

Comparing a Complex-Valued Sinusoidal Process with an Autoregressive Process Using Jeffrey's Divergence

Eric Grivel*, Mahdi Saleh †* and Samir-Mohamad Omar†

*Bordeaux university - INP Bordeaux ENSEIRB-MATMECA - IMS - UMR CNRS 5218, Talence, France

†Department of Physics and Electronics, Faculty of Sciences I, Lebanese University Hadath, Beirut, Lebanon

Abstract—This paper deals with the analysis of the Jeffrey's divergence (JD) between an autoregressive process (AR) and a sum of complex exponentials (SCE), whose magnitudes are Gaussian random values, which is then disturbed by an additive white noise. As interpreting the value of the JD may not be necessarily an easy task, we propose to give an expression of the JD and to analyze the influence of each process parameter on it. More particularly, we show that the ratios between the variance of the additive white noise and the variance of the AR-process driving process on the one hand, and the sum of the ratios between the SCE process power and the AR-process PSD at the normalized angular frequencies on the other hand, has a strong impact on the JD. The 2-norm of the AR-parameter has also an influence. Illustrations confirm the theoretical part.

Index Terms—Jeffrey's divergence, Kullback-Leibler divergence, AR process, Sum of complex exponentials, model comparison.

I. INTRODUCTION

Statistical signal processing based on *a priori* modeling plays a key role in speech processing, radar processing, flood forecasting, etc. However, choosing models is not necessarily an easy task and model comparison can be of interest. As Wold decomposition [1] states that any covariance-stationary process can be decomposed into two mutually uncorrelated component processes, *i.e.* a deterministic part that can be modeled by a harmonic model and a purely non-deterministic part, a great deal of interest has been paid to autoregressive (AR) moving average (MA) processes but also to the sum of complex exponentials whose magnitudes are random and Gaussian distributed, which is then disturbed by an additive white Gaussian noise (SCE). Various signals are often modeled by these types of processes. In spectrum analysis, this leads to subspace methods such as MUSIC, ESPRIT and their various variants and to the high-resolution spectral analysis based on AR modeling [2]. In speech processing, the signal can be modeled by a sum of sinewaves or by an AR process. See [3] for instance. In mobile communication systems, this is the same type of choice that can be done for fading channel modeling [4] [5].

In this paper, we suggest comparing a p^{th} -order AR process with a SCE process. For this purpose, we propose to use the Jeffrey's divergence (JD) which is the symmetric version of the Kullback-Leibler (KL) divergence. Both divergences have been used in a wide range of applications, starting from biomedical applications [6], passing through radar clutter

analysis [7] and model motion comparison [8], to texture analysis [9] [10]. They consist in comparing probability density functions (pdf). When dealing with process comparison, the pdf of k successive samples are considered. The less dissimilar the pdf are, the smaller the divergences are. As a consequence, we could *a priori* guess that the JD between a SCE process, defined by the normalized angular frequencies of its complex exponentials, and an AR process which has sharper and sharper resonances in its spectrum at the same normalized angular frequencies becomes smaller and smaller. Nevertheless, in practice, the results that can be obtained do not necessarily confirm this statement. For this reason, we propose to derive the expression of the JD. Then, we analyze how it evolves when the number of variates k increases. In addition, the role played by the different parameters characterizing the processes is studied.

This paper is organized as follows: in sections II, we briefly recall the definition and the expression of the JD. Then, we present the correlation properties of an AR process and those of the SCE process. Then, the JD between both processes is addressed. In section III, theoretical results are illustrated by some examples.

In the following, I_k is the identity matrix of size k , J_k the "shift" matrix which has ones on the first sub-diagonal and zeros elsewhere, Tr the trace of a matrix, \det the determinant of a matrix, and $diag(x)$ the diagonal matrix whose main diagonal is defined by x . $e_{k,l}$ denotes the column vector of size k with zeros except at the l^{th} row where the element is equal to 1. The upperscripts T , $*$ and H denote the transpose, the conjugate and the hermitian. $x_{k_1:k_2} = (x_{k_1}, \dots, x_{k_2})$.

