COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL
ET SES APPLICATIONS

NICE du 26 au 30 AVRIL 1977

# ESTIMATION DE LA PUISSANCE D'UN PROCESSUS ALEATOIRE GAUSSIEN DE SPECTRE LORENTZIEN PAR LA METHODE DU MAXIMUM DE VRAISEMBLANCE

ERNOULT Max

OFFICE NATIONAL D'ETUDES ET DE RECHERCHES AEROSPATIALES (O.N.E.R.A)
29, avenue de la Division Leclerc - 92320 CHATILLON SOUS BAGNEUX

## RESUME

Dans le cas particulier d'un processus aléatoire $X(t)$ gaussien de spectre lorentzien (fonction de corrélation de la forme $p\,a_1^{|\tau|}$ avec $0 < a_1 < 1$), on étudie les propriétés de l'estimateur du maximum de vraisemblance $\hat{p}$ de la puissance $p$ de $X(t)$ lorsque l'observation a une durée finie $T$ suivant l'information dont on dispose sur le paramètre $a_1$.

Dans le cas où l'on dispose d'une approximation $b_1$, de $a_1$, on compare les propriétés de $\hat{p}$ à celles de l'estimateur usuel $\left(\tilde{p} = \frac{1}{T}\int_0^T X^2(t)\,dt\right)$.

Dans le cas contraire où l'on ne dispose d'aucune information sur $a_1$, les propriétés de $\hat{p}$ sont équivalentes à celles de $\tilde{p}$ ; les hypothèses faites sur la forme du spectre n'apportent alors pas d'amélioration par rapport à la méthode usuelle.

## SUMMARY

*ESTIMATION OF THE POWER OF A GAUSSIAN RANDOM PROCESS OF LORENTZIAN SPECTRUM BY THE MAXIMUM LIKELIHOOD METHOD*

In the special case of a Gaussian random process $X(t)$ of Lorentzian spectrum (the autocorrelation function is $p\,a_1^{|\tau|}$ with $0 < a_1 < 1$), we study the properties of the maximum likelihood estimator $\hat{p}$ of the power $p$ of $X(t)$ for a finite time sample $(T)$ according to the a priori information on the parameter $a_1$.

At first, when we have an approximation $b_1$ of $a_1$, we compare the properties of $\hat{p}$ to those of the usual estimator $\left(\tilde{p} = \frac{1}{T}\int_0^T X^2(t)\,dt\right)$.

Secondly, when we have no information on $a_1$, the properties of $\hat{p}$ are equivalent to those of $\tilde{p}$ : the supplementary hypothesis on the spectrum does not improve the usual method.

# ESTIMATION DE LA PUISSANCE D'UN PROCESSUS ALEATOIRE GAUSSIEN DE SPECTRE LORENTZIEN PAR LA METHODE DU MAXIMUM DE VRAISEMBLANCE

## I - INTRODUCTION

Dans de nombreuses études expérimentales, il est nécessaire d'estimer la puissance $P$ d'un processus aléatoire stationnaire $X(t)$, centré et gaussien en ne disposant que d'un échantillon sur un intervalle de temps fini $[0,T]$.

Lorsque l'on ne dispose d'aucune information sur le processus, l'estimateur utilisé habituellement est :

$$\tilde{P} = \frac{1}{T}\int_0^T X^2(t)\,dt \qquad (1)$$

Si la fonction de corrélation de $X(t)$ est égale à :

$$E\{X(t).X(t-\tau)\} = P\,C(\tau) \qquad (2)$$

avec $C(0) = 1$

$\tilde{P}$ est un estimateur sans biais dont la variance normalisée, définie par :

$$\tilde{V} = \frac{1}{P^2} E\{(\tilde{P} - E\{\tilde{P}\})^2\} \qquad (3)$$

est égale à :

$$\tilde{V} = \frac{2}{T}\int_{-T}^{+T} C^2(\tau)\left(1 - \frac{|\tau|}{T}\right)d\tau \qquad (4)$$

Lorsque $T \to 0$, $\tilde{V} \sim 2$ ;
la variance normalisée de $\tilde{P}$ est donc médiocre quand $T$ est petit.

