STUDYING THREE FAMILIES OF DIVERGENCES TO COMPARE WIDE-SENSE STATIONARY GAUSSIAN ARMA PROCESSES

Eric Grivel

Bordeaux University - INP Bordeaux, IMS - UMR CNRS 5218, FRANCE

ABSTRACT

In this paper, we aim at analyzing the differences between three families of divergences used to compare probability density functions of Gaussian random vectors storing k consecutive samples of wide-sense stationary ARMA processes. There may be various applications: signal classification, statistical change detection, etc. Among the families that are studied, we propose to look at the α-divergence, the β-divergence and the γ-divergence. We first provide the expression of the divergences in the Gaussian case and then express their divergence increments, *i.e.* the differences between the divergences computed for $k+1$ and k consecutive samples. Finally, we analyze how these divergence increments evolve when k increases and tends to infinity.

***Index Terms*—** divergences, rate, ARMA processes.

1. INTRODUCTION

Autoregressive moving average (ARMA) processes are very popular in statistical signal processing for various reasons: it can be useful to model and analyze short-memory signals such as a speech signal, a clutter, a biomedical signal and a channel in a mobile-communication system. The ARMA parameters and the corresponding power spectral density (PSD) can serve as features useful to characterize the signals. Then, distances based on these features can be considered to classify them. An alternative solution is to use divergences whose purpose is to compare probability density functions. Note that divergences have many applications in deep learning methods, image processing, information theory. See [17, 14, 6, 1].

The Kullback-Leibler (KL) divergence is probably the most popular, but many others exist. For some of them, KL divergence is a special case. Thus, families of divergences have emerged such as the γ-divergence, those based on the Chernoff coefficient leading to the α-divergence and the $\alpha - \gamma$ divergence, the Bregman divergences including the β-divergence, and finally the Csiszar divergences. Symmetric versions were also developed and links between all of them exist. See [2] for instance for more details.

The recent theoretical contributions deal with various topics: Nielsen and other authors proposed closed-form expressions of divergences for some families of distributions or studied generalizations of existing divergences [20, 21, 5, 13, 26, 12]. The estimation of the divergence for Gaussian or non-Gaussian data was also investigated [19, 4, 25, 22, 3]. The "rates"[1] of the KL and Rényi divergences for zero-mean wide-sense-stationary (w.s.s.) Gaussian processes were respectively studied in [11] and [7]. They were expressed in terms of integrals of functions depending on the PSDs of the processes. As an alternative, in [16, 10, 15, 23, 9], by combining the definitions of the divergences in the Gaussian case and the statistical properties of the processes to be compared, we derived the rates of the Jeffreys and Rényi divergences between the probability density functions (pdfs) of vectors storing consecutive samples of Gaussian ARMA or AR fractionally integrated MA (ARFIMA) processes -noisy or not-. In addition, we suggested using the divergence rate as a tool to analyze the similarity between stochastic processes. More recently, in [8], we studied the Chernoff, Bhattacharyya and Sharma-Mittal divergences. We showed that the asymptotic normalized increment of the Sharma-Mittal divergence could be used to compare w.s.s. Gaussian ARMA processes. In this paper, our purpose is to see if other divergence increments can be considered to compare w.s.s. Gaussian ARMA processes. Thus, we propose to analyze the asymptotic increments of alternative divergences, such as the family of α-divergences[2], the β-divergence and the γ-divergence.

The remainder of this paper is organized as follows: in section 2, properties of ARMA processes are recalled. In section 3, the definitions of the divergences and their expressions in the Gaussian case are given. The difference between the divergences computed for $k+1$ and k variates is studied. Its value is then deduced when k tends to infinity. Comments are made and an illustration is provided.

2. ABOUT GAUSSIAN REAL ARMA PROCESS

For $i = 1, 2$, the t^{th} sample $x_{t,i}$ of the i^{th} real ARMA process of order (p_i, q_i) is given by:

$$x_{t,i} = -\sum_{j=1}^{p_i} a_{j,i} x_{t-j,i} + \sum_{j=0}^{q_i} b_{j,i} u_{t-j,i} \tag{1}$$

[1] It corresponds to the asymptotic increment, *i.e.* when k tends to infinity.