II. JEFFREY'S DIVERGENCE ANALYSIS

A. Definition of the Jeffrey's divergence

To evaluate the dissimilarities between the distributions of k successive values of two random processes, denoted as $p_1(x_{1:k})$ and $p_2(x_{1:k})$, the KL divergence can be used and satisfies [11]:

$$KL_k^{(1,2)} = \int_{x_{1:k}} p_1(x_{1:k}) \ln \left(\frac{p_1(x_{1:k})}{p_2(x_{1:k})} \right) dx_{1:k} \quad (1)$$

When the processes are both Gaussian and real with means $\mu_{1,k}$ and $\mu_{2,k}$ and covariance matrices $Q_{1,k}$ and $Q_{2,k}$, it can

be easily shown that the KL satisfies¹ [12]:

$$KL_k^{(1,2)} = \frac{1}{2} \left[\text{Tr}(Q_{2,k}^{-1} Q_{1,k}) - k - \ln \frac{\det Q_{1,k}}{\det Q_{2,k}} \right. \\ \left. + (\mu_{2,k} - \mu_{1,k})^T Q_{2,k}^{-1} (\mu_{2,k} - \mu_{1,k}) \right] \quad (2)$$

However, as the KL is not symmetric, the Jeffrey divergence is often used. It is defined as follows:

$$JD_k^{(1,2)} = \frac{1}{2} (KL_k^{(1,2)} + KL_k^{(2,1)}) \quad (3)$$

When dealing with zero-mean processes in (2), (3) becomes:

$$JD_k^{(1,2)} \propto -k + \frac{1}{2} \left[\text{Tr}(Q_{2,k}^{-1} Q_{1,k}) + \text{Tr}(Q_{1,k}^{-1} Q_{2,k}) \right] \quad (4)$$

B. Presentation of the processes under study

Two types of processes are considered:

1) *Autoregressive process*: Let us consider a p^{th} -order AR process. Its n^{th} sample, x_n , is defined as follows:

$$x_n = - \sum_{i=1}^p a_i^p x_{n-i} + u_n^p \quad (5)$$

where u_n^p is the zero-mean Gaussian white driving process with variance $\sigma_{u,p}^2$, $a_0^p = 1$ and $\{a_i^p\}_{i=0,\dots,p}$ are the AR parameters that can be stored in a row vector of size $p+1$, denoted as \underline{a}^p . In addition, as the AR process is assumed to be a strict p^{th} -order AR process, one has: $\{a_i^p\}_{i>p} = 0$.

Given (5), the AR process can be seen as the filtering of a white noise, whose transfer function $H(z) = 1/A^p(z)$ is only defined by its p poles². Its power spectrum density (PSD), evaluated at the normalized angular frequency θ , is denoted as $S_{AR,p}(\theta)$:

$$S_{AR,p}(\theta) = \frac{\sigma_{u,p}^2}{\left| \sum_{k=0}^p a_k^p e^{-jk\theta} \right|^2} = \frac{\sigma_{u,p}^2}{\left| A^p(z) \Big|_{z=e^{j\theta}} \right|^2} \quad (6)$$

When the AR poles are close to the unit circle in the z -plane, the PSD exhibits resonances which are located at the normalized angular frequencies around the pole arguments.

In the following, $Q_{AR,k}$ corresponds to the Toeplitz covariance matrix of the vector storing k consecutive samples of the AR process, namely $x_{n:n+k-1}$. Note that the k elements of its main diagonal correspond to the correlation function of the AR process for a lag equal to 0, i.e. $r_{AR,0}$. The τ^{th} subdiagonal is defined by the correlation function with lag τ , i.e. $r_{AR,\tau}$. There are various ways to compute the inverse of the correlation matrix of the AR process. Eigenvalue decomposition could be considered but there is no explicit form for the eigenvalues of an AR covariance matrix. It is true that there are some approaches that provide approximations of the eigenvalues when the size of the AR correlation matrix is large and when

¹In the complex case, T is replaced by H and $\frac{1}{2}$ disappears in (2). In the following, \propto is used instead of $=$.

