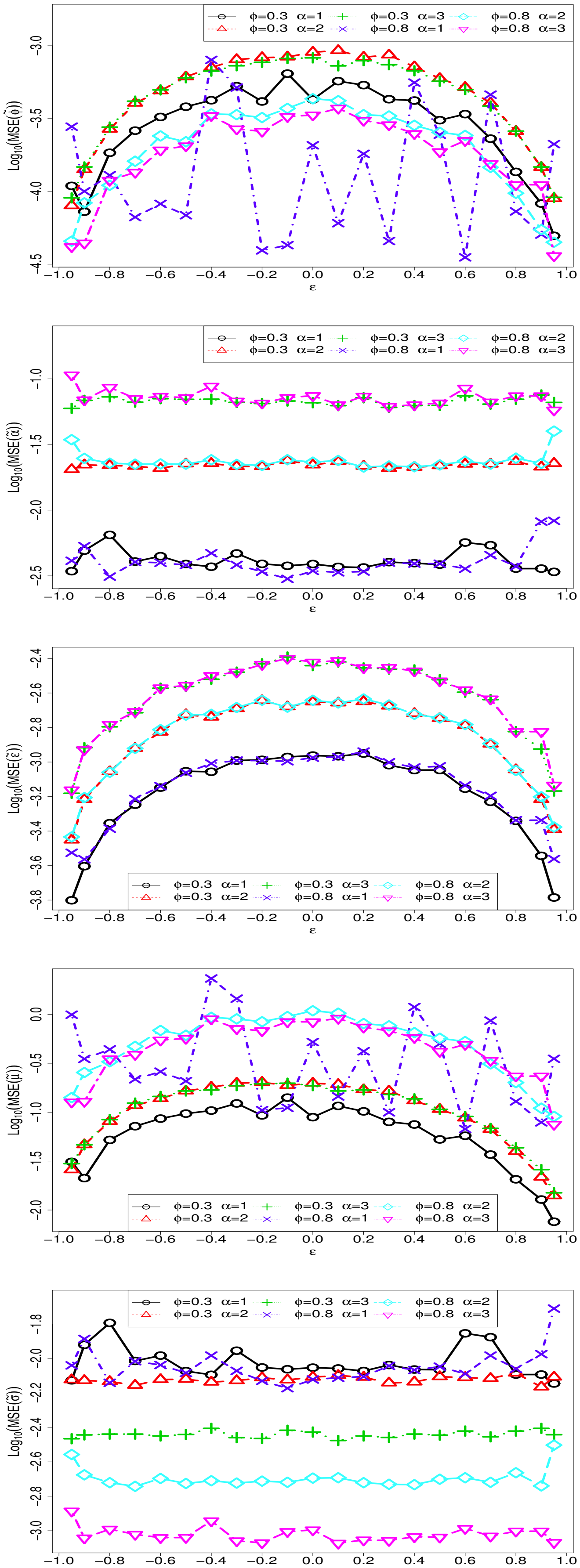

Fig. 1. MSE of $\widetilde{\phi}_n$, $\widetilde{\alpha}_n$, $\widetilde{\epsilon}_n$, $\widetilde{\mu}_n$ and $\widetilde{\sigma}_n$ when $n = 1000$.