Si l'on connaît à priori la fonction de corrélation normalisée $C(\tau)$, c'est-à-dire toute l'information possible sauf la puissance $P$, on peut estimer quel que soit $T$, la puissance par la méthode du maximum de vraisemblance.

Discrétisons le processus et posons :

$$X^T = (X_1, \ldots, X_n)$$

avec $X_i = X(i\Delta t)$ et $n = \frac{T}{\Delta t}$

$X$ est une variable aléatoire à $n$ dimensions qui suit une loi normale $\mathcal{N}(0, P\Gamma)$ avec
$P$ puissance du bruit
$\Gamma = (\Gamma_{ij})$ matrice des corrélations

$$(\Gamma_{ij} = C(|i-j|\Delta t) \quad (5) \quad )$$

L'estimateur $\hat{P}_n$ du maximum de vraisemblance s'obtient en minimisant l'expression (6) qui est égale, à une constante près, à moins deux fois le logarithme de la densité de $X$.

$$n \log P - \log(\det(\Gamma^{-1})) + \frac{1}{P} X^T \Gamma^{-1} X \qquad (6)$$

En annulant la dérivée par rapport à $P$, on obtient le résultat classique

$$\hat{P}_n = \frac{1}{n} X^T \Gamma^{-1} X \qquad (7)$$

Ce résultat peut s'interpréter en cherchant les transformations linéaires $(H)$ :

$$Y = HX \qquad (8)$$

qui "blanchissent" $X$, c'est-à-dire qui imposent à $X$ de suivre une loi normale $\mathcal{N}(0, PI)$ ( $I$ matrice identité).

De la condition $E\{YY^T\} = PI$ on déduit l'équation matricielle définissant $H$.

$$H^T H = \Gamma^{-1}$$

et l'on a :

$$\hat{P}_n = \frac{1}{n} X^T \Gamma^{-1} X = \frac{1}{n} Y^T Y \qquad (9)$$

L'estimateur $\hat{P}_n$ est sans biais ( $E\{\hat{P}_n\} = P$ ) et la variance normalisée $\hat{V}_n$ définie par :

$$\hat{V}_n = \frac{1}{P^2} E\{(\hat{P}_n - E\{\hat{P}_n\})^2\}$$

est égale à :

$$\hat{V}_n = \frac{2}{n} \qquad (10)$$

Pour une durée $T$ quelconque, si l'on fait tendre $n$ vers l'infini, alors $\hat{P}_n$ converge vers $P$ en moyenne quadratique et l'estimateur $\hat{P}$ devient parfait. Dans la mesure où il n'est pas certain qu'on puisse physiquement extraire de l'information pour des échantillons très voisins, le cas limite $(n = \infty)$ n'a pas de sens physique. Par contre l'intérêt possible de cette méthode vient du comportement de $\hat{V}_n$ en $\frac{2}{n}$ qui est rapidement meilleur que la valeur constante de $\tilde{V}$ (4) pour $T$ fixé. Des études expérimentales [1] ont montré que cet estimateur n'est pas robuste, c'est-à-dire qu'une petite erreur sur la fonction de corrélation normalisée peut entraîner une erreur importante sur l'estimation de la puissance. Plus précisément, Picinbono [2] a montré que cette erreur provenait essentiellement d'une erreur sur le comportement à l'infini de la densité spectrale normalisée donc d'une erreur sur la forme de la fonction de corrélation.

On étudie donc, dans cet article, et pour un cas particulièrement simple, les propriétés de l'estimateur du maximum de vraisemblance lorsque la forme de la fonction de corrélation est connue (ici choisie égale à $P a_1^{|\tau|}$ avec $0 < a_1 < 1$ suivant les informations dont on dispose sur $a_1$. Dans le deuxième paragraphe, le biais et la variance sont calculés pour tout $n$ quand on connaît une approximation $b_1$ de $a_1$ et le troisième paragraphe contient l'étude, par simulation, du cas où l'on n'a aucune information sur $a_1$.