² $H(z) = \frac{1}{A^p(z)} = \frac{1}{\prod_{i=1}^p (1-p_i z^{-1})}$ where $\{p_i\}_{i=1,\dots,p}$ denote the AR poles.

the order of the AR process is equal to 1 [13]. An analytical expression based on the AR parameters also exists [14]:

$$Q_{AR,k}^{-1} = \frac{1}{\sigma_{u,p}^2} (FF^H - GG^H) \quad (7)$$

$$\text{with: } F = I_k + \sum_{i=1}^p a_i^p J_k^i \quad \text{and} \quad G = \sum_{i=1}^p a_i^p J_k^{k-i} \quad (8)$$

However, in this paper, we rather suggest using the LDL factorization of $Q_{AR,k}$ which involves the product between a lower unit triangular matrix L and a diagonal matrix D . In this latter case, the inverse of $Q_{AR,k}$ satisfies:

$$Q_{AR,k}^{-1} = (L^H)^{-1} D^{-1} L^{-1} \quad (9)$$

$$\text{where } D^{-1} = \text{diag}\left(\frac{1}{\sigma_{u,k-1}^2} \dots \frac{1}{\sigma_{u,0}^2}\right) \quad (10)$$

$$(L^H)^{-1} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ a_1^{k-1} & 1 & 0 & \dots & 0 \\ a_2^{k-1} & a_1^{k-2} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{k-1}^{k-1} & a_{k-2}^{k-2} & \dots & \dots & 1 \end{bmatrix} \quad (11)$$

However, as the AR process is assumed to be a strict p^{th} -order AR process, the matrices D^{-1} and $(L^H)^{-1}$ become:

$$D^{-1} = \text{diag}\left(\underbrace{\frac{1}{\sigma_{u,p}^2} \dots \frac{1}{\sigma_{u,p}^2}}_{k-p}, \frac{1}{\sigma_{u,p-1}^2}, \dots, \frac{1}{\sigma_{u,0}^2}\right) \quad (12)$$

and

$$(L^H)^{-1} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ a_1^p & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ a_2^p & a_1^p & 1 & 0 & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ a_p^p & a_{p-1}^p & \dots & \dots & 1 & \dots & \dots & \vdots \\ 0 & a_p^p & \dots & \dots & a_1^{p-1} & \dots & \dots & \vdots \\ \vdots & \vdots & \ddots & a_{p-1}^p & \dots & \dots & 1 & 0 \\ 0 & 0 & \dots & a_p^p & a_{p-1}^p & \dots & a_1^p & 1 \end{bmatrix} \quad (13)$$

To end up with this brief presentation about AR processes, let us recall that the AR parameters as well as the variance of the driving process can be estimated by using the Yule-Walker equations, the correlation method, adaptive filters such as the LMS and the APA, Burg's method, etc. [1].

2) *Sum of complex exponentials (SCE) disturbed by an additive noise*: The process is Gaussian zero-mean and defined by its $k \times k$ covariance matrix $Q_{SCE,k}$:

$$Q_{SCE,k} = S_k P S_k^H + \sigma^2 I_k \quad (14)$$

where $S_k = [S_k^1 \dots S_k^M]$ is a matrix which satisfies:

$$S_k = \begin{bmatrix} 1 & \dots & 1 \\ e^{j\theta_1} & \dots & e^{j\theta_M} \\ \vdots & \ddots & \vdots \\ e^{j(k-1)\theta_1} & \dots & e^{j(k-1)\theta_M} \end{bmatrix} \quad (15)$$

with k the number of successive samples that are considered and M the number of complex exponentials with normalized angular frequencies $\{\theta_m\}_{m=1,\dots,M}$ different from each other in the interval $[-\pi, \pi]$. In addition, P is a diagonal matrix of size $M \times M$ whose main diagonal is defined by the variances $\{\gamma_m\}_{m=1,\dots,M}$ of the random zero-mean magnitudes of the complex exponentials. At this stage, let us recall the following properties that the vectors $\{S_k^m\}_{m=1,\dots,M}$ satisfy:

$$\begin{cases} \frac{1}{k} (S_k^m)^H S_k^m = 1 \\ \frac{1}{k} (S_k^m)^H S_k^n = \frac{1}{k} \sum_{t=0}^{k-1} e^{j(\theta_m - \theta_n)t} \\ = \frac{\sin(\frac{k(\theta_m - \theta_n)}{2})}{k \sin(\frac{(\theta_m - \theta_n)}{2})} e^{j(\frac{(k-1)(\theta_m - \theta_n)}{2})} \end{cases} \quad (16)$$