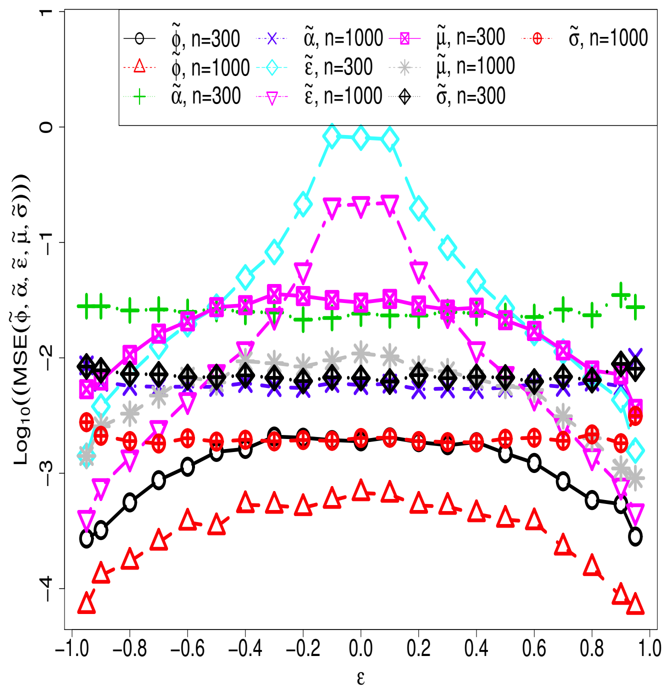

Fig. 2. Relative MSE of $\widetilde{\phi}_n$, $\widetilde{\alpha}_n$, $\widetilde{\epsilon}_n$, $\widetilde{\mu}_n$ and $\widetilde{\sigma}_n$ when $n = 300$ and $n = 1000$, and $\phi = 0.8$, $\alpha = 2$.

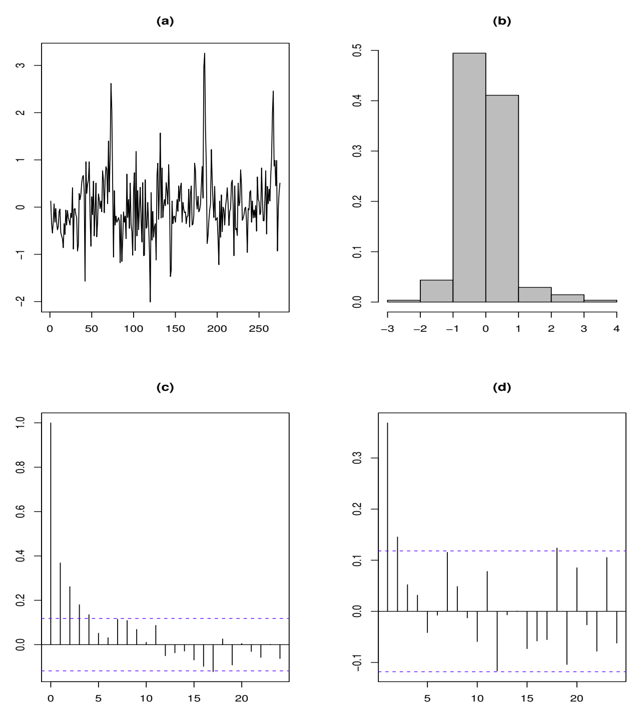

Fig. 3. Differenced series of the Dow-Jones Utilities index (Sep. 20, 1967 - Nov. 26, 1968): (a) Series, (b) Histogram, (c) Sample autocorrelation function, (d) Sample partial autocorrelation function.

η	SEP		Gaussian	
	Estimate	Variance	Estimate	Variance
ϕ_1	0.21	9.28e-5	0.32	1.64e-3
ϕ_2	0.18	9.28e-5	0.15	1.64e-3
α	1.13	1.58e-3	-	-
ϵ	-0.14	4.05e-3	-	-
μ	-0.07	1.06e-3	0.019	1.39e-3
σ	0.37	4.82e-4	0.63	6.93e-4

Table 1. AR(2) models with SEP and Gaussian innovations.

both Gaussianity and symmetry of the innovations are rejected.

Finally, we use the normality tests by Jarque-Bera, Shapiro-Francia, Shapiro-Wilk, Pearson (χ^2), Anderson-Darling, and Cramer-von Mises, to check the residuals of the Gaussian AR(2) model. The corresponding p-values are $2.2 \cdot 10^{-16}$, $8.9 \cdot 10^{-6}$, $6.2 \cdot 10^{-6}$, $1.1 \cdot 10^{-3}$, $6.5 \cdot 10^{-5}$ and $1.7 \cdot 10^{-4}$. Therefore, all tests reject the null hypothesis of normality at the 95% confidence level.

6. CONCLUSIONS

We have proposed an AR process with SEP innovations to model non-Gaussian asymmetric correlated data. The great flexibility of the SEP distribution allows to model a large class of data which are skewed and fat-tailed or short-tailed and have finite higher order moments.

7. REFERENCES

[1] M. T. Subbotin, "On the law of frequency of error," *Mat. Sb.*, vol. 31, no. 2, pp. 296–301, 1923.

[2] G. E. P. Box and G. C. Tiao, *Bayesian inference in statistical analysis*, Addison-Wesley, Reading MA, 1973.