[2] It includes the Square Hellinger discrimination. The α-divergence is also equal to the Tsallis divergence up to a multiplicative factor and is related to the Chernoff and the Rényi divergences through a non-linear function.

where $\{a_{j,i}\}_{j=1,\dots,p_i}$ and $\{b_{j,i}\}_{j=0,\dots,q_i}$ are the ARMA parameters with $b_{0,i}=1$ and the driving process $u_{t,i}$ is a zero-mean w.s.s. Gaussian white sequence with variance $\sigma^2_{u,i}$.

As the ARMA parameters are constant, the ARMA process is also w.s.s. In the following, a non zero-mean w.s.s. ARMA process is obtained by adding $\mu_{x,i}$ to $x_{t,i}$, once generated.

The process $x_{t,i}$ can be seen as the output of an infinite-impulse-response stable linear filter whose input is $u_{t,i}$. Its transfer function $H_i(z) = \frac{\prod_{l=1}^{q_i}(1-z_{l,i}z^{-1})}{\prod_{l=1}^{p_i}(1-p_{l,i}z^{-1})}$ is defined by its zeros $\{z_{l,i}\}_{l=1,\dots,q_i}$ and its poles $\{p_{l,i}\}_{l=1,\dots,p_i}$.

When the moduli of the zeros are smaller than 1, one speaks of minimum-phase ARMA processes. For every non-minimum phase ARMA process, the equivalent minimum-phase process is obtained by replacing the zeros $\{z_{l,i}\}_{l=1,\dots,m_i\leq q_i}$ whose modulus is larger than 1 by $1/z^*_{l,i}$ for $l=1,\dots,m_i$ to get the transfer function $H_{min,i}(z)$. Then, one substitutes the variance $\sigma^2_{u,i}$ with:

$$\sigma^2_{u,min,i} = \sigma^2_{u,i}\prod_{l=1}^{q_i} K_{l,i} \tag{2}$$

with $K_{l,i} = \begin{cases} |z_{l,i}|^2 \text{ for the modified zeros} \\ 1 \text{ otherwise} \end{cases}$.

All the above ARMA processes lead to the same PSD equal to $S_{x,i}(\theta) = \sigma^2_{u,i}\,|H_i(e^{j\theta})|^2 = \sigma^2_{u,min,i}\,|H_{min,i}(e^{j\theta})|^2$ with θ the normalized angular frequency and consequently to the same correlation and covariance functions.

As our purpose is to compare the pdfs of ARMA processes characterized by their means and their covariance matrices, one can consider the ARMA processes associated with the minimum-phase filter [24]. Comparing the minimum-phase ARMA processes is relevant because they can be represented by an infinite-order AR process. The latter can be approximated by a AR model of finite-order $\tau > \max\,(p_i, q_i)$:

$$x_{t,i} \approx -\sum_{j=1}^{\tau} \alpha_{j,\tau,i}x_{t-j,i} + u_{t,\tau,i} \tag{3}$$

where $u_{t,\tau,i}$ is the driving process of the i^{th} process associated with the order τ whose variance is $\sigma^2_{u,\tau,i}$ and $\{\alpha_{j,\tau,i}\}_{j=1,\dots,\tau}$ are the AR parameters that can be estimated using the Yule-Walker equations [18].

In addition, the Toeplitz covariance matrix $Q_{\tau,i}$ of the vector $X_{\tau,i}$ storing τ consecutive samples is non singular even if the PSD of the process is equal to zero at some frequencies. It is hence invertible. Only the infinite-size Toeplitz covariance matrix is not invertible when the transfer function has unit zeros.

By using results on the determinants, such as the matrix determinant lemma, it can be shown that the determinant of the covariance matrix of size $k+1$ can be expressed this way:

$$\begin{aligned} |Q_{k+1,i}| &= \begin{vmatrix} Q_{k,i} & R_{k,i} \\ R^T_{k,i} & r_{0,i} \end{vmatrix} = r_{0,i} \times |Q_{k,i} - R_{k,i}\frac{1}{r_{0,i}}R^T_{k,i}| \\ &= |Q_{k,i}| \times (r_{0,i} - R^T_{k,i}Q^{-1}_{k,i}R_{k,i}) \end{aligned} \tag{4}$$

where $R_{k,i}$ is a column vector storing the covariance function from lags k down to 1 and $r_{0,i}$ is the covariance function for lag 0. Then, after some mathematical development combining (4) and the Yule-Walker equations, one has:

$$\frac{|Q_{k+1,i}|}{|Q_{k,i}|} = r_{0,i} - R^T_{k,i}Q^{-1}_{k,i}R_{k,i} = \sigma^2_{u,k,i} \tag{5}$$

Therefore, when k increases and tends to infinity, one gets:

$$\lim_{k\to+\infty}\frac{|Q_{k,i}|}{|Q_{k-1,i}|} = \sigma^2_{u,min,i} \tag{6}$$

As $\sigma^2_{u,min,i} < \sigma^2_{u,k,i}$ for any k [18], $|Q_{k,i}| > \sigma^{2k}_{u,min,i}$.

