



HAL
open science

Estimation et détection de signaux cyclostationnaires par les méthodes de ré-échantillonnage statistique : applications à l'analyse des signaux biomécaniques

Sofiane Maiz

► **To cite this version:**

Sofiane Maiz. Estimation et détection de signaux cyclostationnaires par les méthodes de ré-échantillonnage statistique : applications à l'analyse des signaux biomécaniques. Biomécanique [physics.med-ph]. Université Jean Monnet - Saint-Etienne, 2014. Français. NNT : 2014STET4020 . tel-01623766

HAL Id: tel-01623766

<https://theses.hal.science/tel-01623766v1>

Submitted on 25 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT
DE L'UNIVERSITE JEAN MONNET – SAINT-ETIENNE-

Présentée par
Monsieur Sofiane Maiz

Pour obtenir le grade de
DOCTEUR EN SCIENCES DE L'INGÉNIEUR

Domaine:
IMAGE, VISION, SIGNAL

Sujet de la thèse :

**ESTIMATION ET DÉTECTION DE SIGNAUX CYCLOSTATIONNAIRES PAR
LES MÉTHODES DE RÉ-ÉCHANTILLONNAGE STATISTIQUE :
APPLICATIONS À L'ANALYSE DES SIGNAUX BIOMÉCANIQUES**

Thèse présentée et soutenue à Roanne le 04 Décembre 2014 devant le jury composé de:

J. Antoni	Professeur à LVA, INSA de Lyon	Rapporteur
K. Abed-Meraim	Professeur à PRISME, Polytech Orléans	Rapporteur
M. El badaoui	Professeur à l'UJM, Saint-Étienne	Directeur
C. Servière	C. Recherche CNRS à Gipsalab, Grenoble	Codirecteur
J. Leskow	Professeur, TU of Cracow, Pologne	Examinateur
F. Roche	Professeur à CHU de Saint Saint-Étienne	Examinateur
F. Guillet	Professeur à l'UJM, Saint-Étienne	Examinateur

Table des matières

Introduction générale	7
1 La cyclostationnarité	13
1.1 Introduction	14
1.2 Genèse de la théorie de la cyclostationnarité	14
1.3 Définitions	17
1.3.1 Les processus stochastiques	17
1.3.2 Les processus stochastiques au sens strict	18
1.3.3 Les propriétés statistiques à l'ordre un et deux	19
1.3.4 Les processus stochastiques au sens large	20
1.4 Ergodicité	24
1.5 Modèles cyclostationnaires	26
1.5.1 Modulation d'amplitude aléatoire	26
1.5.2 Modulation d'amplitude aléatoire bi-composante	28
1.5.3 Modulation d'amplitude aléatoire par une fonction périodique	31
1.5.4 Modulation d'amplitude aléatoire bi-composante avec dépendance	34
1.5.5 Modulation numérique linéaire	36
1.6 De la cyclostationnarité à la cyclostationnarité généralisée	39
1.6.1 Définition d'un processus cyclostationnaire généralisé	40
1.6.2 Cas cyclostationnaire généralisé	42
1.6.3 Étude de modèle à fréquence cyclique variable	42
1.6.4 Étude de simulation	45
1.7 Conclusion	48
2 Le Bootstrap	51
2.1 Introduction	53
2.2 Le bootstrap et ses applications	54
2.2.1 Définition du bootstrap	54
2.2.2 Avantages du bootstrap	54
2.2.3 Cas d'utilisation du bootstrap	55
2.2.4 Les domaines d'application du bootstrap	56
2.3 Le principe du bootstrap	57
2.4 Résultats théoriques du bootstrap pour la moyenne	59

2.4.1	La consistance du bootstrap pour la moyenne	59
2.4.2	La consistance du bootstrap pour la moyenne standardisée	61
2.4.3	La consistance du bootstrap pour la moyenne studentisée	61
2.5	Exemples d'estimations par le bootstrap non paramétrique . .	62
2.5.1	Estimation du biais	62
2.5.2	Estimation de la variance	64
2.5.3	Intervalles de confiance pour la moyenne	65
2.6	Bootstrap paramétrique	66
2.6.1	Intervalles de confiance pour la moyenne via le boots- trap paramétrique	67
2.7	Les limitations du bootstrap	69
2.7.1	Taille d'échantillon très petite	69
2.7.2	Séries m-dépendantes	70
2.7.3	Distribution avec des moments infinis	70
2.7.4	Le sondage	71
2.7.5	Dépendance à long terme	72
2.8	Conclusion	72
3	Ré-échantillonnage statistique dans le cas de signaux cyclo- stationnaires	73
3.1	Introduction	74
3.2	La formulation du problème	75
3.3	Le Moving Block Bootstrap (MBB)	77
3.3.1	Le MBB pour les coefficients de Fourier	78
3.3.2	Algorithme du MBB	79
3.4	Le subsampling	80
3.4.1	Le subsampling pour les coefficients de Fourier	80
3.4.2	Algorithme du subsampling	82
3.5	Etude de simulation et de performances	83
3.6	Application du subsampling sur des signaux biomécaniques . .	91
3.6.1	La description des signaux de GRF	91
3.6.2	Pré-traitement des signaux de GRF	93
3.6.3	Analyse cyclostationnaire basé sur le subsampling	93
3.6.4	Cartographies de significativité fréquentielle cyclosta- tionnaire	96
3.6.5	La comparaison entre les trois niveaux de fatigue du coureur	98
3.7	Conclusion	100

4	Inférence statistique temporelle des signaux cyclostationnaires par la méthode du Bloc Bootstrap Généralisé	103
4.1	Introduction	105
I	Estimation de la cyclostationnarité d'ordre 1 par la méthode du bloc bootstrap généralisé	106
4.2	Présentation de la problématique	108
4.2.1	Définitions basiques	108
4.2.2	Séparation de composantes CS1/CS2	109
4.3	Le GSBB temporel pour l'estimateur de la moyenne synchrone	113
4.3.1	Algorithme du GSBB	114
4.3.2	Les intervalles de confiance GSBB pour la moyenne synchrone	116
4.3.3	Étude de simulation	116
II	Application du GSBB à l'analyse de la marche à pied	123
4.4	Introduction	125
4.5	Méthodologie	126
4.5.1	Acquisition des données	126
4.5.2	Fluctuations de vitesse et description des signaux de marche	127
4.5.3	Le ré-échantillonnage synchrone	128
4.5.4	La séparation CS1/CS2	133
4.5.5	Mesure de la qualité de l'estimation	134
4.5.6	Indicateur de séparabilité CS1/CS2 basé sur le GSBB	134
4.6	Résultats	135
4.6.1	Le ré-échantillonnage synchrone	136
4.6.2	Analyse cyclostationnaire d'ordre deux CS2	136
4.6.3	Estimation de CS1 par GSBB	139
4.6.4	Indicateur de séparabilité CS1/CS2	141
4.7	Application à l'analyse de la marche en double tâche	143
4.7.1	Le paradigme expérimental de la marche avec double tâche	144
4.7.2	Analyse des paramètres temporels de la foulée	145
4.7.3	Analyse d'ordre deux des signaux de marche avec double tâche	147

4.7.4	Étude de groupe	148
4.8	Conclusion	153
5	Le GSBB dans le domaine fréquentiel	155
5.1	Introduction	156
5.2	Le GSBB	156
5.3	Problématiques et consistance du GSBB	157
5.4	Étude de simulation : le GSBB fréquentiel	164
5.5	Application du GSBB sur des données réelles	169
5.6	Conclusion	172
	Conclusion générale	173
A	La relation entre l'analyse d'enveloppe et la corrélation spectrale intégrée	181
B	Consistance du GSBB dans le domaine fréquentiel	184
	Bibliographie	191

Table des figures

1.1	Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)\cos(2\pi f_0 t)$	27
1.2	Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)\cos(2\pi f_0 t)$	28
1.3	Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)\cos(2\pi f_0 t)$	29
1.4	Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$	30
1.5	Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$	31
1.6	Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$	32
1.7	Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)f(t)$	33
1.8	Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)f(t)$	34
1.9	Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)f(t)$	35
1.10	Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$	36
1.11	Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$	37
1.12	Projection sur l'axe des α de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$	38
1.13	Fonction d'autocorrélation cyclique en échelle linéaire du signal modulé en amplitude $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$	39
1.14	Fonction d'autocorrélation cyclique en échelle logarithmique du signal modulé en amplitude $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$	40
1.15	Projection sur l'axe des α de la fonction d'autocorrélation cyclique du signal $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$	41
1.16	Les relations théoriques entre les fréquences cycliques et le paramètre de retard τ pour le modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire	44

1.17	Les relations théoriques entre les fréquences cycliques et le paramètre de retard τ pour le modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire	46
1.18	Fonction d'autocorrélation cyclique en échelle linéaire du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire	47
1.19	Fonction d'autocorrélation cyclique en échelle logarithmique du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire	48
1.20	Fonction d'autocorrélation cyclique en échelle linéaire du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire	49
1.21	Fonction d'autocorrélation cyclique en échelle logarithmique du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire	50
2.1	L'approche du Bootstrap.	55
2.2	Principe du bootstrap non-paramétrique pour l'estimation de la fonction de distribution de $\hat{F}(\hat{\theta})$ du paramètre $\hat{\theta}$ [142].	59
2.3	Exemple de distribution approximée par le bootstrap non paramétrique. Le paramètre d'intérêt est la moyenne. $X \sim \mathcal{N}(0, 1)$ de taille $n = 30$	60
2.4	Exemple d'histogramme de $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$ obtenus par bootstrap paramétrique. $X \sim \mathcal{N}(3, 25)$ de taille $n = 15$	68
3.1	Les valeurs estimées de $\hat{R}_X(\alpha, \tau)$ sur le signal simulé $X(t)$: (a). RSB = -5 dB ; (b). RSB = 0 dB ; (c). RSB = 10 dB ; (d). RSB = 20 dB.	85
3.2	Valeurs des ACP pour le subsampling et le MBB : (a) et (b). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = -5 dB. (c) et (d). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 0 dB.	86
3.3	Valeurs des ACP pour le subsampling et le MBB : (e) et (f). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 10 dB. (g) et (h). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 20 dB.	87
3.4	Valeurs de la fonction d'autocorrélation cyclique de $X(t)$ avec un RSB = -5dB (itération numéro 1). (a). Le module de $\hat{R}_X(\alpha, \tau)$; (b). La partie réelle de $\hat{R}_X(\alpha, \tau)$; (c). La partie imaginaire de $\hat{R}_X(\alpha, \tau)$	88
3.5	Valeurs moyennes des largeurs des intervalles de confiance pour le subsampling et le MBB.	90
3.6	Tapis roulant dynamométrique.	92
3.7	Exemple d'un signal de GRF	92

3.8	Les valeurs estimées de $\hat{R}_X(\alpha, \tau)$ sur les résiduels provenant des signaux de GRF : (a). T = 0h ; (b). T = 12h ; (c). T = 24h.	94
3.9	(a),(c) et (e). Les parties réelles de $\hat{R}_X(\alpha, 0)$, respectivement pour T=0 ; 12 et 24h. (b), (d) et (f). Les parties imaginaires de $\hat{R}_X(\alpha, 0)$ respectivement pour T=0 ; 12 et 24h. Les points et les lignes verticales représentent respectivement les valeurs de l'estimateur et les intervalles de confiance du subsampling pour 95%.	95
3.10	(a),(c) et (e). Parties réelles des cartographies de significativité fréquentielle de $\hat{R}_X(\alpha, \tau)$, respectivement pour T=0 ; 12 et 24h. (b), (d) et (f). Parties imaginaires des cartographies de significativité fréquentielle de $\hat{R}_X(\alpha, \tau)$ respectivement pour T=0 ; 12 et 24h. En fonction de α , τ et δ	97
3.11	(a) et (b). Matrices de probabilité estimées de $\hat{R}_X(\alpha, \tau)_{T_{12}} > \hat{R}_X(\alpha, \tau)_{T_0}$ respectivement pour les parties réelle et imaginaires. (c) et (d). Matrices de probabilité estimées de $\hat{R}_X(\alpha, \tau)_{T_{24}} > \hat{R}_X(\alpha, \tau)_{T_{12}}$ respectivement pour les parties réelle et imaginaires. Les matrices sont seuillées à un niveau de confiance $\delta = 0.05$	99
4.1	Diagramme schématique de la méthode ANC.	110
4.2	Diagramme schématique de la méthode SANC.	111
4.3	Algorithme du GSBB en pratique.	115
4.4	Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_1(t)$. (a) RSB=-5dB ; (b) RSB=10dB.	118
4.5	Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_2(t)$. (a) RSB=-5dB ; (b) RSB=10dB.	119
4.6	Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_1(t)$ avec les intervalles de confiance GSBB : moyenne synchrone (rouge) ; limite supérieure (bleu) et limite inférieure (magenta). (a) RSB=-5dB ; (b) RSB=10dB.	121
4.7	Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_2(t)$ avec les intervalles de confiance GSBB : moyenne synchrone (rouge) ; limite supérieure (bleu) et limite inférieure (magenta). (a) RSB=-5dB ; (b) RSB=10dB.	122
4.8	Dispositif d'acquisition SMTEC FootSwitch	126
4.9	Un signal de marche.	127
4.10	La vitesse instantanée du signal de marche.	129
4.11	La transformée de Fourier du signal de marche.	130
4.12	Les étapes d'un cycle de marche.	131

4.13	Transformées de Fourier de $X(t)$ (en bleu) et $X(\Theta)$ (en rouge).	136
4.14	Les valeurs estimées de la fonction d'autocorrélation cyclique sur le résiduel $\hat{V}(\Theta)$ (après synchronisation) : (a) $\hat{R}(\alpha, \tau)$. (b) $\hat{R}(\alpha, 0)$.	137
4.15	Les valeurs estimées de la fonction d'autocorrélation cyclique sur le résiduel $\hat{V}(t)$ (sans synchronisation) : (a) $\hat{R}(\alpha, \tau)$. (b) $\hat{R}(\alpha, 0)$.	138
4.16	L'estimée de la moyenne synchrone par GSBB avec 95% d'intervalles de confiance -le signal synchronisé $X(\Theta)$ -.	140
4.17	L'estimée de la moyenne synchrone par GSBB avec 95% d'intervalles de confiance -le signal original $X(t)$ -.	141
4.18	Les valeurs du PSM pour $\hat{V}(\Theta)$ issu du signal synchronisé $X(\Theta)$.	142
4.19	Les valeurs du PSM pour $\hat{V}(t)$ issu du signal synchronisé $X(t)$.	143
4.20	Exemple de détection de maximums et minimums locaux sur un signal de marche (les points rouges).	145
4.21	Tests de Newman-Keuls sur les valeurs des temps de foulées moyens : (a) capteur du talon du pied droit; (b) capteur du métatarse/orteils du pied droit; (c) capteur du talon du pied gauche; (d) capteur du métatarse/orteils du pied gauche. Les trois conditions MS, MD et MF sont représentées sur l'axe des ordonnées des figures.	146
4.22	Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche simple (MS).	148
4.23	Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche avec décompte (MD).	149
4.24	Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche avec fluence verbale (MF).	150
4.25	Un exemple du résultat d'analyse d'enveloppe sur un signal de marche.	151
4.26	Les valeurs de la CS2 pour les trois conditions de marche.	151
5.1	Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ sur le signal simulé $X(t)$: (a). $RSB = -5$ dB; (b). $RSB = 0$ dB; (c). $RSB = 10$ dB; (d). $RSB = 20$ dB.	166
5.2	(a) Les valeurs estimées de $ \hat{b}(\gamma) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). $RSB=-5$ dB.	167

5.3	(a) Les valeurs estimées de $ \hat{b}(\gamma) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=0dB.	167
5.4	(a) Les valeurs estimées de $ \hat{b}(\gamma) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=10dB.	168
5.5	(a) Les valeurs estimées de $ \hat{b}(\gamma) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=20dB.	168
5.6	Signal de marche.	169
5.7	(a). Les valeurs estimées de $ \hat{b}(\gamma) $ pour le signal de marche ; (b). Les valeurs estimées de $ \hat{R}(\alpha, \tau) $ pour le résiduel du signal de marche.	170
5.8	(a). Les valeurs estimées de $ b(\gamma) $ (points noirs) avec 95% d'intervalles de confiance GSBB (lignes verticales) pour le signal de marche ; (b). Les valeurs estimées de $ R(\alpha, \tau) $ (points noirs) avec 95% d'intervalles de confiance GSBB (lignes verticales) pour la composante CS2 du signal de marche.	171
A.1	Relation entre l'analyse d'enveloppe et la corrélation spectrale intégrée d'un signal CS2.	182

Introduction générale

Dans le contexte actuel, les nouvelles technologies prennent une grande place dans la vie de tous les jours. En effet, le développement de l'outil informatique et de l'électronique simplifient grandement les différentes tâches du quotidien. L'évolution de ces technologies implique que le traitement du signal est de plus en plus utilisé en mécanique et dans le domaine de la santé. En effet, il est très important de maîtriser les coûts et de les réduire par le biais de développement d'outils de diagnostic précoce et de surveillance. Dans le domaine de la mécanique, la maintenance de la machine de production est primordiale. Un défaut non détecté peut conduire à un arrêt de la production et par conséquent, entraîne des pertes financières qui peuvent s'avérer très conséquentes. Par ailleurs, de plus en plus de travaux en biomécanique s'intéressent à l'étude de la locomotion humaine à des fins de diagnostic et de prévention pour optimiser les coûts de santé. Dans cette thèse, l'intérêt est porté sur ces derniers types de signaux. En effet, ce travail de thèse rentre dans le cadre de collaborations entre le LASPI, le GIPSAlab et le LPE (Laboratoire de Physiologie de l'Exercice) d'une part, et entre le LASPI et le CHU de Saint Etienne d'autre part.

La compétition sportive est de plus en plus technique, il devient important d'une part de mesurer les performances de l'équipement et d'autre part de quantifier l'état physique des sportifs afin d'optimiser leurs performances tout en les protégeant d'éventuelles blessures (musculaires, squelettique... etc.). Dans cette optique, la collaboration étroite entre le LASPI et le LPE avait pour objectif de mesurer et analyser l'impact des contraintes mécaniques répétées intervenant lors de la locomotion prolongée et intense chez les sportifs de haut niveau sur la structure musculo-squelettique. Ceci pourrait être valorisé et développé dans le cadre de l'amélioration des équipements sportifs (par exemple, amorti des chaussures), et du suivi et de la prévention des activités physiques et sportives (exemple : monitoring des coureurs de haut niveau aussi bien que le jogger occasionnel).

D'autre part, et dans la collaboration du LASPI avec le CHU de Saint Etienne, nous nous sommes penchés sur l'étude du phénomène très complexe de la marche humaine, particulièrement, dans le cas de la population des gens âgés. Cette population est très souvent sujette à de nombreux dysfonctionnement moteurs et/ou neuro-moteurs. Quantifier les paramètres spatio-

temporels de la marche chez cette population peut aider à l'identification précoce de la chute ainsi qu'à la caractérisation de quelques troubles et maladies fortement liées à la marche.

La locomotion humaine ainsi que les machines tournantes produisent des signaux cycliques selon un ou plusieurs cycles de base se répétant continuellement. La nature périodique liée à ces phénomènes conduit naturellement à exploiter les propriétés cyclostationnaires de ces signaux, c'est-à-dire, la périodicité de leurs paramètres statistiques.

L'analyse statistique des signaux a pour rôle de construire des estimateurs et des tests. On attend très souvent de ces estimateurs d'avoir la plus petite variance possible et des tests d'être très efficaces afin de minimiser la probabilité de fausse détection. Dans le traitement du signal classique, ces tâches étaient très souvent accomplies en imposant des hypothèses de lois de distributions sur les données ou les signaux étudiés (la Gaussiennité est incontestablement la plus utilisée). Ou bien, en démontrant des propriétés asymptotiques (très grand nombre d'échantillons) des distributions des estimateurs d'intérêt. Les méthodes utilisées pour le traitement des signaux cyclostationnaires ne font pas exception à de telles hypothèses.

Dans le contexte des signaux mécaniques et biomécaniques, l'hypothèse de gaussiennité n'est généralement pas vérifiable. En effet, ceci est dû principalement au caractère sinusoïdal ou impulsionnel de ce type de signaux. De plus, cela est d'autant plus vrai que la présence de défauts, de défaillances ou d'anomalies entraînent des dérivations dans les signaux. Par conséquent, la loi de distribution qui régit les données diverge de plus en plus de la loi normale. D'autre part, l'approximation asymptotique des estimateurs par une loi Gaussienne n'est malheureusement pas très utile en pratique à cause de la forme très complexe des matrices de variance-covariance des distributions approximées.

Une alternative aux problématiques citées ci-dessus est l'utilisation et le développement de méthodes de ré-échantillonnage statistique et de bootstrap. En effet, ce type d'approche permet d'obtenir une approximation de la distribution d'un estimateur d'intérêt sans avoir à imposer des hypothèses de gaussiennité. En outre, ces dernières permettent aussi de construire des tests statistiques valides pour l'étude de la significativité du paramètre étudié. En effet, dans le cas cyclostationnaire, on est souvent confronté à un problème assez complexe qui est de savoir si quelques fréquences estimées sont de vraies fréquences ou bien juste le résultat d'erreurs aléatoires. C'est pourquoi, il est fondamental d'étudier la significativité des fréquences d'ordre un et deux des signaux cyclostationnaires pour une prise de décision adéquate.

Par ailleurs, les signaux mécaniques et biomécaniques sont caractérisés par la présence de composantes cyclostationnaires d'ordre un (CS1) très importantes comparés aux composantes cyclostationnaires d'ordre deux (CS2), contrairement aux signaux issus des télécommunications. Ainsi, l'estimation et la séparation entre ces deux composantes est une issue méthodologique fondamentale. C'est pourquoi, il est très important de développer des méthodes et des indicateurs pour mesurer la qualité de l'estimation des composantes CS1 et CS2 et leur séparabilité.

L'objectif de cette thèse est d'expliquer comment exploiter les méthodes de ré-échantillonnage statistique et de bootstrap dans un cadre cyclostationnaire et leur application sur des signaux issus de la biomécanique. Les approches développées sont illustrées sur des signaux de synthèse afin de les valider puis sur des signaux réels provenant de coureurs de très haut niveau et d'une population de personnes âgées. Les méthodes présentées dans cette thèse ont permis la construction de différents tests statistiques valides pour les signaux cyclostationnaires, et ce, dans le domaine temporel et le domaine fréquentiel. De plus, ces méthodes nous ont permis de fixer des seuils de significativité statistique pour différents paramètres caractérisant la CS1 et la CS2. En outre, nous proposons aussi de nombreux indicateurs basés sur les méthodes développées dans cette thèse afin de décrire la locomotion humaine ainsi que de détecter les défaillances qui y sont liées.

Dans le premier chapitre nous présentons les bases théoriques de la cyclostationnarité ainsi que ses divers descripteurs. Ensuite, nous proposons et étudions différents modèles cyclostationnaires dont certains sont innovants. Les modèles étudiés sont aussi liés à de possibles applications dans différents domaines. Enfin, nous étendons la notion de cyclostationnarité généralisée à l'étude du cas de fréquence cyclique variable dans le temps. Une étude théorique d'un nouveau modèle cyclostationnaire généralisé est présentée et des travaux de simulations sont réalisés. Ces simulations confirment les résultats théoriques obtenus ce qui ouvre des perspectives très prometteuses quant à leur application en pratique.

Dans le deuxième chapitre, nous introduisons la notion du bootstrap. En effet, la définition ainsi que le principe du bootstrap sont présentés. Puis, des exemples d'application du bootstrap dans différents domaines sont exposés. Quelques résultats théoriques sont aussi donnés pour deux types de bootstrap qui sont le bootstrap non-paramétrique et le bootstrap paramétrique. Des exemples d'estimation de la moyenne, la variance, le biais et les intervalles de confiance par bootstrap sont ensuite présentés et quelques simulations sont réalisées.

Dans le troisième chapitre, nous nous concentrons sur la consistance et l'application de deux méthodes de ré-échantillonnage statistique pour l'estimateur des coefficients de Fourier de la fonction d'autocorrélation d'un signal cyclostationnaire. Les deux méthodes sont appelées le Subsampling et le Moving Block Bootstrap (MBB). L'objectif de ce chapitre est de construire des tests statistiques via le subsampling et le MBB pour la présence de cyclostationnarité, c'est-à-dire., tester l'hypothèse de non-cyclostationnarité du signal étudié versus l'hypothèse alternative de présence de structure cyclostationnaire dans ce dernier. Cela est fait en testant si les coefficients de Fourier de la fonction d'autocorrélation du signal sont nulles en dehors d'un ensemble fini de fréquences cycliques, ou pas. Dans un premier temps, une description de la problématique est donnée. Ensuite, les résultats fondamentaux de la consistance des deux méthodes ainsi que leurs algorithmes respectifs sont présentés. Une étude de simulation et de performance est ensuite réalisée. Enfin, quelques applications innovantes sur des signaux réels de course à pied et la construction de nouveaux indicateurs de CS2 sont proposées.

Le quatrième chapitre s'intéresse à l'estimation des composantes CS1 et CS2 dans le domaine temporel. Dans le cas des signaux biomécaniques, une forte composante CS1 est présente. Il est alors indispensable de séparer les composantes CS1 et CS2 pour les analyser indépendamment, d'autant plus qu'en pratique, les signaux CS présentent souvent des fluctuations de vitesse. L'estimation de la composante CS2 est souvent conditionnée par une estimation robuste de la CS1. Pour ce faire, une méthodologie combinant entre des outils de traitement du signal et de la théorie du bootstrap est développée dans ce chapitre. Ce quatrième chapitre est composé de deux parties principales. La première partie est consacrée à une nouvelle méthode de bootstrap appelée le Generalized Seasonal Block Bootstrap (GSBB). Nous utilisons cette méthode dans l'optique d'estimer la composante CS1 d'un signal CS dans le domaine temporel. Ainsi, dans un premier temps, la formulation du problème est d'abord présentée. Ensuite, la méthode du GSBB est décrite et son algorithme est proposé. Une étude de simulation est alors réalisée afin d'illustrer la performance du GSBB. Dans la deuxième partie de ce chapitre, nous présentons différentes applications du GSBB sur un signal réel issu de la marche à pied. Une méthode de ré-échantillonnage synchrone du signal est proposée et une étude comparative via le GSBB est développée afin de mettre en évidence l'influence de la présence de fluctuations de vitesse sur les propriétés CS du signal. De nouveaux indicateurs de séparabilité CS1/CS2 basés sur le GSBB sont aussi proposés pour mesurer la performance de la séparation CS1/CS2. Enfin, une étude de la marche à pied en condition de double tâche d'une im-

portante population de personnes âgées ainsi que les résultats relatifs sont exposées.

Après avoir introduit et défini le GSBB dans le domaine temporel dans le quatrième chapitre, nous proposons d'étendre le GSBB au domaine fréquentiel dans le cinquième chapitre. En effet, ce chapitre se focalise sur l'estimation et la détection de la CS1 et de la CS2 dans le domaine fréquentiel. Les problématiques d'estimation des coefficients de Fourier de la moyenne ainsi que de la fonction d'autocorrélation d'un signal cyclostationnaire sont présentées et l'application du GSBB pour la construction d'intervalles de confiance pour les fréquences d'ordre un et deux d'un signal cyclostationnaire est réalisée. Par conséquent, il devient possible de proposer des tests de présence de structures CS1 et CS2 dans le domaine fréquentiel. Ce chapitre est organisé comme suit : les problématiques sont tout d'abord formulées et les résultats principaux de la consistance du GSBB dans le domaine fréquentiel sont présentés dans le cas de la CS1 et la CS2. Ensuite, une étude de simulation est réalisée afin de valider la méthode et les tests statistiques qui sont construits. Enfin, une application sur un signal réel est exposée et les résultats qui y sont liés sont présentés.

Nous terminons par une conclusion sur l'ensemble des résultats obtenus pour chaque voie d'étude, avant de présenter les perspectives de recherche qui se dégagent de ce travail.

La cyclostationnarité

Sommaire

1.1	Introduction	14
1.2	Genèse de la théorie de la cyclostationnarité	14
1.3	Définitions	17
1.3.1	Les processus stochastiques	17
1.3.2	Les processus stochastiques au sens strict	18
1.3.3	Les propriétés statistiques à l'ordre un et deux	19
1.3.4	Les processus stochastiques au sens large	20
1.4	Ergodicité	24
1.5	Modèles cyclostationnaires	26
1.5.1	Modulation d'amplitude aléatoire	26
1.5.2	Modulation d'amplitude aléatoire bi-composante	28
1.5.3	Modulation d'amplitude aléatoire par une fonction périodique	31
1.5.4	Modulation d'amplitude aléatoire bi-composante avec dépendance	34
1.5.5	Modulation numérique linéaire	36
1.6	De la cyclostationnarité à la cyclostationnarité généralisée	39
1.6.1	Définition d'un processus cyclostationnaire généralisé	40
1.6.2	Cas cyclostationnaire généralisé	42
1.6.3	Étude de modèle à fréquence cyclique variable	42
1.6.4	Étude de simulation	45
1.7	Conclusion	48

1.1 Introduction

Les processus rencontrés dans le traitement statistique du signal ainsi que dans l'analyse des séries temporelles sont très souvent supposés stationnaires. Le nombre important d'algorithmes et de méthodes disponibles pour de tels processus témoigne de la nécessité de leur traitement et de leur analyse spectrale. Cependant, en raison de la nature variante dans le temps de la plupart des phénomènes physique, naturels et synthétiques, l'hypothèse simplificatrice d'invariance temporelle ainsi que la notion de stationnarité associée ne sont pas pleinement valables dans la pratique. Par conséquent, l'étude des systèmes variant dans le temps et des processus non-stationnaires paraît être bien plus en adéquation avec la nature réelle des processus physiques, naturels et synthétiques.

Les travaux de recherche dédiés à l'étude des signaux et des systèmes non-stationnaires ont abouti au développement de divers algorithmes et différents outils très attractifs. Nous pouvons citer en exemple la transformée de Fourier glissante, les différentes représentations temps-fréquence telles que la représentation de Wigner-Ville, l'expansion de Karhunen-Loeve ainsi que les différentes représentations temps-fréquence et temps-échelles basées sur la théorie des ondelettes.

Un cas particulier des processus non-stationnaires est celui des phénomènes cyclostationnaires (CS), communément appelés les processus périodiquement corrélés (PC) par la communauté des statisticiens. Comme nous l'avons cité auparavant, la notion de périodicité est omniprésente dans les phénomènes naturels aussi bien que dans les processus synthétiques.

1.2 Genèse de la théorie de la cyclostationnarité

Les processus CS sont des processus aléatoires dans lesquels il existe un rythme périodique plus compliqué que la périodicité de la fonction de moyenne ou du moment d'ordre 1. De façon plus précise, les processus CS sont des processus aléatoires mais qui ont une structure périodique. Globalement, si un système physique contient de l'aléa couplé à de la périodicité, alors les mesures effectuées sur le système au fil du temps auront très probablement une structure périodiquement stationnaire. Ces périodicités ne sont donc pas accessibles de façon directe, c'est pourquoi on les qualifie de *périodicités cachées*. La cyclostationnarité est une propriété qui caractérise les processus stochastiques dont les propriétés statistiques varient périodiquement par rapport à certaines variables. Par conséquent, la cyclostationnarité permet d'accéder aux

périodicités cachées contenues dans de tels processus.

La notion de CS (ou PC) est apparue avec Bennett [19]. En effet, ce dernier a observé la présence de CS dans un contexte théorique lié aux télécommunications. Gudzenko [80] a ensuite introduit la notion d'analyse spectrale non paramétrique pour les processus CS. Quant à Markelov [111], il s'est intéressé à l'étude des processus CS dans le cadre Gaussien. Quelques temps après, Gladyshev [78] a publié pour la première fois sur le sujet de l'analyse des propriétés spectrales des processus CS et la relation entre les séquences CS et les séquences de vecteurs stationnaires. Ce dernier a montré que le support spectral d'un processus CS était formé d'une famille de lignes parallèles à la diagonale principales et espacées de $2\pi/T$. En 1963 Gladyshev [79] a traité les processus CS en temps continu, de plus, il a introduit la notion de processus presque-cyclostationnaires (*almost cyclostationary ou almost periodically correlated* en anglais).

Dans une série de papiers, Herbst [85, 86, 87, 88, 89, 90] a exploré des processus ayant des variances périodiques ou presque-périodique dans le temps. Toutefois, ces travaux n'ont utilisé aucune notion de cyclostationnarité. Par la suite, Brelford [31] a présenté des méthodes pour l'estimation des coefficients périodiques d'un modèle autorégressif périodique.

Dans d'autres travaux, Kampé de Fériet et Parzen [97, 98, 119, 120, 121] se sont focalisés sur les processus cyclostationnaires dans la nature, d'une part, et sur l'estimation asymptotique de la corrélation et des fonctions de densité spectrale d'autre part.

Les tous premiers exemples d'applications de la CS ont commencé avec Markelov [111]. Ce dernier a montré que le bruit en sortie d'un amplificateur paramétrique est cyclostationnaire si le bruit à l'entrée du système est stationnaire. Un autre exemple est celui cité par Monin [112] qui a suggéré l'utilisation de CS comme modèle pour les données météorologiques. La même chose a été suggéré par Jones et Brelford [96], mais cette fois-ci pour des séries temporelles de données de températures.

Plusieurs ouvrages traitant des différents aspects des processus CS sont apparus. Nous citons à titre d'exemples le livre de Papoulis [118]. Dans ce dernier, les processus stochastiques ainsi que les propriétés des processus CS sont discutés. Franks [65] a quant à lui abordé les processus cyclostationnaires dans un livre dédié à la théorie des télécommunications.

Hurd [91] a étudié la nature du support spectral dans le cas de processus cyclostationnaire à temps continu. Il a aussi établi le lien existant entre les processus CS et les processus pouvant être rendus stationnaires par un décalage temporel indépendant et uniformément distribué (*iid*). De plus, il a

obtenu des résultats de consistance pour l'estimateur de la fonction d'auto-covariance de tels processus. Gardner et Franks [67] ont développé différentes représentations des processus CS à temps continu et les ont utilisés dans divers problèmes d'estimations. Il est aussi important de rappeler qu'après les premiers travaux de Gladyshev, le sujet des processus CS a été beaucoup traité par Dragan et ses collaborateurs. Beaucoup de leurs travaux ont été rassemblés dans trois ouvrages uniquement en russe [46, 47, 48]. Dans Yaglom [141], des références très intéressantes ainsi qu'une claire exposition de nombreuses relations basiques dans la théorie de la CS ont été données.

Beaucoup de travaux sur les processus cyclostationnaires ont par la suite été réalisés. Ces travaux de recherche étaient essentiellement dédiés à des problématiques dans le domaine des télécommunications. Les deux ouvrages de Gardner [68, 69] traitent les processus cyclostationnaires dans deux cadres différents; une approche probabilistique (classique) et une approche appelée *Fraction Of Time* (FOT). Une discussion autour de ces deux approches différentes de traiter un processus CS est détaillée dans [70].

D'autre part, des travaux ultérieurs se sont penchés sur d'autres problématiques liées aux processus cyclostationnaires. De plus des travaux motivés par des problèmes de télécommunications, des travaux en statistique se sont intéressés à la théorie structurale des processus cyclostationnaires et presque-cyclostationnaires [29, 40, 41, 74]. D'autres ont étudié des problèmes d'estimation spectrale et d'estimation de la covariance [1, 2, 33]. Des problématiques de tests statistiques ont été aussi initiés [39, 71, 92] ainsi que l'étude des modèles ARMA périodiques ou *PARMA* [3, 108]. L'article de Gardner et Napolitano [76] résume de façon très intéressante le développement de la théorie de la cyclostationnarité durant les dernières cinquante années.

Par la suite, la cyclostationnarité a connu un grand succès auprès de nombreux praticiens dans des domaines divers et variés. En plus des télécommunications et de la météorologie citées précédemment, elle est très utile en diagnostic et surveillance des machines tournantes, en économétrie, en géophysique, en climatologie, en biologie et plus récemment en biomécanique. Un article bibliographique très complet sur la cyclostationnarité et ses applications dans différents domaines a été publié par Serpedin [133].

1.3 Définitions

1.3.1 Les processus stochastiques

Soit un processus aléatoire (stochastique) à temps continu et à valeurs réelles noté par $\{X(t, \omega), t \in \mathbb{R}, \omega \in \Omega\}$ et défini dans l'espace de probabilités (Ω, A, P) . Les processus stochastiques sont caractérisés par leurs fonctions de répartition et leurs densités de probabilité. Pour une meilleure compréhension de ce qui suit dans ce chapitre, nous rappelons brièvement les définitions de la fonction de répartition d'un processus stochastique ainsi que sa densité de probabilités.

1.3.1.1 La fonction de répartition

La fonction de répartition, encore appelée fonction de distribution cumulative ou fonction de distribution, d'une variable aléatoire $X(t)$ est définie comme étant la probabilité que $X(t) < x$. Donc la fonction de répartition de $X(t)$ est définie comme suit :

$$F_{X(t)}(x) = Prob\{X(t) < x\}$$

Il est important de remarquer que $F_{X(t)}(x)$ est une fonction de x et non de la variable aléatoire $X(t)$. Pour chaque valeur de x , $F_{X(t)}(x)$ représente une probabilité cumulée.

Une description alternative de la probabilité d'une variable aléatoire est souvent utilisée, il s'agit de la dérivée de la fonction de répartition.

1.3.1.2 La fonction de densité de probabilité

La fonction de densité de probabilité d'une variable aléatoire $X(t)$ est définie comme étant la dérivation de la fonction de répartition $F_{X(t)}(x)$ de la variable aléatoire $X(t)$ comme suit :

$$f_{X(t)}(x) = \frac{dF_{X(t)}(x)}{dx}$$

L'extension de la fonction de répartition d'une variable aléatoire à un ensemble de N variables aléatoires est direct. Soit $X(t) = [X(t+t_1), \dots, X(t+t_N)]$ un vecteur de N variables aléatoires. Ce vecteur représente des observations du processus stochastique $X(t, \omega)$ pour les temps $t : t + t_1, \dots, t + t_N$. La loi jointe pour le vecteur des N variables aléatoires évaluée à $x = \{x_1, \dots, x_N\}$ est définie par :

$$F_{X(t)}(x) = \text{Prob}\{X(t + t_1) < x_1, \dots, X(t + t_N) < x_N\}$$

Par conséquent, la fonction de densité de probabilité jointe correspondant est donnée par la dérivative N -ième de la fonction de répartition jointe de $X(t)$:

$$f_{X(t)}(x) = \frac{\partial^N}{\partial x_1, \dots, \partial x_N} F_{X(t)}(x)$$

Après ce petit rappel, nous présentons dans ce qui suit les différentes définitions des divers types de processus stochastiques au sens strict et au sens large.

1.3.2 Les processus stochastiques au sens strict

Dans cette partie, des définitions des différents types de processus stochastiques au sens strict sont données. Soit un processus stochastique à temps continu et à valeurs réelles $X(t)$. Le processus $X(t)$ est dit :

- *Stationnaire au sens strict à l'ordre N* : [72] si et seulement si ses propriétés statistiques sont indépendantes du temps t ou invariantes à un changement d'origine. En d'autres termes, cela signifie que les processus $X(t)$ et $X(t + \Delta\tau)$ ont les mêmes statistiques quelque soit $\Delta\tau$, i.e., $\forall \Delta\tau, t \in \mathbb{R}, \forall x \in \mathbb{R}^N$:

$$F_{X(t)}(x) = F_{X(t+\Delta\tau)}(x)$$

- *Cyclostationnaire au sens strict à l'ordre N* : [72] avec une période T_0 si et seulement si les propriétés statistiques du processus $X(t)$ restent invariantes à un décalage d'origine par des intégrales multiples de T_0 , i.e., $\forall t \in \mathbb{R}, \forall k \in \mathbb{Z}, \forall x \in \mathbb{R}^N, \exists T_0 \in \mathbb{R}^+$:

$$F_{X(t)}(x) = F_{X(t+kT_0)}(x)$$

- *Poly-cyclostationnaire au sens strict à l'ordre N* : [72] avec des périodes T_1, T_2, \dots, T_s si et seulement si les propriétés statistiques du processus $X(t)$ sont périodiques dans le temps avec les périodes T_1, T_2, \dots, T_s , i.e., $\forall t \in \mathbb{R}, \forall x \in \mathbb{R}^N, \exists T_s \in \mathbb{R}^+$:

$$F_{X(t)}(x) = F_{X(t+T_s)}(x)$$

- *Presque-cyclostationnaire au sens strict à l'ordre N* : [76] si et seulement si ses propriétés statistiques sont presque-périodiques dans le temps. Une fonction $g : \mathbb{R} \rightarrow \mathbb{C}$ est dite presque-périodique si pour tout réel $\varepsilon > 0$, il existe un polynôme trigonométrique $p_\varepsilon(t) = \sum_{n=-N}^N c_n e^{j\omega_n t}$ ($c_n \in \mathbb{C}, \omega_n \in \mathbb{R}$) tel que :

$$|g(t) - p_\varepsilon(t)| < \varepsilon, \quad \forall t \in \mathbb{R}$$

Donc, concrètement, un processus est dit presque-cyclostationnaire si :

$$\forall t_0 \in \mathbb{R}, \exists \tau_\varepsilon \in (t_0, t_0 + l_\varepsilon), |F_{X(t+\tau_\varepsilon)}(x) - F_{X(t)}(x)| < \varepsilon$$

1.3.3 Les propriétés statistiques à l'ordre un et deux

Les fonctions de répartitions ainsi que les fonctions de densité de probabilités d'un processus stochastique définis précédemment sont dépendantes du temps. Afin de décrire de façon complète un processus stochastique, la connaissance de sa fonction de répartition $F_{X(t)}(x)$ est obligatoire. Cependant, cette donnée n'est généralement pas connue dans les cas pratiques. C'est pourquoi, il est très rare d'obtenir toutes les fonctions de $F_{X(t)}(x)$. Dans beaucoup de cas pratiques, les processus aléatoires sont caractérisés que par la moyenne, l'énergie ou bien la variance du processus. Ces dernières quantités représentent les propriétés statistiques du processus au premier ordre et au second ordre.

La moyenne d'un processus stochastique $X(t, \omega)$ au temps t est définie comme étant la moyenne de la variable aléatoire $X(t)$ et peut être écrite comme suit :

$$\mu_X(t) = E\{X(t)\}$$

$\mu_X(t)$ est fonction du temps et est communément appelée la moyenne d'ensemble de $X(t)$ ou bien la statistique d'ordre un du processus stochastique $X(t)$.

Une autre propriété statistique est celle de la mesure de la dépendance du processus à deux instants différents t_1 et t_2 . Cette mesure est appelée la fonction d'autocorrélation instantanée du processus $X(t, \omega)$ [73]. Cette dernière mesure la corrélation entre deux variables aléatoires $X(t_1)$ et $X(t_2)$ (ou entre $X(t)$ et $X(t + \tau)$). La fonction d'autocorrélation est notée par $R_X(t_1, t_2)$ (ou $R_X(t, \tau)$) et est donnée comme suit :

$$R_X(t_1, t_2) = E\{X(t_1)X(t_2)\}$$

$$R_X(t, \tau) = E\{X(t)X(t + \tau)\} \quad \text{La version asymétrique}$$

$$R_X(t, \tau) = E\{X(t + \tau/2)X(t - \tau/2)\} \quad \text{La version symétrique}$$

Évidemment, cette fonction d'autocorrélation représente une statistique d'ordre deux du processus stochastique.

1.3.4 Les processus stochastiques au sens large

Dans cette partie, des définitions des différents types de processus stochastiques au sens large sont données. Soit un processus stochastique à temps continu et à valeurs réelles $X(t)$. Le processus $X(t)$ est dit :

- *Stationnaire au sens large* : [76] si et seulement si sa fonction de densité de probabilité du premier ordre est indépendante du temps. Dès lors, sa moyenne $\mu_X(t)$ est indépendante du temps. De plus, si sa fonction de densité de probabilité du second ordre, i.e., sa fonction d'autocorrélation $R_X(t_1, t_2)$ ne dépend que de la différence entre les temps d'observation t_1 et t_2 (ou si $R_X(t, \tau)$ ne dépend que de τ). Donc tout cela peut être défini comme suit : $\forall t_1, t_2 \in \mathbb{R}$ (ou $\forall t, \tau \in \mathbb{R}$) :

$$\mu_X(t) = \text{constante}$$

$$\text{et } R_X(t_1, t_2) = R_X(t_1 - t_2)$$

$$\text{ou } R_X(t, \tau) = R_X(\tau)$$

Il est donc évident que, si un processus aléatoire est stationnaire au sens strict, il est automatiquement stationnaire au sens large. Par contre, la réciproque n'est pas vraie. En outre, il est important de noter qu'il n'existe aucune corrélation entre les composantes spectrales d'un processus aléatoire stationnaire.

- *Cyclostationnaire au sens large (ou periodically correlated en anglais)* : [76] si et seulement si ses statistiques d'ordre un et d'ordre deux varient de façon périodique dans le temps. Cela signifie que sa moyenne $\mu_X(t)$ et sa fonction d'autocorrélation $R_X(t_1, t_2)$ (ou $R_X(t, \tau)$) sont périodiques dans le temps avec une période T_0 . On pourrait donc définir un processus cyclostationnaire au sens large comme suit : $\forall t_1, t_2 \in \mathbb{R}$ (ou $\forall t, \tau \in \mathbb{R}$), $\exists T_0 \in \mathbb{R}^+$ tel que :

$$\mu_X(t) = \mu_X(t + T_0)$$

$$\text{et } R_X(t_1, t_2) = R_X(t_1 + T_0, t_2 + T_0)$$

$$\text{ou } R_X(t, \tau) = R_X(t + T_0, \tau)$$

Alors, si un processus $X(t)$ est cyclostationnaire au sens large, T_0 est appelée *la période cyclique*. Dès lors que la fonction d'autocorrélation $R_X(t, \tau)$ est périodique, avec la période T_0 , elle admet un développement en séries de Fourier selon la variable t comme suit :

$$R_X(t, \tau) = \sum_{\alpha \in \mathcal{A}_\alpha} R_X(\alpha, \tau) e^{j2\pi\alpha t} \quad (1.1)$$

La somme ici est effectuée sur les multiples entiers de la fréquence fondamentale $\alpha_f = \frac{1}{T_0}$. Dans la décomposition en séries de Fourier de la fonction d'autocorrélation, α est appelée fréquence cyclique et l'ensemble $\mathcal{A}_\alpha = \{\alpha = \frac{k}{T_0}, k \in \mathbb{Z}\}$ représente l'ensemble des fréquences cycliques (la fréquence cyclique fondamentale et ses harmoniques). Quant aux coefficients de Fourier $R_X(\alpha, \tau)$, on les appelle les fonctions d'autocorrélations cycliques et on les écrit comme suit :

$$R_X(\alpha, \tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_X(t, \tau) e^{-j2\pi\alpha t} dt$$

Concrètement, la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ est en fonction de deux variables τ et α . Elle est continue par rapport à la variable τ et discrète par rapport à la variable α . De plus, cette fonction est non nulle pour quelques fréquences $\alpha \neq 0$ et est caractérisée par les propriétés de symétrie suivantes :

$$R_X(\alpha, -\tau) = R_X(\alpha, \tau)$$

$$R_X(-\alpha, \tau) = R_X^*(\alpha, \tau)$$

Une remarque très importante est celle de la fonction d'autocorrélation dans le cas de processus stationnaire. La fonction d'autocorrélation cyclique est nulle pour toutes les fréquences $\alpha \neq 0$. En effet, pour $\alpha = 0$, la fonction d'autocorrélation cyclique devient exactement la fonction d'autocorrélation classique. Donc pour un processus stationnaire, la fonction d'autocorrélation cyclique est nulle partout, sauf pour $\alpha = 0$.

Domaine fréquentiel

Le théorème de Wiener-Khintchine permet de définir la densité spectrale de puissance (DSP) d'un processus stationnaire comme étant la transformée de Fourier de sa fonction d'autocorrélation. Ce théorème s'étend au cas cyclostationnaire [69] et permet d'écrire la relation suivante :

$$S_X(t, f) = TF_\tau\{R_X(t, \tau)\} \quad (1.2)$$

Avec TF_τ la transformée de Fourier par rapport à la variable τ et $S_X(t, f)$ la densité spectrale de puissance instantanée. La DSP est fonction du temps t et de la fréquence f . Comme cette DSP est périodique par rapport au temps, elle admet par conséquent un développement en série de Fourier par rapport à la variable t :

$$S_X(t, f) = \sum_{\alpha \in \mathcal{A}_\alpha} S_X(\alpha, f) e^{j2\pi\alpha t} \quad (1.3)$$

De la même manière que dans le développement en série de Fourier de la fonction d'autocorrélation, la somme ici est effectuée sur les multiples entiers de la fréquence cyclique fondamentale α_f . De plus, le coefficient de Fourier dans cette décomposition est noté par $S_X(\alpha, f)$ et est appelé la fonction de corrélation spectrale :

$$S_X(\alpha, f) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} S_X(t, f) e^{-j2\pi\alpha t} dt$$

Par de simples calculs, et en remplaçant la fonction d'autocorrélation $R_X(t, \tau)$ dans l'équation (1.2) par son développement en séries de Fourier. Ensuite, l'identification du résultat obtenu par le développement de l'équation (1.3) nous donne une relation semblable à celle de Wiener-Khintchine dans le domaine des fréquences cycliques :

$$S_X(\alpha, f) = TF_\tau\{R_X(\alpha, \tau)\}$$

Par conséquent, la fonction de corrélation spectrale est tout simplement la transformée de Fourier de la fonction d'autocorrélation cyclique. On peut donc déduire qu'elle présente les mêmes propriétés que cette dernière, c'est à dire, elle est continue par rapport à la fréquence f et est

discrète par rapport à la fréquence cyclique α . La fonction de corrélation spectrale est aussi non nulle pour quelques $\alpha \neq 0$, de plus, elle présente les propriétés de symétrie suivantes :

$$S_X(\alpha, -f) = S_X(\alpha, f)$$

$$S_X(-\alpha, f) = S_X^*(\alpha, f)$$

Comme pour la fonction d'autocorrélation cyclique, dans le cas $\alpha = 0$ la fonction de corrélation spectrale se réduit au spectre classique défini dans le cas des processus stationnaires.

Finalement, on observe bien qu'il suffit d'appliquer une ou plusieurs transformées de Fourier afin de passer d'une fonction à l'autre, ou d'un domaine à l'autre.

- *Presque-cyclostationnaire au sens large (ou almost periodically correlated en anglais) : [74]* si et seulement si ses propriétés statistiques d'ordre un et d'ordre deux sont presque-périodiques dans le temps. Cependant, de façon générale, il est supposé que la moyenne d'un tel processus est nulle [115] et on se focalise sur sa statistique d'ordre deux, c'est à dire sa fonction d'autocorrélation $R_X(t, \tau)$.

Donc un processus est dit presque-cyclostationnaire si sa fonction d'autocorrélation présente plusieurs périodes T_1, T_2, \dots, T_k incommensurables entre elles. La différence entre un processus presque-cyclostationnaire et un processus cyclostationnaire réside dans le fait qu'il y ait plusieurs périodes au lieu d'une seule. La fonction d'autocorrélation $R_X(t, \tau)$ peut tout de même se mettre sous la forme suivante :

$$R_X(t, \tau) = \sum_{\alpha \in \mathcal{A}_\alpha} R_X(\alpha, \tau) e^{j2\pi\alpha t} \quad (1.4)$$

Avec un ensemble de fréquences cycliques dénombrable $\mathcal{A}_\alpha = \{\frac{1}{T_1}, \frac{1}{T_2}, \dots, \frac{1}{T_k}\}$

À première vue, cette décomposition est semblable à celle d'une décomposition en séries de Fourier, mais en réalité, elle ne l'est pas exactement. Dans ce cas, le développement en séries de Fourier n'est pas fait sur les harmoniques d'une seule fréquence cyclique fondamentale α_f . Il s'agit alors d'une décomposition en séries de Fourier-Bohr. Les coefficients $R_X(\alpha, \tau)$ sont alors appelés les coefficients de Fourier-Bohr et sont donnés par une moyenne temporelle comme suit :

$$R_X(\alpha, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t, \tau) e^{-j2\pi\alpha t} dt$$

La fonction de corrélation spectrale d'un processus presque cyclostationnaire au sens large est définie de la façon équivalente suivante :

$$S_X(\alpha, f) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} S_X(t, f) e^{-j2\pi\alpha t} dt$$

- *Poly-cyclostationnaires au sens large* : si et seulement si les statistiques d'ordre un et d'ordre deux du processus sont poly-périodiques dans le temps.

Une fonction $g : \mathbb{R} \rightarrow \mathbb{C}$ est dite poly-périodique avec les fréquences fondamentales associées $\{\nu_i\}_{i=1,\dots,q}$ si elle est presque-périodique, et si les fréquences associées aux coefficients de Fourier-Bohr c_k sont toutes multiples de l'un des ν_i . L'ensemble $\{\nu_i\}_{i=1,\dots,q}$ est tel que q est fini et les ν_i sont incommensurables.

Dès lors, un processus est poly-cyclostationnaire si ses statistiques sont poly-périodiques dans le temps. Ainsi, la somme dans l'équation (1.4) sera faite sur les fréquences cycliques appartenant à l'ensemble $\mathcal{A}_\alpha = \{\frac{m}{T_1}, \frac{m}{T_2}, \dots, \frac{m}{T_k}, m \in \mathbb{Z}^*\}$.

Il est important de noter que dès lors que les fonctions périodiques sont un cas particulier des fonctions poly-périodiques, les processus cyclostationnaires sont ainsi un cas particulier des processus poly cyclostationnaires.

Dans la suite du manuscrit, et sauf indication particulière, le terme cyclostationnaire est utilisé de façon indifférente.

1.4 Ergodicité

Un processus aléatoire est dit ergodique si les valeurs moyennes statistiques d'ensemble sont égales aux valeurs moyennes temporelles sur une seule réalisation [35]. La conséquence de cette hypothèse est très importante en pratique. En effet, les moyennes empiriques, lors de l'étude d'un système sont obtenues en utilisant un grand nombre de systèmes identiques appelés réalisations. Cependant, il est très rare, voire impossible, d'avoir accès à un nombre aussi

important de réalisation de ces systèmes identiques. C'est pourquoi, cette hypothèse d'ergodicité permet de remplacer les calculs de moments statistiques (qui supposent connues les densités de probabilité ou les fonctions de répartition) par les moyennes temporelles sur une seule réalisation (observation). En pratique, faire une estimation de la moyenne temporelle sur une fenêtre de taille infinie est impossible. Il faut se contenter d'une approximation calculée sur une fenêtre de taille finie, qui tendra asymptotiquement, avec la taille de la fenêtre, vers la moyenne temporelle. Donc, l'ergodicité dans le cas d'un système stationnaire permet de mesurer des moyennes à partir de l'observation d'un seul système, pendant un temps suffisamment long. Par conséquent, les moyennes d'ensemble peuvent être approchées par les moyennes temporelles comme suit :

$$E\{X(t)\} = \langle X(t) \rangle_t$$

Avec $\langle \cdot \rangle_t$ est l'opérateur de moyenne temporelle défini comme suit :

$$\langle \cdot \rangle_t = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} (\cdot) dt$$

Par ailleurs, Boyles et Gardner [29] se sont intéressés à la question de l'ergodicité dans le cas de processus cyclostationnaires. Ainsi, la notion d'ergodicité s'est étendue au cas des processus cyclostationnaires donnant lieu au concept de *cycloergodicité*. La cycloergodicité implique qu'asymptotiquement, les moyennes temporelles du processus multipliées par une sinusoïde complexe sont égales aux moyennes temporelles des paramètres statistiques du même processus multipliés par une sinusoïde complexe. Cela donne lieu à la relation suivante :

$$\langle E\{X(t)e^{j2\pi\alpha t}\} \rangle_t = \langle X(t)e^{j2\pi\alpha t} \rangle_t$$

L'application de cette dernière relation dans le cas d'un processus cyclostationnaire et supposé être cycloergodique donne lieu à une nouvelle expression de la fonction d'autocorrélation cyclique, soit :

$$R_X(\alpha, \tau) = \langle X(t - \frac{\tau}{2})X^*(t + \frac{\tau}{2})e^{-j2\pi\alpha t} \rangle_t$$

Ou de façon alternative :

$$R_X(\alpha, \tau) = \langle X(t)X^*(t + \tau)e^{-j2\pi\alpha t} \rangle_t$$

Ces deux expressions sont très pratiques dans la construction d'estimateurs de propriétés statistiques du second ordre des processus cyclostationnaires.

1.5 Modèles cyclostationnaires

Après avoir donné les différentes bases théoriques de la cyclostationnarité ainsi que les propriétés relatives à cette dernière, nous proposons dans cette partie d'étudier différents modèles cyclostationnaires afin de se fixer les idées. Certains des modèles étudiés ci-dessous sont connus alors que d'autres sont assez innovants (à notre connaissance). Nous rappelons aussi que sauf indication particulière, le terme cyclostationnaire est utilisé de façon indifférente.

1.5.1 Modulation d'amplitude aléatoire

Le premier modèle cyclostationnaire étudié est le signal modulé en amplitude. L'étude de ce modèle en premier semble très logique car la modulation d'amplitude est l'exemple le plus basique dans la théorie de la cyclostationnarité [76]. Le simple fait de moduler un signal aléatoire stationnaire par une fréquence porteuse rend le signal final cyclostationnaire. Sa période cyclique est directement liée à la fréquence de l'onde porteuse.

Soit $\{U(t), t \in \mathbb{Z}\}$ un processus stationnaire avec une moyenne nulle et une fonction de covariance $r_U(\cdot)$. Alors, le signal modulé en amplitude sera de la forme suivante :

$$X(t) = U(t)\cos(2\pi f_0 t)$$

Avec $t \in \mathbb{Z}$ et $f_0 \neq 0$. Le signal $X(t)$ est alors cyclostationnaire et sa fonction d'autocorrélation est donnée par :

$$R_X(t, \tau) = E\{X(t)X(t + \tau)\} = \cos(2\pi f_0 t)\cos(2\pi f_0(t + \tau))r_U(\tau)$$

$$R_X(t, \tau) = \frac{1}{2}r_U(\tau)\cos(2\pi f_0 \tau) + \frac{1}{4}r_U(\tau)e^{j2\pi f_0 \tau}e^{-j4\pi f_0 t} + \frac{1}{4}r_U(\tau)e^{-j2\pi f_0 \tau}e^{j4\pi f_0 t}$$

Donc, la fonction d'autocorrélation $R_X(t, \tau)$ du signal $X(t)$ est périodique pour l'ensemble des périodes correspondant aux fréquences cycliques \mathcal{A}_α suivant :

$$\mathcal{A}_\alpha = \{-2f_0, 0, 2f_0\}$$

Avec les valeurs des coefficients de Fourier non nulles (fonctions d'autocorrélations cycliques) suivants :

$$\begin{aligned}
 R_X(0, \tau) &= \frac{1}{2}r_U(\tau)\cos(2\pi f_0\tau) \\
 R_X(2f_0, \tau) &= \frac{1}{4}r_U(\tau)e^{-j2\pi f_0\tau} \\
 R_X(-2f_0, \tau) &= \overline{R_X(2f_0, \tau)}
 \end{aligned}$$

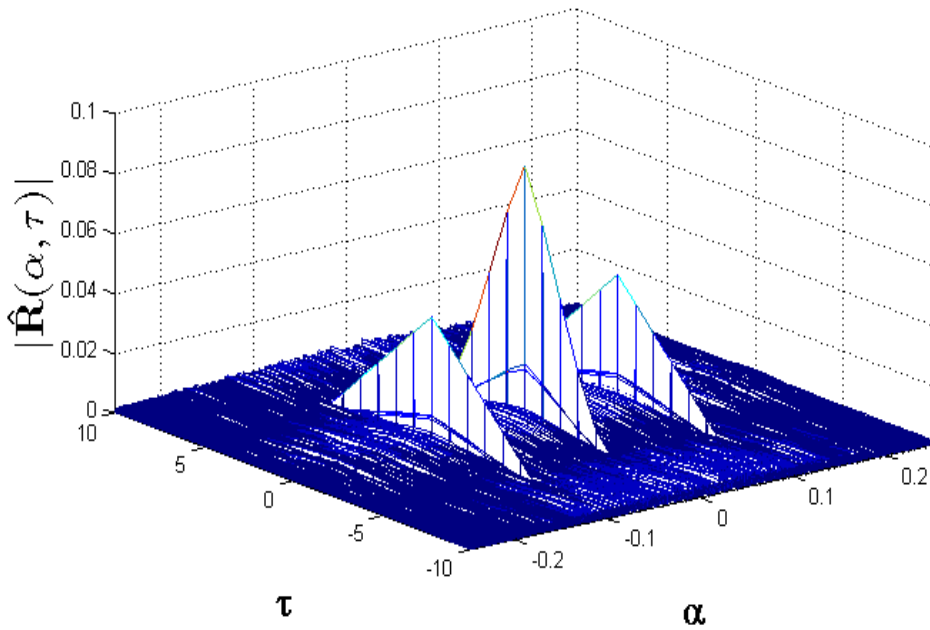


FIGURE 1.1 – Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)\cos(2\pi f_0 t)$.

Les figures 1.1, 1.2 et 1.3 montrent respectivement la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ du signal modulé en amplitude $X(t)$ en échelle linéaire, en échelle logarithmique et sa coupe à $\tau = 0$. La fréquence de la porteuse utilisée est $f_0 = 0.05\text{Hz}$.

Ce type de modèle est basique dans la théorie de la cyclostationnarité et a été très souvent utilisé dans les domaines des télécommunications et de la mécanique [76, 8].

