STATISTICAL ANALYSIS OF A LINEAR RELATION BETWEEN NOISY REPETITIVE SIGNALS

Gérard FAUCON[1], Régine LE BOUQUIN JEANNES[1], Jean-Pierre CONZE[2]

[1]Laboratoire Traitement du Signal et de l'Image, INSERM EMI 9934
Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

[2]Institut de Recherche Mathématique de Rennes
Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

Gerard.Faucon@univ-rennes1.fr

ABSTRACT

This paper deals with the coherence function in order to study relations between channels, in the context of auditory evoked potentials, and more specifically with the statistical analysis of a coherence estimator applied to signals having periodicity. In practice, a stimulus is periodically sent and a sufficient number of individual responses are recorded to increase the signal-to-noise ratio. In this way, a coherence estimator can be defined. Under some hypotheses, we get a simplified version of this quantity and we study its distribution and its mean. Simulations illustrate theoretical results. Then, we estimate the mean value of the coherence computed on a given number of responses. We derive the bias versus the theoretical coherence and the number of individual responses.

1. INTRODUCTION

In biomedical engineering, the understanding of certain cerebral structures goes through the analysis of responses, called evoked potentials, to auditory, visual or somesthetic stimuli. In the present work, we are concerned by audition, the aim being a better knowledge of the auditory cortex through the study of auditory evoked potentials (AEPs) recorded either on the surface of the scalp or using intra-cerebral electrodes (in the case of stereoelectroencephalo-graphy (SEEG) where patients are under a surgical investigation). We are interested in characterizing the recorded signals according to the localization of the recording and to the type of stimulus. Now, undesirable signals, such as electroencephalogram (EEG) or other physiological signals, can be superimposed to the useful signal induced by the external stimulus. In this way, this useful signal is embedded in noise and it becomes essential to emit a great number of identical stimuli and add the ensemble of synchronized individual responses to get an averaged evoked potential whose signal-to-noise ratio is actually increased compared to that of the individual responses. This paper is aimed at analyzing a linear relation between noisy signals using individual responses. This relation is recalled in section 2 and section 3 is devoted to the presentation of the asymptotic estimators. Then, a statistical analysis on the real estimator is conducted in section 4 before giving some concluding remarks.

2. METHOD AND MATERIAL

If the analysis of averaged evoked potentials provides the physiologist with some information on the function of particular structures, the study of relationships between different structures can be viewed as a valuable help. Parametric and non-parametric tools allow to determine the connection degree between different explored channels and to detect some activity changes.

In this paper, our attention is focused on the coherence function which is a frequency-dependent function we recall next. Let $\{x_i\}$, $i = 1,2,...,M$, be M temporal signals recorded on M channels. If these signals are jointly stationary, the ordinary coherence between two signals x_i and x_j, $i \neq j$, measures their linear relation degree and is defined as [1]:

$$\rho_{ij}(f) = \frac{\gamma_{x_i x_j}(f)}{\sqrt{\gamma_{x_i x_i}(f)\gamma_{x_j x_j}(f)}} \qquad (1)$$

where $\gamma_{x_i x_i}(f)$, $\gamma_{x_j x_j}(f)$ and $\gamma_{x_i x_j}(f)$ are the power spectral densities of x_i, x_j and the cross-spectral density between x_i and x_j, at the frequency f, respectively.

In the standard case of random signals (without any repetitive character), the practical computation of the coherence is performed on a temporal length in relation with the signals non-stationarity and each spectral density is obtained by averaging frequential informations on a given number of adjacent or overlapped segments.

In the analysis of AEPs, the same stimulus is emitted L times and the evoked responses are assimilated to noisy repetitive signals. On each channel i, we receive L individual responses to this stimulus, $x_{i,\ell}(n)$, $\ell = 1,...,L$, where $x_{i,\ell}(n)$

represents the time series sampled at the sampling rate f_e. Each response is composed of *N* samples. In the following, we only consider the coherence between two signals $x_1(n)$ and $x_2(n)$, and the estimated coherence, noted $\hat{\rho}(f)$, is derived from the *L* responses by:

$$\hat{\rho}(f)=\frac{\sum_{\ell=1}^{L} X_{1,\ell}(f)\,X_{2,\ell}^{*}(f)}{\left(\sum_{\ell=1}^{L} \left|X_{1,\ell}(f)\right|^2 \sum_{\ell=1}^{L} \left|X_{2,\ell}(f)\right|^2\right)^{1/2}} \qquad (2)$$

where $X_{i,\ell}(f)$ is the *N*-point Fourier transform of the ℓ th response on channel *i*, and the symbol * denotes the conjugate operator. More generally, the relationships may vary on the length of the responses, which forces to estimate the coherence at different times on the responses, and the Fourier transforms are computed on a part of the response. In other contexts such as EEG, no stimulus is emitted and we can compute in the same way the coherence from two observations using (2) where *L* becomes the number of blocks used in the estimation.

3. ASYMPTOTIC ESTIMATORS

In the following, we consider two cases : *i)* no stimulus is sent, *ii)* the observations are responses to a stimulus (AEP context).

Let $x_{i,\ell}(n)$, $i=1,2$, be the observation on channel *i* defined as $x_{i,\ell}(n)=s_\ell(n)+d_{i,\ell}(n)$, where $s_\ell(n)$ represents the useful signal in the ℓ th response or the ℓ th block, which is assumed identical on both channels and $d_{i,\ell}(n)$ the disturbing noise. The signals $s_\ell(n)$, $d_{1,\ell}(n)$ and $d_{2,\ell}(n)$ are decorrelated. In the particular case of AEPs, the useful part $s_\ell(n)$ of the ℓ th response to a stimulus is always the same and noted $s(n)$. The study of the coherence estimated according to equation 2 is quite difficult but it can be simplified when the number of blocks increases: the correlation between signal and noise as well as the correlation between noises tend in the mean to zero. In the following, the asymptotic estimators in cases *i)* and *ii)* are considered.

3.1 Non repetitive signals

If $s_\ell(n)$, $d_{1,\ell}(n)$, $d_{2,\ell}(n)$ are stationary gaussian noises with zero mean value, the asymptotic coherence is defined at each frequency as

$$\rho_{nr}(f)=\frac{\gamma_{ss}(f)}{\gamma_{ss}(f)+\gamma_{dd}(f)} \qquad (3)$$

considering that the noise power spectral densities are identical on both channels ($\gamma_{dd}(f)=\gamma_{d_1d_1}(f)=\gamma_{d_2d_2}(f)$). Given $a=\gamma_{ss}(f)/\gamma_{dd}(f)$ (a is the signal-to-noise ratio at a given frequency f but, in the following, the variable f is omitted for clarity reasons),

$$\rho_{nr}(f)=\frac{a}{a+1}. \qquad (4)$$

3.2 Repetitive signals

If we consider the same hypotheses as in section 3.1 and take the repetitive character of the useful signal into account, $\hat{\rho}(f)$ is asymptotically close to

$$\frac{L\left|S(f)\right|^2}{\left[L\left|S(f)\right|^2+\sum_{\ell=1}^{L}\left|D_{1,\ell}(f)\right|^2\right]^{1/2}\left[L\left|S(f)\right|^2+\sum_{\ell=1}^{L}\left|D_{2,\ell}(f)\right|^2\right]^{1/2}} \qquad (5)$$

where $S(f)$ represents the spectrum of the useful signal and $D_{i,\ell}(f)$ the spectrum of the noise present on channel *i*. If noises are of same power on both channels, the quantities $(1/L).\sum_{\ell=1}^{L}\left|D_{i,\ell}(f)\right|^2$, $i=1,2$, are close to $\gamma_{dd}(f)$ for a large *L*, and we obtain the quantity

$$\hat{\rho}_r(f)=\frac{\left|S(f)\right|^2}{\left|S(f)\right|^2+\gamma_{dd}(f)}. \qquad (6)$$

This coherence is a random variable, due to the random character of $S(f)$.