3. DIVERGENCES UNDER STUDY

3.1. Definitions

Let us first introduce the following two notations:

$$\begin{cases} I_1(\mathsf{p}_1,\mathsf{p}_2,a,b) = \int_{X_k}\mathsf{p}_1^a(X_k)\mathsf{p}_2^b(X_k)dX_k \\ I_2(\mathsf{p}_1,a) = \int_{X_k}\mathsf{p}_1^a(X_k)dX_k \end{cases} \tag{7}$$

To study the dissimilarities between two pdfs denoted[3] p_1 and p_2, the KL divergence is defined by:

$$KL_k^{(1,2)} = \int_{X_k}\mathsf{p}_1(X_k)\ln\left(\frac{\mathsf{p}_1(X_k)}{\mathsf{p}_2(X_k)}\right)\mathrm{d}X_k \tag{8}$$

Three alternative families of methods will be addressed in this paper. Thus, let us first introduce the γ-divergence denoted as $G_k^{(1,2)}(\gamma)$ and defined as:

$$G_k^{(1,2)}(\gamma) = \frac{1}{\gamma(\gamma-1)}\ln\left(\frac{I_2(\mathsf{p}_1,\gamma)I_2^{\gamma-1}(\mathsf{p}_2,\gamma)}{I_1^{\gamma}(\mathsf{p}_1,\mathsf{p}_2,1,\gamma-1)}\right) \tag{9}$$

A second family includes the β-divergence denoted as $BeD_k^{(1,2)}(\beta)$ that can be expressed as follows:

$$\begin{aligned} BeD_k^{(1,2)}(\beta) &= \frac{1}{\beta(\beta-1)}I_2(\mathsf{p}_1,\beta) + \frac{1}{\beta}I_2(\mathsf{p}_2,\beta) \\ &\quad - \frac{1}{\beta-1}I_1(\mathsf{p}_1,\mathsf{p}_2,1,\beta-1) \end{aligned} \tag{10}$$

It should be noted that among the particular cases, the KL divergence is retrieved if β tends to 1.

Finally, the last family of divergences includes the Chernoff and Rényi divergences and can be expressed from the Chernoff coefficient C_α of order $0<\alpha<1$. The latter is given by:

$$C_\alpha\big(\mathsf{p}_1(X_k),\mathsf{p}_2(X_k)\big) = I_1(\mathsf{p}_1,\mathsf{p}_2,\alpha,1-\alpha) \tag{11}$$

Here, we focus our attention on the α-divergence, $A_k^{(1,2)}(\alpha)$:

$$A_k^{(1,2)}(\alpha) = -\frac{1}{\alpha(1-\alpha)}\Big(C_\alpha\big(\mathsf{p}_1(X_k),\mathsf{p}_2(X_k)\big) - 1\Big) \tag{12}$$

Note that the α-divergence, the Chernoff divergence $CD_k^{(1,2)}(\alpha)$, the Rényi divergence $RD_k^{(1,2)}(\alpha)$ are related as follows:

$$\begin{aligned} A_k^{(1,2)}(\alpha) &= -\frac{1}{\alpha(1-\alpha)}\Big(\exp\big(-CD_k^{(1,2)}(\alpha)\big) - 1\Big) \\ &= -\frac{1}{\alpha(1-\alpha)}\Big(\exp\big((\alpha-1)RD_k^{(1,2)}(\alpha)\big) - 1\Big) \end{aligned} \tag{13}$$

[3] p_1 denotes one of the AR order while p_1 denotes a pdf.

