



**HAL**  
open science

# Traitement du signal dans le domaine compressé et quantification sur un bit : deux outils pour les contextes sous contraintes de communication

Augusto Zebadúa

► **To cite this version:**

Augusto Zebadúa. Traitement du signal dans le domaine compressé et quantification sur un bit : deux outils pour les contextes sous contraintes de communication. Traitement du signal et de l'image [eess.SP]. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAT085 . tel-01761603

**HAL Id: tel-01761603**

**<https://theses.hal.science/tel-01761603>**

Submitted on 9 Apr 2018

**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

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : SIGNAL IMAGE PAROLE TELECOMS

Arrêté ministériel : 25 mai 2016

Présentée par

**Augusto ZEBADÚA**

Thèse dirigée par **Pierre-olivier AMBLARD**, Directeur de recherche, CNRS, et co-encadre par **Eric MOISAN**, Maître de conférences, Grenoble INP.

préparée au sein du **Laboratoire Grenoble Images Parole Signal Automatique (GIPSA-lab)** dans l'**École Doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)**

**Traitement du signal dans le domaine compressé et quantification sur un bit : deux outils pour les contextes sous contraintes de communication**

**Compressed-domain signal processing and one-bit quantization: two tools for contexts under communication constraints**

Thèse soutenue publiquement le **11 décembre 2017**,  
devant le jury composé de :

**Jean-Marc BROSSIER**

Professeur, Grenoble INP, Président du Jury.

**Jean-François BERCHER**

Professeur, ESIEE Paris, Rapporteur.

**André FERRARI**

Professeur, Université de Nice, Rapporteur.

**Rodrigo CABRAL FARIAS**

Maître de Conférences, Université de Nice, Examineur.

**Pierre-Olivier AMBLARD**

Directeur de Recherche, CNRS, Directeur de thèse.

**Eric MOISAN**

Maître de Conférences, Grenoble INP, Co-encadrant de thèse.

**Olivier MICHEL**

Professeur, Grenoble INP, Invité.

UNIVERSITÉ GRENOBLE ALPES  
**ÉCOLE DOCTORALE EEATS**  
Électronique, Électrotechnique, Automatique et Traitement du Signal

# THÈSE

pour obtenir le titre de

**Docteur en traitement du signal**

de l'Université Grenoble Alpes

Présentée par

Augusto ZEBADÚA

**Traitement du signal dans le domaine compressé et  
quantification sur un bit : deux outils pour les  
contextes sous contraintes de communication**

Thèse dirigée par Pierre-Olivier AMBLARD  
et coencadrée par Eric MOISAN

préparée au laboratoire Grenoble Image Parole Signal  
Automatique (GIPSA-lab)

soutenue publiquement le 11 Décembre 2017

**Jury :**

<i>Président :</i>	Jean-Marc BROSSIER	- Pr. Gipsa-lab DIS/CICS
<i>Rapporteurs :</i>	Jean-François BERCHER	- Pr ESIEE Paris
	André FERRARI	- Pr. Laboratoire Lagrange
<i>Examineur :</i>	Rodrigo CABRAL FARIAS	- MC. Laboratoire I3S
<i>Directeur :</i>	Pierre-Olivier AMBLARD	- DR. Gipsa-lab DIS/CICS
<i>Encadrant :</i>	Eric MOISAN	- MC. Gipsa-lab DIS/CICS
<i>Invité :</i>	Olivier MICHEL	- Pr. Gipsa-lab DIS/CICS

## Résumé

La surveillance de phénomènes physiques à l'aide d'un réseau de capteurs (autonomes mais communicants) est fortement contrainte en consommation énergétique, principalement pour la transmission de données. Dans ce cadre, cette thèse propose des méthodes de traitement du signal permettant de réduire les communications sans compromettre la précision des calculs ultérieurs. La complexité de ces méthodes est réduite, de façon à ne consommer que peu d'énergie supplémentaire. Deux éléments servent à leur synthèse : la compression dès l'acquisition (Acquisition compressive) et la quantification grossière (sur 1 bit).

D'abord, on étudie le *corrélateur compressé*, un estimateur qui permet d'évaluer les fonctions de corrélation, temps de retard et densités spectrales en exploitant directement des signaux compressés. Ses performances sont comparées au corrélateur usuel. Si le signal à traiter possède un support spectral étroit, l'estimateur proposé s'avère sensiblement meilleur que l'usuel.

Ensuite, inspirés par les corrélateurs à forte quantification des années 50 et 60, deux nouveaux corrélateurs sont étudiés : le *compressé sur 1 bit* et le *compressé hybride*, qui peuvent également surpasser les performances de leurs contreparties non-compressées.

Finalement, on montre la pertinence de ces méthodes pour les applications envisagées à travers l'exploitation de données réelles.

## Abstract

Monitoring physical phenomena by using a network of sensors (autonomous but interconnected) is highly constrained in energy consumption, mainly for data transmission. In this context, this thesis proposes signal processing tools to reduce communications without compromising computational accuracy in subsequent calculations. The complexity of these methods is reduced, so as to consume only little additional energy. Our two building blocks are compression during signal acquisition (Compressive Sensing) and coarse quantization (1 bit).

We first study the *Compressed Correlator*, an estimator which allows for evaluating correlation functions, time-delay, and spectral densities directly from compressed signals. Its performances are compared with the usual correlator. As we show, if the signal of interest has reduced frequency content, the proposed estimator significantly outperforms the conventional correlator.

Then, inspired by the coarse quantization correlators from the 50s and 60s, two new correlators are studied : The *1-bit Compressed* and the *Hybrid Compressed*, which can also outperform their uncompressed counterparts.

Finally, we show the applicability of these methods in the context of interest through the exploitation of real data.

## Remerciements

Cette thèse récapitule le travail effectué pendant un peu plus de trois ans au cours desquels j'ai eu le plaisir d'intégrer le laboratoire GIPSA. J'ai compté sur la contribution précieuse de plusieurs personnes à qui je veux remercier :

Tout d'abord, Olivier MICHEL. de m'avoir fait confiance pour travailler dans ce sujet, et d'avoir apporté des éléments clés dans la synthèse de cette thèse. Deuxièmement, Eric MOISAN qui a contribué énormément à ce travail. Non seulement sur le plan technique, mais aussi sur le plan humain. J'apprécie particulièrement la grande patience avec laquelle il a examiné chaque détail de ce travail y compris en langue Française. Troisièmement, Pierre Olivier AMBLARD. J'apprécie la liberté qu'il m'a donnée et toutes les très bonnes idées qu'il a apportées à cette thèse. Puis je remercie les membres du jury, merci d'avoir accepté d'évaluer mon travail et de toutes vos remarques et conseils..

Je remercie particulièrement Pierre GRANJON pour son soutien à un moment difficile. Sans sa grande perspicacité et la clarté de ses idées, plusieurs résultats de cette thèse seraient probablement absents. Merci aussi à Jean-Louis LACOUME qui m' accordé du temps précieux et qui m'a donné beaucoup de pistes à suivre grâce à sa grande maîtrise de la discipline.

Je remercie mes collègues de bureau, Easter SELVAN, Ewaldo SANTANA, Pierre PIGNEDE et Thibaut TEZENAS qui ont grandement contribué à surmonter les moments d'isolement avec des échanges agréables et intéressants. Ensuite Je remercie ceux qui ont aidé avec leur temps ou bien leur remarques. Laurent ROS, Steve ZOZOR, Florent CHATELAIN, Ronald PHLIPO, Nicolas TREMBLAY, et Jeanne LEFEVRE.

J'ai compté également sur l'apport inestimable de ceux qui beaucoup contribué bien qu'indirectement.

D'abord mes plus profonds remerciements à ma Stephanie, à qui je dédie cette thèse. Elle a absolument tout fait pour me soutenir, surtout dans les moments les plus sombres de ces derniers années. Je suis très reconnaissant de son immense amour, de sa patience infinie et de tous les sacrifices personnels qu'elle a fait afin que je finisse ce manuscrit.

Je remercie mon père, d'être mon exemple, de m'avoir guidé jusqu'ici et de m'avoir toujours encouragé à viser le plus haut. Merci à mon frère Julio et ma sœur Ilsecita qui ont toujours cru en moi et m'ont motivé à donner le meilleur de moi-même.

Je dis merci à mes amis, Ricardo et Aline de leur soutien à des moments difficiles, d'avoir partagé l'expérience de vivre dans un autre pays et de savoir en extraire le meilleur. Enfin, merci à Arturo qui m'a fait venir en France, sans son influence je serais sûrement ailleurs et j'aurais raté l'une des meilleurs expériences de ma vie.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Traitement du signal sous contraintes . . . . .	1
1.2	Estimation passive . . . . .	2
1.2.1	Structural Health Monitoring (SHM) . . . . .	3
1.3	Réseaux de capteurs . . . . .	3
1.4	Modèle et approche suivie . . . . .	4
1.5	Aperçu et contributions . . . . .	7
	Bibliographie . . . . .	9
<b>2</b>	<b>Traitement du signal dans le domaine compressé</b>	<b>11</b>
2.1	Traitement numérique du signal . . . . .	11
2.2	Echantillonnage . . . . .	12
2.3	Acquisition compressive . . . . .	12
2.3.1	Acquisition compressive numérique . . . . .	13
2.3.2	Reconstruction . . . . .	16
2.3.3	Matrice $\Phi$ et son lien avec $\Psi$ . . . . .	18
2.3.4	Défis à relever et tendances de la recherche . . . . .	19
2.4	Traitement du signal compressé TSC . . . . .	20
2.4.1	Outils prédécesseurs : compression des signaux orientée vers les tâches	22
2.4.2	Quelques travaux récents . . . . .	24
2.4.3	Produit interne et corrélation . . . . .	25
	Bibliographie . . . . .	26
<b>3</b>	<b>Estimation dans le domaine compressé</b>	<b>29</b>
3.1	Contexte considéré . . . . .	29
3.2	Estimation de corrélation . . . . .	30
3.2.1	Corrélateur classique . . . . .	31
3.2.2	Corrélateur compressé . . . . .	32
3.2.3	Influence de $\Phi$ dans la performance . . . . .	34
3.2.4	Corrélateur compressé par échantillonnage aléatoire . . . . .	36
3.2.5	Analyse asymptotique à taux de compression fixé . . . . .	36
3.2.6	Compressibilité et performance . . . . .	39
3.2.7	Effet du bruit . . . . .	44
3.3	Estimation du temps de propagation . . . . .	46
3.4	Estimation d'interspectre . . . . .	48
3.4.1	Corrélogramme Blackman-Tukey . . . . .	48
3.4.2	Corrélogramme classique . . . . .	49
3.4.3	Corrélogramme compressé . . . . .	50
3.4.4	Corrélogramme compressé par échantillonnage aléatoire . . . . .	51
3.4.5	Compressibilité et performance . . . . .	52
	Bibliographie . . . . .	60

<b>4</b>	<b>Traitement des signaux quantifiés</b>	<b>61</b>
4.1	Quantification orientée vers la tâche . . . . .	61
4.2	Corrélateurs fortement quantifiés . . . . .	62
4.2.1	Corrélateurs avec des signaux à deux niveaux . . . . .	62
4.2.2	Corrélateurs hybrides . . . . .	63
4.2.3	Corrélateurs généralisés et modifiés . . . . .	65
4.3	Acquisition compressive et quantification . . . . .	67
4.4	Corrélogrammes efficaces . . . . .	68
4.4.1	Densité spectrale de puissance . . . . .	68
4.4.2	Densité spectrale d'interaction . . . . .	69
	Bibliographie . . . . .	70
<b>5</b>	<b>Estimation dans le domaine compressé et quantifié</b>	<b>73</b>
5.1	Corrélateurs sur 1 bit . . . . .	73
5.1.1	Estimateur classique sur 1 bit . . . . .	75
5.1.2	Estimateur compressé sur 1 bit . . . . .	77
5.1.3	Compressibilité et performance . . . . .	77
5.2	Corrélateurs hybrides . . . . .	79
5.2.1	Estimateur classique hybride . . . . .	80
5.2.2	Estimateur compressé hybride . . . . .	81
5.2.3	Compressibilité et performance . . . . .	81
5.3	Corrélogrammes compressés et quantifiés . . . . .	83
5.3.1	Corrélogrammes sur 1 bit . . . . .	84
5.3.2	Corrélogramme classique sur 1 bit . . . . .	84
5.3.3	Corrélogramme compressé sur 1 bit . . . . .	84
5.3.4	Compressibilité et performance . . . . .	85
5.3.5	Corrélogrammes hybrides . . . . .	87
	Bibliographie . . . . .	90
<b>6</b>	<b>Application à des signaux réels</b>	<b>91</b>
6.1	Structure A (Signaux modérément corrélés) . . . . .	91
6.1.1	Estimation de $\Gamma_{xy}(\tau)$ . . . . .	93
6.1.2	Estimation de $\Gamma_{xz}(\tau)$ . . . . .	95
6.2	Structure L (Signaux faiblement corrélés) . . . . .	98
6.2.1	Estimation de $\Gamma_{xy}(\tau)$ . . . . .	99
6.2.2	Estimation de $\gamma_{xy}(\nu)$ . . . . .	101
	<b>Conclusions et perspectives</b>	<b>105</b>
	Bibliographie . . . . .	107
	<b>Annexes</b>	<b>109</b>
	Bibliographie . . . . .	123

# Chapitre 1

## Introduction

### Sommaire

---

<b>1.1</b>	<b>Traitement du signal sous contraintes</b>	<b>1</b>
<b>1.2</b>	<b>Estimation passive</b>	<b>2</b>
1.2.1	Structural Health Monitoring (SHM)	3
<b>1.3</b>	<b>Réseaux de capteurs</b>	<b>3</b>
<b>1.4</b>	<b>Modèle et approche suivie</b>	<b>4</b>
<b>1.5</b>	<b>Aperçu et contributions</b>	<b>7</b>
	<b>Bibliographie</b>	<b>9</b>

---

### 1.1 Traitement du signal sous contraintes

Le traitement du signal, discipline scientifique qui s'intéresse à l'acquisition, l'analyse et l'exploitation des données numériques issues de phénomènes naturels ou humains, a connu un énorme succès ces dernières années et possède à l'heure actuelle une valeur et une importance énorme dans la vie moderne. Les développements dans cette discipline sont à la base du progrès technologique actuel et ont fourni des capacités qui étaient inaccessibles aux appareils analogiques. Les smartphones, les appareils photo et vidéo numériques de haute résolution dont nous disposons aujourd'hui en constituent quelques exemples notables.

Pour y parvenir, la recherche en traitement du signal a concentré majoritairement ses efforts sur deux questions fondamentales : *Quelles nouvelles informations peut-on extraire/modifier d'un jeu de données ?* et *Comment améliorer les méthodes de traitement existantes pour les rendre plus performantes et plus précises ?*. Au cours des dernières années, on a cherché à répondre à ces questions entre autre en développant des techniques innovantes de traitement, en exploitant de nouveaux outils mathématiques et en réduisant la complexité des algorithmes de traitement.

Cependant, s'intéresser exclusivement à ces deux questions, suppose implicitement que les ressources matérielles disponibles sont toujours suffisantes. Or, même si l'on conçoit des dispositifs électroniques toujours plus rapides et performants, ceux-ci peuvent s'avérer insuffisants.

Pensons par exemple au déluge de données que nous subissons à l'heure actuelle : la taille des données engendrées par les appareils électroniques numériques (smartphones, ordinateurs portables, cameras vidéo) a explosé. Bien que notre capacité de stockage soit très grande et en constante augmentation, le volume de ces données est trop important

pour pouvoir tout sauvegarder. De plus, en tant qu'utilisateurs, on souhaite toujours des appareils plus précis, de meilleure résolution, et plus performants. Les exigences sur l'électronique risquent donc de s'aggraver encore plus dans le futur.

Un autre exemple est arrivé avec la croissance des nouvelles technologies comme les réseaux de capteurs, où un très grand nombre de données sont enregistrées à différents endroits par des petits appareils électroniques, puis sont communiquées avant de les traiter. Dans ce contexte, les données peuvent s'avérer trop volumineuses pour être traitées. De même, les ressources de calcul et surtout énergétiques à disposition peuvent être trop limitées pour effectuer le traitement.

Une troisième question fondamentale est donc devenue de plus en plus importante en traitement du signal : *Comment peut-on traiter les signaux lorsque les exigences dépassent les moyens disponibles ?* Cet aspect, qui n'a reçu pendant longtemps que peu d'intérêt, redevient d'actualité.

Une réponse possible à cette question passe par l'acquisition des signaux. Les méthodes classiquement utilisées (l'échantillonnage périodique et la quantification fine) ont été conçues pour l'obtention des représentations précises des signaux ce qui offre une flexibilité énorme : on peut effectuer de nombreux traitements ou études sur ces données numériques. En contrepartie le coût de traitement associé à ces méthodes peut être élevé. L'idée que nous poursuivons dans cette thèse est celle de remplacer cette acquisition précise et flexible mais coûteuse, par une acquisition plus simple et économique, permettant de ne garder que l'information dont on a besoin. La représentation numérique qui en résulte peut s'avérer difficile à interpréter directement, cependant la capacité à effectuer le traitement sur ces données doit être préservée.

Partant d'une représentation numérique différente, le réexamen d'algorithmes conventionnels de traitement du signal dans ce contexte est nécessaire et nous semble une activité de recherche prometteuse dans les années à venir pour deux raisons :

- Premièrement, on réduit l'impact du déluge de données, car de nouveaux algorithmes basés sur une acquisition plus économique étendent la capacité du matériel disponible, ce qui permet d'effectuer des traitements sur des données qui seraient probablement perdues autrement.
- Deuxièmement, acquérir les signaux différemment peut réduire les demandes sur les appareils électroniques qui seront développés dans le futur, ouvrant ainsi la porte à des dispositifs plus compacts, moins consommateurs d'énergie et plus économiques.

Les paragraphes suivants présentent les généralités sur le traitement du signal qui nous intéresse et les contraintes imposées par ce contexte.

## 1.2 Estimation passive

L'application du traitement du signal motivant ces travaux est l'extraction de caractéristiques, physiques ou géométriques (par exemple, géométrie, densité, masse etc), qui ne sont pas directement mesurables dans un environnement. Ces caractéristiques sont obtenues indirectement, à partir de signaux qui portent de l'information sur ces propriétés.

Une possibilité pour les obtenir est de perturber le milieu à étudier par des signaux engendrés par le traicteur du signal. Ensuite, les réponses du milieu à ces perturbations sont enregistrées et exploitées pour en déduire ses propriétés. Les caractéristiques des signaux produits sont optimisées pour mieux extraire ces informations. Le traitement de ces signaux *ad hoc* générés par l'homme est nommé *estimation active*.

Si les ressources énergétiques sont limitées, on peut plutôt extraire ces informations à l'aide de signaux naturels engendrés à l'intérieur du milieu. Cette extraction d'information lorsque les perturbations ne sont ni connues ni contrôlées est appelée *estimation passive*. Les inconvénients liés à cette approche sont :

- Ces excitations ne possèdent pas les caractéristiques optimales
- Le manque d'information *a priori* sur les signaux à traiter

L'estimation passive est exploitable dans différents contextes. Nous nous intéressons particulièrement au suivi de l'état des structures (*Structural Health Monitoring*) qui est présenté au paragraphe suivant.

### 1.2.1 Structural Health Monitoring (SHM)

Le SHM est une problématique, à la frontière du génie civil, de la mécanique et du traitement du signal, qui cherche à établir quantitativement des indices sur l'état des bâtiments et notamment à rendre compte de leur état physique : les forces et charges subies ou d'éventuels dommages ou détériorations. Avoir des indices sur la santé des structures présente un intérêt économique mais est surtout lié à la sécurité des utilisateurs. Voir par exemple [6, 9] et ses références.

Dans cette problématique, les informations à extraire à partir des perturbations inconnues sont par exemple la masse des structures, leur amortissement ou des variations structurelles. Ces dernières s'avèrent très difficiles à mesurer directement. Ces informations peuvent être déduites du comportement dynamique des structures, notamment en étudiant les vibrations qu'elles créent en réponse aux forces externes auxquelles elles sont soumises. Ces vibrations sont mesurées à l'aide d'accéléromètres.

Typiquement, l'estimation passive est mise en place à l'aide d'un réseau de capteurs qui effectuent des mesures locales sur la dynamique. L'exploitation simultanée des résultats locaux permet alors d'extraire les propriétés et l'état global de la structure.

## 1.3 Réseaux de capteurs

Les réseaux de capteurs sont un ensemble de dispositifs électroniques, déployés dans un environnement pour mesurer et suivre le comportement des grandeurs physiques à différents points géographiques. La répartition des tâches sur l'ensemble de ces petits capteurs peut permettre la réalisation d'un traitement qui resterait inaccessible à un seul ordinateur, même puissant [8]

Puisque les variables physiques à étudier sont susceptibles d'évoluer en temps et en espace, il est souhaitable de disperser un grand nombre de capteurs dans le milieu pour avoir une bonne résolution spatiale. Un des buts principaux des concepteurs est donc de construire des capteurs à faible coût pour pouvoir les déployer en grande quantité.

La tâche d'un capteur ne se limite pas à la récolte des données. Il les transmet aussi. Typiquement, chaque capteur est doté d'au moins trois éléments :

- Des transducteurs pour convertir les grandeurs physiques en données numériques.
- Un dispositif numérique de calcul.
- L'électronique nécessaire à la transmission des mesures, la transmission sans fil étant le choix le plus fréquent.

Un point important à souligner est qu'un grand nombre d'environnements d'étude est très difficile d'accès pour l'Homme. Il arrive même qu'il soit impossible de redonner de l'énergie aux capteurs et que la durée d'utilisation du réseau de capteurs dépende de la batterie de ces derniers. Consommer moins d'énergie signifie pouvoir exploiter le réseau plus longtemps.

La performance d'un réseau de capteurs est jugée sous plusieurs aspect : la précision dans les calculs, la tolérance aux erreurs et particulièrement la scalabilité, c'est-à-dire la capacité d'un réseau à grandir autant que souhaité. Ce qui détermine ces caractéristiques est l'architecture (la façon dont les capteurs interagissent). Il en existe deux principales :

- **Centralisée** : constituée de petits capteurs et une unité centrale, celle-ci possède beaucoup de ressources énergétiques et calculatoires. Le principe de fonctionnement est la transmission des observations individuelles des capteurs vers cette centrale. Le plus souvent, le traitement et l'analyse conjoints des données y sont réalisés sans contraintes sur la complexité ou la consommation énergétique des algorithmes de traitement.

Le principal inconvénient d'un réseau de ce type est son manque de scalabilité, puisque toutes les données sont envoyées vers le même endroit. L'énergie requise augmente avec la surface étudiée. En plus, la puissance transmissible dans certains endroits, comme les villes, est restreinte et régulée par la loi, ce qui limite la taille maximale d'un réseau centralisé.

- **Distribuée ou collaborative** : le réseau ne comporte que des capteurs qui se partagent la globalité du traitement. Ici la transmission centralisée est remplacée par le partage d'information entre capteurs. Avec cette architecture, le réseau peut être scalable si les communications de chaque capteur sont limitées à ses voisins. Naturellement, ce sont les capteurs qui effectuent directement tous les calculs de manière collaborative. Les fortes contraintes énergétiques limitent grandement le choix des algorithmes de traitement. Ceux-ci doivent être très simples et d'une complexité réduite.

Nous nous intéressons plutôt à la deuxième architecture, mais le modèle que nous considérons, présenté dans le paragraphe suivant, reste valable dans les deux cas. Ce dernier paragraphe présente également une courte discussion sur les compromis à gérer dans les réseaux de capteurs.

## 1.4 Modèle et approche suivie

Dans le cadre de réseaux de capteurs, l'énergie est un aspect clé. Toutes les tâches effectuées par un capteur en consomment mais la transmission de données est la tâche qui nécessite le plus d'énergie. Cf [10, 11]. ou [2] pour une étude brève mais très illustrative concernant les compromis à gérer dans ce contexte.

Pour s'attaquer à l'aspect critique, les communications, une possibilité consiste à réduire la taille des données, mais on doit prendre en compte le fait que ces données sont utilisées dans des calculs. Une réduction naïve de leur taille peut dégrader excessivement la précision. Différentes manières de mettre en place cette réduction ont été étudiées dans

la littérature. La plupart font intervenir la quantification<sup>1</sup>, d'autres, comme nous, se sont intéressées simultanément à la compression et à la quantification [1, 3]. Ces derniers travaux ont proposé une approche : *compression de données pour des tâches d'inférence* et suggéré un modèle du problème. Nous reprenons leur modèle dans ce manuscrit : Un capteur,  $x$ , comprime les données qu'il mesure avant de les transmettre vers un autre capteur  $y$  où les traitements sont effectués. L'avantage du modèle est qu'il est valable pour des réseaux centralisés et distribués.

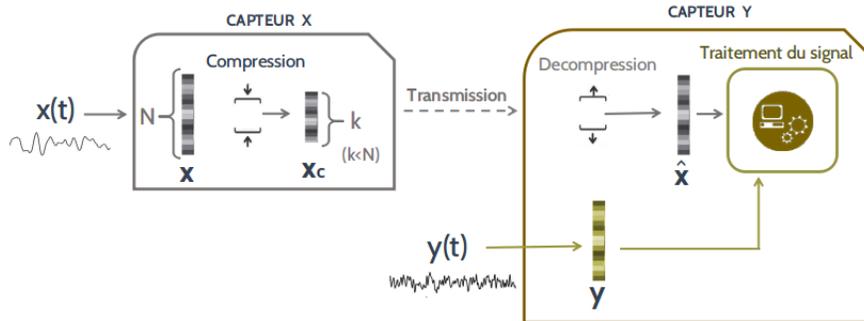


FIGURE 1.1 – Modèle d'étude de Hao et Fowler [1]. Un des capteurs comprime les données qu'il transmet vers un deuxième capteur. Les calculs sont effectués dans ce dernier.

Ces travaux ont souligné deux aspects importants :

- Premièrement, la compression de données peut elle-même nécessiter des calculs, ce qui consomme de l'énergie. Néanmoins, le coût de la transmission de données brutes est plus important que celui d'une compression permettant une moindre transmission. Il en résulte une économie d'énergie.
- Deuxièmement, si le but principal est de préserver la précision calculatoire, il peut être préférable d'orienter la compression vers le traitement souhaité plutôt que vers le type de signal à traiter.

A propos du second point, les résultats [1, 3] montrent, dans le cadre de deux problèmes d'estimation paramétrique : différence de temps d'arrivée, et différence de fréquence d'arrivée (TDOA et FDOA), que la compression orientée vers la tâche permet de bien préserver la performance d'estimation. Pourtant, leur approche de compression possède un désavantage important si plusieurs tâches d'estimation sont nécessaires : étant optimisée pour le traitement à effectuer, les exigences sur la compression peuvent être différentes et même contradictoires pour des tâches distinctes .