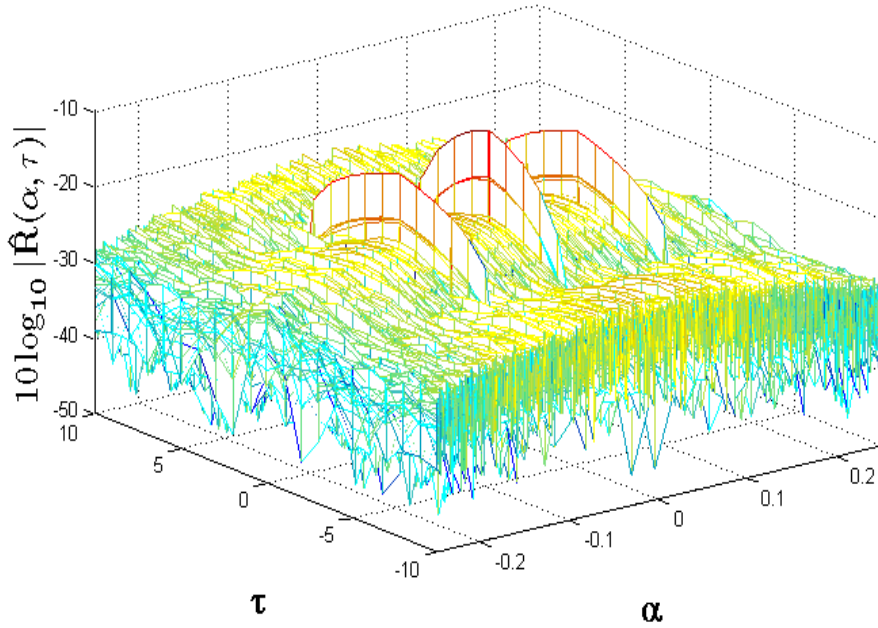


FIGURE 1.2 – Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)\cos(2\pi f_0 t)$.

1.5.2 Modulation d'amplitude aléatoire bi-composante

Le second modèle cyclostationnaire étudié est celui de la somme de deux modulations d'amplitudes. Le signal est donné comme suit :

$$X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$$

Le signal $X(t)$ est une somme de deux modulations d'amplitudes, telle que $\{U_i(t), t \in \mathbb{Z}\}$, $i = 1, 2$, soient deux processus indépendants, stationnaires, de moyennes nulles et avec des fonctions de covariances $r_i(\cdot)$ respectivement pour $i = 1, 2, \dots$. De plus $|f_1| \neq |f_2| \neq 0$

Procédant comme pour la simple modulation d'amplitude, nous retrouvons les résultats théoriques suivant pour la fonction d'autocorrélation de $X(t)$:

$$\begin{aligned} R_X(t, \tau) &= \frac{1}{2}r_1(\tau)\cos(2\pi f_1 \tau) + \frac{1}{2}r_2(\tau)\cos(2\pi f_2 \tau) + \frac{1}{4}r_1(\tau)e^{j2\pi f_1 \tau}e^{-j4\pi f_1 t} \\ &+ \frac{1}{4}r_1(\tau)e^{-j2\pi f_1 \tau}e^{j4\pi f_1 t} + \frac{1}{4}r_2(\tau)e^{j2\pi f_2 \tau}e^{-j4\pi f_2 t} + \frac{1}{4}r_2(\tau)e^{-j2\pi f_2 \tau}e^{j4\pi f_2 t} \end{aligned}$$

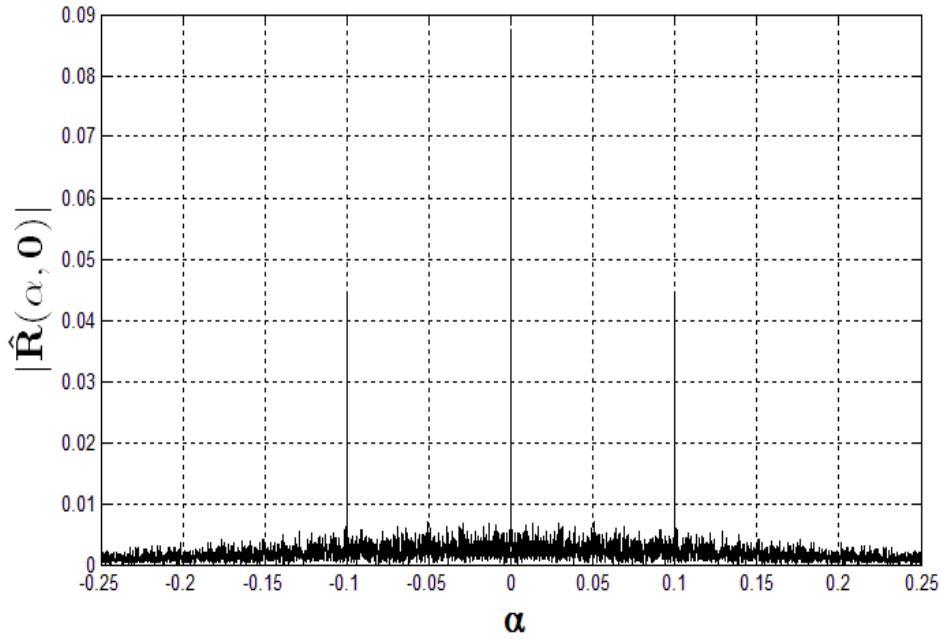


FIGURE 1.3 – Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)\cos(2\pi f_0 t)$.

Cela donne une fonction d'autocorrélation $R_X(t, \tau)$ périodique pour l'ensemble des périodes correspondant aux fréquences cycliques \mathcal{A}_α suivant :

$$\mathcal{A}_\alpha = \{-2f_2, -2f_1, 0, 2f_1, 2f_2\}$$

Avec les valeurs des coefficients de Fourier non nulles $R_X(\alpha, \tau)$ (fonctions d'autocorrélations cycliques) suivants :

$$\begin{aligned} R_X(0, \tau) &= \frac{1}{2}r_1(\tau)\cos(2\pi f_1\tau) + \frac{1}{2}r_2(\tau)\cos(2\pi f_2\tau) \\ R_X(2f_1, \tau) &= \frac{1}{4}r_1(\tau)e^{j2\pi f_1\tau} \\ R_X(2f_2, \tau) &= \frac{1}{4}r_2(\tau)e^{j2\pi f_2\tau} \\ \overline{R_X(2f_1, \tau)} &= R_X(-2f_1, \tau) \\ \overline{R_X(2f_2, \tau)} &= R_X(-2f_2, \tau) \end{aligned}$$

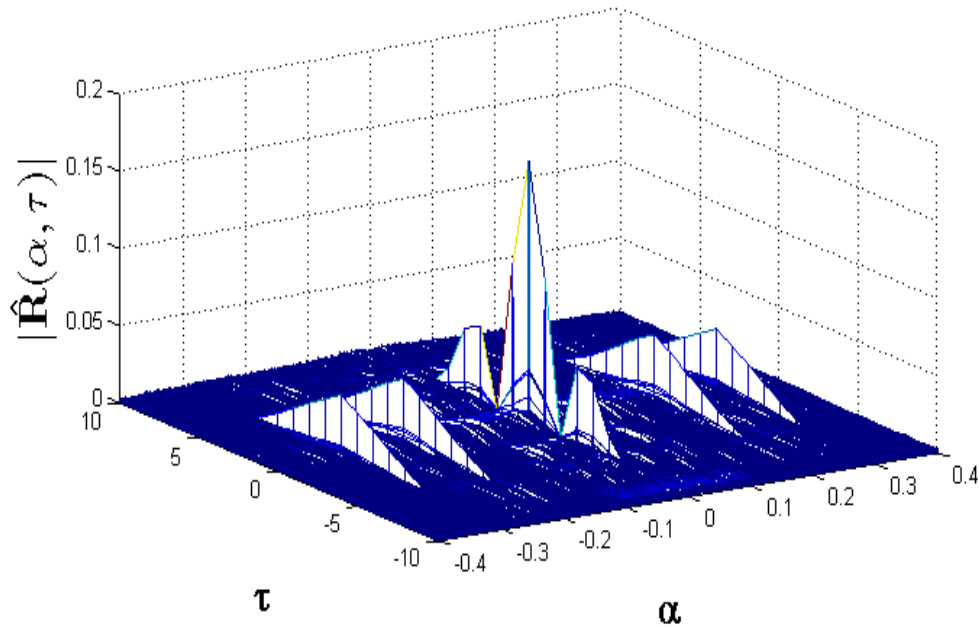


FIGURE 1.4 – Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$.

Les figures 1.4, 1.5 et 1.6 montrent respectivement la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ du signal composé de deux modulations d'amplitudes $X(t)$ en échelle linéaire, en échelle logarithmique et sa coupe à $\tau = 0$. Les fréquences des deux sinusoides utilisées sont respectivement $f_1 = 0.1Hz$ et $f_2 = 0.15Hz$.

Même si ce type de modèle a été très rarement utilisé en pratique. Il semble être un modèle très plausible et en adéquation avec la nature des données météorologiques (température, ensoleillement...). En effet, ce type de données est caractérisé par la présence de deux cycles principaux, en l'occurrence, le cycle journalier et le cycle annuel. Ces données pouvant être perturbées par des phénomènes aléatoires (changements de températures, passages de nuages ...). Le modèle proposé pourrait être utilisé pour caractériser de telles données météorologiques. En outre, on pourrait bien envisager l'exploitation de ce modèle dans le domaine de la mécanique, et ce, dans le cas de présence de défaut d'écaillage sur des roulements à billes.

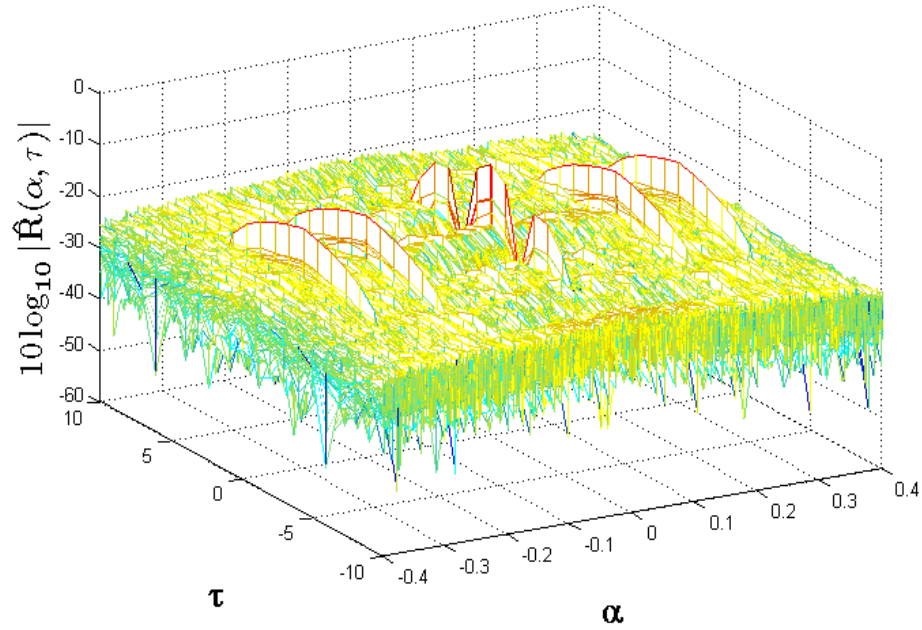


FIGURE 1.5 – Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$.

1.5.3 Modulation d'amplitude aléatoire par une fonction périodique

Le modèle qui suit est un nouveau modèle, il s'agit d'une modulation d'amplitude par un processus stationnaire couplé à une fonction périodique (ou presque périodique).

Soit $X(t) = U(t)f(t)$, avec $f(t) = \sin(2\pi f_1 t)\sin(2\pi f_2 t)$, avec $|f_1| \neq |f_2|$ et $\{U(t), t \in \mathbb{Z}\}$ un processus stationnaire avec une moyenne nulle et une fonction de covariance $r_U(\cdot)$.

L'étude théorique de ce modèle donne lieu à un ensemble de fréquences cycliques \mathcal{A}_α qui se met sous la forme suivante :

$$\mathcal{A}_\alpha = \{0, -2f_1, 2f_1, -2f_2, 2f_2, -2(f_1 - f_2), 2(f_1 - f_2), -2(f_1 + f_2), 2(f_1 + f_2)\}$$

Avec les valeurs des coefficients de Fourier non nulles $R_X(\alpha, \tau)$ (fonctions d'autocorrélations cycliques) suivants :

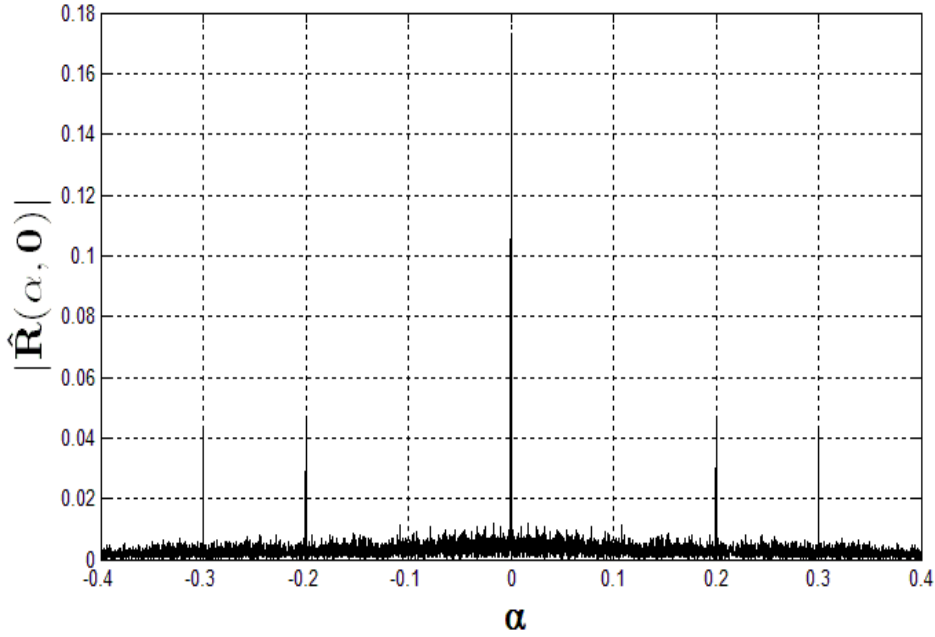


FIGURE 1.6 – Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t)$.

$$\begin{aligned}
 R_X(0, \tau) &= \frac{1}{8}(\cos((2\pi f_1 + 2\pi f_2)\tau) \\
 &\quad + \cos((2\pi f_1 - 2\pi f_2)\tau))r_U(\tau) \\
 &= \frac{1}{4}\cos(2\pi f_1 \tau)\cos(2\pi f_2 \tau)r_U(\tau) \\
 R_X(2f_1, \tau) = \overline{R_X(-2f_1, \tau)} &= \frac{-1}{16}(e^{j2\pi(f_1+f_2)\tau} + e^{j2\pi(f_1-f_2)\tau})r_U(\tau) \\
 &= \frac{-1}{8}e^{j2\pi f_1 \tau}\cos(2\pi f_2 \tau)r_U(\tau) \\
 R_X(2f_2, \tau) = \overline{R_X(-2f_2, \tau)} &= \frac{-1}{16}(e^{j2\pi(f_1+f_2)\tau} + e^{-j2\pi(f_1-f_2)\tau})r_U(\tau) \\
 &= \frac{-1}{8}e^{j2\pi f_2 \tau}\cos(2\pi f_1 \tau)r_U(\tau) \\
 R_X(2(f_1 + f_2), \tau) = \overline{R_X(-2(f_1 + f_2), \tau)} &= \frac{1}{16}e^{j2\pi(f_1+f_2)\tau}r_U(\tau) \\
 R_X(2(f_1 - f_2), \tau) = \overline{R_X(-2(f_1 - f_2), \tau)} &= \frac{1}{16}e^{j2\pi(f_1-f_2)\tau}r_U(\tau)
 \end{aligned}$$

Les figures 1.7, 1.8 et 1.9 montrent respectivement la fonction d'au-

to-corrélation cyclique $R_X(\alpha, \tau)$ du signal $X(t)$ en échelle linéaire, en échelle logarithmique et sa coupe à $\tau = 0$. Les fréquences des deux sinusoides utilisées sont respectivement $f_1 = 0.05Hz$ et $f_2 = 0.07Hz$. L'ensemble des fréquences cycliques théoriques correspondant \mathcal{A}_α s'écrit comme suit :

$$\mathcal{A}_\alpha = \{0, -0.1, 0.1, -0.14, 0.14, 0.04, -0.04, 0.24, -0.24\}$$

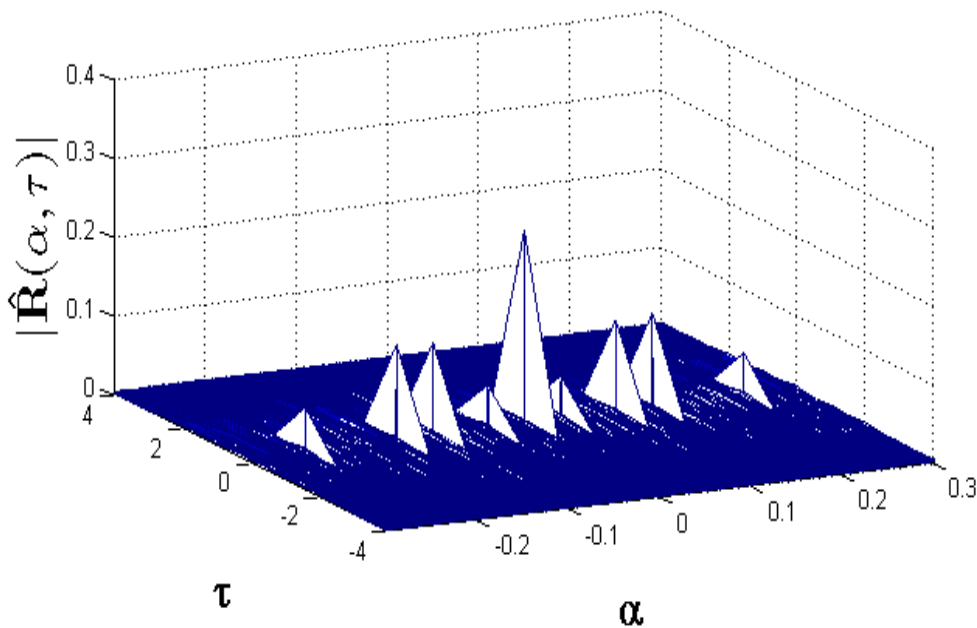


FIGURE 1.7 – Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)f(t)$.

Une possible application possible de ce modèle serait dans le domaine de la mécanique et des machines tournantes. En effet, en envisageant un système mécanique constitué d'engrenages et de roulements défectueux. Il serait très probable que ce modèle soit le plus plausible à ce genre de système mécanique. Les défauts de roulements étant généralement cyclostationnaires à l'ordre deux et modélisés par une simple modulation d'amplitude (le $U(t)\sin(2\pi f_1 t)$ de notre modèle), alors que les engrenages sont généralement déterministes ou cyclostationnaires à l'ordre un (le $\sin(2\pi f_2 t)$ de notre modèle). Le couplage entre ces deux phénomènes donnera lieu à un processus ayant les mêmes caractéristiques que le modèle étudié ci-dessus. Ce modèle reste aussi valable dans le cas d'un couplage entre un balourd et un défaut de roulement.

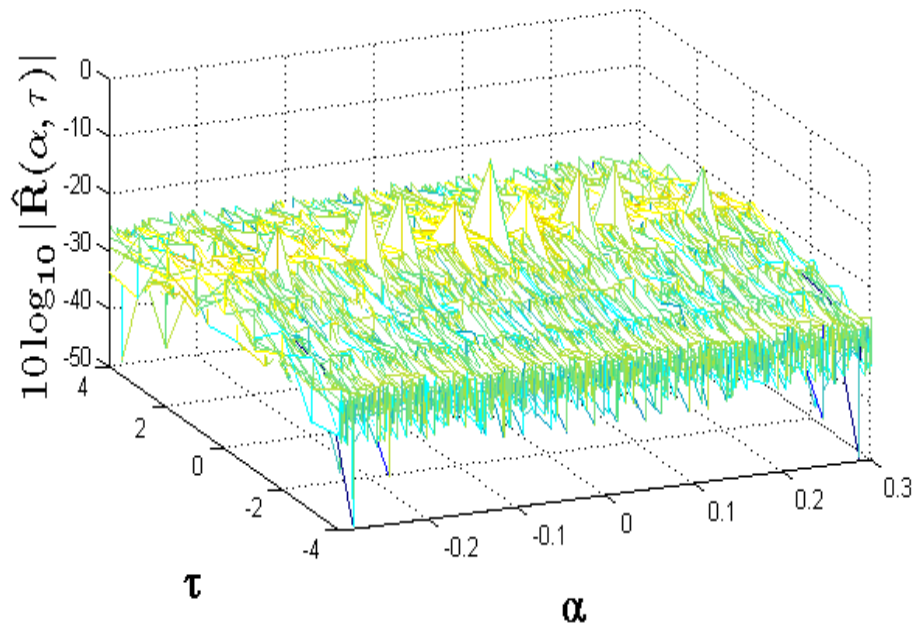


FIGURE 1.8 – Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)f(t)$.

1.5.4 Modulation d'amplitude aléatoire bi-composante avec dépendance

Soit le modèle suivant $X(t) = U(t)\cos(2\pi f_1 t) + U(t-1)\cos(2\pi f_2 t)$, avec $t \in \mathbb{Z}$ et $U(t)$ un processus stationnaire de moyenne nulle et de fonction de covariance $r_U(\cdot)$. À noter la dépendance entre $U(t)$ et $U(t-1)$.

Les calculs théoriques donnent l'ensemble de fréquences cycliques suivant :

$$\mathcal{A}_\alpha = \{-2f_2, -f_1 - f_2, f_1 - f_2, -2f_1, 0, 2f_1, f_2 - f_1, f_1 + f_2, 2f_2\}$$

Avec les valeurs des coefficients de Fourier non nulles $R_X(\alpha, \tau)$ correspondant aux fréquences cycliques suivantes :

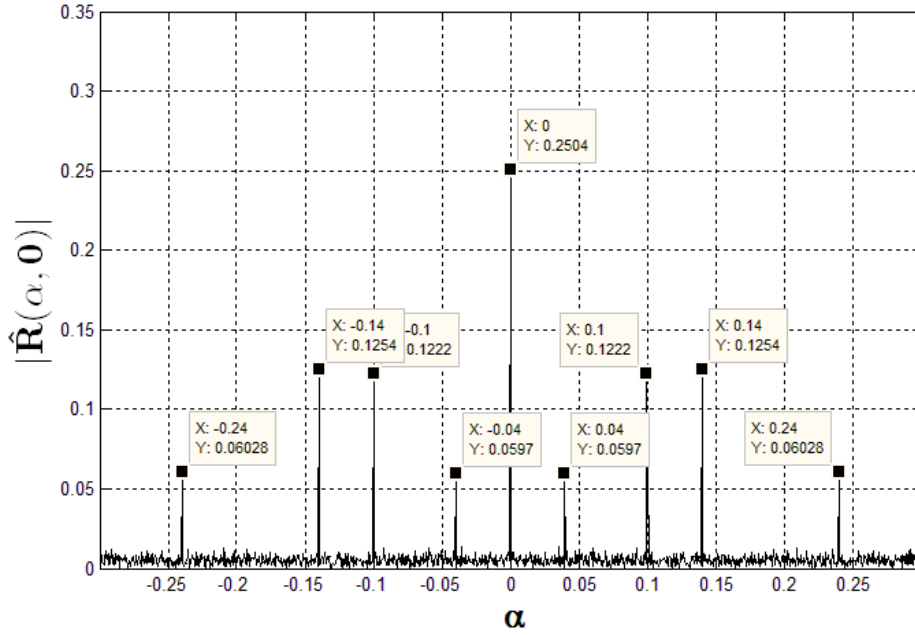


FIGURE 1.9 – Coupe à $\tau = 0$ de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)f(t)$.

$$\begin{aligned}
 R_X(0, \tau) &= \frac{1}{2}(\cos(2\pi f_1\tau) + \cos(2\pi f_2\tau))r_U(\tau) \\
 R_X(2f_1, \tau) = \overline{R_X(-2f_1, \tau)} &= \frac{1}{4}e^{j2\pi f_1\tau}r_U(\tau) \\
 R_X(f_1 - f_2, \tau) = \overline{R_X(-f_1 + f_2, \tau)} &= \frac{1}{4}(e^{-j2\pi f_1\tau}r_U(\tau + 1) + e^{-j2\pi f_2\tau}r_U(\tau - 1)) \\
 R_X(f_1 + f_2, \tau) = \overline{R_X(-f_1 - f_2, \tau)} &= \frac{1}{4}(e^{j2\pi f_1\tau}r_U(\tau + 1) + e^{j2\pi f_2\tau}r_U(\tau - 1)) \\
 R_X(2f_2, \tau) = \overline{R_X(-2f_2, \tau)} &= \frac{1}{4}e^{j2\pi f_2\tau}r_U(\tau)
 \end{aligned}$$

Les figures 1.10, 1.11 et 1.12 montrent respectivement la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ du signal $X(t)$ en échelle linéaire, en échelle logarithmique et sa projection sur l'axe des α . Les fréquences des deux sinusoides utilisées sont respectivement $f_1 = 0.05Hz$ et $f_2 = 0.07Hz$. L'ensemble des fréquences cycliques théoriques correspondant \mathcal{A}_α s'écrit comme suit :

$$\mathcal{A}_\alpha = \{-0.14, -0.12, -0.02, -0.1, 0, 0.1, 0.02, 0.12, 0.14\}$$

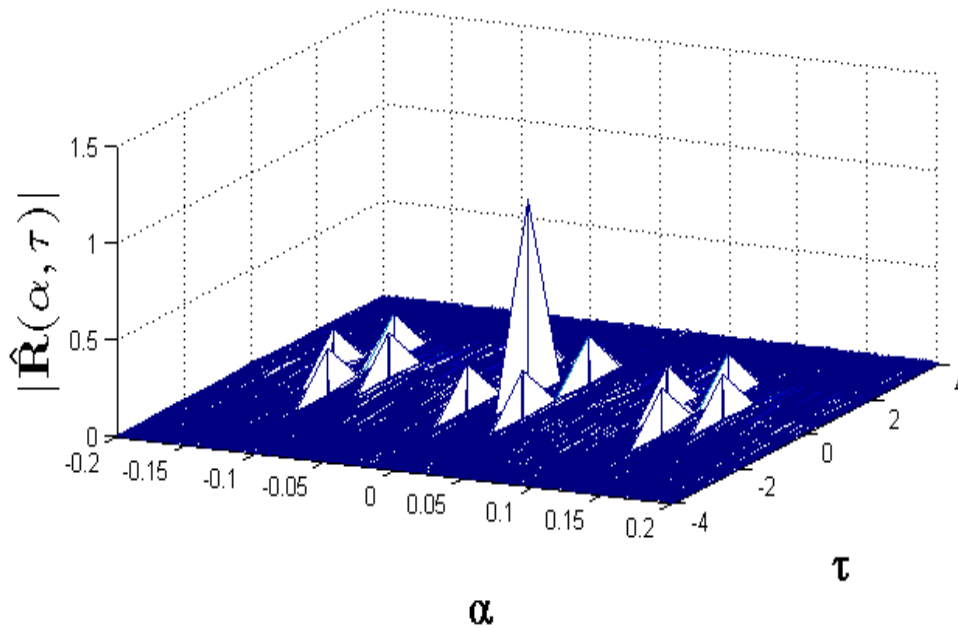


FIGURE 1.10 – Fonction d'autocorrélation cyclique en échelle linéaire du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$.

Ce modèle pourrait trouver quelques applications très intéressantes dans le domaine de l'économétrie ainsi que dans le domaine de la mécanique. En effet, il est souvent supposé que les composantes aléatoires qui viennent moduler d'autres composantes périodiques sont indépendantes. Dans les cas pratiques, cette hypothèse n'est pas forcément vérifiée, d'où la nécessité de prendre en considération la question de la dépendance. Par conséquent, ce modèle paraît un peu plus proche de la nature réelle de quelques données qu'on pourrait rencontrer dans différents domaines d'application.

1.5.5 Modulation numérique linéaire

La modulation numérique linéaire, communément appelée modulation d'amplitude d'un train d'impulsions (PAM) est un modèle de base dans la modulation numérique. Soit le signal $X(t)$ défini comme suit :

$$X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$$

Telle que $\{d(n), n = -\infty, \dots, \infty\}$ est une suite de symboles (complexes)

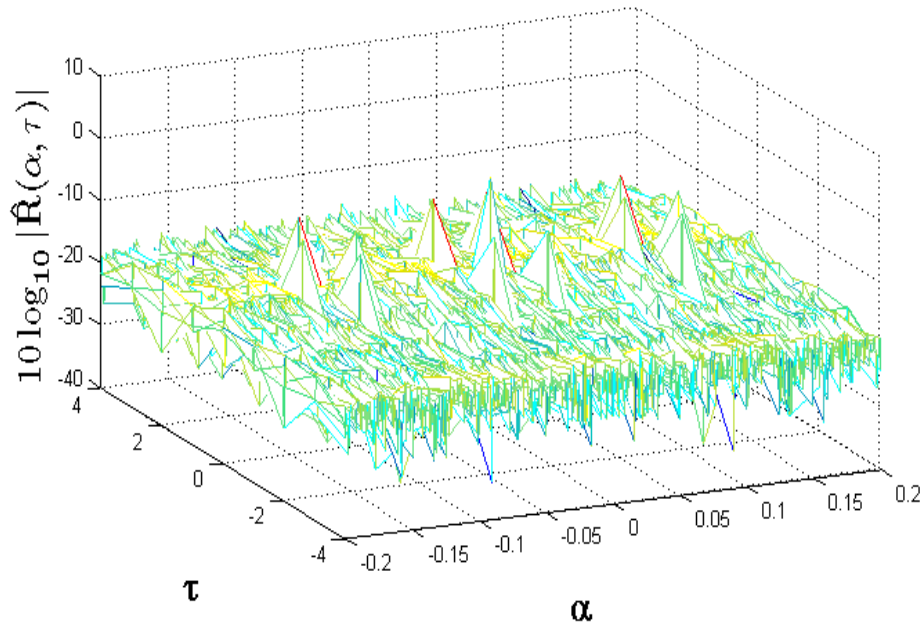


FIGURE 1.11 – Fonction d'autocorrélation cyclique en échelle logarithmique du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$.

aléatoires supposés indépendants et identiquement distribués i.i.d. Cela signifie que tous les symboles $d(n)$ sont supposés suivre la même loi de probabilité et deux symboles pris à des dates différentes devraient être indépendants.

Quant à $g(t)$, elle représente la fonction de mise en forme et T_s est la période symbole.

La fonction d'autocorrélation (conjuguée) d'un tel processus est donnée par :

$$R_X(t, \tau) = E\left\{ \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} d(m)d^*(n)g\left(t - mT_s - \frac{\tau}{2}\right)g^*\left(t - nT_s + \frac{\tau}{2}\right) \right\}$$

Sachant que les symboles sont i.i.d alors : $E\{d(m)d^*(n)\} = \sigma_d^2\delta(n-m)$ avec $\sigma_d^2 = E\{|d(n)|^2\}$. En prenant en compte cela, la fonction d'autocorrélation devient :

$$R_X(t, \tau) = \sum_{n=-\infty}^{+\infty} \sigma_d^2 g\left(t - nT_s - \frac{\tau}{2}\right)g^*\left(t - nT_s + \frac{\tau}{2}\right)$$

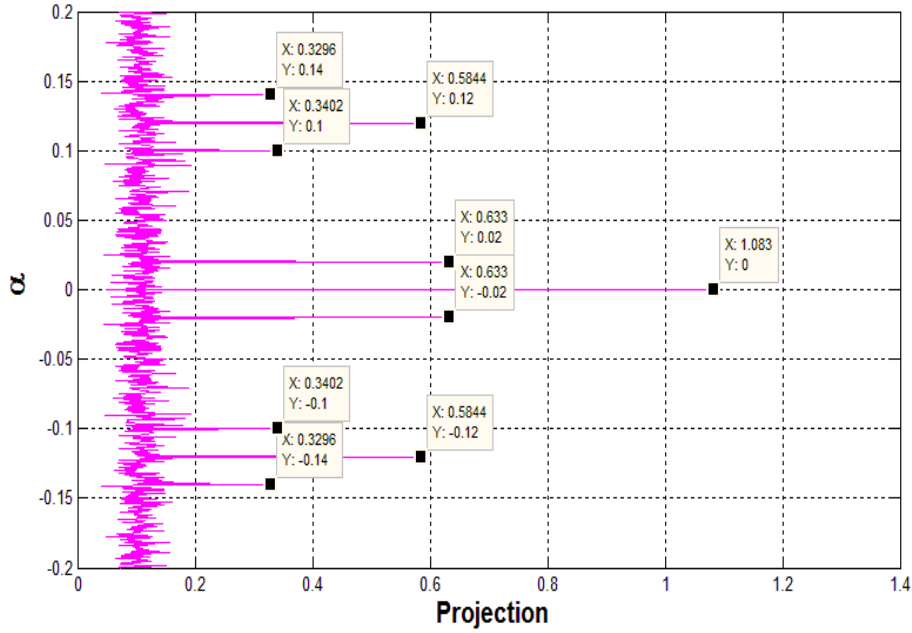


FIGURE 1.12 – Projection sur l'axe des α de la fonction d'autocorrélation cyclique du signal $X(t) = U(t)\cos(2\pi f_0 t) + U(t-1)\cos(2\pi f_1 t)$.

A partir de cette dernière équation, on peut facilement noter que la fonction d'autocorrélation (conjuguée) $R_X(t, \tau)$ de $X(t)$ est périodique, avec une période égale à T_s , i.e, égale à la période du symbole. Donc, ce résultat implique que le signal $X(t)$ est CS de fréquence cyclique fondamentale $\alpha = \frac{1}{T_s}$.

Par conséquent, les coefficients de Fourier de $R_X(t, \tau)$ représentant la fonction d'autocorrélation cyclique sont donnés comme suit :

$$R_X\left(\frac{k}{T_s}\right) = \frac{\sigma_d^2}{T_s} \int_{-\infty}^{+\infty} g\left(t - \frac{\tau}{2}\right) g^*\left(t + \frac{\tau}{2}\right) e^{-i2\pi \frac{k}{T_s} t} dt$$

Donc, les coefficients de Fourier de la fonction d'autocorrélation (conjuguée) sont non nuls pour l'ensemble des fréquences cycliques $\mathcal{A}_\alpha = \left\{\frac{k}{T_s}, k \in \mathbb{Z}\right\}$.

Les figures 1.13, 1.14 et 1.15 montrent respectivement la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ d'un signal PAM simulé en échelle linéaire, en échelle logarithmique et sa projection sur l'axe des α . On peut remarquer de façon très claire que la fonction d'autocorrélation est périodique, avec une période T_s égale à la période du symbole. Par conséquent, le signal est CS avec une fréquence cyclique fondamentale $\alpha = \frac{1}{T_s}$.

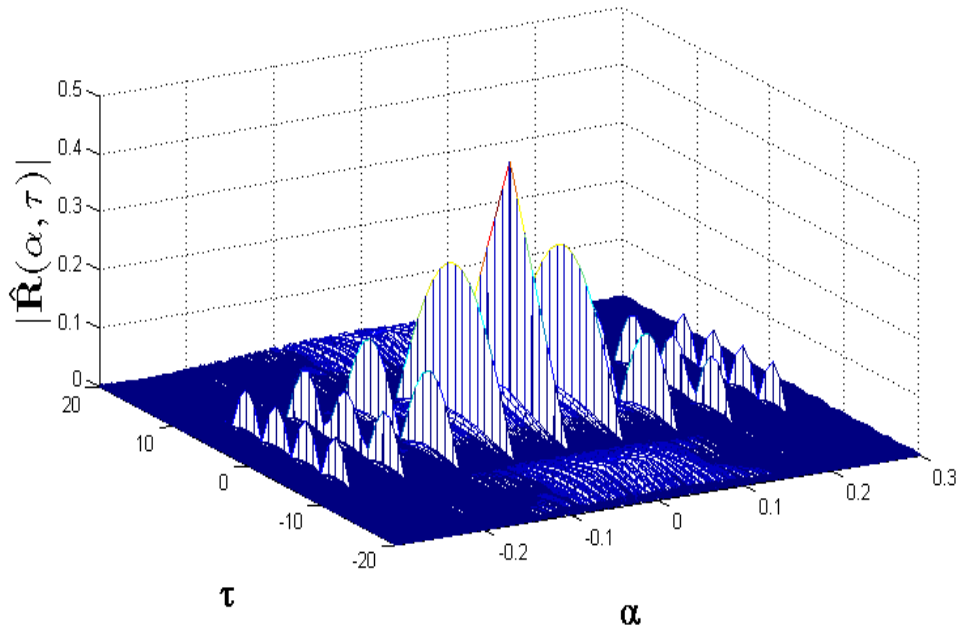


FIGURE 1.13 – Fonction d'autocorrélation cyclique en échelle linéaire du signal modulé en amplitude $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$.

1.6 De la cyclostationnarité à la cyclostationnarité généralisée

Dans différents domaines, la modélisation cyclostationnaire semble être appropriée pour décrire les processus sous étude. En effet, un processus cyclostationnaire peut être défini de façon schématisée comme étant le résultat d'un couplage entre un processus aléatoire stationnaire avec un phénomène périodique. Cependant, dans certains cas pratiques, cette notion de périodicité n'est pas accessible, et nous avons affaire à des phénomènes non périodiques et non cycliques. Donc, cela a conduit certains chercheurs à pousser leurs travaux de recherche et à créer un cadre plus approprié à l'étude de tels phénomènes. Motivé par des cas pratiques liés aux télécommunications, Napolitano [115] a très récemment introduit la notion de *cyclostationnarité généralisée*.

Les processus cyclostationnaires montrent une fonction d'autocorrélation périodique dans le temps. Les coefficients de Fourier d'une fonction de corrélation d'un tel processus dépendent du retard τ , contrairement aux fréquences cycliques, qui elles, ne dépendent pas de ce paramètre de retard. Quant aux processus cyclostationnaires généralisés, ils ont comme caractéristiques des

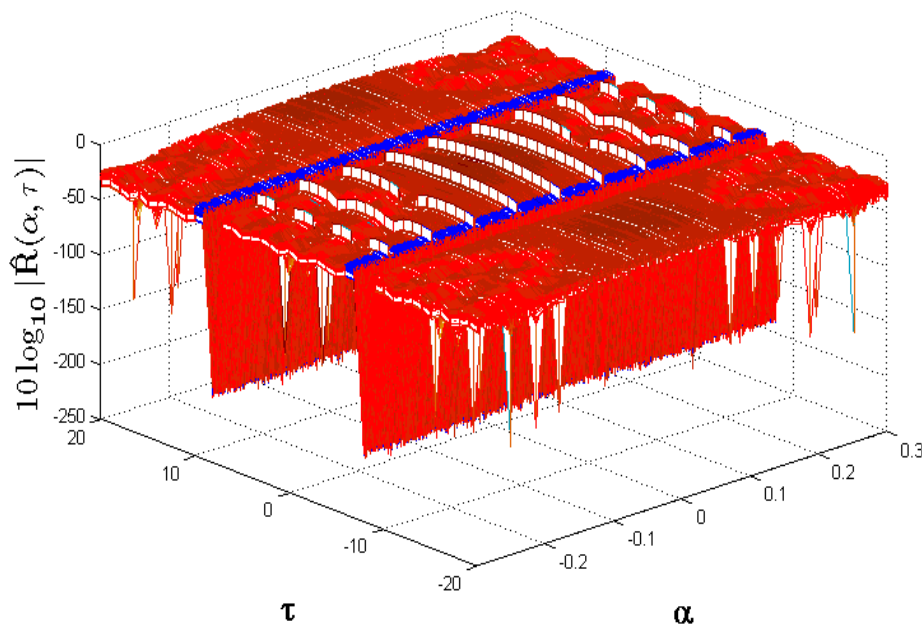


FIGURE 1.14 – Fonction d'autocorrélation cyclique en échelle logarithmique du signal modulé en amplitude $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$.

coefficients de Fourier dépendants du paramètre τ , de plus, les fréquences cycliques correspondant dépendent eux aussi de ce paramètre. Ainsi, les processus cyclostationnaires pourraient être aperçus comme des cas particuliers des processus cyclostationnaires généralisés quand l'ensemble des fréquences cycliques est indépendant du retard τ .

1.6.1 Définition d'un processus cyclostationnaire généralisé

Un processus stochastique $X(t)$ est dit cyclostationnaire généralisé si et seulement si sa fonction d'autocorrélation est périodique et variante dans le temps. De plus les coefficients de Fourier ainsi que les fréquences cycliques correspondant dépendent du paramètre de retard τ du processus. Ainsi, pour chaque valeur de τ , la fonction d'autocorrélation $R_X(t, \tau)$ est représentée comme suit :

$$R_X(t, \tau) = \sum_{\alpha \in \mathcal{A}_{\alpha\tau}} R_X(\alpha, \tau) e^{j2\pi\alpha t}$$

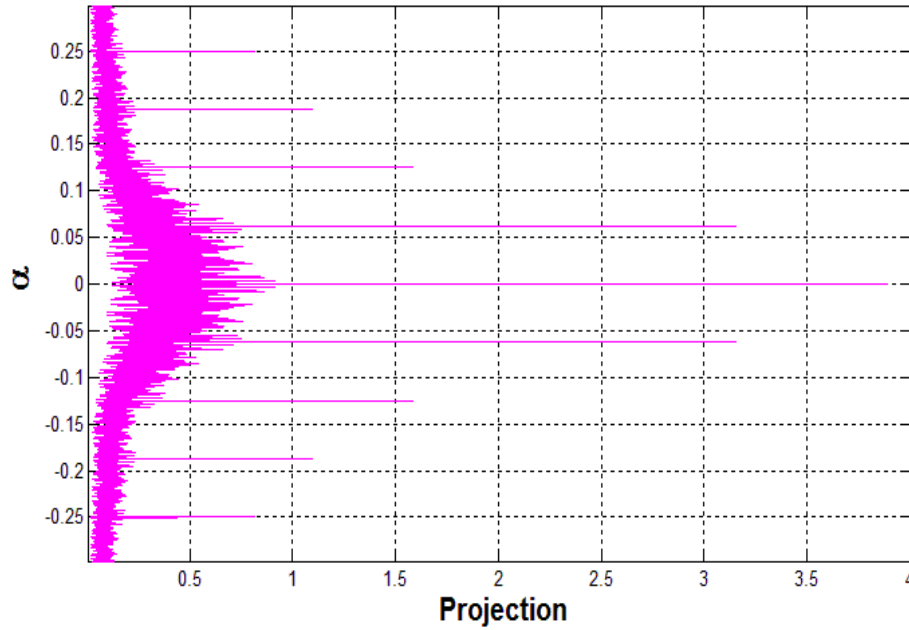


FIGURE 1.15 – Projection sur l’axe des α de la fonction d’autocorrélation cyclique du signal $X(t) = \sum_{n=-\infty}^{+\infty} d(n)g(t - nT_s)$.

Avec $\mathcal{A}_{\alpha\tau}$ un ensemble de fréquences dépendant de τ et est défini comme suit :

$$\begin{aligned} \mathcal{A}_{\alpha\tau} &= \{\alpha \in \mathbb{R} : R_X(\alpha, \tau) \neq 0\} \\ &= \bigcup_{n \in \mathbb{I}} \{\alpha \in \mathbb{R} : \alpha = \alpha_n(\tau)\} \end{aligned}$$

Tels que $\alpha_n(\tau)$ est une fréquence cyclique dépendant de τ et \mathbb{I} est un ensemble fini.

Ainsi, nous pouvons réécrire la fonction d’autocorrélation $R_X(t, \tau)$ d’un processus cyclostationnaire généralisé de la façon suivante :

$$R_X(t, \tau) = \sum_{n \in \mathbb{I}} R_X(\alpha_n(\tau), \tau) e^{j2\pi\alpha_n(\tau)t}$$

Avec $R_X(\alpha_n(\tau), \tau)$ la fonction d’autocorrélation généralisée du processus et qui est donnée par :

$$R_X(\alpha_n(\tau), \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t, \tau) e^{-j2\pi\alpha_n(\tau)t} dt$$

Cette expression montre bien la dépendance des coefficients de Fourier ainsi que des fréquences cycliques de la fonction d'autocorrélation au paramètre de retard τ .

1.6.2 Cas cyclostationnaire généralisé

Dans différents domaines d'application, et spécialement dans la mécanique et la surveillance des machines tournantes, la modélisation cyclostationnaire de quelques phénomènes peut ne pas être suffisante et inappropriée. En effet, dans de nombreux cas, les machines fonctionnent dans des conditions non-stationnaires ou variables dans le temps. Nous pouvons citer le cas de machines tournantes avec charge variable ou avec une vitesse variable dans le temps. Dans ce type de cas, et contrairement au couplage entre un phénomène aléatoire stationnaire (un défaut dans un élément par exemple) et un phénomène périodique déterministe (le cycle de rotation par exemple) qui donne lieu à un phénomène cyclostationnaire, le couplage se fera entre la composante aléatoire et une composante non périodique et non-stationnaire (vitesse de rotation ou charge variable).

Dans ce qui suit, nous proposons l'étude d'un tel modèle. Ce modèle pourrait éventuellement être exploité dans différents cas pratiques, nous citons l'exemple d'un système mécanique comprenant un roulement défectueux avec un arbre de rotation. Le roulement défectueux engendrera une composante aléatoire, supposée stationnaire ou cyclostationnaire et l'arbre tournera à une vitesse de rotation variable. La vitesse de rotation de l'arbre (la fréquence) est supposée suivre une fonction linéaire et strictement croissante (montée de régime).

1.6.3 Étude de modèle à fréquence cyclique variable

Soit $X(t)$ une réalisation d'un processus aléatoire donné comme suit :

$$X(t) = U(t) \cos(2\pi f_r(t)t)$$

tel que $f_r(t) = f_0 + f_1 t$ ce qui signifie une fréquence qui varie linéairement et $U(t)$ un signal aléatoire dont nous étudierons les différents cas de figures par la suite.

1.6. De la cyclostationnarité à la cyclostationnarité généralisée 43

En utilisant la formule d'Euler, le signal $X(t)$ peut s'écrire sous la forme suivant :

$$X(t) = U(t) \frac{1}{2} [e^{j2\pi(f_0 t + f_1 t^2)} + e^{-j2\pi(f_0 t + f_1 t^2)}]$$

A partir de là, on calcule la fonction d'autocorrélation de $X(t)$ qui est donné comme suit :

$$R_X(t, \tau) = E\{X(t + \tau)X(t)\}$$

Après calculs et simplifications, nous obtenons le résultat suivant :

$$\begin{aligned} R_X(t, \tau) &= \frac{1}{4} E\{U(t + \tau)U(t)\} [e^{j2\pi(f_0 \tau + f_1 \tau^2)} e^{j2\pi(2f_0)t} e^{j2\pi(2f_1)t^2} \\ &+ e^{j2\pi(f_0 \tau + f_1 \tau^2)} e^{j2\pi(2f_1)t} \\ &+ e^{-j2\pi(f_0 \tau + f_1 \tau^2)} e^{-j2\pi(2f_1)t} \\ &+ e^{-j2\pi(f_0 \tau + f_1 \tau^2)} e^{-j2\pi(2f_0)t} e^{-j2\pi(2f_1)t^2}] \end{aligned}$$

Et la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ est donnée par :

$$R_X(\alpha, \tau) = \langle R_X(t, \tau) e^{-j2\pi\alpha t} \rangle_t$$

Dans ce qui suit, nous étudions les différents cas de figures de la nature du processus $X(t)$.

Cas numéro un :

Dans un premier temps, nous étudions le cas où le signal $U(t)$ est stationnaire (au sens large). Si $U(t)$ est stationnaire alors :

$$E\{U(t + \tau)U(t)\} = r_U(\tau).$$

Alors nous avons les deux sous-cas suivants en fonction de la fréquence f_1 :

- Si $f_1 = 0$ alors le signal $X(t)$ est une simple modulation d'amplitude, donc cyclostationnaire avec un ensemble de fréquences cycliques $\mathcal{A}_\alpha = \{-2f_0, 0, 2f_0\}$.

- Si $f_1 \neq 0$ alors la fonction d'autocorrélation $R_X(t, \tau)$ du signal $X(t)$ est cyclostationnaire généralisé avec un ensemble de fréquences cycliques qui dépend du paramètre du retard τ . L'ensemble des fréquences cycliques est donné par :

$$\mathcal{A}_{\alpha\tau} = \{-2f_1\tau, 2f_1\tau\}$$

Il faut aussi noter que quand $f_1 \neq 0$, les caractéristiques cycliques autour de $\pm f_0$ disparaissent.

La figure 1.16 montre les relations théoriques obtenues entre les fréquences cycliques et le paramètre de retard τ dans le cas $f_1 \neq 0$

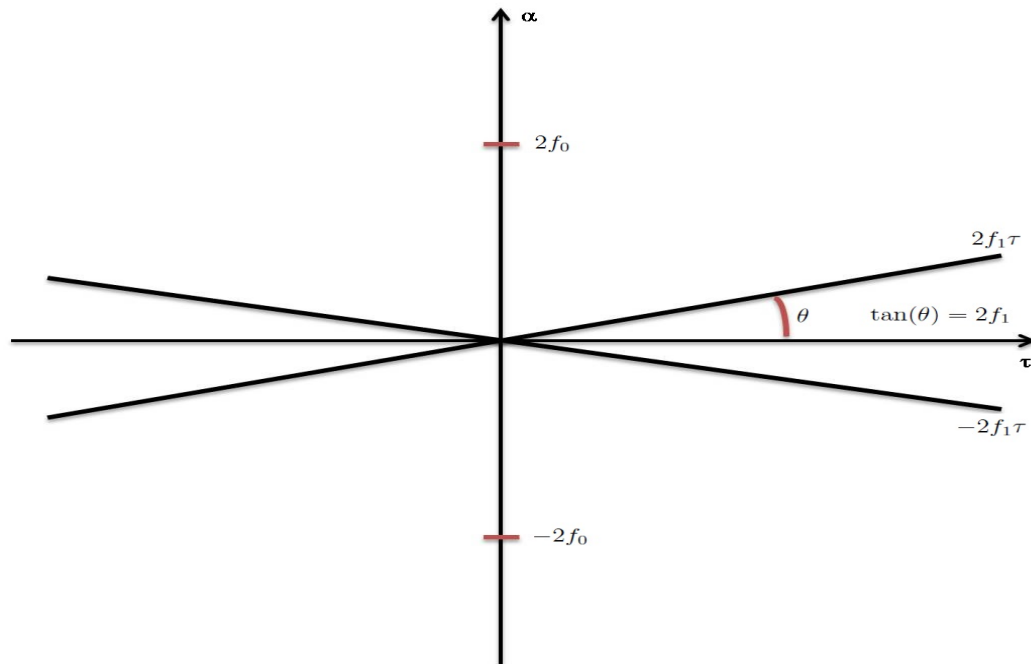


FIGURE 1.16 – Les relations théoriques entre les fréquences cycliques et le paramètre de retard τ pour le modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire

Ainsi, dans le cas où $U(t)$ est stationnaire, nous obtenons un processus cyclostationnaire modulé en amplitude si la fréquence $f_1 = 0$. Sinon, et dans le cas où $f_1 \neq 0$, nous obtenons un processus cyclostationnaire généralisé avec des fréquences cycliques dépendant du retard τ et de la fréquence f_1 , tel que $\mathcal{A}_{\alpha\tau} = \{-2f_1\tau, 2f_1\tau\}$.

Cas numéro deux :

Le deuxième cas de figure est celui où le signal $U(t)$ est cyclostationnaire. Si $U(t)$ est cyclostationnaire alors :

$$E\{U(t + \tau)U(t)\} = \sum_k R_U(\alpha_k, \tau)e^{j2\pi\alpha_k t}.$$

Avec α_k les fréquences cycliques du processus cyclostationnaire $U(t)$.

Comme pour le premier cas de figure, nous avons les deux sous-cas suivants en fonction de la fréquence f_1 :

- Si $f_1 = 0$ alors le signal $X(t)$ est cyclostationnaire avec un ensemble de fréquences cycliques $\mathcal{A}_\alpha = \{\alpha_k, \alpha_k \pm 2f_0\}$.
- Si $f_1 \neq 0$ alors la fonction d'autocorrélation $R_X(t, \tau)$ du signal $X(t)$ est cyclostationnaire généralisé avec un ensemble de fréquences cycliques qui dépend du paramètre du retard τ . L'ensemble des fréquences cycliques est donné par :

$$\mathcal{A}_{\alpha\tau} = \{\alpha_k \pm 2f_1\tau\}$$

La figure 1.17 montre les relations théoriques obtenues entre les fréquences cycliques et le paramètre de retard τ dans le cas $f_1 \neq 0$ et $U(t)$ un processus cyclostationnaire.

Ainsi, dans le cas où $U(t)$ est cyclostationnaire, nous obtenons un processus cyclostationnaire si $f_1 = 0$. Sinon, et dans le cas où $f_1 \neq 0$, nous obtenons un processus cyclostationnaire généralisé avec des fréquences cycliques dépendant du retard τ et de la fréquence f_1 , tel que $\mathcal{A}_{\alpha\tau} = \{\alpha_k \pm 2f_1\tau\}$. Avec α_k les fréquences cycliques de $U(t)$.

1.6.4 Étude de simulation

Dans l'optique de vérifier les résultats théoriques obtenus dans la partie précédente, nous présentons une étude de simulation pour deux cas de figures possibles : (i) $U(t)$ stationnaire ; (ii) $U(t)$ cyclostationnaire. Motivés par l'étude des cas cyclostationnaires généralisés, nous prendrons dans les deux simulations une valeur de fréquence f_1 non nulle.

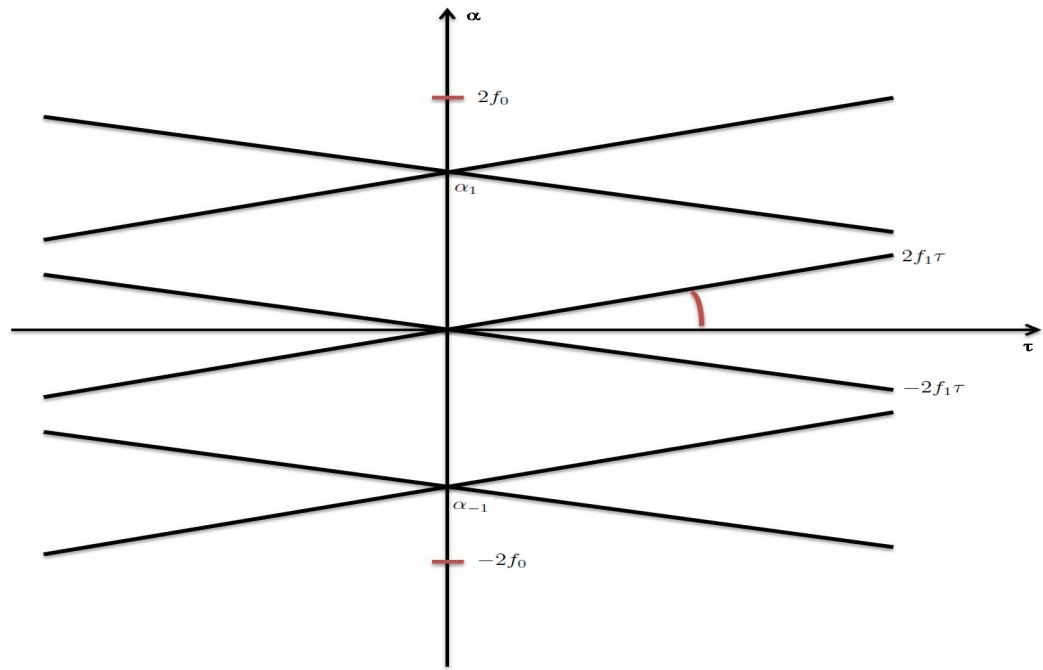


FIGURE 1.17 – Les relations théoriques entre les fréquences cycliques et le paramètre de retard τ pour le modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire

Première simulation

Dans cette première simulation, nous nous intéressons au cas où le processus $U(t)$ est stationnaire. Pour cela, nous rappelons d'abord le modèle de l'étude :

$$X(t) = U(t)\cos(2\pi f_r(t)t)$$

tel que $f_r(t) = f_0 + f_1 t$. Nous prenons $U(t)$ un signal stationnaire (bruit blanc), et comme valeurs de f_0 et f_1 , 0.05Hz et 0.005Hz respectivement.

Les fonctions d'autocorrélations cycliques du signal $X(t)$ sont représentées en échelle linéaire et en échelle logarithmique, respectivement dans les figures 1.18 et 1.19

Les résultats de simulation confirment clairement les résultats théoriques obtenus auparavant. Ainsi, nous remarquons des fréquences cycliques dépen-

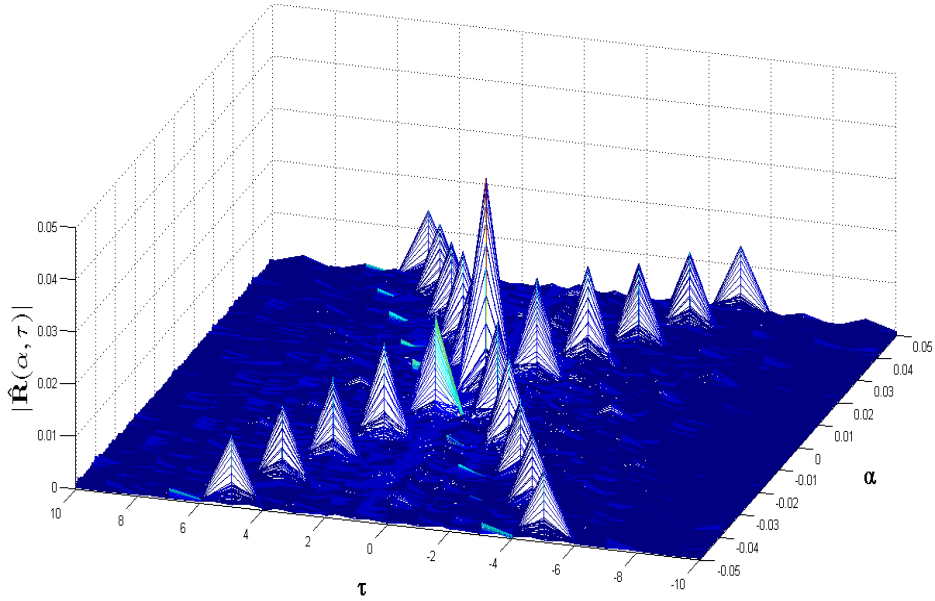


FIGURE 1.18 – Fonction d'autocorrélation cyclique en échelle linéaire du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire

dant du retard τ et de la fréquence f_1 , tel que $\mathcal{A}_{\alpha\tau} = \{-2f_1\tau, 2f_1\tau\}$.

Deuxième simulation

Dans cette seconde simulation, nous nous intéressons au cas où le processus $U(t)$ est cyclostationnaire. Donc soit $X(t)$:

$$X(t) = U(t)\cos(2\pi f_r(t)t)$$

Avec

$$U(t) = a(t)\cos(2\pi f_U t)$$

tel que $f_r(t) = f_0 + f_1 t$ et $a(t)$ un processus stationnaire. Ainsi, $U(t)$ est un signal modulé en amplitude, par conséquent cyclostationnaire avec un ensemble de fréquences cycliques $\alpha_k = \{-2f_U, 0, 2f_U\}$.

Les valeurs des fréquences f_0 , f_1 et f_U sont égales respectivement à 0.05Hz, 0.001Hz et 0.03Hz.

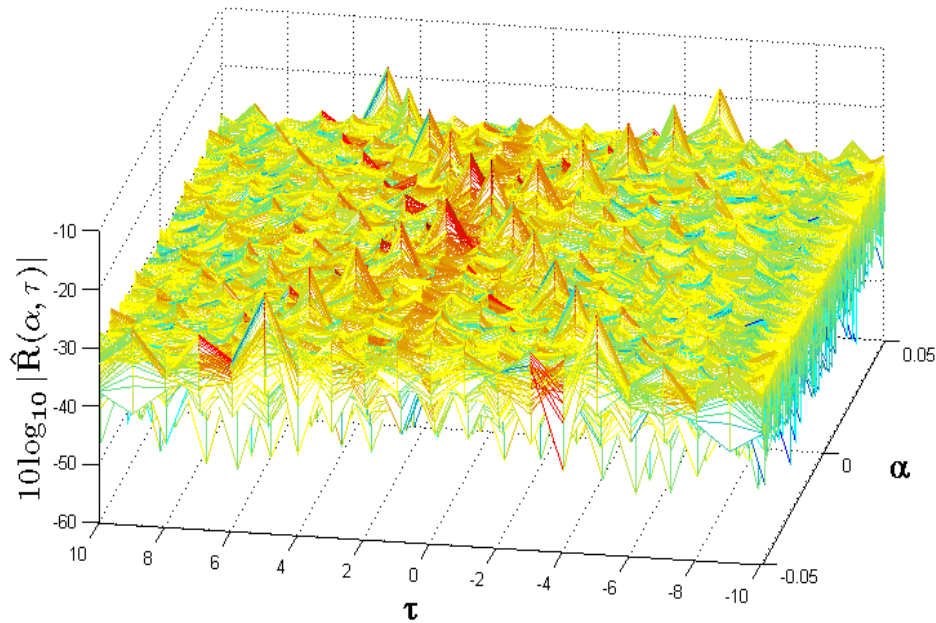


FIGURE 1.19 – Fonction d'autocorrélation cyclique en échelle logarithmique du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ stationnaire

Les fonctions d'autocorrélations cycliques du signal $X(t)$ sont représentées en échelle linéaire et en échelle logarithmique, respectivement dans les figures 1.20 et 1.21.

Les résultats de simulation affirment clairement les résultats théoriques obtenus dans la partie précédente. Ainsi, nous remarquons des fréquences cycliques dépendant du retard τ et de la fréquence f_1 , tel que $\mathcal{A}_{\alpha\tau} = \{\alpha_k \pm 2f_1\tau\}$. Avec α_k les fréquences cycliques de $U(t)$ qui sont égales à $\{-2f_U, 0, 2f_U\}$.