[3] A. Azzalini, "A class of distributions which includes the normal ones," *Scand. J. Statist.*, vol. 12, no. 2, pp. 171–178, 1985.

[4] A. Azzalini, "Further results on a class of distributions which includes the normal ones," *Statistica (Bologna)*, vol. 46, no. 2, pp. 199–208, 1986.

[5] C. Fernández, J. Osiewalski, and M. F. J. Steel, "Modeling and inference with v-spherical distributions," *J. Amer. Statist. Assoc.*, vol. 90, no. 432, pp. 1331–1340, 1995.

[6] G. T. Fechner, *Kollectivmasslehre*, Engleman, Leipzig, 1897.

[7] G. S. Mudholkar and A. D. Hutson, "The epsilon-skew-normal distribution for analyzing near-normal data," *J. Statist. Plann. Inference*, vol. 83, no. 2, pp. 291–309, 2000.

[8] R. B. Arellano-Valle, H. W. Gómez, and F. A. Quintana, "Statistical inference for a general class of asymmetric distributions," *J. Statist. Plann. Inference*, vol. 128, no. 2, pp. 427–443, 2005.

[9] I. Komunjer, "Asymmetric power distribution: theory and applications to risk measurement," *J. Appl. Econometrics*, vol. 22, no. 5, pp. 891–921, 2007.

[10] K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, vol. 10, no. 4, pp. 1195–1208, 1982.

[11] E. Damsleth and A. H. El-Shaarawi, "ARMA models with double-exponentially distributed noise," *J. Roy. Statist. Soc. Ser. B*, vol. 51, no. 1, pp. 61–69, 1989.

[12] L. S. Dewald and P. A. W. Lewis, "A new Laplace second-order autoregressive time-series model—$\mathrm{NLAR}(2)$," *IEEE Trans. Inform. Theory*, vol. 31, no. 5, pp. 645–651, 1985.

[13] K. Jayakumar and A. P. Kuttykrishnan, "A time-series model using asymmetric Laplace distribution," *Statist. Probab. Lett.*, vol. 77, no. 16, pp. 1636–1640, 2007.

[14] B. Picinbono, "ARMA signals with specified symmetric marginal probability distribution," *IEEE Trans. Signal Process.*, vol. 58, no. 3, part 2, pp. 1542–1552, 2010.

[15] C. L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*, John Wiley & Sons Inc., New York, NY, USA, 1995.

[16] A. Achim, E. E. Kuruoglu, and J. Zerubia, "Sar image filtering based on the heavy-tailed rayleigh model," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2686–2693, 2006.

[17] G. Tzagkarakis, B. Beferull-Lozano, and P. Tsakalides, "Rotation-invariant texture retrieval with gaussianized steerable pyramids," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2702–2718, 2006.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice-Hall, Inc., Englewod Cliffs, NJ, USA, 1993.

[19] T. Chonavel, *Statistical Signal Processing : Modelling and Estimation*, Springer, London, UK, 2002.

[20] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*, Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991.

[21] G. K. Grunwald, R. J. Hyndman, L. Tedesco, and R. L. Tweedie, "Non-Gaussian conditional linear $\mathrm{AR}(1)$ models," *Aust. N. Z. J. Stat.*, vol. 42, no. 4, pp. 479–495, 2000.

[22] M. L. Tiku, W.-K. Wong, D. C. Vaughan, and G. Bian, "Time series models in non-normal situations: symmetric innovations," *J. Time Ser. Anal.*, vol. 21, no. 5, pp. 571–596, 2000.

[23] A. A. Trindade, Y. Zhu, and B. Andrews, "Time series models with asymmetric Laplace innovations," *J. Stat. Comput. Simul.*, vol. 80, no. 12, pp. 1317–1333, 2010.

[24] P. Bondon, "Estimation of autoregressive models with epsilon-skew-normal innovations," *J. Multivariate Anal.*, vol. 100, no. 8, pp. 1761–1776, 2009.

[25] P. J. Brockwell and R. A. Davis, *Introduction to time series and forecasting*, Springer Texts in Statistics. Springer-Verlag, New York, second edition, 2002.

[26] J. S. Racine and E. Maasoumi, "A versatile and robust metric entropy test of time-reversibility, and other hypotheses," *J. Econometrics*, vol. 138, no. 2, pp. 547–567, 2007.