Now, we are interested in the theoretical distribution of $\hat{\rho}_r(f)$ when the useful signal is composed of decorrelated samples. The term $S(f)$ may be written

$$S(f)=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} s(n)e^{2i\pi nf}, \qquad (7)$$

where we assume that $s(n), n=0,\ldots,N-1$, is a series of *N* independent random variables of normal law $N\left(0,\sigma_s^2\right)$. The couple $\left(Re(S(f)), Im(S(f))\right)$ follows a gaussian law whose covariance matrix is $\sigma_s^2\begin{pmatrix}\alpha & \gamma\\ \gamma & \beta\end{pmatrix}$

with
$$\begin{cases}\alpha=\dfrac{1}{N}\sum_{n=0}^{N-1}\left(\cos 2\pi nf\right)^2\\ \beta=\dfrac{1}{N}\sum_{n=0}^{N-1}\left(\sin 2\pi nf\right)^2=1-\alpha\\ \gamma=\dfrac{1}{N}\sum_{n=0}^{N-1}\cos 2\pi nf\cdot\sin 2\pi nf.\end{cases} \qquad (8)$$

For frequencies $f=p/N$, $p\in\{0,1,\ldots,N-1\}$, corresponding to spectral analysis by Discrete Fourier Transform, the term γ is zero. Thus, the couple $\left(Re(S(f)), Im(S(f))\right)$ is a couple of independent gaussian variables. The law of $\left|S(f)\right|^2$ is that of a random variable $\alpha U^2+(1-\alpha)V^2$ where *U* and *V* are independent variables of normal law $N\left(0,\sigma_s^2\right)$. Two cases have to be distinguished:

i) if $f=0$ or $f=f_e/2$, $\alpha=1$,

ii) if $f\neq 0$ and $f\neq f_e/2$, $\alpha=1/2$.

In this way, if $f\neq 0$ and $f\neq f_e/2$, $|S(f)|^2$ is of the form $(U^2+V^2)/2$, with U and V defined previously. It is a random variable of exponential law and one can easily show that $\hat{\rho}_r(f)$ (given by equation 6) follows a law whose density is

$$\frac{1}{a(1-x)^2}\exp\left(-\frac{x}{a(1-x)}\right)1_{[0,1[}(x), \qquad (9)$$

where $1_{[0,1[}$ is the unit function on the interval $[0,1[$.

In case $f=0$ or $f=f_e/2$, $\alpha=1$ and $|S(f)|^2$ is the square of a gaussian variable and one can show that the density of $\hat{\rho}_r(f)$ may be written

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{ax(1-x)^3}\right)^{1/2}\exp\left(-\frac{x}{2a(1-x)}\right)1_{]0,1[}(x). \quad (10)$$

So, given the distribution of $\hat{\rho}_r(f)$, we can deduce by numerical computation its mean value for a given frequency.

3.3 Comparison

Figure 1 exhibits on the one hand the theoretical value of the coherence for random non repetitive signals (Fig. 1.a), and on the other hand that same value in the case of repetitive signals for two frequencies ($f=0$ (Fig. 1.b), $f\neq 0$ and $f\neq f_e/2$ (Fig. 1.c)), versus the signal-to-noise ratio *a*.

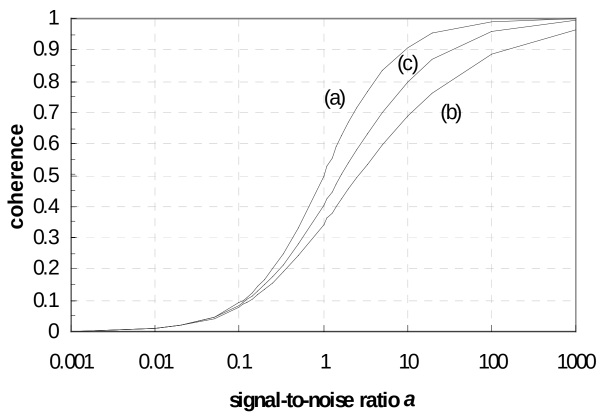

FIG. 1. Theoretical coherence
(a) non repetitive signals, (b) and (c) repetitive signals
((b) $f=0$, (c) $f\neq 0$ and $f\neq f_e/2$)

For the same signal-to-noise ratio, the coherence computed on repetitive signals is lower than that computed on non repetitive signals. In this second configuration, the coherence is not frequency-dependent while, in the first case, the coherence is lower for the two extreme frequencies $f=0$ and $f=f_e/2$.