1.7 Conclusion

Dans ce chapitre, nous avons présenté différentes définitions ainsi que diverses propriétés liés aux processus stochastiques. Par la suite, nous nous sommes focalisés sur les modèles statistiques cyclostationnaires. En effet, les processus cyclostationnaires sont schématiquement caractérisés par des propriétés statistiques périodiques dans le temps contrairement aux processus stationnaires dont les propriétés statistiques sont indépendantes du temps. Ensuite, nous avons proposé puis étudié différents modèles cyclostationnaires dont certains sont innovants. Les modèles étudiés ont aussi été liés à de pos-

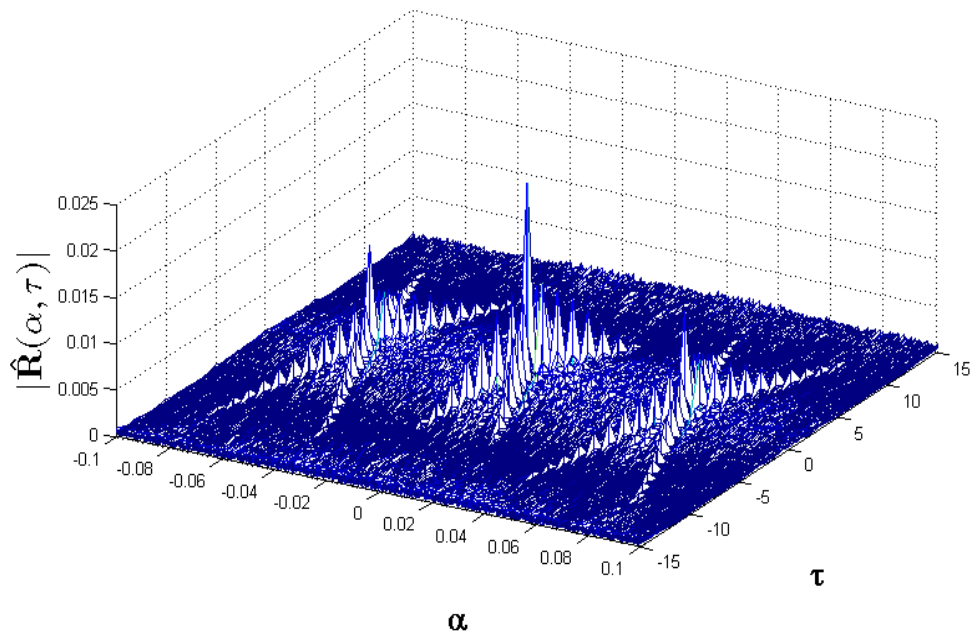


FIGURE 1.20 – Fonction d'autocorrélation cyclique en échelle linéaire du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire

sibles applications dans différents domaines. Enfin, nous avons étendu la notion de cyclostationnarité généralisée à l'étude du cas de fréquence cyclique variable dans le temps. Une étude théorique d'un nouveau modèle cyclostationnaire généralisé a été présentée et des travaux de simulations ont été réalisés. Ces simulations ont confirmé les résultats théoriques obtenus ce qui ouvre des perspectives très prometteuses quant à leurs applications dans les domaines de la mécanique et de la biomécanique.

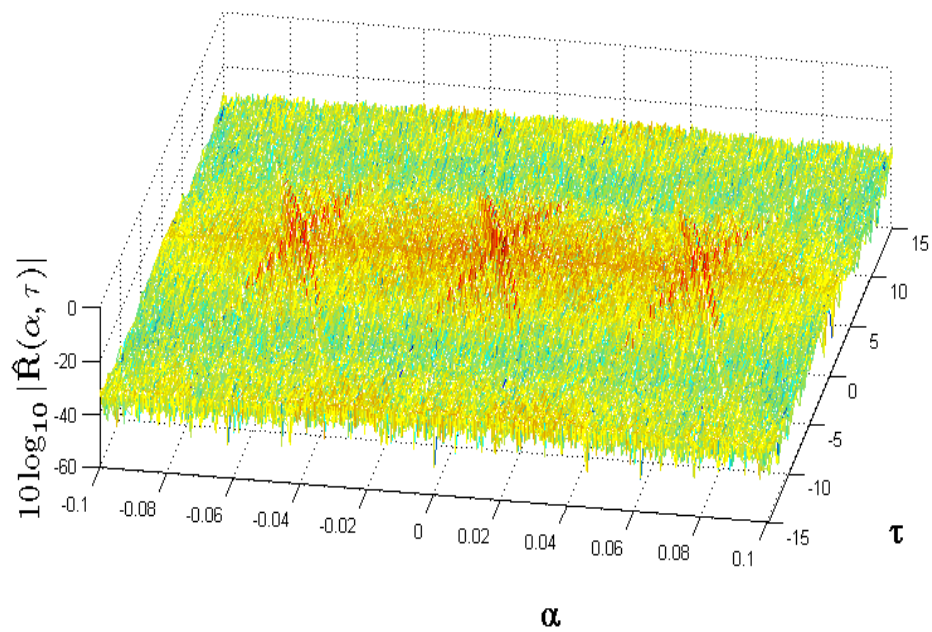


FIGURE 1.21 – Fonction d'autocorrélation cyclique en échelle logarithmique du modèle $X(t) = U(t)\cos(2\pi f_r(t)t)$. $U(t)$ cyclostationnaire

Le Bootstrap

Sommaire

2.1	Introduction	53
2.2	Le bootstrap et ses applications	54
2.2.1	Définition du bootstrap	54
2.2.2	Avantages du bootstrap	54
2.2.3	Cas d'utilisation du bootstrap	55
2.2.4	Les domaines d'application du bootstrap	56
2.3	Le principe du bootstrap	57
2.4	Résultats théoriques du bootstrap pour la moyenne	59
2.4.1	La consistance du bootstrap pour la moyenne	59
2.4.2	La consistance du bootstrap pour la moyenne standardisée	61
2.4.3	La consistance du bootstrap pour la moyenne studentisée	61
2.5	Exemples d'estimations par le bootstrap non paramétrique	62
2.5.1	Estimation du biais	62
2.5.2	Estimation de la variance	64
2.5.3	Intervalles de confiance pour la moyenne	65
2.6	Bootstrap paramétrique	66
2.6.1	Intervalles de confiance pour la moyenne via le bootstrap paramétrique	67
2.7	Les limitations du bootstrap	69
2.7.1	Taille d'échantillon très petite	69
2.7.2	Séries m-dépendantes	70
2.7.3	Distribution avec des moments infinis	70
2.7.4	Le sondage	71
2.7.5	Dépendance à long terme	72
2.8	Conclusion	72

2.1 Introduction

Le traitement statistique du signal est un domaine dans lequel des outils mathématiques et statistiques sont utilisés afin de résoudre des problèmes d'estimation de grandeurs physiques. En effet, de nos jours, il est très rare de rencontrer des applications de traitements de signal où des outils statistiques ne sont pas utilisés. Dans la plupart de ces applications, une analyse statistique rigoureuse des estimateurs de paramètres d'intérêt est fondamentalement nécessaire. Un exemple serait de trouver l'exactitude et la précision d'un estimateur d'un paramètre donné. Dans la pratique, un estimateur est généralement calculé sur la base d'un nombre fini de mesures, communément appelé échantillon. Un autre exemple est celui de déterminer les limites supérieures et inférieures d'un estimateur d'un paramètre d'intérêt à l'intérieur desquelles la valeur réelle du paramètre se trouve, et ce, pour un niveau de confiance donné. Ces limites constituent les intervalles de confiance du paramètre estimé.

Ainsi, il est très important de trouver des réponses à deux questions très importantes. Étant donné un nombre de mesures et un paramètre d'intérêt :

- Quel estimateur doit être utilisé ?
- Ayant décidé d'utiliser un estimateur, quel est sa précision et son exactitude ?

La première solution pour répondre à ces questions serait l'utilisation de la méthode du maximum de vraisemblance ou de la méthode des moindres carrés. En effet, après le calcul de l'estimateur du paramètre, la précision pourrait être mesurée par la variance de l'estimateur ou par des intervalles de confiance de ce dernier. Cependant, dans la majorité des cas, les techniques disponibles pour la caractérisation statistique des estimateurs de paramètres supposent que la taille de l'échantillon observé soit suffisamment grande, de telle sorte à pouvoir appliquer des résultats asymptotiques. Le théorème central limite et l'hypothèse de gaussiennité du processus sous études sont des exemples de telles approches. Cependant, dans de nombreux problèmes de traitement du signal, ces méthodes qui se basent sur l'observation d'un très grand nombre d'échantillons ne sont pas applicables, et ce, à cause de contraintes de temps ou de matériels ou bien à cause du caractère non-stationnaire des signaux. Une alternative pour solutionner ces problèmes est celle de l'application des méthodes dites de *bootstrap*.

Dans ce qui suit, nous donnons la définition du bootstrap, ses applications dans différents domaines ainsi que son principe.

2.2 Le bootstrap et ses applications

Le bootstrap a été tout d'abord introduit par Efron [55, 56, 57, 58] principalement dans l'objectif de calculer des intervalles de confiance d'un paramètre dans le cas où les méthodes standard ne sont plus applicables [59]. Un exemple de cela est le cas où le nombre d'observations serait tellement petit que l'application de résultats asymptotiques ne serait acceptable. D'autres travaux ont ensuite suivis, nous citons ceux de Edgington [52, 53, 54], Manly [109, 110] et [81].

Depuis son apparition, le bootstrap a eu de nombreuses applications et a été utilisé pour résoudre quelques problématiques qui sont très compliquées à résoudre analytiquement [83, 60].

Dans ce qui suit, nous donnons la définition générale du bootstrap, ses avantages ainsi que les cas dans lesquels il pourrait être utilisé.

2.2.1 Définition du bootstrap

Schématiquement, le bootstrap est une méthode qui consiste à réaliser par le biais d'un ordinateur ce que l'expérimentateur pourrait faire en pratique si cela était possible, c'est à dire répéter l'expérimentation un grand nombre de fois. Ainsi, avec le bootstrap, de nouvelles expérimentations ne sont plus nécessaires. En effet, les données disponibles sont réutilisées. Pour être plus précis, les observations originales sont réaffectées de façon aléatoire, puis l'estimateur est recalculé. Ces réaffectations et recalculs d'estimateurs sont réalisés un grand nombre de fois et sont considérés comme des répétitions de l'expérimentation. À première vue, cela ressemble à une simulation Monte Carlo, mais ce n'est pas le cas. En effet, le plus grand avantage du bootstrap comparé avec la méthode de Monte Carlo réside dans le fait que le bootstrap ne nécessite pas la répétition de l'expérimentation, contrairement aux simulations Monte Carlo.

D'un autre point de vue, l'idée principale du bootstrap est de simuler autant que possible la probabilité du *monde réel* en substituant les inconnus par des estimés provenant des données observées. Ainsi, à partir des simulations du *monde bootstrap*, des entités d'intérêt inconnues dans le *monde réel* peuvent être estimées comme il est montré dans la figure 2.1.

2.2.2 Avantages du bootstrap

L'avantage principal du bootstrap réside dans le fait qu'aucune hypothèse de disponibilité d'un grand nombre d'échantillons n'est émise. En effet, le

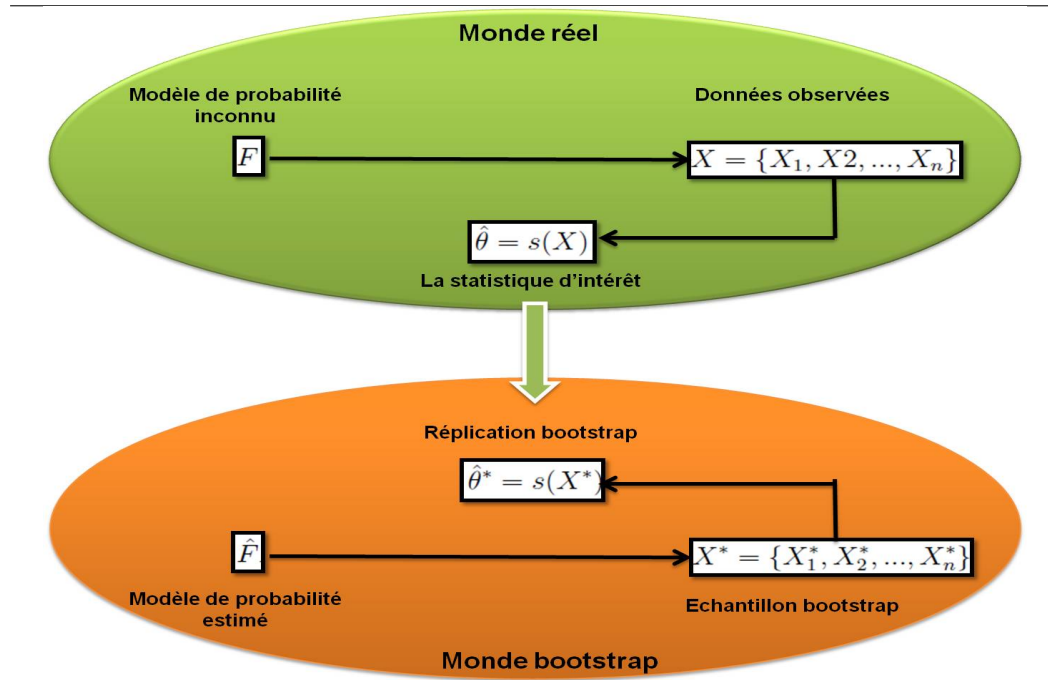


FIGURE 2.1 – L'approche du Bootstrap.

bootstrap peut répondre à de nombreuses questions avec très peu d'information, que ce soit en termes de modélisation ou d'hypothèses. De plus, le bootstrap est assez facilement applicable. Ainsi, dans un contexte où l'outil informatique et les ordinateurs se développent à une vitesse exponentielle, les méthodes basées sur le calcul intensif, telle que le bootstrap deviennent très maniables et pratiques. En outre, la relative simplicité du bootstrap peut s'avérer une solution alternative très efficace à de nombreux problèmes difficiles [83, 134].

2.2.3 Cas d'utilisation du bootstrap

De façon générale, le bootstrap est une méthodologie destinée à répondre à la question que nous avons vu précédemment, c'est à dire, *quel est la précision et l'exactitude d'un estimateur d'un paramètre donné?* Cette question est fondamentale dans de nombreux problèmes de traitement du signal. Nous verrons dans les chapitres suivants de ce manuscrit comment le bootstrap peut résoudre ces problématiques.

2.2.4 Les domaines d'application du bootstrap

Depuis son développement, le bootstrap a connu de nombreuses applications dans divers domaines tels que dans le radar/sonar, la géophysique, l'imagerie biomédicale, le contrôle, la reconnaissance des formes, le traitement d'image, les sciences de l'environnement...etc. Dans la plupart de ces applications, le bootstrap a été utilisé pour approximer la distribution d'un estimateur ou certaines de ses caractéristiques. L'ouvrage de Zoubir et Iskander [144] rassemble de nombreuses applications du bootstrap dans divers domaines d'applications. Dans ce qui suit, nous rappelons quelques travaux liés aux applications du bootstrap.

Dans le domaine du radar/sonar, le bootstrap a été utilisé dans [114, 113] afin d'estimer une distribution à partir de laquelle un indice de risque de collision dans le trafic aérien a été dérivé. Dans [4, 5, 6], le bootstrap fut utilisé pour approximer un model de Markov caché, puis appliqué sur des problèmes de radar transhorizon. Le bootstrap a, par ailleurs été utilisé pour des mines terrestre [143]. Aussi, des procédures de bootstrap ont été utilisées dans la détection des signaux et à l'analyse de données sismiques [22]. En outre, des méthodes de bootstrap ont été utilisées pour évaluer la performance de différentes techniques de localisation [5]. D'autre part, Reid et al [128] ont construit des intervalles de confiance pour de nombreux paramètres d'un avion, et ce, en ayant accès qu'à une seule réalisation acoustique.

Le bootstrap a suscité le même engouement dans d'autres domaines d'applications, y compris le domaine de la géophysique. En effet, Fisher et Hall [62, 63, 64] ont appliqué le bootstrap afin de résoudre des problèmes liés au paléomagnétisme des roches. Dans d'autres travaux, un critère d'information basé sur le bootstrap a été proposé afin d'estimer le nombre d'harmoniques présentes dans des données de géophysique [139]. Quant à Sacchi [132], il a utilisé le bootstrap pour une analyse de la vitesse à haute résolution. En outre, le bootstrap a été utilisé pour améliorer les modèles de prédiction atmosphérique [127].

Le bootstrap a aussi connu de nombreuses applications dans le domaine des sciences biomédicales et le traitement d'image. En effet, dans [84] les auteurs ont utilisé le bootstrap pour estimer la variance dans le cas d'images de tomographie par émission de positons. D'autres travaux ont exploité le bootstrap dans l'objectif d'éliminer la dépendance entre les pixels dans l'imagerie de la rétine [16]. Il a aussi été utilisé dans la reconstruction d'images dans [38]. En imagerie par résonance magnétique (IRM), le bootstrap a été appliqué pour ajuster des p -values dans de nombreux tests de signification de pixels [106].

Bullmore et al [32] ont utilisé les techniques de bootstrap dans le domaine temporel et pour les ondelettes dans l'analyse des données neurophysiologiques. Une autre application très intéressante du bootstrap est celle exposée dans [36] où les auteurs ont utilisé le bootstrap comme outils permettant de faciliter le diagnostic du cancer du sein en imagerie par ultrason.

En plus de ces exemples d'applications cités ci-dessus, il existe de très nombreuses autres applications du bootstrap dans des domaines très variés tels que la médecine, la pharmacie, la chimie, la météorologie, l'économétrie et la finance, les technologies nucléaires ...etc. Dans les chapitres suivants de ce manuscrit, nos applications seront sur des données issues de la biomécanique.

Dans la partie qui suit, le principe du bootstrap ainsi que quelques exemples et résultats de bootstrap pour quelques estimateurs usuels sont donnés.

2.3 Le principe du bootstrap

Afin de mieux expliquer la façon dont le bootstrap fonctionne en pratique, supposons que nous observons un échantillon $X = \{X_1, X_2, \dots, X_n\}$ de taille n tiré aléatoirement d'une distribution complètement inconnue F . Le tirage aléatoire signifie que les X_i sont des variables aléatoires indépendantes et identiquement distribuées (iid) avec une distribution F pour chacune d'elles. Soit θ une caractéristique inconnue de F . Cette caractéristique pourrait être, par exemple, la moyenne ou la variance de F . La problématique ici est de trouver la distribution de $\hat{\theta}$ l'estimateur du paramètre θ dérivé de l'échantillon X . L'objectif est d'inférer θ en se basant sur $\hat{\theta}$.

Un moyen d'obtenir la distribution de $\hat{\theta}$ serait de répéter l'expérimentation un grand nombre de fois, puis approximer la distribution de $\hat{\theta}$ par sa distribution empirique qu'on obtient. Ceci est équivalent à des simulations de Monte Carlo. Cependant, cela n'est pas très applicable en réalité à cause de nombreuses contraintes telles que le temps, les coûts ou bien les conditions de l'expérimentation qui ne sont pas reproductibles.

Ainsi, le bootstrap suggère qu'il est possible de ré-échantillonner à partir d'une distribution choisie qui serait proche de F . Cela peut être la distribution empirique de \hat{F} tel que la mesure de probabilité qui est associée à un ensemble \mathcal{A} appartenant au sous-espace de X une mesure égale à la proportion des valeurs qui sont contenues dans \mathcal{A} . En effet, il est connu que sous certaines hypothèses, \hat{F} converge vers F quand $n \rightarrow \infty$ [83].

Alors, le ré-échantillonnage à partir de \hat{F} est appelé le *bootstrap non-paramétrique*. Ce ré-échantillonnage à partir de \hat{F} est considéré comme le

passage du *monde réel* vers le *monde bootstrap* (figure 2.1).

Le principe du bootstrap non-paramétrique est donc le suivant : calculer $\hat{\theta}$ basé sur l'échantillon $X = \{X_1, X_2, \dots, X_n\}$. Ensuite, construire la distribution empirique \hat{F} . Puis, effectuer un tirage de l'échantillon bootstrap $X^* = \{X_1^*, X_2^*, \dots, X_n^*\}$. Enfin, approximer la distribution de $\hat{\theta}$ à partir de la distribution de $\hat{\theta}^*$ obtenue de l'échantillon bootstrap X^* .

Il est clair que le principe du bootstrap non-paramétrique montre qu'on construit \hat{F} puis on ré-échantillonne à partir de cette dernière. En pratique, il n'est pas obligatoire d'estimer directement F . Cela signifie que la manière dont le bootstrap ré-échantillonne n'implique pas le calcul explicite de la fonction de distribution empirique ou l'histogramme. Aussi, il n'existe pas d'obligation d'estimer la moyenne, la variance, ou les moments d'ordres supérieurs de F . Le bootstrap non-paramétrique utilise tout simplement l'échantillon $X = \{X_1, X_2, \dots, X_n\}$ afin de générer un nouvel échantillon par tirage aléatoire avec remise des éléments de X . Ce nouvel échantillon étant l'échantillon bootstrap.

L'estimation de la distribution de \hat{F} nécessite un grand nombre de répétition B de la procédure du ré-échantillonnage du bootstrap non-paramétrique. En effet, un nombre B d'échantillons bootstrap sont créés X_1^*, \dots, X_B^* . Un échantillon bootstrap $X^* = \{X_1^*, X_2^*, \dots, X_n^*\}$ est obtenu en effectuant des tirages aléatoires avec remise des éléments de X telle que chaque X_i^* a une probabilité de $1/n$ d'être égal à n'importe quel autre X_j , soit :

$$Prob\{X_i^* = X_j/X\} = 1/n, \quad 1 \leq i, j \leq n.$$

Cela signifie que les X_i^* sont iid et que X^* peut contenir des répétitions et l'ordre des éléments n'a aucun rôle [60, 83]. La figure 2.2 explique plus clairement une implémentation pratique du principe du bootstrap non-paramétrique pour l'estimation de la fonction de distribution d'un estimateur donné.

La figure 2.3 montre un exemple d'approximation de la distribution par le bootstrap. La taille de l'échantillon original est égale à $n = 30$ et le nombre de répétitions $B = 1000$.

Dans ce qui suit, nous présentons quelques résultats théoriques du bootstrap pour la moyenne.

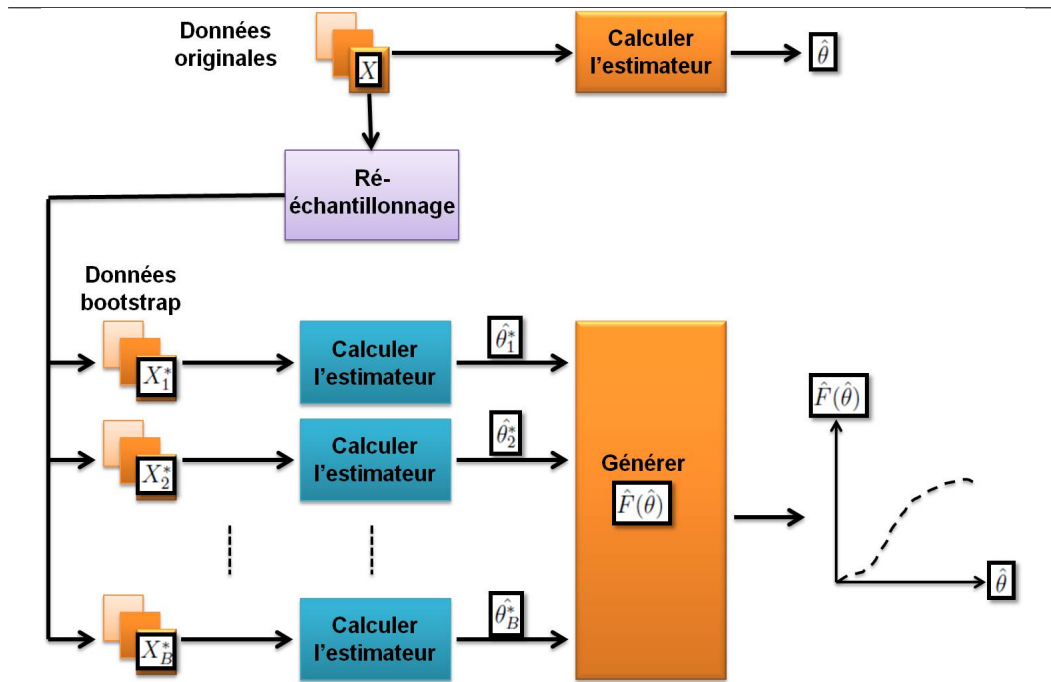


FIGURE 2.2 – Principe du bootstrap non-paramétrique pour l'estimation de la fonction de distribution de $\hat{F}(\hat{\theta})$ du paramètre $\hat{\theta}$ [142].

2.4 Résultats théoriques du bootstrap pour la moyenne

Après avoir donné la définition du bootstrap et son principe de manière générale, il est important maintenant d'évaluer la consistance et la convergence du ré-échantillonnage bootstrap. En effet, cela permettrait d'avoir une idée sur la quantité d'information que les échantillons bootstrap retiennent des échantillons de départ. Dans la littérature, de nombreux résultats de consistance du bootstrap pour diverses statistiques ont été démontrés. Dans ce qui suit, nous nous intéresserons à un seul type de statistique qui est la moyenne et ses variantes.

2.4.1 La consistance du bootstrap pour la moyenne

Soit l'échantillon observé $X = \{X_1, X_2, \dots, X_n\}$. Nous considérons le paramètre d'intérêt $\theta = E(X)$ qui représente l'espérance mathématique de X . L'estimateur naturel de θ est la moyenne, soit :

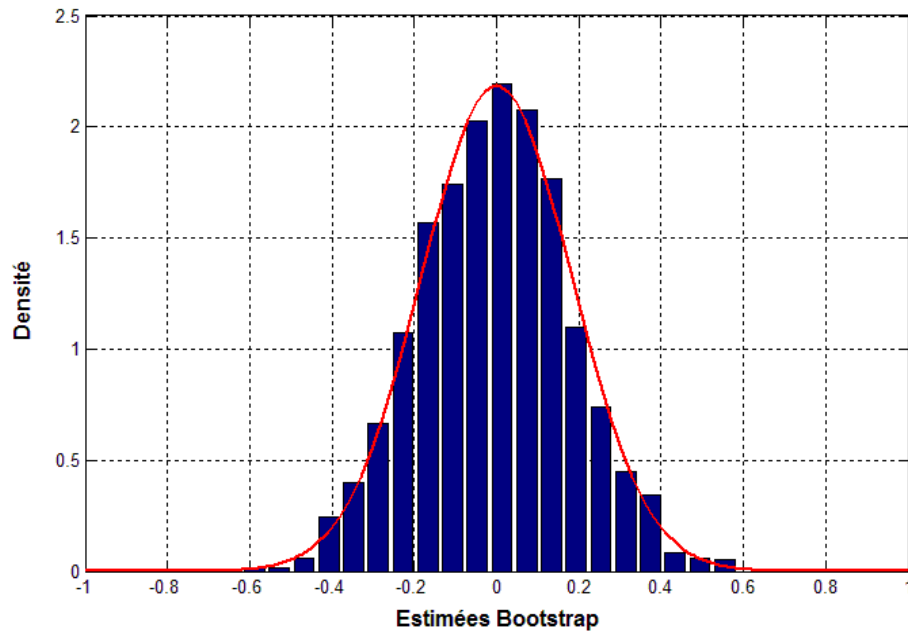


FIGURE 2.3 – Exemple de distribution approximée par le bootstrap non paramétrique. Le paramètre d'intérêt est la moyenne. $X \sim \mathcal{N}(0,1)$ de taille $n = 30$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

Nous définissons aussi la version du bootstrap de cet estimateur comme suit :

$$\hat{\theta}^* = \frac{1}{n} \sum_{i=1}^n X_i^*$$

D'après [135], si $E(X^2) < \infty$ alors :

$$\sup_{x \in \mathcal{R}} \left| P^* \left(\sqrt{n} \left(\hat{\theta}^* - \hat{\theta} \right) \leq x \right) - P \left(\sqrt{n} \left(\hat{\theta} - \theta \right) \leq x \right) \right| \rightarrow 0 \quad \text{p.s.},$$

Si nous notons $T_n^* = \sqrt{n} \left(\hat{\theta}^* - \hat{\theta} \right)$ et $T_n = \left(\sqrt{n} \left(\hat{\theta} - \theta \right) \right)$, alors ce résultat signifie que la distribution de la version bootstrap T_n^* de l'estimateur de la moyenne converge presque-surement vers la distribution T_n de l'estimateur

initial telle que la norme-sup de la différence entre les deux distributions tends vers zéro.

Ainsi, ce résultat montre bien que le bootstrap est consistant pour l'estimateur de la moyenne d'une variable aléatoire X . Des résultats similaires pour d'autres types de moyennes peuvent être dérivés de ce dernier résultat.

2.4.2 La consistance du bootstrap pour la moyenne standardisée

Soient $\theta = E(X)$, $\hat{\theta}$ et $\hat{\theta}^*$ respectivement l'espérance de X , l'estimateur naturel de la moyenne de X et la version bootstrap de l'estimateur de la moyenne de X .

Soit $\sigma^2 = \text{Var}(X)$ la variance de X telle que :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\theta})^2$$

l'estimateur biaisé de la variance.

Si $E(X^3) < \infty$, alors selon [135], nous obtenons le résultat suivant :

$$\sqrt{n} \sup_{x \in \mathcal{R}} \left| P^* \left(\sqrt{n} (\hat{\theta}^* - \hat{\theta}) / \hat{\sigma} \leq x \right) - P \left(\sqrt{n} (\hat{\theta} - \theta) / \sigma \leq x \right) \right| \longrightarrow 0 \quad \text{p.s.},$$

Ce résultat montre encore une fois la convergence des deux distributions, ce qui signifie la consistance du bootstrap pour l'estimateur de la moyenne standardisée. Pour être plus précis, il a été montré que l'approximation de la distribution de $T_n = \left(\sqrt{n} (\hat{\theta} - \theta) / \sigma \right)$ dans le sens de distance norme-sup par la distribution de $T_n^* = \sqrt{n} (\hat{\theta}^* - \hat{\theta}) / \hat{\sigma}$ est de l'ordre de $o(n^{-1/2})$.

2.4.3 La consistance du bootstrap pour la moyenne studentisée

Soient $\theta = E(X)$, $\hat{\theta}$ et $\hat{\theta}^*$ respectivement l'espérance de X , l'estimateur naturel de la moyenne de X et la version bootstrap de l'estimateur de la moyenne de X .

Soient $\hat{\sigma}^2$ l'estimateur biaisé de la variance défini dans 2.4.2 et sa version bootstrap donnée comme suit :

$$\hat{\sigma}^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i^* - \hat{\theta}^*)^2$$

Si $E(X^6) < \infty$, alors on obtient le résultat suivant :

$$\sqrt{n} \sup_{x \in \mathcal{R}} \left| P^* \left(\sqrt{n} \left(\hat{\theta}^* - \hat{\theta} \right) / \hat{\sigma}^* \leq x \right) - P \left(\sqrt{n} \left(\hat{\theta} - \theta \right) / \hat{\sigma} \leq x \right) \right| \rightarrow 0 \quad \text{p.s.},$$

Ce résultat stipule que différence dans le sens de la distance norme-sup entre la distribution de la moyenne studentisée $T_n = \left(\sqrt{n} \left(\hat{\theta} - \theta \right) / \hat{\sigma} \right)$ et la distribution issue du bootstrap $T_n^* = \sqrt{n} \left(\hat{\theta}^* - \hat{\theta} \right) / \hat{\sigma}^*$ est de l'ordre de $o(n^{-1/2})$ [117].

Ainsi, ces résultats pris de la littérature montrent bien que le bootstrap est consistant pour les trois types de moyennes présentées ci-dessus. Bien évidemment, d'autres résultats de consistance pour de nombreux autres estimateurs peuvent être trouvés dans la littérature, mais le point principal à retenir est que de façon générale, les échantillons issus du ré-échantillonnage par bootstrap contiennent la même information statistique que les échantillons originaux. Ce qui signifie la convergence et la consistance du ré-échantillonnage par bootstrap. Nous verrons dans les chapitres 3, 4 et 5 d'autres résultats de consistance de méthodes de bootstrap pour des estimateurs de cyclostationnarité à l'ordre un et à l'ordre deux.

2.5 Exemples d'estimations par le bootstrap non paramétrique

Le bootstrap non-paramétrique est utilisé souvent pour estimer la distribution d'un estimateur d'un paramètre donné provenant d'un échantillon aléatoire. Par conséquent, certains paramètres tels que le biais, la variance et les intervalles de confiance peuvent aussi être dérivés. Afin d'explicitier ces dires, nous montrons trois exemples d'estimations basées sur le bootstrap non-paramétrique.

2.5.1 Estimation du biais

Dans ce premier exemple, nous considérons le problème d'estimation de la variance d'une distribution inconnue notée par $F_{\mu, \sigma}$ qui a pour paramètres : μ et σ .

Soit un échantillon aléatoire $X = \{X_1, X_2, \dots, X_n\}$. Deux types d'estimateurs sont usuellement utilisés :

2.5. Exemples d'estimations par le bootstrap non paramétrique 63

$$\hat{\sigma}_u^2 = \frac{1}{n-1} \sum_{i=1}^n \left(X_i - \frac{1}{n} \sum_{j=1}^n X_j \right)^2 \quad (2.1)$$

et

$$\hat{\sigma}_b^2 = \frac{1}{n} \sum_{i=1}^n \left(X_i - \frac{1}{n} \sum_{j=1}^n X_j \right)^2 \quad (2.2)$$

Sachant que :

$$E(\hat{\sigma}_u^2) = \sigma^2$$

et

$$E(\hat{\sigma}_b^2) = \left(1 - \frac{1}{n}\right) \sigma^2$$

Alors, il est clair que $\hat{\sigma}_b^2$ est un estimateur biaisé de σ^2 alors que $\hat{\sigma}_u^2$ est un estimateur non biaisé de σ^2 .

En utilisant le bootstrap, il est possible d'estimer le biais :

$$b(\hat{\sigma}^2) = E(\hat{\sigma}^2 - \sigma^2)$$

Par :

$$E^*(\hat{\sigma}^{*2} - \hat{\sigma}^2)$$

Avec $\hat{\sigma}^2$ l'estimé du maximum de vraisemblance de σ^2 et $E^*(.)$ l'espérance de l'échantillon bootstrap. Ainsi, et sous ces notations, le bootstrap non-paramétrique est utilisé pour l'estimation du biais de la façon suivante :

- (1) À partir de $X = \{X_1, X_2, \dots, X_n\}$, estimer $\hat{\sigma}_u^2$ et $\hat{\sigma}_b^2$ en utilisant (2.1) et (2.2).
- (2) Effectuer un tirage aléatoire avec remise de n échantillons à partir de X et construire X^* .
- (3) Calculer les estimés bootstrap $\hat{\sigma}_u^{*2}$ et $\hat{\sigma}_b^{*2}$ en utilisant (2.1) et (2.2) basé sur X^* .
- (4) Répéter les étapes (1) et (2) afin d'obtenir un grand nombre B d'estimés bootstrap $\hat{\sigma}_{u,1}^{*2}, \hat{\sigma}_{u,2}^{*2}, \dots, \hat{\sigma}_{u,B}^{*2}$ et $\hat{\sigma}_{b,1}^{*2}, \hat{\sigma}_{b,2}^{*2}, \dots, \hat{\sigma}_{b,B}^{*2}$.

(5) Estimer le biais $b(\hat{\sigma}_u^2)$ en utilisant :

$$b^*(\hat{\sigma}_u^{*2}) = \frac{1}{B} \sum_{i=1}^B \hat{\sigma}_{u,i}^{*2} - \hat{\sigma}_b^2$$

et $b(\hat{\sigma}_b^2)$ par

$$b^*(\hat{\sigma}_b^{*2}) = \frac{1}{B} \sum_{i=1}^B \hat{\sigma}_{b,i}^{*2} - \hat{\sigma}_b^2$$

Ainsi, il est tout à fait possible d'exploiter le bootstrap non-paramétrique pour l'estimation du biais.

2.5.2 Estimation de la variance

Dans ce second exemple, nous nous focalisons sur le problème d'estimation de la variance $\sigma_{\hat{\theta}}^2$ d'un estimateur $\hat{\theta}$ du paramètre θ . Cet estimateur est basé sur l'échantillon $X = \{X_1, X_2, \dots, X_n\}$ provenant d'une distribution inconnue F_{θ} .

Dans certains cas, il est possible d'obtenir analytiquement la forme de $\sigma_{\hat{\theta}}^2$. Sinon, on utilise des résultats asymptotiques afin d'obtenir une estimée $\hat{\sigma}_{\hat{\theta}}^2$ de $\sigma_{\hat{\theta}}^2$. Cependant, cette approche n'est pas toujours valide et applicable à cause de l'hypothèse de n très grand afin d'atteindre le régime asymptotique. De plus, même dans les cas où les résultats asymptotiques sont applicables, les expressions des estimées de la variance sont très souvent très compliquées (voir le chapitre 5). Ainsi, le bootstrap semble une excellente alternative à ce type de problème. En effet, il est possible d'exploiter le bootstrap afin d'approximer $\sigma_{\hat{\theta}}^2$ par $\sigma_{\hat{\theta}}^{*2}$. Donc, de façon un peu similaire à l'estimation du biais, nous allons montrer les différentes étapes nécessaires pour résoudre ce problème. En effet, on utilise l'échantillon original $X = \{X_1, X_2, \dots, X_n\}$ pour générer des échantillons bootstrap. Ensuite, ces échantillons bootstrap sont utilisés pour calculer l'estimée de $\sigma_{\hat{\theta}}^{*2}$. La méthode se résume en quatre étapes comme suit :

- (1) Effectuer un tirage aléatoire avec remise de n échantillons à partir de X et construire X^* .
- (2) Calculer l'estimé bootstrap $\hat{\theta}^*$ basé sur l'échantillon bootstrap X^* .

2.5. Exemples d'estimations par le bootstrap non paramétrique 65

- (3) Répéter les étapes (1) et (2) afin d'obtenir un grand nombre B d'estimés bootstrap $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$.
- (4) Estimer la variance $\sigma_{\hat{\theta}}^2$ de $\hat{\theta}$ comme suit :

$$\sigma_{\hat{\theta}}^{*2} = \frac{1}{B-1} \sum_{i=1}^B \left(\hat{\theta}_i^* - \frac{1}{B} \sum_{j=1}^B \hat{\theta}_j^* \right)^2$$

2.5.3 Intervalles de confiance pour la moyenne

L'une des applications les plus importantes du bootstrap est la construction d'intervalles de confiance pour un estimateur d'intérêt. En effet, les intervalles de confiance représentent une mesure de la précision et de l'exactitude de l'estimateur. De plus, les intervalles de confiance permettent de construire différents tests statistiques qui peuvent s'avérer très utiles comme nous le verrons dans les prochains chapitres de ce manuscrit.

Soit $X = \{X_1, X_2, \dots, X_n\}$ disponible d'une distribution inconnue $F_{\mu, \sigma}$. Dans cette partie, nous souhaitons trouver un estimateur de la moyenne ainsi que ses $100(1 - \delta)\%$ intervalles de confiance. Pour ce faire, l'estimateur de la moyenne est utilisé :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$$

Des intervalles de confiance peuvent être obtenus à partir de l'approximation de la distribution de $\hat{\mu}$. En effet, l'objectif est de trouver les valeurs $\hat{\mu}_L$ et $\hat{\mu}_U$ qui représentent respectivement la borne inférieure et la borne supérieure de l'intervalle de confiance pour $\hat{\mu}$, soit :

$$P(\hat{\mu}_L \leq \hat{\mu} \leq \hat{\mu}_U) = 1 - \delta, \quad \text{avec } \delta \in (0, 1).$$

La distribution de $\hat{\mu}$ dépend de la distribution des X_i qui est inconnue. Dans le cas où n est très grand, la distribution de $\hat{\mu}$ peut être approximée par une distribution Gaussienne en utilisant le théorème central limite. Cependant, une telle approximation serait erronée dans le cas où n ne serait pas très grand.

La consistance du bootstrap implique que l'échantillon de départ $X = \{X_1, X_2, \dots, X_n\}$ constitue par lui-même une distribution sous-jacente. Ainsi, si on ré-échantillonne à partir de X un certain nombre de fois, puis en calculant $\hat{\mu}$ pour chaque nouvel échantillon bootstrap, il est possible d'obtenir une approximation de la distribution de $\hat{\mu}$. Enfin, des intervalles de confiance pour

le paramètre μ peuvent être dérivés à partir de cette distribution et pour un niveau de confiance souhaité δ .

Les différentes étapes de la construction d'intervalles de confiance pour la moyenne μ sont les suivantes :

- (1) Estimer $\hat{\mu}$ sur l'échantillon original X .
- (2) Effectuer un tirage aléatoire avec remise de n échantillons à partir de X et construire l'échantillon bootstrap X^* .
- (3) Calculer l'estimé bootstrap $\hat{\mu}^*$ basé sur l'échantillon bootstrap X^* .
- (4) Répéter les étapes (1) et (2) afin d'obtenir un grand nombre B d'estimés bootstrap $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$.
- (5) Approximer la distribution de $\hat{\mu}$. Arranger dans un ordre croissant les $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$ obtenus dans l'étape précédente telle que $\hat{\mu}_{(1)}^* \leq \hat{\mu}_{(2)}^* \leq \dots \leq \hat{\mu}_{(B)}^*$. Avec $\hat{\mu}_{(k)}^*$ représente le k -ième plus petit de $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$.
- (6) Dériver les intervalles de confiance $(\hat{\mu}_{(m1)}^*, \hat{\mu}_{(m2)}^*)$, avec $m1$ est égale à la partie entière de $B \frac{\delta}{2}$ et $m2$ est égale à $B - m1 + 1$. Sachant que le niveau de confiance $\delta \in (0, 1)$. Ainsi, on obtient les intervalles de confiance pour l'estimateur $\hat{\mu}$ de la moyenne μ .

Il est nécessaire de souligner qu'il existe différentes méthodes pour construire des intervalles de confiance pour un estimateur d'un paramètre d'intérêt. Nous verrons d'autres types d'intervalles de confiance dans les chapitres suivants.

Dans la partie qui suit, l'intérêt est porté sur un autre genre de bootstrap qui est le bootstrap paramétrique. Nous rappelons brièvement son principe et nous donnons un exemple d'application.

2.6 Bootstrap paramétrique

Afin de définir simplement le bootstrap paramétrique, supposons qu'il existe des informations sur F . En effet, si on prend l'exemple d'une distribution F Gaussienne mais avec des paramètres de moyenne et de variance μ et σ^2 inconnus. En utilisant le bootstrap, cela suggère que le tirage de l'échantillon bootstrap de taille n se fait à partir d'une distribution Gaussienne ayant une moyenne $\hat{\mu}$ et une variance $\hat{\sigma}^2$. Contrairement au cas non-paramétrique, ces deux paramètres sont estimés à partir des données originales X . Cette

méthode de bootstrap paramétrique a été introduite par [83, 60] et se base sur la connaissance à priori de la distribution des données originales, d'où sont appellation *paramétrique*.

Alors, étant donné un échantillon original $X = \{X_1, X_2, \dots, X_n\}$ de F_θ , on estime tout d'abord le paramètre d'intérêt θ qui pourrait être un des paramètres d'une Gaussienne (si F est Gaussienne). Puis, on effectue le tirage aléatoire de taille n de \hat{F}_θ et on répète cela un nombre de fois B .

Afin de mieux comprendre le concept du bootstrap paramétrique, un exemple d'application est montré dans ce qui suit.

2.6.1 Intervalles de confiance pour la moyenne via le bootstrap paramétrique

Soit $X = \{X_1, X_2, \dots, X_n\}$ disponible d'une distribution supposée connue $F_{\mu, \sigma}$ (Gaussienne, par exemple). Nous souhaitons trouver les $100(1-\delta)\%$ intervalles de confiance pour la moyenne en utilisant une approche paramétrique. Les différentes étapes sont les suivantes :

- (1) Estimer $\hat{\mu}$ sur l'échantillon original X sachant que $F_{\mu, \sigma}$ est *Gaussienne* $\mathcal{N}(\mu, \sigma^2)$.
- (2) Effectuer un tirage aléatoire avec remise de n échantillons à partir de $F_{\hat{\mu}, \hat{\sigma}}$ et construire l'échantillon bootstrap X^* .
- (3) Calculer l'estimé bootstrap $\hat{\mu}^*$ basé sur l'échantillon bootstrap X^* .
- (4) Répéter les étapes (1) et (2) afin d'obtenir un grand nombre B d'estimés bootstrap $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$.
- (5) Approximer la distribution de $\hat{\mu}$. Arranger dans un ordre croissant les $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$ obtenus dans l'étape précédente telle que $\hat{\mu}_{(1)}^* \leq \hat{\mu}_{(2)}^* \leq \dots \leq \hat{\mu}_{(B)}^*$. Avec $\hat{\mu}_{(k)}^*$ représente le k -ième plus petit de $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$.
- (6) Dériver les intervalles de confiance $(\hat{\mu}_{(m1)}^*, \hat{\mu}_{(m2)}^*)$, avec $m1$ est égale à la partie entière de $B \frac{\delta}{2}$ et $m2$ est égale à $B - m1 + 1$. Sachant que le niveau de confiance $\delta \in (0, 1)$. Ainsi, on obtient les intervalles de confiance pour l'estimateur $\hat{\mu}$ de la moyenne μ .

Il est clair que la plus grande différence entre l'approche non-paramétrique et l'approche paramétrique du bootstrap réside dans le fait d'avoir la connaissance de la distribution des données originales. Le reste des étapes pour l'obtention des intervalles de confiance pour la moyenne est très similaire.

La figure 2.4 montre un exemple de l'histogramme et de la distribution obtenus par cette approche paramétrique. L'échantillon originale X est pris gaussien avec une moyenne $\mu = 3$, une variance $\sigma^2 = 25$ ($\mathcal{N}(3, 25)$) et un nombre d'échantillon $n = 15$. Le nombre de répétitions B est égal à 1000.

La moyenne estimée à partir de X est $\hat{\mu} = 3.201$ et les intervalles de confiances dérivées par la méthode du bootstrap paramétrique sont $\mu_L = 0.8573$ et $\mu_U = 5.2875$, avec μ_L et μ_U qui représentent respectivement la borne inférieure et la borne supérieure de l'intervalle de confiance pour la moyenne.

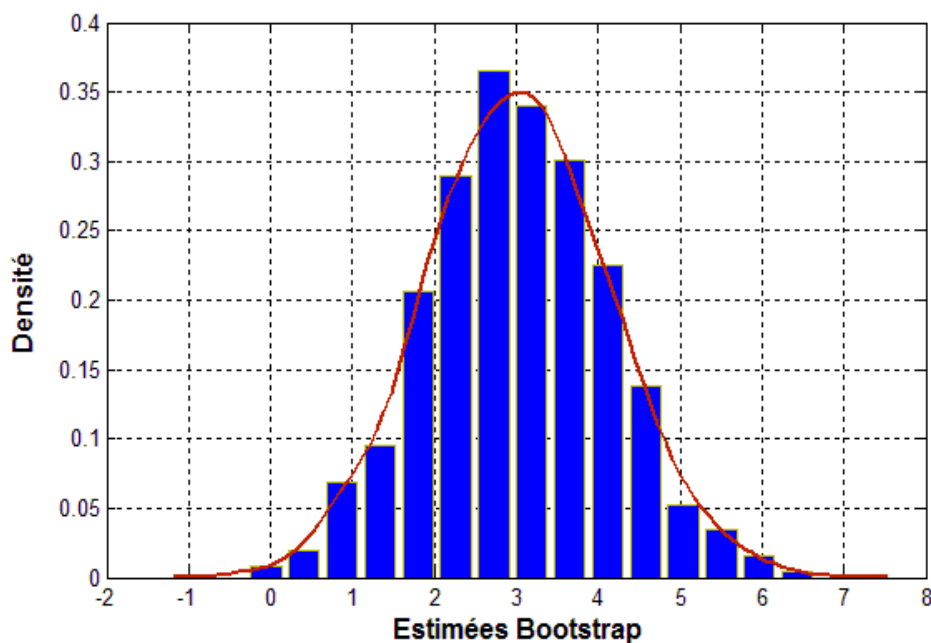


FIGURE 2.4 – Exemple d'histogramme de $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$ obtenus par bootstrap paramétrique. $X \sim \mathcal{N}(3, 25)$ de taille $n = 15$.

Il est très important de noter que la théorie existant sur l'estimation par maximum de vraisemblance est tout à fait adéquate avec ce type de problèmes paramétrique. Par conséquent, le bootstrap paramétrique ajoute que très peu d'information, voire rien à cette théorie. Ainsi, ce type de bootstrap est très peu utilisé dans des problèmes réels. En effet, dans le traitement du signal, les données traitées sont généralement de distributions inconnues. C'est pourquoi, le bootstrap paramétrique présente de sérieuses limitations quant à son application en pratique contrairement au bootstrap non-paramétrique. Ainsi, les méthodes de bootstrap qui sont développées et appliquées dans les chapitres suivants sont de type non-paramétrique.

De nombreuses autres variantes du bootstrap existent dans la littérature, nous pouvons citer le *bootstrap bayésien*, le *bootstrap lissé*, le *double bootstrap* et le *m-out-of-n bootstrap*. Une brève revue de ces méthodes est exposée dans [37].

2.7 Les limitations du bootstrap

Tout d'abord, il est précisé que les limitations qui sont citées dans cette partie concernent uniquement la version non-paramétrique du bootstrap.

Bien que le bootstrap soit une méthode très efficace et très souvent consistante, il existe de nombreux cas pour lesquels le bootstrap échoue. Dans ce qui suit, nous présentons quelques cas où le bootstrap révèle des limitations.

2.7.1 Taille d'échantillon très petite

Pour le bootstrap non-paramétrique, les échantillons bootstrap sont tirés à partir d'un ensemble discret de valeurs. Dans certains cas où la taille de l'échantillon d'origine n est très petite, il existe des arguments pour douter des résultats obtenus par le ré-échantillonnage bootstrap. En effet, dans des échantillons à très petite taille, le choix des valeurs est très limité. Ainsi, les échantillons issus du bootstrap *sous représentent* la réelle variabilité puisque les observations sont très répétées et les échantillons bootstrap, eux mêmes, sont aussi très répétés.

La distribution du bootstrap est définie comme étant celle de tous les échantillons possible de taille n qui sont tirés avec remise à partir de l'échantillon original de taille n aussi. Le nombre d'échantillons possibles est égal logiquement à :

$$\frac{(2n - 1)!}{n!(n - 1)!}$$

Ainsi, il est claire que même pour un n petit, le nombre d'échantillons possibles est très grand. À titre d'exemple, il a été montré dans [83] que pour une taille $n = 20$ et un nombre de répétition bootstrap $B = 2000$ que la probabilité de ne pas obtenir d'échantillons bootstrap répétés était plus grande que 0.95. Cependant, il est rapporté dans [37] que pour des tailles d'échantillons inférieurs à 10, le bootstrap n'est très probablement pas la meilleure des approches. C'est pourquoi, il est nécessaire de disposer d'un échantillon original de taille acceptable.

2.7.2 Séries m-dépendantes

Une série de variables aléatoire infinie $\{X_j\}$ pour $j = 1, 2, \dots, n, \dots, \infty$ est dite m-dépendante si pour chaque j , X_j et X_{j+m-1} sont dépendants, mais X_j et X_{j+m} sont indépendants. Un exemple typique de ce genre de processus est le modèle de moyenne mobile à l'ordre m suivant :

$$X_j - \mu = \sum_{l=1}^m \alpha_l \varepsilon_{j-l}$$

Avec $j \in \mathbb{Z}$.

Le paramètre μ est la moyenne de la séquence stationnaire, $\{X_j\}$ et $\{\alpha_l\}$ sont les coefficients de la moyenne mobile. Quant aux $\{\varepsilon_j\}$, ils ont considérés iid.

La série définie ci-dessus est m-dépendante. En effet, les X_j qui sont séparés par moins de m unités de temps partagent en commun au moins un ε . Cependant, si deux X_j sont séparés par plus de m unités de temps alors ils ne partagent aucun ε en commun, et puisque les ε_j sont iid, alors les deux X_j sont indépendants.

précédemment, nous avons vu le résultat de consistance du bootstrap non-paramétrique pour la moyenne présenté par [135]. Dans le même article, l'auteur s'est aussi intéressé à la consistance du bootstrap pour la moyenne dans le cas de données m-dépendantes. En effet, il a fourni un exemple de non consistance du bootstrap pour la moyenne d'un tel processus. Ainsi, le bootstrap non-paramétrique est inconsistant pour des données m-dépendantes.

2.7.3 Distribution avec des moments infinis

Dans [135], l'auteur à montré que X_1, X_2, \dots, X_n sont des variables iid avec des moments d'ordre deux finis et si Y_1, Y_2, \dots, Y_n sont choisies par ré-échantillonnage statistique aléatoire avec remise à partir de X_1, X_2, \dots, X_n alors :

$$H_n(x, \omega) = P\left(\frac{\bar{Y}_n - \bar{X}_n}{S_n} \leq x | X_1, X_2, \dots, X_n\right),$$

Avec

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i, \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Et soit $\Phi(x)$ la distribution standard normale cumulative, alors on obtient :

$$\sup |H_n(x, \omega) - \Phi(x)| \rightarrow 0, \quad -\infty < x < \infty$$

Avec une probabilité égale à 1 quand $n \rightarrow \infty$

Tels que ω désigne le résultat aléatoire qui mène aux valeurs X_1, X_2, \dots, X_n et la fonction $H_n(x, \omega)$ est une distribution de probabilité aléatoire. Pour des valeurs fixes de X_1, X_2, \dots, X_n ce qui signifie que pour une valeur fixe de ω , $H_n(x, \omega)$ devient une fonction distribution cumulative. Ainsi, par le théorème central limite on a :

$$G_n(x) = P\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \quad \text{converge vers} \quad \Phi(x)$$

Le principe du bootstrap remplace μ et σ par \bar{X}_n et S_n et remplace aussi \bar{X}_n par \bar{Y}_n . Alors, $H_n(x, \omega)$ converge vers $\Phi(x)$ pour toutes les valeurs de ω [135].

Dans [15], l'auteur s'est intéressé à l'obtention du même résultat précédent en considérant que la variance (moment d'ordre 2) est infinie. Les résultats obtenus dans cette étude ont montrés qu'en considérant une variance infinie, la fonction $H_n(x, \omega)$ converge vers une distribution aléatoire, ce qui signifie que le bootstrap n'est pas consistant.

2.7.4 Le sondage

Dans le cas de sondage, la population ciblée est toujours finie. Si la taille de la population est N et la taille de l'échantillon est n , si n/N n'est pas très petit, alors la variance d'une estimée (par exemple la moyenne) est plus petite que dans le cas où on considère une population infinie. Rappelons que pour des variables iid X_i avec $i = 1, 2, \dots, n$ la variance de la moyenne est σ^2/n , avec σ^2 la variance de la population.

Pour une population finie, un échantillon de taille n choisis d'une population de taille N ayant une moyenne μ et une variance σ^2 a une moyenne égale à μ mais une variance égale à $(N - n)\sigma^2/(nN)$. Cette variance est plus petite que dans le cas d'une population infinie avec un facteur égal à $(N - n)/N$. Soit $f = n/N$, alors le facteur $(N - n)/N$ devient $1 - f$. Ce dernier est appelé le facteur de correction. Alors, si le bootstrap est appliqué pour la moyenne, la variance de la moyenne provenant des échantillons bootstrap sera égale à σ^2/n au lieu de la valeur correcte qui devrait être $(1 - f)\sigma^2/n$. C'est pourquoi le bootstrap est inconsistant.

2.7.5 Dépendance à long terme

Un autre cas dans lequel le bootstrap est inconsistant est celui de données qui contiennent des dépendances à long terme. En effet, ce type de processus a pour propriété que ses premières observations continuent à avoir une influence sur des observations beaucoup plus lointaine dans le temps. Une caractérisation mathématique d'un tel processus serait que la somme des autocorrélations divergent. Plus d'exemples et de détails sur les processus dépendant à long terme se trouvent dans [82].

Lahiri [102] a montré l'échec du bootstrap à obtenir une approximation valide de la distribution d'une moyenne normalisée dans le cas de données dépendantes à long terme. Dans ce cas, la difficulté résulte dans le fait que la vitesse de convergence de la moyenne est très lente. Par conséquent, la normalisation mène la moyenne à converger vers une limite dégénérée.

Par ailleurs, d'autres cas d'inconsistance du bootstrap existent, nous pouvons citer par exemple le cas d'estimation de valeurs extrêmes [7] ou bien, le cas de processus autorégressifs instables [102].

Malgré ces quelques limitations du bootstrap non-paramétrique, cette approche reste très utile et consistantes dans de nombreuses autres problématiques liées au traitement du signal.

2.8 Conclusion

Dans ce chapitre, nous avons introduit la notion du bootstrap. En effet, la définition ainsi que le principe du bootstrap ont été présentés. De plus, des exemples d'application du bootstrap dans différents domaines ont été exposés. Quelques résultats théoriques ont aussi été donnés pour deux types de bootstrap qui sont le bootstrap non-paramétrique et le bootstrap paramétrique. Dans les chapitres 3, 4 et 5, nous exploitons les propriétés du bootstrap afin d'étudier les estimateurs de la cyclostationnarité à l'ordre un et à l'ordre deux d'un processus stochastique et nous proposons des méthodologies originales basées sur le ré-échantillonnage statistique pour le traitement des signaux cyclostationnaires.

Ré-échantillonnage statistique dans le cas de signaux cyclostationnaires

Sommaire

3.1	Introduction	74
3.2	La formulation du problème	75
3.3	Le Moving Block Bootstrap (MBB)	77
3.3.1	Le MBB pour les coefficients de Fourier	78
3.3.2	Algorithme du MBB	79
3.4	Le subsampling	80
3.4.1	Le subsampling pour les coefficients de Fourier	80
3.4.2	Algorithme du subsampling	82
3.5	Etude de simulation et de performances	83
3.6	Application du subsampling sur des signaux biomécaniques	91
3.6.1	La description des signaux de GRF	91
3.6.2	Pré-traitement des signaux de GRF	93
3.6.3	Analyse cyclostationnaire basé sur le subsampling	93
3.6.4	Cartographies de significativité fréquentielle cyclostationnaire	96
3.6.5	La comparaison entre les trois niveaux de fatigue du coureur	98
3.7	Conclusion	100

3.1 Introduction

La cyclostationnarité utilisée comme caractéristique des signaux est une technique qui offre des avantages pour étudier et diagnostiquer des défauts et des perturbations liés à un système sous étude. Au cours des dernières décennies, la modélisation des phénomènes périodiques a été relativement bien explorée. Cependant, quelques travaux de recherche se sont intéressés à l'estimation de l'autocovariance (autocorrélation) périodique des processus cyclostationnaires et aux résultats asymptotiques qui pourraient être obtenus. De récents travaux [42, 103, 93] ont montré que ces estimateurs étaient asymptotiquement normaux. Cependant, la matrice de covariance asymptotique est très difficile à estimer et à calculer. En effet, cette dernière dépend des moments d'ordre quatre du processus cyclostationnaire étudié [103]. Par conséquent, l'inférence statistique pour les tests et les intervalles de confiance sont très difficile à construire.

L'un des problèmes majeurs rencontré dans ce genre de cas est *comment détecter ce type de structure non-stationnaire dans le processus sous considération*. Cette problématique a été exploré dans le domaine temporel par différents chercheurs [39, 43, 137, 103]. Durant les dernières années, de nombreux travaux de recherche se sont penchés sur les méthodes de bootstrap et de ré-échantillonnage statistique. Nous citons à titre d'exemple la référence [144] qui contient des résultats très intéressants sur le bootstrap et ses applications dans le domaine du traitement du signal.

Par ailleurs, d'autres travaux se sont intéressés à la consistance de telles méthodes et à leurs possibles applications sur des modèles et signaux non-stationnaires, et spécialement dans le cas des processus cyclostationnaires [42, 103].

Dans ce chapitre, nous nous focalisons sur la consistance de deux méthodes de ré-échantillonnage statistique pour l'estimateur du paramètre $|R_X(\alpha, \tau)|$ d'un signal $\{X(t) : t \in \mathbb{Z}\}$. Les deux méthodes sont dénommées le *Subsampling* (ou sous-échantillonnage statistique) et le *Moving Block Bootstrap* (MBB). Avec les nombres complexes $R_X(\alpha, \tau)$ représentant les coefficients de Fourier de la fonction d'autocorrélation donnée par :

$$R_X(t, \tau) = \sum_{\alpha \in \mathcal{A}_\alpha} R_X(\alpha, \tau) e^{j2\pi\alpha t}$$

La consistance ci-dessus signifie la convergence des distributions des estimateurs des deux méthodes citées à la distribution asymptotique de l'estimateur initial.

L'objectif du travail effectué dans ce chapitre est la construction de tests statistiques via le subsampling et le MBB pour la présence de cyclostationnarité, i.e., tester l'hypothèse de non-cyclostationnarité du processus étudié versus l'hypothèse alternative de présence de structure cyclostationnaire dans ce dernier. Cela sera fait en testant si les coefficients de Fourier $R_X(\alpha, \tau)$ de la fonction d'autocorrélation $R_X(t, \tau)$ sont nulles en dehors d'un ensemble fini de fréquences cycliques, ou pas.

Ce chapitre est organisé de la façon suivante : la Section 2 contient un bref rappel sur quelques définitions basiques et nécessaires ainsi que la formulation du problème. Dans la section 3, les résultats fondamentaux de la consistance du MBB pour la fonction moyenne de l'estimateur $|R_X(\alpha, \tau)|$ ainsi que l'algorithme du MBB sont présentés. La section 4 est dédiée aux résultats fondamentaux de la consistance du subsampling, aussi bien qu'à la description de son algorithme. La section 5 contient une étude de simulation dans le cas d'un signal cyclostationnaire. De plus, une étude comparative entre les deux méthodes est réalisée. Dans la section 6, les applications sur des signaux réels biomécaniques, les résultats ainsi que plusieurs possibles applications pour les tests des fréquences cycliques sont montrés. En outre, quelques nouvelles méthodes pour la construction de signatures fréquentielles cyclostationnaires décrivant et caractérisant les signaux réels sont proposées dans cette section. Enfin, la section 7 contient les conclusions de ce chapitre.