Nos travaux s'appuient aussi sur la compression, mais avec un paradigme différent : l'acquisition compressive (Compressive Sensing). Cette méthode de numérisation, relativement récente, cherche à acquérir et simultanément compresser les signaux à traiter. Elle est l'objet d'un très grand nombre de travaux récents principalement parce qu'elle a permis, dans certains contextes, de dépasser la limite dictée par le critère de Nyquist

1. La référence [3] contient une vue d'ensemble sur ces méthodes

quant au nombre minimum de données nécessaires pour représenter numériquement un signal.

Nous soulignons que la quasi-totalité des travaux en acquisition compressive reconstruisent le signal avant de faire le traitement, mais le coût énergétique associé est très élevé et non envisageable dans notre contexte. De manière intéressante, cette méthode d'acquisition ouvre la porte aussi au *traitement du signal compressé* (Compressive Signal Processing), contournant ainsi le passage par cette coûteuse étape de reconstruction qui précède le traitement. Il a été montré par Davenport, [4] que l'acquisition compressive permet non seulement de compresser lors de l'acquisition, mais aussi d'effectuer des tâches d'inférence classiques, telles que la classification, la détection ou l'estimation, directement sur les signaux *dans le domaine compressé*. Ces deux approches sont illustrées dans la figure 1.2

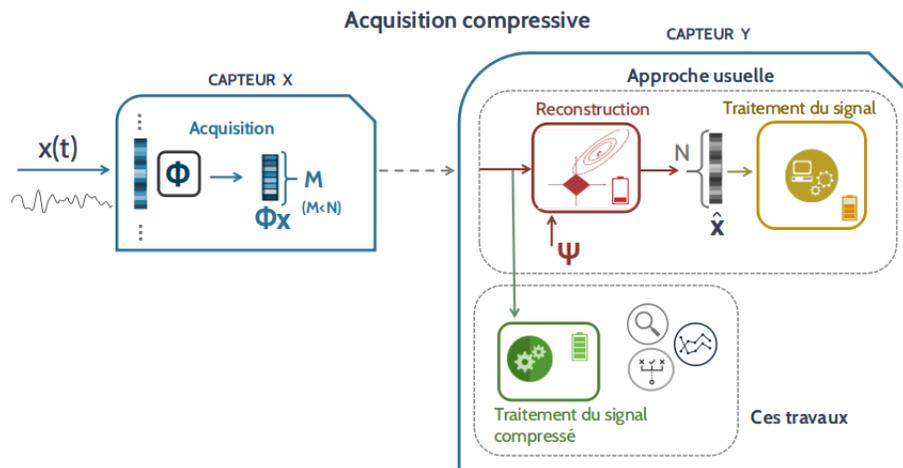


FIGURE 1.2 – Approches pour l’acquisition compressée, la plus connue avec reconstruction (en haut) est très chère en termes énergétiques, le traitement des signaux dans le domaine compressé (en bas) est une solution plus économique.

L’acquisition compressive en tant que méthode de compression possède les caractéristiques suivantes [4] :

Elle est qualifiée d’*universelle*, ce qui veut dire que l’on peut acquérir et compresser des signaux de nature différente avec la même méthode de compression. Au lieu d’un taux de compression élevé pour un type de signal précis (compression *ad hoc*), on obtient une réduction sous-optimale de la taille des données pour des signaux possédant des caractéristiques distinctes.

Elle est également *aveugle*, puisque aucune connaissance sur le signal, ni sur le traitement à effectuer, n’est nécessaire. Il s’agit donc d’un bon choix dans le cadre de l’estimation passive, caractérisée justement par le manque d’information a priori sur le signal. Cette caractéristique permet de contourner les exigences contradictoires qui peuvent apparaître dans la compression orientée vers la tâche, lorsque plusieurs traitements sont nécessaires.

La figure 1.3 montre les trois approches évoquées : la compression classique, la compression orientée vers la tâche et l’acquisition compressive.

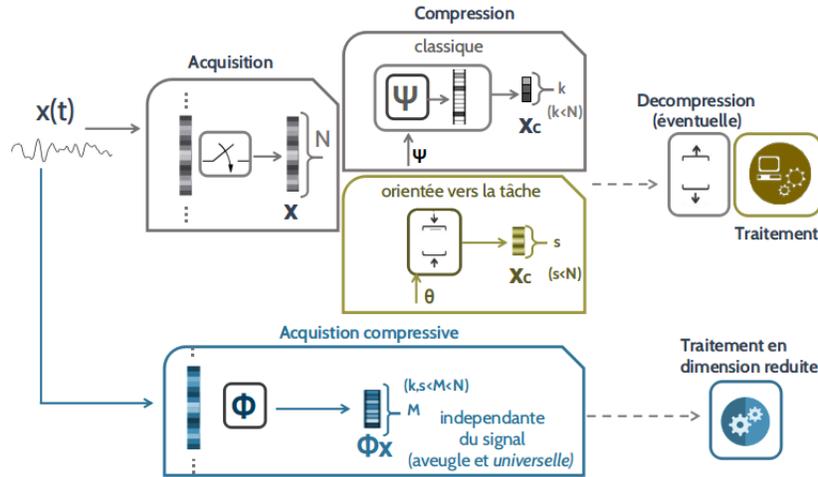


FIGURE 1.3 – Trois approches : compression classique, compression orientée vers la tâche et acquisition compressive. Compresser consomme de l'énergie, mais transmettre les données brutes en consomme davantage. Compresser dégrade la performance mais réduit la consommation énergétique.

Compresser n'est pas la seule façon de réduire la taille des données. Une autre possibilité réside dans la quantification des signaux. Nous nous sommes donc intéressés à un deuxième outil, plus ancien. Développé pendant les années 50 et 60, le traitement de signaux fortement quantifiés a été utilisé dans le passé pour réaliser des calculs avec des ressources très limitées. Le corrélateur à coïncidence de polarité [5] et le corrélateur à relais [7] en sont des exemples notables. En effet, il n'est pas toujours nécessaire d'utiliser un grand nombre de bits pour représenter les valeurs numériques des signaux. On peut même en utiliser très peu pour quantifier astucieusement et ainsi tirer le meilleur parti du matériel disponible.

Les travaux présentés dans ce manuscrit s'appuient sur ces deux outils : compression et quantification pour le développement de méthodes de traitement du signal dans un cadre distribué dans lequel on retrouve trois besoins :

- Limiter drastiquement la transmission de données pour économiser en énergie mais sans compromettre la précision des calculs ultérieurs.
- Garder une complexité raisonnable de traitement.
- Effectuer un traitement aveugle (sans a priori sur le signal)

## 1.5 Aperçu et contributions

Le reste du manuscrit est organisé comme suit :

- Le **chapitre 2** s'intéresse à quelques généralités sur les outils que nous venons d'exposer : l'acquisition compressive et le *traitement des signaux dans le domaine compressé*.
- Le **chapitre 3** est consacré à l'estimation distribuée dans le domaine compressé.

Dans un premier temps nous étudions un nouvel estimateur de corrélation, le *corrélateur compressé*. Sa performance est comparée à celle du corrélateur usuel dans un cadre d'estimation distribuée, c'est-à-dire avec des signaux provenant d'endroits différents. L'influence de la matrice de compression ainsi que celle du

type du signal sont étudiées et illustrées à travers le cas particulier d'un processus AR(1). Dans un deuxième temps, nous étudions deux estimations supplémentaires qui se déduisent de la fonction corrélation : le temps de propagation et la densité spectrale d'interaction.

- Le **chapitre 4** présente quelques notions sur le traitement des signaux fortement quantifiés et notamment sur les corrélateurs à forte quantification des années 50 et 60.
- Au **chapitre 5**, inspirés des corrélateurs à quantification grossière, deux nouveaux estimateurs de corrélation sont étudiés : les corrélateurs *compressé sur 1 bit* et *compressé hybride*. Ces estimateurs exploitent simultanément la compression et la quantification sur 1 bit.
- Puis, le **chapitre 6**, est consacré à des données réelles. Nous comparons les performances des méthodes *classiques* d'estimation à leur version compressé.
- Enfin, nous concluons et proposons quelques pistes de recherche qui méritent d'être poursuivies.

## Publications de l'auteur

### Article de Journal

- ZEBADUA, A. AMBLARD, P.O. MOISAN, E. MICHEL, O. Compressed and quantized correlation estimators. *IEEE Transactions on Signal Processing* 65.1 (2017) : 56-68.

Cet article paru en *Transactions on Signal Processing*, présente l'analyse théorique sur l'estimation de corrélation à partir des signaux compressés, et des simulations Monte-Carlo pour des signaux compressés et quantifiés sur 1 bit.

### Conférences Internationales

- ZEBADUA, A. MICHEL, O. AMBLARD, P.O. MOISAN, E. Signal processing with coarse-quantized measurements : efficient cross-correlation estimation, Proc. IMA conf on Math. and Sig. Proc., Birmingham, UK, dec 2014.

Cet article présenté à *IMA*, montré des résultats préliminaires sur l'estimation de corrélation avec des signaux compressés et quantifiés sur 1 bit.

- ZEBADUA, A. AMBLARD, P.O. MOISAN, E. MICHEL, O. CARMONA, M. et al. Examples of output-only modal identification using compressive sensing techniques. *IWSHM 2015*, Stanford, USA, 2015.

Cet article présenté à *IWSHM 2015*, en collaboration avec le CEA-LETI Grenoble, porte sur quelques résultats préliminaires sur l'estimation spectrale dans le domaine compressé effectué sur des données réelles.

### Conférences Nationales

- ZEBADUA, A. AMBLARD, P.O. MOISAN, E. MICHEL, O. Estimation de corrélation sur signaux compressés. *GRETSI, 2015*, Lyon, France.
- ZEBADUA, A. AMBLARD, P.O. MOISAN, E. Sur le corrélateur compressé : Estimation de temps de retard et d'interspectre *GRETSI, 2017*, Juan-les-Pins, France.

Ces articles de conférence présentés au GRETSI proposent à la communauté française de traitement du signal des résultats préliminaires concernant l'estimation de corrélation, le temps de retard et l'interspectre à partir des signaux compressés.

## Bibliographie

- [1] CHEN, M. *Data compression for inference tasks in wireless sensor networks*. PhD thesis, State University of New York, 2006.
- [2] CHEN, M., AND FOWLER, M. L. Data compression trade-offs in sensor networks. In *Optical Science and Technology, the SPIE 49th Annual Meeting* (2004), International Society for Optics and Photonics, pp. 96–107.
- [3] CHEN, M., AND FOWLER, M. L. Data compression for multi-parameter estimation for emitter location. *IEEE Transactions on Aerospace and Electronic Systems* 46, 1 (2010).
- [4] DAVENPORT, M. A., BOUFONOS, P. T., WAKIN, M. B., AND BARANIUK, R. G. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing* 4, 2 (2010), 445–460.
- [5] FARAN, J., JAMES, J., HILLS, J., ET AL. Correlators for signal reception. Tech. rep., HARVARD UNIV CAMBRIDGE MA ACOUSTICS RESEARCH LAB, 1952.
- [6] MIKAEL, A. *Evaluation des paramètres physiques des bâtiments : amortissement, fréquence et modes de comportement des structures de génie civil : Approche expérimentale*. PhD thesis, Université de Grenoble, 2011.
- [7] NUTTALL, A. H. Theory and application of the separable class of random processes. Tech. rep., Massachusetts Institute of Technology, Research Laboratory of Electronics, 1958.
- [8] PRADHAN, S. S., KUSUMA, J., AND RAMCHANDRAN, K. Distributed compression in a dense microsensor network. *IEEE Signal Processing Magazine* 19, 2 (2002), 51–60.
- [9] VINCENT, R. *Identification passive en acoustique : estimateurs et applications au SHM*. PhD thesis, Université Grenoble Alpes, 2016.
- [10] XIAO, J.-J., RIBEIRO, A., LUO, Z.-Q., AND GIANNAKIS, G. B. Distributed compression-estimation using wireless sensor networks. *IEEE Signal Processing Magazine* 23, 4 (2006), 27–41.
- [11] ZHAO, F., LIU, J., LIU, J., GUIBAS, L., AND REICH, J. Collaborative signal and information processing : an information-directed approach. *Proceedings of the IEEE* 91, 8 (2003), 1199–1209.



## Chapitre 2

# Traitement du signal dans le domaine compressé

### Sommaire

---

<b>2.1</b>	<b>Traitement numérique du signal</b>	<b>11</b>
<b>2.2</b>	<b>Echantillonnage</b>	<b>12</b>
<b>2.3</b>	<b>Acquisition compressive</b>	<b>12</b>
2.3.1	Acquisition compressive numérique	13
2.3.2	Reconstruction	16
2.3.3	Matrice $\Phi$ et son lien avec $\Psi$	18
2.3.4	Défis à relever et tendances de la recherche	19
<b>2.4</b>	<b>Traitement du signal compressé TSC</b>	<b>20</b>
2.4.1	Outils prédécesseurs : compression des signaux orientée vers les tâches	22
2.4.2	Quelques travaux récents	24
2.4.3	Produit interne et corrélation	25
	<b>Bibliographie</b>	<b>26</b>

---

## 2.1 Traitement numérique du signal

Plusieurs raisons se trouvent à l'origine du succès du traitement des signaux sous forme numérique, une qui est très importante est que les circuits numériques ont accès à de nombreux calculs qui ne sont pas réalisables de manière analogique. En plus, un même dispositif électronique peut être utilisé pour effectuer une tâche parmi plusieurs très variées, puis reprogrammé facilement pour réaliser une autre tâche complètement différente. Ce passage important au domaine numérique a été possible grâce aux travaux de Whittaker et Nyquist sur l'échantillonnage des signaux. Celui-ci est devenu un principe fondamental et a longtemps été considéré comme un besoin plutôt que comme une méthode de numérisation parmi plusieurs.

Néanmoins, certaines applications, dont celles en hautes fréquences, ont mis en évidence les limites de l'échantillonnage classique et d'autres approches d'acquisition des signaux ont vu le jour<sup>1</sup>. Par exemple, l'échantillonnage passe-bande [43], les approches par sous-espaces [42], ou plus récemment l'échantillonnage des signaux à taux fini d'innovation [44] et l'acquisition compressive [15, 6]. Toutes ces méthodes de numérisation ont réussi, dans certains cas, à représenter un signal avec un nombre de valeurs numériques

---

1. Cf [29] pour une complète vue d'ensemble.

inférieur à celui exigé par le critère de Nyquist. Comme mentionné dans l'introduction, nous nous focalisons ici uniquement sur le cas de l'acquisition compressive, qui est aujourd'hui un axe de recherche prometteur et très actif dans différentes communautés scientifiques.

Nous commencerons ce chapitre par expliquer les différences et le lien entre échantillonnage et acquisition compressive.

## 2.2 Echantillonnage

De manière générale, on parlera de numérisation d'un signal continu  $x(t)$  de durée finie. Ce qui signifie obtenir une représentation sur un nombre fini  $N$  de valeurs numériques  $x_1, \dots, x_N$ , que l'on écrira toujours sous forme vectorielle,  $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{R}^N$ .

Dans le cas particulier de la numérisation classique, l'échantillonnage des signaux, le lien entre  $\mathbf{x}$  et  $x(t)$  est trivial :  $\mathbf{x}$  contient les valeurs du signal analogique relevées à intervalles réguliers (toutes les  $T_e$  secondes) c'est à dire  $\mathbf{x} = [x(T_e), \dots, x(NT_e)]^T$ . Les contraintes sur la période d'échantillonnage  $T_e$  sont bien connues et données par le critère de Nyquist-Shannon :  $\frac{1}{T_e} = F_e \geq 2F_{max}$ . Ce paradigme est basé sur l'hypothèse que le contenu fréquentiel du signal est nul au delà de la fréquence  $F_{max}$  (ce qu'on appelle signal à bande limitée).

Après échantillonnage, la reconstruction (ou plutôt la synthèse) du signal continu  $x(t)$  à partir de  $\mathbf{x}$ , est relativement simple :

$$x(t) = \sum_n x_n \text{sinc}(\pi(t - nT_e)/T_e)$$

où la fonction sinc est définie pour tout  $t$  non nul par  $\text{sinc}(t) = \frac{\sin(t)}{t}$  et  $\text{sinc}(0) = 1$ . D'autre part, une interprétation de l'échantillonnage, détaillée dans [17, p. 106], consiste à écrire :

$$x_n = x(nT_e) = \frac{1}{T_e} \int_{-\infty}^{\infty} x(t) \text{sinc}(\pi(t - nT_e)/T_e) dt$$

C'est à dire que les composantes du vecteur  $\mathbf{x}$  s'interprètent comme les produits scalaires entre  $x(t)$  et les fonctions  $s_n(t) = \frac{1}{T_e} \text{sinc}(\pi(t - nT_e)/T_e)$  :

$$x_n = \langle s_n(t), x(t) \rangle$$

Notons que l'échantillonnage classique dans ce sens n'est qu'un cas particulier de numérisation :  $\mathbf{x}$  est constitué de  $N$  valeurs numériques  $x_n$ . Chaque  $x_n$  représente la valeur du signal à l'instant  $nT_e$  et quantifie également la similarité entre  $x(t)$  et le *signal de mesure*  $s_n(t)$ . Dans ce sens l'acquisition compressive est plus générale, l'idée étant de remplacer les fonctions orthogonales  $s_n(t)$  par d'autres signaux de mesure.

## 2.3 Acquisition compressive

Puisque ce paradigme permet d'aller au delà de l'échantillonnage de Nyquist, le signal numérisé, dénoté  $\mathbf{y}$ , n'est constitué que de  $M \ll N$  valeurs numériques. Ce sont les produits scalaires obtenus à partir du signal continu  $x(t)$  et  $M$  fonctions de mesure  $\phi_m(t)$  :

$$y_m = \langle \phi_m(t), x(t) \rangle$$

Cette idée est illustrée dans la figure 2.1.

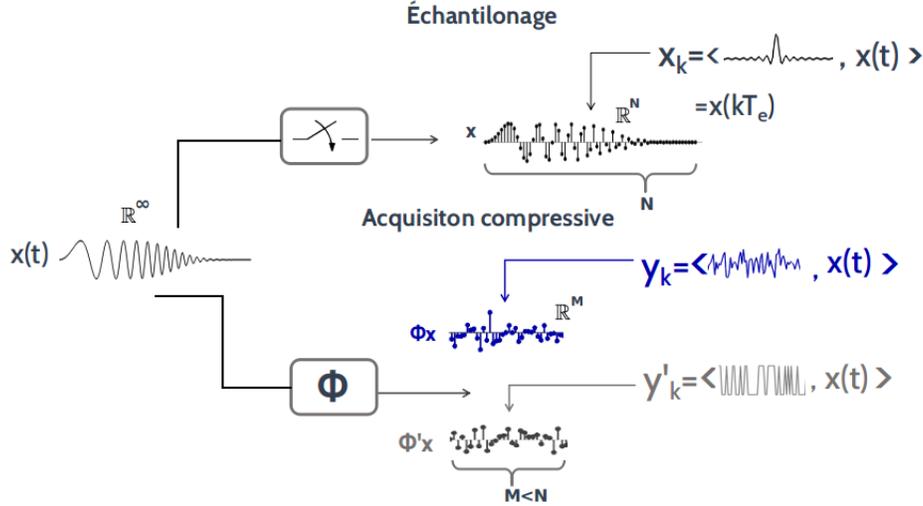


FIGURE 2.1 – L’acquisition compressive est une généralisation de l’échantillonnage classique. Au lieu de prendre directement la valeur des signaux analogiques (interprétable comme un produit interne entre le signal et des fonctions sinus cardinaux), la numérisation est faite sous la forme de produits internes entre le signal et des signaux de mesure  $\phi_m$  plus généraux [7].

La discussion sur les propriétés de ces fonctions est reportée à la section suivante de ce chapitre. Plusieurs travaux ont montré que cette mise en place de l’acquisition compressive est possible dès le domaine analogique : l’échantillonnage non-uniforme, la pré-intégration et modulation aléatoire [7, 3], le démodulateur aléatoire [22], le multiplexeur aléatoire [40], etc, en sont quelques exemples. Le but à long terme est de remplacer les convertisseurs analogiques numériques (CAN) actuels par une électronique différente, permettant d’acquérir  $\mathbf{y}$  directement à partir du signal analogique  $x(t)$ . Cette problématique à la frontière du traitement du signal et du design électronique, souvent nommé **conversion analogique à information** (CAI), est un sujet de recherche actif et reste encore un défi important à relever.

Conformément à l’état de l’art actuel en électronique, nous nous contentons de la mise en place de l’acquisition compressive dans le domaine numérique, qui est faite avec un CAN traditionnel, suivi des calculs pour effectuer la compression. On identifiera cette approche comme acquisition compressive numérique.

### 2.3.1 Acquisition compressive numérique

En pratique, il existe différentes manières de réaliser l’acquisition compressive. Nous utiliserons le terme *acquisition compressive numérique* pour parler du cas que nous avons déjà évoqué : le signal analogique  $x(t)$  est d’abord échantillonné pour construire  $\mathbf{x}$ , puis  $\mathbf{y}$  est obtenu à partir de calculs effectués dans le domaine numérique, c’est à dire les produits internes entre  $\mathbf{x}$  et les *signaux de mesure ou d’acquisition*  $\phi_m$ .

$$\langle \phi_m, \mathbf{x} \rangle = \mathbf{x}^T \phi_m = \sum_{i=1}^N \phi_{i,m} x_i = y_m \quad m = 1 \dots M$$

Les  $M$  valeurs numériques obtenues, sont regroupées dans le vecteur  $\mathbf{y}$ , appelé aussi représentation compressée de  $\mathbf{x}$ . Ce calcul peut s’écrire sous forme matricielle :  $\Phi \mathbf{x} = \mathbf{y}$  où  $\mathbf{x} \in \mathbb{R}^N$ ,  $\mathbf{y} \in \mathbb{R}^M$ ,  $\Phi \in \mathbb{R}^{M \times N}$  et  $M < N$ . Les  $M$  lignes de  $\Phi$  constituent les *vecteurs de mesure* auxquels  $\mathbf{x}$  est comparé. Cette approche est illustrée dans la figure 2.2.

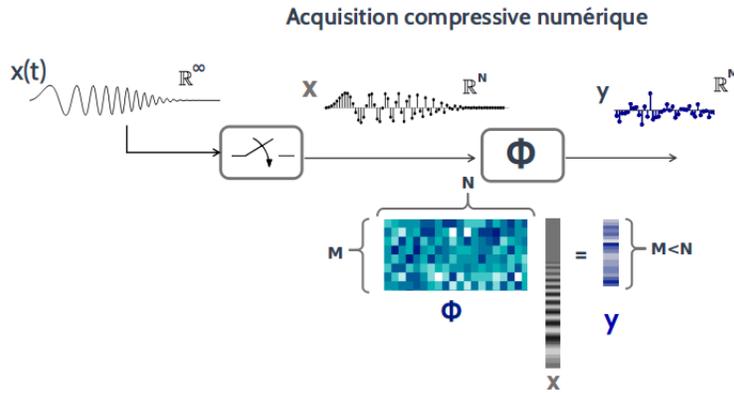


FIGURE 2.2 – Acquisition compressive numérique, les signaux analogiques sont d’abord échantillonnés puis réduits en dimension à l’aide d’une matrice de projection  $\Phi$  dans le domaine numérique.

$$\begin{bmatrix} \Phi \\ \phi_1^T \\ \vdots \\ \phi_M^T \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} \Phi x \\ \langle \phi_1, x \rangle \\ \vdots \\ \langle \phi_M, x \rangle \end{bmatrix} = \begin{bmatrix} y \\ y_1 \\ \vdots \\ y_M \end{bmatrix}$$

L’acquisition compressive s’intéresse essentiellement au problème de reconstruction, à savoir : retrouver le signal  $\mathbf{x}$ , ainsi acquis. Cela signifie résoudre le système linéaire d’équations  $\Phi \mathbf{x} = \mathbf{y}$  qui possède plus d’inconnues (les  $N$  valeurs de  $\mathbf{x}$ ) que d’équations ( $M$ , où  $M < N$ ). Ceci est un système indéterminé, connu pour posséder un nombre infini de solutions. Autrement dit, une infinité de vecteurs  $\mathbf{x} \in \mathbb{R}^N$  engendrent ce même vecteur  $\mathbf{y}$ , lorsqu’ils sont projetés sur  $\Phi$ . Cette problématique a une longue histoire dans le traitement du signal, par exemple dans les *problèmes inverses*. La formulation classique consiste à retrouver  $\mathbf{x}$  à partir des observation bruitées  $\mathbf{y}^b$  : c’est à dire  $\Phi \mathbf{x} + \mathbf{b} = \mathbf{y}^b$ .

En l’absence de bruit, une solution à ce problème s’écrit :

$$\underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x}) \text{ t.q. } \Phi \mathbf{x} = \mathbf{y}$$

En d’autres termes, il faut choisir une fonction  $f()$  capable d’imposer certaines propriétés à la solution, ce qui permet de trier parmi toute l’infinité des solutions possibles.

La plupart du temps, la fonction  $f()$  n’est autre que la norme  $L_2$  :  $\|\mathbf{x}\|_2$

$$\|\mathbf{x}\|_2 = \left( \sum_{i=1}^N |x_i|^2 \right)^{\frac{1}{2}}$$

Deux raisons de ce choix sont : (i) La solution que l’on retrouve possède une énergie minimale (un *a priori* raisonnable puisque les signaux physiques sont bornés). (ii) La solution s’écrit explicitement et peut se calculer directement :  $\hat{\mathbf{x}} = \Phi^T (\Phi \Phi^T)^{-1} \mathbf{y}$ . Nous soulignons que récupérer **exactement** le signal  $\mathbf{x}_0$  n’est possible que si le vecteur projeté sur  $\Phi$  était à l’origine celui d’énergie minimale. Néanmoins, dans la plupart des cas la solution retrouvée est uniquement **approximative**.