As $\lim_{k \rightarrow +\infty} \left| \frac{\sin(\frac{k(\theta_m - \theta_n)}{2})}{k \sin(\frac{(\theta_m - \theta_n)}{2})} \right| = 0$ for any set of normalized angular frequencies $\theta_n \neq \theta_m$, (16) becomes:

$$\begin{cases} \frac{1}{k} (S_k^m)^H S_k^m = 1 \\ \lim_{k \rightarrow +\infty} \frac{1}{k} (S_k^m)^H S_k^n = 0 \end{cases} \quad (17)$$

Given (14) and using the matrix inversion lemma³, one has:

$$Q_{SCE,k}^{-1} = \frac{1}{\sigma^2} \left(I_k - S_k (\sigma^2 P^{-1} + S_k^H S_k)^{-1} S_k^H \right) \quad (18)$$

In (18), let us apply again the inversion matrix lemma on the matrix $(S_k^H S_k + \sigma^2 P^{-1})^{-1}$. This leads to:

$$\begin{aligned} (S_k^H S_k + \sigma^2 P^{-1})^{-1} &= (S_k^H S_k)^{-1} \\ -(S_k^H S_k)^{-1} (\sigma^2 P^{-1} + (S_k^H S_k)^{-1})^{-1} (S_k^H S_k)^{-1} & \end{aligned} \quad (19)$$

Given (19), $(S_k (\sigma^2 P^{-1} + S_k^H S_k)^{-1} S_k^H)$ is rewritten as:

$$T_k = V_k + U_k \quad (20)$$

$$\text{with: } \begin{cases} V_k = S_k (S_k^H S_k)^{-1} S_k^H \\ U_k = -S_k (S_k^H S_k)^{-1} \times \\ (\sigma^2 P^{-1} + (S_k^H S_k)^{-1})^{-1} (S_k^H S_k)^{-1} S_k^H \end{cases} \quad (21)$$

When k increases, due to (16), (21) becomes:

$$\begin{cases} \lim_{k \rightarrow +\infty} V_k = \frac{1}{k} S_k S_k^H \\ \lim_{k \rightarrow +\infty} U_k = -\frac{\sigma^2}{k^2} S_k P^{-1} S_k^H \end{cases} \quad (22)$$

Therefore, given (22), (18) becomes:

$$\lim_{k \rightarrow +\infty} Q_{SCE,k}^{-1} = \frac{1}{\sigma^2} \left(I_k - \frac{1}{k} S_k S_k^H + \frac{\sigma^2}{k^2} S_k P^{-1} S_k^H \right) \quad (23)$$

Finally, let us point out the fact this process is characterized by a power spectrum which has a discrete part due to the complex exponential functions and a continuous part due to the additive white noise.

In the next sections, using the definitions and properties of the JD and the processes under study, we propose to express $Tr(Q_{SCE,k}^{-1} Q_{AR,k})$ and $Tr(Q_{AR,k}^{-1} Q_{SCE,k})$.

C. Expression of the trace $Tr(Q_{SCE,k}^{-1} Q_{AR,k})$

Given (14) and (23), after developing, $Tr(Q_{SCE,k}^{-1} Q_{AR,k})$ is the sum of three terms denoted as

³Given the matrices A , U , C and V where A and C are assumed to be invertible, one has: $(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$

Ω , Ψ and Υ :

$$\Omega = \frac{1}{\sigma^2} Tr(Q_{AR,k}) = \frac{k}{\sigma^2} r_{AR,0} \quad (24)$$

$$\Psi = -\frac{1}{k\sigma^2} Tr(S_k S_k^H Q_{AR,k}) = -\frac{1}{k\sigma^2} Tr(S_k^H Q_{AR,k} S_k) \quad (25)$$

and

$$\Upsilon = \frac{1}{k^2} Tr(S_k P^{-1} S_k^H Q_{AR,k}) = \frac{1}{k^2} Tr(P^{-1} S_k^H Q_{AR,k} S_k) \quad (26)$$

In (25) and (26), $S_k^H Q_{AR,k} S_k$ has to be computed. Let us study what the value of $S_k^H Q_{AR,k} S_k^m$ is, when $m = 1, \dots, M$:

$$S_k^H Q_{AR,k} S_k^m = \quad (27)$$

$$\left[\sum_{\tau=0}^{k-1} r_{AR,\tau} e^{-j\tau\theta_m} \dots \sum_{\tau=1-k}^0 r_{AR,\tau} e^{-j(\tau+k-1)\theta_m} \right] \times [1 \dots e^{j(k-1)\theta_m}]^T$$