## II - ERREUR SUR LE PARAMETRE $a_1$.

On suppose donc que

$$C(\tau) = a_1^{|\tau|} = e^{-\frac{|\tau|}{\tau_a}} \qquad (11)$$

avec $\tau_a = \frac{-1}{\log a_1}$ "durée" de corrélation

et que l'on connaît une réalisation de $X(t)$ sur [0,1] ce qui ne restreint pas la généralité de l'étude car, par changement d'unité de temps, on peut toujours se ramener à cet intervalle. La valeur de $a_1$ n'est pas connue mais on dispose d'une approximation $b_1 \left(\tau_b = \frac{-1}{\log b_1}\right)$ de $a_1$ et l'on blanchit l'observation avec cette valeur $b_1$. On étudie, dans ce cas particulier les propriétés de $\hat{P}_n$ en fonction de $a_1$ et $b_1$ d'abord quand $n$ est fixé puis leurs comportements asymptotiques.

### II.1 - $n$ fixé

On pose $b = b_1^{\frac{1}{n}}$ et $a = a_1^{\frac{1}{n}}$ ; en particularisant (5) et en tenant compte de (11) où l'on substitue à $a_1$ son approximation $b_1$, on a :

$$\Gamma = \begin{bmatrix} 1 & b & b^2 & \cdots & b^{n-1} \\ b & 1 & \ddots & & \vdots \\ b^2 & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & b \\ b^{n-1} & \cdots & & b & 1 \end{bmatrix} \qquad (12)$$

et on trouve que

$$\Gamma^{-1} = \frac{1}{1-b^2} \begin{bmatrix} 1 & -b & 0 & & & \\ -b & 1+b^2 & -b & & 0 & \\ 0 & -b & 1+b^2 & \ddots & & \\ & & \ddots & \ddots & -b & 0 \\ & 0 & & -b & 1+b^2 & -b \\ & & & 0 & -b & 1 \end{bmatrix} \qquad (13)$$

Par récurrence, on montre que :

$$\det(\Gamma^{-1}) = \frac{1}{(1-b^2)^{n-1}} \qquad (14)$$

L'estimateur $\hat{P}_n$ étant toujours défini par $\hat{P}_n = \frac{1}{n} X^T \Gamma^{-1} X$ il vient :

$$\hat{P}_n = \frac{1}{n} \frac{1}{1-b^2} \left[ \sum_{i=1}^{n} X_i^2 + b^2 \sum_{i=2}^{n-1} X_i^2 - 2b \sum_{i=1}^{n-1} X_i X_{i+1} \right] \qquad (15)$$

la relation (8) devient :

$$Y_1 = X_1 \qquad (16)$$

$$Y_j = \frac{1}{\sqrt{1-b^2}} \left( X_j - b X_{j-1} \right) \qquad j = 2, n$$

avec

$$\hat{P}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i^2$$

Dans le cas où $a$ est égale à $b$ , les $Y_i$ sont effectivement des variables aléatoires gaussiennes et indépendantes.

II.1.1. - Calcul de $E\{\hat{P}_n\}$

L'espérance mathématique de $\hat{P}_n$ est donnéepar :

$$\frac{E\{\hat{P}_n\}}{P} = \frac{1}{nP} \sum_{i=1}^{n} E\{Y_i^2\}$$

on trouve d'après (16) :

$$\frac{E\{\hat{P}_n\}}{P} = 1 + \frac{n-1}{n} \cdot \frac{2b(b-a)}{1-b^2} \qquad (17)$$

En général $\hat{P}_n$ est un estimateur biaisé (sauf si $a = b$ ). Dans le cas où $b = 0$ , on a $\tilde{P}_n = \frac{1}{n} \sum_{i=1}^{n} X_i^2$ et c'est donc l'estimateur classique $\tilde{P}$ lorsque l'on discrétise le processus $X(t)$ et on vérifie bien qu'il est sans biais.