Ce qui est remarquable et étonnant dans l'acquisition compressive, c'est le fait de parvenir, sous certaines conditions sur  $\Phi$ , à résoudre ce système et ainsi **retrouver  $\mathbf{x}$  exactement**. Pour y arriver, l'hypothèse de parcimonie sur  $\mathbf{x}$  est une condition suffisante. Autrement dit, ce vecteur doit être constitué majoritairement de zéros, sauf quelques éléments. De manière générale, tout signal  $\mathbf{s} \in \mathbb{R}^N$  possédant uniquement  $K \ll N$  valeurs différentes de zéro est appelé *K-parcimonieux*. Les premiers travaux [6, 15], ont établi les contraintes sur  $\Phi$  :

Une telle matrice doit satisfaire la propriété d'isométrie restreinte (RIP *Restricted Isometry Property*) :

Soit une matrice  $\Phi \in \mathbb{R}^{N \times M}$  avec  $N > M$ , s'il existe un  $\delta > 0$  tel que, pour tout vecteur K-parcimonieux  $\mathbf{x}_K \in \mathbb{R}^N$  avec  $K < N$  :

$$(1 - \delta) \|\mathbf{x}_K\|_2 \leq \|\Phi \mathbf{x}_K\|_2 \leq (1 + \delta) \|\mathbf{x}_K\|_2$$

Alors, cette matrice satisfait la propriété d'isométrie restreinte avec une constante  $\delta$ . En clair, cette condition signifie que la norme  $L_2$  de la projection des vecteurs K-parcimonieux sur  $\Phi$  est quasiment préservée, c'est à dire  $\|\Phi \mathbf{x}_K\|_2$  ne change pas au-delà d'un facteur  $\delta$  par rapport à la norme d'origine  $\|\mathbf{x}_K\|_2$ . Cette propriété s'écrit facilement, mais la prouver pour une matrice donnée est très difficile. Heureusement, les matrices aléatoires dont les éléments sont tirés au hasard suivant certaines lois de probabilité (par exemple toute loi sous-gaussienne) satisfont la RIP avec une forte probabilité si le nombre de lignes de  $\Phi$  est de l'ordre de  $M = \mathcal{O}(K \log(N/K))$ .

Avant d'aller plus loin, il est important de souligner que la parcimonie **ou sparsité** est un *a priori* puissant. Citons Candès[7] : « elle exprime l'idée que le taux d'information d'un signal en temps continu peut être plus petit que celui suggéré par sa largeur de bande, ou qu'un signal en temps discret dépend d'un nombre de degrés de liberté très petit par rapport à sa longueur (finie) ». Bien que la plupart des signaux naturels soient non-parcimonieux dans le domaine temporel, très souvent ils sont bien approchés, après transformation, par des signaux parcimonieux (ces signaux sont alors dits parcimonieux dans le domaine transformé). Cette transformation consiste à exprimer  $\mathbf{x}$  en termes d'une base orthonormale  $\Psi \in \mathbb{R}^{N \times N}$  constituée de  $N$  signaux  $\Psi = [\psi_1 \psi_2 \dots \psi_N]$ . Dans le codage par transformation, tout d'abord l'analyse de  $\mathbf{x}$ , ou sa décomposition selon les colonnes de  $\Psi$ , est effectuée :

$$\boldsymbol{\alpha} = \begin{bmatrix} - & \psi_1^T & - \\ & \vdots & \\ - & \psi_M^T & - \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} \langle \psi_1, \mathbf{x} \rangle \\ \vdots \\ \langle \psi_M, \mathbf{x} \rangle \end{bmatrix}$$

Ensuite, une approximation  $\hat{\mathbf{x}}$  est synthétisée en ne retenant que les  $k \ll N$  coefficients les plus significatifs.

$$\mathbf{x} = \Psi \boldsymbol{\alpha} = \sum_{i=1}^N \alpha_i \psi_i \approx \hat{\mathbf{x}} = \sum_{i=1}^k \alpha_{I(i)} \psi_{I(i)} = \Psi \mathbf{s}$$

où  $I$  indexe les coefficients triés par module décroissant. Ce vecteur  $\mathbf{s}$  est parcimonieux et possède  $k$  éléments non nuls. Voici le principe sous-jacent à la compression de données classique (ou codage par transformation) et la motivation derrière l'acquisition compressive, cette fois-ci exprimé par Donoho[15] : « *Pourquoi faire tant d'efforts pour acquérir toutes les données lorsque la plupart de ce que nous obtenons sera jeté ? Ne pouvons-nous pas simplement mesurer directement la partie qui ne finira pas par être jetée ?* ».

Passons maintenant à la question centrale en acquisition compressive, comment reconstruire les signaux ainsi acquis ?

### 2.3.2 Reconstruction

Comme évoqué précédemment, la recherche du vecteur le plus parcimonieux  $\mathbf{x}_K$  parmi toutes les solutions possibles de  $\Phi \mathbf{x} = \mathbf{y}$  en absence de bruit s'écrit :

$$\operatorname{argmin}_{\mathbf{x}} f(\mathbf{x}) \text{ t.q. } \Phi \mathbf{x} = \mathbf{y}$$

Ici  $f()$  doit favoriser les solutions parcimonieuses. Deux cas importants dans ce contexte sont la norme  $L_1$  et les quasi normes  $L_p (0 \leq p < 1)$ , définies comme :

$$\|\mathbf{x}\|_p = \left( \sum_{i=1}^N |x_i|^p \right)^{\frac{1}{p}} \text{ et } \|\mathbf{x}\|_0 := \operatorname{card}\{i : x_i \neq 0\}$$

Cette dernière compte le nombre d'éléments de  $\mathbf{x}$  qui diffèrent de zéro.

La solution exacte correspond à l'utilisation de la norme  $L_0$  pour fonction  $f$  :

$$(P0) \quad \operatorname{argmin}_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ t.q. } \Phi \mathbf{x} = \mathbf{y}$$

Malheureusement, ce problème, appelé souvent  $P0$ , est de type combinatoire et très difficile à résoudre, par exemple par recherche exhaustive (ou force brute), c'est à dire en examinant toutes les solutions possibles. Ce problème est dit **NP-difficile** : sa complexité, ou le nombre d'opérations dont on a besoin pour le résoudre, est une fonction **Non Polynômiale** de  $N$  (la taille des signaux à reconstruire).

Il est possible, néanmoins de formuler un problème équivalent à  $(P0)$ , mais plus facile à résoudre et dont la solution soit aussi le vecteur le plus parcimonieux :  $\mathbf{x}_K$ . À ses débuts, l'acquisition compressive s'est appuyée sur l'exploitation de la norme  $L_1$  :

$$(P1) \quad \operatorname{argmin}_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ t.q. } \Phi \mathbf{x} = \mathbf{y}$$

Ce problème est convexe, c'est à dire qu'il n'a qu'une solution, puisque cette fonction  $f()$  ne présente qu'un minimum, nécessairement global. Sous certaines conditions sur le système [16],  $(P1)$  est équivalent à  $(P0)$ . Souvent, la littérature de l'acquisition compressive exige la RIP sur la matrice  $\Phi$ , ce qui constitue une condition suffisante pour assurer cette équivalence.

Contrairement au cas de la minimisation de la norme  $L_2$ ,  $(P1)$  n'a pas de solution explicite. Elle doit être recherchée à l'aide d'algorithmes itératifs. Cette approche est néanmoins exploitable puisque, contrairement à  $(P0)$ ,  $(P1)$  n'est pas NP mais **P-difficile**, c'est à dire que sa complexité est une fonction **Polynômiale** de  $N$ , généralement  $\mathcal{O}(N^3)$ .

Afin d'illustrer un peu cette discussion, la figure 2.3 reprend l'interprétation géométrique souvent présentée pour expliquer pourquoi la norme  $L_1$  fournit la bonne réponse mais pas la norme  $L_2$ . Dans la recherche d'un vecteur parcimonieux  $\mathbf{x}_K$  sur  $\mathbb{R}^2$ , minimiser la norme  $L_p$  est équivalent à augmenter progressivement la taille d'une boule  $L_p$  (le sous-ensemble de points ayant la même norme  $L_p$ ) jusqu'à atteindre un point appartenant aux solutions possibles (la ligne  $\Phi \mathbf{x}$ ). La boule  $L_1$  (possédant une forme de diamant) est plus pointue selon ses axes que la boule  $L_2$  (circulaire), ce qui permet d'atteindre les solutions parcimonieuses, souvent positionnées sur les axes. Bien que sur  $\mathbb{R}^2$  les deux solutions semblent proches, sur  $\mathbb{R}^N$ , ce n'est plus le cas. Un exemple de ce qui arrive dans  $\mathbb{R}^{100}$  est présenté à droite de cette figure.

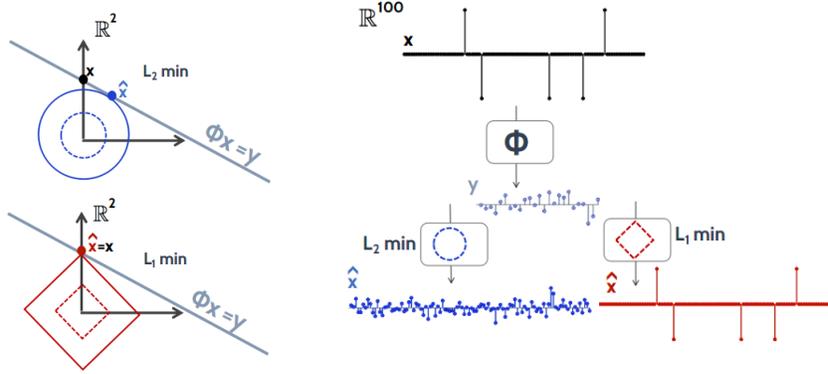


FIGURE 2.3 – Gauche : la minimisation de la norme  $L_1$  favorise les solutions sur les axes, alors que la norme  $L_2$  produit la solution d'énergie minimale, qui n'est pas parcimonieuse. Droite : reconstruction d'un signal parcimonieux par minimisation de normes. Seule la norme  $L_1$  fournit des signaux parcimonieux.

Le tableau suivant présente un bilan de cette problématique :

$f(x)$	solution	caractéristique	complexité
$\ x\ _2$	$\hat{x} = \Phi^T (\Phi \Phi^T)^{-1} y$	énergie minimale : $\sum_i  x_i ^2$	P
$\ x\ _1$	algorithmes itératifs	minimum $\sum_i  x_i $	P
$\ x\ _0$	algorithmes combinatoires	parcimonie	NP

Or, si la norme  $L_1$ , qui approche la norme  $L_0$ , privilégie les solutions parcimonieuses, pourquoi ne pas étudier les (pseudo) normes  $L_p$  où  $0 < p < 1$ ? En effet, ces dernières possédant des boules plus pointues que la norme  $L_1$ , elles approchent mieux la norme  $L_0$ . Plusieurs travaux s'y sont intéressés. Dans la littérature on trouve des résultats empiriques qui montrent qu'il est en effet possible de surpasser la performance de la norme  $L_1$  avec des algorithmes itératifs qui minimisent ces pseudo-normes [18, 8]. Malheureusement ces problèmes ne sont pas convexes et n'ont pas de solution unique : les critères  $f()$  à minimiser possèdent plusieurs minima locaux. Ces algorithmes itératifs dépendent donc fortement d'un bon choix du point d'initialisation. Le lecteur intéressé trouvera une discussion sur ce problème dans [47]. La figure 2.4 illustre cet aspect.

### Inexactitudes du modèle d'acquisition

Le modèle d'acquisition présenté précédemment montre que l'on peut reconstruire exactement des signaux parcimonieux. Cependant, ce modèle est très idéaliste. En particulier, il peut y avoir du bruit sur les mesures ou bien les signaux naturels ne sont en réalité qu'approximativement parcimonieux. Pour prendre cela en compte, le problème à résoudre est modifié. Il s'écrit :

$$\operatorname{argmin}_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{t.q.} \quad \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 < \lambda \quad (\lambda > 0)$$

ou bien :

$$\operatorname{argmin}_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

Ce problème est connu sous le nom de poursuite de base et débruitage (*Basis Pursuit Denoising*). Les tous premiers travaux en acquisition compressive s'appuient sur des méthodes d'optimisation convexe. Néanmoins celles-ci ont été petit à petit remplacées

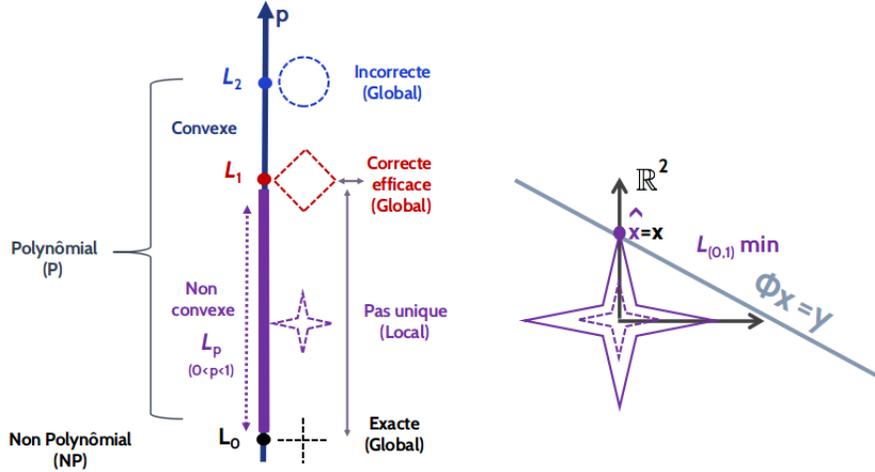


FIGURE 2.4 – Gauche : Les différentes boules  $L_p$  tracées sur l’axe  $p$ , la convexité du problème et type de solution. Droite : Une boule  $L_p$  ( $0 < p < 1$ ) approche mieux la norme  $L_0$  que la norme  $L_1$  et induise ainsi des solutions parcimonieuses mais le problème est non-convexe, il n’a donc pas de solution unique.

par des algorithmes itératifs gloutons, qui s’avèrent plus performants lorsque la taille des signaux est grande [27]. Le lecteur intéressé trouvera une vue d’ensemble des algorithmes de reconstruction dans cette dernière référence. Ces inexactitudes du modèle font que la solution que l’on trouve en pratique n’est pas exacte. Néanmoins dans la plupart des cas elle est satisfaisante. Par exemple, en présence d’un bruit  $\mathbf{b}$  borné :  $\Phi \mathbf{x} + \mathbf{b} = \mathbf{y}$  l’erreur de reconstruction est elle aussi bornée [2] :

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq C \|\mathbf{b}\|_2$$

Ou bien, lorsque les signaux ne sont pas exactement  $K$ -parcimonieux mais bien approchés par ceux-ci, l’erreur de reconstruction en présence d’un bruit borné s’écrit : [2].

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq C_1 \|\mathbf{x} - \mathbf{x}_K\|_2 + C_2 \frac{1}{\sqrt{K}} \|\mathbf{x} - \mathbf{x}_K\|_1 + C_3 \|\mathbf{b}\|_2$$

Ici  $\mathbf{x}_K$  est le signal  $K$ -parcimonieux qui approche au mieux  $\mathbf{x}$ .

### 2.3.3 Matrice $\Phi$ et son lien avec $\Psi$

Il existe différentes méthodes de construction des matrices d’acquisition, aléatoires ou déterministes. Dans le cas de l’acquisition compressée numérique, projeter  $\mathbf{x}$  sur  $\Phi$  de manière directe implique le calcul de  $M$  produits internes avec des vecteurs à  $N$  termes, c’est à dire des calculs de l’ordre de  $\mathcal{O}(MN)$  et un stockage semblable. Pour cette raison, les matrices denses ne sont pas optimales du point de vue pratique. Les améliorations qui existent sont de différents types : dans le cas de matrices aléatoires, on peut (i) limiter le nombre de valeurs possibles, par exemple avec une distribution binaire ou ternaire, ou bien (ii) faire apparaître des zéros dans  $\Phi$  et ainsi réduire les calculs. Dans le cas de matrices déterministes, on a des constructions structurées, ce qui permet de stocker moins de données et d’exploiter des algorithmes rapides de multiplication matrice vecteur.

Le seul aspect qui peut limiter le choix de la matrice d’acquisition  $\Phi$  est la nature des signaux à traiter. Ce point nécessite l’introduction de la notion de cohérence.

### Lien avec $\Psi$

La cohérence établit le lien entre la matrice de compression  $\Phi$  et le dictionnaire  $\Psi$  sur lequel le signal présente une décomposition parcimonieuse

$$\mu(\Phi, \Psi) = \max_{1 \leq i, j \leq N} \sqrt{N} |\langle \phi_i, \psi_j \rangle|$$

C'est le maximum des produits scalaires entre les lignes de  $\Phi$  et les colonnes de  $\Psi$ . On montre que cette quantité est bornée

$$1 \leq \mu(\Phi, \Psi) \leq \sqrt{N}$$

En clair, elle quantifie la similarité entre les signaux de mesure et la base sur laquelle les signaux à traiter sont parcimonieux. Moins ces deux ensembles sont cohérents, mieux  $\Phi$  capture l'information des signaux que l'on peut écrire de manière parcimonieuse sur  $\Psi$ . D'après Candès, qui a introduit cette notion [7] : « L'incohérence étend la dualité temps-fréquence et exprime l'idée que les signaux qui ont une représentation parcimonieuse dans une base  $\Psi$  doivent être étalés dans le domaine dans lequel ils sont acquis [...] autrement dit, l'incohérence signifie que, contrairement au signal d'intérêt, les fonctions de mesure ont une représentation très dense (non-parcimonieuse) dans  $\Psi$  ».

Formulé autrement, on montre que pour un signal parcimonieux sur une base  $\Psi$ , acquis avec une matrice  $\Phi$ , le nombre minimum de mesures pour le reconstruire est :

$$M \geq C \mu^2(\Phi, \Psi) K \log N$$

Évidemment, on souhaite  $\mu^2(\Phi, \Psi)$  le plus faible possible, ce qui arrive lorsque  $\Psi$  et  $\Phi$  sont incohérentes. Nous privilégierons les constructions aléatoires des matrices  $\Phi$  dans ces travaux parce qu'elles sont incohérentes avec quasiment toute base  $\Psi$ , ce qui permet d'obtenir une acquisition la plus générale possible.

### 2.3.4 Défis à relever et tendances de la recherche

L'acquisition compressée a évolué rapidement depuis qu'elle est apparue. Les défis à relever aujourd'hui se focalisent surtout sur la reconstruction des signaux et notamment dans la performance et la précision. Trois objectifs typiques dans différents contextes sont :

- i) Réduire la complexité lors de la reconstruction, inférieure à  $\mathcal{O}(N^3)$
- ii) Réduire la taille des mesures  $M \leq \mathcal{O}(K \log(K/N))$
- iii) Améliorer la précision de la reconstruction.

Pour nous, la recherche dans le CS peut être classée dans trois catégories différentes :

- La première est **l'amélioration des algorithmes de reconstruction**. Elle concerne typiquement les objectifs i) et ii). Par exemple, l'optimisation non convexe, mentionnée précédemment, permet de reconstruire des signaux à partir d'un nombre de mesures inférieur au nombre minimal exigé dans la minimisation de la norme  $L_1$ . Mais ceci n'est qu'une approche possible parmi d'autres. Le lecteur intéressé trouvera une vue d'ensemble sur l'état de l'art dans [27].
- La deuxième est **la généralisation du modèle** au delà des seuls signaux parcimonieux, ce qu'on appelle acquisition compressée généralisée (Generalized Compressed Sensing). Cette approche cherche des réponses aux défis ii) et iii). Souvent, l'hypothèse de parcimonie ne fournit pas la précision souhaitée. Des a priori plus sophistiqués que la seule parcimonie, qui exploitent une information supplémentaire sur le signal (par exemple la parcimonie par bloc ou les arbres

d'ondelettes [2]) permettent une reconstruction avec  $M = \mathcal{O}(K)$  au lieu de  $\mathcal{O}(K \log(K/N))$  mesures. Plus récemment [4] a étudié d'autres modèles, comme la co-parcimonie ou l'union de sous-espaces, avec lesquels on peut obtenir une erreur de reconstruction inférieure, ou bien encore une réduction du nombre minimal de mesures pour reconstruire. Cependant, un compromis important doit être géré avec ces modèles : le prix à payer pour ces avantages est une augmentation de la complexité des algorithmes de reconstruction.

- La troisième, plus récente, est inspirée par la recherche en machine learning et notamment en apprentissage profond. Il s'agit de **l'introduction de modèles et d'algorithmes guidés par les données**, qui promet de répondre aux questions i) ii) et iii). Voir par exemple [31, 30]. Cette approche s'intéresse à la similarité entre le problème de reconstruction des signaux et celui des *autoencoders* dans le contexte des réseaux de neurones. Une perspective historique et un bilan de cette approche sont disponibles dans [28]. De la même manière que l'apprentissage profond, cette approche reste encore très peu comprise et en attente de développements théoriques solides qui expliquent les premiers résultats empiriques encourageants.

Malgré tous ces efforts, le problème de réduction de la complexité algorithmique dans la reconstruction reste encore un défi à relever. Dans certains cas, par exemple lorsque les contextes d'application sont contraints en termes de matériel ou d'énergie, le coût de reconstruction des signaux est trop important pour utiliser l'acquisition compressée telle qu'elle a été présentée. Naturellement, la communauté concentre jusqu'à aujourd'hui quasiment tous ses efforts sur ce problème de reconstruction. Cependant, dans beaucoup d'applications, retrouver le signal n'est pas le but principal, mais plutôt extraire uniquement certaines informations (ou bien les supprimer). En plus : « Reconstruire les signaux, en augmentant le nombre de données pour ensuite exploiter les techniques de traitement du signal classique peut s'avérer non optimal, ni en termes de précision ni en termes d'efficacité »[11]. Alors, *pourquoi ne pas profiter de cette réduction de dimensionnalité pour procéder directement au traitement ?*

Une approche qui aborde cette question s'est développée en parallèle à l'acquisition compressée classique. On la nomme traitement compressif du signal (CSP, Compressive Signal Processing) [11], ou bien *traitement des signaux dans le domaine compressé ou encore traitement du signal compressé*.

## 2.4 Traitement du signal compressé TSC

En acquisition compressive, le plus souvent, on restreint la matrice d'acquisition  $\Phi$ , qui doit satisfaire la RIP, c'est à dire garantir la quasi-préservation de la norme  $L_2$  des signaux parcimonieux en basse dimension. C'est le cas d'un très grand nombre de matrices aléatoires.

De manière intéressante, ces matrices aléatoires sont connues dans différents domaines pour satisfaire une autre propriété, plus générale que la RIP, car elle ne se limite pas aux signaux parcimonieux : le lemme de Johnson-Linderaus(JLL) [20].

Datant des années 80, ce lemme garantit l'existence d'applications linéaires réduisant les dimensions d'un sous-ensemble quelconque de vecteurs, capables de préserver quasiment la distance euclidienne entre eux :

$$(1 - \delta) \|\mathbf{x}_1 - \mathbf{x}_2\|_2 \leq \|\Phi \mathbf{x}_1 - \Phi \mathbf{x}_2\|_2 \leq (1 + \delta) \|\mathbf{x}_1 - \mathbf{x}_2\|_2 \quad (2.1)$$

En d'autres termes, les projections des vecteurs  $\mathbf{x}_1$  et  $\mathbf{x}_2$  dans une matrice  $\Phi : \mathbb{R}^N \mapsto \mathbb{R}^M$ , respectivement  $\Phi\mathbf{x}_1$  et  $\Phi\mathbf{x}_2$ , sont séparées d'une distance euclidienne qui diffère, au maximum d'un facteur delta, de celle des vecteurs dans l'espace d'origine. Selon le contexte, ce

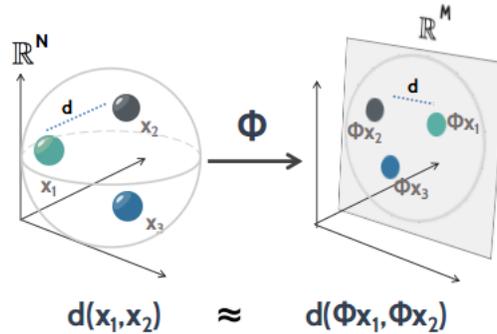


FIGURE 2.5 – Lemme de Johnson et Lindenstrauss, garantie de l'existence d'une transformation linéaire  $\Phi : \mathbb{R}^N \mapsto \mathbb{R}^M$  qui réduit la dimensionnalité tout en préservant la distance euclidienne des vecteurs de grande dimension.

lemme peut aussi s'exprimer comme la conservation approximative et contrôlée des produits internes, après la projection des vecteurs dans des espaces de dimension inférieure. Plus précisément, une matrice aléatoire  $\Phi$  est une transformation JL avec paramètres  $\delta, \epsilon, n$  [46] si, avec une probabilité au moins  $\epsilon$ , tout sous-ensemble de  $n$  points  $V \subset \mathbb{R}^N$  vérifie :

$$|\langle \Phi\mathbf{x}_1, \Phi\mathbf{x}_2 \rangle - \langle \mathbf{x}_1, \mathbf{x}_2 \rangle|^2 \leq \delta \|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 \text{ pour tout } (\mathbf{x}_1, \mathbf{x}_2) \in V^2.$$

Cette préservation de géométrie, après réduction de dimension, permet l'exploitation des algorithmes de détection, classification et estimation directement sur les signaux sous forme compressée, c'est à dire en dimension réduite  $\mathbb{R}^M (M \ll N)$ . Cf fig. 2.6. Les travaux de Davenport et al.[11] ont souligné la bivalence des matrices JLL en CS. D'un côté en permettant la reconstruction des signaux parcimonieux (parce que JLL  $\implies$  RIP) et de l'autre, en donnant la possibilité, si nécessaire, de réaliser un traitement efficace en dimension réduite, sans passer par une coûteuse étape de reconstruction. Des années plus tard, Kramer et Ward ont apporté une autre contribution importante, en établissant une procédure pour aller dans le sens inverse, c'est à dire construire des matrices JLL à partir des matrices RIP [23], ouvrant ainsi la porte à l'exploitation des matrices RIP développées dans le cadre du CS classique, pour faire des traitements dans le domaine compressé.

Les avantages de ce point de vue ont été clairement soulignés [11] :

- Un même protocole d'acquisition reste utile pour de nombreux signaux/tâches de traitement.
- La perte en performances par rapport aux solutions *ad hoc* peut être justifiée par le gain en simplicité, universalité et flexibilité.

Dans ces premiers travaux figure une discussion sur la performance de quatre tâches canoniques : détection, estimation, classification et filtrage. Néanmoins, l'analyse est limitée à l'étude des signaux déterministes.