By rearranging the terms, (27) can be rewritten as follows:

$$S_k^H Q_{AR,k} S_k^m = \sum_{i=0}^{k-1} \sum_{\tau=-i}^i r_{AR,\tau} e^{-j\tau\theta_m} \quad (28)$$

Then, let us introduce a truncated correlogram $C_{AR,i}$ based on a window of size $2i + 1$ and computed for the normalized angular frequency θ_m :

$$C_{AR,i}(\theta_m) = \sum_{\tau=-i}^i r_{AR,\tau} e^{-j\tau\theta_m} \quad (29)$$

By combining (25) with (28) and (29) on the one hand and (26) with (28) and (29) on the other hand, one has:

$$\begin{cases} \Psi = -\frac{1}{k\sigma^2} \sum_{m=1}^M \sum_{i=0}^{k-1} C_{AR,i}(\theta_m) \\ \Upsilon = \frac{1}{k^2} \sum_{m=1}^M \frac{1}{\gamma_m} \sum_{i=0}^{k-1} C_{AR,i}(\theta_m) \end{cases} \quad (30)$$

Given (24) and (30), $Tr(Q_{SCE,k}^{-1} Q_{AR,k})$ can be expressed as follows when k tends to infinity:

$$\begin{aligned} Tr(Q_{SCE,k}^{-1} Q_{AR,k}) &\approx \frac{k}{\sigma^2} r_{AR,0} + \sum_{m=1}^M \frac{1}{k} \left(\frac{1}{k\gamma_m} - \frac{1}{\sigma^2} \right) \sum_{i=0}^{k-1} C_{AR,i}(\theta_m) \end{aligned} \quad (31)$$

D. Expression of the trace $Tr(Q_{AR,k}^{-1} Q_{SCE,k})$

Given (9) and (14), $Tr(Q_{AR,k}^{-1} Q_{SCE,k})$ can be expressed as the sum of two traces, denoted as Δ and Θ :

$$\Delta = Tr(\sigma^2 (L^H)^{-1} D^{-1} L^{-1}) = \sigma^2 Tr(L^{-1} (L^H)^{-1} D^{-1}) \quad (32)$$

Using (13), this leads to:

$$\begin{aligned} \Delta &= (k-p) \frac{\sigma^2}{\sigma_{u,p}^2} \sum_{i=0}^p |a_i^p|^2 + \sum_{j=0}^{p-1} \frac{\sigma^2}{\sigma_{u,j}^2} \sum_{i=0}^j |a_i^j|^2 \\ &= (k-p) \frac{\sigma^2}{\sigma_{u,p}^2} \|a^p\|^2 + \sum_{j=0}^{p-1} \frac{\sigma^2}{\sigma_{u,j}^2} \|a^j\|^2 \end{aligned} \quad (33)$$

with $\|\underline{a}^j\|^2$ the 2-norm of the j^{th} -order AR-parameter vector.

Remark: When considering a 1st-order AR process, $Q_{AR,k}^{-1}$ has a simple expression:

$$Q_{AR,k}^{-1} = \frac{1}{\sigma_{u,1}^2} (a_1^1 J_k + a_1^{1*} J_k^T + (1 + |a_1^1|^2) I_k - |a_1^1|^2 (e_{k,1} e_{k,1}^T + e_{k,k} e_{k,k}^T)). \quad (34)$$

Therefore, $Tr(\sigma^2 Q_{AR,k}^{-1}) = \frac{\sigma^2}{\sigma_{u,1}^2} (k + (k-2)|a_1^1|^2)$. This is also confirmed by (33).

