



Conditions d'Application et Validation des Algorithmes d'Identification Fondés sur les Spectres d'Ordre Supérieur

J. Le Roux and D. Rossille

I3S, University of Nice CNRS, bat. 4, 250 rue A. Einstein, Sophia Antipolis, 06560 Valbonne France
leroux@mimosa.unice.fr

RÉSUMÉ

Cette communication propose une méthode pour engendrer un signal aléatoire qui peut être considéré comme l'échantillonné d'un signal stationnaire au troisième ordre à bande limitée, vérifiant les hypothèses de Brillinger et Rosenblatt. Nous insistons sur la différence des effets de l'échantillonnage appliqué aux signaux stationnaires et non stationnaires. Nous discutons les conséquences de cette différence sur l'utilisation des algorithmes d'identification fondés sur l'analyse des spectres d'ordre supérieur.

ABSTRACT

This communication proposes a method for generating a random signal that can be interpreted as the sampling of a third order stationary continuous-time signal, corresponding to the hypotheses of Brillinger and Rosenblatt. We insist on the difference between stationary and non-stationary signals when they are sampled. We discuss the consequences of this difference when using identification algorithms based on the analysis of higher order spectra.

INTRODUCTION

De nombreux algorithmes d'identification fondés sur les statistiques d'ordre supérieur [2][3][9][10] sont appliqués à des signaux aléatoires obtenus en appliquant un filtre numérique à une séquence d'échantillons indépendants. Une telle séquence ne peut pas être interprétée comme étant le résultat de l'échantillonnage d'un signal stationnaire au troisième ordre: Le spectre d'ordre trois (TOS) de la séquence d'échantillons indépendants a un module constant, il ne s'annule pas dans les triangles ABC et EFG de la figure 1. Ceci contredit les résultats de Brillinger et Rosenblatt [1] caractérisant les effets à l'ordre trois de l'échantillonnage sur les signaux stationnaires. Par conséquent, les algorithmes d'identification de ce type doivent être utilisés avec précaution. Ce point sera discuté dans le second paragraphe. Pour analyser les effets d'un algorithme avant de traiter des signaux réels stationnaires, il peut être utile de disposer d'un signal simulé qui puisse être considéré comme le résultat de l'échantillonnage d'un signal analogique stationnaire au troisième ordre. C'est ce que nous proposons dans le premier paragraphe.

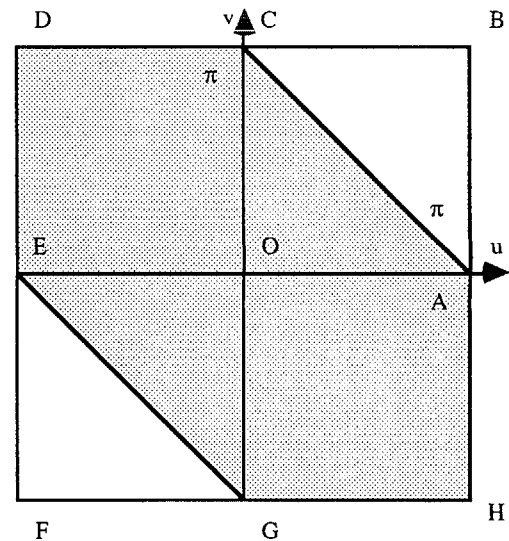


Fig. 1. Support du TOS d'un signal stationnaire à bande limitée (ACDEGH) ; après échantillonnage le support est la périodisation carrée de cet hexagone : le TOS est nul en ABC et EFG.