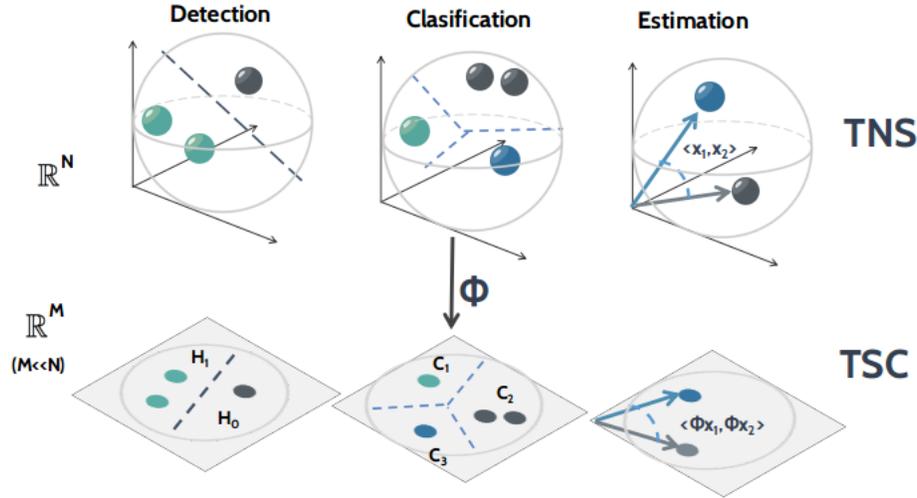


FIGURE 2.6 – Traitement des signaux dans le domaine compressé. La réduction de dimension préserve la géométrie et, en conséquence, la capacité à effectuer la détection, la classification et l'estimation.

Ces travaux ont introduit ces idées en traitement du signal, mais il nous semble pertinent de souligner que JLL a une histoire plus longue dans d'autres domaines, notamment en informatique, dans la communauté de *Data Streaming Algorithms*, connus en français sous le nom d'*algorithmes de fouille de flots de données*, cf par exemple [32]. La problématique est similaire : en l'absence des ressources informatiques nécessaires pour fournir une réponse exacte, on réduit la dimension des données pour obtenir un résultat approximatif, tout en contrôlant la perte en précision. Cependant, une différence importante existe : dans ce contexte, la réduction de dimension est souvent ajustée au calcul d'intérêt [11], ce qui n'est pas forcément souhaitable en traitement du signal, car plusieurs calculs peuvent être désirés, y compris la reconstruction des signaux.

Par ailleurs, une approche du traitement des signaux compressés existait bien avant dans la communauté d'*apprentissage automatique*, qui s'intéresse aussi très souvent à la réduction de dimension. Des techniques très proches, connues sous le nom de *projections aléatoires* (Random Projections) sont apparues avant [1] le TSC et montraient déjà les avantages de cette approche, ainsi que des méthodes pour réduire le nombre de calculs, par exemple [26].

Ensuite, avant de présenter ce qui a été fait dans le cadre des traitements dans le domaine compressé, on présentera certains outils développés auparavant en traitement du signal avec leurs motivations. Cela nous semble pertinent pour donner au lecteur une perspective plus générale.

#### 2.4.1 Outils prédécesseurs : compression des signaux orientée vers les tâches

Comme mentionné précédemment, la réduction de la taille des signaux, que l'on pourrait appeler compression classique, ignore souvent que ceux-ci sont l'objet de traitements dont la performance dépend de différents facteurs, parmi lesquels on trouve le rapport signal sur bruit (RSB). Cependant, la compression classique s'intéresse uniquement à ce dernier et s'appuie sur le critère de l'erreur de reconstruction, ou de décompression (c'est à dire  $e = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$ ) pour évaluer la qualité de la compression.

Nous parlerons de *compression des signaux orientée vers les tâches*, pour identifier les travaux qui utilisent plutôt la performance des traitements pour guider la réduction de la taille des signaux. Nous les classons en deux groupes :

#### Compression *ad hoc*.

Le principe est justement de maximiser la performance d'une ou plusieurs tâches particulières de traitement. Un exemple notable est la compression adaptée à l'estimation de paramètres, déjà mentionnée dans l'introduction [10, 9], qui s'intéresse à la préservation de l'information de Fisher (proportionnelle à la précision de l'estimation du paramètre). Plus récemment, inspirée aussi par la réduction de dimension par projection aléatoire, une approche bayésienne, appelée transformation préservatrice d'information a été développée [41]. Le lecteur intéressé y trouvera également de nombreuses références sur la réduction de dimension *ad hoc*.

Nous ne rentrerons pas dans le détail dans ce cas car, même si ces approches sont performantes dans des cas particuliers, elles ont par nature des difficultés lorsque plusieurs traitements sont nécessaires. Nous nous intéressons plutôt au cadre le plus général possible.

#### Réduction ou sélection aléatoire de données.

Les travaux sur cette ligne de recherche sont plus proches des traitements dans le domaine compressé et partagent nos motivations : des ressources énergétiques limitées et le besoin de les économiser. L'idée est simple : supprimer une partie des données disponibles.

Naturellement, il est possible de choisir les données à supprimer de manière optimale (suppression *ad hoc*), mais cela exige des a priori sur le signal et le bruit [39]. En l'absence d'information fiable sur les signaux, la sélection aléatoire est alors un bon choix et très simple à mettre en œuvre.

Un premier exemple de cette famille d'outils est l'échantillonnage aléatoire en temps discret [38]. L'objectif est double : réduction de l'impact du repliement spectral et diminution du nombre d'opérations lors d'un filtrage numérique ultérieur. Inspiré par les techniques d'échantillonnage aléatoire analogique, mais effectué dans le domaine numérique, il s'agit d'une conversion analogique numérique classique, suivie d'un sous-échantillonnage aléatoire. Le vecteur sous-échantillonné  $\mathbf{x}_g$  s'écrit :

$$\mathbf{x}_g = \mathbf{x} \odot \mathbf{g}$$

Ici  $\odot$  représente le produit terme à terme. Le signal  $\mathbf{x}$  est multiplié par un processus aléatoire binaire, dit d'échantillonnage,  $\mathbf{g} \in \{0, 1\}^N$ , qui est indépendant du signal et qui possède en moyenne uniquement  $M < N$  éléments non nuls.

Tout filtrage sur ce signal *compressé*, ne nécessite plus que  $M$  opérations au lieu de  $N$  si on avait filtré le signal initial. Le processus échantillonnage  $\mathbf{g}$  a été étudié dans deux cas, blanc[38] et coloré[13, 14].

Un autre exemple de cette famille de méthodes est la réduction sélective de données dans le cadre d'un problème de détection [39].

Les données fournies au détecteur sont sélectionnées de façon aléatoire, selon le modèle :

$$\mathbf{x}_g = \mathbf{G}\mathbf{x}$$

Ici,  $\mathbf{G}$  est une matrice  $M \times N$  construite aléatoirement, qui joue le rôle de réducteur de dimension. Elle possède un seul élément non nul par ligne. La sélection aléatoire se fait sans remise si bien qu'il y a au plus un élément non nul par colonne. Les résultats de [39] établissent que échantillonnage aléatoire est plus performante que le sous-échantillonnage simple, parce que repliement de spectre est moins important. Passons finalement à ce

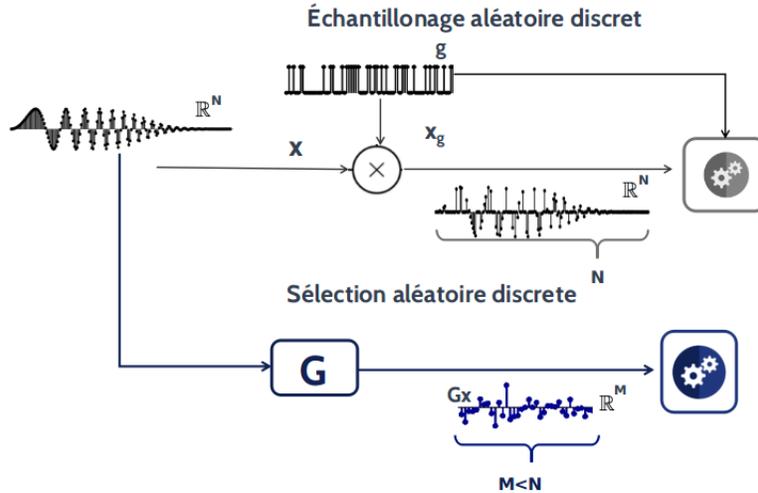


FIGURE 2.7 – Haut : le sous-échantillonnage aléatoire discret [38, 14] peut être considéré comme une réduction de dimension effective et non pas explicite, car le signal en sortie a un nombre très important d'éléments nuls. Néanmoins, la longueur des vecteurs et la grille d'échantillonnage restent fixes. Bas : la sélection aléatoire des échantillons est une réduction effective [39].

qui a été fait récemment concernant les traitements dans le domaine compressé.

## 2.4.2 Quelques travaux récents

Le traitement des signaux compressés s'est développé dans différentes directions, il nous semble important de les mentionner.

Premièrement, concernant l'estimation paramétrique dans le domaine compressé, des résultats intéressants existent sur la perte d'information en termes de borne de Cramer-Rao et Ziv-Zakai [35] [34] et sur l'effet de seuil [33]. Ces travaux montrent que l'estimation des paramètres à partir des signaux compressés reste possible, le coût associé au passage dans le domaine compressé se traduisant par une simple pénalité d'un facteur  $\sqrt{\frac{M}{N}}$  dans le RSB.

Plusieurs autres travaux méritent d'être mentionnés, qui ont étendu cette approche aux signaux aléatoires. Notamment dans le cas de la détection [24] et de la classification [36, 45]. Dans la deuxième référence on montre que d'autres mesures, comme la divergence de Kullback-Leibler et les distances de Chernoff, sont également préservées par projection aléatoire et demeurent ainsi exploitables.

Signalons également l'existence d'autres travaux qui portent sur le TSC, certains s'intéressent à l'estimation des propriétés à l'ordre deux des signaux aléatoires : la corrélation [37] et la densité spectrale de puissance [25]. D'autres portent sur l'apprentissage statistique compressif : par exemple pour l'estimation de mélange de Gaussiennes

[4, 5], et le partitionnement en k-moyennes [21]. Ce type d'algorithmes sont basés sur des méthodes d'optimisation et trop complexes pour être envisagé dans le cadre qui nous intéresse.

### 2.4.3 Produit interne et corrélation

Les traitements des signaux qui nous intéressent s'appuient tous sur l'estimation de la fonction d'intercorrélation entre les signaux enregistrés par deux capteurs différents. Cette estimation peut s'interpréter en termes de produits internes. Nous nous appuyons alors sur la préservation du produit interne dans le domaine compressé pour étudier la performance du calcul effectué directement avec des signaux issus d'une acquisition compressée. Des résultats sur la préservation du produit interne existent uniquement dans le cas des signaux déterministes. En particulier, [11] a abordé le problème d'estimation du produit interne  $\langle \mathbf{s}, \mathbf{l} \rangle$ ,  $\mathbf{l}$  étant fixé et  $\mathbf{s}$  un signal déterministe que l'on ne connaît qu'à travers ses projections  $\mathbf{y} = \Phi \mathbf{s}$ . On montre qu'il existe une certaine perte de précision et que la performance qu'on peut attendre pour cet estimateur dépend du  $\delta$  associé à la matrice de compression  $\Phi$  utilisée :

$$|\langle \Phi \mathbf{l}, \Phi \mathbf{s} \rangle - \langle \mathbf{l}, \mathbf{s} \rangle| \leq 2\delta \|\mathbf{l}\|_2 \|\mathbf{s}\|_2$$

Dans notre cas, on considère les deux vecteurs inconnus et aléatoires.

La fonction d'intercorrélation entre deux signaux  $\mathbf{x}$  et  $\mathbf{y}_\tau$  est estimée à l'aide de leur simple produit scalaire. Ces vecteurs sont constitués d'échantillons successifs des signaux :

$$\mathbf{x} = [x(t), \dots, x(t + T_e(N - 1))]^\top, \mathbf{y}_\tau = [y(t + T_e\tau), \dots, y(t + T_e(\tau + N - 1))]^\top$$

$T_e$  est la période échantillonnage.

Pour alléger l'écriture, on notera dans le reste du manuscrit :

$$\mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

#### Corrélateur classique

$$\mathcal{C}_N(\tau) = \langle \mathbf{x}, \mathbf{y}_\tau \rangle = \frac{1}{N} \mathbf{x}^\top \mathbf{y}_\tau \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N$$

$$\mathbf{x} = [x_1, \dots, x_N]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Dans le domaine compressé l'estimation de corrélation s'écrit :

#### Corrélateur compressé

$$C_N(\tau) = \langle \Phi \mathbf{x}, \Phi \mathbf{y}_\tau \rangle = (\Phi \mathbf{x})^\top (\Phi \mathbf{y}_\tau) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

$$\mathbf{x} = [x_1, \dots, x_N]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Dans le chapitre 3, nous étudions les performances de ces deux estimateurs. On commence le chapitre avec quelques précisions sur contexte considéré. Puis, les propriétés du *corrélateur classique* étant connues, elles sont brièvement rappelées. Les performances du *corrélateur compressé* sont ensuite évaluées dans un cadre distribué. Notre étude considère d'abord la corrélation et puis deux estimations qui en découlent : le temps de propagation et l'interspectre.

## Bibliographie

- [1] ACHLIOPTAS, D. Database-friendly random projections. In *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (2001), ACM, pp. 274–281.
- [2] BARANIUK, R. G., CEVHER, V., DUARTE, M. F., AND HEGDE, C. Model-based compressive sensing. *IEEE Transactions on Information Theory* 56, 4 (2010), 1982–2001.
- [3] BECKER, S. R. *Practical compressed sensing : modern data acquisition and signal processing*. California Institute of Technology, 2011.
- [4] BOURRIER, A., DAVIES, M. E., PELEG, T., PÉREZ, P., AND GRIBONVAL, R. Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems. *IEEE Transactions on Information Theory* 60, 12 (2014), 7928–7946.
- [5] BOURRIER, A., GRIBONVAL, R., AND PÉREZ, P. Compressive gaussian mixture estimation. In *Compressed Sensing and its Applications*. Springer, 2015, pp. 239–258.
- [6] CANDÈS, E. J., ROMBERG, J. K., AND TAO, T. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics* 59, 8 (2006), 1207–1223.
- [7] CANDÈS, E. J., AND WAKIN, M. B. An introduction to compressive sampling. *IEEE signal processing magazine* 25, 2 (2008), 21–30.
- [8] CHARTRAND, R. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters* 14, 10 (2007), 707–710.
- [9] CHEN, M. *Data compression for inference tasks in wireless sensor networks*. PhD thesis, State University of New York, 2006.
- [10] CHEN, M., AND FOWLER, M. L. Data compression for multi-parameter estimation for emitter location. *IEEE Transactions on Aerospace and Electronic Systems* 46, 1 (2010).
- [11] DAVENPORT, M. A., BOUFONOS, P. T., WAKIN, M. B., AND BARANIUK, R. G. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing* 4, 2 (2010), 445–460.
- [12] DAVENPORT, M. A., DUARTE, M. F., WAKIN, M. B., LASKA, J. N., TAKHAR, D., KELLY, K. F., AND BARANIUK, R. G. The smashed filter for compressive classification and target recognition. In *Electronic Imaging 2007* (2007), International Society for Optics and Photonics, pp. 64980H–64980H.
- [13] DEY, S. R. *Randomized Sampling and Multiplier-Less Filtering*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [14] DEY, S. R., AND OPPENHEIM, A. V. Frequency-shaped randomized sampling. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (2007), vol. 3, IEEE, pp. III–1493.
- [15] DONOHO, D. L. Compressed sensing. *IEEE Transactions on information theory* 52, 4 (2006), 1289–1306.
- [16] DONOHO, D. L. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics* 59, 6 (2006), 797–829.
- [17] ELДАР, Y. C. *Sampling Theory : Beyond Bandlimited Systems*. Cambridge University Press, 2015.
- [18] GASSO, G., RAKOTOMAMONJY, A., AND CANU, S. Recovering sparse signals with a certain family of nonconvex penalties and dc programming. *IEEE Transactions on Signal Processing* 57, 12 (2009), 4686–4698.

- [19] GIRYES, R., SAPIRO, G., AND BRONSTEIN, A. M. Deep neural networks with random gaussian weights : A universal classification strategy. *CoRR*, *abs/1504.08291* (2015).
- [20] JOHNSON, W. B., AND LINDENSTRAUSS, J. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics* 26, 189-206 (1984), 1.
- [21] KERIVEN, N., TREMBLAY, N., TRAONMILIN, Y., AND GRIBONVAL, R. Compressive K-means. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (New Orleans, United States, Mar. 2017).
- [22] KIROLOS, S., LASKA, J., WAKIN, M., DUARTE, M., BARON, D., RAGHEB, T., MASSOUD, Y., AND BARANIUK, R. Analog-to-information conversion via random demodulation. In *Design, Applications, Integration and Software, 2006 IEEE Dallas/CAS Workshop on* (2006), IEEE, pp. 71–74.
- [23] KRAHMER, F., AND WARD, R. New and improved johnson–lindenstrauss embeddings via the restricted isometry property. *SIAM Journal on Mathematical Analysis* 43, 3 (2011), 1269–1281.
- [24] LAGUNAS, E., SHARMA, S. K., CHATZINOTAS, S., AND OTTERSTEN, B. Compressive sensing based energy detector. In *Signal Processing Conference (EUSIPCO), 2016 24th European* (2016), IEEE, pp. 1678–1682.
- [25] LEXA, M. A., DAVIES, M. E., THOMPSON, J. S., AND NIKOLIC, J. Compressive power spectral density estimation. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (2011), IEEE, pp. 3884–3887.
- [26] LI, P., HASTIE, T. J., AND CHURCH, K. W. Very sparse random projections. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (2006), ACM, pp. 287–296.
- [27] METZLER, C. A., MALEKI, A., AND BARANIUK, R. G. From denoising to compressed sensing. *IEEE Transactions on Information Theory* 62, 9 (2016), 5117–5144.
- [28] METZLER, C. A., MOUSAVI, A., AND BARANIUK, R. G. Learned d-amp : A principled cnn-based compressive image recovery algorithm. *arXiv preprint arXiv :1704.06625* (2017).
- [29] MISHALI, M., AND ELДАР, Y. C. From theory to practice : Sub-nyquist sampling of sparse wideband analog signals. *IEEE Journal of Selected Topics in Signal Processing* 4, 2 (2010), 375–391.
- [30] MOUSAVI, A., AND BARANIUK, R. G. Learning to invert : Signal recovery via deep convolutional networks. *arXiv preprint arXiv :1701.03891* (2017).
- [31] MOUSAVI, A., PATEL, A. B., AND BARANIUK, R. G. A deep learning approach to structured signal recovery. In *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on* (2015), IEEE, pp. 1336–1343.
- [32] MUTHUKRISHNAN, S., ET AL. Data streams : Algorithms and applications. *Foundations and Trends® in Theoretical Computer Science* 1, 2 (2005), 117–236.
- [33] PAKROOH, P., PEZESHKI, A., AND SCHARF, L. L. Threshold effects in parameter estimation from compressed data. In *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE* (2013), IEEE, pp. 997–1000.
- [34] PAKROOH, P., PEZESHKI, A., SCHARF, L. L., COCHRAN, D., AND HOWARD, S. D. Analysis of fisher information and the cramér–rao bound for nonlinear parameter estimation after random compression. *IEEE Transactions on Signal Processing* 63, 23 (2015), 6423–6428.
- [35] RAMASAMY, D., VENKATESWARAN, S., AND MADHOW, U. Compressive parameter estimation in awgn. *IEEE Trans. Signal Processing* 62, 8 (2014), 2012–2027.
- [36] REBOREDO, H., RENNA, F., CALDERBANK, R., AND RODRIGUES, M. R. Bounds on the number of measurements for reliable compressive classification. *IEEE Transactions on Signal Processing* 64, 22 (2016), 5778–5793.

- [37] ROMERO, D., ARIANANDA, D. D., TIAN, Z., AND LEUS, G. Compressive covariance sensing : Structure-based compressive sensing beyond sparsity. *IEEE signal processing magazine* 33, 1 (2016), 78–93.
- [38] SAID, M. R., AND OPPENHEIM, A. V. Discrete-time randomized sampling. In *Electronics, Circuits and Systems, 2001. ICECS 2001. The 8th IEEE International Conference on* (2001), vol. 3, IEEE, pp. 1407–1411.
- [39] SESTOK, C. K., SAID, M. R., AND OPPENHEIM, A. V. Randomized data selection in detection with applications to distributed signal processing. *Proceedings of the IEEE* 91, 8 (2003), 1184–1198.
- [40] SLAVINSKY, J. P., LASKA, J. N., DAVENPORT, M. A., AND BARANIUK, R. G. The compressive multiplexer for multi-channel compressive sensing. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (2011), IEEE, pp. 3980–3983.
- [41] STEIN, M., CASTANEDA, M., MEZGHANI, A., AND NOSSEK, J. A. Information-preserving transformations for signal parameter estimation. *IEEE Signal Processing Letters* 21, 7 (2014), 866–870.
- [42] UNSER, M. Sampling-50 years after shannon. *Proceedings of the IEEE* 88, 4 (2000), 569–587.
- [43] VAUGHAN, R. G., SCOTT, N. L., AND WHITE, D. R. The theory of bandpass sampling. *IEEE Transactions on signal processing* 39, 9 (1991), 1973–1984.
- [44] VETTERLI, M., MARZILIANO, P., AND BLU, T. Sampling signals with finite rate of innovation. *IEEE transactions on Signal Processing* 50, 6 (2002), 1417–1428.
- [45] WIMALAJEEWA, T., CHEN, H., AND VARSHNEY, P. K. Performance limits of compressive sensing-based signal classification. *IEEE Transactions on Signal Processing* 60, 6 (2012), 2758–2770.
- [46] WOODRUFF, D. P., ET AL. Sketching as a tool for numerical linear algebra. *Foundations and Trends® in Theoretical Computer Science* 10, 1–2 (2014), 1–157.
- [47] ZHENG, L., MALEKI, A., WENG, H., WANG, X., AND LONG, T. Does lp minimization outperform l1 minimization. *arXiv preprint arXiv :1501.03704* (2015).

# Chapitre 3

## Estimation dans le domaine compressé

### Sommaire

---

<b>3.1</b>	<b>Contexte considéré</b>	<b>29</b>
<b>3.2</b>	<b>Estimation de corrélation</b>	<b>30</b>
3.2.1	Corrélateur classique	31
3.2.2	Corrélateur compressé	32
3.2.3	Influence de $\Phi$ dans la performance	34
3.2.4	Corrélateur compressé par échantillonnage aléatoire	36
3.2.5	Analyse asymptotique à taux de compression fixé	36
3.2.6	Compressibilité et performance	39
3.2.7	Effet du bruit	44
<b>3.3</b>	<b>Estimation du temps de propagation</b>	<b>46</b>
<b>3.4</b>	<b>Estimation d'interspectre</b>	<b>48</b>
3.4.1	Corrélogramme Blackman-Tukey	48
3.4.2	Corrélogramme classique	49
3.4.3	Corrélogramme compressé	50
3.4.4	Corrélogramme compressé par échantillonnage aléatoire	51
3.4.5	Compressibilité et performance	52
	<b>Bibliographie</b>	<b>60</b>

---

### 3.1 Contexte considéré

Dans le cas général, le nombre de capteurs d'un réseau est grand, voire très grand. Nous nous limitons ici à l'étude de deux capteurs voisins, notés  $x$  et  $y$ , qui mesurent respectivement les signaux  $x(t)$  et  $y(t)$ . Ces signaux provenant d'endroits géographiques différents, toute estimation conjointe à partir du couple, par exemple  $\hat{\Gamma}_{xy}(\tau)$ , ou  $\hat{\gamma}_{xy}(\nu)$ , requiert la transmission d'au moins un de ces signaux. Comme expliqué précédemment, les calculs peuvent avoir lieu soit à l'un de ces deux endroits (réseau collaboratif), soit dans un troisième (réseau centralisé). Cf la figure 3.1

Le choix que nous avons fait et le problème que nous étudions sont aussi illustrés sur cette même figure : seul  $x$  transmet ses mesures vers  $y$  et ce dernier effectue l'estimation conjointe. Les données sont disponibles dans  $x$ , en très grand nombre (noté  $N$ ). Le coût énergétique requis pour leur transmission étant élevé, nous étudions différentes méthodes visant à limiter cette transmission à  $M$  valeurs, où  $M < N$ , sans pour autant dégrader la précision des calculs de manière importante.

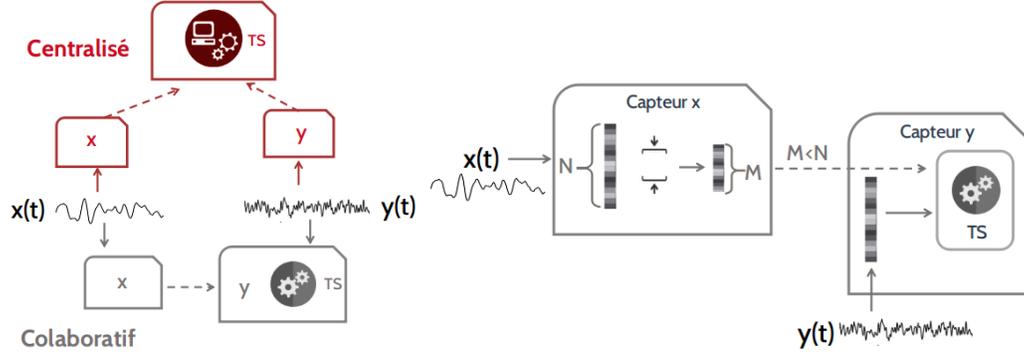


FIGURE 3.1 – Gauche : approches d’estimation distribuée dans un réseau à deux capteurs. Un des signaux, au moins, doit être transmis. Droite : problème étudié : dû au coût énergétique élevé associé à la transmission de données, l’un des capteurs ne transmet qu’une partie de ses observations vers l’autre capteur, qui effectue les calculs. Nous étudions différentes approches de réduction de données dans ces travaux.

### 3.2 Estimation de corrélation

Les signaux observés dans les capteurs sont modélisés comme étant aléatoires, à valeurs réelles, de moyenne nulle, conjointement stationnaires et ergodiques. D’abord, on s’intéresse à l’estimation de la fonction d’intercorrélation pour le retard  $\tau$ , notée  $\Gamma_{xy}(\tau) = E[x(t)y(t+\tau)]$ . Nous admettrons la sommabilité absolue de cette dernière [15, 8], ainsi que celle du cumulatif d’ordre quatre  $C_{xyxy}(a, b, c) = \text{Cum}[x(t), y(t+a), x(t+b), y(t+c)]$  :

$$\int_{\mathbb{R}} |\Gamma_{xy}(\tau)| d\tau < \infty \text{ et } \int_{\mathbb{R}^3} |C_{xyxy}(a, b, c)| da db dc < \infty$$

Parmi les estimateurs de corrélation proposés dans la littérature, par exemple [3, 6, 8, 10, 11, 17], figurent deux estimateurs *standards* donnés pour tout retard  $\tau$  de valeur absolue inférieure à  $N$ .