3.2 La formulation du problème

Nous rappelons qu'un signal $\{X(t) : t \in \mathbb{Z}\}$ est dit cyclostationnaire si ses moments d'ordre un et d'ordre deux sont périodiques. Plus précisément si sa moyenne et sa fonction d'autocorrélation sont toutes deux périodiques dans le temps t , avec la période T :

$$\mu_X(t) = \mu_X(t + T)$$

$$R_X(t, \tau) = R_X(t + T, \tau)$$

Puisque notre intérêt se porte dans ce chapitre sur les propriétés cyclostationnaires d'ordre deux, i.e., la fonction d'autocorrélation $R_X(t, \tau)$, nous assumons tout au long de ce chapitre que le moment d'ordre un est nul. De plus, la représentation de Fourier de la fonction d'autocorrélation est donnée comme suit :

$$R_X(t, \tau) = \sum_{\alpha \in \mathcal{A}_\alpha} R_X(\alpha, \tau) e^{j2\pi\alpha t} \quad (3.1)$$

Avec

$$R_X(\alpha, \tau) = \frac{1}{T} \sum_{t=1}^T R_X(t, \tau) e^{-j2\pi\alpha t},$$

Et l'ensemble

$$\mathcal{A}_\alpha = \{\alpha : R_X(\alpha, \tau) \neq 0\}$$

est supposé être fini.

Il est important de rappeler que les nombres réels α et les valeurs complexes $R_X(\alpha, \tau)$, appelés fréquences cycliques et fonctions d'autocorrélation cyclique représentent respectivement les fréquences et les coefficients de Fourier de la fonction d'autocorrélation $R_X(t, \tau)$.

Cette représentation de Fourier est très utile pour estimer la fonction d'autocorrélation $R_X(t, \tau)$ à travers l'estimation de ses coefficients de Fourier complexes $R_X(\alpha, \tau)$. Par ailleurs, si la période T de la fonction d'autocorrélation $R_X(t, \tau)$ est connue, alors l'ensemble \mathcal{A}_α contient les multiples de la fréquence cyclique fondamentale α correspondant à la période T . Cependant, en pratique, les éléments constituant l'ensemble \mathcal{A}_α sont très souvent inconnus. Pour cela, le travail effectué dans ce chapitre tend à démontrer l'utilité et la nécessité d'analyses statistiques modernes telles que le subsampling et le MBB pour l'estimation et l'identification des éléments de l'ensemble \mathcal{A}_α d'un processus cyclostationnaire.

Soit un échantillon observé $\{X(1), X(2), \dots, X(n)\}$ de $\{X(t) : t \in \mathbb{Z}\}$. Alors l'estimateur naturel des paramètres $R_X(\alpha, \tau)$ est donné comme suit [94] :

$$\hat{R}_n(\alpha, \tau) = \frac{1}{n - |\tau|} \sum_{t=1-\min\{\tau,0\}}^{n-\max\{\tau,0\}} X(t+\tau)X(t)e^{-j2\pi\alpha t} \quad (3.2)$$

Avec $\hat{R}_n(\alpha, \tau)$ l'estimateur de la fonction d'autocorrélation cyclique. Cet estimateur combine l'information sur le retard, représenté par l'argument τ et la fréquence cyclique, représentée par l'argument α . Il faut noter que l'analyse statistique de l'estimateur (3.2) est fondamentale afin d'identifier le caractère cyclostationnaire du processus sous étude. Cela signifie que si pour toute fréquence α non nulle, le paramètre $\hat{R}_n(\alpha, \tau)$ n'est pas significativement différent de zéro, alors, il serait raisonnable d'admettre que $X(t)$ ne montre pas

de structure cyclostationnaire. Cependant, si le paramètre $\hat{R}_n(\alpha, \tau)$ est significativement différent de zéro pour au moins une fréquence α non nulle, alors $X(t)$ est cyclostationnaire.

3.3 Le Moving Block Bootstrap (MBB)

Le MBB est une méthode de bootstrap non paramétrique qui peut être appliquée sur des données présentant une dépendance. L'idée du MBB a été proposée par [101, 105] dans le cas stationnaire. Cette idée consiste à calculer l'estimateur d'intérêt sur des pseudo-données répliquées. Ces dernières sont obtenues en faisant un tirage aléatoire avec remise à partir de blocs consécutives de données. Pour être plus précis, à partir des n échantillons disponibles de $\{X(t) : t \in \mathbb{Z}\}$, une sélection aléatoire de blocs de taille b à partir des données originales est réalisée. Ensuite, ces blocs sont concaténés afin d'obtenir des pseudo-données répliquées de bootstrap. Par conséquent, la procédure du MBB nous permet d'obtenir une approximation des distributions inconnues des statistiques d'intérêt. En effet, ces approximations sont très importantes car elles permettent la construction d'intervalles de confiance valides ainsi que l'élaboration de différents tests statistiques.

La méthode de MBB utilisée dans cette partie est inspirée de [137]. En effet, cette méthode est dédiée à un cas spécial de processus non-stationnaire, en l'occurrence, les processus cyclostationnaires. L'idée générale du MBB est d'investiguer la valeur moyenne $\mu = \langle EX_t \rangle_t$ du signal $\{X(t) : t \in \mathbb{Z}\}$. Avec $\langle \cdot \rangle_t$ le moyennage sur la variable de temps t . Ainsi, l'estimateur obtenu à partir de l'observation $\{X(1), \dots, X(n)\}$ a la forme suivante :

$$\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$$

Afin de montrer le principe du MBB pour la valeur moyenne μ , nous supposons disposer d'une taille de bloc b , avec b entier. Ensuite, on définit $B_{t,b} = (X_t, \dots, X_{t+b-1})$ et l'entier $k = \lfloor n/b \rfloor$. Soit i_1, \dots, i_k des variables aléatoires indépendantes et identiquement distribuées avec une distribution uniforme sur l'ensemble $\{1, 2, \dots, n - b + 1\}$. Ainsi, en joignant les k blocs $B_{i_1,b}, \dots, B_{i_k,b}$, sachant que chaque bloc est de taille égale à b , nous obtenons l'échantillon bootstrap $\{X(1)^*, \dots, X(n)^*\}$ de taille n . Par conséquent, la version MBB de l'estimateur de départ sera sous la forme suivante :

$$\bar{X}_n^* = \frac{1}{n} \sum_{t=1}^n X_t^*$$

Finalemment, la répétition de la même procédure du MBB un certain nombre de fois permet le calcul des quantiles sans pour autant utiliser la forme de la variance asymptotique.

3.3.1 Le MBB pour les coefficients de Fourier

Dans le cadre cyclostationnaire, l'inférence statistique se focalise seulement sur les propriétés statistiques du second ordre du processus cyclostationnaire étudié. La représentation de Fourier de la fonction d'autocorrélation $R_X(t, \tau)$ est donnée par l'équation (3.1). Soit la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$ suivante :

$$R_X(\alpha, \tau) = \langle R_X(t, \tau) e^{-j2\pi\alpha t} \rangle_t$$

Tel que l'estimateur naturel des paramètres $R_X(\alpha, \tau)$ est donné par l'équation (3.2). Dans [137], la consistance du MBB dans le cas d'un processus cyclostationnaire est démontrée, avec comme paramètres d'intérêt les coefficients de Fourier de la fonction d'autocorrélation. Afin d'éclaircir tout cela, soit le signal $\{X(t) : t \in \mathbb{Z}\}$. Un nouveau vecteur complexe $U_t(\alpha, \tau)$ est défini pour des valeurs de α et τ fixes, soit :

$$U_t(\alpha, \tau) = X(t)X(t + \tau)e^{-j2\pi\alpha t} \quad (3.3)$$

Ainsi, nous avons

$$\langle E(U_t(\alpha, \tau)) \rangle_t = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n E(X(t + \tau)X(t))e^{-j2\pi\alpha t} = R_X(\alpha, \tau) \quad (3.4)$$

Par conséquent, nous obtenons

$$\bar{U}_n(\alpha, \tau) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} U_t = \hat{R}_n(\alpha, \tau) \quad (3.5)$$

Ainsi, la procédure du MBB peut être appliquée, comme définie dans la section 3.3, sur la fonction de moyenne donnant l'estimateur $\hat{R}_n(\alpha, \tau)$. Cela est réalisé de la manière suivante :

Soit l'échantillon original $U_t(\alpha, \tau)$ donné comme suit :

$$(U_1(\alpha, \tau), \dots, U_{n-\tau}(\alpha, \tau)) = (X(1)X(1 + \tau)e^{-j2\pi\alpha}, \dots, X(n - \tau)X(n)e^{-j2\pi\alpha(n - \tau)})$$

Et soit l'échantillon MBB $U_t^*(\alpha, \tau)$ obtenu à partir de $U_t(\alpha, \tau)$:

$$(U_1^*(\alpha, \tau), \dots, U_{n-\tau}^*(\alpha, \tau))$$

Ainsi, nous définissons la version MBB $\hat{R}_n^*(\alpha, \tau)$ de l'estimateur $\hat{R}_n(\alpha, \tau)$ de la manière suivante :

$$\hat{R}_n^*(\alpha, \tau) = \bar{U}_n^*(\alpha, \tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} U_t^*(\alpha, \tau)$$

Enfin, après B répétitions de la procédure du MBB pour l'estimateur $\hat{R}_n(\alpha, \tau)$, il est tout à fait possible de construire des intervalles de confiance valides à partir des quantiles de la distribution du MBB. De plus, des tests statistiques pour la détermination des fréquences cycliques significatives peuvent être dérivés et élaborés.

Dans ce qui suit, une description détaillée de l'algorithme du MBB pour les coefficients de Fourier est exposée.

3.3.2 Algorithme du MBB

L'algorithme du MBB pour l'estimation des coefficients de Fourier de la fonction d'autocorrélation et pour tester les fréquences cycliques est décrit ci-dessous. La procédure d'estimation est basée sur les échantillons disponibles $\{X(1), X(2), \dots, X(n)\}$ de $\{X(t) : t \in \mathbb{Z}\}$.

Pour des valeurs fixés de α et τ , l'algorithme du MBB est donné comme suit :

Étape 1 : Définir $U_t(\alpha, \tau) = X(t)X(t + \tau)e^{-j2\pi\alpha t}$

Étape 2 : Estimer $\hat{R}_n(\alpha, \tau) = \bar{U}_n(\alpha, \tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} U_t$ basé sur les données originales.

Étape 3 : Choisir une taille de bloc $b (< n - \tau)$, et un niveau de confiance $\delta \in (0, 1)$

Étape 4 : -Appliquer le MBB sur $U(t)$ pour la fonction moyenne de l'estimateur $\hat{R}_n(\alpha, \tau)$.

-Calculer le nouvel estimateur MBB

$$\hat{R}_n^*(\alpha, \tau) = \bar{U}_n^*(\alpha, \tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} U_t^*(\alpha, \tau)$$

-Définir

$$V_{re_t}^{n,t} = \Re[\sqrt{n}(\hat{R}_n^*(\alpha, \tau) - \hat{R}_n(\alpha, \tau))], \text{ et } V_{im_t}^{n,t} = \Im[\sqrt{n}(\hat{R}_n^*(\alpha, \tau) - \hat{R}_n(\alpha, \tau))]$$

tels que $V_{re_t}^{n,t}$ et $V_{im_t}^{n,t}$ soient respectivement la partie réelle et la partie imaginaire de la statistique définie juste au dessus.

Étape 5 : Répéter l'étape 4 B fois afin d'obtenir B répétitions de l'estimateur original $\hat{R}_n(\alpha, \tau)$

Étape 6 : Ordonner les valeurs de $V_{re_t}^{n,B,t}$ et $V_{im_t}^{n,B,t}$ afin d'obtenir les quantiles réels et imaginaires $c_{n,b}(1 - \frac{\delta}{2})$ et $c_{n,b}(\frac{\delta}{2})$ dérivés du MBB.

Étape 7 : Définir les intervalles de confiance basés sur le MBB pour les deux parties, réelle et imaginaire

$$\left(\hat{R}_n(\alpha, \tau) - \frac{c_{n,b}(1 - \frac{\delta}{2})}{\sqrt{n}}, \hat{R}_n(\alpha, \tau) - \frac{c_{n,b}(\frac{\delta}{2})}{\sqrt{n}} \right)$$

3.4 Le subsampling

L'idée du subsampling a été introduite par Politis [122] et est brièvement décrite dans ce qui suit. Soit P la loi de probabilité qui gouverne la variable $\{X(t) : t \in \mathbb{Z}\}$. La distribution de P est supposée inconnue et appartenant à une certaine classe de mesure de probabilité \mathcal{P} . L'objectif est de construire des intervalles de confiance pour un paramètre $\theta(P) \in \{\theta(P) : P \in \mathcal{P}\}$.

Soit $\hat{\theta}_n = \hat{\theta}_n(X(1), X(2), \dots, X(n))$ un estimateur de $\theta(P)$. Cet estimateur est basé sur l'échantillon disponible $\{X(1), X(2), \dots, X(n)\}$ de $\{X(t) : t \in \mathbb{Z}\}$. Soit une séquence quelconque b tendant à l'infini avec n , tel que $b < n$. De plus, nous prenons $t \in \{1, \dots, n - b + 1\}$ et définissons

$$\hat{\theta}_{n,b,t} = \hat{\theta}_b(X(t), \dots, X(t + b - 1)) \tag{3.6}$$

comme étant l'estimateur du paramètre $\theta(P)$ basé sur le sous-échantillon $\{X(t), X(t+1), \dots, X(t+b-1)\}$. Le paramètre t peut être interprété comme le point de départ du re-calcul de $\hat{\theta}_n$ sur un bloc de taille égale à b . De cette façon, les quantiles de la distribution empirique des répliques de l'estimateur ($n - b + 1$ répliques) sont utilisés pour remplacer les quantiles de la distribution initiale.

3.4.1 Le subsampling pour les coefficients de Fourier

Nous rappelons que l'estimateur $\hat{R}_n(\alpha, \tau)$ défini par l'équation (3.2) pour chaque valeur de τ et $\alpha \in \mathcal{A}_\alpha$ est notre point central d'intérêt. Ce dernier

est basé sur l'échantillon $\{X(1), \dots, X(n)\}$ disponible de $\{X(t) : t \in \mathbb{Z}\}$. Cet estimateur complexe est traité comme un vecteur à deux dimensions, avec des coordonnées égales à sa partie imaginaire et à sa partie réelle.

Pour la commodité du lecteur, il est rappelé que le théorème traitant de la consistance du subsampling pour les coefficients de Fourier de la fonction d'autocorrélation a été présenté par [103]. Dans ce papier, l'auteur définit l'estimateur du subsampling suivant :

$$\hat{R}_{n,b,t}(\alpha, \tau) = \frac{1}{b} \sum_{k=1}^{b-\tau} X(t-1+k)X(t-1+k+\tau)e^{-j2\pi\alpha k} \quad (3.7)$$

Cependant, cet estimateur de subsampling n'est pas proprement défini par l'équation (3.7) et est mal conditionné. En effet, ce dernier ne prend pas en compte les indices temps des sous-échantillons. L'indice temps k est fixe quelque soit le sous-échantillon utilisé ($k = 1, \dots, b - \tau$). C'est pourquoi, nous proposons une version modifiée et corrigée de cet estimateur. Notre nouvel estimateur prend en compte les indices temps de chaque sous-échantillon de telle façon à respecter les points de départs et de fins de chaque sous-échantillon. Par conséquent, notre nouvel estimateur garde et conserve les structures temporelles des sous-échantillons utilisés dans le processus d'estimation. Ce nouvel estimateur est donné comme suit :

$$\hat{R}_{n,b,t}(\alpha, \tau) = \frac{1}{b} \sum_{k=t}^{t+b-1} X(k)X(k+\tau)e^{-j2\pi\alpha k} \quad (3.8)$$

Avec $t+b-1 < n-\tau$, une taille de bloc égale à b et $t \in \{1, \dots, n-b+1\}$. Ainsi, l'équation (3.8) donne l'estimateur de $R_X(\alpha, \tau)$ basé sur $(X(t), \dots, X(t+b-1))$. De plus, il est important de noter que $\hat{R}_{n,b,t}(\alpha, \tau)$ est un complexe et que ses parties réelles et imaginaires seront traitées.

Ainsi, à partir des $n + b - 1$ répliques de l'estimateur initial, une approximation valide de la distribution empirique du subsampling est obtenue et a la forme suivante :

$$L_{n,b}(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} \mathbf{1}\{\sqrt{b}(|\hat{R}_{n,b,t}(\alpha, \tau)| - |\hat{R}_n(\alpha, \tau)| \leq x)\} \quad (3.9)$$

telle que $\mathbf{1}\{\cdot\}$ représente la fonction indicatrice (ou caractéristique) de l'événement $\{\cdot\}$. Sous ces notations, le subsampling pour l'estimateur $|\hat{R}_n(\alpha, \tau)|$ du paramètre $|R(\alpha, \tau)|$ est consistant. De plus, des intervalles de confiance valides pour le paramètre $|R(\alpha, \tau)|$ sont obtenus et ont la forme suivante :

$$\left(\left| \hat{R}_n(\alpha, \tau) \right| - \frac{c_{n,b}(1 - \frac{\delta}{2})}{\sqrt{n}}, \left| \hat{R}_n(\alpha, \tau) \right| - \frac{c_{n,b}(\frac{\delta}{2})}{\sqrt{n}} \right),$$

tels que $\delta \in (0, 1)$ est le niveau de confiance, $c_{n,b}(1 - \frac{\delta}{2})$ et $c_{n,b}(\frac{\delta}{2})$ les quantiles dérivés à partir de la procédure de subsampling.

Dans ce qui suit, une description pratique de l'algorithme de la procédure du subsampling est présentée. En outre, quelques applications du MBB et du subsampling sur des signaux simulés ainsi qu'une étude comparative entre ces deux méthodes sont exposées.

3.4.2 Algorithme du subsampling

L'algorithme du subsampling pour l'estimation des coefficients de Fourier de la fonction d'autocorrélation et pour tester les fréquences cycliques est décrit ci-dessous. La procédure d'estimation est basée sur les échantillons disponibles $\{X(1), X(2), \dots, X(n)\}$ de $\{X(t) : t \in \mathbb{Z}\}$.

L'algorithme du subsampling est composé des différentes étapes suivantes :

Étape 1 : Estimer $\hat{R}_n(\alpha, \tau)$ basé sur la totalité de la longueur de l'échantillon n . L'estimation de $\hat{R}_n(\alpha, \tau)$ est réalisée suivant (3.2).

Étape 2 : Choisir une taille de bloc $b(< n)$, et un niveau de confiance $\delta \in (0, 1)$

Étape : Pour $t = 1, 2, \dots, n - b + 1$,

Utiliser (3.8) pour estimer $\hat{R}_{n,b,t}(\alpha, \tau)$, avec $\hat{R}_{n,b,t}(\alpha, \tau)$ les estimateurs basés sur les sous-échantillons.

Étape 4 : Définir

$$V_{ret}^{n,b,t} = \Re[\sqrt{b}(\hat{R}_{n,b,t}(\alpha, \tau) - \hat{R}_n(\alpha, \tau))], \text{ et } V_{imt}^{n,b,t} = \Im[\sqrt{b}(\hat{R}_{n,b,t}(\alpha, \tau) - \hat{R}_n(\alpha, \tau))]$$

tels que $V_{ret}^{n,b,t}$ and $V_{imt}^{n,b,t}$ soient respectivement la partie réelle et la partie imaginaire de la statistique définie juste au dessus. Ils sont de longueur égale à $n - b + 1$ correspondant aux $n - b + 1$ sous-échantillons.

Étape 5 : Ordonner les valeurs de $V_{ret}^{n,b,t}$ et $V_{imt}^{n,b,t}$ afin d'obtenir les quantiles réels et imaginaires $c_{n,b}(1 - \frac{\delta}{2})$ et $c_{n,b}(\frac{\delta}{2})$ dérivés à partir du subsampling.

Étape 6 : Définir les intervalles de confiance basés sur le subsampling pour les deux parties, réelle et imaginaire

$$\left(\hat{R}_n(\alpha, \tau) - \frac{c_{n,b}(1 - \frac{\delta}{2})}{\sqrt{n}}, \hat{R}_n(\alpha, \tau) - \frac{c_{n,b}(\frac{\delta}{2})}{\sqrt{n}} \right)$$

Ainsi, à partir de l'algorithme de la procédure du subsampling décrit ci-dessus, il devient assez simple de construire des tests statistiques pour la significativité des fréquences cycliques. En effet, si pour chaque α et τ , zéro appartient à l'intervalle de confiance, alors $R_X(\alpha, \tau)$ n'est pas significatif. Par conséquent, la fréquence α testée est dite non significative. Cependant, si pour un α donné et un τ donné, zéro n'appartient pas à l'intervalle de confiance, alors $R_X(\alpha, \tau)$ est significatif. Dans ce cas, la fréquence cyclique α est dite significative.

Dans ce qui suit, une étude de simulation consacrée à l'application du subsampling et du MBB ainsi qu'à l'investigation de leurs performances respectives est exposée. Une étude comparative entre les deux méthodes est réalisée. Enfin, l'approche la plus efficace et la plus robuste est gardée et est appliquée sur des signaux biomécaniques réels.

3.5 Etude de simulation et de performances

Dans cette partie, une étude de simulation est présentée. En effet, une investigation de l'efficacité ainsi que de la robustesse des estimations basées sur le MBB et le subsampling est réalisée. Les études ci-dessous sont effectuées en considérant différents rapports de signal à bruit (RSB). Les valeurs du RSB utilisées sont les suivantes : -5 dB, 0 dB, 10 dB et 20 dB.

Afin d'illustrer la manière dont le subsampling, le MBB et la construction de tests sont réalisés en pratique, l'attention est focalisée sur un modèle cyclostationnaire théorique. Soit $\{U_i(t), t \in \mathbb{Z}\}$, $i = 1, 2,$, deux processus aléatoires stationnaires, indépendants et de moyennes nulles. Soit $b(t)$ un bruit additif, gaussien et corrélé. Alors, le modèle cyclostationnaire étudié est la somme de deux modulations d'amplitudes, soit :

$$X(t) = U_1(t)\cos(2\pi f_1 t) + U_2(t)\cos(2\pi f_2 t) + b(t), \quad (3.10)$$

Les caractéristiques du signal simulé $X(t)$ sont les suivantes : un nombre d'échantillons (longueur des données) $n = 4096$ et les fréquences f_1 et f_2 sont égales respectivement à $0.1Hz$ et à $0.11Hz$.

L'ensemble des fréquences cycliques théorique est sous la forme suivante :

$$\mathcal{A}_\alpha = \{-2f_2, -2f_1, 0, 2f_1, 2f_2\},$$

Soit, numériquement :

$$\mathcal{A}_\alpha = \{-0.22, -0.2, 0, 0.2, 0.22\},$$

De plus, nous définissons le problème du test de la manière suivante :

Pour $\alpha \neq 0$

$$\begin{cases} H_0 & : R_X(\alpha, \tau) = 0 \\ H_1 & : R_X(\alpha, \tau) \neq 0. \end{cases}$$

Où H_0 et H_1 représentent les deux hypothèses de notre test.

Les figures 3.1a, b, c et d montrent les valeurs absolues de la fonction d'autocorrélation cyclique $R_X(\alpha, \tau)$, estimées respectivement pour des valeurs de RSB égales à -5dB , 0dB , 10dB et 20dB . Comme il a été mentionné auparavant, l'ensemble des fréquences cycliques est sous la forme $\mathcal{A}_\alpha = \{-0.22, -0.2, 0, 0.2, 0.22\}$. Étant donné que la fonction d'autocorrélation cyclique présente des propriétés de symétrie, nous nous intéressons que sur les parties positives des fréquences cycliques α .

Afin d'étudier la robustesse ainsi que l'efficacité des deux méthodes, les probabilités de couvrage (*actual coverage probability ACP* en anglais) et les moyennes des largeurs des intervalles de confiance sont calculées pour les deux fréquences cycliques positives $\alpha_1 = 2f_1$ et $\alpha_2 = 2f_2$, et ce, pour les différents RSB et pour différentes tailles de bloc b . L'ensemble des tailles (en nombre d'échantillons) de bloc b sont choisies de l'ensemble :

$$b \in \{10, 16, 25, 50, 64, 80, 100, 150, 200, 300, 500, 1000, 1500, 1800, 2000\}$$

Habituellement, on construit des intervalles de confiance pour un niveau visé $1 - \delta$, par exemple, pour une valeur de $\delta = 0.05$, le niveau de confiance nominal est égale à 95%. Cependant, la vraie probabilité de couverture (ACP) est différente et inconnue. Théoriquement, on devrait obtenir une probabilité de couverture d'au moins $1 - \delta$.

Les valeurs des ACP et des largeurs des intervalles de confiance sont calculées à partir de 1000 itérations ; i.e., un nouveau signal est simulé à chaque

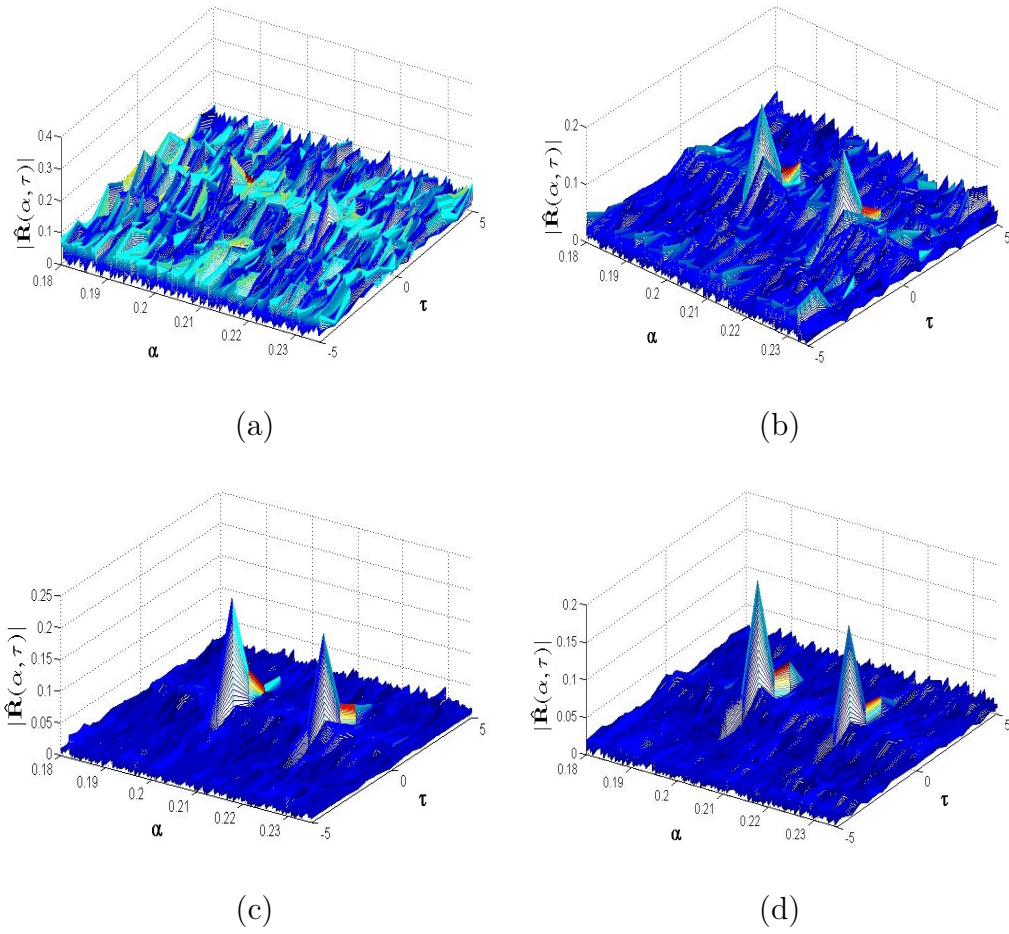


FIGURE 3.1 – Les valeurs estimées de $\hat{R}_X(\alpha, \tau)$ sur le signal simulé $X(t)$: (a). RSB = -5 dB ; (b). RSB = 0 dB ; (c). RSB = 10 dB ; (d). RSB = 20 dB.

itération, puis le test de l'hypothèse H_0 par rapport à l'hypothèse alternative H_1 est réalisé aux moyens du subsampling et du MBB. Le nombre de répétitions pour la procédure du MBB est fixé à $B = 1000$.

Durant toute cette étude de simulation, les intervalles de confiance calculés à partir des deux méthodes correspondent à un niveau de confiance $\delta = 0.05$. De plus, le paramètre de retard τ est fixé à 0 ; i.e., la coupe de $\hat{R}_X(\alpha, \tau)$ à $\tau = 0$. La motivation essentielle derrière le choix de $\tau = 0$ est purement pratique. En effet, cela nous permet de réduire le temps de calcul de manière très significative. De plus, le test statistique reste très valide car il suffit de montrer que pour en moins une valeur de $\alpha \neq 0$ et en moins une valeur (quelconque) de τ , si $R_X(\alpha, \tau)$ est significativement différent de zéro, alors le signal $X(t)$ est cyclostationnaire.

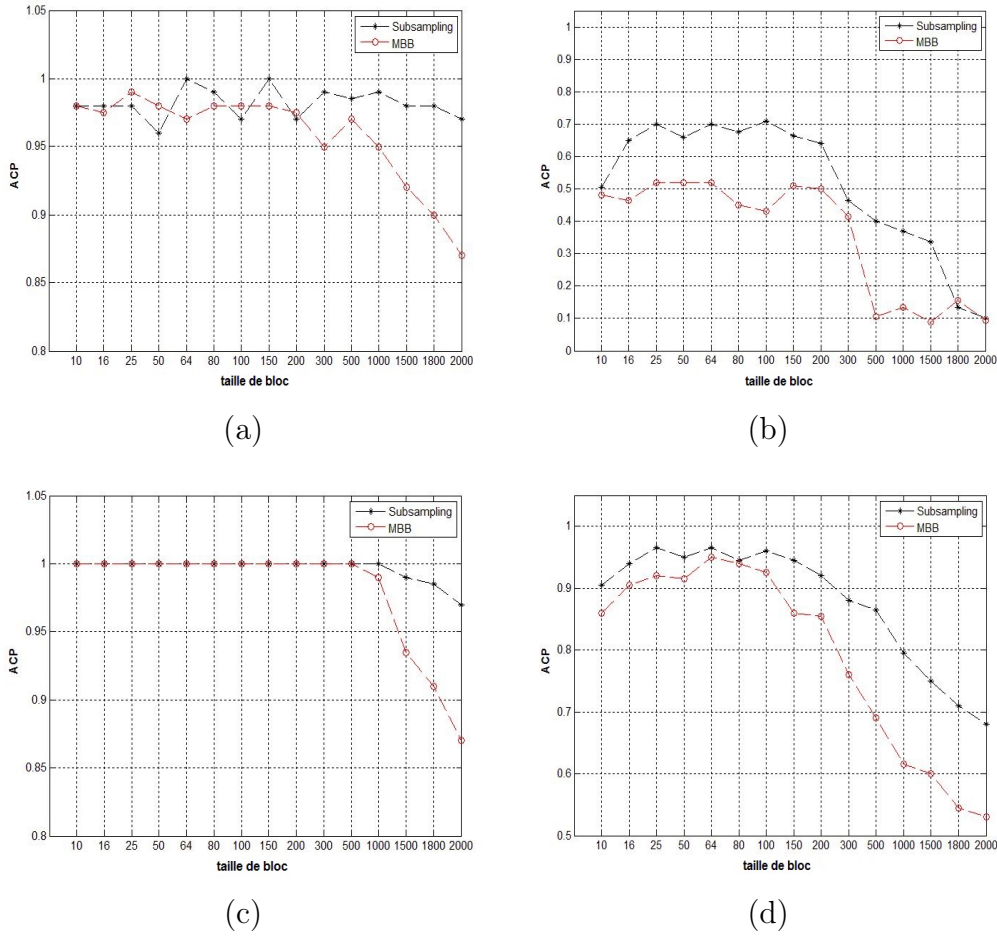


FIGURE 3.2 – Valeurs des ACP pour le subsampling et le MBB : (a) et (b). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = -5 dB. (c) et (d). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 0 dB.

Les figures 3.2 et 3.3 représentent les valeurs moyennes des ACP calculées pour les fréquences cycliques $\alpha_1 = 0.2Hz$ et $\alpha_2 = 0.22Hz$, et ce, pour les quatre différentes valeurs du RSB.

La lecture des résultats obtenus dans chaque cas de figure est montrée ci-dessous :

- Dans le cas où le RSB = -5dB (figures 3.2 a et b), les valeurs moyennes des ACP sont globalement hautes (> 95%) pour la partie réelle de $R_X(\alpha, \tau)$. Ce résultat est valide autant pour le subsampling que pour le MBB. Cependant, on peut remarquer que le subsampling semble être plus robuste dans la détection des deux fréquences cycliques α_1 et α_2 par rapport à longueur de bloc utilisé. En effet, Les valeurs moyennes

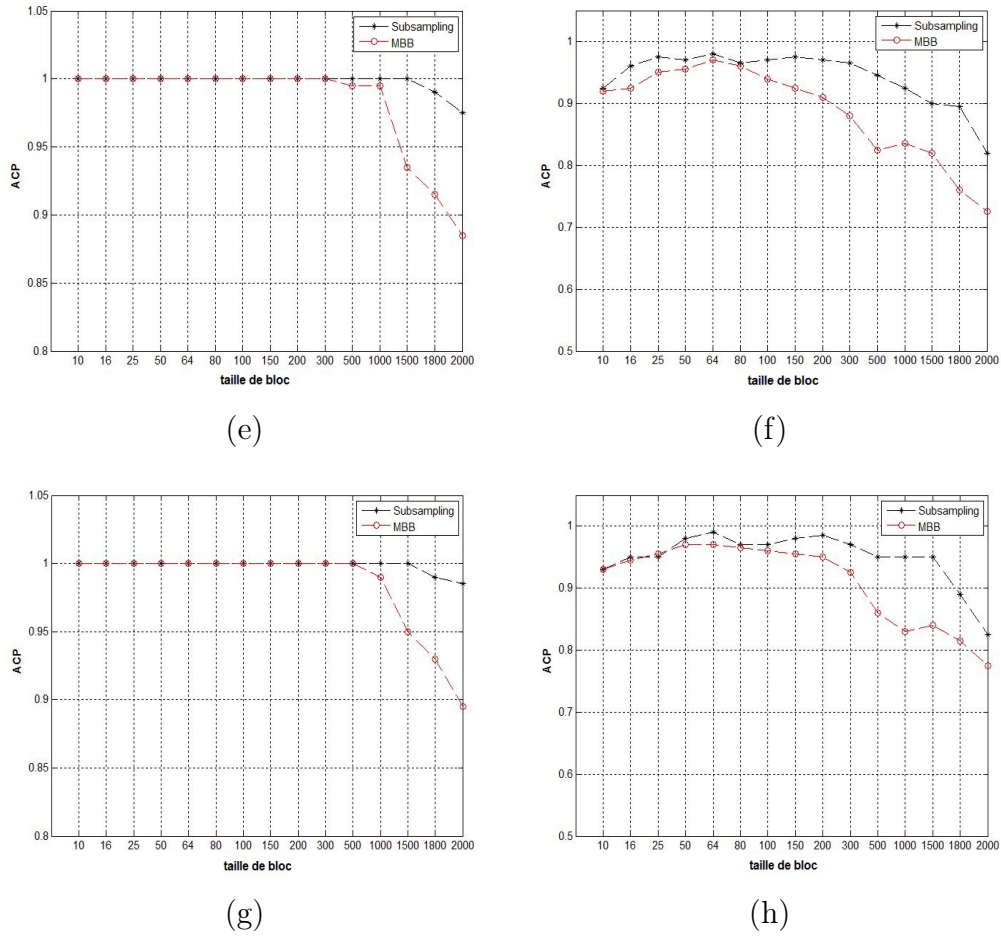
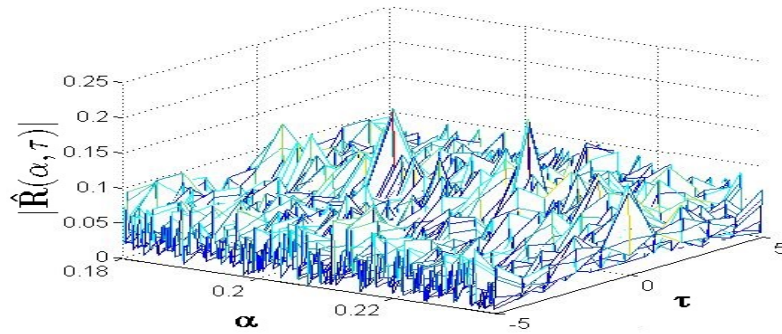


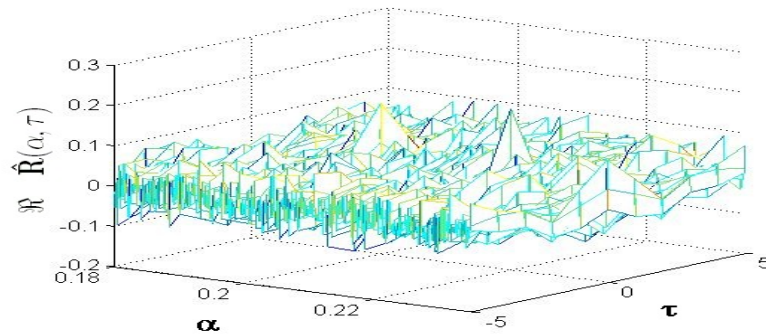
FIGURE 3.3 – Valeurs des ACP pour le subsampling et le MBB : (e) et (f). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 10 dB. (g) et (h). ACP pour les parties réelle et imaginaire de $\hat{R}_X(\alpha, \tau)$. RSB = 20 dB.

des ACP dans le cas du MBB chutent à partir de la longueur de bloc $b = 1000$.

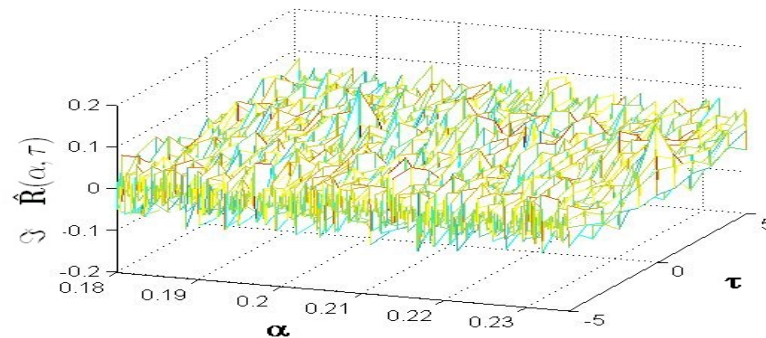
En revanche, les valeurs moyennes des ACP dans les deux méthodes sont relativement basses pour la partie imaginaire de $R_X(\alpha, \tau)$. Nous pouvons tout de même noter que ces valeurs sont plus hautes dans le subsampling comparé au MBB. Les valeurs basses des ACP pour la partie imaginaire de $R_X(\alpha, \tau)$ peuvent être dues au fait que les valeurs des parties imaginaires de l'estimateur pour les deux fréquences cycliques soient moins importantes que celles des parties réelles. En effet, il semblerait qu'à un RSB égal à -5 dB, les valeurs de la partie imaginaire de $R_X(\alpha, \tau)$ soient noyées dans le bruit (voir la figure 3.4)



(a)



(b)



(c)

FIGURE 3.4 – Valeurs de la fonction d'autocorrélation cyclique de $X(t)$ avec un RSB $= -5$ dB (itération numéro 1). (a). Le module de $\hat{R}_X(\alpha, \tau)$; (b). La partie réelle de $\hat{R}_X(\alpha, \tau)$; (c). La partie imaginaire de $\hat{R}_X(\alpha, \tau)$.

- Dans le cas où le RSB $= 0$ dB (figures 3.2 c et d), il s'avère que les résultats soient globalement améliorés. Les valeurs moyennes des ACP

sont très hautes ($>95\%$) pour la partie réelle de $R_X(\alpha, \tau)$, et ce, pour des valeurs de blocs comprises entre 10 et 1000. De plus, il semble que le subsampling soit toujours plus robuste et efficace comparé au MBB. En effet, à partir de la taille de bloc $b = 1000$, les valeurs des ACP dans le cas du MBB chutent très rapidement.

Concernant la partie imaginaire, on observe clairement une amélioration des résultats en comparaison avec le cas où le RSB était égal à -5dB . Cela dit, le subsampling semble encore être plus robuste à la détection des fréquences cyclique, comparé au MBB.

- Finalement, dans les cas où le RSB = 10dB et 20dB (figure 3.3), les valeurs moyennes des ACP sont très hautes pour la partie réelle de $R_X(\alpha, \tau)$. Cependant, le subsampling semble encore une fois plus robuste que le MBB. En outre, les valeurs moyennes des ACP semblent augmenter avec l'amélioration du ratio RSB pour la partie imaginaire de $R_X(\alpha, \tau)$. Par conséquent, l'estimation et la détection des fréquences cycliques sont améliorées. Néanmoins, il est noté que la méthode du subsampling paraît être plus robuste et plus efficace que la méthode du MBB. En effet, les résultats obtenus avec le subsampling montrent moins de dépendance au choix de la taille du bloc utilisé, contrairement au MBB.

Nous pouvons résumer les résultats obtenus dans cette première partie de simulation comme suit : (1) La robustesse et l'efficacité du subsampling et du MBB semblent être nettement meilleures avec l'amélioration des rapports signal à bruit et (2) Le subsampling est plus robuste à l'estimation et la détection des fréquences cycliques. De plus, il semble être moins sensible à la taille du bloc b choisie.

Une autre partie de cette simulation est dédiée à l'étude de la largeur des intervalles de confiance générés par les deux méthodes. Ainsi, les valeurs moyennes des largeurs des intervalles de confiance en fonction de la taille du bloc utilisée, et ce, pour le subsampling et le MBB sont représentés dans la figure 3.5.

Les résultats obtenus montrent que les largeurs des intervalles de confiance dérivés à partir du subsampling sont plus étroites en comparaison avec ceux obtenus à partir du MBB. Ce résultat semble être plus évident lorsque les longueurs des blocs deviennent de plus en plus importantes ($b > 500$). Cela peut expliquer les meilleurs résultats d'estimation et de détection de fréquences

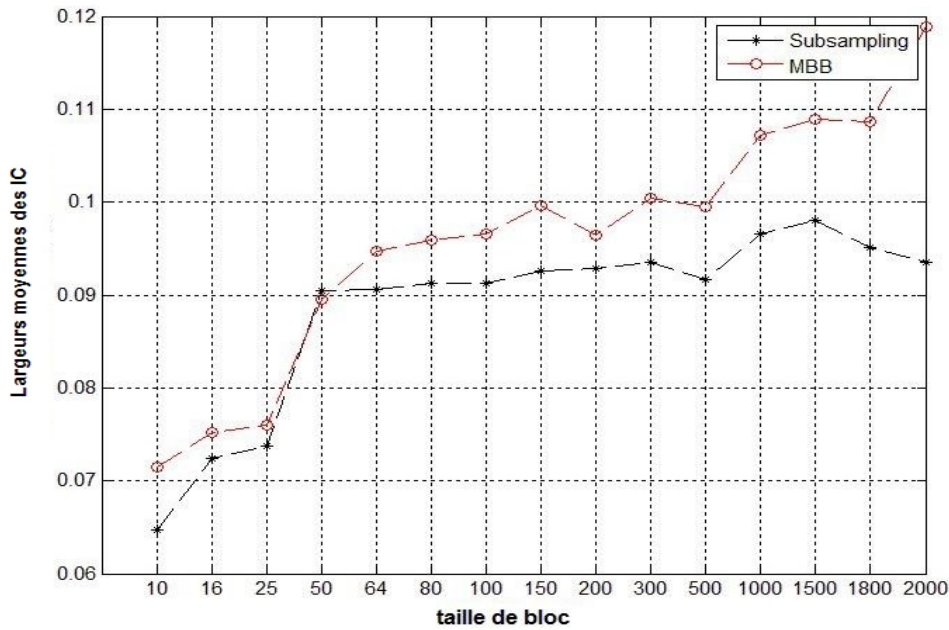


FIGURE 3.5 – Valeurs moyennes des largeurs des intervalles de confiance pour le subsampling et le MBB.

cycliques obtenus avec le subsampling lors de la première partie de simulation. En effet, plus les intervalles de confiance sont larges et plus ils peuvent contenir le zéro.

Ainsi, la comparaison entre le subsampling et le MBB a montré que le MBB était moins robuste et moins efficace que le subsampling dans l'estimation et la détection des fréquences cycliques. De plus, il est évident que le MBB est plus sensible à la taille du bloc utilisé. Par ailleurs, d'un point de vue purement pratique, la procédure du MBB est très couteuse en termes de temps de calcul par comparaison avec le subsampling. Cet inconvénient est dû au fait que pour chaque fréquence α et chaque retard τ à être testés, il y a un besoin de définir un nouveau vecteur complexe $U(t) = X(t)X(t + \tau)e^{-j2\pi\alpha t}$, puis appliquer l'algorithme du MBB, plus ses B répétitions. Par conséquent, cela rend l'exploitation du MBB dans le cas des signaux réels très contraignante.

Pour toutes les différentes raisons citées ci-dessus (théoriques et pratiques), nous avons opté pour l'utilisation de la méthode du subsampling pour traiter des signaux réels biomécaniques.

3.6 Application du subsampling sur des signaux biomécaniques

Malgré le nombre important de travaux de recherches dédiés à l'étude des phénomènes cyclostationnaires, les méthodes de ré-échantillonnage statistique, telle que le subsampling n'ont pas été appliquées aux problèmes d'identification et d'estimation de fréquences cycliques sur des signaux réels. Dans ce qui suit, des outils d'identification et d'estimation statistiques de fréquences cycliques significatives sont appliqués afin d'analyser et de caractériser des signaux réels issus de la biomécanique. De plus, des indicateurs basés sur la méthode du subsampling sont proposés et utilisés pour l'analyse et le traitement de ces signaux.

L'analyse des troubles de la locomotion humaine, en particulier dans les cas de fatigue musculaire ou osseuse peut apporter des informations utiles concernant le diagnostic clinique, l'évaluation des gestes sportifs, la rééducation, etc. A cette fin, nous proposons de nouvelles approches basées sur le subsampling afin d'étudier et analyser les signaux d'un coureur professionnel d'ultra marathon. Ces signaux ont été acquis sur un tapis de course expérimental. Le coureur a été sujet à un test de course continue de 24 heures.

Les signaux acquis sont appelés *forces de réaction au sol* (*Ground Reaction Forces (GRF)* en anglais). Le traitement de ces signaux comprend l'identification des caractéristiques de la variabilité liée à la fatigue. La fatigue du coureur peut être associée à différentes signatures fréquentielles des caractéristiques du second ordre des signaux de GRF.

3.6.1 La description des signaux de GRF

Les acquisitions des signaux de GRF sont réalisées sur un tapis roulant dynamométrique qui permet la mesure des forces de réactions au sol [18]. Le tapis est équipé de 12 capteurs placés sur ses 4 coins (voir la figure 3.6). Les données sont acquises : avant la fatigue ; i.e., au début de la course ($T = 0h$), après 12 heures de course ($T = 12h$) et après 24 heures de course ($T = 24h$). Les signaux acquis sont d'une durée de 20s chacun et la fréquence d'échantillonnage des signaux est fixée à 1000Hz. Enfin, la vitesse du tapis est maintenue à 10km/h.

Un signal de GRF est comme étant un ensemble d'impacts et de forces de propulsion consécutives des deux jambes. En effet, un tel signal représente les pas du coureur. Un pas est défini comme étant le temps entre le contact du premier pied avec le sol et le contact du second pied avec le sol. La foulée

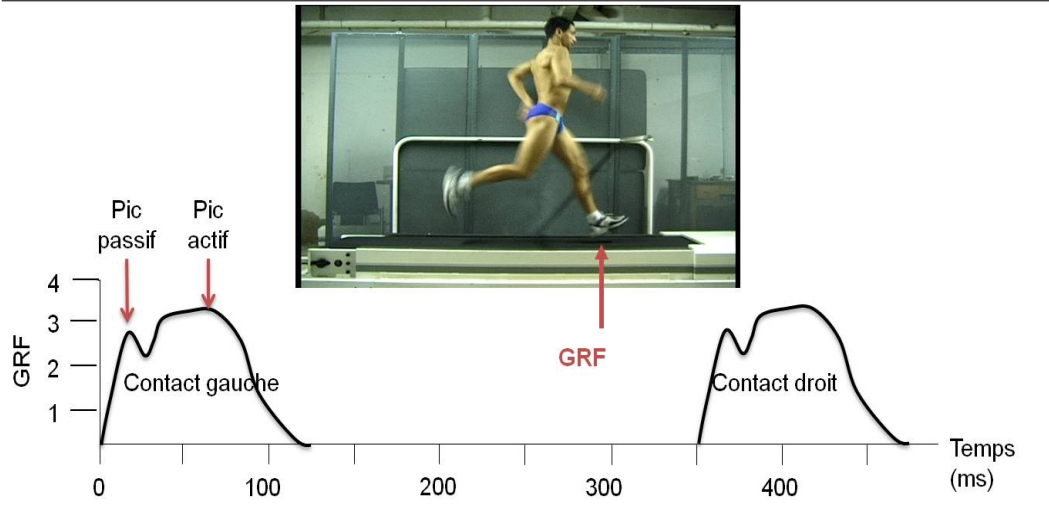


FIGURE 3.6 – Tapis roulant dynamométrique.

correspondra à deux pas consécutifs. Aussi, il est important de noter qu'un pas est constitué de 2 phases : une phase d'impact (pic passif) et une phase de propulsion (pic actif). La figure 3.7 montre un exemple d'un signal de GRF.

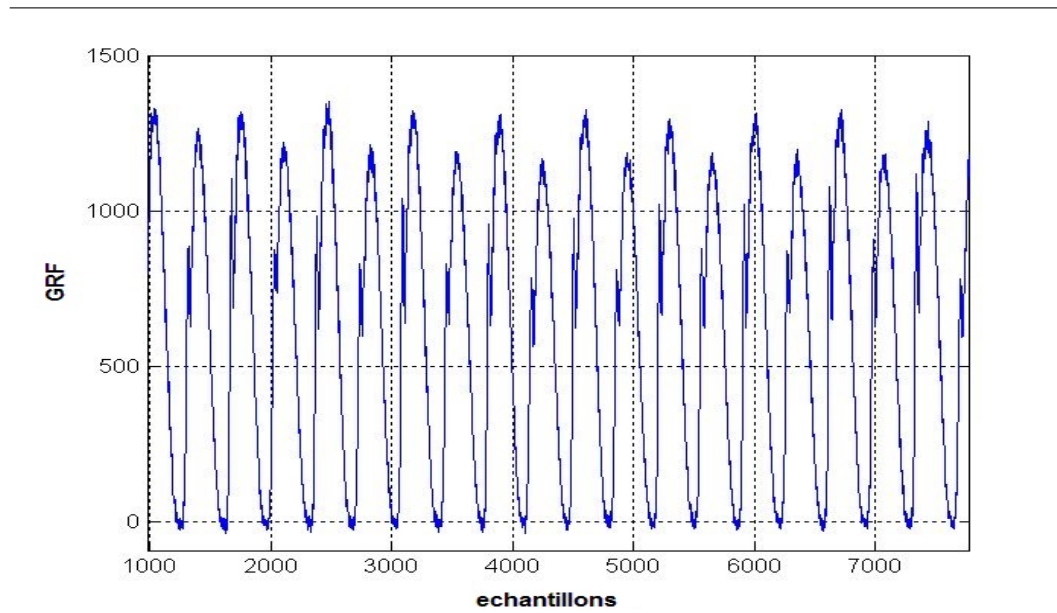


FIGURE 3.7 – Exemple d'un signal de GRF

Ainsi, l'objectif de cette étude expérimentale est d'identifier et caractériser

3.6. Application du subsampling sur des signaux biomécaniques 93

l'état de fatigue du coureur au cours du temps, et ce, en utilisant une approche basée sur la méthode du subsampling.

3.6.2 Pré-traitement des signaux de GRF

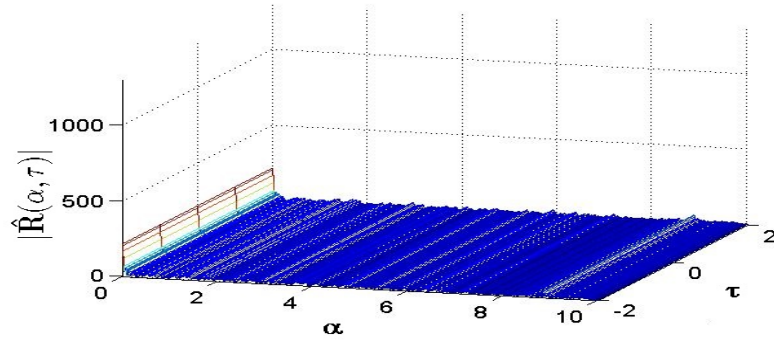
Dans certains travaux de recherche [131, 23], les auteurs ont montré le caractère cyclostationnaire de ce type de signaux. La première étape avant l'utilisation de l'approche du subsampling est très essentielle. En effet, le pré-traitement des signaux de GRF par le biais de la méthodologie exposée dans le chapitre 4. Les signaux ont tout d'abord été synchronisés, puis le GSBB utilisé afin de vérifier la séparation entre les composantes cyclostationnaires à l'ordre un et à l'ordre deux. L'intérêt étant sur l'étude de la composante d'ordre deux (les composantes résiduelles), nous nous sommes focalisés sur cette dernière dans la suite du travail.

3.6.3 Analyse cyclostationnaire basé sur le subsampling

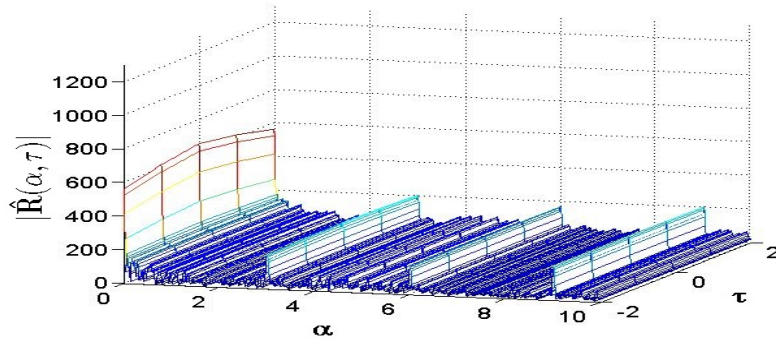
Les figures 3.8a, b et c montrent les valeurs estimées des fonctions d'autocorrélation cyclique $\hat{R}_X(\alpha, \tau)$, respectivement pour $T = 0h$, $T = 12h$ et $T = 24h$. Une analyse préliminaire de ces figures suggèrent que les composantes résiduelles des signaux provenant du sujet à $T = 12h$ et $T = 24h$ montrent un caractère cyclostationnaire. Afin de vérifier ces résultats et tester la présence de structures cyclostationnaires dans les trois signaux, la méthode du subsampling est appliquée. Les paramètres du subsampling utilisés sont les suivants : une taille de bloc $b = 1000$ échantillons et un niveau de confiance $\delta = 0.05$ correspondant à 95% de probabilité de signification. Le choix de la taille du bloc est conditionné par la taille de la période possible (ou des périodes possibles) de la fonction d'autocorrélation du signal. En effet, la longueur du bloc est choisie supérieure à la taille de la période la plus élevée de l'ensemble des périodes possibles. Dans le cas de l'étude des signaux de GRF, les périodes possibles de la fonction d'autocorrélation sont les périodes cycliques de la foulée et du pas qui sont respectivement égales à 666.67 et 333.33 échantillons.

Les résultats de l'analyse via le subsampling pour le test de l'hypothèse H_0 contre l'hypothèse alternative H_1 des trois signaux à $T = 0h$, $T = 12h$ et $T = 24h$ sont représentés dans la figure 3.9

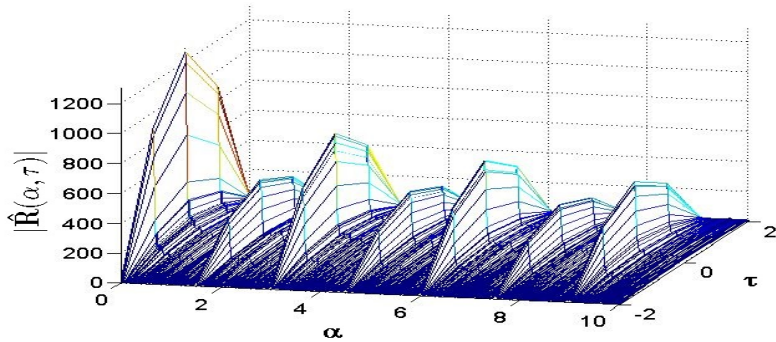
Les figures 3.9a et b montrent clairement la non détection de fréquences cycliques (non nulles) pour les deux parties, réelle et imaginaire de l'estimateur de $R_X(\alpha, 0)$. Cela signifie le non rejet de l'hypothèse H_0 . Par conséquent,



(a) $T = 0h$



(b) $T = 12h$



(c) $T = 24h$

FIGURE 3.8 – Les valeurs estimées de $\hat{R}_X(\alpha, \tau)$ sur les résiduels provenant des signaux de GRF : (a). $T = 0h$; (b). $T = 12h$; (c). $T = 24h$.

on peut affirmer que le signal de GRF du coureur à $T = 0h$ ne montre pas de caractère cyclostationnaire. Les figures 3.9c et d révèlent le succès du subsampling à détecter la fréquence cyclique $\alpha = 3Hz$ et ses harmoniques. Alors,

l'hypothèse H_0 est rejeté dans ce cas et le signal à $T = 12h$ présente une structure cyclostationnaire. La fréquence cyclique $\alpha = 3Hz$ correspond à la fréquence du pas du coureur. Enfin, les figures 3.9e et f montrent nettement le succès du subsampling à détecter deux fréquences cycliques $\alpha_1 = 1.5Hz$ et $\alpha_2 = 3Hz$ ainsi que leurs harmoniques respectives. Cela signifie que l'hypothèse H_0 est encore une fois rejetée pour $T = 24h$. Par conséquent, le signal de GRF à $T = 24h$ est cyclostationnaire. La fréquence cyclique $\alpha_1 = 1.5Hz$ correspond à la fréquence de foulée du coureur.

Puisque l'analyse des par le subsampling a permis de tester avec succès le caractère cyclostationnaire des signaux sous étude, quelques indicateurs issues du subsampling sont présentés dans ce qui suit. L'objectif étant de mieux caractériser et décrire les trois différents états de fatigue du coureur.

3.6.4 Cartographies de significativité fréquentielle cyclostationnaire

Afin de décrire plus exactement le comportement du coureur au cours du temps, des cartographies de significativité fréquentielle cyclostationnaire sont construites à partir du subsampling. Ces cartographies combinent les informations extraites de l'estimée de $\hat{R}_X(\alpha, \tau)$, du paramètre de retard τ et du niveau de confiance δ .

Soit $\Delta = \{\delta_1, \dots, \delta_k\}$, avec $k \in \mathbb{N}$. Les cartographies sont alors construites de la façon suivante : pour chaque niveau de confiance δ_i de Δ , avec $i = 1, \dots, k$, la méthode du subsampling est appliquée, ensuite, seulement les niveaux de confiances les plus élevés sont gardés. Cette démarche est réalisée pour chaque fréquences cycliques testée α et pour chaque retard τ de $\hat{R}_X(\alpha, \tau)$. Pour être plus précis, nous définissons numériquement l'ensemble des niveaux de confiance $\Delta = \{0.05, \dots, 0.0001\}$, tels que $\delta = 0.05$ correspond à 95% de significativité et $\delta = 0.0001$ correspond à 99.99% de significativité. L'objectif étant d'associer à chaque α et à chaque τ de $\hat{R}_X(\alpha, \tau)$ leur plus haut degrés de significativité.

Les cartographies de la figure 3.10 donnent une description complète des trois différents états de fatigue du coureur. En effet, les fréquences cycliques α , le paramètre de retard τ ainsi que le niveau de confiance δ sont pris en considération dans la construction de ces cartographies.