1. SIMULATION ET ÉCHANTILLONNAGE DE SIGNAUX STATIONNAIRES À BANDE LIMITÉE

Tout d'abord nous engendrons une séquence régulière d'échantillons indépendants à la cadence Δt *a priori* différente de la cadence d'échantillonnage, centrés et d'amplitude aléatoire non gaussienne $s(p\Delta t)$. Ensuite nous construisons un signal à bande limitée en filtrant cette séquence par $\sin \pi t / \pi t$ pour l'échantillonner à la cadence 1:

$$z(t) = \sum_{p=-\infty}^{\infty} s(p\Delta t) \frac{\sin \pi(t-p\Delta t)}{\pi(t-p\Delta t)} \quad (1)$$

1.1. Stationnarité de $z(t)$

$z(t)$ est stationnaire au troisième ordre seulement si $\Delta t < 2/3$. En effet,

$$E\{z(t)z(t')z(t'')\} = E\left\{ \sum_{p=-\infty}^{\infty} s(p\Delta t) \frac{\sin \pi(t-p\Delta t)}{\pi(t-p\Delta t)} \right. \\ \left. \times \sum_{p'=-\infty}^{\infty} s(p'\Delta t) \frac{\sin \pi(t'-p'\Delta t)}{\pi(t'-p'\Delta t)} \times \sum_{p''=-\infty}^{\infty} s(p''\Delta t) \frac{\sin \pi(t''-p''\Delta t)}{\pi(t''-p''\Delta t)} \right\} \quad (2)$$



$$E\{z(t)z(t')z(t'')\} = E\{s^3\} \quad (3)$$

$$\times \sum_{p=-\infty}^{\infty} \frac{\sin \pi(t-p\Delta t)}{\pi(t-p\Delta t)} \frac{\sin \pi(t'-p\Delta t)}{\pi(t'-p\Delta t)} \frac{\sin \pi(t''-p\Delta t)}{\pi(t''-p\Delta t)}$$

La transformée de Fourier de $E\{z(t)z(t')z(t'')\}$ est

$$R_a(u,v,w) = E\{s^3\} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \quad (4)$$

$$\times \frac{\sin \pi(t-p\Delta t)}{\pi(t-p\Delta t)} \frac{\sin \pi(t'-p\Delta t)}{\pi(t'-p\Delta t)} \frac{\sin \pi(t''-p\Delta t)}{\pi(t''-p\Delta t)}$$

$$\times \exp -j (ut+vt'+wt'')dt dt' dt''$$

En effectuant une translation sur t, t' et t'',

$$R_a(u,v,w) = E\{s^3\} H_a(u)H_a(v)H_a(w) \quad (5)$$

$$\times \sum_{p=-\infty}^{\infty} \exp j (u+v+w)p\Delta t$$

$$\text{où } H_a(u) = 1 \text{ si } |u| < \pi \text{ et } H_a(u) = 0 \text{ si } |u| > \pi. \quad (6)$$

Donc

$$R_a(u,v,w) = 0, \quad (7)$$

$$\text{sauf si } |u|, |v|, |w| < \pi \text{ et } (u+v+w)\Delta t = 2q\pi, \quad (8)$$

où q est entier. Si $\Delta t < 2/3$, $R_a(u,v,w) \neq 0$ seulement lorsque $q = 0$ soit dans $u+v+w = 0$ et à l'intérieur du cube $-\pi < u, v, w < \pi$.

Dans ce cas, $E\{z(t)z(t')z(t'')\}$ est la transformée de Fourier inverse d'une fonction qui est nulle en dehors du plan $u+v+w = 0$. C'est donc une fonction invariante par translation. Alors $E\{z(t+\tau)z(t'+\tau)z(t''+\tau)\}$ est indépendant de τ et $z(t)$ est stationnaire. Son TOS est la projection de $R_a(u,v,w)$ sur le plan $w = 0$:

$$B_a(u,v) = R_a(u,v,-u-v) = H_a(u)H_a(v)H_a(-u-v)E\{s^3\}, \quad (9)$$

non nul dans ACDEGH seulement.

Si $\Delta t > 2/3$, $R_a(u,v,w) \neq 0$ est possible pour $q \neq 0$, $R_a(u,v,w)$ n'est pas nul en dehors de $u+v+w = 0$ et par conséquent $z(t)$ n'est pas stationnaire.