Using the L'Hôspital rule, it can be shown that the α-divergence tends to the KL between p_1 and p_2 when α tends to 1 and to the KL between p_2 and p_1 when α tends to 0. Given (12), the Tsallis divergence $T_k^{(1,2)}(\alpha)$ and the reversed Tsallis divergence $rT_k^{(1,2)}(\alpha)$ are related to the α-divergence:

$$\begin{cases} T_k^{(1,2)}(\alpha) & = -\frac{C_\alpha\left(\mathsf{p}_1(X_k),\mathsf{p}_2(X_k)\right)-1}{1-\alpha} = \alpha A_k^{(1,2)}(\alpha) \\ rT_k^{(1,2)}(\alpha) & = -\frac{C_\alpha\left(\mathsf{p}_2(X_k),\mathsf{p}_1(X_k)\right)-1}{1-\alpha} \end{cases} \tag{14}$$

Given the definitions of these various divergences, let us now look at their expressions in the Gaussian case.

3.2. Expressions in the Gaussian case

Starting from the expression of the pdf in the Gaussian case for a real random vector of size k characterized by $\mu_{k,i} = E[X_{k,i}]$ the mean vector and $|Q_{k,i}|$ the determinant of the covariance matrix $Q_{k,i}$, let us introduce three terms: the column vector corresponding to the mean difference, the weighted sum of $Q_{k,1}$ and $Q_{k,2}$ and a quadratic form:

$$\begin{cases} \Delta\mu_k = \mu_{k,2} - \mu_{k,1} \\ Q_{k,wei,1-\alpha,\alpha} = (1-\alpha)Q_{k,1} + \alpha Q_{k,2} \\ h_{k,1-\alpha,\alpha} = \Delta\mu_k^T Q_{k,wei,1-\alpha,\alpha}^{-1}\Delta\mu_k \end{cases} \tag{15}$$

Note that similar definitions for $Q_{k,wei,1-\alpha,\alpha}$ and $h_{k,1-\alpha,\alpha}$ can be obtained by replacing $1-\alpha$ and α by other values.

After some mathematical developments we will not detail in this conference paper, we can show that for $0 < \alpha < 1$:

$$A_k^{(1,2)}(\alpha) = \frac{1}{\alpha(1-\alpha)}\Big(1 - \frac{|Q_{k,1}|^{\frac{1-\alpha}{2}}|Q_{k,2}|^{\frac{\alpha}{2}}}{|Q_{k,wei,1-\alpha,\alpha}|^{\frac{1}{2}}} \times \exp\left[\frac{\alpha(\alpha-1)}{2}h_{k,1-\alpha,\alpha}\right]\Big) \tag{16}$$

Given (14), $T_k^{(1,2)}(\alpha)$ can be then easily deduced. Moreover, one has:

$$rT_k^{(1,2)}(\alpha) = \frac{1}{1-\alpha}\times \Big(1 - \frac{|Q_{k,1}|^{\frac{\alpha}{2}}|Q_{k,2}|^{\frac{1-\alpha}{2}}}{|Q_{k,wei,\alpha,1-\alpha}|^{\frac{1}{2}}}\exp\left[\frac{\alpha(\alpha-1)h_{k,\alpha,1-\alpha}}{2}\right]\Big) \tag{17}$$

For $\beta > 1$, one can show that the β-divergence is given by:

$$\begin{aligned} &BeD_k^{(1,2)}(\beta) = \\ &\frac{1}{\beta^{\frac{k}{2}+1}(2\pi)^{\frac{k(\beta-1)}{2}}}\left(\frac{1}{(\beta-1)|Q_{k,1}|^{\frac{(\beta-1)}{2}}} + \frac{1}{|Q_{k,2}|^{\frac{(\beta-1)}{2}}}\right) \\ &-\frac{1}{\beta-1}\frac{1}{(2\pi)^{\frac{k(\beta-1)}{2}}|Q_{k,wei,\beta-1,1}|^{\frac{1}{2}}|Q_{k,2}|^{\frac{\beta-2}{2}}} \\ &\times\exp\left(-\frac{(\beta-1)}{2}h_{k,\beta-1,1}\right) \end{aligned} \tag{18}$$

Finally, with $\gamma > 1$, one has:

$$G_k^{(1,2)}(\gamma) = \ln\left(\frac{|Q_{k,wei,\gamma-1,1}|^{\frac{1}{2(\gamma-1)}}}{|Q_{k,1}|^{\frac{1}{2\gamma}}|Q_{k,2}|^{\frac{1}{2\gamma(\gamma-1)}}}\right) - \frac{k\ln\gamma}{2(\gamma-1)} + \frac{1}{2}h_{k,\gamma-1,1} \tag{19}$$

In the next subsection, our purpose is to express the difference between the divergences computed for $k+1$ and k variates. This is the increment. Our purpose will be then to analyze how it evolves when k increases and tends to infinity.