Les figures 3.10a et b confirment bien la non détection de fréquences cycliques, non seulement pour $\tau = 0$, mais pour toutes les valeurs du paramètre de retard testés. Dans les figures 3.10c et d, les cartographies de significativité fréquentielle indiquent une fréquence cyclique significative avec ses harmo-

3.6. Application du subsampling sur des signaux biomécaniques 97

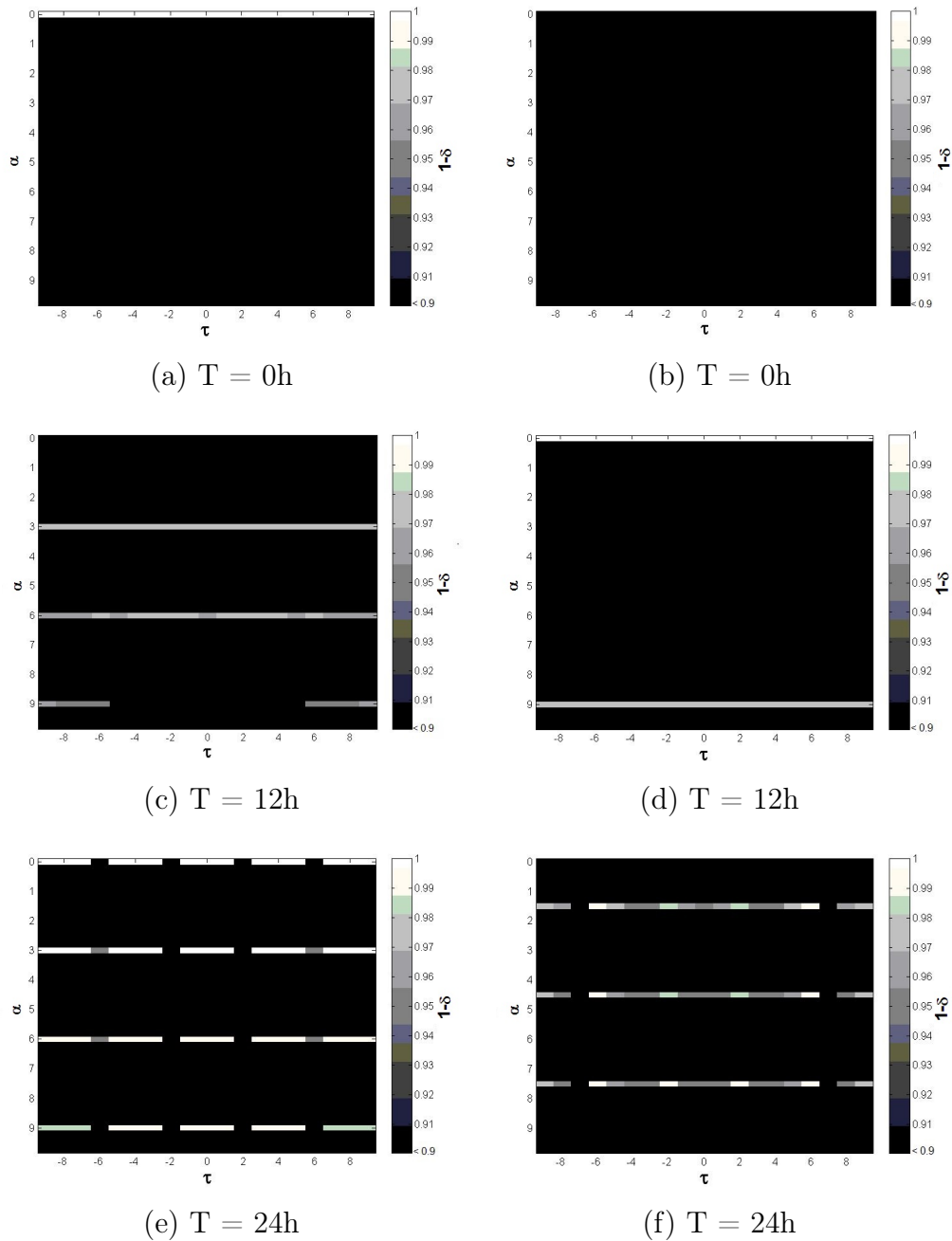


FIGURE 3.10 – (a),(c) et (e). Parties réelles des cartographies de significativité fréquentielle de $\hat{R}_X(\alpha, \tau)$, respectivement pour $T=0$; 12 et 24h. (b), (d) et (f). Parties imaginaires des cartographies de significativité fréquentielle de $\hat{R}_X(\alpha, \tau)$ respectivement pour $T=0$; 12 et 24h. En fonction de α , τ et δ .

riques. Les degrés de significativité de ces fréquences sont très hauts. Enfin,

les figures 3.10e et f révèlent bien la présence de deux fréquences significatives, ainsi que leurs harmoniques respectives. Les degrés de significativité des ces fréquences semblent être beaucoup plus élevés que pour $T = 12h$. En effet, la fréquence cyclique $\alpha_2 = 3Hz$ et ses harmoniques à $T = 24h$ semblent avoir des degrés de significativité plus importants qu'à $T = 12h$. De plus, on note l'émergence de la fréquence cyclique $\alpha_1 = 1.5Hz$ et ses harmoniques, ces dernières ayant des degrés de significativité élevés. Aussi, nous remarquons que la fréquence cyclique $\alpha_2 = 3Hz$ (fréquence du pas) et ses harmoniques semblent être plus significatives que la fréquence cyclique $\alpha_1 = 1.5Hz$ (fréquence de la foulée) et ses harmoniques à $T = 24h$.

3.6.5 La comparaison entre les trois niveaux de fatigue du coureur

L'objectif de cette partie est la comparaison entre les différents états de fatigue du coureur au cours du temps. Pour ce faire, nous exploitons le subsampling afin de construire des *matrices de probabilité*. En effet, pour un niveau de confiance donné $\delta = 0.05$, pour chaque α et chaque τ , la probabilité que $\hat{R}_X(\alpha, \tau)_{T_{12}}$ provenant des sous-échantillons estimés à $T = 12h$ soit supérieur à $\hat{R}_X(\alpha, \tau)_{T_0}$, provenant des sous-échantillons estimés à $T = 0h$ est approximée comme suit :

Pour chaque α et chaque τ :

$$P(\hat{R}_X(\alpha, \tau)_{T_{12}} > \hat{R}_X(\alpha, \tau)_{T_0}) = \frac{1}{n - b + 1} \# \{t : \hat{R}_{n,b,t}(\alpha, \tau)_{T_{12}} > \hat{R}_{n,b,t}(\alpha, \tau)_{T_0}\}$$

avec $t = 1, \dots, n - b + 1$ et $\#$ la fonction indicatrice (caractéristique) de l'ensemble $\{.\}$; i.e., sur les $n - b + 1$ cas, le nombre de fois où la proposition $\{.\}$ est vérifiée ($\leq n - b + 1$).

De la même façon, nous définissons la probabilité que $\hat{R}_X(\alpha, \tau)_{T_{24}}$ provenant des sous-échantillons estimés à $T = 24h$ soit supérieur à $\hat{R}_X(\alpha, \tau)_{T_{12}}$, provenant des sous-échantillons estimés à $T = 12h$:

$$P(\hat{R}_X(\alpha, \tau)_{T_{24}} > \hat{R}_X(\alpha, \tau)_{T_{12}}) = \frac{1}{n - b + 1} \# \{t : \hat{R}_{n,b,t}(\alpha, \tau)_{T_{24}} > \hat{R}_{n,b,t}(\alpha, \tau)_{T_{12}}\}$$

Les matrices de probabilité de la figure 3.11 montrent très clairement que la probabilité que $\hat{R}_X(\alpha, \tau)_{T_{12}}$ provenant des sous-échantillons estimés à $T = 12h$ soit supérieur à $\hat{R}_X(\alpha, \tau)_{T_0}$, provenant des sous-échantillons estimés

3.6. Application du subsampling sur des signaux biomécaniques 99

à $T = 0h$ est très importante. De même, la probabilité que $\hat{R}_X(\alpha, \tau)_{T_{24}}$ provenant des sous-échantillons estimés à $T = 24h$ soit supérieur à $\hat{R}_X(\alpha, \tau)_{T_{12}}$, provenant des sous-échantillons estimés à $T = 12h$ est très importante aussi. Ces matrices de probabilité construites à partir du subsampling apportent une analyse complémentaire à celle des cartographies fréquentielles. En effet, elles permettent de confirmer les résultats obtenus. De plus, elles quantifient les différences entre les états de fatigue en associant des probabilités de différence.

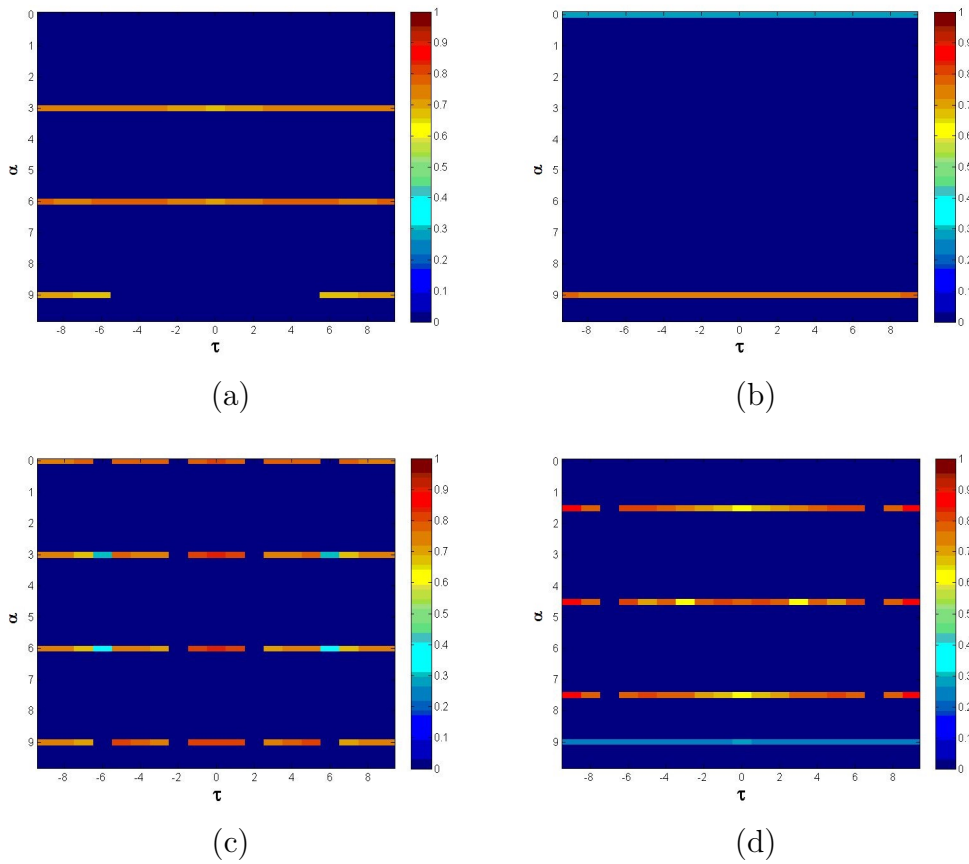


FIGURE 3.11 – (a) et (b). Matrices de probabilité estimées de $\hat{R}_X(\alpha, \tau)_{T_{12}} > \hat{R}_X(\alpha, \tau)_{T_0}$ respectivement pour les parties réelle et imaginaires . (c) et (d). Matrices de probabilité estimées de $\hat{R}_X(\alpha, \tau)_{T_{24}} > \hat{R}_X(\alpha, \tau)_{T_{12}}$ respectivement pour les parties réelle et imaginaires. Les matrices sont seuillées à un niveau de confiance $\delta = 0.05$.

Ainsi, tous les résultats obtenus lors de l'analyse des signaux de GRF suggèrent qu'en cas de fatigue, il y a émergence d'une fréquence cyclique correspondant à la fréquence du pas. Ceci est du à l'apparition de fluctua-

tions de pas quand le coureur commence à sentir la fatigue. En outre, les degrés de significativité de la fréquence cyclique correspondant au pas et ses harmoniques deviennent de plus en plus importants avec l'augmentation du niveau de fatigue du coureur. Par ailleurs, nous avons constaté la présence de fréquences cycliques correspondant à la fréquence de la foulée et ses harmoniques à $T = 24h$. Cela implique une variabilité au niveau des foulées du coureur quand l'état de fatigue est très important. Enfin, les degrés de significativité de la fréquence cyclique à 3Hz et de ses harmoniques semblent être plus élevés que ceux à 1.5Hz et ses harmoniques. Cela signifie qu'il existe plus de variabilité au niveau du pas qu'au niveau de la foulée (des fluctuations de pas en pas plus importantes que celles de foulée en foulée).

3.7 Conclusion

Dans ce chapitre, nous avons vu et démontré que les méthodes de ré-échantillonnage statistiques, telles que le subsampling et le MBB simplifiaient significativement l'inférence pour les processus cyclostationnaires. Cela a permis la construction de tests statistiques valides avec des intervalles de confiance pour l'estimateur de $R_X(\alpha, \tau)$. L'applicabilité de ces deux méthodes sur des signaux de simulation avec différents rapports de signal à bruit a été démontrée. La comparaison entre les deux approches a permis de mettre en évidence que le subsampling semblait être plus robuste au bruit, moins sensible au choix de la taille de bloc utilisé b et plus efficace pour l'estimation et la détection des fréquences cycliques α . De plus, le subsampling semble être plus dédié à l'analyse pratique des signaux réels.

La méthode du subsampling a aussi permis une analyse complète et innovante de différents états de fatigue d'un coureur expérimenté. La méthode du subsampling, ainsi que les cartographies de significativité fréquentielle et les matrices de probabilités présentées dans ce chapitre peuvent être d'une grande utilité dans d'autres domaines d'application, y compris, la mécanique et la surveillance des machines tournantes. En effet, il est bien connu que beaucoup de défauts mécaniques apparaissent sur les propriétés statistiques du second ordre des signaux étudiés. C'est pourquoi, les différentes approches basées sur le subsampling présentées dans ce chapitre pourraient ouvrir de nouvelles perspectives pour l'analyse des signaux issus des machines tournantes. En effet, des cartographies de significativité fréquentielle pourraient être obtenues pour différents défauts (exemple : les différents défauts de roulements) et la comparaison entre ces défauts quantifiée par les matrices de probabilité propo-

sées dans ce chapitre. Enfin, l'étude du degrés de signification des fréquences cycliques pourrait être un outil très intéressant et utile pour fixer des seuils critiques (ou alarmes) dans la surveillance des machines tournantes.

Inférence statistique temporelle des signaux cyclostationnaires par la méthode du Bloc Bootstrap Généralisé

Sommaire

4.1 Introduction	105
4.2 Présentation de la problématique	108
4.2.1 Définitions basiques	108
4.2.2 Séparation de composantes CS1/CS2	109
4.3 Le GSBB temporel pour l'estimateur de la moyenne synchrone	113
4.3.1 Algorithme du GSBB	114
4.3.2 Les intervalles de confiance GSBB pour la moyenne syn- chrone	116
4.3.3 Étude de simulation	116
4.4 Introduction	125
4.5 Méthodologie	126
4.5.1 Acquisition des données	126
4.5.2 Fluctuations de vitesse et description des signaux de marche	127
4.5.3 Le ré-échantillonnage synchrone	128
4.5.4 La séparation CS1/CS2	133
4.5.5 Mesure de la qualité de l'estimation	134
4.5.6 Indicateur de séparabilité CS1/CS2 basé sur le GSBB	134
4.6 Résultats	135
4.6.1 Le ré-échantillonnage synchrone	136

**Inférence statistique temporelle des signaux cyclostationnaires par
104 la méthode du Bloc Bootstrap Généralisé**

4.6.2	Analyse cyclostationnaire d'ordre deux CS2	136
4.6.3	Estimation de CS1 par GSBB	139
4.6.4	Indicateur de séparabilité CS1/CS2	141
4.7	Application à l'analyse de la marche en double tâche	143
4.7.1	Le paradigme expérimental de la marche avec double tâche	144
4.7.2	Analyse des paramètres temporels de la foulée	145
4.7.3	Analyse d'ordre deux des signaux de marche avec double tâche	147
4.7.4	Étude de groupe	148
4.8	Conclusion	153

4.1 Introduction

Dans ce chapitre, on s'intéresse à l'estimation des composantes CS1 et CS2 dans le domaine temporel. Cette problématique a été très peu développée dans la littérature scientifique car dans le domaine des statistiques ou bien des télécommunications, les signaux d'intérêts sont supposés centrés avec une composante CS1 nulle. Dans le cas des signaux mécaniques ou biomécaniques, cette composante CS1 est présente et est loin d'être nulle. En effet, elle peut représenter jusqu'à 90% de l'énergie du signal, voire beaucoup plus. Il est donc indispensable dans ce cas de séparer les composantes CS1 et CS2 pour les analyser indépendamment. Ces composantes ayant des natures différentes peuvent révéler et apporter des informations sur l'état d'un système donné, il est donc nécessaire de traiter cette issue avec le plus de précaution possible, d'autant plus qu'en pratique, les signaux CS présentent souvent des fluctuations de vitesse. Ces fluctuations sont dues à la nature non *purement* cyclique de la plupart des phénomènes appelés *périodiques*, qu'ils soient naturels ou artificiels.

Dans quelques domaines d'application tels que la mécanique et la biomécanique, une première étape fondamentale est l'estimation des composantes périodiques et déterministes. En effet, l'estimation de la CS2 est conditionnée par une estimation robuste de la CS1. Pour ce faire, une méthodologie combinant entre des outils de traitement du signal et de la théorie du bootstrap est développée dans ce qui suit.

Ainsi, ce chapitre est composé de deux parties principales. La première partie est consacrée à une nouvelle méthode de bootstrap appelée le *Generalized Seasonal Block Bootstrap* (GSBB). Cette méthode est utilisée dans le cadre de l'estimation de la CS1 d'un signal CS dans le domaine temporel. Dans cette partie, la formulation du problème est d'abord présentée. Ensuite, la méthode du GSBB est décrite et son algorithme est proposé. Enfin, une étude de simulation est réalisée afin d'illustrer la performance du GSBB. Dans la deuxième partie de ce chapitre, différentes applications du GSBB sur un signal réel issu de la marche à pied sont présentées. Le dispositif d'acquisition du signal est présenté. Ensuite, une méthode de ré-échantillonnage synchrone du signal est proposée et une étude comparative via le GSBB est développée afin de mettre en évidence l'influence de la présence de fluctuations de vitesse sur les propriétés CS du signal. Des indicateurs de séparabilité CS1/CS2 basés sur le GSBB sont aussi proposés pour mesurer la performance de la séparation CS1/CS2. Enfin, une étude de la marche à pied en condition de double tâche d'une population de 345 sujets ainsi que les résultats relatifs sont exposées.

Première partie

Estimation de la cyclostationnarité d'ordre 1 par la méthode du bloc bootstrap généralisé

4.2 Présentation de la problématique

4.2.1 Définitions basiques

Dans cette partie, nous donnons les principes du modèle cyclostationnaire des signaux rencontrés dans de nombreux domaines d'application.

La nature cyclique et périodique est omniprésente dans différents phénomènes physiques. Cependant, des perturbations aléatoires qui peuvent être dues à différents facteurs viennent impacter ces phénomènes cycliques. Par conséquent, la nature des données issues de tels phénomènes sera un couplage et une somme entre une grandeur périodique et une grandeur aléatoire. Afin d'éclaircir ces propos, un modèle basique est proposé.

Soit $X = \{X(t) : t \in \mathbb{Z}^+\}$ un signal modélisant le type de données citées ci-dessus. Le point de départ sera la décomposition de $X(t)$ en trois parties, soit :

$$X(t) = P(t) + V(t) + b(t)$$

Avec $P(t)$ un signal CS1, i.e., pour lequel la moyenne $\mu(t)$ est périodique dans le temps avec la période T :

$$\mu(t) = \mathbb{E}[X(t)] = \mathbb{E}[X(t+T)] = \mu(t+T)$$

Le signal $P(t)$ est appelé la composante d'ordre un du signal $X(t)$ et a une moyenne périodique, i.e., il est composé de fréquences déterministes, ce qui signifie que $P(t)$ est composée entièrement de composantes sinusoïdales.

Le signal $V(t)$ est aléatoire, il est appelé la composante du second ordre ou (CS2) du signal $X(t)$. Les moments d'ordre deux de cette composante sont périodiques dans le temps, en particulier sa fonction d'autocorrélation, soit :

$$\begin{aligned} R_V(t, \tau) &= \mathbb{E}[V(t+\tau)V(t)] = R_V(t+T, \tau+T) \\ &= \sum_{\alpha \in \mathcal{A}_\alpha} R_V(\alpha, \tau) e^{j2\pi\alpha t} \end{aligned}$$

Avec $R_V(\alpha, \tau)$ les coefficients de Fourier de la fonction d'autocorrélation de $V(t)$. Ces coefficients représentent la fonction d'autocorrélation cycliques comme nous l'avons vu dans le chapitre 1.

Il est donc très important de souligner qu'il existe une différence fondamentale entre les natures de ces deux composantes. En effet, $P(t)$ est de nature purement déterministe, contrairement à $V(t)$, qui elle est de nature purement aléatoire (stochastique).

Enfin, $b(t)$ est un bruit additif aléatoire, stationnaire et indépendant de $V(t)$.

Ainsi, la problématique de l'estimation et de la séparation entre les différentes composantes (CS1 et CS2) qui forment ce type de signaux est fondamentale en pratique. Dans ce qui suit, l'intérêt est porté sur ce type de problème.

4.2.2 Séparation de composantes CS1/CS2

Les différentes composantes qui forment un signal CS sont généralement liées à des sources physiques faisant partie du système analysé. De ce fait, la séparation entre les contributions déterministes (CS1) et aléatoires (ou CS2) est essentielle afin de mener à bien l'analyse d'un système donné. Nous pouvons citer à titre d'exemple la séparation entre des signaux issus d'engrenages et des signaux issus de roulement défectueux dans le domaine de la surveillance des machines tournantes. Un autre exemple est celui de séparer entre des composantes déterministes et des composantes aléatoires dans le domaine de la biomécanique et l'étude de la locomotion humaine.

Il existe peu de travaux relatifs à la séparation des composantes CS1 et CS2. Néanmoins, on en trouve quelques uns dans le domaine de l'analyse vibratoire. Dans ce qui suit, nous présentons quelques unes de ces méthodes ainsi que leurs brèves descriptions.

Suppression de bruit adaptative (Adaptive Noise Cancellation (ANC) en anglais :

La méthode *ANC* est une méthode où un signal contenant deux composantes non corrélées peuvent être séparées. Cela est réalisé par le biais de l'utilisation d'un signal de référence contenant seulement l'une des deux composantes. Le signal de référence ne doit pas nécessairement être identique à la partie correspondante du signal préliminaire. Cependant, il est nécessaire que ces deux derniers soient liés par une fonction de transfert linéaire. La méthode du *ANC* identifie cette fonction de transfert par un algorithme adaptatif, puis soustrait le signal de référence filtré du signal préliminaire laissant ainsi l'autre composante [140]. Le filtre adaptatif ajuste ses paramètres dans le but de minimiser la variance du signal d'erreur ε (la différence). Ensuite, et puisque les deux composantes sont supposées non corrélées, la variance du signal total sera donc la somme des deux variances des composantes. Ainsi, la séparation sera achevée lorsque la variance du signal de différence est réduite au minimum.

La figure 4.1 montre le schéma de la méthode ANC.

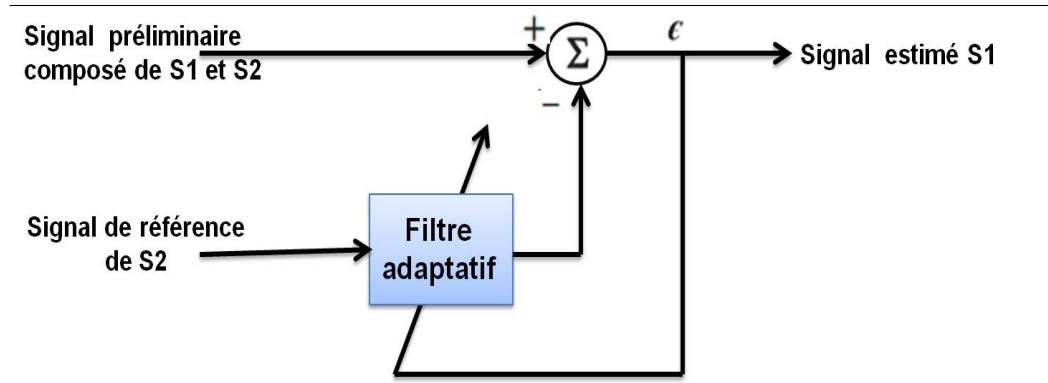


FIGURE 4.1 – Diagramme schématique de la méthode ANC.

Suppression de bruit auto-adaptative (Self-Adaptive Noise Cancellation (SANC) en anglais :

Dans le cas où l'une des deux composantes à séparer est déterministe et la seconde est aléatoire, il est possible de prendre une version décalée du signal préliminaire comme signal de référence. En effet, si le retard est choisi de telle sorte qu'il soit plus long que la longueur de la corrélation du signal aléatoire, alors le filtre adaptatif identifiera la fonction de transfert entre la partie déterministe du signal et sa version décalée. Ainsi, la séparation est réalisée en utilisant seulement un signal [140]. Le filtre adaptatif utilisé dans cette méthode est récursif avec un nombre de pondérations à déterminer. De plus, l'algorithme récursif utilisé est souvent le *LMS* (Least Mean Squares) avec des paramètres à choisir avec précaution. La figure 4.2 montre le schéma de la méthode SANC.

Prédiction linéaire :

La méthode de prédiction linéaire consiste à obtenir un modèle de la partie déterministe du signal à partir d'un certain nombre de ses précédents échantillons. Ensuite, on utilise ce modèle afin de prédire les échantillons suivants du signal. Ainsi, la partie résiduelle (aléatoire) est obtenue après soustraction du signal prédit (déterministe) du signal préliminaire.

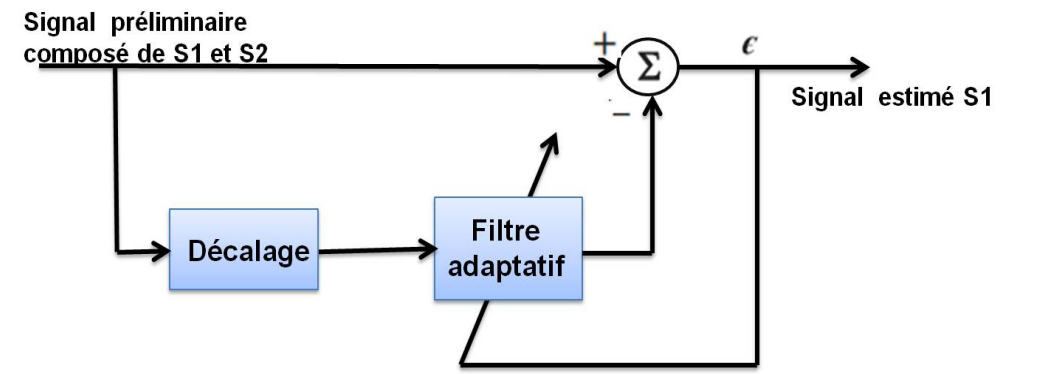


FIGURE 4.2 – Diagramme schématisé de la méthode SANC.

Le modèle utilisé dans cette méthode est un modèle auto-régressif (AR), soit :

$$\hat{x}(n) = - \sum_{k=1}^p a(k)x(n-k)$$

Telle que la valeur à prédire $\hat{x}(n)$ est obtenue à partir d'une somme pondérée de p valeurs précédentes.

Alors, on aura :

$$x(n) = \hat{x}(n) + e(n)$$

Et les coefficients de pondération $a(k)$ sont obtenus en utilisant les équations de Yule-Walker et l'algorithme récursif de Levinson-Durban (LDR) [28].

Le moyennage synchrone (Time Synchronous Averaging (TSA) en anglais :

Une autre méthode qui a pour objectif de séparer entre la composante déterministe CS1 et la composante aléatoire CS2 est la méthode du moyennage synchrone. Il est rappelé que le signal d'intérêt est modélisé par une somme de deux contributions différentes, soit :

1. $P(t)$ est déterministe et périodique (CS1) avec la période T ;
2. $V(t)$ est aléatoire cyclostationnaire à l'ordre deux (CS2) ;

La partie périodique $P(t)$ est estimée par le moyen d'un moyennage synchrone du signal $X(t)$.

La moyenne synchrone $\mu(t)$ de $X(t)$ est donnée par la formule suivante :

$$\mu(t) = \frac{1}{K} \sum_{k=0}^{K-1} X(t + kT)$$

Dès lors, la partie périodique $P(t)$ est donnée comme suit :

$$\hat{P}(t) = \mu(t) * \Pi_T(t)$$

Où $\Pi_T(t)$ est une fonction de Dirac T -périodique et $*$ le produit de convolution. Ceci est équivalent à une convolution de $\mu(t)$ avec un train d'impulsions de Dirac espacées de multiples entiers de la période T . Dans le domaine fréquentiel, la partie périodique $\hat{P}(t)$ montre un spectre de raie aux fréquences $f_k = kf_{cyc}$, tels que $k \in \mathbb{N}$ et f_{cyc} la fréquence du cycle ($1/T$).

La partie résiduelle $V(t)$, qui est supposée CS2, est quant à elle estimée par la soustraction de $\hat{P}(t)$ du signal entier $X(t)$, soit :

$$\hat{V}(t) = X(t) - \hat{P}(t)$$

Dans le domaine fréquentiel, cette soustraction est synonyme de suppression des composantes spectrales $f_k = kf_{cyc}$ tel que $k \in \mathbb{N}$. Ainsi, la partie résiduelle $\hat{V}(t)$ montre un spectre continu où aucune composante fréquentielle n'est clairement présente.

Bien évidemment, il existe d'autres méthodes de séparation entre composantes déterministes et composantes aléatoires. Cependant, le moyennage synchrone est incontestablement la méthode la plus largement utilisée dans ce type de problématiques et dans divers domaines d'application. Néanmoins, aucune étude statistique valide, ni aucun test n'existent pour mesurer l'efficacité et l'exactitude des résultats obtenus. En effet, la première étape qui consiste à l'estimation de la composante CS1 (moyenne synchrone) est primordiale et conditionne les résultats de la séparation CS1/CS2. Motivé par cette problématique, nous proposons dans ce qui suit une nouvelle méthode statistique qui permet de vérifier l'exactitude de l'estimation de la moyenne synchrone. Cette nouvelle méthode s'appelle le *Generalized Seasonal Block Bootstrap (GSBB)* (Bloc Bootstrap Généralisé en français). Cette méthode est basée sur la théorie du bootstrap.

4.3 Le GSBB temporel pour l'estimateur de la moyenne synchrone

Dans le domaine de l'analyse des signaux et des séries temporelles périodiques, de nombreux travaux sur le bootstrap ont été réalisés. Cependant, la plupart de ces travaux sont applicables sous certaines restrictions ayant un rapport avec la longueur de la période ou avec la taille du bloc utilisé lors de l'utilisation de ces différentes méthodes. C'est pourquoi leur application est très limitée. La méthode de bootstrap la plus connue est certainement le *Moving Block Bootstrap* (MBB). Cette méthode a été introduite par [101] et [105] et peut être appliquée pour l'estimation de la moyenne globale, mais pas pour l'estimation de la moyenne synchrone. En effet, le MBB a comme principal inconvénient la destruction totale de la périodicité des données. Les deux méthodes suivantes peuvent être utilisées dans le cas de la moyenne synchrone mais comportent tout de même d'autres inconvénients. En effet, le *Seasonal Block Bootstrap* (SBB) proposé par Politis [123] requiert une taille de bloc b égale à un multiple entier de la taille de la période T . Par conséquent, la taille des données doit être assez longue dans certains cas. En outre, le *Periodic Block Bootstrap* (PBB) proposé dans [34] peut être utilisé lorsque la longueur de la période T accroît avec l'augmentation de la taille des données (voir [104]). De plus, la taille du bloc b doit être inférieure à celle de la période T . D'autres exemples peuvent être trouvés dans [124, 102].

Dans ce chapitre, la méthode du GSBB est utilisée [51]. L'avantage principal du GSBB réside dans le fait qu'il permet un choix de bloc b arbitraire et indépendant de la longueur de la période, contrairement aux méthodes cités dans le paragraphe précédent. En effet, plutôt que de diviser chaque période des données en petits blocs de taille b (PBB) ou de choisir b comme un entier multiple de la période (SBB), puis ré-échantillonner ces blocs de telle sorte à construire un nouvel échantillon de pseudo-observations, le GSBB permet de diviser les données en choisissant une taille de bloc arbitraire (mais raisonnable) qui ne dépend pas de la longueur de la période, puis ré-échantillonner les blocs obtenus de telle sorte à garder intacte la périodicité des données.

Dans ce contexte, l'attention est donc portée sur l'estimation de moyenne synchrone $\mu(t)$. Par conséquent, l'estimation de la composante CS1. Nous rappelons que la moyenne synchrone est donnée par :

$$\mu(t) = \frac{1}{K} \sum_{k=0}^{K-1} X(t + kT).$$

Il est constaté que la moyenne synchrone est en réalité un vecteur de taille de la période T . Alors notons par $\mu = (\mu(1), \dots, \mu(T))$ la moyenne synchrone et par $\hat{\mu} = (\hat{\mu}(1), \dots, \hat{\mu}(T))$ son estimateur. De plus, soit $\mathcal{L}(\sqrt{K}(\hat{\mu} - \mu))$ la distribution de probabilité de $\sqrt{K}(\hat{\mu} - \mu)$ et $\mathcal{L}^*(\sqrt{K}(\hat{\mu}^* - E^*\hat{\mu}^*))$ la distribution sa contrepartie bootstrap.

Dans [51], la consistance du GSBB pour la moyenne synchrone est montrée et le résultat principal est le suivant :

$$d_0 \left(\mathcal{L} \left(\sqrt{K} (\hat{\mu} - \mu) \right), \mathcal{L}^* \left(\sqrt{K} (\hat{\mu}^* - E^* \hat{\mu}^*) \right) \right) \xrightarrow{p} 0, \quad (4.1)$$

Avec $K = \lfloor n/T \rfloor$, d_0 une distance (ou métrique) de mesure de convergence faible en probabilité sur \mathcal{R}^d et $\hat{\mu}^* = (\hat{\mu}^*(1), \dots, \hat{\mu}^*(T))$ la contrepartie issue du bootstrap de $\hat{\mu}$.

Ce résultat implique la consistance du GSBB pour l'estimateur de la moyenne synchrone. Cela signifie que la distribution issue du GSBB converge à la distribution asymptotique de l'estimateur initial.

4.3.1 Algorithme du GSBB

Dans cette partie, l'algorithme du GSBB est décrit. Nous précisons que nous nous focalisons sur une version circulaire du GSBB afin d'éviter de possibles effets de bords. De plus, pour des raisons de simplicité, nous supposons que la taille des données n est un multiple entier de la période T ($n = KT$) et de taille de bloc b ($n = lb$). Ces hypothèses permettent seulement de simplifier les notations et ne sont aucunement essentielles. L'algorithme du GSBB est donc le suivant :

(i) Choisir un entier positif comme taille de bloc $b (< n)$ tel que $l = n/b$ est entier.

(ii) Pour $t = 1, b + 1, 2b + 1, \dots, lb$ soit

$$(X^*(t), X^*(t+1), \dots, X^*(t+b-1)) = (X(h_t), X(h_t+1), \dots, X(h_t+b-1))$$

Avec h_t est iid avec une distribution uniforme telle que :

$$P(h_t = t + vT) = \frac{1}{K} \quad \text{pour } v = 0, 1, \dots, K - 1.$$

Puisque nous considérons la version circulaire du GSBB, alors quand $t + vT > n$ on prend $h_t = t + vT - n$.

- (iii) Joindre les l blocs obtenus après tirage afin de former la nouvelle série de pseudo-observations bootstrap $X^*(1), X^*(2), \dots, X^*(n)$ ayant la même taille n que les données de départ et préservant aussi leur périodicité.

La figure 4.3 schématise les différentes étapes de l'algorithme du GSBB pour un signal périodique.

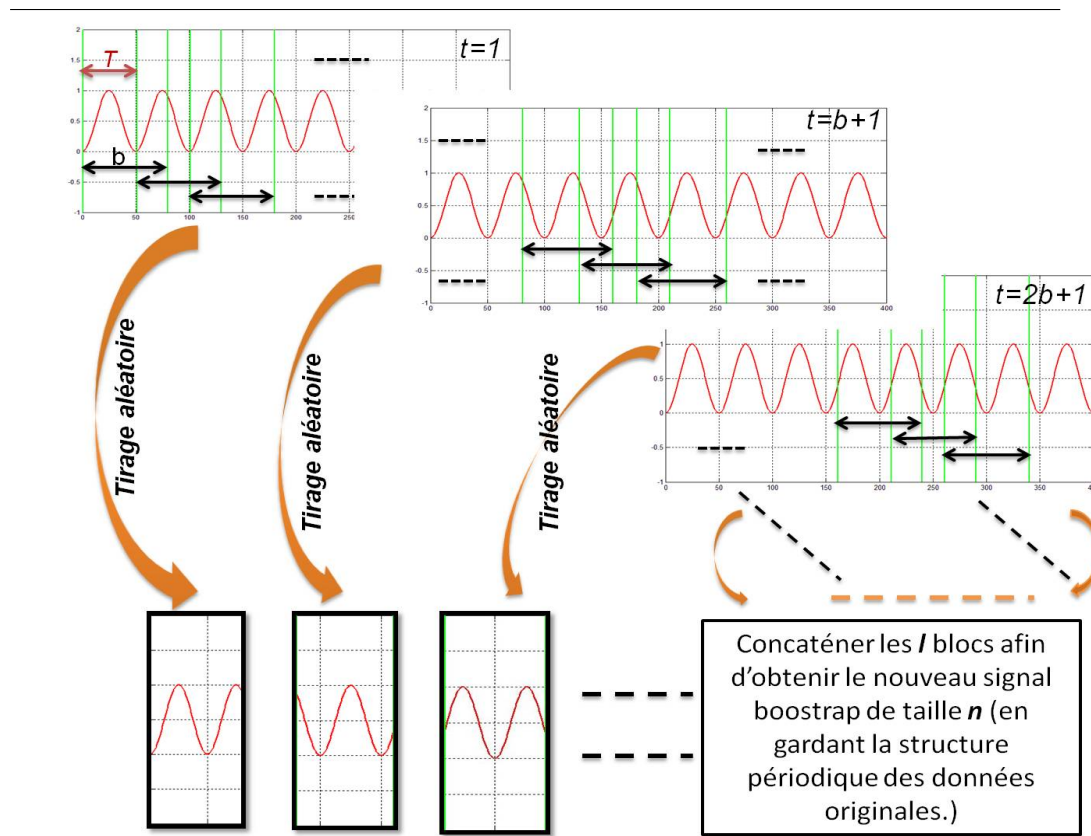


FIGURE 4.3 – Algorithme du GSBB en pratique.

Remarque :

Afin de maintenir la structure périodique du signal, les blocs candidats aux tirages aléatoires sont, pour ainsi dire, synchronisés par rapport à la période cyclique T (voir la figure 4.3).

Par ailleurs, nous précisons que la notion de généralisation dans le GSBB est due au fait que pour certaines valeurs précises de la taille du bloc b utilisé, nous obtenons d'autres méthodes de bootstrap en bloc. En effet, si la taille

du bloc b est inférieure à la longueur de la période T , on obtient le *Periodic Block Bootstrap* (PBB) et pour une taille de bloc est un entier multiple de la période T on obtient le *Seasonal Block Bootstrap* (SBB).

4.3.2 Les intervalles de confiance GSBB pour la moyenne synchrone

L'une des applications les plus importantes du GSBB est la construction d'intervalles de confiance pour $\mu_i, i = 1, \dots, T$. Le résultat (4.1) implique qu'asymptotiquement, des intervalles de confiance valides peuvent être construits pour chaque μ_i . En effet, il est tout à fait possible de dériver des intervalles de confiance à partir des quantiles de la distribution du GSBB. De plus, la consistance de la distribution jointe GSBB de μ^* signifie que des intervalles de confiance simultanés pour $\mu_i, i = 1, \dots, T$ peuvent aussi être construits. Ce fait est une conséquence directe de ce qui suit :

$$\sup_t \left| P \left(\sqrt{K} \max_i |\hat{\mu}_i - \mu_i| \leq t \right) - P^* \left(\sqrt{K} \max_i |\hat{\mu}_i^* - E^* \hat{\mu}_i^*| \leq t \right) \right| \xrightarrow{p} 0 \quad (4.2)$$

Afin de calculer les intervalles de confiance pour l'estimateur de $\mu = (\mu_1, \dots, \mu_T)$, un grand nombre B de pseudo-données bootstrap $X_1^*, X_2^*, \dots, X_n^*$ doit être obtenu par répétition de l'algorithme du GSBB. Ensuite, pour chaque pseudo-données bootstrap, $\hat{\mu}_i^*$ est estimé ($i = 1, \dots, T$). Puis, le résultat 4.2 est utilisé afin d'obtenir les quantiles de la distribution du bootstrap correspondant aux $\frac{\delta}{2}$ et $1 - \frac{\delta}{2}$, avec δ le niveau de confiance désiré. Notons maintenant ces quantiles par $c_{K,B}(\frac{\delta}{2})$ et $c_{K,B}(1 - \frac{\delta}{2})$, alors les intervalles de confiances pour la moyenne synchrone et qui sont dérivés du GSBB sont de la forme suivante :

$$\left(\hat{\mu}_i - \frac{c_{K,B}(1 - \frac{\delta}{2})}{\sqrt{K}}, \quad \hat{\mu}_i - \frac{c_{K,B}(\frac{\delta}{2})}{\sqrt{K}} \right),$$

Avec $i = 1, \dots, T$ et le niveau de confiance $\delta \in (0, 1)$.

Dans ce qui suit, une étude de simulation est réalisée afin de prouver l'applicabilité de la méthode du GSBB.

4.3.3 Étude de simulation

Afin d'illustrer la manière dont le GSBB temporel pour l'estimation de la moyenne synchrone est réalisé en pratique, une étude de simulation est réalisée

en considérant deux différents rapports de RSB. Les valeurs du RSB utilisés sont respectivement -5dB et 10dB.

L'attention est focalisée sur deux types de modèles $X_1(t)$ et $X_2(t)$, soit :

$$X_1(t) = \sin(2\pi f_0 t) + a(t)\sin(2\pi f_0 t) + b(t)$$

et

$$X_2(t) = \sin(2\pi f_0 t) + \sin(2\pi f_1 t) + a(t)\sin(2\pi f_0 t) + b(t)$$

Il est très important de souligner que les signaux simulés $X_1(t)$ et $X_2(t)$ suivent le même modèle qui est défini dans 4.2.

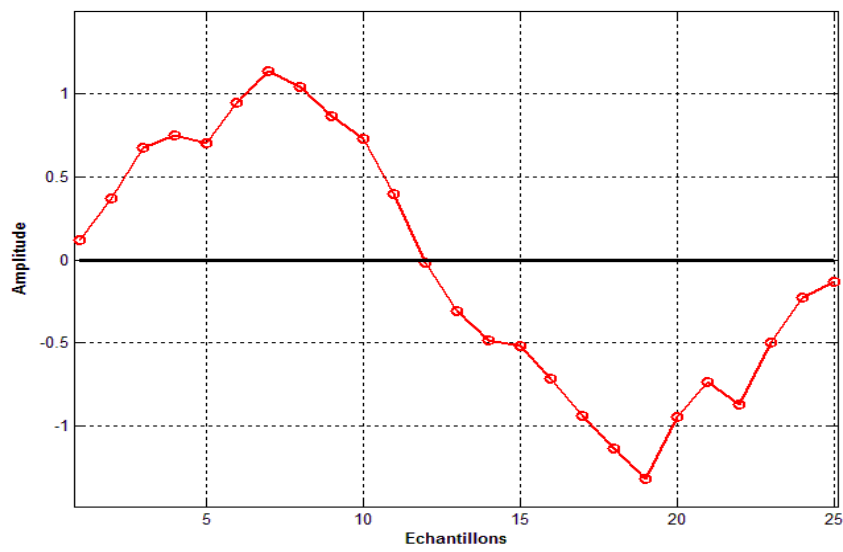
Les caractéristiques de $X_1(t)$ sont les suivantes : le nombre d'échantillons $n = 5000$, la fréquence $f_0 = 1/25Hz$, $a(t)$ et $b(t)$ sont des bruits blancs stationnaires gaussiens.

Les caractéristiques de $X_2(t)$ sont les suivantes : le nombre d'échantillons $n = 15000$, la fréquence $f_0 = 1/25Hz$, la fréquence $f_1 = 1/30Hz$. Enfin, $a(t)$ et $b(t)$ sont des bruits blancs stationnaires gaussiens.

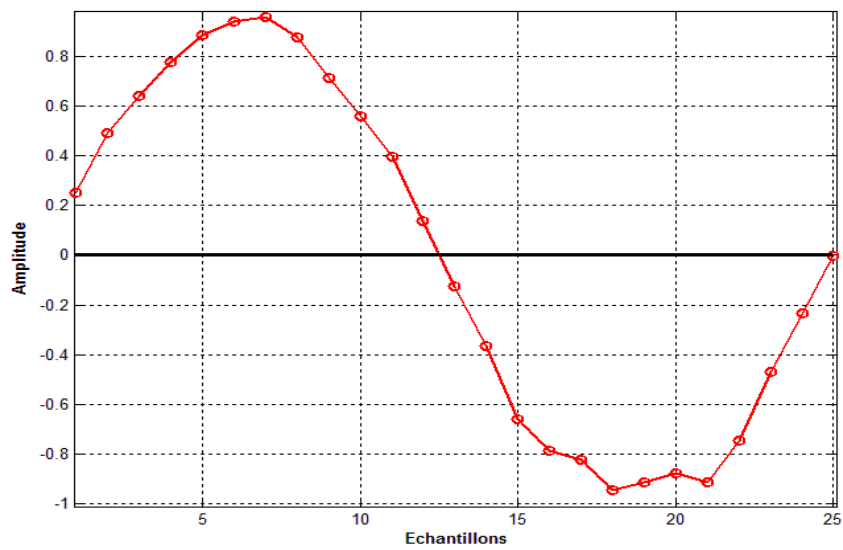
Dans un premier temps, les moyennes synchrones sont estimées pour chaque modèle et pour chaque valeur de RSB. Le moyennage synchrone pour le modèle $X_1(t)$ est réalisé suivant la période $p1 = 25$ échantillons. Quant au modèle $X_2(t)$, le moyennage synchrone est réalisé en prenant la période commune correspondant aux fréquences f_0 et f_1 . En d'autres termes, le plus petit multiple commun des périodes correspondant aux fréquences f_0 et f_1 , c'est-à-dire, $p2 = ppmc(25, 30) = 150$ échantillons.

Les figures 4.4 et 4.5 montrent les estimées de la moyenne synchrone par le biais du moyennage synchrone respectivement pour le modèle $X_1(t)$ et $X_2(t)$, et ce, pour des valeurs de RSB égales à -5dB et 10dB. Globalement, les figures 4.4a et b montrent peu de différence par rapport à la valeur du RSB. La même remarque reste valable pour les figures 4.5a et b. En effet, l'analyse de ces résultats suggérerait que quelque soit la valeur du RSB choisie, la significativité des échantillons qui forment les moyennes synchrones restent la même. Par conséquent, et dans l'optique d'approfondir cette question et étudier la significativité statistique de chaque échantillon formant les moyennes synchrones, nous proposons l'application de la méthode du GSBB. En effet, la génération d'intervalles de confiance par le biais du GSBB permet de construire un test statistique valide pour vérifier la significativité de chaque échantillon de la moyenne synchrone. Cette étude de significativité est fondamentale, particulièrement dans l'analyse des signaux réels.

La méthode du GSBB est appliquée sur les signaux simulés $X_1(t)$ et $X_2(t)$ pour l'estimation de la moyenne synchrone en utilisant les paramètres sui-



(a)

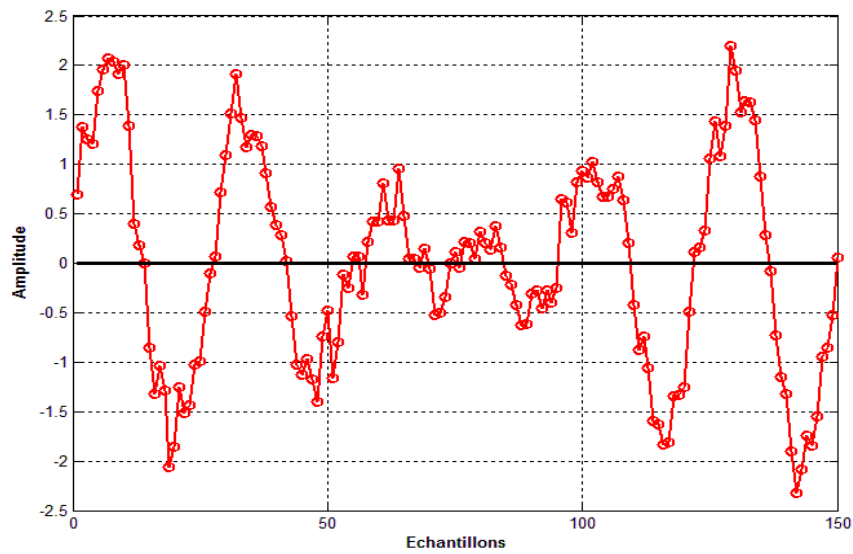


(b)

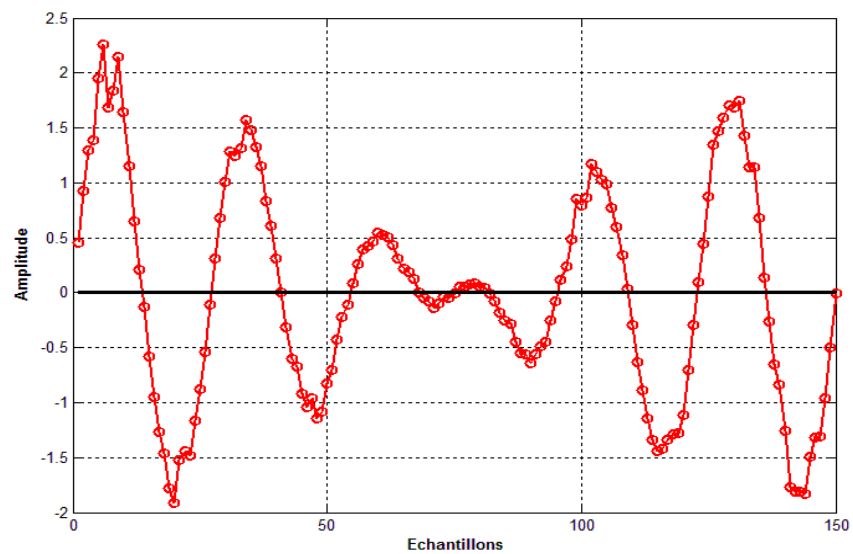
FIGURE 4.4 – Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_1(t)$.
(a) RSB=-5dB ; (b) RSB=10dB.

vants :

- $B = 1000$ le nombre de répétitions du GSBB ;



(a)



(b)

FIGURE 4.5 – Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_2(t)$.
 (a) RSB=-5dB; (b) RSB=10dB.

- $b = 200$ la taille du bloc utilisé (en nombre échantillons) ;
- $\delta = 0.05$ le niveau de confiance ;

Le test statistique pour la moyenne synchrone $\mu_s = (\mu(1), \dots, \mu(p_s))$ construit à partir du GSBB est le suivant :

Pour $i = 1, \dots, p_s$, avec p_s la période d'ordre un de $X_s(t)$ et $s = 1, 2$

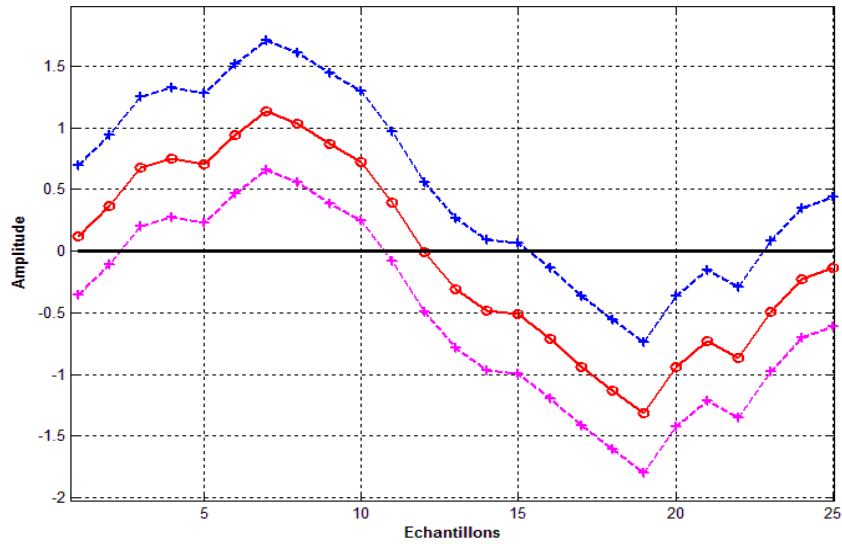
$$\begin{cases} H_0 & : \mu(i) = 0 \\ H_1 & : \mu(i) \neq 0. \end{cases}$$

Où H_0 et H_1 représentent les deux hypothèses de notre test.

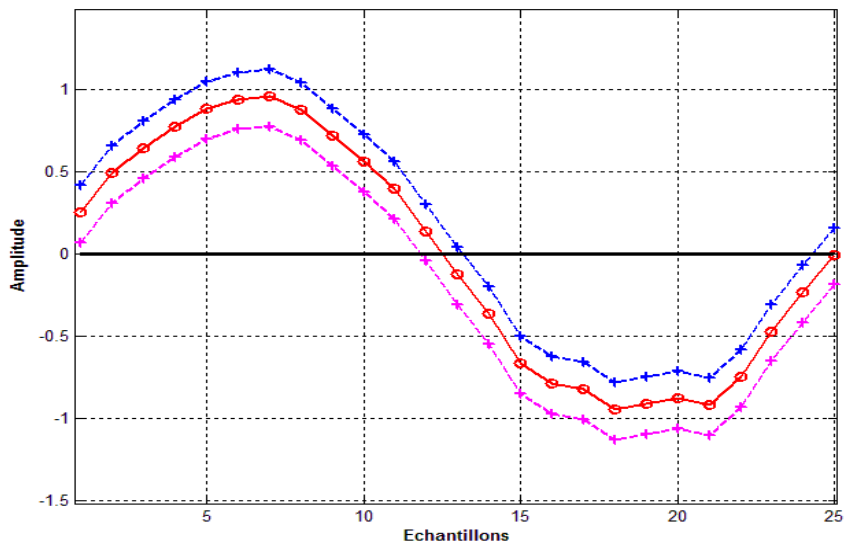
Les figures 4.6 et 4.7 montrent les estimées de la moyenne synchrone par le biais du moyennage synchrone avec les intervalles de confiance construits à partir du GSBB respectivement pour le modèle $X_1(t)$ et $X_2(t)$, et ce, pour des valeurs de RSB égales à -5dB et 10dB.

Contrairement aux résultats précédents obtenus sans l'application du GSBB, nous notons l'existence de différences significatives entre les différentes moyennes synchrones. En effet, la comparaison entre les figures 4.6a et 4.6b montre clairement qu'il existe plus d'échantillons significatifs dans le cas où le RSB est égal à 10dB que dans le cas où le RSB est égal à -5dB. En d'autres termes, l'hypothèse H_0 de notre test statistique est plus souvent rejetée dans le cas où la valeur du RSB est égale à 10dB que dans le cas où la valeur du RSB est égale à -5dB. Les figures 4.7a et 4.7b confirment bien les résultats obtenus pour le signal $X_1(t)$, c'est à dire, les échantillons formant la moyenne synchrone du signal $X_2(t)$ sont globalement plus significatifs quand le RSB est égal à 10dB que quand le RSB est égal à -5dB. Par ailleurs, Il est rappelé que l'hypothèse H_0 est retenue pour un échantillon donné si les intervalles de confiance de cet échantillon contiennent le zéro.

Enfin, et pour conclure cette première partie de ce chapitre, la nouvelle méthode du GSBB présentée ainsi que les résultats de simulation obtenus montrent bien la nécessité et l'intérêt de l'étude statistique de l'estimateur de la moyenne synchrone. En effet, la caractérisation de la significativité des échantillons qui forment la CS1 par le biais de méthodes telle que le GSBB devient fondamentale pour l'étude des signaux CS. La deuxième partie de ce chapitre montre bien l'utilité et la nécessité d'une telle approche dans le cas de l'analyse de signaux réels. En effet, différentes applications de la CS et du GSBB sur des signaux de marche à pied sont présentées dans la partie qui suit.

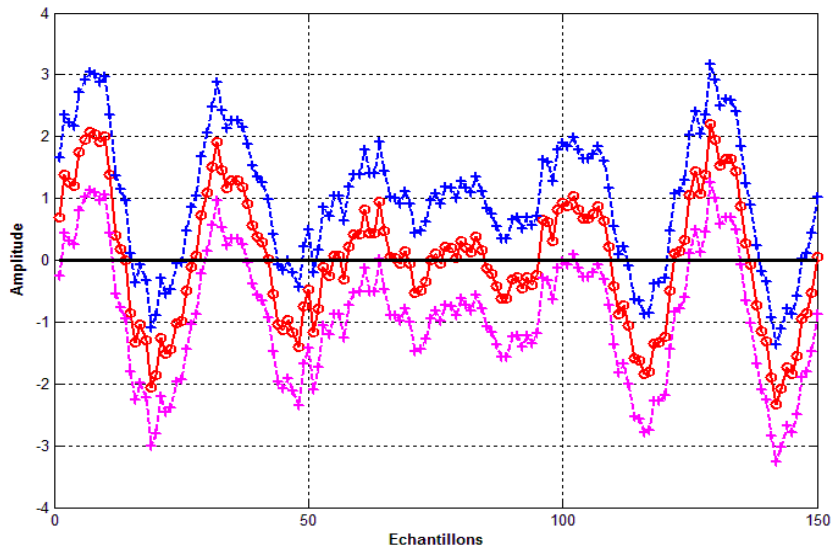


(a)

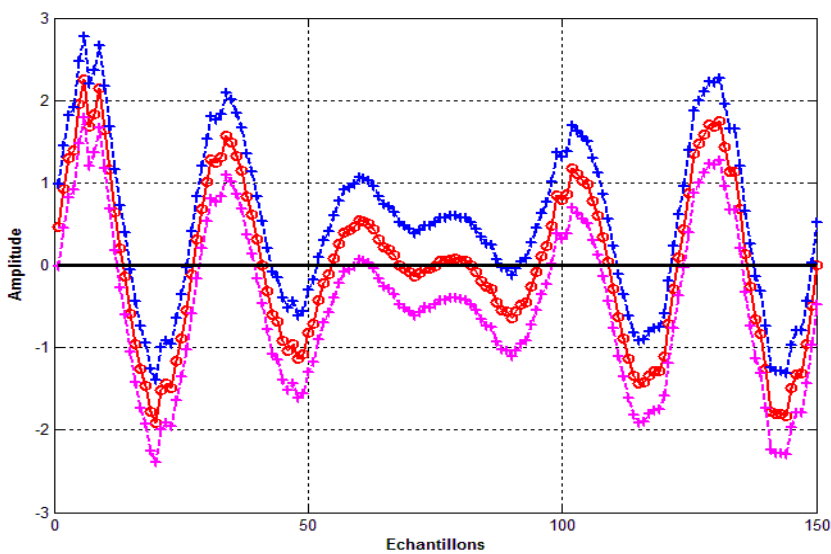


(b)

FIGURE 4.6 – Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_1(t)$ avec les intervalles de confiance GSBB : moyenne synchrone (rouge) ; limite supérieure (bleu) et limite inférieure (magenta). (a) RSB=-5dB ; (b) RSB=10dB.



(a)



(b)

FIGURE 4.7 – Les estimées de la moyenne de synchrone $\hat{\mu}(t)$ du modèle $X_2(t)$ avec les intervalles de confiance GSBB : moyenne synchrone (rouge) ; limite supérieure (bleu) et limite inférieure (magenta). (a) RSB=-5dB ; (b) RSB=10dB.

Deuxième partie

Application du GSBB à l'analyse de la marche à pied

4.4 Introduction

Dans cette deuxième partie, l'intérêt se porte sur un type de phénomène qui n'a, jusqu'à présent pas été étudié en utilisant la modélisation cyclostationnaire. La *marche humaine* est définie par des séquences cycliques et périodiques. Ces séquences sont souvent parasitées par des perturbations qui sont dues à différents facteurs de natures diverses. Par conséquent, la marche humaine est considérée comme un système complexe. Ce système est le résultat de millions d'années d'évolution et d'apprentissage. Il existe deux approches pour étudier la marche humaine. La première consiste à observer les effets de la marche sur des variables mesurées qui sont principalement associées à un modèle descriptif. Cette approche est très utile dans de nombreux domaines tels que la biomécanique, l'analyse du mouvement, mesure des performances, la caractérisation des pathologies, la rééducation, etc.

La seconde approche est plus exploratoire. Elle se concentre sur les causes supposées du déplacement afin d'améliorer les connaissances sur le fonctionnement du cerveau, du système nerveux central et le système sensori-moteur. Ceci conduit à une meilleure compréhension des pathologies associées. Ce domaine est essentiellement celui des neurosciences.

Dans cette recherche, nous nous intéresserons à l'analyse et la caractérisation de la marche dans le cas des personnes âgées. En effet, cette catégorie de personnes est souvent sujette à de nombreux dysfonctionnements moteur/neuromoteur liés à l'âge. La quantification des paramètres spatio-temporelle de la marche peut contribuer à la caractérisation de certains troubles de mobilité et de maladies directement liées à la marche. L'étude de la marche chez cette population peut permettre d'augmenter notre compréhension de la nature du mouvement de la locomotion humaine et contribuer à mieux cibler les interventions préventives. Par conséquent, il devient très important d'étudier les séquences de marche chez les personnes âgées afin de mieux caractériser ce mécanisme, puis proposer des paramètres fiables et spécifiques qui décrivent différents types de mouvements et de troubles du système moteur/neuromoteur.

Dans ce qui suit, une méthodologie associant l'analyse cyclostationnaire au GSBB pour le traitement des signaux de marche est présentée. En outre, une étude comparative via le GSBB est développée afin de mettre en évidence l'influence de la présence de fluctuations de vitesse sur les propriétés CS du signal.

4.5 Méthodologie

4.5.1 Acquisition des données

Un sujet âgé sain (âge : 73 ans ; taille : 174 cm ; poids : 76 kg) a été recruté par le CHU de Saint-Étienne. Le sujet ne présentait aucun problème de type physique ou neurologique. Ensuite, il a été demandé au participant de marcher de façon usuelle et à une vitesse de marche naturelle. L'expérience s'est déroulée dans un long couloir, rectiligne avec un sol régulier.

Afin de réduire d'éventuels facteurs qui pourraient influencer les mesures, quelques règles préliminaires ont été imposées avant le début des acquisitions : un bon éclairage, le port des chaussures quotidiennes, pas de bruit ou autres stimulations externes, marcher en ligne droite en effectuant une distance suffisante et un test préliminaire (sans enregistrements de données) a été réalisé.

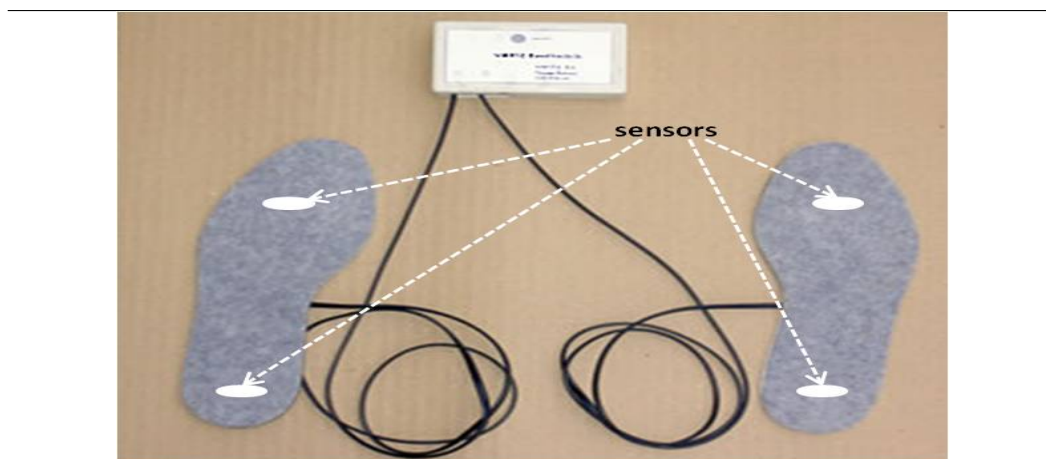


FIGURE 4.8 – Dispositif d'acquisition SMTEC FootSwitch .