1.2. Effet de l'échantillonnage

Nous vérifions que les résultats de Brillinger et Rosenblatt [1] sur l'échantillonnage des signaux stationnaires s'appliquent à $z(t)$. Soit $z(k)$ obtenu en échantillonnant $z(t)$ à la cadence unité. Sa covariane d'ordre trois est

$$E\{z(k)z(k')z(k'')\} = E\{s^3\} \quad (10)$$

$$\times \sum_{p=-\infty}^{\infty} \frac{\sin \pi(k-p\Delta t)}{\pi(k-p\Delta t)} \frac{\sin \pi(k'-p\Delta t)}{\pi(k'-p\Delta t)} \frac{\sin \pi(k''-p\Delta t)}{\pi(k''-p\Delta t)}$$

Elle a pour transformée de Fourier

$$R(u,v,w) = E\{s^3\} \sum_{k=-\infty}^{\infty} \sum_{k'=-\infty}^{\infty} \sum_{k''=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \quad (11)$$

$$\times \frac{\sin \pi(k-p\Delta t)}{\pi(k-p\Delta t)} \frac{\sin \pi(k'-p\Delta t)}{\pi(k'-p\Delta t)} \frac{\sin \pi(k''-p\Delta t)}{\pi(k''-p\Delta t)}$$

$$\times \exp -j (uk+vk'+wk'')$$

La transformée de Fourier du signal échantillonné $\sin \pi(k-p\Delta t)/\pi(k-p\Delta t)$ est la périodisation de $\exp(jup\Delta t)$ dans $-\pi \leq u < \pi$. Par conséquent, $R(u,v,w)$ est la périodisation en u, v et w de la fonction égale à

$$R_a(u,v,w) = E\{s^3\} \sum_{p=-\infty}^{\infty} \exp j (u+v+w)p\Delta t \quad (12)$$

dans le cube $-\pi \leq u, v, w < \pi$.

Si $\Delta t < 2/3$, $R_a(u,v,w)$ est non nul seulement dans $u+v+w=0$. Alors, $z(k)$ est un signal échantillonné stationnaire et son TOS, $B(u,v)$ s'obtient en projetant $R(u,v,w)$ dans le plan $w = 0$. C'est la périodisation de la fonction constante dans ACDEGH et nulle dans les triangles ABC et EFG :

$$B(u,v) = R_a(u,v,w) \text{ pour } -\pi < u, v < \pi. \quad (13)$$

1.3. Exemple

La figure 2 montre $\log|B(u,v)|$ lorsque $z(k)$ est calculé pour $\Delta t = 1/2$. $s(p\Delta t)$ vaut $(1-m)/m$ avec la probabilité $1/m$ et $-1/m$ avec la probabilité $(1-m)/m$ (dans l'exemple $m = 3$).

Remarque : si $B(u,v)$ est estimé à partir d'une seule séquence, il est factorisable par construction ; par conséquent il ne s'annule pas dans ABC. $B(u,v)$ décroît de manière significative seulement quand on calcule des moyennes sur plusieurs séquences (ici 100.000 séquences de 128 échantillons).

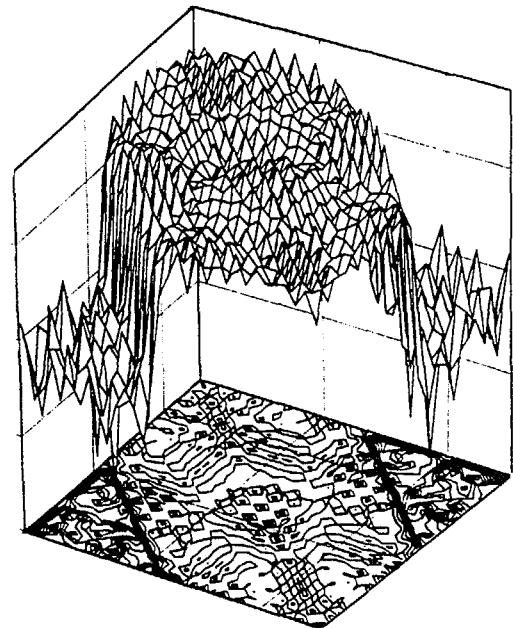


Fig. 2. module du spectre d'ordre trois de $z(k)$ (échelle 10dB)