Then, let us give some details about the second trace. Taking into account the expression of the PSD of the AR process for any order $i = 0, \dots, p$ evaluated at normalized angular frequencies θ_m with $m = 1, \dots, M$, one has:

$$\Theta = Tr(S_k^H (L^H)^{-1} D^{-1} L^{-1} S_k P) = \sum_{m=1}^M \gamma_m \left(\frac{k-p}{S_{AR,p}(\theta_m)} + \sum_{i=0}^{p-1} \frac{1}{S_{AR,i}(\theta_m)} \right) \quad (35)$$

It should be noted that in the above equation (35), we can organize the terms as follows:

$$\Theta = \sum_{m=1}^M \gamma_m \left(\frac{k-p-1}{S_{AR,p}(\theta_m)} + \frac{1}{S_{capon}(\theta_m)} \right) \quad (36)$$

Indeed, the pseudo-spectrum S_{capon} that can be deduced using Capon's method [2] can be expressed by:

$$S_{capon}(\theta_m) = \frac{\left[\sum_{i=0}^p \frac{1}{S_{AR,i}(\theta_m)} \right]^{-1}}{[1 \dots e^{-jp\theta_m}] Q_{AR,p+1}^{-1} [1 \dots e^{-jp\theta_m}]^H} \quad (37)$$

E. Expression of the JD

Combining (31), (33) and (35), the JD can be approximated when k tends to infinity as follows:

$$JD_k \propto -k + \frac{1}{2} \left(\frac{k}{\sigma^2} r_{AR,0} + \sum_{m=1}^M \sum_{i=0}^{k-1} \frac{1}{k} \left(\frac{1}{k\gamma_m} - \frac{1}{\sigma^2} \right) C_{AR,i}(\theta_m) \right) \quad (38)$$

$$+ \left((k-p) \frac{\sigma^2}{\sigma_{u,p}^2} \|\underline{a}^p\|^2 + \sum_{j=0}^{p-1} \frac{\sigma^2}{\sigma_{u,j}^2} \|\underline{a}^j\|^2 \right) + \sum_{m=1}^M \gamma_m \left(\frac{k-p}{S_{AR,p}(\theta_m)} + \sum_{i=0}^{p-1} \frac{1}{S_{AR,i}(\theta_m)} \right)$$

F. Analysis of the increment of the JD

Given (38), let us now deduce the asymptotic increment, i.e. the increment of the JD when k tends to infinity:

$$\Delta JD = \lim_{k \rightarrow +\infty} JD_k - JD_{k-1} \quad (39)$$

When k tends to infinity, one has:

$$\Delta JD \propto -1 + \frac{1}{2} \left(\underbrace{\frac{r_{AR,0}}{\sigma^2}}_{1^{\text{st}} \text{ term}} + \underbrace{\frac{\sigma^2}{\sigma_{u,p}^2} \|\underline{a}^p\|^2}_{2^{\text{nd}} \text{ term}} + \underbrace{\sum_{m=1}^M \frac{\gamma_m}{S_{AR,p}(\theta_m)}}_{3^{\text{rd}} \text{ term}} \right) \quad (40)$$

According to (40), we can notice that the asymptotic JD increment depends on three main terms:

- 1) $\frac{r_{AR,0}}{\sigma^2}$ is the power of the AR process filtered by an all-pass filter whose transfer function is $\frac{1}{\sigma}$,
- 2) $\frac{\sigma^2}{\sigma_{u,p}^2} \|\underline{a}^p\|^2$ is the power of the additive white noise with variance σ^2 that has been filtered by the finite-impulse response (FIR) "AR inverse filter" whose transfer function is $\frac{A^p(z)}{\sigma_{u,p}}$,
- 3) the third term depends on the sum of the ratios between the SCE process power and the AR-process PSD at the normalized angular frequencies $\{\theta_m\}_{m=1, \dots, M}$. Nevertheless, by using (6), each term $\frac{\gamma_m}{S_{AR,p}(\theta_m)}$ can be also seen as the power of the complex exponential at θ_m which has been filtered by the AR inverse filter. Therefore, this third term corresponds to the power of the deterministic part of the SCE process filtered by the AR inverse filter.

The above results lead to some comments:

- 1) when $M = p = 0$, (40) reduces to:

$$\Delta JD \propto -1 + \frac{1}{2} \left(\frac{\sigma_{u,1}^2}{\sigma^2} + \frac{\sigma^2}{\sigma_{u,1}^2} \right) \quad (41)$$

It corresponds to the JD increment between white noises with variance $\sigma_{u,1}^2$ and σ^2 .