II.1.2. - Calcul de $\hat{V}_n$

En utilisant une propriété vérifiéepar des variables aléatoires $Y_i$ gaussiennes et centrées :

$$E\{Y_i^2 Y_j^2\} = E\{Y_i^2\} E\{Y_j^2\} + 2 \left( E\{Y_i Y_j\} \right)^2$$

on en déduit que :

$$\hat{V}_n = \frac{2}{n^2 P^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( E\{Y_i Y_j\} \right)^2$$

$$\hat{V}_n = \frac{2}{n^2 P^2} \left[ \sum_{i=1}^{n} \left( E\{Y_i^2\} \right)^2 + 2 \sum_{1 \le i < j \le n} \left( E\{Y_i Y_j\} \right)^2 \right]$$

D'après les relations (16) on a :

$$\left( E\{Y_1 Y_j\} \right)^2 = \frac{P^2}{1-b^2} (a-b)^2 a^{2(j-2)} \qquad j = 2, n$$

$$\left( E\{Y_i Y_j\} \right)^2 = \frac{P^2}{(1-b^2)^2} (b-a)^2 (1-ab)^2 a^{2(j-i-1)} \qquad 2 \le i < j \le n$$

et on en déduit :

$$\hat{V}_n = \frac{2}{n^2} \left[ 1 + (n-1) \left( \frac{1+b^2-2ab}{1-b^2} \right)^2 + \frac{2(b-a)^2}{1-b^2} \sum_{j=2}^{n} a^{2(j-2)} \right.$$

$$\left. + 2 \frac{(b-a)^2 (1-ab)^2}{(1-b^2)^2} \sum_{2 \le i < j \le n} a^{2(j-i-1)} \right]$$

En évaluant les progressions géométriques, on obtient :

$$\hat{V}_n = \frac{2}{n^2} \left[ 1 + (n-1) \left( \frac{1+b^2-2ab}{1-b^2} \right)^2 + \frac{2(b-a)^2}{1-b^2} \frac{1-a^{2(n-1)}}{1-a^2} \right.$$

$$\left. + 2 \frac{(b-a)^2 (1-ab)^2}{(1-b^2)^2} \left( \frac{n-1}{1-a^2} - \frac{1-a^{2(n-1)}}{(1-a^2)^2} \right) \right] \qquad (18)$$

Dans le cas où $b$ est égale à zéro, la variance normalisée $\tilde{V}_n$ de $\tilde{P}_n$ est donnée par :

$$\tilde{V}_n = \frac{2}{n^2} \left( n \frac{1+a^2}{1-a^2} - 2a^2 \frac{1-a^{2n}}{(1-a^2)^2} \right) \qquad (19)$$

II.1.3. Remarque

Pour un pas d'échantillonnage $\Delta t$ fixé, si l'on fait tendre le temps d'observation vers l'infini $n$ tend vers l'infini, $a$ et $b$ restent constants et l'estimateur du maximum de vraisemblance, qui est biaisé, devient dans ce cas moins intéressant que l'estimateur usuel.

II.2 - Cas asymptotique pour $T$ fixé.

On détermine maintenant, pour $T$ fixé, le comportement asymptotique des deux estimateurs quand $\Delta t$ tend vers zéro ( $n$ tend alors vers l'infini). En remplaçant $a$ et $b$ par $e^{\frac{1}{n}\log a_1}$ et $e^{\frac{1}{n}\log b_1}$ et par un développement limité au premier ordre de (17), (18) et (19), on obtient par un calcul simple :

pour $b_1 \neq 0$

$$E\{\hat{P}_n\} \sim P \frac{\log a_1}{\log b_1} = P \frac{\tau_b}{\tau_a} \qquad (17')$$