4. STATISTICAL ANALYSIS

Until now, only studies on random non repetitive signals can be found in the literature [2]. To judge the performance of our estimator on signals having periodic components, a statistical analysis is conducted. First of all, our aim was to compare simulation and theoretical results related to the expression 6 of $\hat{\rho}_r(f)$, and compare them to values obtained using equation 2. Signals *s*, $d_{1,\ell}$, $d_{2,\ell}$ are white gaussian noises and the short time Fourier transform is 256-point long. The theoretical study indicates that we have to distinguish frequencies $f=0$ and $f=f_e/2$ from the others. First simulations described in [3] show that the histogram of the coherence $\hat{\rho}_r(f)$ is comparable to the theoretical distribution of the coherence given by equations 9 and 10. Figures 2 and 3 represent respectively for $f=0$ (or $f=f_e/2$) and for $f\neq 0, f\neq f_e/2$ the histogram of the coherence $\hat{\rho}(f)$ as well as the theoretical distribution of the coherence $\hat{\rho}_r(f)$, for $L=100$ and $a=1$.

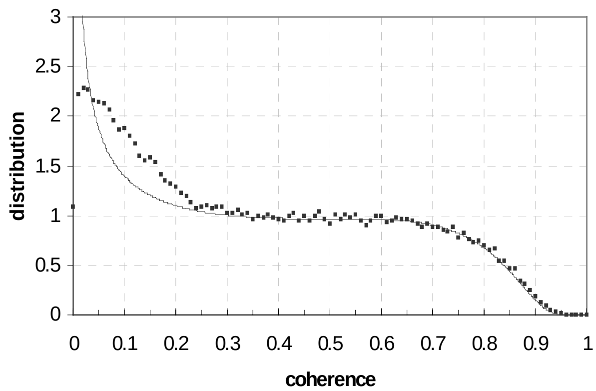

FIG. 2. Coherence distribution, $a=1$, $f=0$
histogram of the coherence $\hat{\rho}(f)$, $L=100$ (dotted line)
theoretical distribution of the coherence $\hat{\rho}_r(f)$
according to (10) (solid line)

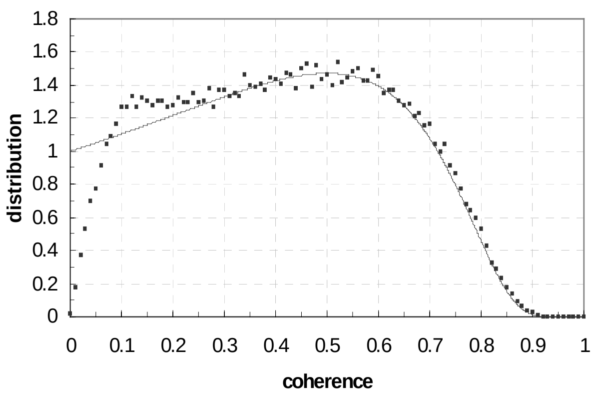

FIG. 3. Coherence distribution, $a=1$, $f\neq 0$ and $f\neq f_e/2$
histogram of the coherence $\hat{\rho}(f)$, $L=100$ (dotted line)
theoretical distribution of the coherence $\hat{\rho}_r(f)$
according to (9) (solid line)

To complete this study, we evaluate the bias of the estimator according to the number of responses L (Figures 4 and 5). It is obvious that the bias is all the more reduced as the number of responses is great. Nevertheless, in some applications, the value of L is limited and the knowledge of the bias may be introduced in a bias correcting algorithm to approach the true value if necessary.

Now, if we compare the bias obtained for $f = 0$ and $f \neq 0$ (and $f \neq f_e/2$), we observe that the curves are comparable whatever the number of responses is.