3.3. Expressions of the increments in the Gaussian case

First, we propose to look at the increment of the α-divergence $\Delta A_k^{(1,2)}(\alpha) = A_{k+1}^{(1,2)}(\alpha) - A_k^{(1,2)}(\alpha)$. Of course (16) could be considered. Otherwise, using (13), one has:

$$\Delta A_k^{(1,2)}(\alpha) = -\frac{\exp\left[(\alpha-1)RD_k^{(1,2)}(\alpha)\right]}{\alpha(1-\alpha)} \times\Big(\exp\left[(\alpha-1)\Delta RD_k^{(1,2)}(\alpha)\right]-1\Big) \tag{20}$$

Regarding the analytical expression of the β-divergence increment, we propose to write it this way:

$$\Delta BeD_k^{(1,2)}(\beta) = \sum_{i=1}^{3}\delta BeD_{k,i}^{(1,2)} \tag{21}$$

with the three terms $\{\delta BeD_{k,i}^{(1,2)}(\beta)\} = i = 1, ..., 3$ being respectively defined by:

$$\delta BeD_{k,1}^{(1,2)} = B_{k,1}\left(\frac{1}{\beta^{\frac{1}{2}}(2\pi)^{\frac{(\beta-1)}{2}}}\frac{|Q_{k,1}|^{\frac{(\beta-1)}{2}}}{|Q_{k+1,1}|^{\frac{(\beta-1)}{2}}} - 1\right) \tag{22}$$

with $B_{k,1} = \frac{1}{(\beta-1)\beta^{\frac{k}{2}+1}(2\pi)^{\frac{k(\beta-1)}{2}}|Q_{k,1}|^{\frac{(\beta-1)}{2}}}$.

$$\delta BeD_{k,2}^{(1,2)} = B_{k,2}\left(\frac{1}{\beta^{\frac{1}{2}+1}(2\pi)^{\frac{(\beta-1)}{2}}}\frac{|Q_{k,2}|^{\frac{(\beta-1)}{2}}}{|Q_{k+1,2}|^{\frac{(\beta-1)}{2}}} - 1\right) \tag{23}$$

with $B_{k,2} = \frac{1}{\beta^{\frac{k}{2}+1}(2\pi)^{\frac{k(\beta-1)}{2}}|Q_{k,2}|^{\frac{(\beta-1)}{2}}}$.

$$\begin{aligned} \delta BeD_{k,3}^{(1,2)} &= B_{k,3}\exp\left(-\frac{(\beta-1)}{2}h_{k,\beta-1,1}\right)\times \\ &\Big[\frac{|Q_{k,wei,\beta-1,1}|^{\frac{1}{2}}|Q_{k,2}|^{\frac{\beta-2}{2}}}{(2\pi)^{\frac{(\beta-1)}{2}}|Q_{k+1,wei,\beta-1,1}|^{\frac{1}{2}}|Q_{k+1,2}|^{\frac{\beta-2}{2}}} \\ &\exp\big(-\frac{(\beta-1)}{2}(h_{k+1,\beta-1,1}-h_{k,\beta-1,1})\big)-1\Big] \end{aligned} \tag{24}$$

with $B_{k,3} = -\frac{1}{\beta-1}\frac{1}{(2\pi)^{\frac{k(\beta-1)}{2}}|Q_{k,wei,\beta-1,1}|^{\frac{1}{2}}|Q_{k,2}|^{\frac{\beta-2}{2}}}$.

Finally, provided $\gamma > 1$, the increment $\Delta G_k^{(1,2)}(\gamma)$ for the γ-divergence is equal to:

$$\Delta G_k^{(1,2)}(\gamma) = \frac{h_{k+1,\gamma-1,1} - h_{k,\gamma-1,1}}{2} - \frac{\ln\gamma}{2(\gamma-1)} + \frac{1}{2\gamma(\gamma-1)}\ln\left(\frac{|Q_{k,2}||Q_{k+1,wei,\gamma-1,1}|^{\gamma}|Q_{k,1}|^{\gamma-1}}{|Q_{k+1,2}||Q_{k,wei,\gamma-1,1}|^{\gamma}|Q_{k+1,1}|^{\gamma-1}}\right) \tag{25}$$

In the next section, let us analyze how these increments evolve when k increases when comparing ARMA processes. To this end, the results presented in section 2 are used.