Les données ont été acquises au moyen du dispositif *SMTEC FootSwitch* (figure 4.8). Ce système offre la possibilité de mesurer de façon continue et simultanée les paramètres temporels des foulées, et ce, pour de longues distances et périodes. Les chaussures du participant ont été équipées avec des semelles expérimentales. Chaque semelle est ainsi munie de deux capteurs de pression comme indiqué sur la figure 4.8. Une pression plus grande que $40\text{g}/\text{cm}^2$ active les capteurs et définit l'état de contact. Le premier contact est défini par l'activation du capteur localisé au niveau du talon, alors que le dernier contact est défini par le dernier orteil soulevé du capteur localisé au niveau du métatarse du pied. Les capteurs sont reliés à un boîtier électronique d'acquisition par le biais de deux câbles. La communication entre le boîtier

et l'ordinateur se fait par bluetooth. Enfin, la fréquence d'échantillonnage des signaux est égale à 80Hz.

Il s'agit d'un appareil facile à utiliser, tout à fait commode et pas très coûteux, ce qui permet d'envisager une utilisation de routine.

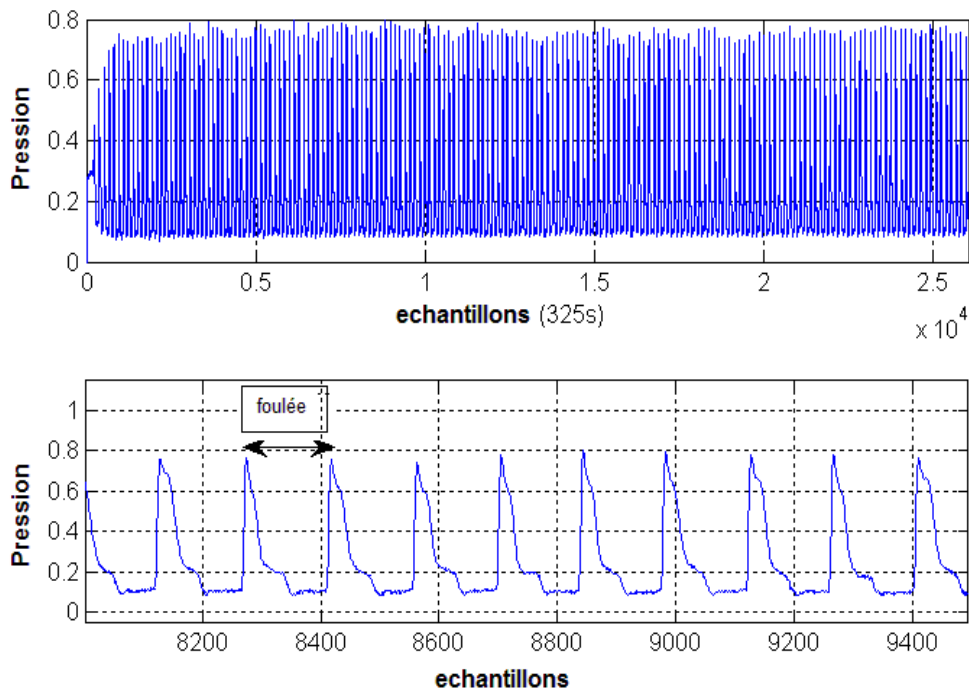


FIGURE 4.9 – Un signal de marche.

La figure 4.9 montre le signal de marche acquis et utilisé dans cette étude. Ce signal représente les amplitudes de pression produites par le talon du pied droit du participant lors de la marche. Ainsi, ce signal décrit un ensemble de foulées consécutives. Enfin, la durée d'acquisition du signal est de 325s.

4.5.2 Fluctuations de vitesse et description des signaux de marche

Les signaux de marche sont, d'une certaine façon, similaires aux signaux issus de machines tournantes. En effet, avec des machines qui tournent avec des vitesses nominales constantes, les signaux sont directement liés à la vitesse de rotation de l'arbre. Cependant, il existe dans la plupart des cas (si ce n'est tous), des fluctuations de vitesse aléatoires (généralement faibles). Par

conséquent, la vitesse de rotation de l'arbre n'est plus purement déterministe (périodiques). Le même effet est observé dans les signaux biomécaniques et dans les données relatives à la locomotion humaine de façon générale. En effet, de faibles variations aléatoires de vitesse sont constatées.

Typiquement, les signaux sont acquis dans le domaine temporel. Si le sujet marche à une vitesse constante $\dot{\theta}(t)$, le signal serait échantillonné à pas temporels et angulaires constants et la relation entre le domaine temporel et le domaine angulaire serait la suivante :

$$\theta(t) = \omega_{cyc}t = 2\pi f_{cyc}t \Rightarrow \dot{\theta} = \omega_{cyc} \quad (4.3)$$

Où f_{cyc} est la fréquence de la foulée (le cycle) et ω_{cyc} la pulsation correspondante. Cela serait le cas idéal, mais en réalité, des fluctuations de vitesse $\dot{\xi}(t)$ existent toujours, et ce à cause de la nature pas tout à fait cyclique de la locomotion humaine. Alors la relation (4.3) devient comme suit :

$$\theta(t) = \omega_{cyc}t + \xi(t) \Rightarrow \dot{\theta} = \omega_{cyc} + \dot{\xi}(t) \quad (4.4)$$

Ainsi, il devient clair que le signal de marche n'est pas échantillonné avec une fréquence d'échantillonnage synchrone de la fréquence de la foulée. Les dispersions de vitesse sont typiquement faibles (voir la figure 4.10, mais aucun moyen n'existe pour vérifier si ces fluctuations de vitesse sont assez importantes pour impacter les propriétés cyclostationnaires du signal.

C'est pourquoi, et afin de vérifier et d'étudier l'effet des fluctuations de vitesse sur les propriétés cyclostationnaires d'ordre un et deux, le même signal est traité dans deux cas de figures : (i) après la compensation des fluctuations de vitesse ; (ii) sans compensation des fluctuations de vitesse.

4.5.3 Le ré-échantillonnage synchrone

Afin d'éviter l'étalement et la dispersion des fréquences à cause des fluctuations de vitesse, une synchronisation du signal est nécessaire. En effet, le signal acquis est observé en fonction du temps et non pas en fonction du cycle de marche du participant. C'est pourquoi, il est nécessaire de synchroniser le signal sur la fréquence correspondant au cycle de la foulée du sujet. Ainsi, la meilleure méthode pour réaliser cela est de ré-échantillonner chaque échantillon du signal par rapport à la période de la foulée. Par conséquent, les fluctuations de vitesse seront compensées et le signal synchronisé sur la foulée du sujet.

Dans le domaine de la surveillance des machines tournantes, la composante de vitesse de l'arbre est mesurée grâce à un tachymètre ou un capteur de

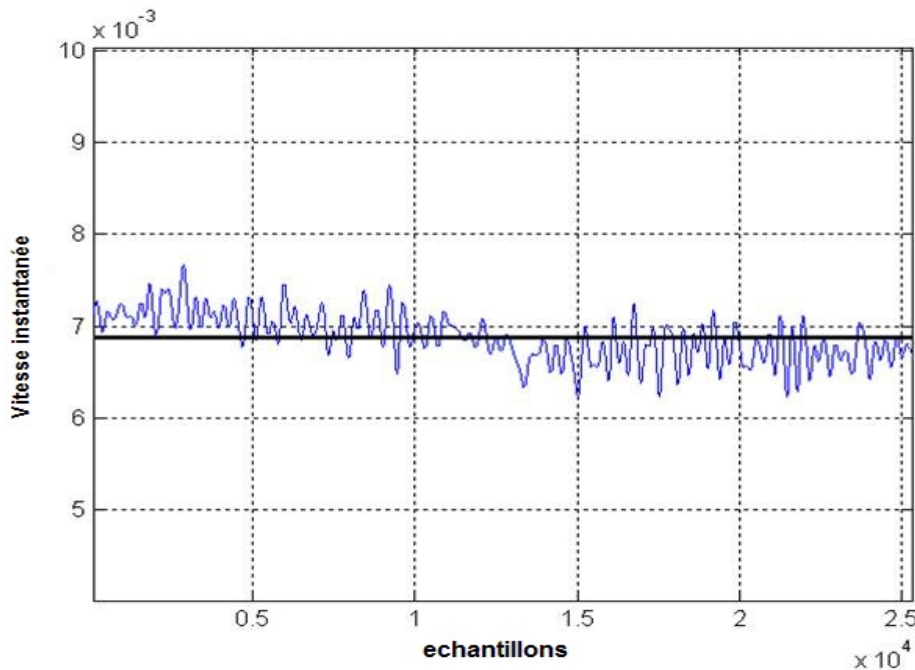


FIGURE 4.10 – La vitesse instantanée du signal de marche.

vitesse/position [24]. Dans notre cas, nous n'avons aucun capteur de mesure du cycle de la foulée.

Si des variations de vitesse existent, le délai entre chaque pic fréquentiel actif varie. En suivant ces variations, il est tout à fait possible d'estimer la position en fonction du temps afin de compenser les fluctuations de vitesse et de ré-échantillonner le signal par rapport au cycle de la foulée, et ce, en utilisant que le signal lui même.

Dans la figure 4.11, nous pouvons distinguer des pics qui correspondent à la fréquence de la foulée et ses harmoniques. Il est très important de souligner l'impact des fluctuations de vitesse se traduit par un élargissement des pics. Cet élargissement est proportionnel au rang de l'harmonique. De plus, la présence de fluctuations de vitesse dissipe l'énergie autour des pics. Ainsi, en suivant ces fluctuations autour des pics, puis en les compensant, il est possible d'avoir des pics plus nets.

L'objectif de la partie ci-dessous est de synchroniser le signal de marche (qu'on appellera $X(t)$) sur le cycle de la foulée du sujet. Une description de la méthode qui se décompose en deux principales étapes est exposée.

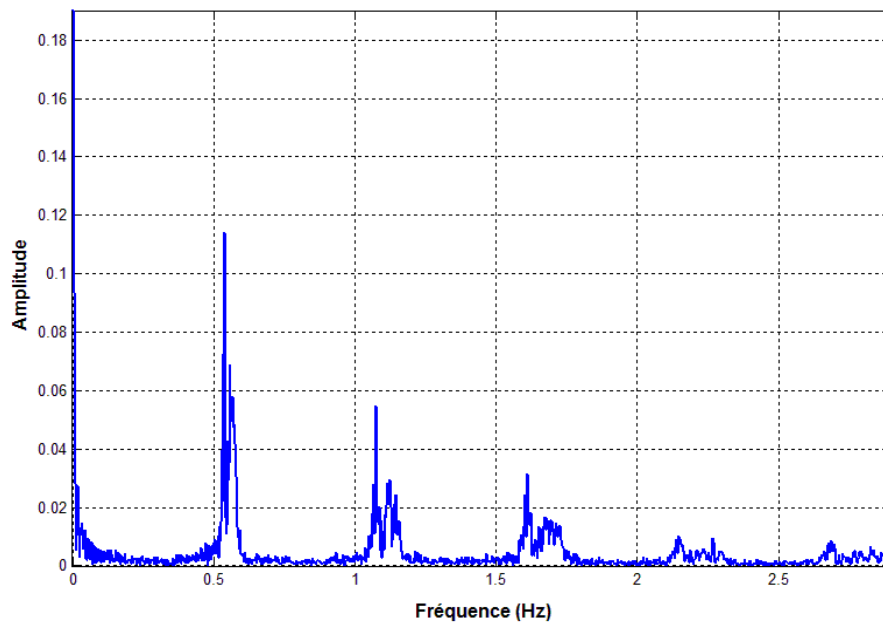


FIGURE 4.11 – La transformée de Fourier du signal de marche.

1. Estimation de la vitesse instantanée et la position instantanée du pied

Afin d'établir la relation entre le domaine angulaire (la phase) et le mouvement du sujet, la figure 4.12 montre les différentes étapes d'un cycle de marche (foulée). Ainsi, chaque position du pied correspondra à un pourcentage de cycle consommé. Par conséquent, la phase du signal de marche nous fournira l'information sur la position du pied θ dans le cycle de la foulée, soit :

$$2\pi \equiv 1\text{cycle (foulée)} \equiv 100\% \text{ de cycle consommé.}$$

Les foulées sont alors vues comme étant des patterns répétitifs avec de faibles fluctuations de fréquence causées par les variations de la vitesse de marche du sujet.

Alors, dans le cas idéal, le signal $X(\theta)$ serait composé entièrement de somme de patterns répétitifs et serait représenté comme suit :

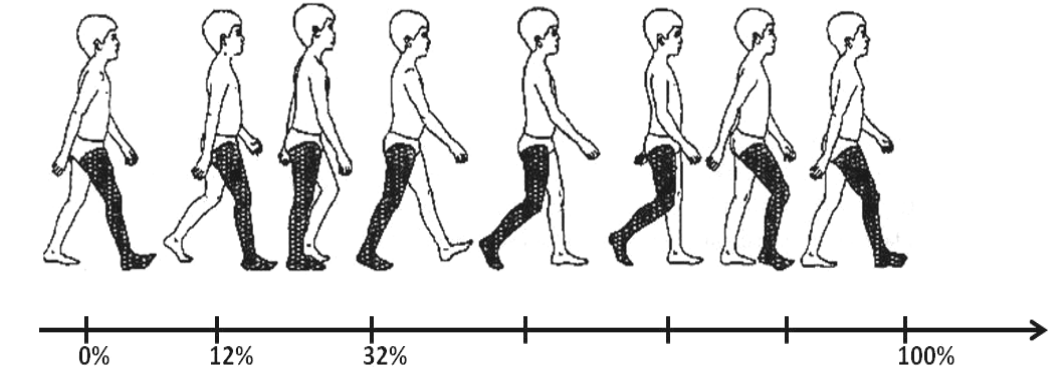


FIGURE 4.12 – Les étapes d'un cycle de marche.

$$X(\theta) = \Re\left(\sum_k A_k \exp(jk\theta(t))\right)$$

et le modèle serait tout simplement une somme de sinusoides.

Cependant en pratique, la locomotion humaine, et spécialement durant la marche à pied, de faibles variations de vitesse existent toujours. Cela signifie que $\theta(t) \neq \omega_{cyc}t$ mais $\theta(t) = \omega_{cyc}t + \xi(t)$.

Dans notre cas, l'acquisition du signal est réalisée en fonction du temps, et non pas en fonction de la progression du mouvement de marche. Alors, nous avons :

$$\begin{aligned} X[\theta(t)] &= X[\omega_{cyc}t + \xi(t)] \\ &= \Re\left(\sum_k A_k \exp(jk[\omega_{cyc}t + \xi(t)])\right) \\ &= \Re\left(\sum_k A_k \exp(jk\theta(t))\right) \end{aligned}$$

Alors, nous obtenons un signal modulé en phase. Contrairement au modèle de somme de sinusoides précédent, la représentation fréquentielle du signal modulé en phase ne sera pas par des pics de fréquence déterministes et nets, mais plutôt des distributions autour des fréquences

moyennes (figure 4.11) causées par les variations de vitesse. Cependant, et comme signalé auparavant, ces variations ne sont pas assez importantes pour noyer les fréquences d'intérêt.

Le signal observé $X(t)$ contient à travers sa phase l'information sur la position du pied. Ainsi, pour estimer la position du pied durant le mouvement ainsi que la vitesse instantanée du sujet, une démodulation analytique du signal $X(t)$ est réalisée en exploitant la fréquence de la foulée. Un filtre passe bande centré sur la fréquence de la foulée (ou une de ses harmoniques) excluant les bandes latérales et tout autre perturbation permet l'isolation de la fréquence de la foulée (ou une de ses harmoniques). Il est important de préciser que l'ajustement du filtre passe bande doit être réalisé en observant le spectre du signal avec une très bonne résolution fréquentielle.

De plus, il est nécessaire que la condition de faibles fluctuations de vitesse soit respectée. Sinon, des collisions entre les harmoniques et/ou entre les bandes latérales du spectre sont possibles. En effet, dans le cas de fluctuations de vitesse très importantes, l'estimation par le moyen de cette méthode n'est plus possible.

L'extraction de la phase instantanée correspondant à la position du pied est réalisée par le biais du signal analytique. Le signal filtré $\hat{X}[\theta(t)]$ contient seulement une composante fréquentielle, comme requis dans [21] :

Pour un k fixe :

$$\begin{aligned}\hat{X}[\theta(t)] &= A_1 \exp(j[\omega_{cyc}t + \xi(t)]) \\ &= A_1 \exp(j\theta(t))\end{aligned}$$

Ainsi, la phase instantanée $\theta(t)$ avec la phase déroulée $\varphi(\hat{X}[\theta(t)])$ sont extraites à partir du signal filtré $\hat{X}[\theta(t)]$. Dès lors, pour un angle donné Θ_i correspondant à une position donnée du pied à l'intérieur du cycle de la foulée, on utilise $\theta(t)$ estimée précédemment afin d'avoir les t_i correspondant tel que : $t_i = \theta^{-1}(\Theta_i)$, puis les $X(t_i)$ correspondant aux $X(\Theta_i)$. Cela signifie que nous sommes désormais dans la capacité d'associer une position à chaque échantillon du signal acquis $X(t)$, et ce, en procédant à une étape d'interpolation.

2. Interpolation

Le seconde et la dernière étape de la méthode de synchronisation du signal de marche sur la période de la foulée du sujet est l'interpolation. En effet, afin d'associer une position de pied durant la marche à chaque échantillon temporel du signal $X(t)$, une interpolation de $\theta(t)$ est réalisée.

Durant l'interpolation, le signal $X(t)$ est estimé pour différentes positions du pied correspondant à un pas constant $\varphi_a(\Theta_i) = \frac{2\pi}{T}\Theta_i$. Avec T la période de la foulée exprimée en nombre d'échantillons. La méthode des splines est utilisée pour l'interpolation à cause de ses propriétés qui imposent quelques contraintes de continuité.

Alors, dans l'optique d'étudier les propriétés cyclostationnaires CS1 et CS2 ainsi que l'effet des faibles fluctuations de vitesse sur ces dernières, le signal original et le signal synchronisé seront analysés séparément.

4.5.4 La séparation CS1/CS2

Soit le signal $X(t)$ le signal original sans synchronisation et soit $X(\Theta)$ sa version synchronisée.

L'objectif de cette étape est de séparer entre la composante déterministe CS1 et la composante aléatoire CS2. Il est rappelé que le signal est modélisé par une somme de deux contributions différentes (section 4.2), soit :

$$X(t) = P(t) + V(t) + b(t),$$

et

$$X(\Theta) = P(\Theta) + V(\Theta) + b(\Theta)$$

Nous rappelons aussi que :

1. $P(t)$ (et $P(\Theta)$) est déterministe et périodique (CS1) avec la période T
2. $V(t)$ (et $P(\Theta)$) est aléatoire cyclostationnaire à l'ordre deux (CS2)

Les parties périodiques $P(t)$ et $P(\Theta)$ sont estimées par le moyen d'un moyennage synchrone des signaux $X(t)$ et $X(\Theta)$. Il est rappelé que la méthode du moyennage synchrone est décrite dans la section 4.2.2.

Cette étape d'estimation et de séparation entre la composante CS1 et CS2 du signal est fondamentale. Cependant, aucune étude statistique valide, ni aucun test n'existent pour mesurer l'efficacité et l'exactitude des résultats obtenus après la séparation. Ceci est particulièrement vrai quand de faibles fluctuations de vitesse existent. C'est pourquoi, l'objectif principal dans la suite est d'étudier la question de séparation des composantes CS1/CS2 du signal de marche par une approche originale issue du GSBB. En effet, de nouveaux indicateurs caractérisant l'estimation CS1/CS2 sont présentés. De plus, l'influence de la présence de faibles fluctuations de vitesse sur les propriétés CS du signal de marche est démontrée.

4.5.5 Mesure de la qualité de l'estimation

La consistance de la méthode du GSBB pour $\hat{\mu}_i$ (section 4.3) a pour conséquence l'obtention d'intervalles de confiance valides pour l'estimateur de la moyenne synchrone $\mu(t)$. Notons ces intervalles de confiance par $[L_T(i), U_T(i)]$, tels que $L_T(i)$ et $U_T(i)$ correspondent respectivement aux limites inférieurs et supérieurs des intervalles de confiance obtenus grâce au GSBB.

Afin de mesurer la qualité de l'estimation de la moyenne synchrone, nous définissons le paramètre suivant :

$$MQ(\hat{\mu}_i) = \frac{1}{T} \sum_{i=1}^T [U_T(i) - L_T(i)]^2 \quad (4.5)$$

Avec T la période.

Plus la valeur de $MQ(\hat{\mu}_i)$ est faible, plus les intervalles de confiance sont étroits et meilleure est la qualité de l'estimée de la moyenne synchrone.

4.5.6 Indicateur de séparabilité CS1/CS2 basé sur le GSBB

Après le moyennage synchrone et la soustraction de la composante d'ordre un (CS1) du signal, nous supposons très souvent que le signal résiduel est CS2 ou purement aléatoire. Cependant, cette hypothèse peut s'avérer très risquée et pourrait conduire à une utilisation inappropriée des outils d'analyse de la CS2. En effet, l'un des problèmes les plus courants est celui de la présence ou non de structures périodiques au premier ordre dans le signal résiduel. En d'autres mots, *est ce que la partie déterministe qui représente la CS1 a été*

proprement enlevée du signal brute ?. Ainsi, l'hypothèse émise après la procédure de soustraction de la composante CS1 est que *le signal résiduel obtenu n'est plus périodique en moyenne*. Dans l'optique d'explorer et d'étudier cette question pour les deux signaux considérés, à savoir : le signal original et le signal synchronisé, nous proposons un indicateur dérivé du GSBB pour mesurer la séparabilité CS1/CS2. Cet indicateur est défini de la façon suivante :

Soit $\Gamma = \{T_1, \dots, T_m\}$ un ensemble de périodes différentes dans le voisinage de la période réelle. Pour chaque période de l'ensemble Γ , la procédure du GSBB est appliquée sur les signaux *résiduels* $\hat{V}(t)$ et $\hat{V}(\Theta)$ qui représentent respectivement la partie résiduelle du signal original et la partie résiduelle du signal synchronisé.

Ensuite, la moyenne synchrone est estimée pour chaque période de l'ensemble Γ et pour chaque signal $X(t)$ et $X(\Theta)$. Les intervalles de confiance sont alors dérivés par le biais du GSBB pour chaque période T_s avec $s = 1, \dots, m$. Puis, le pourcentage d'échantillons significatifs en moyenne $PSM(s)$ est calculé comme suit :

$$PSM(s) = \frac{1}{T_s} \cdot \# \{ \hat{\mu}_{P_{T_s}}(s), L_{P_{T_s}}(s), U_{P_{T_s}}(s) > 0 \\ \text{et } L_{P_{T_s}}(s) < \hat{\mu}_{P_{T_s}}(s) < U_{P_{T_s}}(s); \text{ ou } \\ \hat{\mu}_{P_{T_s}}(s), L_{P_{T_s}}(s), U_{P_{T_s}}(s) < 0 \\ \text{et } L_{P_{T_s}}(s) < \hat{\mu}_{P_{T_s}}(s) < U_{P_{T_s}}(s) \}$$
 (4.6)

Tels que :

- $\#$ la fonction indicatrice (ou caractéristique) de l'ensemble $\{.\}$;
- $P_{T_s} = 1, \dots, T_s$;
- $\hat{\mu}(s)$ l'estimée de la moyenne synchrone pour $s = 1, \dots, m$;
- $L(s)$ et $U(s)$ sont respectivement les bornes inférieures et supérieures des intervalles de confiance ;

4.6 Résultats

Cette partie est consacrée aux résultats obtenus et à leur discussion.

4.6.1 Le ré-échantillonnage synchrone

Le résultat du ré-échantillonnage synchrone du signal $X(t)$ sur la période de la foulée est montré sur la figure 4.13. Cette figure présente les transformées de Fourier du signal original $X(t)$ et de sa version synchronisée $X(\Theta)$.

Il est très clair que les pics qui correspondent à la fréquence de la foulée et ses harmoniques sont nettement plus étroits après la procédure de synchronisation du signal. Contrairement au cas du signal original où les pics présents dans le spectre sont dissipés et élargis. Ainsi, le résultat obtenu montre qu'à priori, les fluctuations de vitesse observées sur le signal original sont compensées après le ré-échantillonnage synchrone du signal.

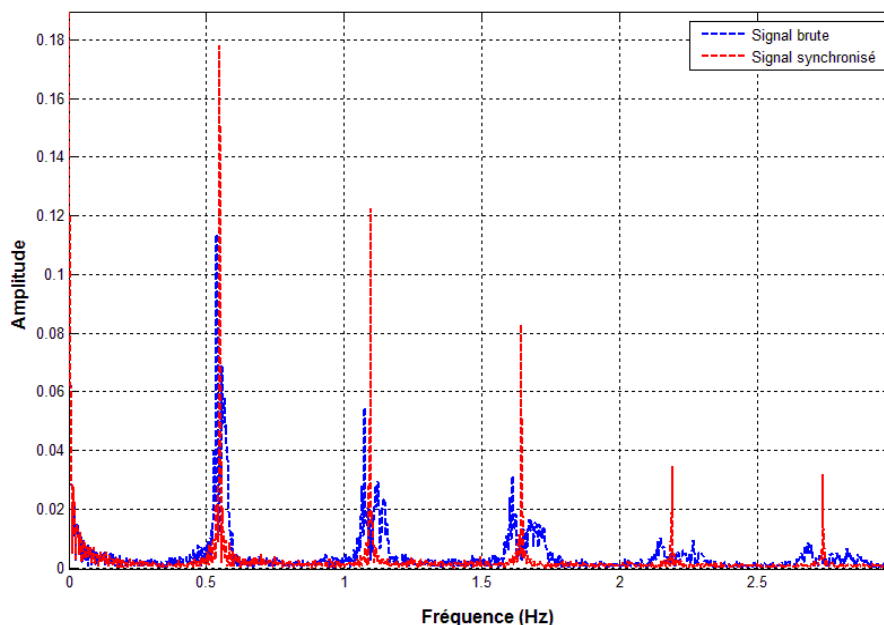


FIGURE 4.13 – Transformées de Fourier de $X(t)$ (en bleu) et $X(\Theta)$ (en rouge).

Afin d'explorer l'effet de la synchronisation du signal et l'effet des faibles fluctuations de vitesse sur les propriétés cyclostationnaires CS1 et CS2 du signal, nous montrons dans la suite les résultats de ces nouvelles approches sur le signal original $X(t)$ et sa version synchronisée $X(\Theta)$.

4.6.2 Analyse cyclostationnaire d'ordre deux CS2

Les figures 4.14 et 4.15 représentent les valeurs estimées des fonctions d'autocorrélation cyclique $R(\alpha, \tau)$ ainsi que leurs coupes à $\tau = 0$, respecti-

vement pour le résiduels provenant du signal synchronisé $X(\Theta)$ et du signal original $X(t)$.

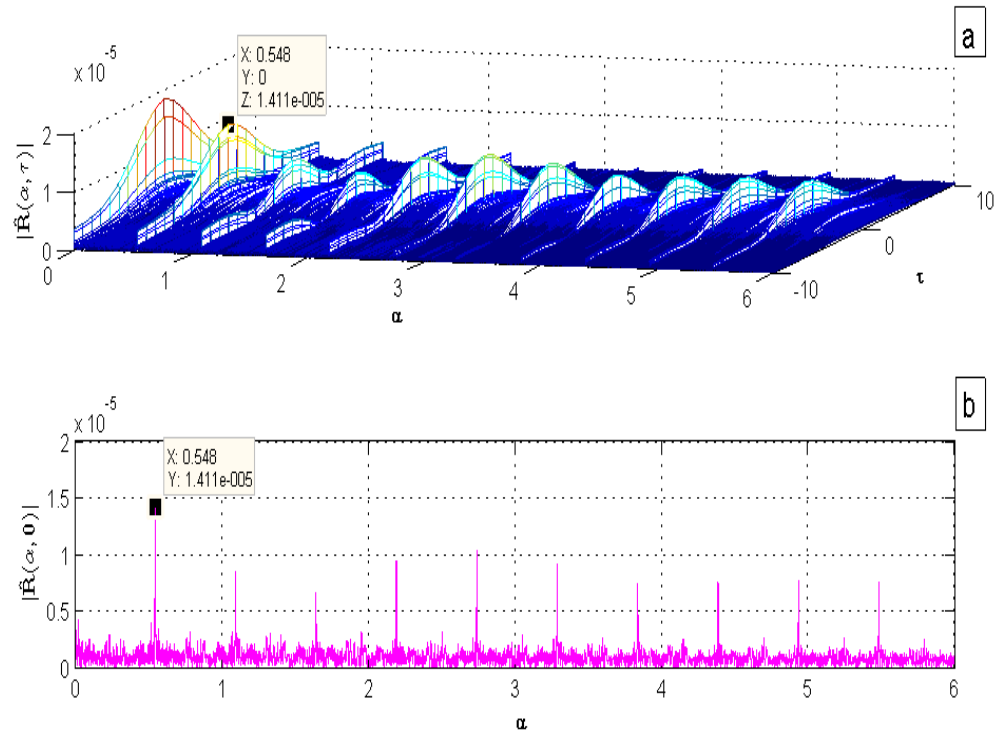


FIGURE 4.14 – Les valeurs estimées de la fonction d'autocorrélation cyclique sur le résiduel $\hat{V}(\Theta)$ (après synchronisation) : (a) $\hat{R}(\alpha, \tau)$. (b) $\hat{R}(\alpha, 0)$.

Les résultats obtenus montrent clairement que le signal résiduel estimé $\hat{V}(\Theta)$ issu du signal synchronisé $X(\Theta)$ montre des propriétés cyclostationnaires à l'ordre deux. En effet, ce signal est CS2 à la fréquence de la foulée du sujet. Cependant, le signal résiduel $\hat{V}(t)$ issu du signal original $X(t)$ semble ne montrer aucune cyclostationnarité à l'ordre 2.

Les fréquences cyclique observées dans le résiduel $\hat{V}(\Theta)$ pourraient être une conséquence de variabilité au niveau de la foulée du sujet. En effet, plus les fluctuations de foulées (de foulée en foulée) sont importantes et plus la CS2 est conséquente. Cela a pour effet de rendre la démarche de moins en moins déterministe, donc aléatoire.

Les personnes âgées présentent généralement des troubles de la marche qui sont dues à des fluctuations de foulées. Ces modifications des paramètres de

marche spatio-temporelles peuvent être une conséquence de déficits de performances moteurs provoquées par un dysfonctionnement du système nerveux central, du système périphérique, ainsi que du système neuromusculaire [125]. De plus, il a été montré que l'atrophie du volume du cerveau dans le cas des gens âgés, en particulier la dégénérescence ganglionnaire du lobe frontal-basale [129], pourrait être directement liée à des déficits moteurs, y compris la marche à pied. Dans [130], les auteurs ont montré l'existence d'une corrélation positive entre le volume de la matière grise et de nombreux paramètres de la marche. En effet, un volume plus important de la matière grise entraînerait de bien meilleures performances de la marche.

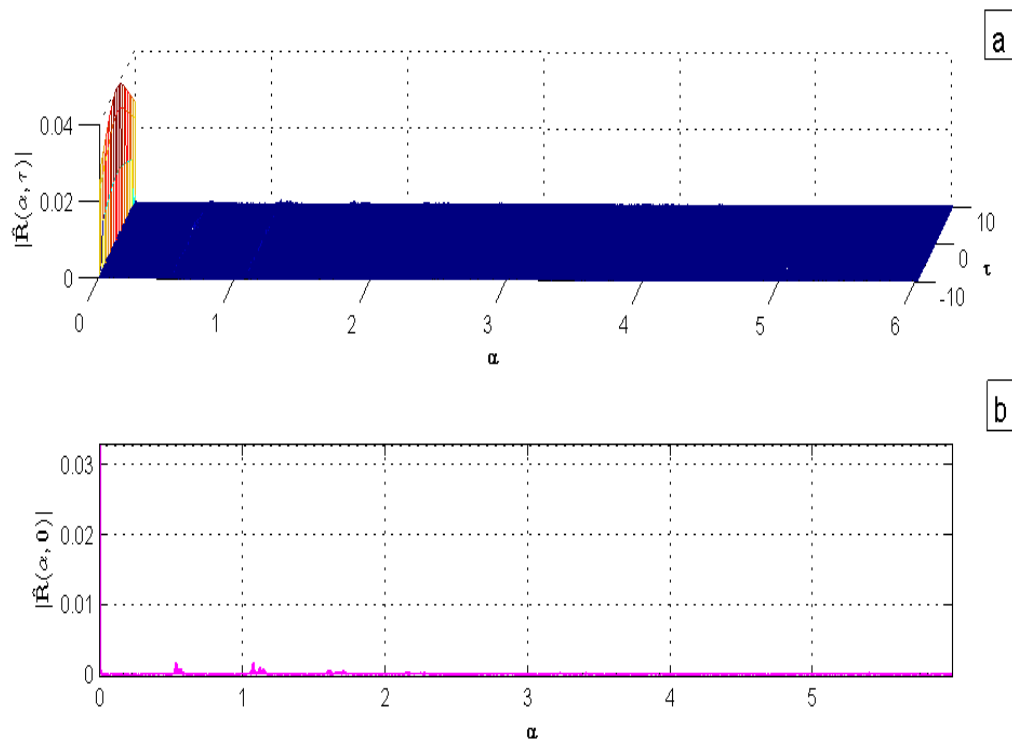


FIGURE 4.15 – Les valeurs estimées de la fonction d'autocorrélation cyclique sur le résiduel $\hat{V}(t)$ (sans synchronisation) : (a) $\hat{R}(\alpha, \tau)$. (b) $\hat{R}(\alpha, 0)$.

Par ailleurs, la comparaison entre les estimées des deux fonctions d'autocorrélation cyclique montre que l'amplitude à $\alpha = 0$ est significativement plus importante dans le cas du signal résiduel $\hat{V}(t)$ issu du signal original $X(t)$. Ce résultat pourrait être la conséquence de la présence de structures pure-

ment périodiques en moyenne dans le signal résiduel provenant du signal non synchronisé.

Afin de vérifier cette dernière hypothèse ainsi que les résultats de CS2 obtenus, dans ce qui suit, l'attention est focalisée sur les propriétés CS1 des deux signaux par le moyen du GSBB.

4.6.3 Estimation de CS1 par GSBB

L'estimation de la contribution déterministe et périodique des pieds (CS1) est réalisée grâce à un moyennage synchrone basé sur le GSBB. La longueur des deux périodes estimées dans les deux cas est $T = 146$ échantillons par cycle, ce qui correspond à la longueur de la foulée du sujet. Les paramètres du GSBB utilisés pour l'estimation des moyennes synchrones et les intervalles de confiance sont les suivants :

- $B = 1000$ le nombre de répétitions du GSBB ;
- $b = 300$ la taille du bloc utilisé (en nombre échantillons) ;
- $T = 146$ la période de la foulée ;
- $\delta = 0.05$ le niveau de confiance ;

Les figures 4.16 et 4.17 montrent les moyennes synchrones estimées ainsi que leurs intervalles de confiance construits à partir du GSBB, et ce, respectivement pour le signal synchronisé $X(\Theta)$ et le signal original $X(t)$.

Tout d'abord, nous remarquons clairement de façon visuelle que les deux moyennes synchrones sont de formes différentes. De plus, la figure 4.17 montre des intervalles de confiance relativement assez larges. En effet, la valeur du MQ défini par (4.5) est $MQ = 3,01 \cdot 10^{-3}$.

Par ailleurs, la figure 4.16 montre des intervalles de confiance très étroits comparés avec ceux de la figure 4.17. En effet, la valeur du MQ est significativement moins importante que dans le premier cas ($MQ = 2,11 \cdot 10^{-4}$).

Sachant que nous observons pour les deux signaux le même nombre d'échantillons n et que les mêmes paramètres du GSBB sont respectés dans les deux cas. Les amplitudes moins importantes des échantillons formant la moyenne synchrone ainsi que les intervalles de confiance assez larges avec une valeur de MQ élevée indiquent une variabilité (variance) plus importante dans le cas du signal original. Cela signifie que la qualité de l'estimée de la moyenne

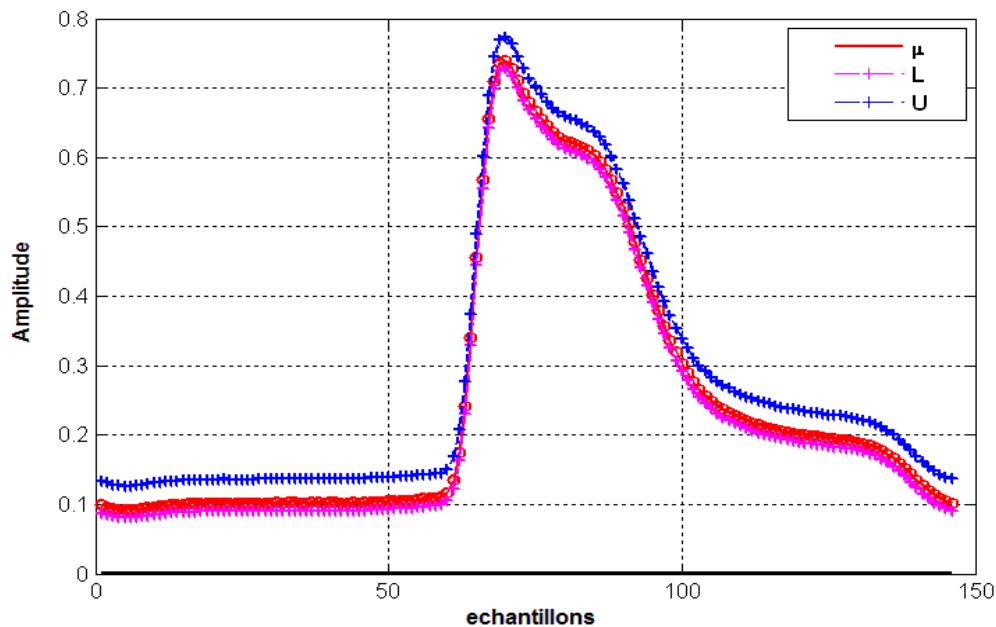


FIGURE 4.16 – L'estimée de la moyenne synchrone par GSBB avec 95% d'intervalles de confiance -le signal synchronisé $X(\Theta)$ -.

synchrone est très moyenne dans ce cas. Cependant, les amplitudes plus importantes des échantillons formant la moyenne synchrone ainsi que les intervalles de confiance très étroits avec une valeur de MQ basse indiquent une variabilité moins importante que dans le premier cas. Ainsi, on obtient une bien meilleure estimation de la moyenne synchrone après la synchronisation du signal.

Par ailleurs, il est très important de souligner que puisque la totalité des échantillons formant les deux moyennes synchrones sont significatifs (aucun intervalle de confiance ne contient le zéro dans les deux cas, figures 4.16 et 4.17), alors les résultats obtenus ne permettent pas d'expliquer entièrement les différences observées des propriétés CS2 entre les deux signaux sous étude. Donc, le moyen le plus raisonnable serait d'explorer les deux composantes résiduelles provenant respectivement du signal original et du signal synchronisé. En effet, les propriétés CS2 des deux signaux sont supposées être contenues dans leurs parties résiduelles respectives.

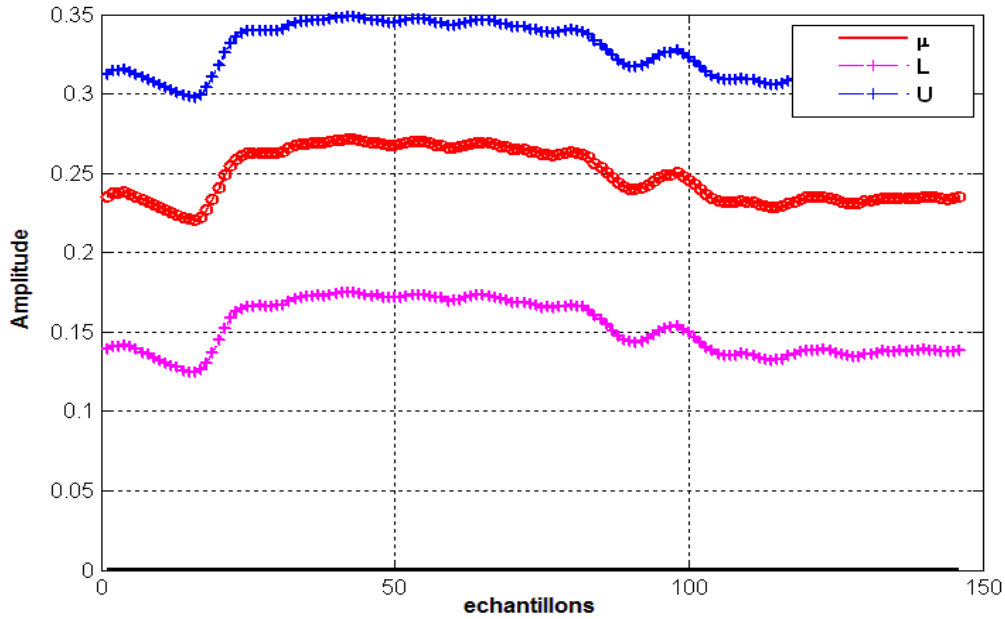


FIGURE 4.17 – L’estimée de la moyenne synchrone par GSBB avec 95% d’intervalles de confiance -le signal original $X(t)$ -.

4.6.4 Indicateur de séparabilité CS1/CS2

Après la soustraction de la composante d’ordre un qui représente la CS1 du signal, il est supposé que la composante résiduelle obtenue contient de la CS2. Afin d’explorer cette hypothèse, l’indicateur de séparabilité CS1/CS2 PSM défini dans la section 4.5.6 est exploité. En effet, le PSM est un indicateur basé sur le GSBB et donne le pourcentage d’échantillons significatifs en moyenne pour un ensemble Γ de différentes périodes se situant dans le voisinage de la période principale de la foulée du sujet.

Soit dans notre cas $\Gamma = \{120, 121, \dots, 146, \dots, 170\}$ (en nombre d’échantillons). L’application de la méthode présentée dans 4.5.6 pour le calcul des PSM est appliquée sur les signaux résiduels provenant respectivement du signal original et du signal synchronisé. Les résultats sont montrés dans les figures 4.18 et 4.19

La figure 4.18 montre nettement aucun échantillon significatif dans les moyennes synchrones correspondant aux périodes testées de l’ensemble Γ . Ces moyennes synchrone étant estimées sur la partie résiduel $\hat{V}(\Theta)$ obtenue après la soustraction de la composante CS1 qui représente la période de la foulée

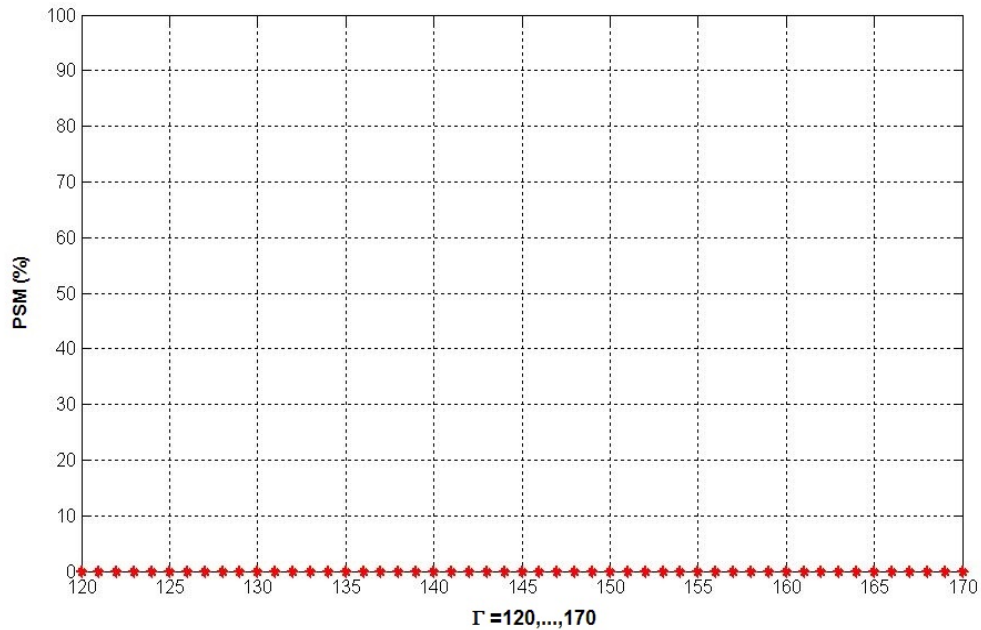


FIGURE 4.18 – Les valeurs du PSM pour $\hat{V}(\Theta)$ issu du signal synchronisé $X(\Theta)$.

du sujet. Ces résultats permettent de conclure que la composante CS1 a été tout à fait bien estimée et enlevée du signal synchronisé. Par conséquent, le résiduel obtenu montre un caractère CS2 comme il a été montré sur la figure 4.14.

Cependant, la figure 4.19 montre de nombreuses périodes de l'ensemble Γ pour lesquelles il existe des échantillons significatifs en moyenne. Cela signifie que des structures purement périodiques en moyenne sont encore présentes dans le signal résiduel $\hat{V}(t)$ du signal original. Ces résultats permettent de conclure que même si la composante CS1 correspondant à la foulée du sujet ($T = 146$) a été bien enlevée ($PSM(T = 146) = 0$), il existe de nombreuses autres structures périodiques en moyenne qui n'ont pas été enlevées. Ainsi, le signal résiduel contient encore des composantes ayant des structures purement périodiques. Par conséquent, les propriétés CS2 ne peuvent pas être estimées proprement et restent cachées et noyées par les différentes structures périodiques en moyenne qui sont contenues dans le signal résiduel (voir la figure 4.15)

Enfin, tous les résultats obtenus dans cette partie grâce à l'utilisation de méthodes combinant entre la théorie de la cyclostationnarité, la synchronisa-

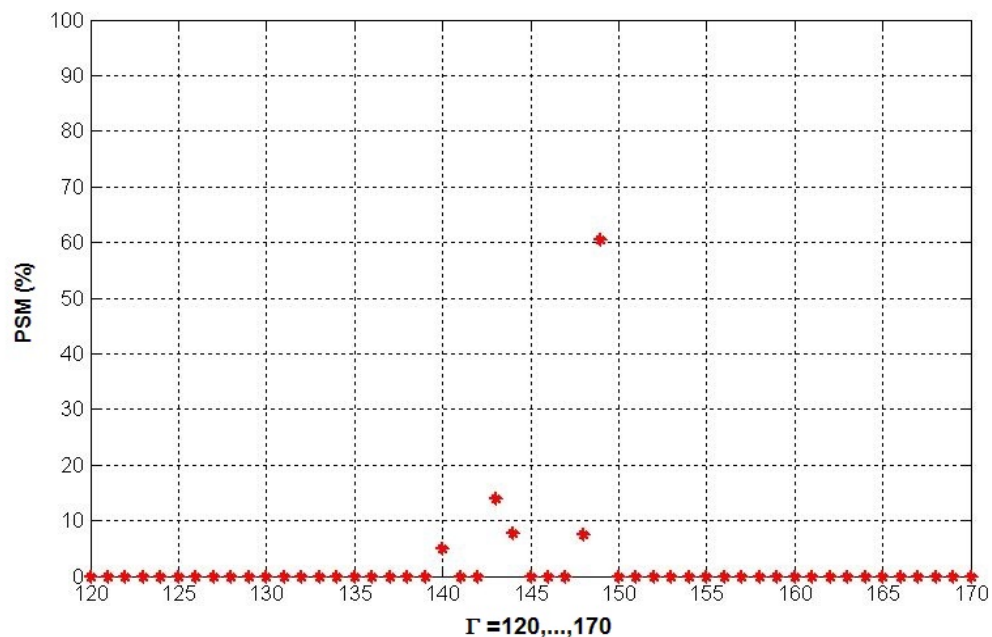


FIGURE 4.19 – Les valeurs du PSM pour $\hat{V}(t)$ issu du signal synchronisé $X(t)$.

tion des signaux et le GSBB montrent le très important effet des fluctuations de vitesse. En effet, même si ces fluctuations sont faibles et de l'ordre de quelques pourcents, elles sont assez pour biaiser les propriétés CS1 et CS2 des signaux.

4.7 Application à l'analyse de la marche en double tâche

Dans cette dernière partie de ce chapitre, une étude de la marche avec double tâche est présentée. Les outils et méthodes présentés auparavant sont appliqués.

Traditionnellement, la marche est considérée comme un acte de réflexe sollicitant au minimum les ressources intentionnelles. Par conséquent, il est très important d'étudier cette hypothèse, et ce, par le moyen de ce qu'on appelle *la marche avec double tâche*. Le paradigme de la marche avec double tâche consiste à réaliser simultanément deux tâches et est utilisé pour l'étude de la relation entre la marche et l'attention. Selon ce paradigme, il pourrait y

avoir une interférence entre les deux tâches réalisée si ces dernières font appel au même système fonctionnel. Ainsi, de nombreux travaux de recherche, relevant essentiellement du domaine de la médecine, se sont intéressés à ce genre de paradigme [49, 136, 107, 100]. Cependant, aucune de ces recherches n'a exploité le traitement statistique du signal tels que le bootstrap et la modélisation cyclostationnaire. Donc, l'originalité du travail effectué dans cette dernière partie de ce chapitre réside dans l'exploitation de la modélisation CS des signaux de marche ainsi que dans l'utilisation du GSBB. Ainsi, l'approche adoptée dans ce travail est très différente des approches rencontrées dans la littérature.

Dans ce qui suit, une description du paradigme expérimental est présentée.

4.7.1 Le paradigme expérimental de la marche avec double tâche

L'expérience présentée dans cette partie a été réalisée au CHU de Saint-Étienne sur une période de plus de 10 ans. Durant ce paradigme, 345 sujets âgés, sains et consentants ont été recrutés. Tous les participants ne présentaient aucun problème d'ordre physique ou mental. L'âge moyen des participants est de 73 ± 1 ans (avoir une population homogène par rapport à l'âge était très important pour d'autres études médicales).

Les mêmes règles préliminaires citées dans la section 4.5.1 sont imposées et respectées.

Ensuite, il a été demandé aux participants d'effectuer différentes marches avec tâche. Les trois types de marche sont les suivants :

- **Marche simple (MS)** : marcher de façon habituelle ;
- **Marche avec un décompte (MD)** : marcher en décomptant à haute voix de 59 à 0 ;
- **Marche avec fluence verbale (MF)** : marcher en citant à haute voix des noms d'animaux commençant par une lettre imposée ;

Les acquisitions des signaux ont été réalisées grâce au même système SM-TEC FootSwitch décrit dans la section 4.5.1.

Ainsi, pour chaque condition et pour chaque sujet, il existe 4 signaux comme suit :

- *Pied droit* : un signal issue du capteur placé au niveau du talon et un autre issu du capteur placé au niveau du métatarse/orteils ;

- *Pied gauche* : un signal issue du capteur placé au niveau du talon et un autre issu du capteur placé au niveau du métatarse/orteils ;

4.7.2 Analyse des paramètres temporels de la foulée

Afin de mener une étude complète sur les données de marche acquies, il est très important de commencer par un paramètre basique mais essentiel à ce type d'étude. Ce paramètre est la période de la foulée. En effet, étudier la période de la foulée revient en quelque sorte à étudier les propriétés statistiques d'ordre un de ces signaux. De plus, c'est un paramètre très important pour caractériser la marche [17].

Pour ce faire, une détection des minimums et maximums locaux de chaque signal est réalisée. En effet, l'écart entre deux maximums locaux consécutifs (ou minimums locaux) représente la période de la foulée comme il est indiqué sur la figure 4.20.

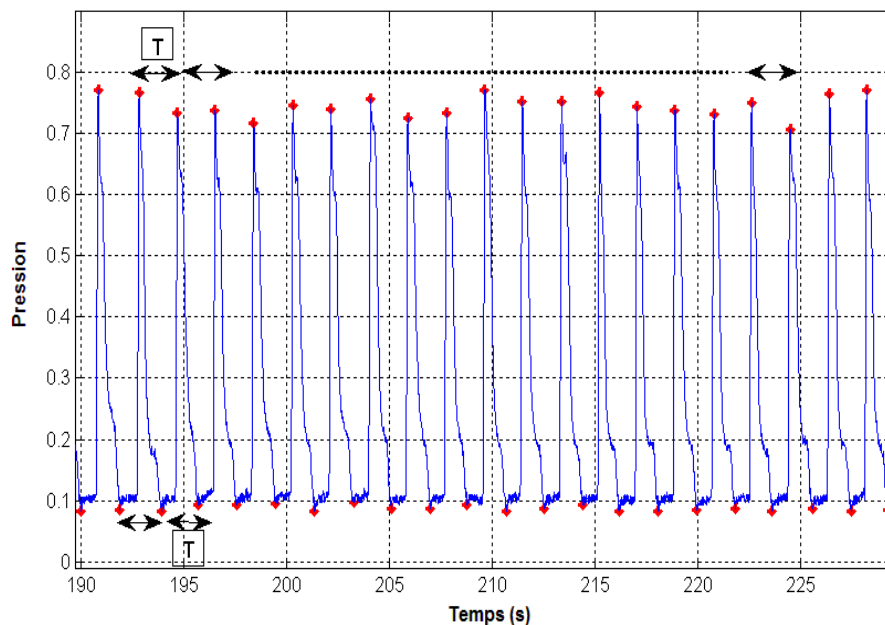


FIGURE 4.20 – Exemple de détection de maximums et minimums locaux sur un signal de marche (les points rouges) .

Ainsi, la période moyenne de chaque signal est estimée, puis, pour chaque capteur, un test statistique de Newman-Keuls est appliqué sur ces périodes

moyennes afin d'étudier s'il existe des différences significatives entre les foulées moyennes dans les trois conditions différentes de marche.

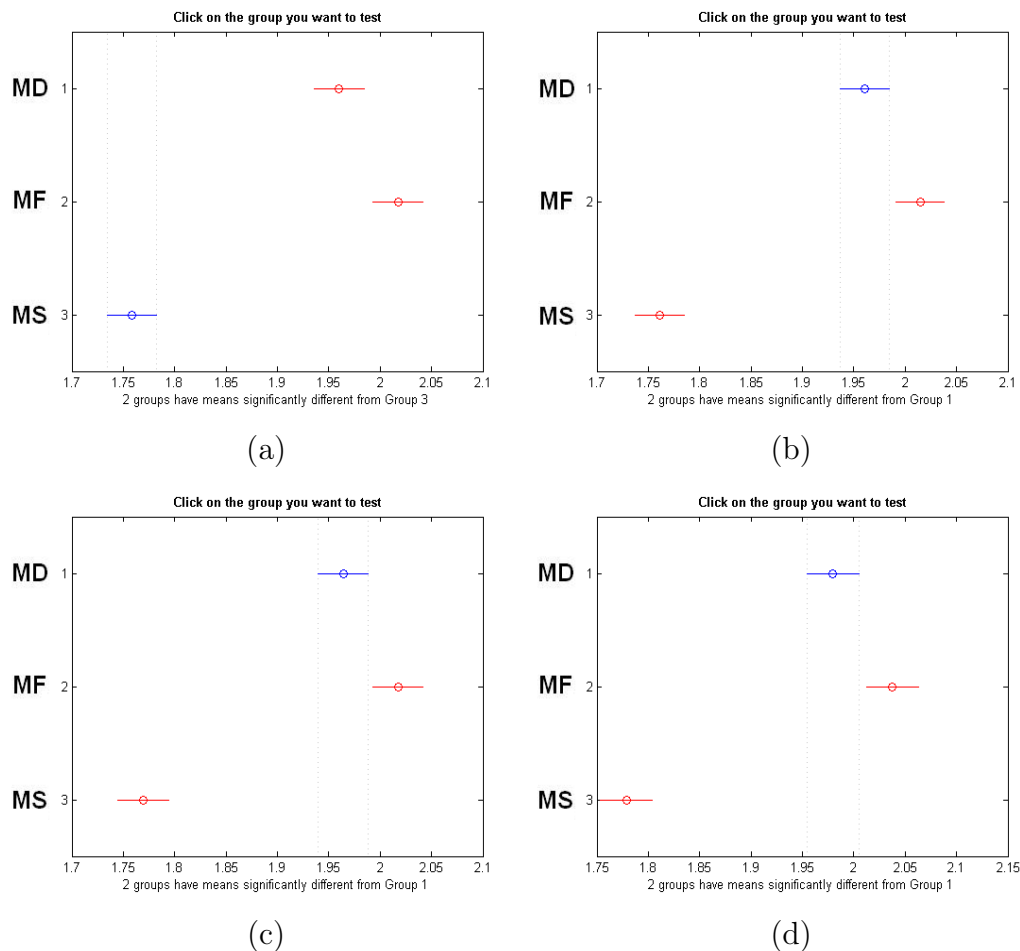


FIGURE 4.21 – Tests de Newman-Keuls sur les valeurs des temps de foulées moyens : (a) capteur du talon du pied droit ; (b) capteur du métatarse/orteils du pied droit ; (c) capteur du talon du pied gauche ; (d) capteur du métatarse/orteils du pied gauche. Les trois conditions MS, MD et MF sont représentées sur l'axe des ordonnées des figures.

Les figures 4.21a, b, c et d représentent les résultats du test statistique de Newman-Keuls pour les 3 conditions de marche respectivement pour : le capteur du talon du pied droit, le capteur du métatarse/orteils du pied droit, le capteur du talon du pied gauche et le capteur du métatarse/orteils du pied gauche. Les résultats obtenus montrent des différences significatives entre les temps de foulée moyens des 3 conditions, et ce, pour les signaux issus des 4 capteurs. En effet, le temps de foulée moyen de la condition de marche (MF)

est significativement plus grand que celui de la condition (MD). Ce dernier est quant à lui significativement plus grand que le temps de foulée moyen de la condition (MS). Ces résultats sont valables pour l'ensemble des capteurs.

Les résultats obtenus suggèrent que lorsque les sujets effectuent la tâche verbale tout en marchant, leurs temps de foulée semblent être significativement plus longs que lors de la réalisation de la marche simple ou de la marche avec décompte. En outre, les temps de foulée des sujets lors de la réalisation de la marche avec décompte semblent significativement plus élevés que lors de la marche simple. Ainsi, ces résultats pourraient signifier que la vitesse de la marche des sujets varie avec le type de tâche effectuée. En effet, il s'avère que plus la tâche cognitive est compliquée à réaliser, plus la vitesse de la marche baisse.

4.7.3 Analyse d'ordre deux des signaux de marche avec double tâche

Afin d'avoir un aperçu de ce que pourraient être les effets de la double tâche sur les propriétés CS2 des signaux, nous proposons tout d'abord de traiter les signaux d'un seul sujet choisis aléatoirement. Les signaux correspondant à un capteur placé au talon droit du sujet choisis sont traités, et ce, dans les 3 différentes conditions de marche. La méthodologie utilisée pour le traitement de ces signaux est celle que nous avons présentée dans les sections précédentes.

Ainsi, chaque signal est d'abord ré-échantillonné de façon synchrone par rapport à sa période de la foulée. Ensuite, l'estimation de la CS1 basée sur le GSBB est appliquée et la séparation CS1/CS2 réalisée. Les valeurs de l'indicateur PSM sont nulles pour les trois signaux qui correspondent aux trois conditions de marche, ce qui signifie que les composantes CS1 et CS2 sont estimées correctement. Enfin, la fonction d'autocorrélation cyclique $R(\alpha, \tau)$ est estimée pour chaque composante résiduelle.

Les figures 4.22, 4.23 et 4.24 représentent les valeurs estimées de la fonction d'autocorrélation cyclique $R(\alpha, \tau)$ sur les résiduels provenant respectivement des signaux de marche avec les conditions (MS), (MD) et (MF).

Les résultats obtenus montrent la présence de CS2 à la fréquence de la foulée du sujet et de ses harmoniques pour les trois conditions de marche. Ce caractère CS2 à la fréquence de la foulée et de ses harmoniques montre que des fluctuations de foulée existent lors de la marche, quelque soit la condition. Pour être plus précis, il existe des variabilités et des fluctuations aléatoires d'une foulée à l'autre.

Par ailleurs, la comparaison entre les trois cas montre que les valeurs les

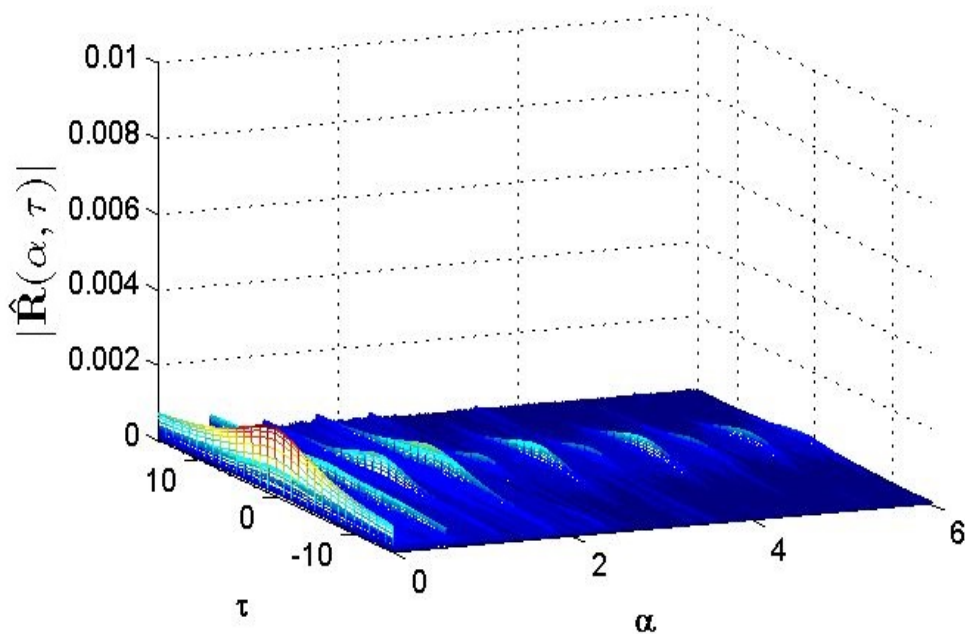


FIGURE 4.22 – Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche simple (MS).

plus élevées de la CS2 sont dans le cas de la marche avec tâche verbale (MF), suivie par la marche avec décompte (MD), et enfin par la marche simple (MS). Ces résultats sous entendent que les fluctuations de foulée sont les plus importantes lorsque le sujet réalise sa marche tout en citant à haute voix des noms d'animaux commençant par une lettre imposée. Ces fluctuations deviennent moins importantes lorsque le sujet marche en décomptant. Enfin, les fluctuations les moins importantes sont observées lorsque le sujet effectue une marche simple.