Cette façon de simuler l'échantillonnage d'un signal analogique stationnaire au troisième ordre permet d'illustrer le résultat de Brillinger et Rosenblatt sur les effets de l'échantillonnage [1]. Il est important de remarquer que lorsque $\Delta t = 1$, $z(k)$ est aussi un signal échantillonné stationnaire bien que n'étant pas l'échantillonné d'un signal analogique stationnaire, mais d'un signal analogique cyclo-stationnaire dont la période des fluctuations statistiques est égale à la cadence d'échantillonnage.

1.4. Remarque sur la qualité des statistiques du troisième ordre des séquences pseudo-aléatoires

Notons que la qualité médiocre des TOS obtenus en simulation est due pour une part aux imperfections des générateurs de nombres pseudo-aléatoires. Ces imperfections s'observent dans la corrélation du troisième ordre d'une séquence d'échantillons indépendants $s(p)$ engendrés aussi bien par un IBM RS 6000 que par un DEC 5500.

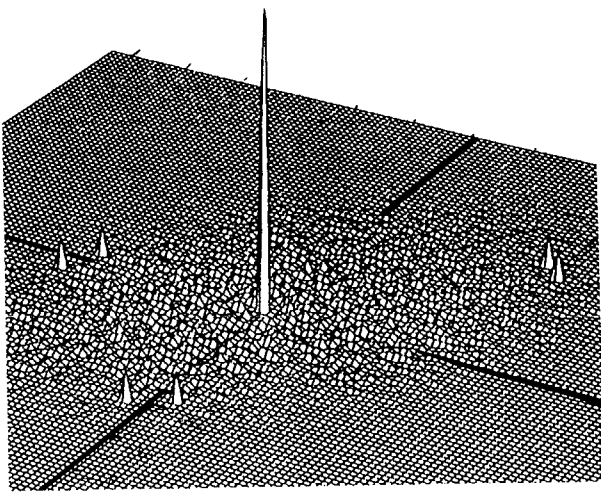


Fig. 3. Corrélation d'ordre 3 d'une séquence d'échantillons pseudo-aléatoires. On note un pic significatif en $(-31, -3)$ et aux cinq points qui en sont déduits par symétrie

Cette corrélation présente des pics significatifs en $(-31, -3)$ et dans les cinq autres points déduits de celui-là par symétrie (fig. 3). On peut sans doute déduire les caractéristiques d'un générateur de nombres pseudo-aléatoires en analysant ses statistiques d'ordre supérieur.

2. SUR LES CONDITIONS D'APPLICATION AUX SIGNAUX RÉELS DES ALGORITHMES D'IDENTIFICATION HOS

2.1. Stationnarité (test de Hinich [4])

Plusieurs algorithmes d'identification fondés sur les HOS ([2],[3],[9],[10]) supposent que ces HOS de signaux de sortie de systèmes échantillonnés soient factorisables en

$$B(u,v) = G(u)G(v)G(-u-v). \quad (14)$$

Cette supposition n'est pas compatible avec l'hypothèse de stationnarité du signal analogique sauf si sa bande est limitée au tiers de la fréquence d'échantillonnage [5]. Si on envisage d'appliquer une méthode d'identification fondée sur la factorisabilité du TOS, il est nécessaire de savoir si le signal analogique est stationnaire ou non, c'est à dire si le TOS est

nul ou non dans le triangle ABC. Cette étude correspond au test de Hinich de détection de transitoires [4]. Notons cependant que la notion de signal transitoire doit inclure celle de signal cyclo-stationnaire et que les signaux non stationnaires dont la bande passante est inférieure au tiers de la fréquence d'échantillonnage ne sont pas détectés par ce test [5].