- 2) The smaller the three terms in (40) are, the smaller the asymptotic JD increment is. However, they do not necessarily have the same behavior with respect to the process parameters. Thus, if $p = M$ and if the arguments of the AR poles correspond to the normalized angular frequencies $\{\theta_m\}_{m=1, \dots, M}$, resonances (resp. rejections) may appear in the transfer function $H(z)$ (resp. in $A^p(z)$). The closer to the unit circle in the z -plane the poles $\{p_i\}_{i=1, \dots, p}$ are, the sharper the resonances of $H(z)$ (resp. the more important the rejections of $A^p(z)$) are. In this case, the 1st term in (40) can grow because of $r_{AR,0} \propto \int S_{AR,p}(\theta) d\theta$ and $\|\underline{a}^p\|^2$. Meanwhile, the 2nd and 3rd term in (40) tend to be smaller and smaller. We will illustrate this phenomenon in the next section.

III. ILLUSTRATIONS AND COMMENTS

In this section, for the sake of simplicity, we suggest comparing a 1st-order AR process with a SCE process. The influences of the different parameters can be also illustrated.

Note that $r_{AR,0} = \frac{\sigma_{u,1}^2}{1 - |a_1^1|^2}$.

A. Influence of the AR-parameter argument

The simulation protocol is the following: $|a_1^1| = 0.95$, $\sigma_{u,1}^2 = 1$, $M = 1$, $\gamma_1 = 10$, $\theta_1 = \pi/2$ and $\sigma^2 = 1$. The argument of a_1^1 is in the interval $] -\pi, 0[$. Based on Fig. 1, which shows the asymptotic JD increment as a function of the argument of a_1^1 , the asymptotic JD increment reaches its minimum when the argument of a_1^1 is equal to $-\pi/2$. In this case, the resonance of AR process appears at the same angular frequency as the one of the SCE process.

B. Influence of the AR-parameter modulus

The simulation protocol is the following: a_1^1 is in the interval $] -1, 0[$. It first increases from -0.95 to -0.70 with a step

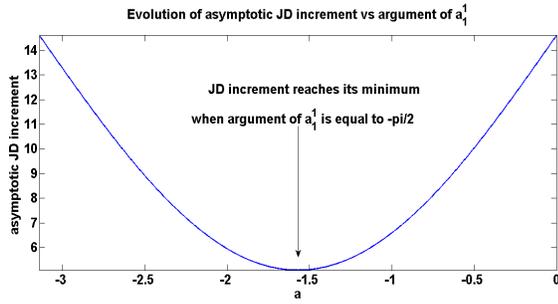


Fig. 1: Evolution of the asymptotic JD increment as a function of the argument of a_1^1

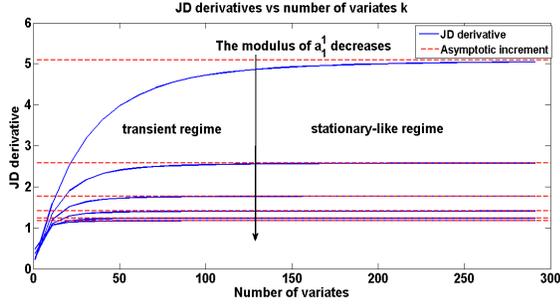


Fig. 2: JD derivative vs number of variates, second simulation with six cases where the modulus of a_1^1 varies

equal to 0.05, which means that the PSD exhibits a resonance at the normalized angular frequency $\theta = 0$ which is less and less sharp. $\sigma_{u,1}^2 = 1$, $M = 1$, $\gamma_1 = 10$, $\theta_1 = 0$ and $\sigma^2 = 1$. In Fig. 2, the derivative of the JD obtained from the expressions of $Q_{AR,k}$ and $Q_{SCE,k}$ tends to the asymptotic increment whatever the case under study. Two regimes appear: a transient one and then a stationary-like one. This confirms our theoretical study.

Fig. 3 shows the asymptotic increment as a function of a_1^1 where the 6 cases addressed above are pointed out. The asymptotic increment reaches its minimum when the modulus of a_1^1 is not necessarily close to the unit-circle but around 0.7. It illustrates our comments of section F.