Afin de confirmer ces résultats obtenus sur un seul sujet, nous proposons une étude du groupe de tous les participants dans ce qui suit.

4.7.4 Étude de groupe

Dans cette partie, nous procédons à l'étude des propriétés de la CS2 sur l'ensemble des sujets, à savoir, 345 participants. Des comparaisons entre les trois différentes conditions de marche sont exposées.

Tout d'abord, chaque signal de pression est normalisé par rapport à sa valeur maximale afin d'éliminer tout éventuel effet du poids des sujets sur les

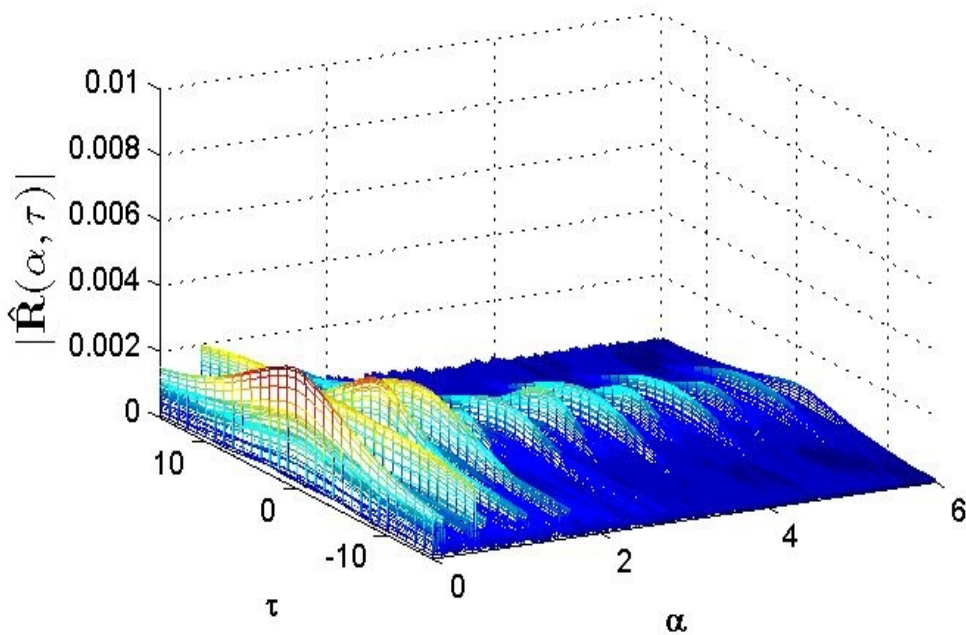


FIGURE 4.23 – Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche avec décompte (MD).

résultats des comparaisons. Ensuite, nous avons appliqué notre méthode de ré-échantillonnage synchrone sur tous les signaux (4140 signaux), puis l'estimation de la CS1 basée sur le GSBB est appliquée et la séparation CS1/CS2 réalisée tout en vérifiant que les composante CS1 aient été proprement enlevées des signaux.

Pour des raisons purement pratiques, nous utilisons la relation existant entre la corrélation spectrale intégrée et l'analyse d'enveloppe qui a été introduite par [126] afin d'estimer la CS2 sur l'ensemble des signaux. En effet, il a été prouvé dans [126] que l'intégration de la corrélation spectrale suivant l'axe des fréquences f fournit un spectre discret sur les fréquences cycliques α (équivalent à une projection de la corrélation spectrale sur l'axe des fréquences cycliques α). Aussi, si le signal analysé est de nature CS2, la corrélation spectrale intégrée est équivalente à élever l'amplitude de ce signal au carré, puis à appliquer sa transformée de Fourier (voir l'annexe A pour plus de détails). Cependant, et contrairement à ce qui est fait habituellement dans l'analyse d'enveloppe, à savoir, un filtrage passe bande ou passe haut pour l'obtention de la composante contenant la CS2, nos composantes CS2 sont les résiduelles $\hat{V}(\Theta)$ obtenues après l'utilisation de notre méthodologie. Un exemple de ré-

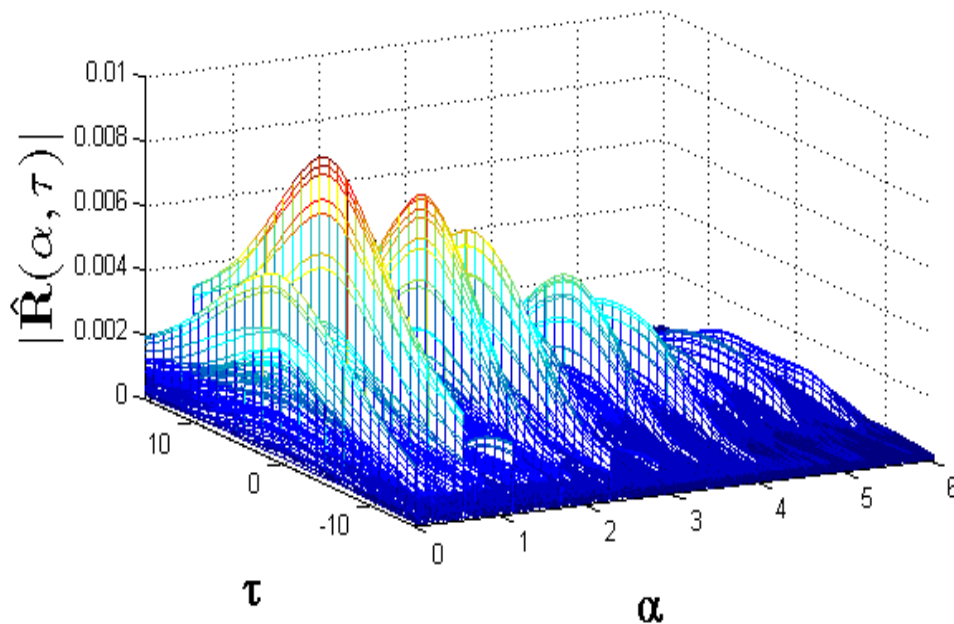


FIGURE 4.24 – Les valeurs estimées de la fonction d'autocorrélation cyclique : Marche avec fluence verbale (MF).

sultat d'analyse d'enveloppe d'un signal tiré arbitrairement est représenté sur la figure 4.25.

Les signaux provenant des deux capteurs placés au niveau des talons sont traités simultanément, ce qui signifie qu'il n'existe pas de distinction entre les pieds. De même pour les deux capteurs placés au niveau des métatarses/orteils. Ainsi, pour chaque signal, l'analyse d'enveloppe est appliquée. Puis, les amplitudes des pics correspondant aux fréquences cycliques (fréquence de la foulée et ses harmoniques) sont sommées. Enfin, les moyennes de ces sommes ainsi que leurs écarts types sont calculés pour les trois conditions de marche et pour les deux emplacements des capteurs. Les résultats de cette analyse sont présentés dans le tableau 4.26.

Les résultats obtenus montrent que d'une part les valeurs les plus importantes de la CS2 sont dans le cas de la marche avec fluence verbale (MF), suivie de la marche avec décompte (MD), et enfin de la marche simple (MS). D'autre part, la CS2 au niveau des talons semblent être plus élevée que celle au niveau du métatarse/orteils. En effet, ce dernier résultat suggère qu'il y aurait plus de variabilité et de fluctuations lors de la pause du talon que lors

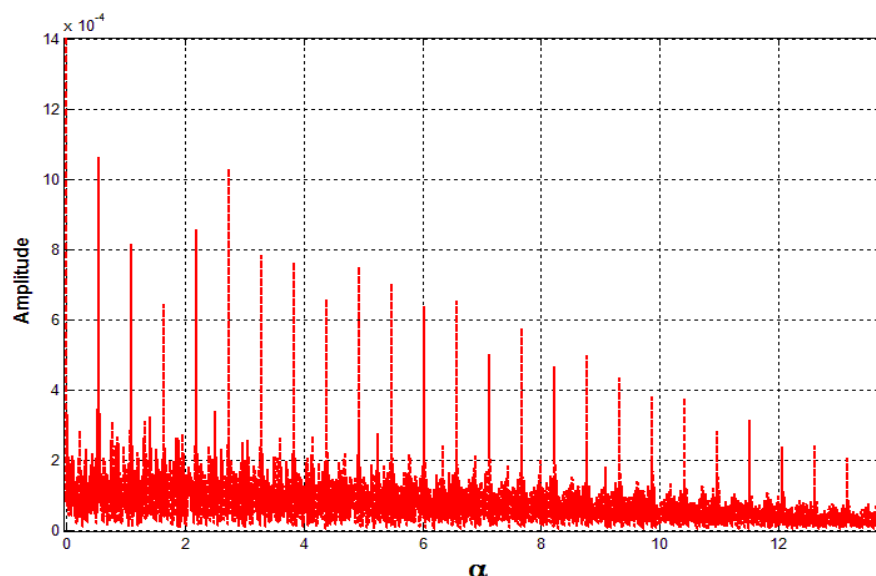


FIGURE 4.25 – Un exemple du résultat d'analyse d'enveloppe sur un signal de marche.

du levé du pied.

Ainsi, les résultats obtenus lors de cette étude suggèrent que le contrôle et l'équilibre de la marche nécessitent de l'attention. C'est pourquoi, sous un paradigme de double tâche, il existe une augmentation des fluctuations de foulée. Cette augmentation des fluctuations de foulée peut être le résultat d'une interférence entre l'attention requise pour le contrôle de la marche d'une part, et l'attention nécessaire pour effectuer la tâche cognitive d'autre part. En outre, la performance des systèmes sensori-moteurs chez les personnes âgées est diminuée ; par conséquent, le contrôle de la marche devient plus complexe et nécessite plus d'attention chez les gens âgés. Ainsi, il serait tout à fait pro-

Capteur \ Condition	MS	MD	MF
Talon	$8.03.10^{-3}$ $\pm 2.02.10^{-4}$	$32.2.10^{-3}$ $\pm 2.83.10^{-4}$	$68.3.10^{-3}$ $\pm 15.1.10^{-4}$
Métatarse/Orteils	$6.02.10^{-4}$ $\pm 3.2.10^{-4}$	$9.12.10^{-4}$ $\pm 7.11.10^{-4}$	$17.8.10^{-4}$ $\pm 7.79.10^{-4}$

FIGURE 4.26 – Les valeurs de la CS2 pour les trois conditions de marche.

bable qu'à cause du coût important de l'attention que génère la réalisation de deux tâches simultanément, l'attention soit divisée entre les deux tâches. Par conséquent, la performance de la réalisation de l'une ou des deux tâches diminue. Par ailleurs, l'augmentation des fluctuations de foulée lors de la marche avec fluence verbale, en comparaison avec la marche avec décompte, pourrait être due à la complexité de la tâche. De toute évidence, la tâche de décompte est relativement facile et est moins exigeante en termes d'attention demandée, contrairement à la tâche verbale assez complexe qui consistait à citer à haute voix des noms d'animaux commençant par une lettre imposée.

Enfin, pour résumer les résultats obtenus, il semblerait que les gens âgés auraient tendance à avoir des foulées plus longues accompagnées de fluctuations plus importantes en condition de double tâche. L'augmentation des longueurs des foulées et des fluctuations seraient proportionnelles à la difficulté de la tâche cognitive réalisée.

4.8 Conclusion

Dans ce chapitre, une nouvelle méthode de bootstrap pour l'étude de la CS a été présentée. En effet, le GSBB temporel pour l'estimateur de la moyenne synchrone $\mu(t)$ a été introduit puis décrit. Une étude de simulation a montré l'applicabilité de cette méthode sur deux modèles CS qui sont souvent rencontrés en pratique. Un test statistique a aussi été construit à partir du GSBB et l'importance de l'étude de la significativité de la moyenne synchrone a été démontrée.

Ensuite, l'application de la méthode du GSBB a permis une analyse détaillée et originale de signaux de marche de gens âgés. Cette approche a aussi permis d'étudier la très importante influence des fluctuations de vitesse sur les propriétés CS d'un signal. De plus, des indicateurs de performance de séparation de composantes CS1/CS2 ont été proposés.

Par ailleurs, des résultats très intéressants et innovants ont été présentés dans ce chapitre. En effet, le caractère CS2 des signaux de marche de gens âgés a été montré, aussi bien dans la marche simple que dans la marche à double tâche. Il a aussi été montré que la double tâche augmentait le caractère CS2 des signaux de marche. De plus, la variabilité des foulées (CS2) semble être liée au niveau de difficulté de la tâche cognitive effectuée. En outre, la CS2 semble être plus importante au niveau du talon qu'au niveau de l'avant pied.

La méthode du GSBB présentée dans ce chapitre pourrait trouver des applications dans différents domaines d'application, y compris dans la mécanique et la surveillance des machines tournantes. De nombreuses méthodes existent pour le diagnostic et la surveillance des machines, mais la première étape qui consiste à la séparation CS1/CS2 est cruciale. Malheureusement, dans beaucoup de cas, il n'y a aucune mesure de l'exactitude de la méthode utilisée, ni d'inspection de présence de structures périodiques dans les signaux supposés être CS2. En effet, cela pourrait causer la non visualisation de quelques défauts, ce qui entraînerait de faux diagnostics. La méthode ainsi que les indicateurs proposés dans ce chapitre pourraient se révéler très utiles dans ce genre de cas.

Enfin, il est tout à fait envisageable d'étendre le GSBB à l'étude de l'estimateur de la variance synchrone. Cette dernière statistique caractérise le second ordre des signaux et est très utile dans le cas de présence de CS2. En effet, la variance synchrone apporte une information sur la variabilité et les fluctuations de cycle en cycle. Ainsi, elle caractérise en quelque sorte la CS2 dans le domaine temporel en quantifiant et en localisant temporellement les différentes fluctuations (aléatoires) cycliques. Par conséquent, l'intérêt de

l'étude de l'estimateur de la variance synchrone par le biais de méthodes statistiques telle que le GSB_B devient grandissant. De plus, le développement de nouveaux indicateurs, comme ceux qui ont été proposés dans ce chapitre, pourrait être d'une très grande utilité dans l'analyse et la caractérisation des propriétés CS des signaux.

Le GSBB dans le domaine fréquentiel

Sommaire

5.1	Introduction	156
5.2	Le GSBB	156
5.3	Problématiques et consistance du GSBB	157
5.4	Étude de simulation : le GSBB fréquentiel	164
5.5	Application du GSBB sur des données réelles	169
5.6	Conclusion	172

5.1 Introduction

Dans le chapitre 4, nous avons défini le GSBB et l'avons introduit dans le domaine temporel. Dans ce présent chapitre, nous proposons d'étendre la notion du GSBB au domaine fréquentiel. En effet, la consistance du GSBB a été montrée que dans la problématique d'estimation de la CS1 dans le domaine temporel. Dans ce qui suit, nous nous focalisons sur l'estimation et la détection de la CS1 et de la CS2 dans le domaine fréquentiel. Le problème d'estimation des coefficients de Fourier de la moyenne ainsi que de la fonction d'autocorrélation d'un signal CS est présenté. Les résultats obtenus permettent non seulement d'obtenir les estimateurs d'intérêt, mais aussi de construire des intervalles de confiance valides pour les caractéristiques CS d'un signal dans le domaine fréquentiel. Par conséquent, il devient possible de construire des tests de présence de structures CS1 et CS2 dans un signal donné. De plus, dans le cas de signaux réels, le nombre d'échantillons disponible est fini et n'est généralement pas assez grand pour assurer que le régime asymptotique est atteint ce qui se traduit par une erreur estimée gaussienne. Nous verrons par la suite que puisque les intervalles de confiance du GSBB ne sont pas basés sur la distribution asymptotique, ils peuvent ne pas être symétriques par rapport à la valeur estimée (voir les résultats de la section 5.4 ce qui reflète l'asymétrie de la distribution du ré-échantillonnage).

Ce dernier chapitre est organisé comme suit : dans la section 2, la méthode du GSBB est rappelée. Dans la section 3, la problématique est formulée et les résultats principaux de la consistance du GSBB dans le domaine fréquentiel sont présentés. La section 4 contient une étude de simulation pour valider la méthode. Dans la section 5, une application sur un signal réel est exposée. Enfin, la section 6 résume les principaux résultats obtenus et présente quelques problèmes ouverts pour de futurs travaux.

5.2 Le GSBB

Dans cette partie, le GSBB est décrit. De plus, quelques nouvelles notations sont introduites. Pour $\{X(1), \dots, X(n)\}$ disponible de $X = \{X(t) : t \in \mathbb{Z}^+\}$, soit B_i le bloc de taille b qui a pour point de départ l'observation $X(i)$; i.e., $B_i = (X(i), \dots, X(i + b - 1))$.

Nous rappelons que nous nous focalisons sur une version circulaire du GSBB afin d'éviter de possibles effets de bords et que pour des raisons de simplicité, nous supposons que la taille des données n est un multiple entier de la période T ($n = KT$) et de taille de bloc b ($n = lb$). Ces hypothèses

permettent seulement de simplifier les notations et ne sont aucunement essentielles. L'algorithme du GSBB suivant ces notations est donc le suivant :

- (1) Choisir un entier positif comme taille de bloc $b (< n)$ tel que $l = n/b$ est entier.
- (2) Pour $t = 1, b + 1, 2b + 1, \dots, lb$ soit

$$(X^*(t), X^*(t+1), \dots, X^*(t+b-1)) = (X(h_t), X(h_t+1), \dots, X(h_t+b-1))$$

Avec h_t est iid avec une distribution uniforme telle que :

$$P(h_t = t + vT) = \frac{1}{K} \quad \text{pour } v = 0, 1, \dots, K - 1.$$

Puisque nous considérons la version circulaire du GSBB, alors quand $t + vT > n$ on prend $h_t = t + vT - n$.

- (3) Joindre les l blocs B_1^*, \dots, B_l^* ensemble afin de former la nouvelle série de pseudo-observations bootstrap $X^*(1), X^*(2), \dots, X^*(n)$ ayant la même taille n que les données de départ et préservant aussi leur périodicité.

Ainsi, les problématiques d'estimation ainsi que les résultats principaux de la consistance du GSBB pour la CS1 et la CS2 dans le domaine fréquentiel sont présentés dans ce qui suit.

5.3 Problématiques et consistance du GSBB

Dans cette partie, une brève description des problématiques d'estimation des coefficients de Fourier de la moyenne et de la fonction d'autocorrélation d'un processus CS est présentée. Ensuite, les résultats de consistance du GSBB sont exposés.

L'inférence pour les signaux est souvent basée sur le domaine fréquentiel. En effet, l'analyse fréquentielle des signaux non-stationnaires est complexe et nécessite beaucoup de précaution. L'analyse de tels signaux dans le domaine fréquentiel est très souvent basée sur l'usage de fenêtres temporelles sur lesquelles le signal étudié est considéré stationnaire. Cette approche, bien

qu'efficace présente quelques problématiques. Ces problèmes sont essentiellement dus aux tailles des fenêtres choisies. Ainsi, les résultats obtenus dépendent souvent de ces tailles de fenêtres [95]. La CS étant un cas particulier de la non-stationnarité, dans ce qui suit, nous proposons d'appliquer la méthode du GSBB qui est décrite dans la section 5.2 afin d'accomplir l'inférence dans le domaine fréquentiel. Cette approche permettra d'effectuer des tests fréquentiels et de construire des intervalles de confiance valides.

Les décompositions en séries de Fourier de la fonction de moyenne et de la fonction d'autocovariance d'un signal cyclostationnaire $X(t)$ sont données par :

$$\begin{aligned}\mu(t) &= \sum_{\gamma \in \Gamma} b(\gamma) e^{j2\pi\gamma t}, \\ B_X(t, \tau) &= \sum_{\alpha \in \mathcal{A}_\alpha} B_X(\alpha, \tau) e^{j2\pi\alpha t},\end{aligned}$$

Avec $B_X(t, \tau) = \text{Cov}(X(t), X(t + \tau))$ et

$$\begin{aligned}\Gamma &= \{\gamma : b(\gamma) \neq 0\}, \\ \mathcal{A}_\alpha &= \{\alpha : B_X(\alpha, \tau) \neq 0\}.\end{aligned}$$

Alors, Γ et \mathcal{A}_α sont des ensembles finis des fréquences de la CS1 et la CS2. Ces ensembles sont tous deux des sous ensembles de l'ensemble $\{k/T, k = 0, \dots, T - 1\}$.

Les coefficients de Fourier $b(\gamma)$ et $B_X(\alpha, \tau)$ sont définis par les moyennages temporels suivants :

$$b(\gamma) = \langle \mu(t) e^{-j2\pi\gamma t} \rangle_t, \quad (5.1)$$

$$B_X(\alpha, \tau) = \langle B_X(t, \tau) e^{-j2\pi\alpha t} \rangle_t \quad (5.2)$$

Ainsi, leurs estimateurs respectifs ont les formes suivantes :

$$\begin{aligned}\hat{b}_n(\gamma) &= \frac{1}{n} \sum_{t=1}^n X(t) e^{-j2\pi\gamma t}, \\ \hat{B}_n(\alpha, \tau) &= \frac{1}{n - |\tau|} \sum_{t=1 - \min\{\tau, 0\}}^{n - \max\{\tau, 0\}} (X(t + \tau) - \hat{\mu}_n(t + \tau)) (X(t) - \hat{\mu}_n(t)) e^{-j2\pi\alpha t},\end{aligned}$$

Avec

$$\widehat{\mu}_n(t) = \sum_{\gamma \in \Gamma} \widehat{b}_n(\gamma) e^{-j2\pi\gamma t}.$$

Dans la suite, nous considérons une version simplifiée de $\widehat{B}_n(\alpha, \tau)$ en supposant que $E(X(t)) \equiv 0$ (ce qui implique que $b(\gamma) \equiv 0$) afin d'obtenir l'estimateur de la fonction d'autocorrélation cyclique $\widehat{R}_n(\alpha, \tau)$ (en prenant en compte que la composante CS2). Donc pour $\tau \geq 0$:

$$\widehat{R}_n(\alpha, \tau) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} X(t)X(t + \tau) e^{-j2\pi\alpha t}.$$

Dès lors, les versions bootstrap de $\widehat{b}_n(\gamma)$ et $\widehat{R}_n(\alpha, \tau)$ sont donnés comme suit :

$$\widehat{b}_n^*(\gamma) = \frac{1}{n} \sum_{t=1}^n X^*(t) e^{-j2\pi\gamma t} \quad \text{et} \quad (5.3)$$

$$\widehat{R}_n^*(\alpha, \tau) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} X^*(t)X^*(t + \tau) e^{-j2\pi\alpha t}, \quad (5.4)$$

Tel que $(X(1)^*, \dots, X(n)^*)$ est obtenu après l'étape (3) de la méthode du GSBB en joignant les l blocs B_1^*, \dots, B_{lb}^* .

Il est fondamental de respecter les indices de temps. En effet, si :

$$(X^*(t), X^*(t + 1), \dots, X^*(t + b - 1)) = (X(h), X(h + 1), \dots, X(h + b - 1))$$

Pour $h \in \{1, \dots, n\}$, alors $(t^*, \dots, (t + b - 1)^*) = (h, \dots, h + b - 1)$. De cette façon, l'indice de temps t^* correspondra à l'indice de temps du bloc choisis et peut avoir des valeurs dans $\{1, \dots, n\}$. Ce dernier point est très essentiel. En effet, en procédant de la sorte, les indices du temps sont respectés et les échantillons formant les blocs utilisés dans la procédure du bootstrap gardent leurs indices de temps originaux.

Avant de présenter les résultats de consistance, des notations supplémentaires sont introduites. En effet, on définit les vecteurs suivants :

$$\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_r)^T, \quad \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)^T, \quad \boldsymbol{\tau} = (\tau_1, \dots, \tau_r)^T,$$

$$b(\boldsymbol{\gamma}) = (\Re(b(\gamma_1)), \Im(b(\gamma_1)), \dots, \Re(b(\gamma_r)), \Im(b(\gamma_r)))^T,$$

$$R(\boldsymbol{\alpha}, \boldsymbol{\tau}) = (\Re(R(\alpha_1, \tau_1)), \Im(R(\alpha_1, \tau_1)), \dots, \Re(R(\alpha_r, \tau_r)), \Im(R(\alpha_r, \tau_r)))^T.$$

De plus, nous notons par $\widehat{b}_n(\boldsymbol{\gamma})$ et $\widehat{R}_n(\boldsymbol{\alpha}, \boldsymbol{\tau})$ les estimateurs respectifs de $b(\boldsymbol{\gamma})$, $R(\boldsymbol{\alpha}, \boldsymbol{\tau})$ et par $\widehat{b}_n^*(\boldsymbol{\gamma})$ et $\widehat{R}_n^*(\boldsymbol{\alpha}, \boldsymbol{\tau})$ leurs estimateurs issus du bootstrap.

En outre, nous supposons que $X(t)$ satisfait quelques conditions de mélange α (α -mixing, à ne pas confondre avec les fréquences cycliques α); i.e., $\alpha_X(k) \rightarrow 0$ avec $k \rightarrow \infty$, tel que :

$$\alpha_X(k) = \sup_t \sup_{\substack{A \in \mathcal{F}_X(-\infty, t) \\ B \in \mathcal{F}_X(t+k, \infty)}} |P(A \cap B) - P(A)P(B)|$$

Avec $\mathcal{F}_X(t_1, t_2)$ une σ -algèbre (ou tribu) générée par $\{X(t) : t_1 \leq t \leq t_2\}$ et $A, B \subset \mathcal{F}_X(t_1, t_2)$.

Nous rappelons que α -mixing est un type de mesure de dépendance faible. En effet, si $\alpha(k) = 0$, alors les observations qui sont distantes d'au moins k unités de temps sont indépendantes. Pour plus de détails, nous suggérons [45].

Dans ce qui suit, les résultats asymptotiques de l'estimateur des coefficients de Fourier de la fonction de moyenne sont présentés pour $r = 1$ (unidimensionnel). Ensuite, les résultats dédiés à la consistance multidimensionnelle du GSBB pour cet estimateur sont donnés.

Theorème 5.3.1 *Soit $\{X(t), t \in \mathbb{Z}\}$ un signal CS et :*

$$(i) \sup_{t \in \mathbb{Z}} \mathbb{E}|X(t)|^{4+\delta} < \infty \text{ pour } \delta > 0;$$

$$(ii) \sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(4+\delta)}(\tau) < \infty;$$

Si $b \rightarrow \infty$ quand $n \rightarrow \infty$ tel que $b = o(n)$, alors

$$\text{Var}^* \left(\sqrt{n} \widehat{b}_n^*(\boldsymbol{\gamma}) \right) \xrightarrow{p} \Sigma(\boldsymbol{\gamma}), \quad (5.5)$$

$$\sup_{(x,y) \in \mathcal{R}^2} \left| P \left(\sqrt{n} \left(\widehat{b}_n(\boldsymbol{\gamma}) - b(\boldsymbol{\gamma}) \right) \leq (x, y) \right) - P^* \left(\sqrt{n} \left(\widehat{b}_n^*(\boldsymbol{\gamma}) - \mathbb{E}^* \widehat{b}_n(\boldsymbol{\gamma}) \right) \leq (x, y) \right) \right| \xrightarrow{p} 0, \quad (5.6)$$

Avec

$$\Sigma(\gamma) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

$$\sigma_1^2 = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov}(X(s), X(s+\tau)) \cos(2\pi\gamma s) \cos(2\pi\gamma k),$$

$$\sigma_2^2 = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov}(X(s), X(s+\tau)) \sin(2\pi\gamma s) \sin(2\pi\gamma k),$$

$$\sigma_{12} = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov}(X(s), X(s+\tau)) \cos(2\pi\gamma s) \sin(2\pi\gamma k)$$

Avec $\det(\Sigma(\gamma)) \neq 0$.

Alors, ces résultats montrent que l'estimateur des coefficients de Fourier de la fonction de moyenne est asymptotiquement normal avec la matrice de variance-covariance $\Sigma(\gamma)$. Il est important de souligner que cette matrice est assez complexe. Par conséquent, son estimation en pratique l'est aussi.

Par ailleurs, l'argument de P doit être lu $\left(\mathfrak{R} \left(\sqrt{n} \left(\widehat{b}_n(\gamma) - b(\gamma) \right) \right) \leq x \right)$, $\left(\mathfrak{S} \left(\sqrt{n} \left(\widehat{b}_n(\gamma) - b(\gamma) \right) \right) \leq y \right)$. De plus, P^* représente la probabilité conditionnelle. En effet, étant donné un échantillon $(X(1), \dots, X(n))$, E^* , Var^* représentent respectivement l'espérance conditionnelle et la variance conditionnelle.

À noter que les conditions nécessaires pour l'obtention du théorème central limite pour $\widehat{b}(\gamma)$ sont plus faibles que dans (i) et (ii). En effet, la normalité asymptotique de l'estimateur a été obtenue et montrée dans [138] pour des moments finis d'ordre $2 + \delta$ et des conditions de α -mixing de la forme $\sum_k \alpha(k)^{\delta/(2+\delta)} < \infty$.

D'autre part, le résultat de consistance du GSBB pour les coefficients de Fourier de la fonction de moyenne est comme suit :

Théorème 5.3.2 *Sous les hypothèses du théorème 5.3.3 on a*

$$\rho \left(\mathcal{L} \left(\sqrt{n} \left(\widehat{b}_n(\gamma) - b(\gamma) \right) \right), \mathcal{L}^* \left(\sqrt{n} \left(\widehat{b}_n^*(\gamma) - E^* \widehat{b}_n^*(\gamma) \right) \right) \right) \xrightarrow{P} 0 \quad (5.7)$$

Avec ρ une métrique de mesure de convergence faible en probabilité sur \mathcal{R}^{2r} . De plus, $\mathcal{L} \left(\sqrt{n} \left(\widehat{b}_n(\gamma) - b(\gamma) \right) \right)$ est une loi de probabilité de $\sqrt{n} \left(\widehat{b}_n(\gamma) - b(\gamma) \right)$ et $\mathcal{L}^* \left(\sqrt{n} \left(\widehat{b}_n^*(\gamma) - E^* \widehat{b}_n^*(\gamma) \right) \right)$ sa version de bootstrap.

Ainsi, la consistance du GSBB pour l'estimateur de $b(\gamma)$ est obtenue.

Dans la partie suivante, les résultats asymptotiques de l'estimateur des coefficients de Fourier de la fonction d'autocorrélation sont présentés. Puis, les résultats de consistance du GSBB pour cet estimateur sont donnés.

Theorème 5.3.3 *Soit $\{X(t), t \in \mathbb{Z}\}$ un signal CS avec $E(X(t)) \equiv 0$ et $PF(4)$ tels que :*

$$(i) \sup_{t \in \mathbb{Z}} E|X(t)|^{8+2\delta} < \infty \text{ pour } \delta > 0;$$

$$(ii) \sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(4+\delta)}(\tau) < \infty;$$

Si $b \rightarrow \infty$ quand $n \rightarrow \infty$ tel que $b = o(n)$, alors

$$\text{Var}^* \left(\sqrt{n} \widehat{R}_n^*(\alpha, \tau) \right) \xrightarrow{p} \Sigma(\alpha, \tau), \quad (5.8)$$

$$\sup_{(x,y) \in \mathbb{R}^2} \left| P \left(\sqrt{n} \left(\widehat{R}_n(\alpha, \tau) - R(\alpha, \tau) \right) \leq (x, y) \right) - P^* \left(\sqrt{n} \left(\widehat{R}_n^*(\alpha, \tau) - E^* \widehat{R}_n^*(\alpha, \tau) \right) \leq (x, y) \right) \right| \xrightarrow{p} 0, \quad (5.9)$$

Avec

$$\Sigma(\alpha, \tau) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

$$\sigma_1^2 = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov} (X(s)X(s+\tau), X(s+k)X(s+k+\tau)) \beta_1,$$

$$\sigma_2^2 = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov} (X(s)X(s+\tau), X(s+k)X(s+k+\tau)) \beta_2,$$

$$\sigma_{12} = \frac{1}{T} \sum_{s=1}^T \sum_{k=-\infty}^{\infty} \text{Cov} (X(s)X(s+\tau), X(s+k)X(s+k+\tau)) \beta_3$$

Tels que

$$\beta_1 = \cos(2\pi\alpha s) \cos(2\pi\alpha k),$$

$$\beta_2 = \sin(2\pi\alpha s) \sin(2\pi\alpha k),$$

$$\beta_3 = \cos(2\pi\alpha s) \sin(2\pi\alpha k)$$

Et $\det(\Sigma(\alpha, \tau)) \neq 0$.

Nous précisons que PF(4) fait référence à un processus périodique faiblement à l'ordre 4, avec la période T . Cela signifie que les moments d'ordre 4 de $X(t)$ sont supposés être périodiques ; i.e., $\forall t, \tau_1, \tau_2, \tau_3$ on a :

$$\begin{aligned} & \mathbb{E} (X(t)X(t + \tau_1)X(t + \tau_2)X(t + \tau_3)) \\ &= \mathbb{E} (X(t + T)X(t + \tau_1 + T)X(t + \tau_2 + T)X(t + \tau_3 + T)). \end{aligned}$$

Pour plus de détails sur les processus périodiques faiblement, se référer à [137].

Ainsi, l'estimateur des coefficients de Fourier de la fonction d'autocorrélation est asymptotiquement normal avec la matrice de variance-covariance $\Sigma(\alpha, \tau)$. Comme pour le cas de l'estimateur des coefficients de Fourier de la fonction de moyenne, il est important de noter que la matrice de variance-covariance asymptotique $\Sigma(\alpha, \tau)$ a une forme très complexe, ce qui rend son utilité en pratique délicate, voire quasi-impossible. Ces formes complexes des matrices de variance-covariance asymptotiques justifient correctement l'utilisation de méthodes de ré-échantillonnage statistique, tel que le bootstrap.

Sous ces notations et hypothèses, on obtient les résultats suivants de consistance du GSBB pour l'estimateur des coefficients de Fourier de la fonction d'autocorrélation :

Théorème 5.3.4 *Sous les hypothèses du théorème 5.3.3 on a*

$$\begin{aligned} & \rho \left(\mathcal{L} \left(\sqrt{n} \left(\widehat{R}_n(\alpha, \tau) - R(\alpha, \tau) \right) \right) \right), \\ & \mathcal{L}^* \left(\sqrt{n} \left(\widehat{R}_n^*(\alpha, \tau) - \mathbb{E}^* \widehat{R}_n^*(\alpha, \tau) \right) \right) \xrightarrow{p} 0, \end{aligned}$$

Des exemples de théorème central limite pour $\widehat{R}_n(\alpha, \tau)$ sous des hypothèses de α -mixing pour des processus CS peuvent être trouvés dans [103] et [138]. En effet, il a été montré que les conditions nécessaires sont des moments d'ordre $4 + \delta$ finis et des conditions de α -mixing de la forme $\sum_k \alpha(k)^{\delta/(2+\delta)} < \infty$. Par conséquent, sous nos conditions (i) et (ii) le théorème central limite reste valide. Nous précisons que l'annexe B contient plus de détails sur les résultats de consistance du GSBB présentés ci-dessus.

Enfin, la consistance du GSBB pour les coefficients de Fourier de la fonction de moyenne ainsi que de la fonction d'autocorrélation permet d'obtenir

des intervalles de confiances valides. Par exemple, soit τ le paramètre de retard et l'ensemble $\{\alpha_1, \dots, \alpha_k\}$ des fréquences cycliques. Supposons que nous considérons la partie réelle de $R(\alpha, \tau)$. Alors nous avons :

$$\begin{aligned} & \sup_t \left| P \left(\sqrt{n} \max_i \Re \left(\hat{R}_n(\alpha_i, \tau) - R(\alpha, \tau) \right) \leq t \right) \right. \\ & \left. - P^* \left(\sqrt{n} \max_i \Re \left(\hat{R}_n^*(\alpha_i, \tau) - E^* \hat{R}_n^*(\alpha_i, \tau) \right) \leq t \right) \right| \xrightarrow{p} 0. \end{aligned} \quad (5.10)$$

Ainsi, il devient simple de dériver des intervalles de confiance simultanés pour l'estimateur de $R(\alpha, \tau)$.

Nous précisons que de la même façon, nous obtenons les intervalles de confiance simultanés pour l'estimateur de $b(\gamma)$.

5.4 Étude de simulation : le GSBB fréquentiel

Cette partie est consacrée à la simulation et à l'illustration de la performance ainsi que la robustesse du GSBB dans le domaine fréquentiel. L'étude ci-dessous est effectuée en considérant différents rapports de signal à bruit (RSB). Les valeurs du RSB utilisées sont les suivantes : -5 dB, 0 dB, 10 dB et 20 dB.

Afin d'illustrer l'estimation de la CS1 et la CS2 basée sur le GSBB dans le domaine fréquentiel, l'attention est focalisée sur un modèle cyclostationnaire théorique. Soit $\{U(t), t \in \mathbb{Z}\}$ un processus stationnaire gaussien et de moyenne nulle. Soit $b(t)$ un bruit additif, gaussien et corrélé tels que $U(t)$ et $b(t)$ soient indépendants. Alors, le modèle cyclostationnaire étudié est la somme de deux contributions avec un bruit, la première est déterministe alors que la seconde est CS2, soit :

$$X(t) = \cos(2\pi f_0 t) + u(t) \cos(2\pi f_0 t) + b(t),$$

Avec $t \in \mathbb{Z}$; Un nombre d'échantillons $n = 2048$ et f_0 est égale à $0.1Hz$.

Les ensembles théoriques des fréquences de $X(t)$ respectivement pour le premier et le second ordre sont :

$$\Gamma = \{f_0\}, \quad \mathcal{A}_\alpha = \{-2f_0, 0, 2f_0\},$$

Soit, numériquement :

$$\Gamma = \{0.1\}, \quad \mathcal{A}_\alpha = \{-0.2, 0, 0.2\},$$

Enfin, les problèmes de testing de CS1 et CS2 sont définis comme suit :

$$\begin{cases} H_0 & : b(\gamma) = 0 \\ H_1 & : b(\gamma) \neq 0 \end{cases}$$

et pour $\alpha \neq 0$

$$\begin{cases} H_0 & : R(\alpha, \tau) = 0 \\ H_1 & : R(\alpha, \tau) \neq 0. \end{cases}$$

Afin de réaliser ces tests, des intervalles de confiance simultanés sont construits pour les estimateurs de $|b(\gamma)|$ et de $|R(\alpha, \tau)|$. Pour ce faire, un grand nombre $B = 500$ de pseudo-données bootstrap $X_1^*, X_2^*, \dots, X_n^*$ est obtenu par répétition de l'algorithme du GSBB. Ensuite, pour chaque pseudo-données bootstrap, $|\hat{b}(\gamma)^*|$ et $|\hat{R}(\alpha, \tau)^*|$ sont estimés. Puis, le résultat 5.10 est utilisé afin d'obtenir les quantiles des distributions du bootstrap correspondant aux $\frac{\delta}{2}$ et $1 - \frac{\delta}{2}$, avec δ le niveau de confiance désiré, et ce, pour chaque estimateur. Notons maintenant ces quantiles par $c_{\hat{b}, B}(\frac{\delta}{2})$ et $c_{\hat{b}, B}(1 - \frac{\delta}{2})$, pour l'estimateur de $|b(\gamma)|$ et par $c_{\hat{R}, B}(\frac{\delta}{2})$ et $c_{\hat{R}, B}(1 - \frac{\delta}{2})$ pour l'estimateur de $|R(\alpha, \tau)|$, alors les intervalles de confiances sont les suivants :

$$\left(|\hat{b}(\gamma)| - \frac{c_{\hat{b}, B}(1 - \frac{\delta}{2})}{\sqrt{n}}, \quad |\hat{b}(\gamma)| - \frac{c_{\hat{b}, B}(\frac{\delta}{2})}{\sqrt{n}} \right),$$

$$\left(|\hat{R}(\alpha, \tau)| - \frac{c_{\hat{R}, B}(1 - \frac{\delta}{2})}{\sqrt{n}}, \quad |\hat{R}(\alpha, \tau)| - \frac{c_{\hat{R}, B}(\frac{\delta}{2})}{\sqrt{n}} \right),$$

respectivement pour $|\hat{b}(\gamma)|$ et $|\hat{R}(\alpha, \tau)|$.

Ainsi, des intervalles de confiance pour les estimateurs de $|\hat{b}(\gamma)|$ et $|\hat{R}(\alpha, \tau)|$ sont construits pour chaque valeur de RSB.

Les figures 5.1a, b, c et d montrent les valeurs estimées de la fonction d'autocorrélation cyclique $|\hat{R}(\alpha, \tau)|$, respectivement pour des valeurs de RSB égales à -5dB , 0dB , 10dB et 20dB . Comme il a été mentionné auparavant, l'ensemble des fréquences cycliques est sous la forme $\mathcal{A}_\alpha = \{-0.2, 0, 0.2\}$. Etant donné que la fonction d'autocorrélation cyclique présente des propriétés de symétrie, nous nous intéressons que sur les parties positives des fréquences cycliques α . De plus, il est important de souligner que la fonction d'autocorrélation cyclique est estimée après la soustraction de la composante d'ordre un du signal $X(t)$.

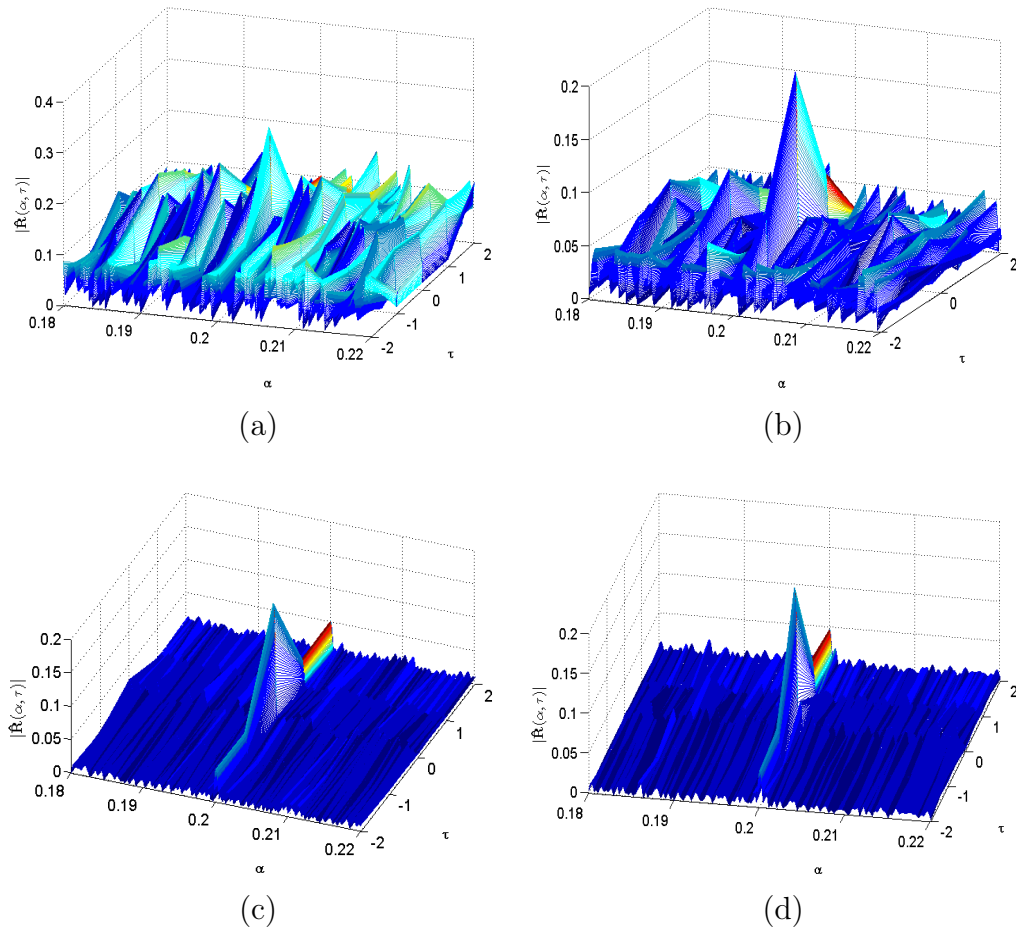


FIGURE 5.1 – Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ sur le signal simulé $X(t)$: (a). RSB = -5 dB ; (b). RSB = 0 dB ; (c). RSB = 10 dB ; (d). RSB = 20 dB.

Les paramètres du GSBB utilisés dans cette simulations sont les suivants : un niveau de confiance $\delta = 0.05$, le nombre de répétitions du GSBB $B = 500$ et une taille de bloc b égale à $\lfloor \sqrt{n} \rfloor$.

Les figures 5.2, 5.3, 5.4 et 5.5 montrent les valeurs estimées de $|\hat{b}(\gamma)|$ et $|\hat{R}(\alpha, \tau)|$ ainsi que leurs intervalles de confiance construits à partir du GSBB respectivement pour les valeurs de RSB : -5dB, 0dB, 10dB et 20dB. Nous tenons à remarquer que pour des raisons de simplicité et de visualisation, les résultats relatifs à $|\hat{R}(\alpha, \tau)|$ sont montrés pour une valeur fixe de $\tau = 0$.

Les figures 5.2a, 5.3a, 5.4a et 5.5a montrent nettement le non rejet de l'hypothèse H_0 pour toutes les fréquences γ testées, excepté dans le cas où $\gamma = 0.1Hz$. Cela signifie que la procédure du GSBB a clairement réussi à détecter la fréquence significatives au premier ordre $\gamma = f_0 = 0.1Hz$, et ce,

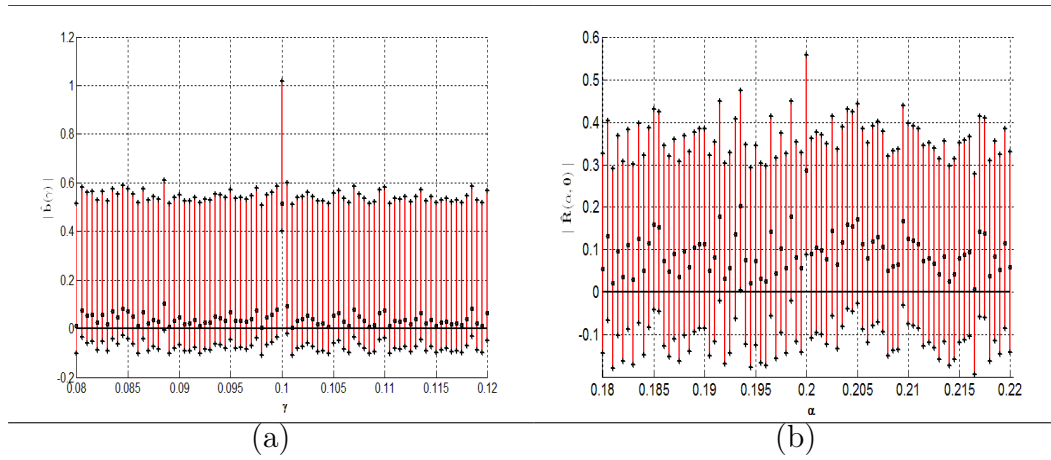


FIGURE 5.2 – (a) Les valeurs estimées de $|\hat{b}(\gamma)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=-5dB.

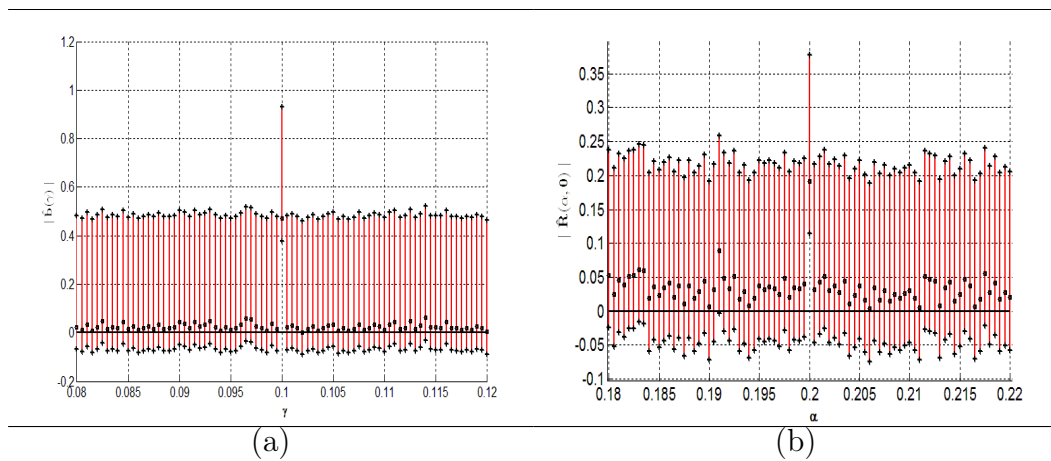


FIGURE 5.3 – (a) Les valeurs estimées de $|\hat{b}(\gamma)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=0dB.

pour toutes les valeurs du RSB.

Par ailleurs, les figures 5.2b, 5.3b, 5.4b et 5.5b montrent clairement le non rejet de l'hypothèse H_0 pour toutes les fréquences cycliques α testées, excepté dans le cas où $\alpha = 0.2Hz$. Ce résultat signifie que le GSBB a réussi à détecter la fréquence cyclique (CS2) $\alpha = 2f_0 = 0.2Hz$, et ce, pour toutes les

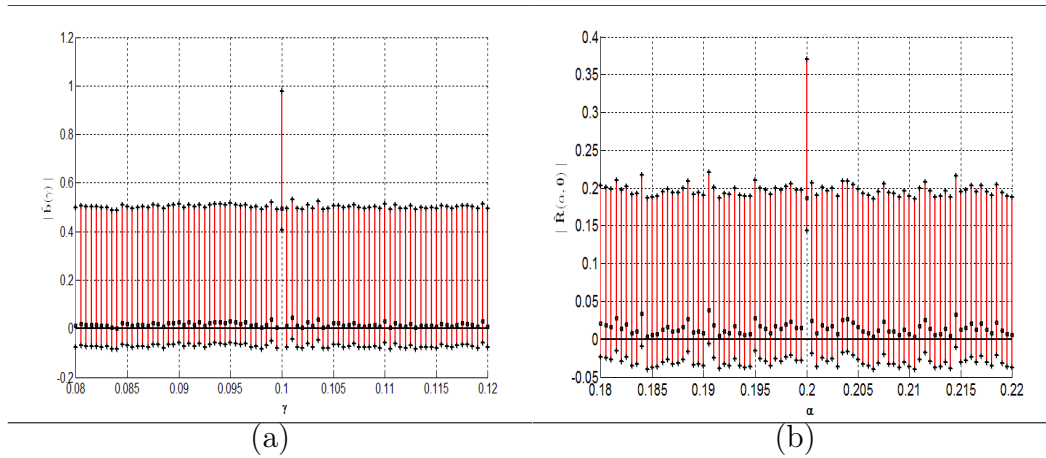


FIGURE 5.4 – (a) Les valeurs estimées de $|\hat{b}(\gamma)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=10dB.

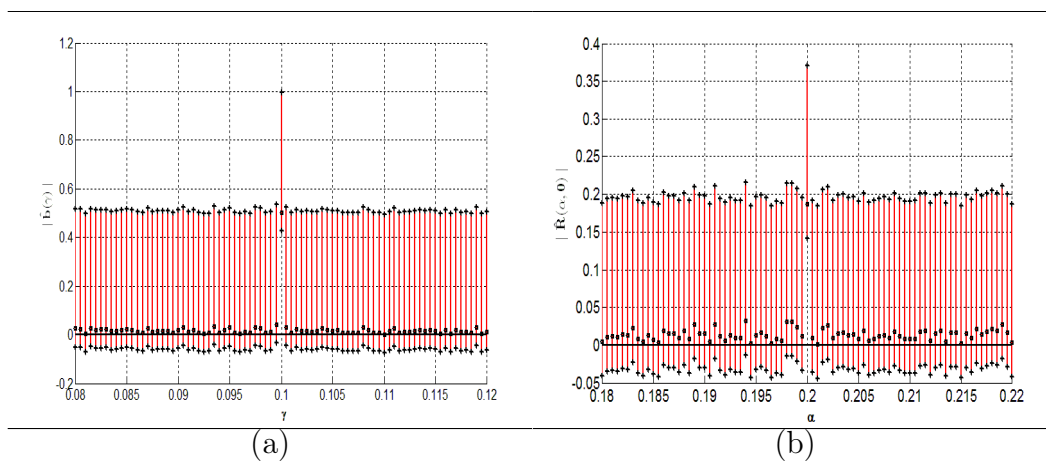


FIGURE 5.5 – (a) Les valeurs estimées de $|\hat{b}(\gamma)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). (b) Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ (points noirs) avec les 95% d'intervalles de confiance issus du GSBB (les lignes verticales). RSB=20dB.

valeurs considérées du RSB.

Les résultats ci-dessus suggèrent que la méthode du GSBB est très efficace dans la détection des fréquences d'un processus CS, que ce soit à l'ordre un (CS1) ou à l'ordre deux (CS2). En outre, le GSBB semble être très robuste au bruit, puisqu'il a réussi à détecter les fréquences théoriquement significatives

malgré des rapports signal à bruit assez bas (RSB=-5dB, 0dB).

Afin d'étudier l'applicabilité du GSBB sur des signaux réels ainsi que son efficacité, la section suivante est dédiée à l'application du GSBB sur un signal réel. Le signal sous étude est un signal de marche.

5.5 Application du GSBB sur des données réelles

Dans le chapitre précédent, nous avons vu le caractère cyclostationnaire à l'ordre un et à l'ordre deux des signaux de marche de gens âgés et nous avons appliqué le GSBB dans le domaine temporel et développé divers indicateurs. Dans cette dernière partie de ce chapitre, nous appliquons le GSBB sur un signal de marche, mais dans le domaine fréquentiel. En effet, les propriétés CS1 et CS2 d'un signal de marche sont étudiées et le GSBB est validé sur des signaux réels.

La figure 5.6 montre le signal de marche utilisé dans cette étude. Ce signal a été acquis dans les mêmes conditions expliquées dans le chapitre 4.

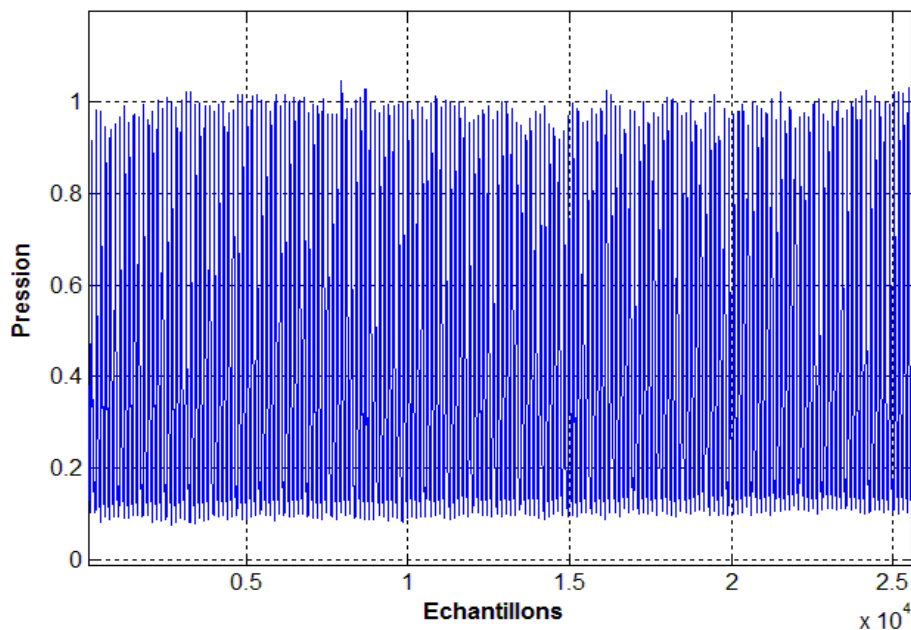


FIGURE 5.6 – Signal de marche.

Tout d'abord, les valeurs de $|b(\gamma)|$ et $|R(\alpha, \tau)|$ sont estimées pour le signal de marche. Nous rappelons que pour l'estimation de $|\hat{R}(\alpha, \tau)|$,

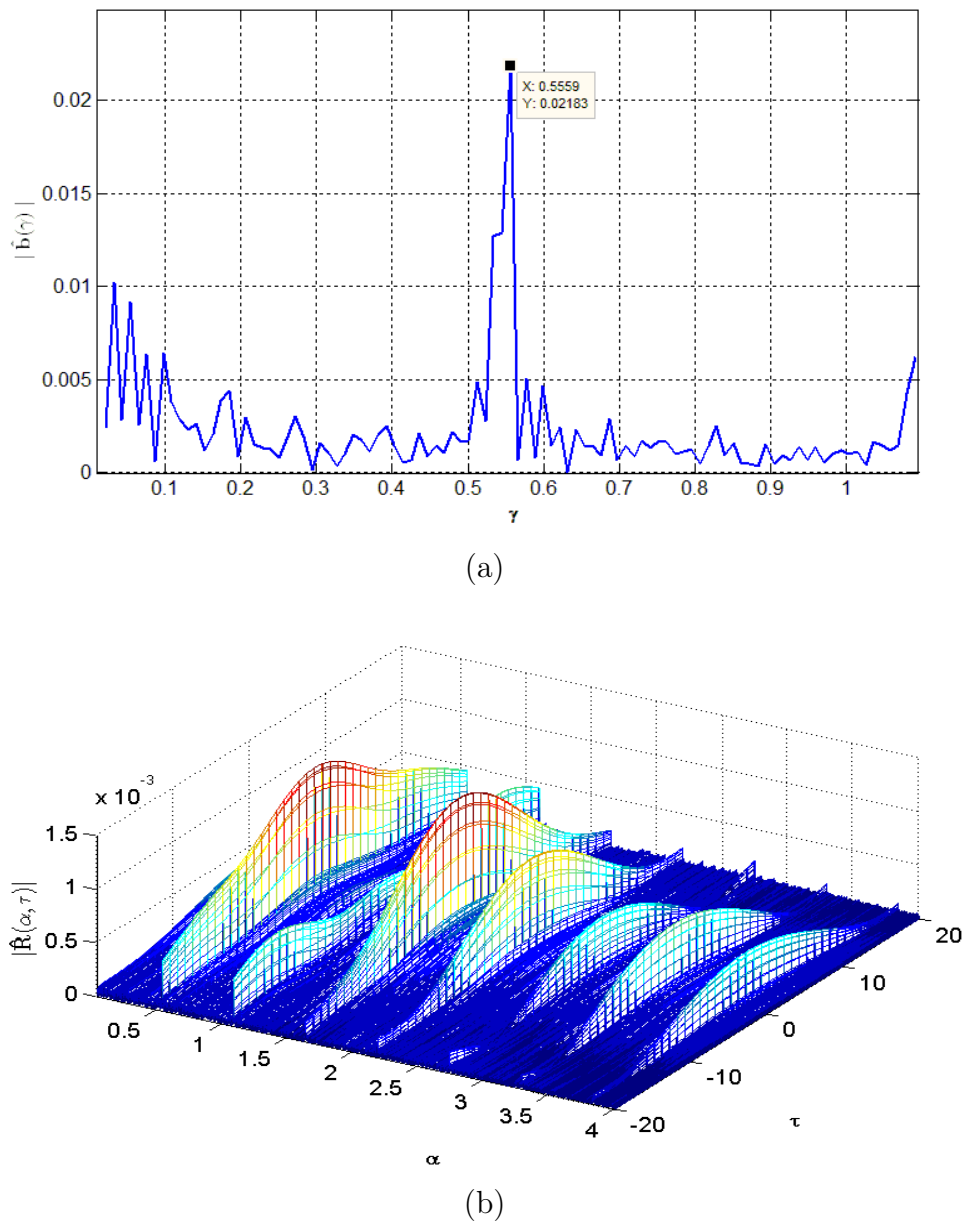


FIGURE 5.7 – (a). Les valeurs estimées de $|\hat{b}(\gamma)|$ pour le signal de marche ; (b). Les valeurs estimées de $|\hat{R}(\alpha, \tau)|$ pour le résiduel du signal de marche.

la méthodologie présentée dans le chapitre 4 est utilisée. Les résultats sont montrés dans la figure 5.7.

Dans la figure 5.7a, les valeurs estimées de $|b(\gamma)|$ sont présentées. La simple visualisation de la figure ne permet pas de reconnaître quelles fréquences sont significatives et quelles fréquences ne le sont pas. En effet, il

existe plusieurs fréquences qui sont candidates à être significatives. Par conséquent, l'ensemble des fréquences significatives du premier ordre semble être assez large.

Par ailleurs, la figure 5.7a montre les valeurs de $|R(\alpha, \tau)|$ estimées sur la partie résiduelle du signal de marche. Les résultats présentés cette figure suggèrent que les fréquences significatives au second ordre sont les multiples de la fréquence fondamentale $\alpha = 0.548Hz$. Cette dernière fréquence correspond à la fréquence de la foulée du sujet étudié.

Maintenant, nous appliquons le GSBB puis nous construisons des intervalles de confiance à 95% de significativité pour les estimateurs de $|b(\gamma)|$ et $|R(\alpha, \tau)|$. La taille du bloc utilisé b est égale à $\lfloor \sqrt{n} \rfloor$ ($= 160$).