L’estimateur biaisé *standard* :

$$c_{xy}(\tau) = \frac{1}{N} \sum_{i=1}^{N-|\tau|} x_i y_{i+|\tau|}$$

et l’estimateur non biaisé *standard* :

$$c'_{xy}(\tau) = \frac{1}{N-|\tau|} \sum_{i=1}^{N-|\tau|} x_i y_{i+|\tau|}$$

Dans le cas de l’estimateur biaisé, le biais apparaît par construction.  $c_{xy}(\tau)$  a la même forme qu’une moyenne empirique sur  $N$  valeurs mais le nombre d’éléments additionnés change selon la valeur d’intérêt de  $\tau$  (et devient très petit lorsque  $|\tau| \approx N$ ). Notons cependant que l’on divise par  $N$  quelque soit le retard à estimer.

En contrepartie, l’estimateur non biaisé élimine le biais en tenant compte du nombre de termes effectivement présents dans la somme. On divise alors par le nombre d’éléments additionnés ( $N - |\tau|$ ). En conséquence, sa variance augmente avec le retard, si bien que, dans la littérature, on a tendance à lui préférer l’estimateur biaisé [17, 14]. Alors, pour réduire l’impact du biais, une règle empirique [10], attribuée à Box et Jenkins [4], suggère que  $N$  soit au moins égal à 50 et que  $|\tau|$  reste inférieur à  $N/4$ .

Or l'estimateur de corrélation que nous avons présenté dans le chapitre 1, appelé estimateur classique (par opposition à l'estimateur compressé), est une variante de ces estimateurs usuels. Il réduit le nombre de données transmises de la façon la plus brutale qui soit, à savoir la restriction pure et simple des  $N$  échantillons observés aux  $M$  premières valeurs. Une deuxième possibilité est le sous-échantillonnage périodique, mais il est exclu de cette étude car il peut engendrer repliement du spectre, ce qui changerait l'allure de la fonction que l'on cherche à estimer.

#### Corrélateur classique :

Transmettre (de  $x$  vers  $y$ ) uniquement un sous-ensemble  $\mathbf{x}$ , de  $M$  échantillons successifs parmi les  $N$  disponibles, puis estimer la corrélation dans le capteur  $y$  selon :

$$\mathcal{C}_M(\tau) = \frac{1}{M} \langle \mathbf{x}, \mathbf{y}_\tau \rangle = \frac{1}{M} \mathbf{x}^\top \mathbf{y}_\tau \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^M$$

$$\mathbf{x} = [x_1, \dots, x_M]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{M+\tau}]^\top$$

Cet estimateur est non biaisé et de variance finie, quelque soit le retard considéré. En effet, le capteur  $y$  observe le signal sur une durée suffisamment longue pour toujours disposer de  $M$  points, quelque soit la valeur du retard  $\tau$ .

D'autre part, l'estimateur que nous nous proposons d'étudier réduit la taille des données par projection aléatoire. Les composantes de la matrice de compression sont supposées centrées, indépendantes et identiquement distribuées (i.i.d.).

#### Corrélateur compressé

Envoyer  $M$  valeurs numériques obtenues à partir des  $N$  valeurs de  $\mathbf{x}$ , puis utiliser le *corrélateur compressé* dans  $y$  :

$$C_N(\tau) = \langle \Phi \mathbf{x}, \Phi \mathbf{y}_\tau \rangle = (\Phi \mathbf{x})^\top (\Phi \mathbf{y}_\tau) \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

$$\mathbf{x} = [x_1, \dots, x_N]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Dans le cas de cet estimateur, nous admettons que le coût énergétique lié aux opérations nécessaires pour calculer les projections est faible par rapport à celui de la transmission de données.

Un cas particulier de matrice  $\Phi$ , qui nous intéresse justement pour des raisons de réduction des calculs, équivaut à sous-échantillonner aléatoirement sans remise les vecteurs  $\mathbf{x}$  et  $\mathbf{y}$ . Une telle matrice possédant uniquement un élément non nul par ligne, il n'y a pas de calcul numérique à effectuer dans ce cas, mais comme les lignes de cette matrice ne sont pas indépendantes, ce cas particulier sera étudié séparément. Dans la suite, cette approche sera identifiée comme échantillonnage aléatoire et l'estimateur sera noté  $\hat{C}_N(\tau)$ .

Les deux approches considérées sont illustrées dans la figure 3.2. Le premier capteur,  $x$ , transmet au deuxième,  $y$ , qui procède à l'estimation.

### 3.2.1 Corrélateur classique

Les statistiques de  $\mathcal{C}_M$  s'établissent de la même manière que celles de l'estimateur biaisé standard, [8, 15], mais les résultats diffèrent quelque peu.

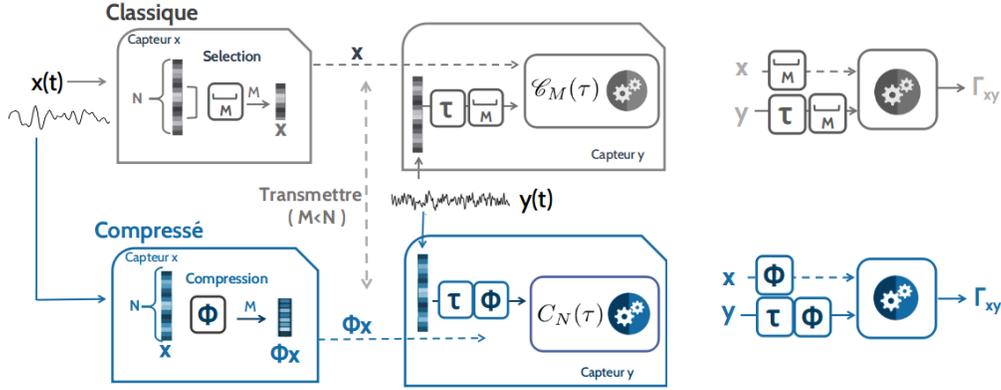


FIGURE 3.2 – Gauche : approches de réduction de données dans la transmission. L’approche classique (haut) consiste à sélectionner et transmettre  $M$  échantillons successifs. L’approche compressée (bas) envoie à la place  $M$  mesures calculées sur  $N$  échantillons. Droite : blocs-diagrammes des corrélateurs.

### Biais

Cet estimateur est non biaisé :

$$\mathbf{E}[\mathcal{C}_M(\tau)] = \Gamma_{xy}(\tau)$$

### Covariance

Pour deux retards quelconques  $\ell$ , et  $k$  on a :

$$\mathbf{Cov}\{\mathcal{C}_M(k), \mathcal{C}_M(\ell)\} = \frac{1}{M} \sum_{r=-M}^M \left(1 - \frac{|r|}{M}\right) f_{xy}(r, k, \ell) \quad (3.1)$$

où  $f_{xy}(r, k, \ell) = C_{xyxy}(k, r, \ell + r) + \Gamma_{xx}(r)\Gamma_{yy}(r + \ell - k) + \Gamma_{xy}(\ell + r)\Gamma_{xy}(k - r)$

La variance s’en déduit en se plaçant en  $\ell = k = \tau$

### Variance

$$\mathbf{Var}\{\mathcal{C}_M(\tau)\} = \frac{1}{M} \sum_{r=-M}^M \left(1 - \frac{|r|}{M}\right) f_{xy}(r, \tau, \tau) \quad (3.2)$$

où  $f_{xy}(r, \tau, \tau) = C_{xyxy}(\tau, r, \tau + r) + \Gamma_{xx}(r)\Gamma_{yy}(r) + \Gamma_{xy}(\tau + r)\Gamma_{xy}(\tau - r)$

## 3.2.2 Corrélateur compressé

Rappelons tout d’abord que les éléments  $\varphi_{ij}$  qui constituent la matrice de compression  $\Phi$  sont supposés centrés, indépendants et identiquement distribués. Les calculs présentés en annexe permettent d’établir les statistiques de l’estimateur construit sur les signaux compressés. Il apparaît en particulier que l’estimateur est non biaisé, dès lors que l’on fixe la variance des termes  $\varphi_{ij}$  à  $\frac{1}{MN}$ . Dans ces conditions, le Kurtosis de ces coefficients

devient

$$\kappa[\varphi] = \frac{\text{Cum}_4[\varphi]}{\mathbf{E}[\varphi^2]^2} = \frac{\mathbf{E}[\varphi^4]}{\mathbf{E}[\varphi^2]^2} - 3 = M^2 N^2 \mathbf{E}[\varphi^4] - 3$$

### Biais

L'estimateur est non biaisé

$$\mathbf{E}[C_N(\tau)] = \Gamma_{xy}(\tau)$$

### Covariance

Pour deux retards quelconques  $\ell$  et  $k$  on a :

$$\begin{aligned} \mathbf{Cov}\{C_N(k), C_N(\ell)\} &= \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} + \frac{\kappa[\varphi]}{MN^2} \sum_{\alpha} \mathbf{E}[x_{\alpha}^2 y_{\alpha+k} y_{\alpha+\ell}] \\ &\quad + \frac{1}{MN^2} (\mathbf{E}[\|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_{\ell})] + \mathbf{E}[(\mathbf{x}^T \mathbf{y}_k)(\mathbf{x}^T \mathbf{y}_{\ell})]) \quad (3.3) \end{aligned}$$

Ou bien :

$$\begin{aligned} \mathbf{Cov}\{C_N(k), C_N(\ell)\} &= \left(1 + \frac{1}{M}\right) \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} \\ &\quad + \frac{\kappa[\varphi]}{MN} (C_{xyxy}(k, 0, \ell) + 2\Gamma_{xy}(k)\Gamma_{xy}(\ell) + \Gamma_{xx}(0)\Gamma_{yy}(k - \ell)) \\ &\quad + \frac{1}{M} (\Gamma_{xx}(0)\Gamma_{yy}(k - \ell) + \Gamma_{xy}(k)\Gamma_{xy}(\ell)) \\ &\quad + \frac{1}{MN} \left( \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(r + k, 0, r + \ell) + 2\Gamma_{xy}(\ell + r)\Gamma_{xy}(r + k)\} \right) \end{aligned}$$

Le cas particulier  $\ell = k = \tau$  conduit à la variance

### Variance

$$\begin{aligned} \mathbf{Var}\{C_N(\tau)\} &= \mathbf{Var}\{\mathcal{C}_N(\tau)\} + \frac{\kappa[\varphi]}{MN^2} \sum_{\alpha} \mathbf{E}[x_{\alpha}^2 y_{\alpha+\tau}^2] \\ &\quad + \frac{1}{MN^2} (\mathbf{E}[\|\mathbf{x}\|_2^2 \|\mathbf{y}_{\tau}\|_2^2] + \mathbf{E}[(\mathbf{x}^T \mathbf{y}_{\tau})^2]) \quad (3.4) \end{aligned}$$

Que l'on peut réécrire :

$$\begin{aligned} \mathbf{Var}\{C_N(\tau)\} &= \left(1 + \frac{1}{M}\right) \mathbf{Var}\{\mathcal{C}_N(\tau)\} \\ &\quad + \frac{\kappa[\varphi]}{MN} (C_{xyxy}(\tau, 0, \tau) + 2\Gamma_{xy}^2(\tau) + \Gamma_{xx}(0)\Gamma_{yy}(0)) \\ &\quad + \frac{1}{M} (\Gamma_{xx}(0)\Gamma_{yy}(0) + \Gamma_{xy}^2(\tau)) \\ &\quad + \frac{1}{MN} \left( \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(r + \tau, 0, r + \tau) + 2\Gamma_{xy}^2(\tau + r)\} \right) \end{aligned}$$

Notons que l'expression (3.4) est plus générale que les résultats connus dans la littérature relative aux projections aléatoires et ce pour deux raisons :

- D’abord, ces études se focalisent sur certains types de matrice de compression  $\Phi$ , alors que nous n’avons pas limité ce choix.
- Deuxièmement, les vecteurs habituellement considérés sont déterministes, alors que nous étudions le cas aléatoire.

Par exemple, [1, 9] ont établi des résultats sous la forme :

$$\mathbf{Var}\{C_N(\tau)\} = \frac{\kappa[\varphi]}{MN^2} \sum_{\alpha} \mathbf{E} [x_{\alpha}^2 y_{\alpha+\tau}^2] + \frac{1}{MN^2} ([\|\mathbf{x}\|_2^2 \|\mathbf{y}_{\tau}\|_2^2] + [(\mathbf{x}^T \mathbf{y}_{\tau})^2])$$

Mais ces études considèrent uniquement des matrices de compression ternaires parcimonieuses.

D’autre part, le terme  $\mathbf{Var}\{\mathcal{C}_N(\tau)\}$ , retrouvé dans (3.4), n’apparaît pas dans ces travaux. Ce terme capture l’imprécision intrinsèque aux données aléatoires. Même avec  $N$  données à disposition, l’estimation de la corrélation n’est pas exacte.

Dans la section suivante, nous étudions l’impact de la matrice de compression sur la précision de l’estimation, ainsi que la complexité des calculs inhérents à cette compression.

### 3.2.3 Influence de $\Phi$ dans la performance

Les expressions de variance et covariance obtenues dans la section précédente, font apparaître le kurtosis  $\kappa[\varphi]$  de la distribution de probabilité des éléments de la matrice. Le kurtosis est souvent interprété comme l’écart entre une densité de probabilité quelconque et la distribution Gaussienne (pour laquelle  $\kappa[\varphi] = 0$ ). Les distributions qui possèdent un kurtosis négatif sont appelées *sous gaussiennes*. Leurs queues décroissent plus rapidement que celles de la Gaussienne. Ces distributions sont plates ou multimodales [7, p. 39]. Par contre, si  $\kappa[\varphi] > 0$ , on parle de distributions *sur gaussiennes*. Celles-ci sont souvent pointues et ont des queues lourdes, c’est-à-dire qu’elles décroissent moins vite que la Gaussienne.

Puisque  $\kappa[\varphi]$  peut prendre des valeurs positives ou négatives, la loi de probabilité que l’on choisit pour tirer les éléments de la matrice de compression, détermine d’une part la précision statistique obtenue et d’autre part, comme nous le préciserons dans la suite, la nature et le nombre des calculs à effectuer pour compresser les signaux.

Voyons maintenant quelques distributions ainsi que leurs avantages et inconvénients respectifs.

- Matrices à éléments Gaussiens :  $\varphi \sim \mathcal{N}(0, 1/MN)$ . C’est le cas que l’on retrouve le plus souvent dans la littérature. Puisque  $\kappa[\varphi] = 0$ , la contribution du terme associé au kurtosis est nulle.

L’inconvénient est que ces matrices sont très denses et à éléments réels. Une haute complexité est donc nécessaire en termes de stockage (il faut garder en mémoire  $\mathcal{O}(MN)$  valeurs numériques) et en termes calculatoires ( $\mathcal{O}(MN)$  additions et produits sont nécessaires pour évaluer chaque vecteur compressé).

- Matrices de Bernoulli :  $\varphi = \pm 1/\sqrt{MN}$  équiprobables. Ces matrices ont deux avantages par rapport aux gaussiennes. Le premier est que les multiplications, très coûteuses en termes énergétiques, ne sont pas nécessaires. Le deuxième est que  $\kappa[\varphi] = -2$ . La variance d’estimation est donc inférieure au cas gaussien. Cette matrice est néanmoins aussi très dense et, par conséquent, également chère quant

au nombre de calculs (additions) et de stockages  $\mathcal{O}(MN)$ .

- Matrices ternaires parcimonieuses (*Very sparse random projections*, établies par Li et. al. [9]). Le remplissage de ces matrices est contrôlé par un paramètre  $s$  supérieur ou égal à 1. En moyenne, le nombre de termes non nuls est égal à  $MN/s \leq MN$ .

Chaque élément de la matrice est tiré au sort de manière indépendante suivant la loi :

$$\varphi = \frac{\sqrt{s}}{\sqrt{MN}} \begin{cases} 1 & \text{avec probabilité } \frac{1}{2s} \\ -1 & \text{avec probabilité } \frac{1}{2s} \\ 0 & \text{avec probabilité } 1 - \frac{1}{s} \end{cases}$$

La division par le terme  $\sqrt{MN}$  qui apparaît ici vise à respecter la contrainte de non biais sur l'estimateur de corrélation.

On retrouve dans cette famille les deux distributions étudiées par Achlioptas [1] (les cas  $s = 1$  et  $s = 3$ ). Ses travaux sont parmi les premiers à étudier les projections aléatoires en vue d'effectuer des calculs en dimension réduite. Le Kurtosis des éléments d'une telle matrice est donné par

$$\kappa[\varphi] = \frac{\mathbf{E}[\varphi^4]}{\mathbf{E}[\varphi^2]^2} - 3 = s - 3$$

Ceci permet de comprendre l'influence du paramètre  $s$  : plus il augmente, plus la matrice engendrée est creuse, ce qui simplifie les calculs. Néanmoins le kurtosis, et par conséquent la perte en précision statistique, augmenteront également. Chose intéressante,  $s = 3$ , a le même kurtosis qu'une matrice Gaussienne, ce qui conduit à la même précision statistique, mais au prix de seulement  $\frac{MN}{3}$  additions.

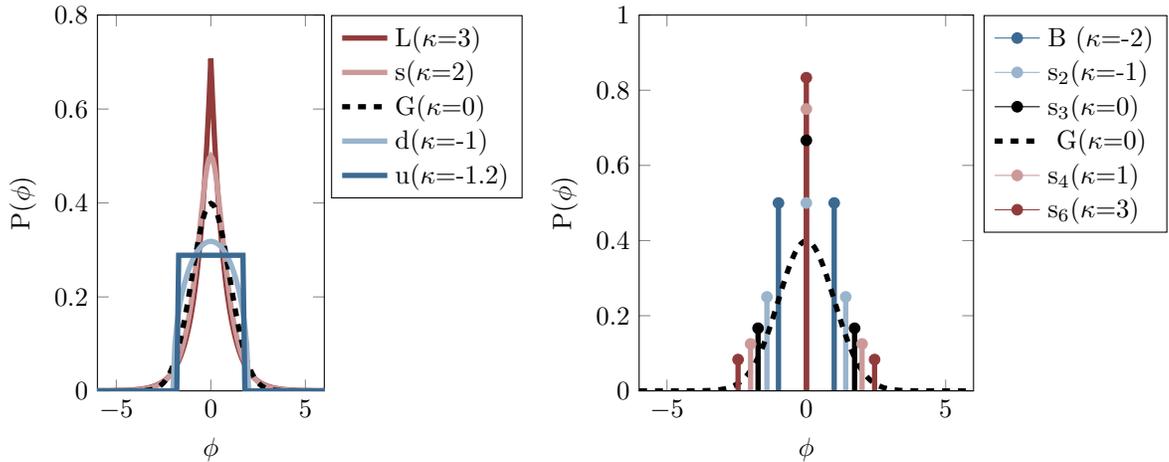


FIGURE 3.3 – Kurtosis normalisé ( $\kappa[\varphi]$ ) pour différentes distributions de probabilité. Gauche : distributions continues : L :Laplacienne, sh :sinus hyperbolique ,G : Gaussienne, d :demi-cercle , u : uniforme. Droite : distributions discrètes : B :Bernoulli, et parcimonieuse de paramètre  $s = 2, 3, 4$  ou  $6$ . Les distributions sur-Gaussiennes apparaissent en rouge et les sous-Gaussiennes en bleu.

Comme mentionné précédemment, l'échantillonnage aléatoire sans remise est un cas particulier de matrice de compression, dont les lignes ne sont pas indépendantes. Il est donc nécessaire de l'étudier séparément.

### 3.2.4 Corrélateur compressé par échantillonnage aléatoire

Soient  $\mathbf{x} = [x_1, \dots, x_N]^\top$  et  $\mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$ , les données mesurées par chaque capteur. Posons  $\mathbf{s} = \mathbf{x} \circ \mathbf{y}_\tau$ , le produit terme à terme de ces deux vecteurs. Alors que le corrélateur classique moyenne les  $N$  éléments du vecteur  $\mathbf{s}$ , le compressé par échantillonnage aléatoire sans remise se contente d'en prélever  $M$  au hasard. On peut écrire l'estimateur correspondant comme :

$$\mathring{C}_N(\tau) = \frac{1}{M} \sum_{i=1}^N \varepsilon_i s_i$$

où  $M$  coefficients  $\varepsilon_i$ , tirés au sort sans remise, sont égaux à 1 tandis que les  $N - M$  autres sont nuls. En annexe, nous établissons les propriétés de cet estimateur :

#### Biais

Cet estimateur est non biaisé :

$$\mathbf{E}\{\mathring{C}_N\} = \Gamma_{xy}(\tau)$$

#### Covariance

Pour deux retards  $\ell, k$  on a :

$$\begin{aligned} \mathbf{Cov}\{\mathring{C}_N(k), \mathring{C}_N(\ell)\} &= \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} \\ &+ \frac{N/M - 1}{N - 1} (\mathbf{C}_{xyxy}(k, 0, \ell) + \Gamma_{xx}(0)\Gamma_{yy}(k - \ell) \\ &\quad + \Gamma_{xy}(\ell)\Gamma_{xy}(k) - \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\}) \quad (3.5) \end{aligned}$$

Enfin, pour  $\ell = k = \tau$

#### Variance

$$\begin{aligned} \mathbf{Var}\{\mathring{C}_N(\tau)\} &= \mathbf{Var}\{\mathcal{C}_N(\tau)\} \\ &+ \frac{N/M - 1}{N - 1} (\mathbf{C}_{xyxy}(\tau, 0, \tau) + \Gamma_{xx}(0)\Gamma_{yy}(0) + \Gamma_{xy}^2(\tau) - \mathbf{Var}\{\mathcal{C}_N(\tau)\}) \quad (3.6) \end{aligned}$$

Pour simplifier l'analyse du comportement des variances des estimateurs, il est courant d'effectuer ce que l'on appelle une analyse asymptotique. C'est à dire que l'on étudie le comportement lorsque le nombre de données exploitées devient grand. Cette analyse fait l'objet de la section suivante.

### 3.2.5 Analyse asymptotique à taux de compression fixé

On considère que les nombres de points  $N$  et  $M$  tendent vers l'infini, mais sous la contrainte d'un taux de compression  $\alpha = N/M \geq 1$  constant. Etablissons d'abord les résultats généraux pour les trois corrélateurs envisagés.

#### Corrélateur classique

Pour cet estimateur, sur  $N$  valeurs :

$$N \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} = \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) f_{xy}^\tau(r, k, \ell)$$

qui, sous réserve que la série des  $f_{xy}^\tau(r, k, \ell)$  converge (ce qui est garanti par la sommabilité absolue des fonctions de corrélation de deuxième et quatrième ordre) tend vers

$$v(k, \ell) := \sum_{r \in \mathbb{Z}} f_{xy}^\tau(r, k, \ell)$$

lorsque  $N$  tend vers l'infini.

### Corrélateur compressé

De la même manière, pour l'estimateur compressé de la corrélation :

$$\begin{aligned} \lim_{N \rightarrow \infty} N \mathbf{Cov}\{C_N(k), C_N(\ell)\} &= v(k, \ell) + C_{4,\varphi}(2\Gamma_{xy}(k)\Gamma_{xy}(\ell) + \Gamma_{xx}(0)\Gamma_{yy}(k-\ell) + C_{xyxy}(k, 0, \ell)) \\ &\quad + \alpha(\Gamma_{xx}(0)\Gamma_{yy}(k-\ell) + \Gamma_{xy}(k)\Gamma_{xy}(\ell)) \end{aligned}$$

On suppose ici l'existence de la limite

$$C_{4,\varphi} = \lim_{N \rightarrow \infty} N \frac{\kappa[\varphi]}{NM} = \lim_{N \rightarrow \infty} \alpha \frac{\kappa[\varphi]}{N}$$

Le Kurtosis des termes des matrices gaussiennes et de Bernoulli étant fini, cette limite est nulle dans ces deux cas. Par contre, pour des matrices creuses ternaires, rien n'interdit de choisir le paramètre  $s$  proportionnel à  $M$  par exemple, de sorte que la limite vaut alors

$$C_{4,\varphi} = \lim_{N \rightarrow \infty} \frac{s-3}{M} = \frac{s}{M}$$

### Échantillonnage aléatoire

$$\lim_{N \rightarrow \infty} N \mathbf{Cov}\{\hat{C}_N(k), \hat{C}_N(\ell)\} = v(k, \ell) + (\alpha-1)(\Gamma_{xx}(0)\Gamma_{yy}(k-\ell) + \Gamma_{xy}(k)\Gamma_{xy}(\ell) + C_{xyxy}(k, 0, \ell))$$

Pour illustrer ces résultats, appliquons les maintenant à un signal autorégressif d'ordre 1, modèle simple qui permet de comprendre et comparer le comportement de ces estimateurs.

### Processus gaussien AR(1)

Nous considérons le modèle des processus gaussiens, autorégressifs d'ordre 1  $x_t = ax_{t-1} + \sqrt{1-a^2}\epsilon_t$  où  $\epsilon_t$  désigne une séquence i.i.d. de variables aléatoires gaussiennes centrées réduites, et  $a \in ]-1, 1[$ . Pour tout retard entier  $\tau$ , l'autocorrélation de ce processus vaut :  $\Gamma_{xx}(\tau) = a^{|\tau|}$   $|a| < 1$ .

L'allure d'un tel signal est pilotée par le seul paramètre réel  $a$ . Pour  $a = 0$ , le signal obtenu est un bruit blanc, l'archétype du signal non compressible. Au fur et à mesure que  $|a|$  s'éloigne de 0, le signal devient de plus en plus compressible, en vertu de la redondance apportée par la corrélation qui lie ses échantillons successifs. Quelques cas particuliers sont présentés dans la figure 3.4.