3.4. Asymptotic analysis when k increases and tends to infinity when comparing Gaussian w.s.s. ARMA processes

3.4.1. Preamble

The ARMA processes being w.s.s., their statistical means are unchanged over time. If $\Delta\mu_x$ is the difference between the process means, various approaches for instance presented in [9] showed that $\Delta h_{k,1-\alpha,\alpha} = h_{k+1,1-\alpha,\alpha} - h_{k,1-\alpha,\alpha}$ satisfies:

$$\Delta h_{k,1-\alpha,\alpha} = \frac{|\sum_{i=0}^{k} \alpha_{i,k,wei,1-\alpha,\alpha}|^2}{\sigma_{u,k,1}^2} \Delta\mu_x^2 \quad (26)$$

where $\alpha_{i,k,wei,1-\alpha,\alpha}$ denotes the i^{th} AR parameter of the k^{th}-order AR process approximating the ARMA processes whose covariance matrix is $Q_{k,wei,1-\alpha,\alpha}$.
Taking the limit of the difference $\Delta h_{k,1-\alpha,\alpha}$ when k tends to infinity, we can see that the asymptotic difference is equal to the power $P^{\Delta\mu_x,wei,1-\alpha,\alpha}$ of a constant signal equal to $\Delta\mu_x$ filtered by the inverse filter associated with the minimum-phase ARMA process whose covariance matrix is $Q_{k,wei,1-\alpha,\alpha}$:

$$\lim_{k\to+\infty} \Delta h_{k,1-\alpha,\alpha} = P^{\Delta\mu_x,wei,1-\alpha,\alpha} \quad (27)$$

Based on the above result, let us now analyze the asymptotic increments for the divergences under study.

3.4.2. Asymptotic increment for the γ-divergence

Given the preamble 3.4.1 and the information about ARMA processes presented in section 2 such as (6), one obtains the asymptotic increment for the γ-divergence. It corresponds also to the divergence rate:

$$\Delta G^{(1,2)}(\gamma) = \lim_{k\to+\infty} \Delta G_k^{(1,2)}(\gamma) = \frac{1}{2} P^{\Delta\mu_x,wei,\gamma-1,1} \quad (28)$$
$$+ \frac{1}{\gamma(\gamma-1)} \ln\left(\frac{\sigma_{u,wei,min,\gamma-1,1}^{\gamma}}{\sigma_{u,2,min}\sigma_{u,1,min}^{\gamma-1}}\right) - \frac{\ln\gamma}{2(\gamma-1)}$$

where $\sigma_{u,wei,min,\gamma-1,1}$ is the standard deviation of the driving process of the minimum-phase ARMA process whose correlation matrix for any k is equal to $(\gamma-1)Q_{k,1} + Q_{k,2}$. In addition, $P^{\Delta\mu_x,wei,\gamma-1,1}$ is the power of the constant signal equal to the mean difference that has been filtered by the inverse filter associated with the minimum phase ARMA process whose covariance matrix is $Q_{k,wei,\gamma-1,1}$.
Note that the above asymptotic increment depends on the ARMA parameters of both processes because $\sigma_{u,wei,min,\gamma-1,1}^2$ depends on the ARMA parameters and the variances of the driving processes of the two processes to be compared.

3.4.3. Asymptotic increment of the β-divergence

Let us focus our attention on the limit $\delta BeD_1^{(1,2)}$ of $\delta BeD_{k,1}^{(1,2)}$ when k increases and tends to infinity. Given (6), one has:

$$\delta BeD_1^{(1,2)} = \frac{1}{\beta(\beta-1)}\left(\frac{1}{\beta^{\frac{1}{2}}(2\pi)^{\frac{(\beta-1)}{2}}\sigma_{u,min,1}^{\beta-1}} - 1\right) \quad (29)$$
$$\times \lim_{k\to+\infty} \frac{1}{\beta^{\frac{k}{2}}(2\pi)^{\frac{k(\beta-1)}{2}}|Q_{k,1}|^{\frac{(\beta-1)}{2}}}$$