$$\hat{V}_n \sim \frac{2}{n} \left( \frac{\log a_1}{\log b_1} \right)^2 = \frac{2}{n} \left( \frac{\tau_b}{\tau_a} \right)^2 \qquad (18')$$

et pour $b_1 = 0$

$$\tilde{V}_n \sim - \frac{2}{\log a_1} \left( 1 + \frac{1}{2} \frac{1-a_1^2}{\log a_1} \right)$$

soit

$$\tilde{V}_n \sim 2\tau_a \left[ 1 - \frac{\tau_a}{2} \left( 1 - \exp\left(-\frac{2}{\tau_a}\right) \right) \right] \qquad (19')$$

La relation (19') peut s'obtenir directement en utilisant la relation (4).

A titre d'illustration, on a tracé, en utilisant les expressions (17), (18) et (19) pour $a_1 = 10^{-3}$ / $\tau_a = 0.144$ ), $\tilde{V}_n$ et, pour trois valeurs différentes de $\tau_b$ le biais au carré et l'erreur quadratique moyenne en fonction de $n$ sur la figure 1

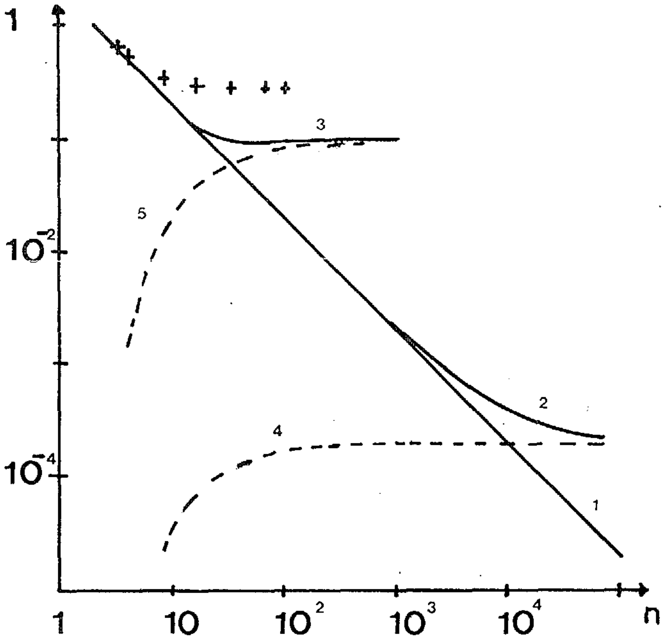


Figure 1 : propriétés des estimateurs (pour $\tau_a = 0.144$) en fonction de $n$ .

++++ $\tilde{V}_n$ variance normalisée de $\tilde{P}_n$

—— erreur quadratique moyenne normalisée de $\hat{P}_n$

pour

| | | |
|---|---|---|
| $\tau_b = \tau_a$ | | 1 |
| $\tau_b = 0,146$ | | 2 |
| $\tau_b = 0,188$ | | 3 |
| - - - biais au carré pour | $\tau_b = 0,146$ | 4 |
| | $\tau_b = 0,188$ | 5 |

Malgré la limitation physique mentionnée dans l'introduction sur la possibilité du passage à la limite, on peut dire que l'estimateur du maximum de vraisemblance permet de mesurer une quantité certaine ( $E\{\hat{p}\} = p\,\frac{\tau_b}{\tau_a}$ ) et biaisée. La connaissance de la forme du modèle nous permet donc théoriquement de mesurer sans erreur la quantité $\frac{p}{\tau_a}$ et, si l'on connaît une approximation $\tau_b$ de $\tau_a$, on en déduit alors une valeur de $p$ .