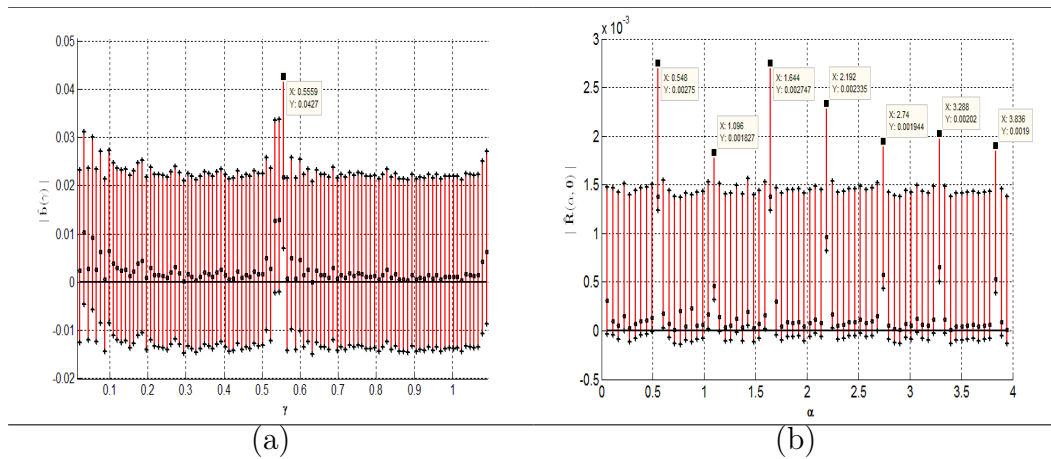


FIGURE 5.8 – (a). Les valeurs estimées de $|b(\gamma)|$ (points noirs) avec 95% d'intervalles de confiance GSBB (lignes verticales) pour le signal de marche; (b). Les valeurs estimées de $|R(\alpha, \tau)|$ (points noirs) avec 95% d'intervalles de confiance GSBB (lignes verticales) pour la composante CS2 du signal de marche.

La figure 5.8 montre les valeurs estimées de $|b(\gamma)|$ et $|R(\alpha, \tau)|$ ensemble avec leurs intervalles de confiance respectifs construits à l'aide du GSBB.

La figure 5.8a montre que le GSBB a réussi à détecter la fréquence $\gamma = 0.5559Hz$ comme significative. Cela signifie que l'hypothèse H_0 n'est pas rejetée sauf pour $\gamma = 0.5559Hz$.

Par ailleurs, la figure 5.8b indique distinctement le succès du GSBB à détecter la fréquence cyclique fondamentale $\alpha = 0.548Hz$ ainsi que ses harmoniques. Alors, l'hypothèse H_0 est rejetée pour toutes les autres fréquences, excepté pour cette fréquence cyclique et ses harmoniques.

Ainsi, la validation de la méthode du GSBB sur un signal réel est confirmée.

5.6 Conclusion

Dans ce chapitre, nous avons présenté de nouvelles applications du GSBB dans l'analyse fréquentielle des signaux cyclostationnaires. La consistance du GSBB a été prouvée pour les coefficients de Fourier des fonctions de moyenne et d'autocorrélation d'un processus CS. En outre, nous avons montré des applications très intéressantes dans la détection de fréquences significatives sur des signaux simulés et réels. La méthode proposée dans ce chapitre pourrait être exploitée dans différents domaines d'application y compris le domaine de la mécanique, le domaine du sonars/radars passifs et le très récent domaine de la radio cognitive afin de détecter des sources cyclostationnaires. Cependant, il existe encore quelques problématiques qui doivent être traitées dans le futur. En effet, il serait très intéressant d'étendre cette approche aux processus cyclostationnaires généralisés que nous avons étudié dans le chapitre 1. De plus, il serait très important d'étudier profondément la question du choix de la taille du bloc b utilisé dans le GSBB.

Par ailleurs, une étude comparative entre le GSBB proposé dans ce chapitre et d'autres méthodes de ré-échantillonnage statistiques serait très intéressante. En effet, pour l'instant il existe que trois méthodes de ré-échantillonnage statistiques qui peuvent être appliquées à l'étude de la cyclostationnarité d'ordre deux qui sont : le subsampling, le MBB et le GSBB. Le plus grand inconvénient du subsampling et du MBB réside dans le fait que c'est des méthodes qui ne prennent pas en compte la périodicité des données. Par conséquent, ces deux dernières ne peuvent être utilisées dans l'estimation de la cyclostationnarité à l'ordre un. Le GSBB est la seule méthode qui permet l'analyse complète des propriétés CS1 et CS2 d'un signal cyclostationnaire.

Enfin, les cartographies de significativité fréquentielle ainsi que les matrices de probabilités proposées dans le chapitre 1 peuvent être construites par le biais du GSBB et exploitées dans des problèmes d'analyse et de comparaison entre différents cas de figure.

Conclusion générale

Dans le cadre de l'analyse de signaux mécaniques ou biomécaniques les outils d'aide à la décision reposent sur des hypothèses statistiques fortes : loi de probabilité normale, stationnarité des variables, variables centrées, variables indépendantes, . . . Or ces hypothèses sont parfois non vérifiées et engendrent des décisions erronées. Ce travail a pour objectif de proposer des méthodes qui font abstractions de certaines hypothèses et notamment de la stationnarité et de la gaussiennité des variables aléatoires. Nous avons revisité certaines méthodes de ré échantillonnages statistiques en incluant la cyclostationnarité des signaux.

Pour cela, nous avons commencé par présenter et proposer quelques exemples de modèles de signaux cyclostationnaires. Ces modèles ont été étudiés et des simulations ont été réalisées pour valider les résultats théoriques. Nous avons également proposé quelques cas pratiques issus de différents domaines d'application qui pourraient exploiter les modèles que nous avons proposés.

Ensuite, nous nous sommes penchés sur la notion de cyclostationnarité généralisée qui a été introduite en premier lieu par Napolitano. Les processus cyclostationnaires sont caractérisés par une fonction d'autocorrélation périodique dans le temps. Les coefficients de Fourier de la fonction de corrélation d'un tel processus dépendent du retard τ . Les processus cyclostationnaires généralisés ont comme particularité d'avoir des coefficients de Fourier dépendants du retard τ . De plus, les fréquences cycliques correspondant dépendent elles aussi de ce paramètre. Ainsi, les processus cyclostationnaires pourraient être perçus comme des cas particuliers des processus cyclostationnaires généralisés lorsque l'ensemble des fréquences cycliques est indépendant du retard τ .

Ainsi, nous avons proposé un nouveau modèle cyclostationnaire généralisé à fréquence cyclique variant linéairement dans le temps. Ensuite, nous avons effectué des simulations qui ont validé les résultats théoriques obtenus.

Une autre voie de recherche a aussi été présentée dans ce mémoire. En effet, nous avons cherché à développer et exploiter des méthodes de ré-échantillonnage statistiques et de bootstrap dans le cadre cyclostationnaire.

L'utilisation de telles méthodes statistiques nous a permis d'obtenir des approximations des distributions des estimateurs caractérisant les propriétés cyclostationnaires d'un signal sans avoir à imposer des hypothèses de Gaussien-

nité d'une part, et sans utiliser de résultats asymptotiques complexes d'autre part. Ainsi, cela nous a permis de proposer de nouveaux algorithmes dans le domaine fréquentiel et temporel. Nous avons effectué des études de performance de ces algorithmes et les avons appliqués sur des signaux de simulation avant de les valider sur des signaux réels issus de la biomécanique.

Nous avons vu et démontré que les méthodes de ré-échantillonnage statistiques, telles que le subsampling et le MBB simplifiaient significativement l'inférence pour les processus cyclostationnaires. Cela nous a permis la construction de tests statistiques valides avec des intervalles de confiance pour l'estimateur des coefficients de Fourier de la fonction d'autocorrélation d'un signal cyclostationnaire. Nous avons également démontré l'applicabilité de ces deux méthodes et étudié leur performance sur des signaux de simulation avec différents rapports de signal à bruit. La comparaison entre les deux approches nous a permis de mettre en évidence que le subsampling semblait être plus robuste au bruit, moins sensible au choix de la taille de bloc utilisé et plus efficace pour l'estimation et la détection des fréquences cycliques.

Par la suite, nous avons appliqué la méthode du subsampling sur des signaux réels provenant d'un coureur de haut niveau. Cette approche, nous a permis d'effectuer une analyse complète et innovante de différents états de fatigue du sportif. En effet, nous avons proposé à partir du subsampling des indicateurs innovants et originaux qui sont les cartographies de significativité fréquentielle cyclostationnaire et les matrices de probabilité des différences. Ces derniers indicateurs ont permis de donner une description exhaustive du contenu fréquentiel d'ordre deux des signaux de course et de quantifier les différences entre les divers états de fatigue du sportif.

Puis, nous nous sommes intéressés à l'estimation des composantes CS1 et CS2 et leur séparation dans le domaine temporel. Nous avons présenté une nouvelle méthode de bootstrap généralisé appelée le GSBB. Cette méthode de bootstrap a l'avantage de garder la structure périodique des données intacte, contrairement à d'autres méthodes. Ainsi, le GSBB temporel pour l'estimateur de la moyenne synchrone a été introduit et décrit. Ensuite, nous avons réalisé une étude de simulation afin de montrer son applicabilité sur des modèles cyclostationnaires que nous rencontrons souvent en pratique. Nous avons également construit un test statistique à partir du GSBB et nous avons montré l'importance de l'étude de la significativité de la moyenne synchrone (CS1). Puis, nous avons appliqué cette méthode sur des signaux de marche à pied provenant de personnes âgées. Cela nous a permis de réaliser une analyse originale de ces signaux d'une part, et d'étudier l'influence des fluctuations de vitesse sur les propriétés cyclostationnaires d'ordres 1 et 2. De plus, nous

avons proposé des indicateurs à partir du GSBB pour mesurer la performance de la séparation entre les composantes CS1 et CS2.

Pour illustrer l'intérêt de notre approche, nous avons étudié des signaux réels issus du CHU de St-Etienne. L'objectif de cette étude était de vérifier si l'analyse cyclostationnaire permettrait de caractériser la marche de personnes âgées. Les résultats obtenus sont très intéressants et ouvrent des perspectives importantes. En effet, nous avons d'abord montré le caractère CS2 de ces signaux, aussi bien dans la marche simple que dans la marche à double tâche. Nous avons également montré que la condition de double tâche augmentait le caractère CS2 des signaux de marche. De plus, la variabilité des foulées (CS2) semblait être liée au niveau de difficulté de la tâche cognitive effectuée. En outre, la CS2 semblait être plus importante au niveau du talon qu'au niveau de l'avant pied.

Nous avons ensuite, proposé d'étendre le GSBB au domaine fréquentiel. En effet, nous nous sommes focalisés sur la problématique d'estimation des coefficients de Fourier de la moyenne, d'une part, et de la fonction d'autocorrélation d'autre part. Cela revenait à caractériser la CS1 et la CS2 dans le domaine fréquentiel via le GSBB.

Nous avons présenté les résultats relatifs à la consistance du GSBB pour les coefficients de Fourier des fonctions de moyenne et d'autocorrélation d'un processus cyclostationnaire. Ensuite, nous avons montré des applications très intéressantes dans la détection de fréquences significatives (CS1 et CS2) sur des signaux simulés et réels. Jusqu'à présent, uniquement trois méthodes de ré-échantillonnage statistiques peuvent être appliquées à l'étude de la cyclostationnarité d'ordre deux et sont : le subsampling, le MBB et le GSBB. Le plus grand inconvénient du subsampling et du MBB réside dans le fait que ce sont des méthodes qui ne prennent pas en compte la périodicité des données. Par conséquent, ces deux dernières ne peuvent être utilisées dans l'estimation de la cyclostationnarité à l'ordre un. Le GSBB est la seule méthode qui permet l'analyse complète des propriétés CS1 et CS2 d'un signal cyclostationnaire.

Perspectives et travaux futurs

Cette thèse ouvre de nombreuses perspectives prometteuses, que ce soit sur le plan méthodologique et applicatif ou sur le plan expérimental.

Nous résumons les différentes perspectives méthodologiques et applicatives comme suit :

- La première repose sur la continuité des travaux de la cyclostationnarité généralisée. Nous avons présenté des résultats originaux concernant cette partie. Cependant, il reste encore des travaux à envisager pour l'avenir. En effet, il serait très intéressant de proposer puis d'étudier d'autres modèles cyclostationnaire généralisés afin d'explorer les limites de la théorie de Fourier pour l'étude de tels phénomènes.

Dans le domaine des machines tournantes, il arrive souvent que les machines fonctionnent dans des conditions non-stationnaires ou variables dans le temps. Nous pouvons citer le cas de machines tournantes avec charge variable ou avec une vitesse variable dans le temps. Dans ce cas, et contrairement au couplage entre un phénomène aléatoire stationnaire (un défaut dans un élément par exemple) et un phénomène périodique déterministe (le cycle de rotation par exemple) qui donne lieu à un phénomène cyclostationnaire, le couplage se fera entre la composante aléatoire stationnaire et une composante non périodique (vitesse de rotation ou charge variable). Par conséquent, nous obtenons des signaux avec des caractéristiques cyclostationnaires généralisées. Ainsi, il serait très intéressant d'appliquer les résultats que nous avons obtenus sur ce type de signaux.

Un autre cas très captivant à étudier est celui du couplage entre une composante aléatoire non-stationnaire et une composante non périodique (exemple : vitesse de rotation variable, charge variable avec la présence d'un défaut de roulement). Ce type de couplage donne lieu à une famille de signaux assez complexes et fortement non-stationnaire pour lesquels il serait très intéressant de formuler un cadre mathématique. *Est-ce que l'approche cyclostationnarité généralisée est suffisante ?*

- La méthode du subsampling, ainsi que les cartographies de significativité fréquentielle et les matrices de probabilité que nous avons proposées peuvent être d'une grande utilité dans d'autres domaines d'application, y compris, la mécanique et la surveillance des machines tournantes. Ces méthodes pourraient ouvrir de nouvelles perspectives pour l'analyse de tels signaux. En effet, des cartographies de significativité fréquentielle pourraient être obtenues pour différents défauts (exemple : les différents défauts de roulements) et la comparaison entre ces défauts pourrait être quantifiée par les matrices de probabilité que nous avons proposées. De plus, l'étude du degré de signification des fréquences cycliques pourrait être un outil très intéressant et utile pour fixer des seuils critiques (ou alarmes) dans la surveillance des machines tournantes.

- Concernant le GSBB dans le domaine temporel, des applications dans différents domaines pourraient être envisagées, y compris dans la mécanique. De nombreuses méthodes existent pour le diagnostic et la surveillance des machines, mais la première étape qui consiste à la séparation CS1/CS2 est cruciale. Malheureusement, dans beaucoup de cas, il n'y a aucune mesure de l'exactitude de la méthode utilisée, ni d'inspection de présence de structures périodiques dans les signaux supposés être CS2. En effet, cela pourrait causer la non visualisation de quelques défauts, ce qui entraînerait de faux diagnostics. La méthode ainsi que les indicateurs que nous avons proposés pourraient se révéler très efficaces dans ce genre de cas.

Aussi, il est tout à fait envisageable d'étendre le GSBB temporel à l'étude de l'estimateur de la variance synchrone. Cette dernière statistique caractérise le second ordre des signaux et est très utile dans le cas de présence de CS2. En effet, la variance synchrone apporte une information sur la variabilité et les fluctuations de cycle en cycle. Ainsi, elle caractérise en quelque sorte la CS2 dans le domaine temporel en quantifiant et en localisant temporellement les différentes fluctuations (aléatoires) cycliques. Par conséquent, l'intérêt de l'étude de l'estimateur de la variance synchrone par le biais de méthodes statistiques telle que le GSBB devient grandissant. De plus, le développement de nouveaux indicateurs, comme ceux qui ont été proposés dans cette thèse, pourrait être d'une très grande utilité dans l'analyse et la caractérisation des propriétés cyclostationnaires des signaux.

Par ailleurs, il serait important d'envisager une étude comparative entre les différentes méthodes de séparation CS1/CS2 existantes afin d'étudier leurs performances respectives. Nous pensons qu'il est aussi fondamental de proposer une étude théorique plus profonde des méthodes de ré-échantillonnage synchrone et angulaire afin de proposer de nouvelles méthodes plus performantes que celles qui existent déjà. Cela est d'autant plus vrai que les fluctuations de vitesses sont omniprésentes et que leur impact sur les propriétés cyclostationnaire des signaux s'avère très important, comme nous l'avons vu dans cette thèse. C'est pourquoi, il est très important de traiter cette issue au mieux, spécialement dans le cas de l'étude de systèmes un peu plus complexes qui contiennent de nombreux éléments tournants (signaux poly-cyclostationnaires).

- Concernant le GSBB dans le domaine fréquentiel, cette méthode pourrait être exploitée dans différents domaines d'application y compris le

domaine de la mécanique, et le domaine du sonars/radars passifs et le très récent domaine de la radio cognitive afin de détecter des sources cyclostationnaires à l'ordre un et deux. De plus, les cartographies de significativité fréquentielle ainsi que les matrices de probabilités proposées dans le chapitre 1 peuvent être construites par le biais du GSBB et exploitées dans des problèmes d'analyse et de comparaison entre différents cas de figure.

- D'autres perspectives de travaux concernant de façon générale les méthodes de ré-échantillonnage statistiques présentées dans cette thèse sont à prévoir dans le futur. En effet, il serait très important d'étudier profondément la question du choix de la taille du bloc utilisé dans ces méthodes ainsi que son éventuel effet sur les résultats obtenus.

Une autre perspective très intéressante est celle de la proposition de méthodes de bootstrap multidimensionnels. En effet, dans certains cas, nous avons accès à des données multidimensionnelles comme dans le domaine de la météorologie ou dans le cas d'étude de groupes sous différentes conditions. C'est pourquoi il serait très intéressant de pouvoir exploiter toutes les données disponibles par le biais de méthodes de bootstrap multidimensionnels.

En plus des perspectives méthodologiques et applicatives que nous avons cité ci-dessus, des perspectives expérimentales se dégagent aussi de ce travail. Nous les résumons comme suit :

- Les résultats que nous avons obtenus dans le cadre de l'analyse de la marche à pied chez les personnes âgées sont tous originaux et très encourageants. Par conséquent, il serait très envisageable d'étendre cette étude à d'autres conditions. En effet, il est très important de différencier entre la marche des personnes âgées et celle des personnes jeunes. Si bien qu'une campagne d'acquisition de signaux sur une population de sujets jeunes et sains est primordiale afin d'isoler l'effet de l'âge sur la marche à pied.

D'autres travaux sont aussi envisageables dans le futur afin d'analyser les effets de divers paramètres sur la marche à pied, nous pouvons citer les effets des paramètres suivants : le sexe, la prise de médicaments, la consommation de tabac ou d'alcool, le poids, la pratique du sport ou d'une activité physique, les chaussures portées, la vue et l'ouïe. Cela nous permettrait d'étudier et de caractériser les effets de tous ces paramètres sur la marche à pied.

- Une autre voie de recherche concerne la marche en multi tâches. Les résultats obtenus dans cette thèse ont montré que la tâche cognitive effectuée durant la marche affectait cette dernière de façon significative. Il serait alors très intéressant de créer un paradigme expérimental à multi tâches. Nous proposons comme exemple de tels paradigmes la marche à pied en rédigeant un SMS, ou la marche à pied en comptant ses pas tout en parlant au téléphone. Ce type de paradigmes multi tâches est plus complexe et fait appel à de nombreuses régions du cerveau. Par conséquent, les effets de telles tâches sur la marche à pied augmenteraient notre compréhension du phénomène de la marche.
- Par ailleurs, il serait très intéressant de disposer dans l'avenir d'imageries cérébrales des sujets participants aux différents paradigmes expérimentaux. En effet, le lien entre la marche à pied en double (ou multi tâches) et le cerveau est direct. C'est pourquoi, il est envisageable de traiter les imageries par résonance magnétique fonctionnelle (IRMf) afin de détecter les réseaux fonctionnels cérébraux et d'étudier les différentes interactions entre ces derniers (l'étude des connectivités fonctionnelles) lors de la réalisation d'une ou plusieurs tâches. Cela permettrait d'isoler l'effet des tâches réalisées et améliorerait notre compréhension de la marche à double ou à multi tâches.

D'autre part, il est connu que l'âge entraîne l'atrophie de quelques régions cérébrales, ainsi, le traitement d'un autre type d'imagerie est envisageable. En effet, l'imagerie par résonance magnétique anatomique pourrait nous permettre d'explorer plus précisément ces atrophies. Par conséquent, cela nous permettrait d'analyser les aptitudes des personnes âgées à réaliser une marche à double ou multi tâches d'une part, et de justifier les changements observés chez les personnes âgées lors de la marche à pied en double ou multi tâches d'autre part.

- Enfin, un objectif plus lointain est d'arriver à détecter et diagnostiquer des maladies neuro-dégénératives telles que la maladie d'Alzheimer, la maladie de Parkinson ou bien la maladie de Huntington à partir de la marche à pied. Pour ce faire, il serait primordial de mettre en œuvre des marqueurs fiables et spécifiques afin de discriminer entre chaque type de maladie.

La relation entre l'analyse d'enveloppe et la corrélation spectrale intégrée

Dans cette annexe, nous rappelons les résultats obtenus dans [126] relatifs à la relation entre l'analyse d'enveloppe et la corrélation spectrale intégrée pour les processus cyclostationnaires. Il est important de souligner que les cartographies des corrélations spectrales contiennent quelques inconvénients dans la pratique. En effet, même si elles sont très utiles et apportent l'information sur les fréquences cycliques, leur estimation est assez coûteuse en termes de temps de calcul. De plus, les résultats obtenus par ces dernières dépendent souvent des paramètres utilisés lors de l'estimation tels que le type de la fenêtre d'estimation et son type ou bien le taux de recouvrement.

Ainsi, les auteurs dans [126] ont proposé une solution pour résorber ces problématiques. L'idée de base est d'intégrer la corrélation spectrale suivant l'axe des fréquences f afin d'obtenir un spectre discret sur l'axe des fréquences cycliques α . Pour ce faire, les auteurs de [126] ont montré une relation théorique entre la corrélation spectrale intégrée et l'analyse d'enveloppe. Le raisonnement suivi est expliqué ci-dessous.

Pour un signal $X(t)$ cyclostationnaire à l'ordre deux, soit :

$$M_X(\alpha) = \int_{\mathbb{R}} S_X(\alpha, f) df \quad (\text{A.1})$$

l'intégrale de la corrélation spectrale $S_X(\alpha, f)$ suivant l'axe des fréquences f .

Ensuite, en remplaçant $S_X(\alpha, f)$ par sa définition, on obtient ce qui suit :

$$M_X(\alpha) = \int_{\mathbb{R}} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{\mathbb{R}} \int_{-T/2}^{T/2} R_X(t, \tau) e^{-j2\pi(f\tau + \alpha t)} dt d\tau df \quad (\text{A.2})$$

Avec $R_X(t, \tau)$ la fonction d'autocorrélation du signal $X(t)$.

Puis, en changeant l'ordre de l'intégration, l'expression devient :

$$\begin{aligned}
 &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{\mathbb{R}} \int_{-T/2}^{T/2} R_X(t, \tau) e^{-j2\pi\alpha t} dt \int_{\mathbb{R}} e^{-j2\pi f\tau} df d\tau, \\
 &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{\mathbb{R}} \int_{-T/2}^{T/2} R_X(t, \tau) e^{-j2\pi\alpha t} dt \delta(\tau) d\tau, \\
 &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t, 0) e^{-j2\pi\alpha t} dt,
 \end{aligned}$$

Le signal $X(t)$ étant supposé cyclostationnaire à l'ordre deux, donc de nature stochastique, alors :

$$R_X(t, 0) = E\{|X(t)|^2\} \quad (\text{A.3})$$

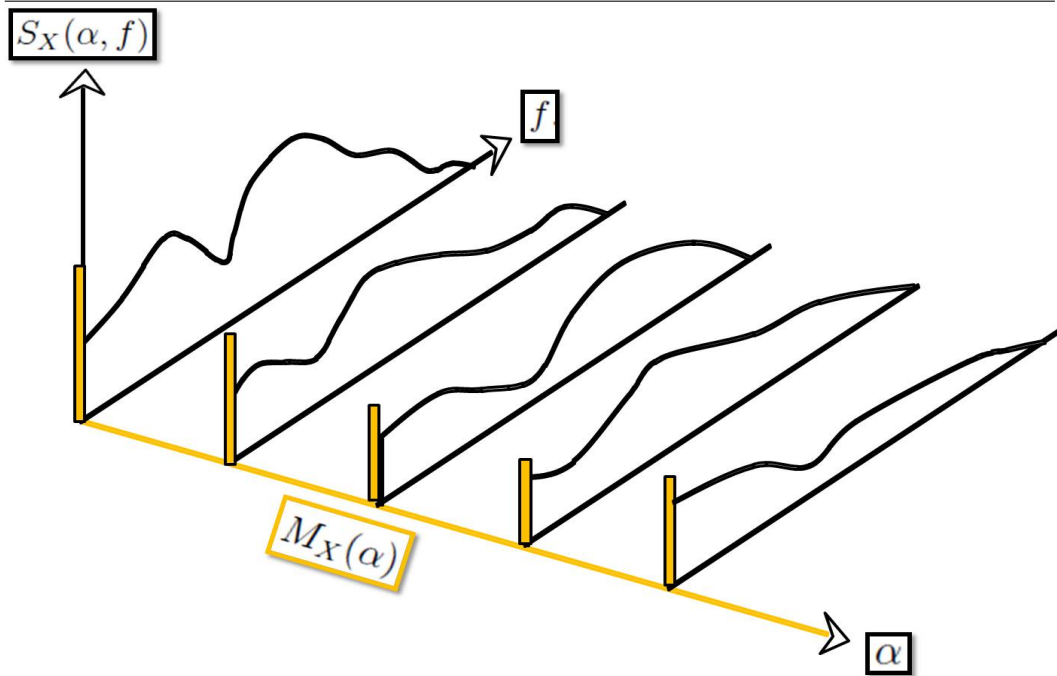


FIGURE A.1 – Relation entre l'analyse d'enveloppe et la corrélation spectrale intégrée d'un signal CS2.

Ce dernier résultat signifie que la corrélation spectrale intégrée $M_X(\alpha)$ d'un signal cyclostationnaire à l'ordre deux $X(t)$ correspond tout simplement à la transformée de Fourier du carré du signal $X(t)$. Ainsi, $M_X(\alpha)$ est donc une fonction quasi-périodique avec un spectre discret qui est nul partout sauf aux fréquences cycliques α .

La figure [A.1](#) schématise la relation établie entre l'analyse d'enveloppe et la corrélation spectrale intégrée. On peut alors dire que le spectre de l'enveloppe d'un signal cyclostationnaire peut être aperçue comme étant la projection de la corrélation spectrale sur l'axe des fréquences cycliques α .

ANNEXE B

Consistance du GSBB dans le domaine fréquentiel

Dans cette partie, nous présentons plus de détails sur les preuves des théorèmes du chapitre 5.

Preuves des théorèmes 5.3.1 et 5.3.2 :

Les preuves des théorèmes 5.3.1 et 5.3.2 sont des versions simplifiées des théorèmes 5.3.3 et 5.3.4 qui sont présentées ci-dessous. La différence principale est que dans ce cas, il n'y a pas besoin de définir l'estimateur asymptotique équivalent (regarder la preuve du théorème 5.3.3) et le raisonnement est basé sur les résultats de [14].

Preuves des théorèmes 5.3.3 et 5.3.4 :

Tout d'abord, nous rappelons que le nombre d'échantillons n est supposé être un entier multiple de la taille du bloc b ($n = lb, l \in \mathbb{N}$) et n est un entier multiple de la longueur de la période T ($n = KT$). Aussi, nous changeons la notation des indices pour des raisons pratiques (espace) de telle sorte à avoir des petits indices (exemple : $X(t)$ devient X_t).

Nous commençons avec la preuve du théorème 5.3.1. Nous montrons la consistance pour la partie réelle de $\widehat{R}_n^*(\alpha, \tau)$ en premier :

$$\sup_{x \in \mathcal{R}} \left| P \left(\sqrt{n} \left(\Re \left(\widehat{R}_n(\alpha, \tau) \right) - \Re(R(\alpha, \tau)) \right) \leq x \right) - P^* \left(\sqrt{n} \left(\Re \left(\widehat{R}_n^*(\alpha, \tau) \right) - E^* \left(\Re \left(\widehat{R}_n^*(\alpha, \tau) \right) \right) \right) \leq x \right) \right| \xrightarrow{p} 0.$$

Comme dans [50], au lieu de l'estimateur original, nous considérons l'estimateur modifié suivant de $\Re\left(\tilde{R}_n^*(\alpha, \tau)\right)$:

$$\Re\left(\tilde{R}_n^*(\alpha, \tau)\right) = \frac{1}{n} \sum_{m=0}^{l-1} \tilde{Z}_{1+bm, b}^*$$

Avec

$$\tilde{Z}_{t, b}^* = X_t^* X_{t+\tau}^* \cos(2\pi\alpha t^*) + \cdots + X_{t+b-\tau-1}^* X_{t+b-1}^* \cos(2\pi\alpha(t+b-\tau-1)^*).$$

Sous ces notations et en utilisant les résultats de [51] et [50] qui ont montré que ces deux estimateurs sont asymptotiquement équivalents quand on utilise le bloc bootstrap circulaire. Nous rappelons ce résultat ci-dessous :

$$\begin{aligned} & \sqrt{n} \left| \Re\left(\hat{R}_n^*(\alpha, \tau)\right) - \Re\left(\tilde{R}_n^*(\alpha, \tau)\right) \right. \\ & \left. - \mathbb{E}^*\left(\Re\left(\hat{R}_n^*(\alpha, \tau)\right)\right) - \mathbb{E}^*\left(\Re\left(\tilde{R}_n^*(\alpha, \tau)\right)\right) \right| \xrightarrow{p^*} 0. \end{aligned}$$

(B.2)

L'utilisation de l'inégalité de Tchebychev est ensuite suffisante pour montrer la convergence de la variance

$$n \text{Var}^*\left(\Re\left(\hat{R}_n^*(\alpha, \tau)\right) - \Re\left(\tilde{R}_n^*(\alpha, \tau)\right)\right) \xrightarrow{p^*} 0. \quad (\text{B.3})$$

Et cette variance a la forme suivante :

$$\begin{aligned} & \frac{2}{n} \sum_{m=0}^{l-2} \text{Var}^*\left(\sum_{i=1}^{\tau} X_{(m+1)b-\tau+i}^* X_{(m+1)b+i}^* \cos(2\pi\alpha((m+1)b-\tau+i)^*)\right) \\ & + \frac{2}{n} \sum_{m=0}^{l-2} \sum_{m'>m} \text{Cov}^*\left(\sum_{i=1}^{\tau} X_{(m+1)b-\tau+i}^* X_{(m+1)b+i}^* \cos(2\pi\alpha((m+1)b-\tau+i)^*)\right. \\ & \left. \times \sum_{j=1}^{\tau} X_{(m'+1)b-\tau+j}^* X_{(m'+1)b+j}^* \cos(2\pi\alpha((m'+1)b-\tau+j)^*)\right), \quad (\text{B.4}) \end{aligned}$$

Puisque les deux estimateurs diffèrent seulement quand X_t^* et $X_{t+\tau}^*$ appartiennent à deux blocs différents, alors dans la dernière somme, la covariance

est nulle quand m et m' diffèrent de plus de 1. Cela signifie que nous devons juste considérer $m' = m + 1$. Ainsi, le terme de la variance dans (B.4) devient égal à :

$$\begin{aligned} & \frac{2}{n} \sum_{m=0}^{l-2} \frac{1}{K} \sum_{j_1=1}^K \frac{1}{K} \sum_{j_2=1}^K \left(\sum_{i=1}^{\tau} X_{(j_1-1)T+g_m^1+i-1} X_{(j_2-1)T+g_m^2+i-1} \right. \\ & \qquad \qquad \qquad \left. \cos(2\pi\alpha((j_1-1)T + g_m^1 + i - 1)) \right)^2 \\ & - \frac{2}{n} \sum_{m=0}^{l-2} \left(\frac{1}{K} \sum_{j_3=1}^K \frac{1}{K} \sum_{j_4=1}^K \left(\sum_{i=1}^{\tau} X_{(j_3-1)T+g_m^1+i-1} X_{(j_4-1)T+g_m^2+i-1} \right. \right. \\ & \qquad \qquad \qquad \left. \left. \cos(2\pi\alpha((j_3-1)T + g_m^1 + i - 1)) \right) \right)^2, \end{aligned}$$

Tels que g_m^1 et g_m^2 pour $m = 0, \dots, l-2$ représentent :

$$\begin{aligned} g_m^1 &= ((m+1)b - \tau + 1) \bmod T, \\ g_m^2 &= ((m+1)b + 1) \bmod T. \end{aligned}$$

De plus, si $g_m^i = 0$ pour $i = 1, 2$ nous mettons $g_m^i = T$. La valeur absolue de la dernière expression se réduit à $O(1/b)$.

En prenant $m' = m + 1$ dans la seconde somme de (B.4), on obtient :

$$\begin{aligned} & \frac{2}{n} \sum_{m=0}^{l-3} \mathbb{E}^* \left(\sum_{i_1=1}^{\tau} X_{(m+1)b-\tau+i_1}^* X_{(m+1)b+i_1}^* \cos(2\pi\alpha((m+1)b - \tau + i_1)^*) \cdot \right. \\ & \qquad \qquad \qquad \left. \cdot \sum_{i_2=1}^{\tau} X_{(m+2)b-\tau+i_2}^* X_{(m+2)b+i_2}^* \cos(2\pi\alpha((m+2)b - \tau + i_2)^*) \right)^2 \\ & = \frac{2}{n} \sum_{m=0}^{l-3} \frac{1}{K} \sum_{j_1=1}^K \frac{1}{K} \sum_{j_2=1}^K \frac{1}{K} \sum_{j_3=1}^K \left(\sum_{i_1=1}^{\tau} X_{(j_1-1)T+g_m^1+i_1-1} X_{(j_2-1)T+g_m^2+i_1-1} \right. \\ & \qquad \qquad \qquad \left. \cos(2\pi\alpha((j_1-1)T + g_m^1 + i_1 - 1)) \cdot \sum_{i_2=1}^{\tau} X_{(j_2-1)T+g_m^2+b-\tau+i_2-1} X_{(j_3-1)T+g_m^3+i_1-1} \right. \\ & \qquad \qquad \qquad \left. \cos(2\pi\alpha(((j_2-1)T + g_m^2 + b - \tau + i_2 - 1))) \right)^2, \end{aligned}$$

Tels que pour $m = 0, \dots, l-3$ on note

$$g_m^3 = ((m+2)b+1) \bmod T.$$

La valeur absolue de l'expression à droite est $O(1/b)$. Ainsi, la preuve de (B.3) est terminée.

Alors, pour obtenir la consistance qu'on cherche, il faudrait montrer que :

$$\begin{aligned} & \sup_{x \in \mathcal{R}} \left| P \left(\sqrt{n} \left(\Re \left(\widehat{R}_n(\alpha, \tau) \right) - \Re(R(\alpha, \tau)) \right) \leq x \right) - \right. \\ & \left. - P^* \left(\sqrt{n} \left(\Re \left(\widetilde{R}_n^*(\alpha, \tau) \right) - E^* \left(\Re \left(\widetilde{R}_n^*(\alpha, \tau) \right) \right) \right) \leq x \right) \right| \xrightarrow{p} 0. \end{aligned} \quad (\text{B.5})$$

Suivant les notations de [50] on définit :

$$Z_{t,b} = \widetilde{Z}_{t,b} - b\Re(R(\alpha, \tau))$$

et sa version issue du bootstrap

$$Z_{t,b}^* = \widetilde{Z}_{t,b}^* - b\Re(R(\alpha, \tau)).$$

Pour $t = 1, b+1, \dots, lb$ et les variables aléatoires $Z_{t,b}^*$ sont conditionnellement indépendantes avec une distribution commune :

$$P^* (Z_{t,b}^* = Z_{t+mT,b}) = P^* (h_t = t + mT) = \frac{1}{K} \quad \text{pour } m = 0, \dots, K-1.$$

Tel que h_t est défini dans l'étape (2) de l'algorithme du GSBB.

Afin d'obtenir (B.5) les résultats de [14] sont utilisés. Ainsi, pour $\nu > 0$ les conditions suivantes sont remplies :

$$\begin{aligned} & \sum_{m=0}^{l-1} P^* \left(\frac{1}{\sqrt{n}} |Z_{1+mb,b}^*| > \nu \right) \xrightarrow{p} 0, \\ & \sum_{m=0}^{l-1} E^* \left(\frac{1}{\sqrt{n}} Z_{1+mb,b}^* \mathbf{1}_{|Z_{1+mb,b}^*| > \sqrt{n}\nu} \right) \xrightarrow{p} 0, \\ & \sum_{m=0}^{l-1} \text{Var}^* \left(\frac{1}{\sqrt{n}} Z_{1+mb,b}^* \mathbf{1}_{|Z_{1+mb,b}^*| \leq \sqrt{n}\nu} \right) \xrightarrow{p} \sigma_1^2. \end{aligned}$$

À partir de là, le raisonnement est exactement le même que dans [51]. Aussi, la preuve pour la partie imaginaire est similaire à celle de la partie réelle.

edefYour command was ignored. Type I <command> <return> to replace it with another command, or <return> to continue without it.

Bibliographie

- [1] Alekseev, V. G. Estimating the spectral densities of a Gaussian periodically correlated stochastic process. *Prob. Inf. Dansm.* 24, pages : 109-115, 1988.
- [2] Allen, J., Hobbs, S. Detecting target motion by frequency-plane smoothing. In *Proceedings of the Twenty-Sixth Asilomar Conference on Systems and Computers*. Pacific Grove, CA, pages : 1042-1047, 1992.
- [3] Anderson, P. L., Meerschaert, M. M., Vecchia, A. Innovations algorithm for periodically stationary time series. *Stoch. Proc. Appl.*, 83, pages : 149-169, 1999.
- [4] Anderson, R. H., Krolik, J. L. Over-the-Horizon Radar Target Localization Using a Hidden Markov Model Estimated from Ionosonde Data. *Radio-Science*, 33(4), pages : 1199-1213, 1998.
- [5] Anderson, R. H., Krolik, J. L. The Performance of Maximum Likelihood Over-The-Horizon Radar Coordinate Registration. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP98*, 4, pages : 2481-2484, Seattle, USA, 1998.
- [6] Anderson, R. H., Krolik, J. L. Multipath Track Association for Over-The-Horizon Radar Using a Bootstrapped Statistical Ionospheric Model. In *Proceedings of 33rd Asilomar Conference on Signals, Systems and Computers*, 1, pages : 8-14, Pacific Grove, California, 1999.
- [7] Angus, J. E. A note on the central limit theorem for the bootstrap mean. *Commun. Statist. Theory Methods* 18, pages : 1979-1982, 1989.
- [8] Antoni, J. Cyclostationarity by examples. *MSSP*, 23 (4), pages : 987-1036, 2009.
- [9] Antoni, J., Randall, R. B. Differential diagnosis of gears and bearing faults, *ASME Journal of Vibration and Acoustics* 124 (2), pages : 165-171, 2002.
- [10] Antoni, J., Randall, R. B. On the use of the cyclic power spectrum in rolling element bearing diagnostics, *Journal of Sound and Vibration* 281 (1-2), pages : 463-468, 2005.

- [11] Antoni, J., Danière, J., Guillet, F. Effective vibration analysis of IC engines using cyclostationarity. Part I : a methodology for condition monitoring, *Journal of Sound and Vibration* 257 (5), pages : 815-837, 2002.
- [12] Antoni, J., Guillet, F., ElBadaoui, M., Bonnardot, F. Blind separation of convolved cyclostationary processes, *Signal Processing* 85 (1), pages : 51-66, 2005.
- [13] Antoni, J., Bonnardot, F., Raad, A., El Badaoui, M. Cyclostationary modelling of rotating machine vibration signals *Mechanical Systems and Signal Processing*, 18(6), pages : 1285-1314, 2004.
- [14] Araujo, A., Giné, E. The central limit theorem for real and Banach valued random variables. New York : Wiley. 1980.
- [15] Athreya, K. B. Bootstrap estimation of the mean in the infinite variance case. *Ann. Statist.* 15, pages : 724-731, 1987.
- [16] Banga, C., Ghorbel, F. Optimal Bootstrap Sampling for Fast Image Segmentation : Application to Retina Image. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP93*, 5, pages : 638-641, Minnesota, 1993.
- [17] Beauchet, O., Herrmann, F. R., Grandjean, R., Dubost, V., Allali, G. Concurrent validity of SMTEC footswitches system for the measurement of temporal gait parameters, *Gait & Posture*, 27 (1), pages : 156-159, 2008.
- [18] Belli, A., Bui, P., Berger, A., Geysant, A., Lacour, J. R. A treadmill ergometer for three-dimensional ground reaction forces measurement during walking, *Journal of Biomechanics*, 34 (1), pages : 105-112, 2001.
- [19] Bennett, W. R. Statistics of regenerative digital transmission. *Bell Syst Tech. J.*, 37, pages : 1501-1542, 1958.
- [20] Bloomfield, P., Hurd, H. L., Lund, R. B. Periodic correlation in stratospheric ozone data, *Time Series Analysis*, 15, pages : 127-150, 1994.
- [21] Boashash, B. Estimating and interpreting the instantaneous frequency of a signal -part 1 : Fundamentals, *Proceedings of the IEEE*, 80(4), 1992.
- [22] Böhme, J. F., Maiwald, D. Multiple Wideband Signal Detection and Tracking from Towed Array Data, *SYSID 94, 10th IFAC Symposium on System Identification*, 1, pages : 107-112, Copenhagen, 1994.

- [23] Bonnardot, F., El Badaoui, M. Etude de la fatigue d'un coureur : de l'instrumentation à l'analyse vibratoire, 10ème congrès français d'acoustique, Lyon, 2010.
- [24] Bonnardot, F., Randall, R. B., Guillet, F. Extraction of second order cyclostationary sources, application to vibration analysis, *Mechanical Systems and Signal Processing* 19 (6), pages : 1230-1244, 2005.
- [25] Bonnardot, F., El Badaoui, M., Randall, R. B., Danière, J., Guillet, F. Use of the acceleration signal of a gearbox in order to perform angular resampling (with limited speed fluctuation), *Mechanical Systems and Signal Processing* 19 (4), pages : 766-785, 2005.
- [26] Bonnardot, F., Antoni, J., Randall, R. B., El Badaoui, M. Enhancement of second-order cyclostationary signals : application to vibration analysis, *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP04, Montreal, Quebec, Canada, 2*, pages : 781-784, 2004.
- [27] Bonnardot, F. Comparaison entre les analyses angulaires et temporelles des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarité floue, Thèse de doctorat, Institut National de Polytechnique de Grenoble, 2004.
- [28] Box, G., Jenkins, G. M., Reinsel, G. C. *Time Series Analysis : Forecasting and Control*, 4th Edition. Hoboken, N.J., New York, 2008.
- [29] Boyles R. A., Gardner, W. A. Cycloergodic properties of discrete-parameter nonstationary stochastic processes. *IEEE Trans. In Theory, IT-29*, pages : 105-114, 1983.
- [30] Braun, S., Rotberg, J., Lenz, E. Signal processing for single tooth milling monitoring, *Mechanical Systems and Signal Processing* 1 (2), pages : 185-196, 19987.
- [31] Brelsford, W. S. Probability predictions and time series with periodic structure, Thèse de doctorat, Johns Hopkins University, Baltimore, 1967.
- [32] Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T. A., Brammer, M. Colored Noise and Computational Inference in Neurophysiological (fMRI) Time Series Analysis : Resampling Methods in Time and Wavelet Domains. *Human Brain Mapping*, 12, pages : 61-78, 2001.

- [33] Cambanis, S., Houdré, C. H., Hurd, H. L., Leskow, J. Laws of large numbers for periodically and almost periodically correlated processes. *Stoch. Proc. App.*, 53, pages : 37-54, 1994.
- [34] Chan, V., Lahiri, S. N., Meeker, W. Q. Block bootstrap estimation of the distribution of cumulative outdoor degradation, *Technometrics*, 46, pages : 215-224, 2004.
- [35] Charbit, M. *Elements de théorie du signal : les signaux aléatoires*. Ellipses, 1990.
- [36] Chen, D. R., Kuo, W. J., Chang, R. F., Moon, W. K., Lee, C. C. Use of the Bootstrap Technique with Small Training Sets for Computer-Aided Diagnosis in Breast Ultrasound. *Ultrasound in Medicine and Biology*, 28(7), pages : 897-902, 2002.
- [37] Chernick. M. R. *Bootstrap Methods : A Guide for Practitioners and Researchers*, 2nd Edition. A John Wiley Sons, Newtown PA, 2007.
- [38] Coakley, K. J. Bootstrap Method for Nonlinear Filtering of EM-ML Reconstructions of PET Images. *International Journal of Imaging Systems and Technology*, 7(1), pages : 54-61, 1996.
- [39] Dandawaté, A. V., Giannakis, G. B. Statistical tests for presence of cyclostationarity. *IEEE Transaction on Signal Processing*. 42(9), pages : 2355-2369, 1994.
- [40] Dehay, D. Spectral analysis of the covariance of the almost periodically correlated processes. *Stoch. Proc. Appl.* 50, pages : 315-330, 1994.
- [41] Dehay, D. Hurd, H. L. Representation and estimation for periodically and almost periodically correlated random processes. In *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner. Ed. IEEE Press. New York, 1993.
- [42] Dehay, D., Leśkow, J. Functional limit theory for the spectral covariance estimator. *Journal of Applied Probability*. 33, pages : 1077-1092, 1996.
- [43] Dehay, D., and Leśkow, J. Testing stationarity for stock market data. *Economics Letters*. 50(2), pages : 205-212, 1996.
- [44] Donohue, K. D., Varghese, T. Spectral cross correlation for tissue characterization, *IEEE Ultrasonic Symposium*, Tucson, AZ, 2, pages : 1049-1052, 1992.

- [45] Doukhan, P. Mixing. Properties and examples. Lecture Notes Statist. Springer, 1994.
- [46] Dragan, Y. P. Structure and Representation of Stochastic Signal Models (version originale en russe), Naukova Dumka, Kiev, 1980.
- [47] Dragan, Y. P., Yavorskiy, I. N. Rhythms of Sea Waves and Underwater Acoustic Signals (version originale en russe), Naukova Dumka, Kiev, 1982.
- [48] Dragan, Y. P., Rozhkov, V. A., Yavorskiy, I. N. Methods of Probabilistic Analysis of Rhythms of Oceanological Processes (version originale en russe), Gidrometeoizdat, Leningrad, 1987.
- [49] Dubost, V., Kressig, R. W., Gonthier, R., Herrmann, F. R., Aminian, K., Najafi, B., Beauchet, O. Relationships between dual-task related changes in stride velocity and stride time variability in healthy older adults. *Human Movement Science* 25, pages : 372-382, 2006.
- [50] Dudek, A. E., Circular block bootstrap for coefficients of autocovariance function of APC time series - sous presse.
- [51] Dudek, A. E., Leśkow, J., Politis, D., Paparoditis, E. A generalized block bootstrap for seasonal time series, *J. Time Ser. Anal*, 35, pages : 89-114, 2014.
- [52] Edgington, E. S. Randomization Tests, Marcel Dekker, New York, 1980.
- [53] Edgington, E. S. Randomization Tests, 2nd ed. Marcel Dekker, New York, 1987.
- [54] Edgington, E. S. Randomization Tests, 3rd ed. Marcel Dekker, New York, 1995.
- [55] Efron, B. Bootstrap Methods. Another Look at the Jackknife. *The Annals of Statistics*, 7, pages : 1-26, 1979.
- [56] Efron, B. Computers and the Theory of Statistics : Thinking the Unthinkable. *SIAM Review*, 4, pages : 460-480, 1979.
- [57] Efron, B. Nonparametric Standard Errors and Confidence Intervals. *The Canadian Journal of Statistics*, 9, pages : 1-26, 1981.
- [58] Efron, B. The Jackknife, the Bootstrap and Other Resampling Plans. CBMS Monograph 38, Society for Industrial and Applied Mathematics, Philadelphia, 1982.

- [59] Efron, B., Gong, G. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, 37, pages : 36-48, 1983.
- [60] Efron, B., Tibshirani, R. *An Introduction to the Bootstrap*. Chapman and Hall, New York, 1993.
- [61] El Badaoui, M., Danière, J., Guillet, F., Servièrre, C. Separation of combustion noise and piston-slap in diesel engine, part I : separation of combustion noise and piston-slap in diesel engine by cyclic wiener filtering, *Mechanical Systems and Signal Processing* 19(6), pages : 1209-1217, 2005.
- [62] Fisher, N. I., Hall, P. Bootstrap Confidence Regions for Directional Data. *Journal of the American Statistical Society*, 84, pages : 996-1002, 1989.
- [63] Fisher, N. I., Hall, P. New Statistical Methods for Directional Data. Bootstrap Comparison of Mean Directions and the Fold Test in Palaeomagnetism. *Geophysics Journal International*, 101, pages : 305-313, 1990.
- [64] Fisher, N. I., Hall, P. General Statistical Test for the Effect of Folding. *Geophysics Journal International*, 105, pages : 419-427, 1991.
- [65] Franks, L. E. *Signal Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1969.
- [66] Fusco, T., Izzo, L., Napolitano, A., Tanda, M. On the Second-Order Cyclostationarity Properties of Long-Code DS-SS Signals. *IEEE Transactions on Communications*, 54(10), pages : 1741-1746, 2006.
- [67] Gardner, W. A., Franks, L. E. Characterization of cyclostationary random signal processes. *IEEE Trans. Inf. Theory*, IT-21, pages : 4-14, 1975.
- [68] Gardner, W. A. *Introduction to Random Processes with Application to Signals and Systems*, Macmillan, New York, 1985.
- [69] Gardner, W. A. *Statistical Spectral Analysis : A Nonprobabilistic Theory*, Prentice Hall, Englewood Cliffs, NJ, 1987.
- [70] Gardner, W. A. Two alternative philosophies for estimation of the parameters of time-series. *IEEE Trans. Inf. Theory*, 37, pages : 216-218, 1991.
- [71] Gardner, W. A., Spooner, C. M. Detection and source location of weak cyclostationary signals : simplification of the maximum likelihood receiver. *IEEE Trans. on Commun*, 41, pages : 905-916, 1993.

- [72] Gardner, W. A. Cyclostationarity in communications and signal processing. IEEE Press, New York, 504 pages, 1994..
- [73] Gardner, W. A. Introduction to Random Processes with Applications to Signals and Systems, second edition. McGraw-Hill Publishing Company, New York, USA, 1990.
- [74] Gardner, W. A. Stationarizable random processes. IEEE Transactions on Information Theory, 24, pages : 8–22, 1978.
- [75] Gardner, W. A. An introduction to Cyclostationary signals. Chapter 1 in Cyclostationarity in Communication and Signal Processing, IEEE Press, pages : 1-90, 1994.
- [76] Gardner, W. A., Napolitano, A. Cyclostationarity : Half a century of research. Signal Processing 86, pages : 639-697, 2006.
- [77] Gardner, W. A., Spooner, C. M. The cumulant theory of cyclostationary time-series, Part I : foundations, IEEE Trans. on Signal Processing, 42(12), pages : 3387-3408, 1994.
- [78] Gladyshev, E. G. Periodically correlated random sequences, Sow. Math, 2, pages : 385-388, 1961.
- [79] Gladyshev, E. G. Periodically and almost periodically correlated random processes with continuous time parameter, Theory Probab. Appl, 8, pages : 173-177, 1963.
- [80] Gudzenko, L. I. On periodically nonstationary processes. Radiotekhnika i elektronika, 4(6), pages : 1062-1064, 1959.
- [81] Good, P. Resampling Methods : A Practical Guide to Data Analysis. Birkhauser, Boston, 1998.
- [82] Hall, P. Defining and measuring long-range dependence. In Nonlinear Dynamics and Time Series. (C. D. Cutler and D. T. Kaplan, editors), 11, Fields Communications, pages : 153-160, 1997.
- [83] Hall, P. The Bootstrap and Edgeworth Expansion. Springer-Verlag New York, Inc, New York, 1992.
- [84] Haynor, D. R., Woods, S. D. Resampling Estimates of Precision in Emission Tomography. IEEE Transactions on Medical Imaging, 8, pages : 337-343, 1989.

- [85] Herbst, L. J. Almost periodic variances. *Ann. Math. Stat.*, 34, pages : 1549-1557, 1963.
- [86] Herbst, L. J. Periodogram analysis and variance fluctuations, *J. R. Stat. Soc.* 25, pages : 442-450, 1963.
- [87] Herbst, L. J. A test for variance heterogeneity in the residuals of a Gaussian moving average. *J. R. Stat. Soc.* 25, pages : 451-454, 1963.
- [88] Herbst, L. J. Spectral analysis in the presence of variance fluctuations. *J. R. Stat. Soc.* 26, pages : 354-360, 1964.
- [89] Herbst, L. J. Stationary amplitude fluctuations in a time series. *J. R. Stat. Soc.* 26, pages : 361-364, 1964.
- [90] Herbst, L. J. The statistical fourier analysis of variances. *J. R. Stat. Soc.* 27, pages : 159-165, 1965.
- [91] Hurd, H. L. An investigation of periodically correlated stochastic processes. Thèse de doctorat, Duke University department of Electrical Engineering, 1969.
- [92] Hurd, H. L., Gerr, N. L. Graphical methods for determining the presence of periodic correlation in time series. *J. Time Series Anal.* 12, pages : 337-350, 1991.
- [93] Hurd, H., Miamee, A. *Periodically Correlated Random Sequences : Spectral Theory and Practice.* John Wiley & Sons, Inc. Hoboken, New Jersey, 384 pages, 2007.
- [94] Hurd, H. L., Leśkow. J. Estimation of the Fourier coefficient functions and their spectral densities for φ -mixing almost periodically correlated processes. *Statistics and Probability Letters*, 14, pages : 299-306, 1992.
- [95] Jabloun, M. Modélisation de signaux fortement non stationnaires à phase et à amplitude locales polynomiales. Thèse de doctorat, Institut national polytechnique de Grenoble, 2007.
- [96] Jones, R. H., Brelsford, W. M. Time series with periodic structure. *Biometrika*, 54, pages : 403-408, 1967.
- [97] Kampé de Fériet, J. Correlation and spectrum of asymptotically stationary random functions. *Math. Stud.* 30, pages : 55-67, 1962.

- [98] Kampé de Fériet, J., Frenkiel, F. N. Correlation and spectra for nonstationary random functions. *Math. Comp*, 16, pages : 1-21, 1962.
- [99] Koenig, D., Boehme, J. Application of cyclostationarity and time-frequency analysis to engine car diagnostics, *Proc. Intl. Conf. on ASSP*, pages : 149-152, Adelaide, Australia, 1994.
- [100] Krampe, R. T., Schaefer, S., Lindenberger, U., Baltes, P. B. Lifespan changes in multi-tasking : concurrent walking and memory search in children, young, and older adults, *Gait and Posture*, 33(3), pages : 401-405, 2011.
- [101] Künsch, H. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*. 17(3), pages : 1217-1241, 1989.
- [102] Lahiri, S. N. *Resampling Methods for Dependent Data*, Springer Science Business Media, 374 pages, 2003.
- [103] Lenart, Ł., Leśkow, J., Synowiecki, R. Subsampling in estimation of autocovariance for PC time series, *Journal of Time Series Anal*, 29(6), pages : 995-1018, 2008.
- [104] Leśkow, J., Synowiecki, R. On bootstrapping periodic random arrays with increasing period, 71(3), pages : 253-279, 2010.
- [105] Liu, R. Y., Singh, K. Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap* (R. LePage and L. Billard, eds.), New York : Wiley, pages : 225-248, 1992.
- [106] Locascio, J. J., Jennings, P. J., Moore, C. I., Corkin, S. Time Series Analysis in the Time Domain and Resampling Methods for Studies of Functional Magnetic Resonance Imaging. *Human Brain Mapping*, 5, pages : 168-193, 1997.
- [107] Loevdén, M., Schaefer, S., Pohlmeier, A. E., Lindenberger, U. Walking variability and working-memory load in aging : a dual-process account relating cognitive control to motor control performance. *The Journals of Gerontology*, 63(3), pages : 121-128, 2008.
- [108] Lund, R., Basawa, I. V. Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Series Anal*, 21, pages : 75-93, 2000.
- [109] Manly, B. F. J. Bootstrapping for determining sample sizes in biological studies. *J. Exp. Mar. Biol. Ecol.* 158, pages : 189-196, 1992.

- [110] Manly, B. F. J. *Randomization and Monte Carlo Methods in Biology*. Chapman Hall, London, 1991.
- [111] Markelov, V. A. Axis crossings and relative time of existence of periodically nonstationary random processes, *Sov. Radiophys.* 9, pages : 440-443, 1966.
- [112] Monin, A. S. Stationary and periodic time series in the general circulation of the atmosphere. in *Proceedings of Symposium on Time Series Analysis*, M. Rosenblatt. Ed., Wiley, Hoboken. NJ, 1963.
- [113] Nagaoka, S., Amai, O. Estimation Accuracy of Close Approach Probability for Establishing a Radar Separation Minimum. *Journal of Navigation*, 44, pages : 110-121, 1991.
- [114] Nagaoka, S., Amai, O. A Method for Establishing a Separation in Air Traffic Control Using a Radar Estimation Accuracy of Close Approach Probability. *Journal of the Japan Institute of Navigation*, 82, pages : 53-60, 1990.
- [115] Napolitano, A. Estimation of second-order cross-moments of generalized almostcyclostationary processes. *IEEE Transactions on Information Theory*, 53(6), pages : 2204–2228, 2007.
- [116] Napolitano, A., Spooner, C. M. Cyclic spectral analysis of continuous-phase modulated signals. *IEEE Transactions on Signal Processing*, 49(1), pages : 30-44, 2001.
- [117] Ong, H. T. *Bootstrap Methods for Signal Detection*. Thèse de doctorat, Curtin University of Technology, Australia, 2000.
- [118] Papoulis, A. *Probability, Random Variables and Stochastic Processes*. McGraw Hill. New York, 1962.
- [119] Parzen, E. Spectral analysis of asymptotically stationary time series. *Bull. Int. Stat. Inst.* 39(2), pages : 87-103, 1962.
- [120] Parzen, E. *Stochastic Processes*, Holden-Day. San Francisco, 1962.
- [121] Parzen, E. On Spectral analysis with missing observations. *Sankhya*, Ser. A, 25, pages : 383-392, 1963.
- [122] Politis, D., Romano, J., Wolf, M. *Subsampling*, Springer Series in Statistics, New-York : Springer, 1999.

- [123] Politis, D. N. Resampling time series with seasonal components, in *Frontiers in Data Mining and Bioinformatics : Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics*, Orange County, California, pages : 619-621, 2001.
- [124] Politis, D. N. The impact of bootstrap methods on time series analysis, *Statistical Science*, 18(2), pages : 219-230, 2003.
- [125] Rachael, D. S., Jessica, A. B., Taritonye, B. B., Brett, W. F., Mark, T. G., Joseph, T. G., Youngbin, K., David B. L. Motor Control and Aging : Links to Age-Related Brain Structural, Functional, and Biochemical Effects. *Neurosci Biobehav Rev*, 34(5), pages : 721-733, 2010.
- [126] Randall, R. B., Antoni, J., Chobsaard. S. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5), pages : 945–962, 2001.
- [127] Rao, J. S. Bootstrapping to Assess and Improve Atmospheric Prediction Models. *Data Mining and Knowledge Discovery*, 4(1), pages : 29-41, 2000.
- [128] Reid, D. C., Zoubir, A. M., Boashash, B. The Bootstrap Applied to Passive Acoustic Aircraft Parameter Estimation. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing, ICASSP96*, 6, pages : 3154-3157, Atlanta, 1996.
- [129] Allan, H. R., Martin, A. S. Adams & Victor's Principles of Neurology, 9th edition : The neurology of aging, chapter 29. March 12, 1572 pages, Publisher : McGraw-Hill Professional ; 9 edition, 2009.
- [130] Rosano, C., Aizenstein, H., Brach, J., Longenberger, A, S. Gait measures indicate underlying focal gray matter atrophy in the brain of older adults. *J Gerontol A Biol Sci Med Sci*, 63(12), pages : 1380-8, 2008.
- [131] Sabri, K., El Badaoui, M., Guillet, F., Belli, A., Millet, G., Morin, J. B. Cyclostationary modelling of ground reaction force signals, *Signal Processing*, 90(4), pages : 1146-1152, 2010.
- [132] Sacchi, M. D. A Bootstrap Procedure for High-Resolution Velocity Analysis. *Geophysics*, 63(5), pages : 1716-1725, 1998.
- [133] Serpedin, E., Panduru, F., San, I., Giannakis, G. B. Bibliography on cyclostationarity. *Signal Processing* 85(12), pages : 2233-2303, 2005.

- [134] Shao, J., Tu, D. The Jackknife and Bootstrap. Springer Verlag, New York, 1995.
- [135] Singh, K. On the Asymptotic Accuracy of Efron's Bootstrap. The Annals of Statistics, 9(6), pages : 1187-1195, 1981.
- [136] Springer, S., Giladi, N., Peretz, C., Yogev, G., Simon, E. S., Hausdorff, J. M. Dual-tasking effects on gait variability : the role of aging, falls, and executive function. Movement Disorders, 21(7), pages : 950-957, 2006.
- [137] Synowiecki, R. Consistency and application of moving block bootstrap for nonstationary time series with periodic and almost periodic structure. Bernoulli, 13(4), pages : 1151-1178, 2007.
- [138] Synowiecki, R. Resampling methods in the time domain for nonstationary time series with periodic and almost periodic structures (PL). Thèse de doctorat, département des mathématiques appliquées, AGH University of Science and Technology, Cracovie, Pologne, 2008.
- [139] Ulrych, T. J., Sacchi, M. D. Sompi, Pisarenko and the Extended Information Criterion. Geophysical Journal International, 122(3), pages : 719-724, 1995.
- [140] Widrow, B., Glover, J. R., McCool, J. M., Kaunitz, J. Adaptive noise cancelling : Principles and applications, Proceedings of the IEEE, 63(12), pages : 1692-1716, 1975.
- [141] Yaglom, A. M. Correlation Theory of Stationary and Related Random Functions, Springer Verlag, New York, 1987.
- [142] Zoubir, A. M., Bouashash, B. The bootstrap and its application in signal processing. Signal Processing Magazine, IEEE, 15(1), pages : 56-76, 1998.
- [143] Zoubir, A. M., Iskander, D. R., Chant, I., Carevic, D. Detection of Landmines Using Ground Penetrating Radar. In Detection and Remediation Technologies for Mines and Minelike Targets IV, 3710, pages : 1301-1312, San Diego, 1999.
- [144] Zoubir, A., Iskander, R. Bootstrap techniques for signal processing, Cambridge University Press, 232 pages, 2004.