2.2. Application des techniques d'identification aux signaux stationnaires

Lorsque le signal analogique est stationnaire, on peut adapter les techniques numériques d'identification HOS de différentes manières :

- On peut échantillonner à une cadence supérieure au triple de la bande passante du signal; alors les signaux analogiques et échantillonnés seront tous les deux factorisables [5] : on pourra leur appliquer les algorithmes développés dans le cas des HOS de signaux déterministes échantillonnés ou de séquences d'échantillons filtrés numériquement (c'est à dire la plupart des algorithmes traitant les données dans le domaine temporel [2][3][9][10]); Cependant, ceci peut entraîner des difficultés numériques car le TOS est théoriquement nul dans 66% du domaine des fréquences.

- Lorsque le signal est échantillonné à une fréquence double de la bande passante, $B(u,v)$ s'annule dans ABC. Le problème de l'identification revient alors à factoriser

$$B(u,v) = G(u)G(v)G(-u-v)H(u,v), \quad (15)$$

où $H(u,v)$ est la périodisation de la fonction égale à 1 dans ACDEGH et égale à zéro dans ABC et EFG. Dans le domaine des fréquences, les techniques sont celles qui ont été développées initialement pour le traitement des signaux analogiques [6][8]. Dans le domaine temporel, il s'agit de trouver $f(k)$ tel que

$$r(m,n) = \sum_{p,q} \left[\sum_k f(k)f(k+p)f(k+q) \right] h(m-p,n-q) \quad (16)$$

où $r(m,n)$ est la corrélation d'ordre trois et $h(m,n)$ est la transformée de Fourier inverse de $H(u,v)$, soit

$$\text{pour } m=n=0 : h(0,0) = 3/4 \quad (17)$$

$$\text{pour } m \neq 0 : h(m,0) = h(0,m) = h(m,m) = \frac{1-(-1)^m}{2\pi^2 m^2};$$

$$\text{pour } mn(m-n) \neq 0 : h(m,n) = \frac{1}{2\pi^2} \frac{1}{nm(m-n)} \times [n(-1)^n - m(-1)^m + (m-n)(-1)^{m-n}].$$

Par exemple une méthode d'identification itérative pourrait s'inspirer du schéma de la figure 4.

- Si le TOS est factorisable, il est possible de déduire sa valeur dans ABC et EFG de sa valeur dans ACDEGH, ce qui permet d'utiliser les algorithmes applicables aux signaux déterministes : La valeur de la phase de $B(u,v)$ dans ABC et EFG se déduit des valeurs dans ACDEGH par une extension de la formule proposée par Marron, Sanchez et Sullivan pour résoudre le problème du déroulement de la phase [7]. Si $B(u,v)$ est factorisable,



$$\text{phase } B(u,v) = \text{phase} \frac{B(u,\pi-u)B(v,\pi-u-v)}{B(u+v,\pi-u-v)} \quad (18)$$

Ainsi une valeur de phase $B(u,v)$ dans ABC se calcule à partir de deux valeurs sur le bord de l'hexagone, phase $B(u,\pi-u)$ sur AC et phase $B(u+v,\pi-u-v)$ sur EG et une troisième à l'intérieur de ACDEGH, phase $B(v,\pi-u-v)$. Cependant l'estimation de phase $B(u,\pi-u)$ peut soulever des difficultés d'ordre théorique et pratique.