Par souci de simplicité et sans perte de généralité, on considère dans la suite que les signaux reçus sur les deux capteurs sont égaux, ce qui revient à estimer l'autocorrélation  $\Gamma_{xx}(\tau) = a^{|\tau|}$ . Les covariances asymptotiques s'écrivent alors comme suit

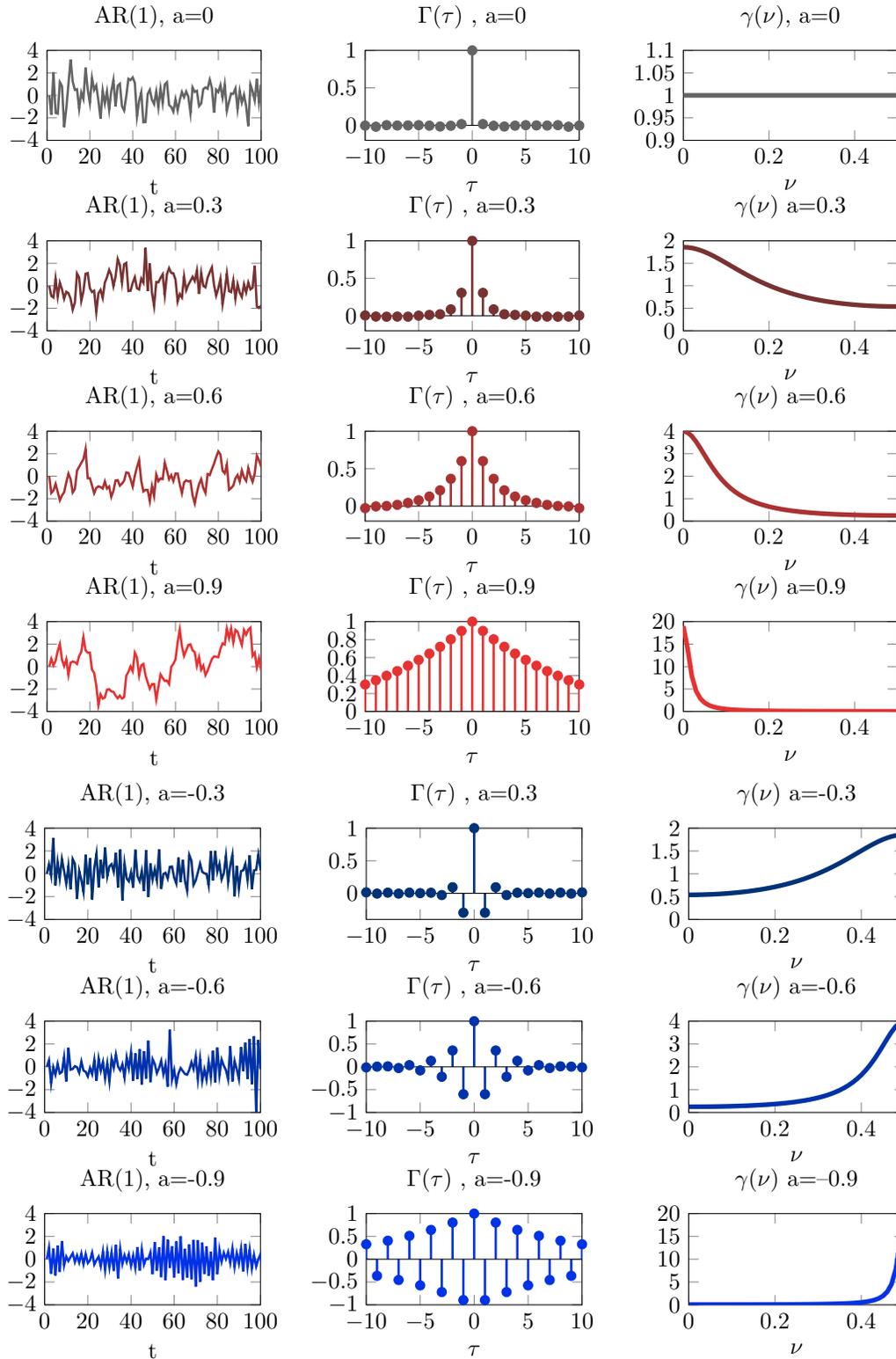


FIGURE 3.4 – Processus AR(1) gaussien pour différentes valeurs de  $a$ . Gauche : réalisation du processus. Centre : fonction d'autocorrélation. Pour  $a = 0$ , on retrouve un processus dont les échantillons sont décorrélés. Lorsque  $|a| \rightarrow 1$  la corrélation entre ses échantillons successifs (et donc sa redondance) augmente. Droite : densité spectrale de puissance. Si  $a = 0$ , le processus possède un support fréquentiel plein (un bruit blanc) qui est un signal non compressible. Par contre, lorsque  $|a| \rightarrow 1$ , le support fréquentiel diminue, dans ce sens, un signal plus *compressible*.

Dans le cas de l'estimateur classique sur  $N$  échantillons, on a :

$$\begin{aligned} \lim_{N \rightarrow \infty} N \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} \\ = 2a^2 \left( \frac{a^{|k-\ell|} + a^{|k+\ell|}}{1-a^2} \right) + (|k-\ell|+1)a^{|k-\ell|} + (|k+\ell|+1)a^{|k+\ell|} \\ = v(k, \ell) \end{aligned}$$

si bien que pour un horizon restreint à  $M$  points

$$\begin{aligned} \lim_{N \rightarrow \infty} N \mathbf{Cov}\{\mathcal{C}_M(k), \mathcal{C}_M(\ell)\} \\ = \alpha v(k, \ell) \end{aligned}$$

Pour l'estimateur compressé par projection aléatoire, on obtient :

$$\lim_{N \rightarrow \infty} N \mathbf{Cov}\{C_N(k), C_N(\ell)\} = v(k, \ell) + \alpha(a^{|k|}a^{|\ell|} + a^{|k-\ell|}) + C_{4,\varphi}(2a^{|k|}a^{|\ell|} + a^{|k-\ell|})$$

et enfin pour l'échantillonnage aléatoire sans remise :

$$\lim_{N \rightarrow \infty} N \mathbf{Cov}\{\mathring{C}_N(k), \mathring{C}_N(\ell)\} = v(k, \ell) + (\alpha - 1)(a^{|k|}a^{|\ell|} + a^{|k-\ell|})$$

Quant aux variances, elles se déduisent des expressions précédentes dans lesquelles il suffit de poser  $k = \ell = \tau$  :

$$\lim_{N \rightarrow \infty} N \mathbf{Var}\{\mathcal{C}_N(\tau)\} = 2a^2 \left( \frac{1 + a^{2|\tau|}}{1 - a^2} \right) + (2|\tau| + 1)a^{2|\tau|} + 1 = v(\tau, \tau)$$

$$\lim_{N \rightarrow \infty} N \mathbf{Var}\{\mathcal{C}_M(\tau)\} = \alpha v(\tau, \tau)$$

$$\lim_{N \rightarrow \infty} N \mathbf{Var}\{C_N(\tau)\} = v(\tau, \tau) + \alpha(a^{2|\tau|} + 1) + C_{4,\varphi}(2a^{2|\tau|} + 1)$$

$$\lim_{N \rightarrow \infty} N \mathbf{Var}\{\mathring{C}_N(\tau)\} = v(\tau, \tau) + (\alpha - 1)(a^{2|\tau|} + 1)$$

Notons que les résultats relatifs aux variances ne dépendent de  $a$  qu'au travers de son carré. Il devient donc inutile par la suite de distinguer  $a$  positif ou négatif.

Dans le paragraphe suivant, ces variances asymptotiques sont exploitées pour comparer la performance des estimateurs et plus précisément l'influence de la nature du signal (sa *compressibilité*) sur la qualité du résultat.

### 3.2.6 Compressibilité et performance

Jusqu'à présent nous avons considéré trois estimateurs différents de la corrélation :  $\mathcal{C}_M(\tau)$ ,  $C_N(\tau)$ , et  $\mathring{C}_N(\tau)$ . Pour justifier la comparaison qui est faite dans la suite, nous rappelons que ces estimateurs présentent un même coût énergétique en termes d'échange de données. Tous trois requièrent la transmission de  $M$  valeurs numériques, d'un capteur à l'autre pour effectuer l'estimation. Nous étudierons donc le comportement des rapports :

$$\rho_C = \frac{\mathbf{Var}_\infty\{\mathcal{C}_M(\tau)\}}{\mathbf{Var}_\infty\{C_N(\tau)\}} \text{ et } \hat{\rho}_C = \frac{\mathbf{Var}_\infty\{\mathcal{C}_M(\tau)\}}{\mathbf{Var}_\infty\{\mathring{C}_N(\tau)\}}$$

où  $\mathbf{Var}_\infty$  dénote la variance asymptotique.

Nous soulignons que lorsque ces rapports ont une valeur supérieure à 1, ils quantifient l'éventuel gain en termes de variance d'estimation des méthodes compressées par rapport à l'estimateur classique. Par contre, quand ces rapports sont inférieurs à 1, les méthodes proposées sont moins performantes que l'estimation classique de corrélation.

En vue de préciser l'influence du type de signal traité dans le comportement de ces estimateurs, nous étudions ces rapports pour le processus AR(1) gaussien.

### Cas AR(1) gaussien

Dans ce cas particulier on a :

$$\rho_C = \frac{\alpha v(\tau, \tau)}{v(\tau, \tau) + \alpha(a^{2|\tau|} + 1) + C_{4,\varphi}(2a^{2|\tau|} + 1)}$$

et

$$\hat{\rho}_C = \frac{\alpha v(\tau, \tau)}{v(\tau, \tau) + (\alpha - 1)(a^{2|\tau|} + 1)}$$

$$\text{où } v(\tau, \tau) = 2a^2 \left( \frac{1 + a^{2|\tau|}}{1 - a^2} \right) + (2|\tau| + 1)a^{2|\tau|} + 1$$

Dans ces expressions, on voit que le comportement des rapports est déterminé par  $\tau$  et  $\alpha$ . On illustrera d'abord le cas  $\tau = 0$  pour observer l'influence de  $\alpha$  (la *compressibilité* du signal). Ensuite, on verra que ce comportement est représentatif, parce qu'il varie peu en fonction de  $\tau$ .

Notre comparaison considère trois estimateurs compressés. Le premier d'entre eux repose sur une projection aléatoire via une matrice gaussienne dense qui nécessite  $\mathcal{O}(MN)$  opérations. Pour cette matrice gaussienne, la limite ( $C_{\varphi,4}$ ) est nulle. Le second s'appuie sur une matrice ternaire de paramètre  $s = M/2$ , qui ne nécessite plus que  $\mathcal{O}(N)$  opérations et pour laquelle la limite ( $C_{\varphi,4}$ ) vaut 1/2. Le troisième estimateur compressé exploite l'échantillonnage aléatoire, auquel cas plus aucun calcul numérique n'est nécessaire lors de la compression.

La figure 3.5 illustre les résultats obtenus dans les trois cas. On observe de manière générale que les comportements des trois estimateurs sont très proches. Cependant, pour les estimateurs par projection aléatoire les rapports de variance s'avèrent inférieurs à 1 pour des signaux peu corrélés et très riches en contenu spectral ( $|a| \rightarrow 0$ ). Dans ce cas, l'estimateur classique est plus performant que les deux estimateurs compressés.

Cela n'arrive pas dans le cas de l'estimateur par échantillonnage aléatoire qui, dans le pire des cas, reste aussi précis que le classique. Par contre, ces rapports de variance présentent des valeurs supérieures à 1 pour des signaux très redondants et donc compressibles (ce qui arrive quand  $|a| \rightarrow 1$ ). Dans ce cas, la valeur de ces rapports exprime directement le gain en variance que l'on obtient avec l'estimateur compressé par rapport au cas classique. Notons que lorsque  $a$  est très proche de 1, ce gain est du même ordre de grandeur que le taux de compression. Cela signifie que pour ce type de signaux, la compression ne dégrade quasiment pas la précision de l'estimation. Ce qui est remarquable, c'est que l'échantillonnage aléatoire ne requiert aucun calcul pour effectuer la compression et en même temps cette approche s'avère aussi la plus performante !

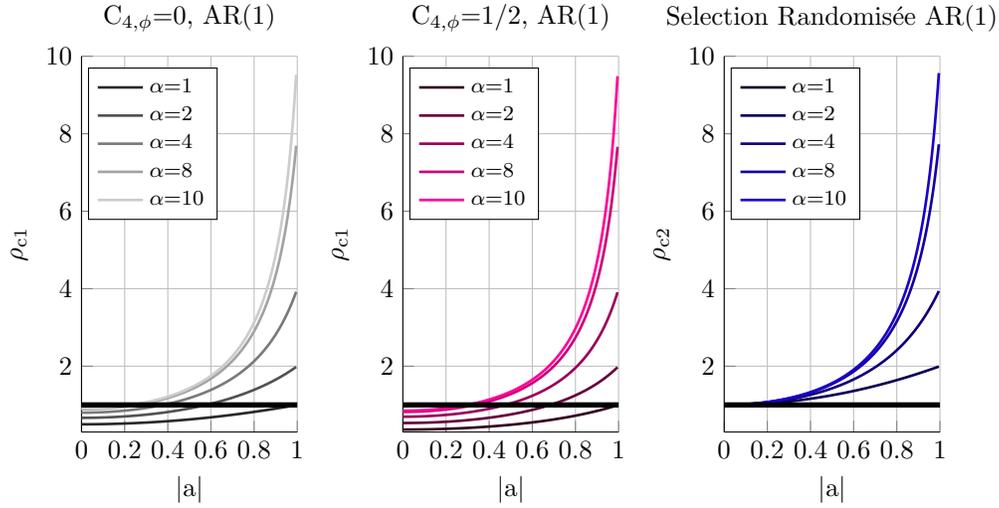


FIGURE 3.5 – Rapport ( $\rho$ ) de variance pour  $\tau = 0$  en fonction de  $|a|$  et différents types de compression : Gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s_M = M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. Pour  $|a| \rightarrow 0$  les signaux obtenus sont proches d'un bruit blanc, l'archétype d'un signal non compressible. Comme ( $\rho < 1$ ), l'estimateur classique s'avère plus performant pour ce type de signaux. Par contre, lorsque  $|a| \rightarrow 1$ , le signal que l'on obtient devient de plus en plus compressible. L'estimateur compressé peut s'avérer significativement meilleur ( $\rho > 1$ ). Dans le meilleur des cas, le gain en variance est d'un facteur très proche du taux de compression ( $\rho \approx \alpha$ ).

Pour observer l'impact du choix de  $\Phi$  dans la performance, la figure 3.6 compare les différents rapports entre eux pour deux taux de compression différents  $\alpha = 2$  et  $\alpha = 10$ , à nouveau en fonction de  $a$ .

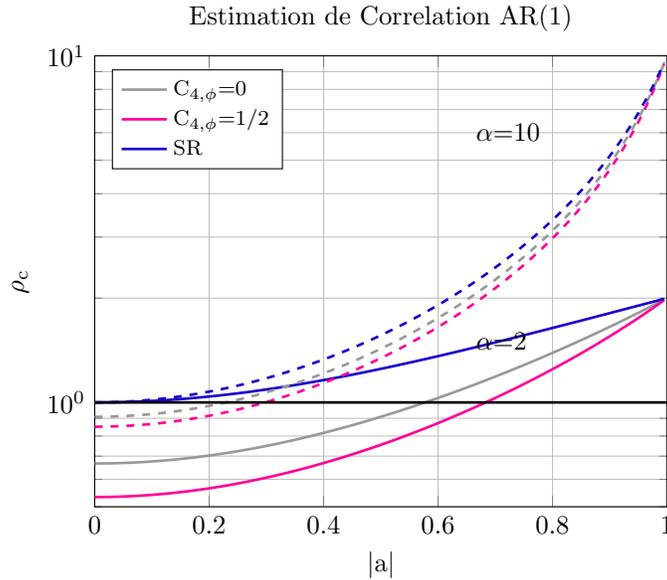


FIGURE 3.6 – Comparaison des comportements obtenus pour différents types de matrice : gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s_M = M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire.  $N/M = 10$  en trait pointillé et  $N/M = 2$  en continu. Le rapport de variance associé au cas gaussien se situe entre l'échantillonnage aléatoire (le plus performant) et le cas des matrices creuses (le moins performant). À taux de compression élevé ( $N/M = 10$ ), le choix de la matrice  $\Phi$  est peu influent.

On voit que si le taux de compression est faible ( $\alpha = 2$ ), on a un écart significatif entre les trois estimateurs. L'échantillonnage aléatoire est la méthode la plus performante (le rapport et donc le gain sont les plus élevés), suivie des matrices gaussiennes et enfin des matrices creuses. Cependant, à taux de compression élevé ( $\alpha = 10$ ), l'écart entre ces méthodes diminue, auquel cas le choix de la matrice a très peu d'influence dans la performance atteinte.

Passons maintenant à l'étude de l'influence de  $\tau$ . Pour cela, la figure 3.8 illustre le comportement des rapports en fonction de  $\tau$  lorsque le coefficient  $a$  est fixé. Trois cas déjà illustrés précédemment dans la figure 3.4 ( $a = 0.3, 0.6, 0.9$ ) sont considérés. Ces figures montrent que ces rapports présentent des variations selon la valeur de  $\tau$ , mais celles-ci sont peu significatives, les rapports restant quasiment constants pour toute valeur de  $\tau$ . Ceci permet de voir pourquoi le cas  $\tau = 0$  est représentatif.

Finalement, pour illustrer le comportement des estimateurs, la figure 3.7 présente quelques estimations de la fonction d'autocorrélation de deux processus AR(1)  $a = 0.2$  et  $a = 0.8$ . Ces valeurs correspondent respectivement à des signaux peu et moyennement compressibles. On présente également leurs statistiques empiriques, obtenues à partir des estimations. Nous comparons l'estimateur classique et l'estimateur compressé le moins performant (celui qui utilise une matrice creuse). On constate premièrement que ces deux estimateurs n'ont pas de biais, puisque la moyenne empirique des estimations coïncide avec la fonction de corrélation théorique. Deuxièmement, on observe que pour des signaux peu compressibles ( $a = 0.2$ ), ces deux estimateurs sont comparables. Par contre, pour un signal plus structuré ( $a = 0.8$ ), on constate que l'écart type empirique, et donc la variance, du classique est supérieur à celui du compressé, quelque soit la valeur du retard  $\tau$ .

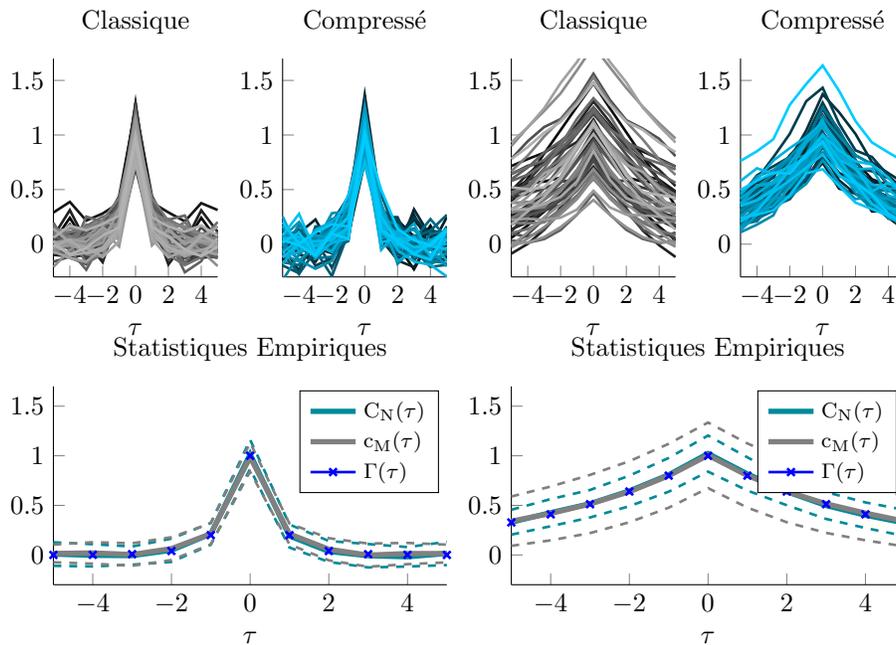


FIGURE 3.7 – Estimation de corrélation pour deux processus AR(1) Gaussiens. Gauche :  $a = 0.2$ , droite :  $a = 0.8$ . Pour ces simulations  $N = 1000, \alpha = 10$ . L'estimateur compressé utilise une matrice creuse. Haut : superposition de 50 estimations. Bas : statistiques empiriques, moyenne en trait continu, écart-type en pointillé. Pour un signal compressible, l'estimateur compressé s'avère plus performant que le classique, puisque son écart-type est inférieur.

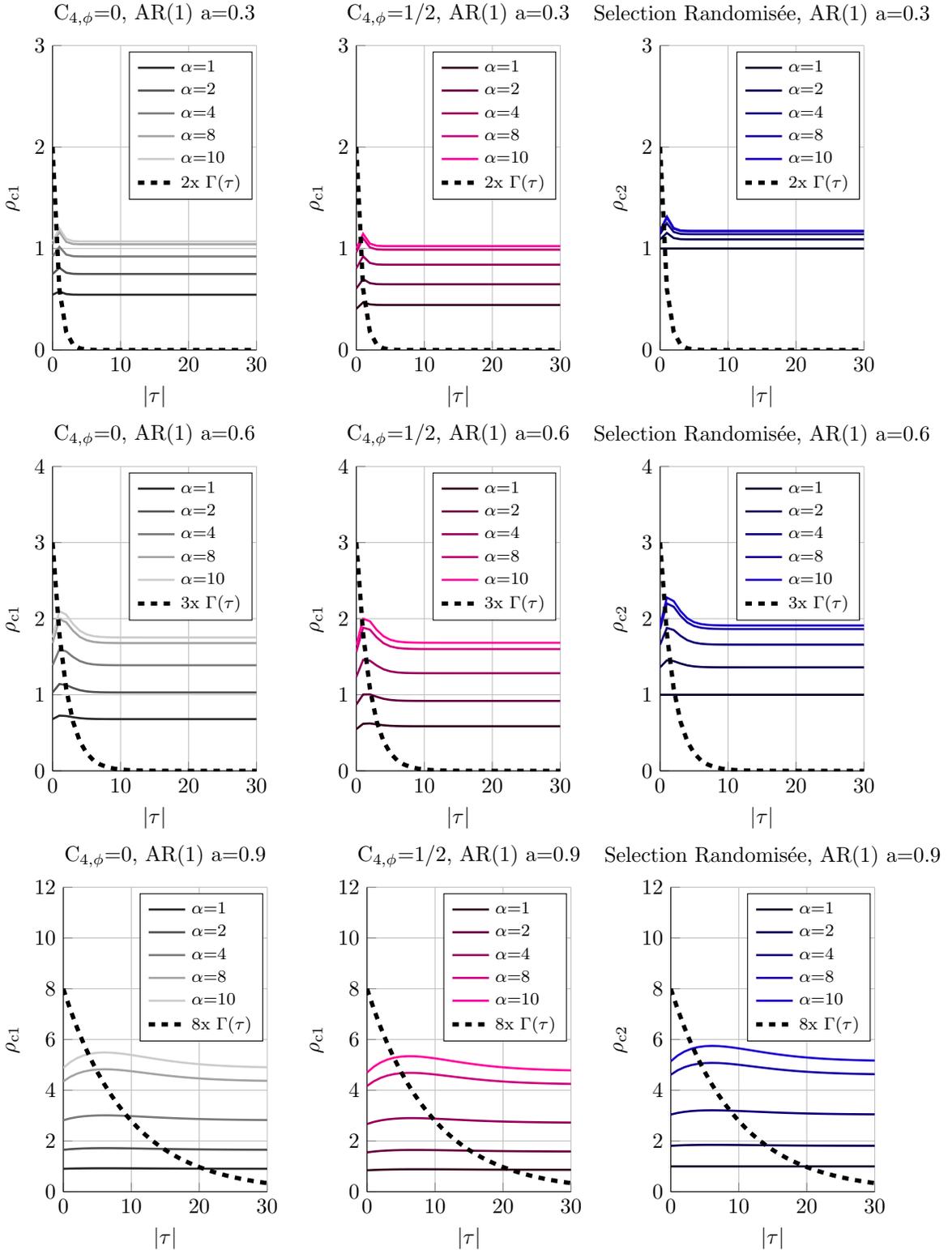


FIGURE 3.8 – Rapports de variance d'estimation de corrélation ( $\rho_{c1}, \rho_{c2}$ ) pour un processus AR(1) Gaussien en fonction de  $\tau$  et différents types de compression : matrice Gaussienne ( $C_{4,\phi} = 0$ ), matrice creuse avec  $s_M = M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. Puisque ( $\rho_{c1}, \rho_{c2} > 1$ ) l'estimateur compressé s'avère meilleur que le classique, quelle que soit la valeur de  $\tau$  pour des processus avec ces valeurs du coefficient  $a$ . Haut 0.3, milieu 0.6 et bas 0.9. Les gains en variance augmentent lorsque  $|a| \rightarrow 1$ . Ils atteignent respectivement  $\approx 1.1, 2$  et  $5$  pour  $\alpha = 10$ .

Ayant présenté le comportement des estimateurs compressés de corrélation, nous nous intéressons maintenant à l'étude de l'influence du bruit sur leurs performances.

### 3.2.7 Effet du bruit

Pour cette analyse, on considère que les deux signaux sont perturbés par des bruits additifs, notés  $\mathbf{b}_1$  et  $\mathbf{b}_2$ . Ces bruits sont supposés blancs, gaussiens, de moyenne nulle et variance  $\sigma_b^2$ , indépendants entre eux et indépendants des signaux. On écrit donc :

$$\mathbf{x}_b = \mathbf{x} + \mathbf{b}_1$$

$$\mathbf{y}_b = \mathbf{y} + \mathbf{b}_2$$

Pour simplifier, nous supposons le même niveau de bruit dans les deux signaux observés :  $\sigma_{b_1}^2 = \sigma_{b_2}^2 = \sigma_b^2$ .