## II.3 - Comparaison des estimateurs $\hat{p}$ et $\tilde{p}$ .

Si l'on compare les deux estimateurs à l'aide du critère de l'erreur quadratique moyenne, l'estimateur $\hat{p}$ sera préférable à $\tilde{p}$ tant que le biais élevé au carré de $\hat{p}$ sera plus petit que la variance de $\tilde{p}$ soit :

$$\left(\frac{\tau_b}{\tau_a}-1\right)^2 \leq 2\,\tau_a\left[1-\frac{\tau_a}{2}\left(1-\exp\left(-\frac{2}{\tau_a}\right)\right)\right]$$

Si $\delta$ est l'erreur relative sur $\tau_a$ ( $\delta = \left|\frac{\tau_b-\tau_a}{\tau_a}\right|$ ), $\hat{p}$ est préférable à $\tilde{p}$ pour

$$\delta \leq \sqrt{2\tau_a}\left[1-\frac{\tau_a}{2}\left(1-\exp\left(-\frac{2}{\tau_a}\right)\right)\right]^{1/2} \qquad (20)$$

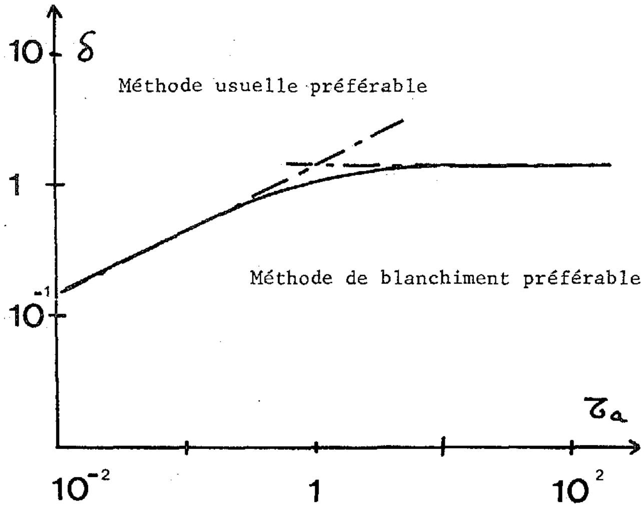


Figure 2 : Choix de l'estimateur en fonction de $\delta$ erreur relative de $\tau_a$ .

La courbe de la figure (2) permet donc de choisir l'estimateur le plus intéressant et suivant la valeur de $\tau_a$ on peut distinguer :

$$\tau_a >> 1 \quad , \quad \delta \leq \sqrt{2}$$

Pour un temps d'observation court (ici $T = 1$ ) devant la "durée" de la corrélation de $X(t)$ la méthode du maximum de vraisemblance (qui s'interprète comme une méthode de blanchiment) est plus intéressante que la méthode classique car elle ne nécessite pas une bonne précision relative sur $\tau_a$ .

$$\tau_a << 1 \; , \quad \delta < \sqrt{2\tau_a}$$

Pour un temps d'observation long devant la "durée" de corrélation, il faut une erreur relative petite, donc difficile à obtenir pour que la méthode de blanchiment soit préférable à la méthode classique, qui retrouve donc ici tout son intérêt.

## III - $a_1$ INCONNU

Si l'on ne dispose d'aucune information sur $a_1$, on peut encore appliquer la méthode du maximum de vraisemblance dans le cas discret en minimisant pour $0 \leq a < 1$ et $p > 0$ l'expression (6) qui s'écrit :

$$(n-1)\log(1-a^2) + n\log p + \frac{1}{p(1-a^2)}\left[\sum_{i=1}^{n} X_i^2 + a^2\sum_{i=2}^{n-1} X_i^2 - 2a\sum_{i=1}^{n-1} X_i X_{i+1}\right] \qquad (21)$$

En annulant la dérivée partielle de (21) par rapport à $p$ , on en déduit le minimum de (21) en fonction de $a$ et il faut trouver la valeur de $a$ ( $0 \leq a < 1$ ) qui minimise :

$$n \log\left[\sum_{i=1}^{n} X_i^2 + a^2\sum_{i=2}^{n-1} X_i^2 - 2a\sum_{i=1}^{n-1} X_i X_{i+1}\right] - \log(1-a^2) \quad (22)$$