Dans le cas où il est possible de contrôler les signaux d'entrée (par exemple, en communication), on peut utiliser un signal non stationnaire constitué d'impulsions aléatoires indépendantes à la fréquence d'échantillonnage et appliquer sur le signal de sortie les algorithmes développés pour les signaux déterministes. Notons que la plupart des algorithmes développés dans le domaine temporel ne s'appliquent correctement que dans cette situation.

Rappelons que pour simuler un signal correspondant à l'échantillonnage d'un signal stationnaire du troisième ordre à bande limitée, on peut utiliser la méthode proposée au début de la communication par exemple pour $\Delta t = 1/2$:

$$z(k) = \sum_{p=-\infty}^{\infty} s(p/2) \frac{\sin \pi(k-p/2)}{\pi(k-p/2)}, \quad (19)$$

où $s(p\Delta t)$ est une séquence d'échantillons indépendants non gaussiens. Alors, le TOS du signal analysé s'annulera en théorie dans ABC et EFG. Rappelons que la médiocre qualité des générateurs de nombres aléatoires peut causer des difficultés supplémentaires.

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] D. Brillinger and M. Rosenblatt, "Computation and interpretation of Kth-order spectra", Spectral analysis of time series, B. Harris, ed., Wiley, New York, 1967, pp 189-232.
- [2] G. B. Giannakis, "Signal reconstruction from multiple correlations : frequency and time domain approaches", J. Opt. Soc. Am. A/vol. 6, n° 5, May 1989, pp 682-697.
- [3] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics", IEEE Trans. on ASSP, vol. 37, No 3, March 1989, pp 360-377.
- [4] M. J. Hinich, "Detecting a Transient Signal by Bispectral Analysis", IEEE Trans. on ASSP, vol. 38, n° 7, July 1990, pp 1277-1283.
- [5] J. Le Roux, "Reconstruction of a sampled signal Fourier transform from its bispectrum", International Signal Processing Workshop on Higher Order Statistics, Chamrousse, France, 10-12 July 1992, pp 233-236.
- [6] A. W. Lohmann, G. Weigelt and B. Wirmitzer, "Speckle masking in astronomy : triple correlation, theory and applications," Applied Optics, vol. 22, 1983, pp 4028-4037.
- [7] J. C. Marron, P. P. Sanchez and R. C. Sullivan, "Unwrapping algorithm for least-squares phase recovery from the modulo 2π bispectrum phase", J. Opt. Soc. Am., vol. 7, n° 1, January 1990, pp 14-20.

[8] T. Matsuoka and T. J. Ulrych, "Phase estimation using the bispectrum", Proceedings of the IEEE, vol. 79, n° 10, October 1984, pp 1404-1411.

[9] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory : theoretical results and applications", Proceedings of the IEEE, vol. 79, n° 3, March 1991, pp 278-305.

[10] C. L. Nikias and M. Raghuveer, "Bispectrum estimation: a digital signal processing framework", Proc. of the IEEE, vol. 75, 1987, pp 869-891.

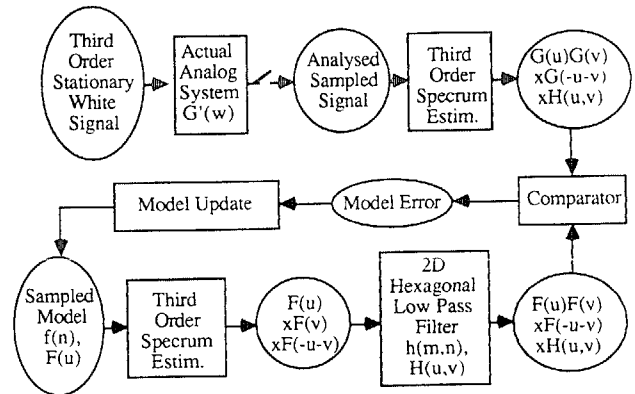


Fig. 4. Schéma algorithmique permettant l'utilisation des techniques bispectrales pour la modélisation des systèmes dans le cas des signaux analogiques stationnaires