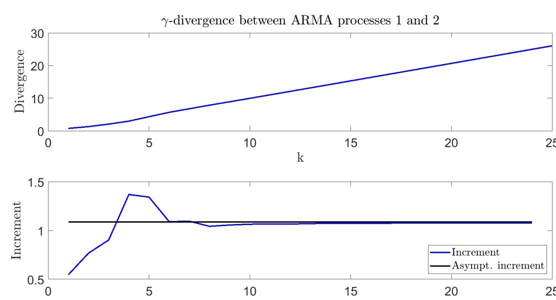
Given the remark after (6), one has $\frac{1}{|Q_{k,i}|^{\frac{(\beta-1)}{2}}} < \frac{1}{\sigma_{u,min,i}^{k(\beta-1)}}$. Therefore, if $\beta^{\frac{1}{2}}(2\pi)^{\frac{\beta-1}{2}}\sigma_{u,min,1}^{\beta-1} > 1$, $\delta BeD_1^{(1,2)}$ is null. Similarly, if $\beta^{\frac{1}{2}}(2\pi)^{\frac{\beta-1}{2}}\sigma_{u,min,2}^{\beta-1} > 1$, $\delta BeD_2^{(1,2)}$ is also null. When both constraints are satisfied, $\delta BeD_3^{(1,2)}$ is also null. So, $\Delta BeD^{(1,2)}(\beta)$ the asymptotic increment of the β-divergence becomes also null. As the null value can be often obtained, the asymptotic increment is not relevant to discriminate ARMA processes.

3.4.4. Asymptotic increment for the α-divergence

Concerning the asymptotic increment of the α-divergence, as the Rényi divergence tends to infinity when the processes are dissimilar and k tends to infinity [9], the α-divergence tends to $\frac{1}{\alpha(1-\alpha)}$ due to (13). Therefore, the asymptotic increment $\Delta A^{(1,2)}$ is equal to 0. Moreover, The Tsallis and reversed Tsallis divergences tend to $\frac{1}{1-\alpha}$ when k increases due to (13) and (14). Therefore, the asymptotic increments $\Delta T^{(1,2)}(\alpha)$ and $\Delta rT^{(1,2)}(\alpha)$ also tend to 0, for $\alpha \in]0,1[$.
Using these features to compare processes is not relevant.

3.4.5. Illustration

Two real ARMA(4,6) processes are compared. Their parameters have been chosen arbitrarily. The first one is defined by two zeros $0.8e^{j\frac{\pi}{4}}$ and $0.9e^{j\frac{3\pi}{4}}$ and their conjugates and three poles $0.9e^{j\frac{\pi}{6}}$, $0.7e^{j\frac{\pi}{5}}$ and $0.9e^{j\frac{\pi}{4}}$ and their conjugates. $\sigma_{u,1}^2 = 1$ and $\mu_{x,1} = 5$. The second one is defined by two zeros $0.8e^{j\frac{2\pi}{3}}$ and $0.6e^{j\frac{3\pi}{7}}$ and their conjugates and three poles $0.95e^{j\frac{5\pi}{6}}$, $0.7e^{j\frac{3\pi}{5}}$ and $0.9e^{j\frac{\pi}{4}}$ and their conjugates. $\sigma_{u,2}^2 = 4$ and $\mu_{x,2} = 1$. The correlation matrices and the means are estimated from the data. The divergences for different values of k are then estimated, their increments are deduced and compared with the limit presented in the paper. For the sake of space, only the γ-divergence is presented below with $\gamma = 2$:



4. CONCLUSIONS AND PERSPECTIVES

For many divergences, the asymptotic increment is not null and can serve as a tool to compare w.s.s. ARMA processes. The γ-divergence also has this property. When the ARMA processes have a zero mean or the same mean, the divergence rate depends on γ and the ARMA parameters of the processes to be compared. As the asymptotic increments of the β-divergence and the α-divergence can be often null, they cannot be used to compare Gaussian ARMA processes.

5. REFERENCES

[1] L. Bombrun, N. E. Lasmar, Y. Berthoumieu, and G. Verdoolaege. Multivariate texture retrieval using the SIRV representation and the geodesic distance. *IEEE ICASSP*, pages 865–868, 2011.

[2] A. Cichocki and S.-I. Amari. Families of alpha- beta- and gamma- divergences: Flexible and robust measures of similarities. *Entropy*, 12:1532–1568, 2010.