Par une étude de fonction élémentaire on montre que si $\sum_{i=1}^{n-1} X_i X_{i+1} \leq 0$ le minimum est atteint pour

$$\hat{a} = 0 \qquad \hat{P}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \qquad (23)$$

si $\sum_{i=1}^{n-1} X_i X_{i+1} > 0$ il existe un et un seul minimum pour $a$ compris strictement entre 0 et 1 et que, pour ce minimum, on a les relations :

$$\hat{P}_n = \frac{1}{n(1-\hat{a}^2)}\left[\sum_{i=1}^{n} X_i^2 + \hat{a}^2\sum_{i=2}^{n-1} X_i^2 - 2\hat{a}\sum_{i=1}^{n-1} X_i X_{i+1}\right] \qquad (23')$$

$$\hat{a} = \frac{\sum_{i=1}^{n-1} X_i X_{i+1}}{\sum_{i=2}^{n-1} X_i^2 + \hat{P}_n}$$

On montre que $\hat{a}$ est donné par une équation du troisième degré et il n'est pas possible d'avoir une expression analytique de $\hat{P}_n$ qui permette d'en déduire ses propriétés. Aussi a-t-on simulé, pour différentes valeurs de $a$ et $n$, 1 000 réalisations de $X$ ce qui nous a permis de calculer des estimations de $\frac{E\{\hat{P}_n\}}{P}$ et $\hat{V}_n$.

Sur le tableau des résultats et en tenant compte des erreurs dues à la statistique, on remarque que :

- $\hat{P}_n$ est un estimateur sans biais
- la variance de $\hat{P}_n$ est sensiblement égale à celle de $\tilde{P}_n$.

Les deux estimateurs ont donc des propriétés statistiques équivalentes et on doit préférer l'estimateur usuel à cause de la simplicité des calculs.

## IV - CONCLUSION.

Lorsque l'on dispose d'une approximation de $\tau_a$ (erreur relative $\delta$) on choisit l'estimateur du maximum de vraisemblance si :

- pour $\tau_a >> T$ , $\delta < \sqrt{2}$
- pour $\tau_a << T$ , $\delta < \sqrt{\frac{2\tau_a}{T}}$ .

Si $\tau_a$ est complètement inconnu, on n'a pas réussi à améliorer, contrairement à l'intuition, les propriétés de l'estimateur usuel malgré les hypothèses supplémentaires faites sur la fonction de corrélation.

| n \ a | $10^{-2}$ | 0,316 | 0,749 | 0.930 |
|---|---|---|---|---|
| 10 | 0,0217 (1)<br>1,00 (2)<br>0,203 (3)<br>0,200 (4) | 0,0868<br>1,02<br>0,248<br>0,239 | 0,347<br>1,04<br>0,618<br>0,597 | 1,39<br>0,986<br>1,22<br>1,31 |
| 20 | 0,0108<br>1,01<br>0,0963<br>0,100 | 0,0434<br>0,993<br>0,122<br>0,120 | 0,174<br>1,01<br>0,337<br>0,327 | 0,694<br>1,00<br>0,864<br>0,937 |
| 40 | 0,00542<br>1,00<br>0,0490<br>0,0500 | 0,0217<br>1,00<br>0,0623<br>0,0608 | 0,0868<br>0,996<br>0,168<br>0,171 | 0,347<br>1,02<br>0,600<br>0,576 |

Tableau des résultats.

(1) $\tau_a$
(2) estimation de $\frac{E\{\hat{P}_n\}}{P}$
(3) estimation de $\hat{V}_n$
(4) valeur théorique de $\tilde{V}_n$.

[1] ERNOULT, HAY, MACCHI
Etude de la stabilité d'un estimateur en métrologie des bruits d'avions.
Annales des Télécommunications, T30, n° 7-8, p. 298-303 (1975)

[2] PICINBONO
Interprétation spectrale du CAG parfait
(Communication personnelle).