C. Influence of the additive-noise variance

The simulation protocol is similar to that of section III.A, except that $a_1^1 = -0.95$ while σ^2 varies between 0.5 and 10. According to Fig. 4, (40) tends to infinity if σ^2 tends to zero. This illustrates the main difference between the spectrum

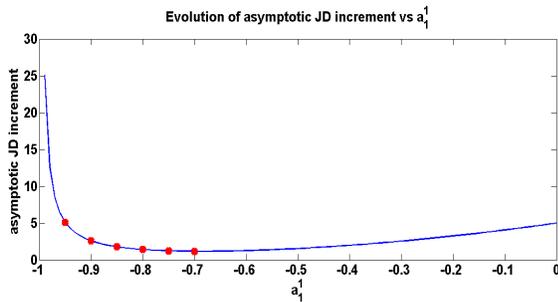


Fig. 3: Evolution of the asymptotic JD increment as a function of a_1^1 where the dotted points corresponds to the cases addressed in Fig.2

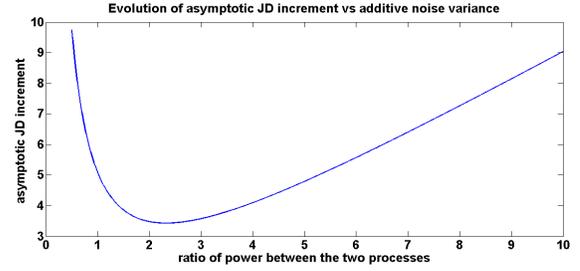


Fig. 4: Evolution of the asymptotic JD increment as a function of σ^2 properties of both processes: one tends to be discrete whereas the other is continuous.

IV. CONCLUSIONS AND PERSPECTIVES

In this paper, our purpose was to study the JD when comparing an AR process with a sum of complex exponentials and an additive white noise. Selecting a too small value for the number of variates k can be questionable because the JD has a transient period before converging to a stationary regime characterized by an asymptotic increment. The expression of this latter depends on the parameters of both processes. Our work is a new step to interpret the value of the JD because the expression that is provided is a good basis to see how the parameters can influence the value of the JD. In practical case, as the process parameters can be estimated from the data, the asymptotic increment can be estimated directly from the data.

REFERENCES

- [1] M. Najim, *Modeling, estimation and optimal filtering in signal processing*. Wiley, 2010.
- [2] C. Therrien, *Discrete random signals and statistical signal processing*. Prentice Hall, 1992.
- [3] W. Bobillet, R. Diversi, E. Grivel, R. Guidorzi, M. Najim, and U. Soverini, "Speech enhancement combining optimal smoothing and errors-in-variables identification of noisy AR processes," *IEEE Trans. on Signal Processing*, vol. 55, pp. 5564–5578, December 2007.
- [4] M. F. Pop and N. C. Beaulieu, "Limitations of sum-of-sinusoids fading channel simulator," *IEEE Trans. on Communications*, vol. 49, pp. 699–708, 2001.
- [5] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. on Communications*, vol. 4, pp. 1650–1662, 2005.
- [6] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis, "Touchless monitoring of breathing function," *IEEE EMBS*, pp. 1196–1199, 2004.
- [7] C. Magnant, E. Grivel, A. Giremus, L. Ratton, and B. Joseph, "Classifying autoregressive models using dissimilarity measures: A comparative study," *EUSIPCO*, pp. 998–1002, 2015.
- [8] C. Magnant, A. Giremus, and E. Grivel, "Jeffreys divergence between state models: Application to target tracking using multiple models," *EUSIPCO*, pp. 1–5, 2013.
- [9] L. Bombrun, N.-E. Lasmar, Y. Berthoumieu, and G. Verdoolaege, "Multivariate texture retrieval using the SIRV representation and the geodesic distance," *IEEE ICASSP*, pp. 865–868, 2011.
- [10] A. Schutz, L. Bombrun, Y. Berthoumieu, and M. Najim, "Centroid-based texture classification using the generalized gamma distribution," *EUSIPCO*, pp. 1–5, 2013.
- [11] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.
- [12] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.
- [13] R. J. Strooker, "Approximations of the eigenvalues of the covariance matrix of a first-order autoregressive process," *Journal of Econometrics*, pp. 269–279–81, 1983.
- [14] B. Cernuschi-Frias, "A derivation of the gohberg-semencul relation [signal analysis]," *IEEE Trans. on Signal Processing*, vol. 39, Issue: 1, pp. 190–192, 1991.