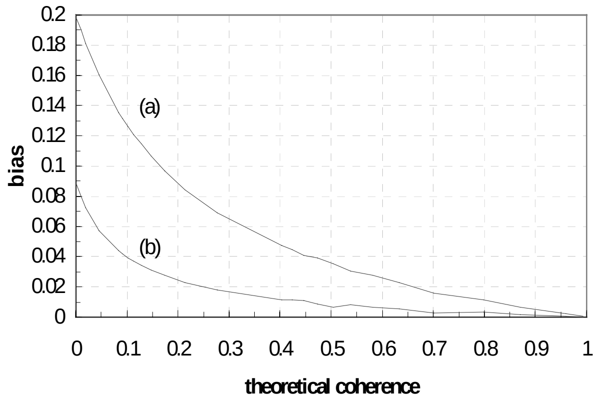

FIG. 4. Bias, $f \neq 0$ and $f \neq f_e/2$
(a) $L = 20$, (b) $L = 100$

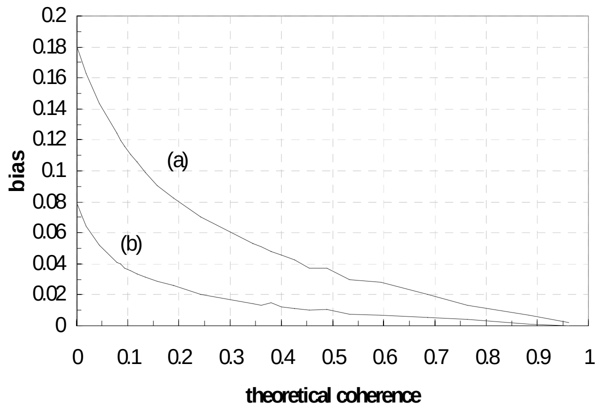

FIG. 5. Bias, $f = 0$, (a) $L = 20$, (b) $L = 100$

The analysis of the bias reported in [2] on non repetitive signals shows that the bias is inversely proportional to the number of blocks used in the coherence estimation. The same conclusion holds with repetitive signals. Now, we consider the situation where L equals 20, $f \neq 0$ and $f \neq f_e/2$, to compare the histograms. Figure 6 exhibits the theoretical coherence according to (9) as well as the histogram of $\hat{\rho}(f)$ for $L = 20$. Compared to Figure 3, it is obvious that the estimated coherence deviates from the theoretical value. We can note that the distribution of $\hat{\rho}(f)$ decreases to zero for low values of coherence and there is a shift towards higher values which makes the bias increase.

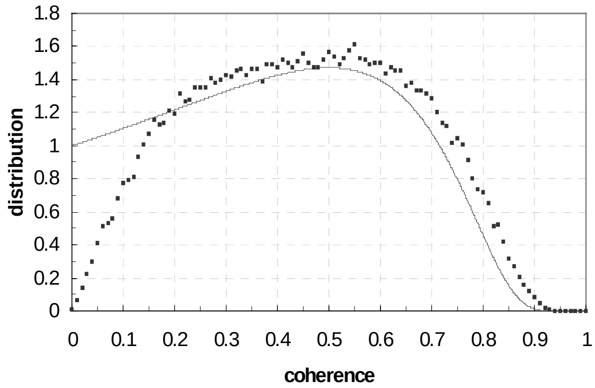

FIG. 6. Coherence distribution, $a = 1$, $f \neq 0$ and $f \neq f_e/2$
histogram of the coherence $\hat{\rho}(f)$, $L = 20$ (dotted line)
theoretical distribution of the coherence $\hat{\rho}_r(f)$
according to (9) (solid line)

5. CONCLUSION

In this paper, we discussed the problem of the coherence estimation when computed from signals whose useful component is repetitive and composed of decorrelated samples, and we showed the influence of this periodicity. The results must be taken into consideration in the interpretation of relations between signals recorded on different channels such as auditory evoked potentials. This study has to be extended to the case where the useful repetitive signal is constituted of correlated samples.

References

[1] J.S. Bendat, A.G. Piersol, *"Random Data: Analysis and Measurement Procedures"*, Wiley-Interscience, 1971.

[2] G.C. Carter, C.H. Knapp, A.H. Nutall, "*Estimation of the Magnitude-Squared Coherence via Overlapped Fast Fourier Transform Processing"*, IEEE Trans. on Audio and Electroacoustics, vol. AU-21, n°4, pp. 337-344, Aug. 1973.

[3] R. Le Bouquin Jeannès, G. Faucon, J.P. Conze, *"Etude Théorique d'un Estimateur de Cohérence entre Signaux Périodiques Bruités"*, GRETSI, Toulouse, 2001.