#### Corrélateur classique

Dans le cas de l'estimateur classique en présence de bruit, noté  $\mathcal{C}_N^b(\tau)$ , on a :

$$N\mathcal{C}_N^b(\tau) = \mathbf{x}_b^\top \mathbf{y}_b = \mathbf{x}^\top \mathbf{y} + \mathbf{b}_1^\top \mathbf{y} + \mathbf{x}^\top \mathbf{b}_2 + \mathbf{b}_1^\top \mathbf{b}_2$$

*Biais*

$$N\mathbf{E}[\mathcal{C}_N^b(\tau)] = \mathbf{E}[\mathbf{x}^\top \mathbf{y}] + \mathbf{E}[\mathbf{b}_1^\top \mathbf{y}] + \mathbf{E}[\mathbf{x}^\top \mathbf{b}_2] + \mathbf{E}[\mathbf{b}_1^\top \mathbf{b}_2]$$

Comme les signaux sont indépendants des bruits, seule  $\mathbf{E}[\mathbf{x}^\top \mathbf{y}]$  est différente de zéro, le caractère non biaisé est donc préservé.

$$\mathbf{E}[\mathcal{C}_N^b(\tau)] = \mathbf{E}[\mathcal{C}_N(\tau)]$$

*Variance*

Dans le cas général, la variance d'une somme de variables aléatoires s'écrit :

$$\mathbf{Var} \left\{ \sum_i x_i \right\} = \sum_i \mathbf{Var}\{x_i\} + \sum_{i \neq j} \mathbf{Cov}\{x_i, x_j\}$$

pour  $\mathbf{Var}\{\mathcal{C}_N^b(\tau)\}$  on obtient :

$$N^2 \mathbf{Var}\{\mathcal{C}_N^b(\tau)\} = \mathbf{Var}\{\mathbf{x}^\top \mathbf{y}\} + \mathbf{Var}\{\mathbf{b}_1^\top \mathbf{y}\} + \mathbf{Var}\{\mathbf{x}^\top \mathbf{b}_2\} + \mathbf{Var}\{\mathbf{b}_1^\top \mathbf{b}_2\}$$

Les termes de covariance sont nuls à cause de l'indépendance entre les signaux et les bruits. Le premier terme de la somme obtenue correspond à la variance de cet estimateur en l'absence de bruit.

$$\frac{1}{N^2} \mathbf{Var}\{\mathbf{x}^\top \mathbf{y}\} = \mathbf{Var}\{\mathcal{C}_N(\tau)\}$$

Ensuite, pour les trois autres termes, on a :

$$\frac{1}{N^2} \mathbf{Var}\{\mathbf{b}_1^\top \mathbf{y}\} = \frac{1}{N} \Gamma_{yy}(0) \Gamma_{bb}(0) = \frac{\sigma_y^2 \sigma_b^2}{N}, \quad \frac{1}{N^2} \mathbf{Var}\{\mathbf{x}^\top \mathbf{b}_2\} = \frac{\sigma_x^2 \sigma_b^2}{N}$$

et

$$\frac{1}{N^2} \mathbf{Var}\{\mathbf{b}_1^\top \mathbf{b}_2\} = \frac{1}{N} \Gamma_{bb}^2(0) = \frac{\sigma_b^4}{N}$$

Finalement :

$$\mathbf{Var}\{\mathcal{C}_N^b(\tau)\} = \mathbf{Var}\{\mathcal{C}_N(\tau)\} + \frac{\sigma_x^2 \sigma_b^2 + \sigma_y^2 \sigma_b^2 + \sigma_b^4}{N}$$

### Corrélateur compressé

Pour l'estimateur compressé en présence de bruit, noté  $C_N^b(\tau)$ , on considère que les observations sont bruitées avant la compression, c'est-à-dire :

$$C_N^b(\tau) = (\Phi \mathbf{x}_b)^\top (\Phi \mathbf{y}_b) = (\Phi \mathbf{x})^\top (\Phi \mathbf{y}) + (\Phi \mathbf{b}_1)^\top (\Phi \mathbf{y}) \\ + (\Phi \mathbf{x})^\top (\Phi \mathbf{b}_2) + (\Phi \mathbf{b}_1)^\top (\Phi \mathbf{b}_2)$$

$$\mathbf{E}[C_N^b(\tau)] = \mathbf{E}[(\Phi \mathbf{x})^\top (\Phi \mathbf{y})] + \mathbf{E}[(\Phi \mathbf{b}_1)^\top (\Phi \mathbf{y})] \\ + \mathbf{E}[(\Phi \mathbf{x})^\top (\Phi \mathbf{b}_2)] + \mathbf{E}[(\Phi \mathbf{b}_1)^\top (\Phi \mathbf{b}_2)]$$

Puisque seule  $\mathbf{E}[(\Phi \mathbf{x})^\top (\Phi \mathbf{y})]$  est non nulle, le biais reste nul.

$$\mathbf{E}[C_N^b(\tau)] = \mathbf{E}[C_N(\tau)]$$

Ensuite,

$$\mathbf{Var}\{C_N^b(\tau)\} = \mathbf{Var}\{(\Phi \mathbf{x})^\top (\Phi \mathbf{y})\} + \mathbf{Var}\{(\Phi \mathbf{b}_1)^\top (\Phi \mathbf{y})\} \\ + \mathbf{Var}\{(\Phi \mathbf{x})^\top (\Phi \mathbf{b}_2)\} + \mathbf{Var}\{(\Phi \mathbf{b}_1)^\top (\Phi \mathbf{b}_2)\}$$

À nouveau, les termes de covariance sont nuls puisque les signaux et les bruits sont supposés indépendants.

Le premier terme correspond à la variance de cet estimateur en l'absence de bruit.

$$\mathbf{Var}\{(\Phi \mathbf{x})^\top (\Phi \mathbf{y})\} = \mathbf{Var}\{C_N(\tau)\}$$

Pour les trois autres termes, de simples substitutions dans l'expression de la variance permettent d'aboutir à :

$$\mathbf{Var}\{C_N^b(\tau)\} = \mathbf{Var}\{C_N(\tau)\} + \left( \frac{1}{N} + \frac{1}{MN} + \frac{\kappa[\varphi]}{MN} + \frac{1}{M} \right) (\sigma_x^2 \sigma_b^2 + \sigma_y^2 \sigma_b^2 + \sigma_b^4)$$

L'estimateur comparable à  $C_N^b(\tau)$  en termes de transmission est le classique sur  $M$  points, pour lequel :

$$\mathbf{Var}\{\mathcal{C}_M^b(\tau)\} = \mathbf{Var}\{\mathcal{C}_M(\tau)\} + \frac{1}{M} (\sigma_x^2 \sigma_b^2 + \sigma_y^2 \sigma_b^2 + \sigma_b^4)$$

À partir de ces deux dernières expressions, on voit que le bruit n'a pas le même effet dans les estimateurs. L'impact est plus fort sur l'estimateur compressé, car son terme de variance dû au bruit est  $\frac{1}{N} + \frac{1}{MN} + \frac{\kappa[\varphi]}{MN}$  fois plus grand que celui du classique. L'estimateur classique est donc plus robuste au bruit que le compressé. Ce fait est illustré par les courbes dans la figure 3.9 pour le cas d'un processus AR(1) Gaussien et deux taux de compression différents. Le rapport de variances  $\rho_c$ , qui exprime le gain en variance du compressé par rapport au classique, se dégrade ( $|\rho_c| \rightarrow 0$ ) au fur et à mesure que le niveau de bruit augmente (ou bien que le rapport signal sur bruit diminue).

Néanmoins, nous soulignons que lorsque le niveau de bruit est faible, voire moyen, l'estimateur compressé peut être pourtant plus performant que le classique.

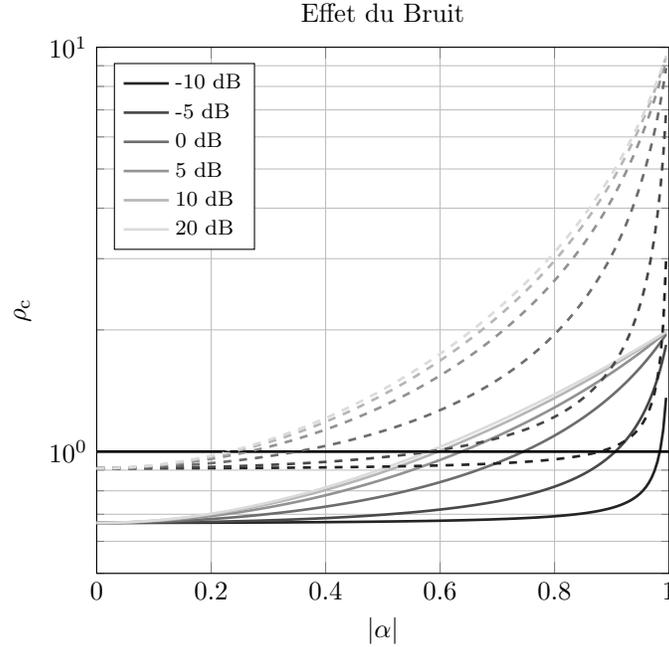


FIGURE 3.9 – Rapports de variance d’estimation de corrélation pour un AR(1) Gaussien en fonction de  $a$ . Différents niveaux de bruit additif sont étudiés pour deux taux de compression différents :  $\alpha = 2$  en trait continu et  $\alpha = 10$  en pointillé. Dans la zone où  $\rho_c > 1$  l’estimateur compressé s’avère plus performant que le classique. Cela peut arriver même en présence de bruit. Néanmoins l’estimateur compressé est plus sensible au bruit que le classique, puisque  $\rho_c$  diminue lorsque le niveau du bruit augmente.

Estimer une fonction de corrélation permet d’effectuer des calculs supplémentaires. Maintenant, nous nous focalisons sur l’étude des problèmes d’estimation pour lesquels l’estimation de la fonction de corrélation est un pas intermédiaire. Le plus courant est l’estimation du temps de propagation entre deux capteurs, étudié dans la section suivante.

### 3.3 Estimation du temps de propagation

Pour estimer un temps de propagation entre deux signaux  $x(t)$  et  $y(t)$ , à partir d’une estimation de la fonction de corrélation, il suffit de trouver l’argument pour lequel cette fonction présente son maximum.

$$\hat{\tau} = \underset{\tau}{\operatorname{argmax}} \hat{\Gamma}_{xy}(\tau)$$

Dans le cas général, la loi de probabilité de cette variable est inaccessible. Mais à temps continu, l’abscisse recherchée correspond également au point où la dérivée de la fonction de corrélation s’annule, pour peu que la fonction en question soit dérivable. Ce n’est évidemment pas le cas de la fonction à temps discret ! Toutefois, il est courant d’approcher la dérivée d’une telle fonction par la moyenne des incréments à gauche et à droite :

$$\hat{\Gamma}'_{xy}(\tau) \approx \frac{\hat{\Gamma}_{xy}(\tau + 1) - \hat{\Gamma}_{xy}(\tau - 1)}{2}$$

si bien que :

$$\operatorname{Var}(\hat{\Gamma}'(\tau)) = \frac{1}{4} \operatorname{Var}(\hat{\Gamma}_{xy}(\tau + 1)) + \frac{1}{4} \operatorname{Var}(\hat{\Gamma}_{xy}(\tau - 1)) - \frac{1}{2} \operatorname{Cov}(\hat{\Gamma}_{xy}(\tau - 1), \hat{\Gamma}_{xy}(\tau + 1)) \quad (3.7)$$

Faute de mieux, nous admettrons que cette quantité détermine la variance du temps de propagation estimé, même dans le cas particulier d'un processus AR(1), dont la fonction de corrélation n'est pourtant pas dérivable (si ce n'est au sens des distributions). Appliqué à l'autocorrélation d'un processus AR(1), dont le maximum apparaît en  $\tau = 0$ , on obtient respectivement

$$\mathbf{Var}\{\mathcal{C}_M'(0)\} = \frac{1}{4} \mathbf{Var}\{\mathcal{C}_M(1)\} + \frac{1}{4} \mathbf{Var}\{\mathcal{C}_M(-1)\} - \frac{1}{2} \mathbf{Cov}\{\mathcal{C}_M(1), \mathcal{C}_M(-1)\}$$

$$\mathbf{Var}\{C'_N(0)\} = \frac{1}{4} \mathbf{Var}\{C_N(1)\} + \frac{1}{4} \mathbf{Var}\{C_N(-1)\} - \frac{1}{2} \mathbf{Cov}\{C_N(1), C_N(-1)\}$$

Après de simples substitutions, on aboutit finalement à :

$$\begin{aligned} \mathbf{Var}\{\mathcal{C}_M'(0)\} &= \frac{1 + \alpha^2 - 2\alpha^{2M}}{2M^2} \\ \mathbf{Var}\{C'_N(0)\} &\approx \frac{1 - \alpha^2}{2M} \left(1 + \frac{\kappa[\varphi] + 2}{N}\right) + \frac{1 + \alpha^2 - 2\alpha^{2N}}{2N^2} \end{aligned}$$

La première expression décroît en  $\mathcal{O}(1/M^2)$  contre  $\mathcal{O}(1/M)$  pour la seconde. L'estimation de  $\hat{\tau}$  est donc plus performante dans le cas classique. Sa variance est toujours inférieure à celle établie à partir des signaux compressés. Cette affirmation est conforme aux résultats observés sur les simulations numériques présentées ci-dessous :

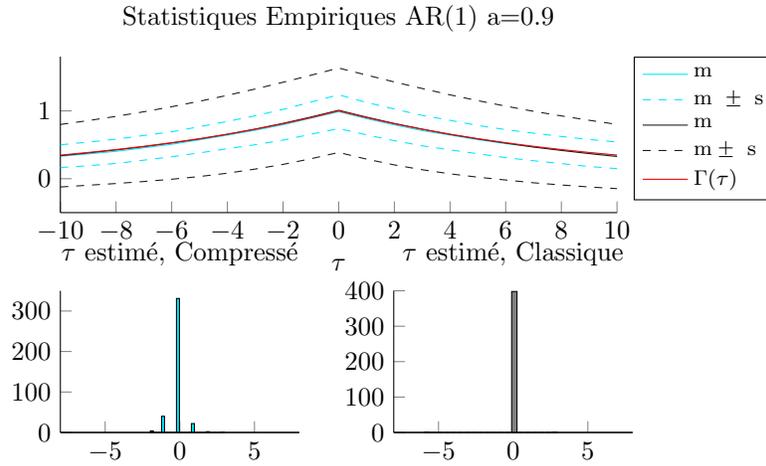


FIGURE 3.10 – Haut : Fonction de corrélation estimée et statistiques des estimateurs du temps de propagation  $\hat{\tau}$ . On constate que l'estimateur compressé de corrélation dégrade la performance d'estimation.

Ce résultat est paradoxal et contraire à ce que nous imaginions. Pour chaque retard  $\tau$  l'estimateur compressé est plus précis que le classique. Nous nous attendions donc à ce que le maximum de la fonction fluctue moins en compressé qu'en classique. C'était sans compter sur la dépendance entre les estimations évaluées pour des retards voisins !

En classique les erreurs pour des retards proches sont très corrélées. Tout se passe comme si toute la courbe se décalait vers le haut ou vers le bas. A l'inverse, en compressé, les erreurs sont moins dépendantes (Cf fig. 3.7), ce qui fait que l'abscisse du maximum peut fluctuer légèrement. C'est ce que signifie l'équation 3.7 : la somme des deux premiers termes est plus petite dans le cas compressé que dans le cas classique mais manque de chance, le terme soustrait dû à la covariance est beaucoup plus fort en classique qu'en compressé.

Estimer un temps de propagation, n'est pas le seul calcul secondaire envisageable, l'évaluation de la densité spectrale d'interaction ou interspectre est un autre problème d'estimation assez courant. Intéressons nous maintenant à l'étude des estimateurs de spectre d'interaction, construits à l'aide des estimateurs de corrélation précédemment évoqués.

### 3.4 Estimation d'interspectre

Il existe différentes manières d'estimer des densités spectrales de puissance et d'interaction des signaux aléatoires. Au lieu d'une description exhaustive des méthodes, déjà disponible dans la vaste littérature sur l'analyse spectrale, nous présentons dans la suite quelques généralités qui nous semblent pertinentes pour contextualiser notre étude sur l'estimation spectrale d'interaction. Le lecteur intéressé pourra par exemple se référer à [10, 6, 17, 11]. En particulier, [8, 3, 15] étudient en détail les propriétés des différents estimateurs d'interspectre.

Il y a deux grandes familles de méthodes d'estimation de spectre :

- Les méthodes **paramétriques**, plus récentes et qui possèdent les meilleures performances. Celles-ci s'appuient sur les connaissances a priori que l'on a sur le signal : on établit d'abord un modèle qui dépend d'un certain nombre de paramètres pour décrire le comportement du spectre et ensuite, on estime la valeur des paramètres au lieu d'estimer le spectre. Pour cette raison, leur performance est bonne à condition d'avoir des connaissances fiables et précises sur le signal.
- En contrepartie, les méthodes **non-paramétriques**, appelées aussi d'estimation spectrale classique, sont des méthodes qui estiment directement le spectre du signal. Leur performance n'est donc pas conditionnée par les a priori sur les signaux. Cette approche constitue un bon choix en l'absence de connaissances sur les signaux à traiter. Justement le type de contexte que nous envisageons.

Les méthodes non-paramétriques peuvent à leur tour être classées en deux groupes. D'une part les méthodes *directes* (ou périodogrammes), qui exploitent directement les transformées de Fourier des signaux, et d'autre part les méthodes *indirectes* (ou corrélogrammes), où l'estimation du spectre est précédée d'une estimation de la fonction de corrélation. Au niveau théorique ces deux méthodes sont souvent considérées comme équivalentes. En pratique il y a pourtant des différences importantes qui favorisent dans certains cas les corrélogrammes (par exemple, lorsque la fréquence des signaux est haute), le lecteur intéressé par cet aspect trouvera une discussion détaillée dans [12].

Or un aspect très important à souligner est que, dans la littérature, on retrouve très souvent l'estimateur biaisé standard de corrélation dans la définition du corrélogramme. Cependant, cet estimateur de corrélation n'est pas obligatoire. Cette approche indirecte permet d'estimer la corrélation avec d'autres méthodes. En fait, certaines méthodes qui estiment les fonctions de corrélation différemment ont été proposées dans le passé, par exemple à partir des signaux très fortement quantifiés. Suivant cette même logique, nous nous sommes intéressés à l'étude des propriétés des estimateurs d'interspectre que l'on obtient lorsque l'estimateur usuel de corrélation est remplacé par le corrélateur compressé.

#### 3.4.1 Corrélogramme Blackman-Tukey

Par définition, la densité spectrale d'interaction  $S_{xy}(\nu)$  est la transformée de Fourier de la corrélation croisée.

$$\gamma_{xy}(\nu) = \sum_{k \in \mathbb{Z}} \Gamma_{xy}(k) e^{-j2\pi\nu k} \quad (\text{où } j^2 = -1)$$

L'estimateur spectral étudié ici, est le corrélogramme de Blackman-Tukey (BT). Dans cette méthode, on calcule la transformée de Fourier de la fonction de corrélation estimée, souvent pondérée par une fenêtre symétrique  $w(k)$  de largeur **finie**  $2L + 1$ .

$$\hat{\gamma}_{BT}(\nu) = \sum_{k=-L}^L \hat{\Gamma}_{xy}(k) w(k) e^{-j2\pi\nu k}$$

Une règle empirique [11] attribuée à Blackman et Tukey, suggère que  $L$  soit choisi de l'ordre de  $N/10$ , si  $N$  désigne le nombre de points disponibles pour estimer la fonction de corrélation.

### 3.4.2 Corrélogramme classique

Nous parlerons du corrélogramme classique lorsque l'autocorrélation  $\Gamma_{xy}(\tau)$  est estimée avec le corrélateur classique.

Pour étudier les propriétés de cet estimateur, nous allons suivre l'analyse asymptotique proposée par Parzen, Hannan et Anderson [13, 5, 2]. Dans cette optique, le corrélogramme BT classique peut se réécrire de la manière suivante :

#### Corrélogramme classique

$$\mathcal{S}_M(\nu) = \sum_{k=-L}^L \mathcal{C}_M(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

Ici,  $w(x)$  est une fonction paire, bornée, de carré intégrable et normalisée, c'est-à-dire  $w(0) = 1$ .

Quelques exemples de fenêtres sont présentés dans le tableau suivant :

	$w\left(\frac{k}{L}\right)$	$\mathcal{S}_M(\nu)$
Rectangulaire	$\begin{cases} 1 &  k  \leq L \\ 0 &  k  > L \end{cases}$	$\sum_{k=-L}^L \mathcal{C}_M(k) e^{-j2\pi\nu k}$
Triangulaire	$\begin{cases} 1 - \frac{ k }{L} &  k  \leq L \\ 0 &  k  > L \end{cases}$	$\sum_{k=-L}^L \mathcal{C}_M(k) \left(1 - \frac{ k }{L}\right) e^{-j2\pi\nu k}$
Hamming	$\begin{cases} \frac{1}{2} \left(1 + \cos\left(\pi \frac{k}{L}\right)\right) &  k  \leq L \\ 0 &  k  > L \end{cases}$	$\sum_{k=-L}^L \frac{\mathcal{C}_M(k)}{2} \left(1 + \cos\left(\pi \frac{k}{L}\right)\right) e^{-j2\pi\nu k}$

D'après ce tableau, on voit que l'estimation de la fonction de corrélation (effectuée à partir de  $M$  points) est toujours limitée à  $2L + 1 < M$  valeurs. Différentes études asymptotiques sur cet estimateur [13, 5, 2] considèrent que ces deux paramètres,  $L$  et  $M$  doivent être liés tout en conservant  $L$  petit devant  $M$ , c'est-à-dire :  $M \rightarrow \infty$ ,  $L \rightarrow \infty$  mais  $\frac{L}{M} \rightarrow 0$ .

Dans le domaine fréquentiel, cet estimateur s'écrit :

$$\mathcal{S}_M(\nu) = \int_{-\pi}^{\pi} W_M(\nu - \lambda) \gamma_{xy}(\lambda) d\lambda$$

Lorsque  $M$  tend vers l'infini, la transformée de Fourier de ces fonctions  $W_M(\nu)$  se concentre autour de l'origine pour approcher la distribution de Dirac et

$$\mathcal{S}_M(\nu) \rightarrow \gamma_{xy}(\nu)$$

**Biais**

L'estimateur est biaisé

$$\mathbf{E}[\mathcal{S}_M(\nu)] = \gamma_{xy}(\nu) * W_M(\nu)$$

mais asymptotiquement non biaisé, lorsque  $M \rightarrow \infty$ ,  $L \rightarrow \infty$  et  $\frac{L}{M} \rightarrow 0$

$$\mathbf{E}[\mathcal{S}_M(\nu)] = \gamma_{xy}(\nu)$$

En annexe, on montre que :

**Covariance**

$$\lim_{M \rightarrow +\infty} \frac{M}{L} \mathbf{Cov}\{\mathcal{S}_M(\nu_1), \mathcal{S}_M(\nu_2)\} =$$

$$\begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ \gamma_{xx}(\nu)\gamma_{yy}(\nu) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 \neq 0, \pm \frac{1}{2} \\ (|\gamma_{xy}(\nu)|^2 + \gamma_{xx}(\nu)\gamma_{yy}(\nu)) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 = 0, \pm \frac{1}{2} \end{cases}$$

**Variance**

$$\lim_{M \rightarrow +\infty} \frac{M}{L} \mathbf{Var}\{\mathcal{S}_M(\nu)\} = \lim_{M \rightarrow +\infty} \frac{M}{L} \mathbf{Cov}\{\mathcal{S}_M(\nu), \mathcal{S}_M(\nu)\} =$$

$$\begin{cases} \gamma_{xx}(\nu)\gamma_{yy}(\nu) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu \neq 0, \pm \frac{1}{2} \\ (|\gamma_{xy}(\nu)|^2 + \gamma_{xx}(\nu)\gamma_{yy}(\nu)) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu = 0, \pm \frac{1}{2} \end{cases}$$

**3.4.3 Corrélogramme compressé**

De même, lorsque la fonction de corrélation est estimée à l'aide du corrélateur compressé

**Corrélogramme compressé**

Par projection aléatoire

$$S_N(\nu) = \sum_{k=-L}^L C_N(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

A nouveau,  $w$  est une fonction paire, bornée, de carré intégrable avec  $w(0) = 1$ .

**Biais**

Cet estimateur est biaisé :

$$\mathbf{E}[S(\nu)] = \mathbf{E}[\hat{S}(\nu)] = \gamma_{xy}(\nu) * W(\nu)$$

mais asymptotiquement non biaisé lorsque  $N, L$  et  $M \rightarrow \infty$ ,  $\frac{N}{M} = \alpha$ , et  $\frac{L}{M} \rightarrow 0$

$$\mathbf{E}[S(\nu)] = \mathbf{E}[\hat{S}(\nu)] = \gamma_{xy}(\nu)$$

**Covariance**

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{S_N(\nu_1), S_N(\nu_2)\} &= \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \\ &+ \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ (C_{4,\varphi} + \alpha)\sigma_x^2 \gamma_{yy}(\nu) \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 = \nu \end{cases} \end{aligned}$$

**Variance**

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{S_N(\nu)\} =$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_N(\nu)\} + (C_{4,\varphi} + \alpha)\sigma_x^2 \gamma_{yy}(\nu) \left[ \int_{-1}^1 w^2(x) dx \right]$$

**3.4.4 Corrélogramme compressé par échantillonnage aléatoire****Corrélogramme compressé par échantillonnage aléatoire sans remise**

$$\hat{S}_N(\nu) = \sum_{k=-N}^N \hat{C}_N(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

où  $w$  est une fonction paire, bornée, de carré intégrable avec  $w(0) = 1$ .

**Covariance**

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{\hat{S}_N(\nu_1), \hat{S}_N(\nu_2)\} &= \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \\ &+ \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ (\alpha - 1)\sigma_x^2 \gamma_{yy}(\nu) \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 = \nu \end{cases} \end{aligned}$$

## Variance

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathring{S}_N(\nu)\} =$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_N(\nu)\} + (\alpha - 1) \sigma_x^2 \gamma_{yy}(\nu) \left[ \int_{-1}^1 w^2(x) dx \right]$$

## 3.4.5 Compressibilité et performance

## Cas AR(1)

En vue de comprendre le fonctionnement des estimateurs d'interspectre selon la nature du signal traité, nous étudions leur comportement pour des processus autorégressifs d'ordre 1. Rappelons qu'un tel processus, s'écrit :

$$y_t = x_t = ax_{t-1} + c_1 \epsilon_t$$

$$\text{où } \sigma_\epsilon^2 = 1 \text{ et } \sigma_x^2 = \frac{c_1^2 \sigma_\epsilon^2}{1 - a^2}$$

Pour normaliser la puissance moyenne du processus ( $\sigma_x^2 = 1$ ), il suffit de fixer  $c_1 = \sqrt{1 - a^2}$ . C'est le choix que nous faisons pour toute la suite de ce paragraphe. La densité spectrale de puissance de ce processus s'écrit[16] :

$$\gamma_{xx}(\nu) = \frac{1 - a^2}{1 - 2a \cos(2\pi\nu) + a^2}, \quad |\nu| < 1/2$$

Les variances asymptotiques des différents corrélogrammes sont :

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_N(\nu)\} = g(\nu)$$

$$= \delta_\nu \frac{(1 - a^2)^2}{(1 - 2a \cos(2\pi\nu) + a^2)^2} \int_{-\infty}^{\infty} w^2(x) dx$$

$$\text{où } \delta_\nu = \begin{cases} 1 & \text{si } \nu \neq 0, \pm \frac{1}{2} \\ 2 & \text{si } \nu = 0, \pm \frac{1}{2} \end{cases}$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_M(\nu)\} = \alpha g(\nu)$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{S_N(\nu)\} = g(\nu) + \alpha \left( \frac{1 - a^2}{1 - 2a \cos(2\pi\nu) + a^2} \right) + C_{4,\varphi} \left( \frac{2(1 - a^2)}{1 - 2a \cos(2\pi\nu) + a^2} \right)$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathring{S}_N(\nu)\} = g(\nu) + (\alpha - 1) \left( \frac{1 - a^2}{1 - 2a \cos(2\pi\nu) + a^2} \right)$$

Comme précédemment, nous étudions les rapports des variances.  $\mathbf{Var}_\infty$  dénote variance asymptotique.

$$\rho_S = \frac{\mathbf{Var}_\infty\{\mathcal{S}_M(\nu)\}}{\mathbf{Var}_\infty\{S_N(\nu)\}} \text{ et } \hat{\rho}_S = \frac{\mathbf{Var}_\infty\{\mathcal{S}_M(\nu)\}}{\mathbf{Var}_\infty\{\mathring{S}_N(\nu)\}}$$

A nouveau, ces trois estimateurs ont un même coût en termes de transmission de données. Ils requièrent tous d'échanger  $M$  valeurs numériques d'un capteur à l'autre pour effectuer l'estimation.