[3] J. E. Contreras-Reyes. Analyzing fish condition factor index through skew-gaussian information theory quantifiers. *Fluctuation and Noise Letters*, 15 (2):1–16, 2016.

[4] J. E. Contreras-Reyes and R. B. Arellano-Valle. Kullback–Leibler divergence measure for multivariate skew-normal distributions. *Entropy*, 14 (9):1606–1626, 2012.

[5] D. C. de Souza, R. F. Vigelis, and C. C. Cavalcante. Geometry induced by a generalization of Rényi divergence. *Entropy*, 18, 407, 2015.

[6] T. T. Georgiou and A. Lindquist. A convex optimization approach to ARMA modeling. *IEEE Transactions on Automatic Control*, 53:1108–1119, 2008.

[7] M. Gil. On Rényi divergence measures for continuous alphabet sources. *PhD Thesis, Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada*, 2011.

[8] E. Grivel. Chernoff, Bhattacharyya, Rényi and Sharma-Mittal divergence analysis for Gaussian stationary ARMA processes. *GSI*, pages 487–495, 2021.

[9] E. Grivel, R. Diversi, and F. Merchan. Kullback-Leibler and Rényi divergence rate for Gaussian stationary ARMA processes comparison. *Elsevier Digital Signal Processing*, 116:103089, 2021.

[10] E. Grivel, M. Saleh, and S.-M. Omar. Interpreting the asymptotic increment of Jeffrey's divergence between some random processes. *Elsevier Digital Signal Processing*, 75, (4):120–133, 2018.

[11] S. Ihara. *Information Theory for Continuous Systems*. World Scientific, 1993.

[12] P. Kluza. On Jensen-Rényi and Jeffreys-Rényi type f-divergences induced by convex functions. *Physica A: Statistical Mechanics and its Applications*, 548:122527, 2020.

[13] P. Kluza and M. Niezgoda. Generalizations of Crooks and Lin's results on Jeffreys-Csiszár and Jensen-Csiszár f-divergences. *Physica A: Statistical Mechanics and its Applications*, 463:383–393, 2016.

[14] M. Lasta and R. Shumway. Detecting abrupt changes in a piecewise locally stationary time series. *Journal of Multivariate Analysis*, 99:191–214, 2008.

[15] L. Legrand and E. Grivel. Jeffrey's divergence between autoregressive processes disturbed by additive white noises. *Signal Processing*, 149:162–178, 2018.

[16] C. Magnant, E. Grivel, A. Giremus, B. Joseph, and L. Ratton. Jeffrey's divergence for state-space model comparison. *Signal Processing*, 114:61–74, September 2015.

[17] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis. Touchless monitoring of breathing function. *IEEE EMBS*, pages 1196–1199, 2004.

[18] M. Najim. *Modeling, estimation and optimal filtering in signal processing*. Wiley, 2010.

[19] X. Nguyen, M. J. Wainwright, and M. I. Jordan. Estimating divergence functionals and the likelihood ratio by convex risk minimization. *IEEE Transactions on Information Theory*, 56, n°11:5847–5861, 2010.

[20] F. Nielsen and R. Nock. A closed-form expression for the Sharma-Mittal entropy of exponential families. *Journal of Physics A: Mathematical and Theoretical*, 45, 2012.

[21] F. Nielsen and R. Nock. On Rényi and Tsallis entropies and divergences for exponential families. *Journal of Physics A: Mathematical and Theoretical*, 45, (3), 2012.

[22] F. Nielsen and R. Nock. On the chi square and higher-order chi distances for approximating f-divergences. *IEEE Signal Processing Letters*, 21, (1):10–13, 2014.

[23] M. Saleh, E. Grivel, and S.-M. Omar. Jeffrey's divergence between ARFIMA processes. *Digital Signal Processing*, 82:175–186, 2018.

[24] T. Soderstrom and P. Stoica. *System Identification*. Prentice Hall, 1989.

[25] M. Sugiyama, S. Liu, M. Christoffel du Plessis, M. Yamanaka, T. Suzuki, and T. Kanamori. Direct divergence approximation between probability distributions and its applications in machine learning. *Journal of Computing Science and Engineering*, 7, no. 2:99–111, 2013.

[26] R. F. Vigelis, L. H.F. de Andrade, and C. C. Cavalcante. Conditions for the existence of a generalization of Rényi divergence. *Physica A*, 558:124953, 2020.