Nous rappelons que dans le cas de l'estimateur de corrélation, le comportement des rapports de variance était quasiment constant pour toute valeur du retard  $\tau$ . Cela n'est pas le cas de l'estimateur spectral. La performance de celui-ci est différente selon la valeur de la fréquence  $\nu$ . Pour le montrer, le comportement de ces rapports de variance en fonction de  $\nu$  est illustré, dans la figure 3.12, pour les cas considérés précédemment ( $a = 0.3, 0.6, 0.9$ ). Dans cette figure, ces rapports (dessinés en trait continu) sont superposés au spectre théorique du signal (en trait pointillé). On constate l'influence de la densité spectrale du processus étudié : dans la région spectrale où la puissance du processus se concentre,  $\rho$  peut devenir supérieur à 1, auquel cas l'estimateur compressé s'avère plus performant que le classique. Par contre, la tendance s'inverse ( $\rho < 1$ ) dans la zone du spectre où la puissance du processus est faible. À ces endroits l'estimateur classique reste plus performant. Notons que plus le processus est compressible ( $|a| \rightarrow 1$ ), plus le gain en variance est important. Il atteint par exemple  $\approx 2.5$  si  $a = 0.6$  et  $\approx 6$  si  $a = 0.9$  pour un taux de compression de 10.

Or ce comportement de l'estimateur compressé n'est pas forcément un gros inconvénient en pratique. Premièrement parce que dans la région où le spectre concentre sa puissance (celle qui porte le plus d'information sur le processus) l'estimation est plus précise dans le cas compressé. Deuxièmement parce que dans la région où la puissance du processus est faible, le rapport de variance lui-même peut devenir élevé mais les variances sont deux quantités qui auront des valeurs faibles. Ce fait est montré dans la figure 3.11, qui illustre le comportement des variances asymptotiques. Dans ce sens l'estimateur compressé du spectre permet de diminuer l'erreur dans la région qui porte le plus de puissance et le prix à payer est une augmentation modérée de l'imprécision dans la zone où le spectre est faible. Dans cette étude, nous avons délibérément exclu les cas  $\nu = 0$  et  $0.5$  pour lesquels l'estimateur classique présente une variance deux fois plus grande que pour les autres fréquences. Puisque le modèle AR(1) concentre inévitablement son énergie dans l'une de ces deux fréquences lorsque  $|a|$  tend vers 1, l'étude sur l'effet de la compressibilité des signaux sur la performance est effectuée pour un autre modèle, un processus AR(2).

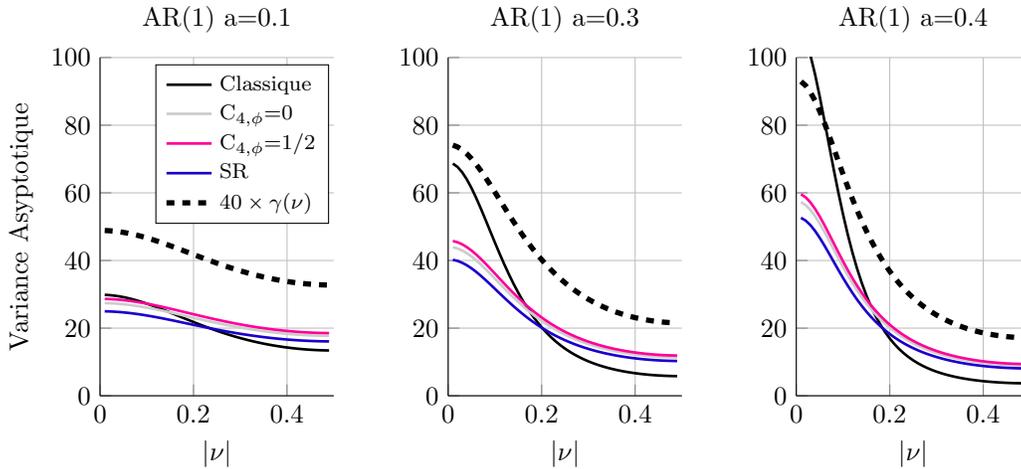


FIGURE 3.11 – Variance d'estimation du spectre, pour différents types de compression : Gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s=M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. L'estimateur compressé peut s'avérer plus performant (moins de variance) mais uniquement dans la région du spectre où le signal concentre sa puissance. Cependant dans la région où l'estimateur classique est meilleur, les deux variances ont une valeur faible. L'effet de la matrice de compression est peu significatif.

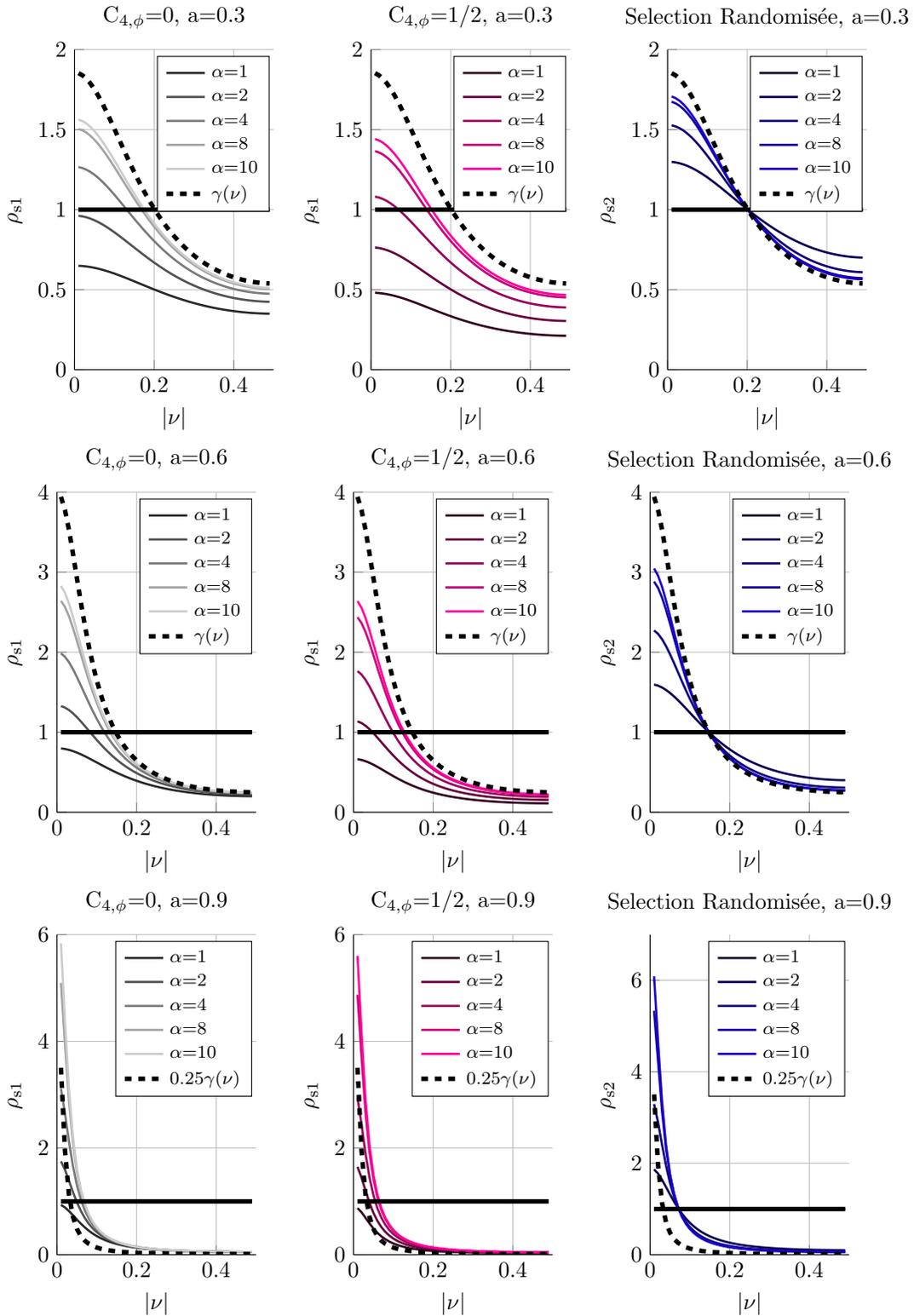


FIGURE 3.12 – Rapports de variance en fonction de  $\nu$  pour trois processus AR(1) (Haut :  $a = 0.3$  milieu :  $a = 0.6$ , bas :  $a = 0.9$ ) et différents types de compression : ( $C_{4,\phi} = 0$ ), ( $C_{4,\phi} = 1/2$ ) en échantillonnage aléatoire. L'estimateur compressé peut s'avérer plus performant ( $\rho > 1$ ), mais uniquement dans la région du spectre où le signal concentre sa puissance. La tendance est opposée dans la région du spectre où la puissance du processus est faible. Le gain en variance est plus important pour des signaux à spectre étroit, donc plus compressibles.

**Cas AR(2)**

Dans ce deuxième cas étudié, deux paramètres  $a_1$  et  $a_2$  permettent de piloter le comportement du signal.

$$y_t = x_t = a_2 x_{t-2} + a_1 x_{t-1} + \sqrt{G} \epsilon_t$$

où  $\sigma_\epsilon^2 = 1$  et

$$G = \frac{(1+a_2)(1-a_1-a_2)(1+a_1-a_2)}{(1-a_2)} > 0$$

Ainsi, le processus  $x$  est de puissance moyenne unitaire  $\sigma_x^2 = 1$ . La densité spectrale de puissance du processus AR(2) s'écrit[16] :

$$\gamma_{xx}(\nu) = \frac{G}{|1 - a_1 e^{-j2\pi\nu} - a_2 e^{-j4\pi\nu}|^2} \quad |\nu| < 1/2$$

Sa fonction de corrélation, s'écrit sous la forme :

$$\Gamma_{xx}(\tau) = A_1(\lambda_1)^\tau + A_2(\lambda_2)^\tau \quad A_1, A_2 \in \mathbb{R}$$

Pour des coefficients  $a_1$  et  $a_2$  réels,  $\lambda_1$  et  $\lambda_2$  s'écrivent :

$$\lambda_1, \lambda_2 = \frac{a_2 \pm \sqrt{a_1^2 + 4a_2}}{2}$$

De sorte que pour ce processus,  $\Gamma_{xx}(\tau)$  consiste en la somme de deux exponentielles décroissantes ( $\lambda_1, \lambda_2$  réels), ou bien deux sinusoides amorties ( $\lambda_1, \lambda_2$  complexes conjugués).

Or dans le cas du processus AR(1), nous avons modifié la nature du signal, sa *compressibilité*, à l'aide du coefficient  $a$ . Pour le processus AR(2), nous allons le faire en nous limitant au cas de  $\lambda$  complexes conjugués, soit sous forme polaire :  $\lambda_{1,2} = r e^{\pm j\theta}$ . Ici,  $\theta$  détermine la valeur de la fréquence centrale du spectre du signal, alors que  $r$  joue un rôle semblable à celui de  $|a|$  pour le modèle AR(1). Pour  $r = 0$ , on retrouve un bruit blanc (un signal non compressible) tandis que quand  $r$  tend vers 1, on obtient un processus qui concentre de plus en plus son énergie sur la fréquence  $\theta$  (dans ce sens un signal plus compressible). La figure 3.13 illustre les processus obtenus pour  $\theta = 0.125$  et différentes valeurs de  $r$ .

Passons maintenant à l'étude des rapports de variance. Pour un tel processus, les variances asymptotiques des différents estimateurs de spectre sont :

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_N(\nu)\} =$$

$$\delta_\nu \frac{G^2}{|1 - a_1 e^{-j2\pi\nu} - a_2 e^{-j4\pi\nu}|^4} \int_{-\infty}^{\infty} w^2(x) dx$$

$$\text{où } \delta_\nu = \begin{cases} 1 & \text{si } \nu \neq 0, \frac{1}{2} \\ 2 & \text{si } \nu = 0, \frac{1}{2} \end{cases}$$

$$= g(\nu)$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\mathcal{S}_M(\nu)\} = \alpha g(\nu)$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{S_N(\nu)\} = g(\nu) + \alpha \frac{G}{|1 - a_1 e^{-j2\pi\nu} - a_2 e^{-j4\pi\nu}|^2} + C_{4,\varphi} \frac{2G}{|1 - a_1 e^{-j2\pi\nu} - a_2 e^{-j4\pi\nu}|^2}$$

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{\hat{S}_N(\nu)\} = g(\nu) + (\alpha - 1) \frac{G}{|1 - a_1 e^{-j2\pi\nu} - a_2 e^{-j4\pi\nu}|^2}$$

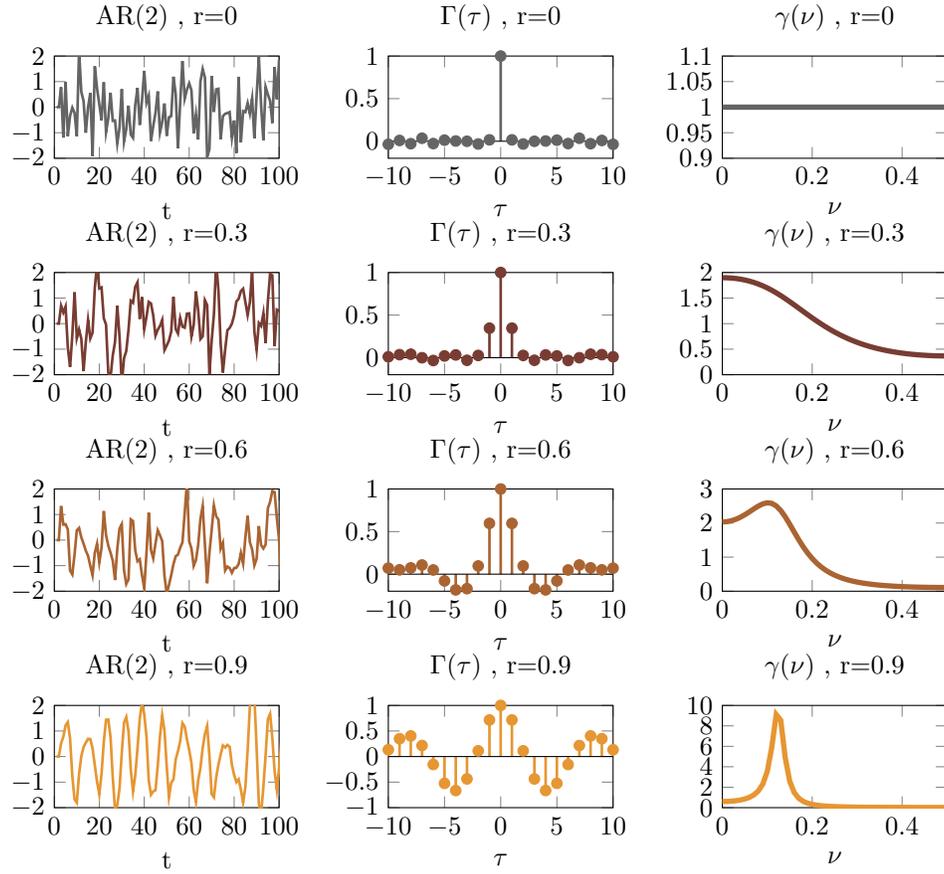


FIGURE 3.13 – Processus AR(2) gaussien pour différentes valeurs de  $r$ . Gauche : réalisations des différents processus considérés. Centre : fonction d'autocorrélation : Pour  $r = 0$ , on retrouve un bruit blanc. Lorsque  $|r| \rightarrow 1$  les valeurs successives du processus sont de plus en plus corrélées. Droite : Densité spectrale de puissance, pour  $r = 0$ , le processus possède un support fréquentiel plein. Par contre, lorsque  $r \rightarrow 1$  le support fréquentiel diminue, dans ce sens, un signal qui porte moins d'information et plus approprié pour la compression.

La figure 3.16 illustre l'effet de la nature du signal dans la performance d'estimation pour le cas des processus AR(2). On retrouve les mêmes observations que pour les AR(1). D'abord, la performance de l'estimateur compressé peut s'avérer meilleure que celle de l'estimateur classique, mais uniquement dans la zone qui concentre la puissance du signal. Par exemple le gain en variance atteint  $\approx 2$  si  $r = 0.6$  et  $\approx 5$  si  $r = 0.9$ , pour un taux de compression  $\alpha = 10$ .

Ensuite, la performance de l'estimateur compressé est meilleure pour des signaux compressibles, ici lorsque  $r$  approche 1. D'autre part, dans la figure 3.14 on voit également que même si les rapports de variances peuvent s'avérer défavorables pour l'estimateur compressé, les variances elles-mêmes ont des valeurs faibles et n'ont que peu d'effet dans l'estimation du spectre. Finalement la figure 3.15 montre le gain maximal en variance en fonction de  $r$ , pour deux taux de compression différents  $\alpha = 2$  et 10. Le type de compression détermine la performance, mais l'écart diminue à taux de compression élevé.

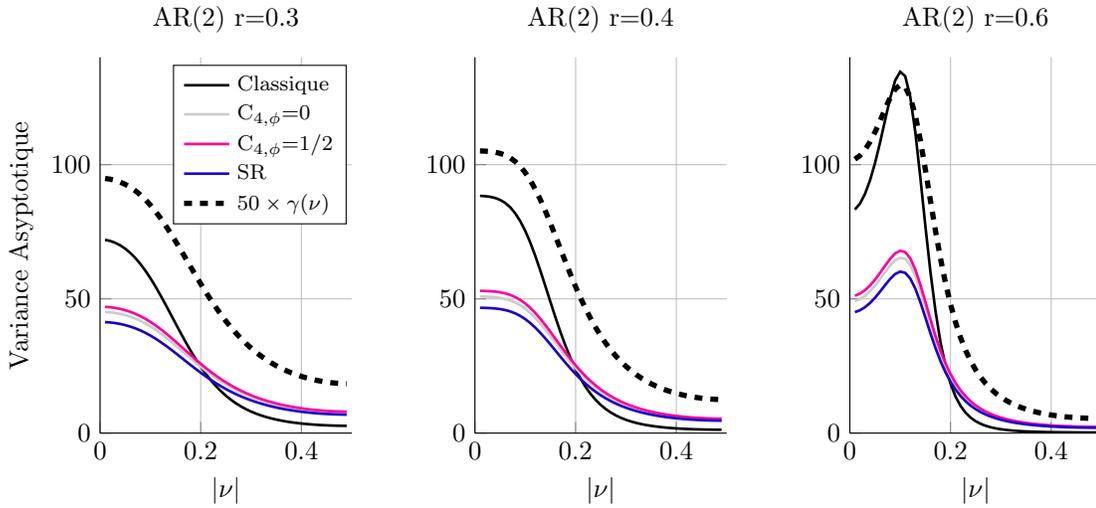


FIGURE 3.14 – Variance d’estimation du spectre pour différents types de compression : Gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s=M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. Deux taux de compression sont considérés :  $N/M = 10$  en trait pointillé et  $N/M = 2$  en continu. Le gain associé au cas Gaussien se situe entre l’échantillonnage aléatoire (le plus performant) et le cas creux (le moins performant). À taux de compression élevé ( $N/M = 10$ ), l’effet de la matrice de compression est peu significatif.

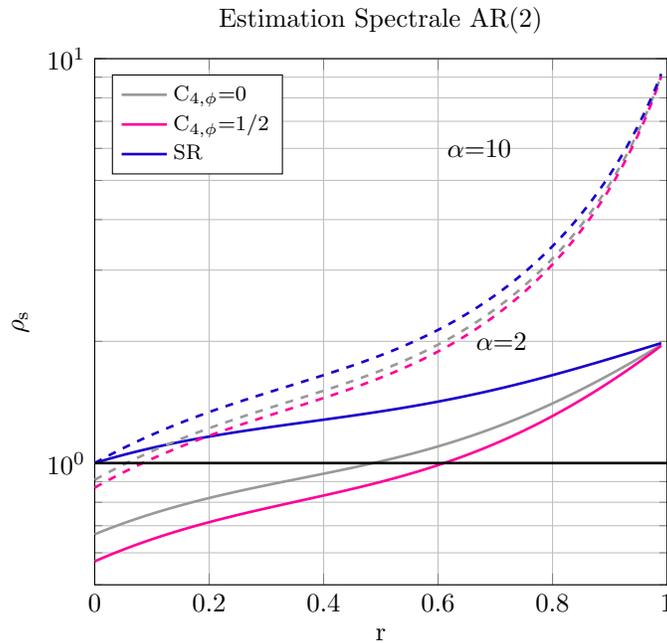


FIGURE 3.15 – Comparaison des rapports de variance d’estimation du spectre, pour différents types de compression : Gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s=M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. Deux taux de compression sont considérés :  $N/M = 10$  en trait pointillé et  $N/M = 2$  en continu. Le gain associé au cas Gaussien se situe entre l’échantillonnage aléatoire (le plus performant) et le cas creux (le moins performant). À taux de compression élevé ( $N/M = 10$ ), l’effet de la matrice de compression est peu significatif.

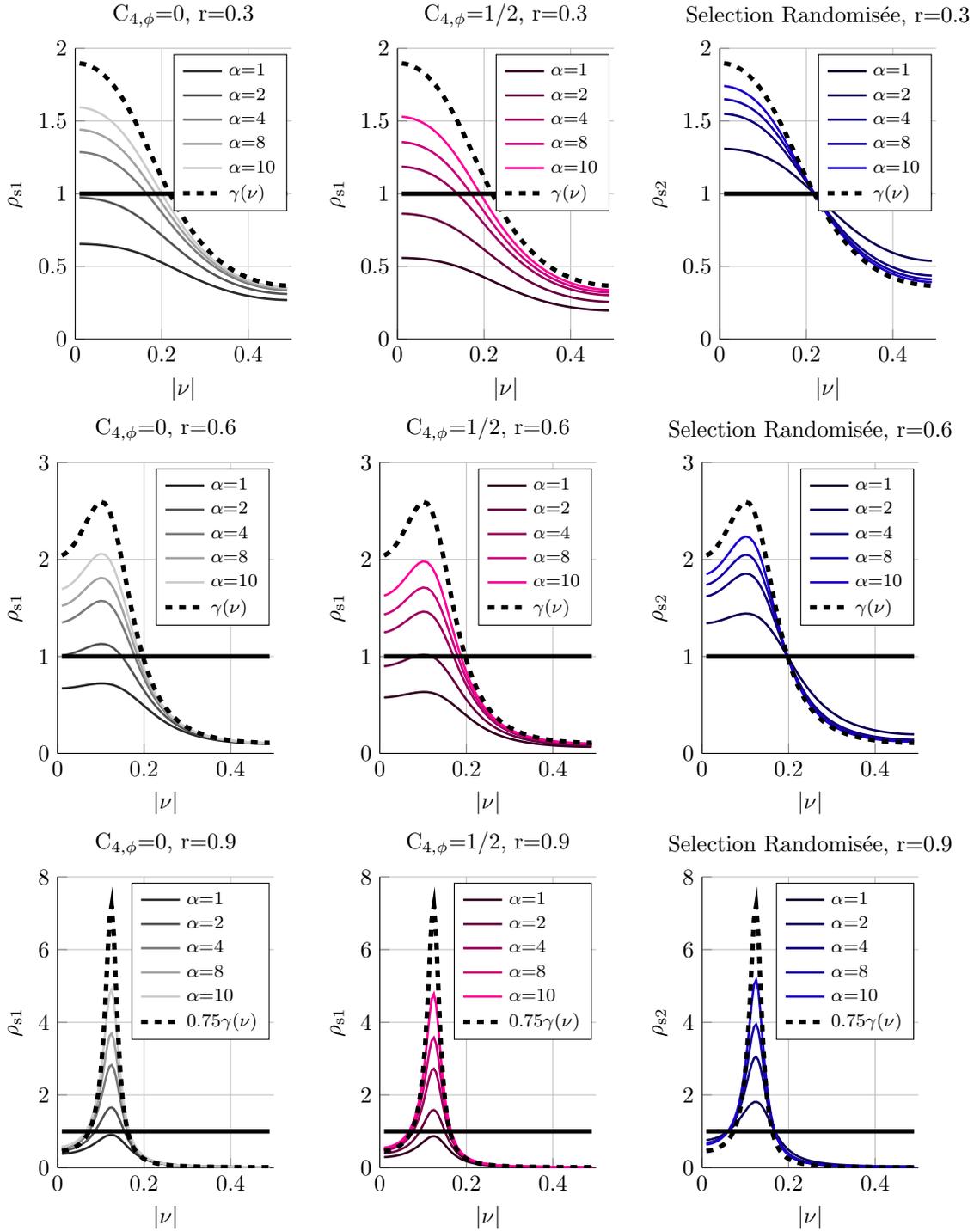


FIGURE 3.16 – Rapports de variance en fonction de  $\nu$  pour un AR(2) ( $r = 0.3$ ) et différents types de compression : Gaussienne ( $C_{4,\phi} = 0$ ), creuse avec  $s=M/2$  ( $C_{4,\phi} = 1/2$ ) et échantillonnage aléatoire. L'estimateur compressé peut s'avérer plus performant ( $\rho > 1$ ), mais uniquement dans la région du spectre où le signal concentre sa puissance, la tendance est opposée dans dans la région du spectre où la puissance du processus est faible.

Finalement, pour conclure cette présentation et montrer le comportement des estimateurs, la figure 3.17 présente la superposition de quelques estimations du spectre d'un

processus AR(2) avec  $r = 0.85$  et  $\nu = 0.125$ . On considère  $N = 2000$ ,  $\alpha = 10$  et  $L = 10$  avec une fenêtre de Hamming. Dans la partie haute, en échelle linéaire on voit que, comme prévu, l'estimateur compressé a une variance plus faible dans la zone qui concentre la puissance du processus. En bas de cette figure on retrouve la zone à faible puissance en échelle logarithmique. On constate que l'estimateur classique, ayant une variance plus faible, est meilleur dans cette région. Cependant, l'influence globale est peu significative. On constate également par ces simulations que, à nombre fini d'échantillons, l'estimateur compressé possède aussi un biais plus faible.

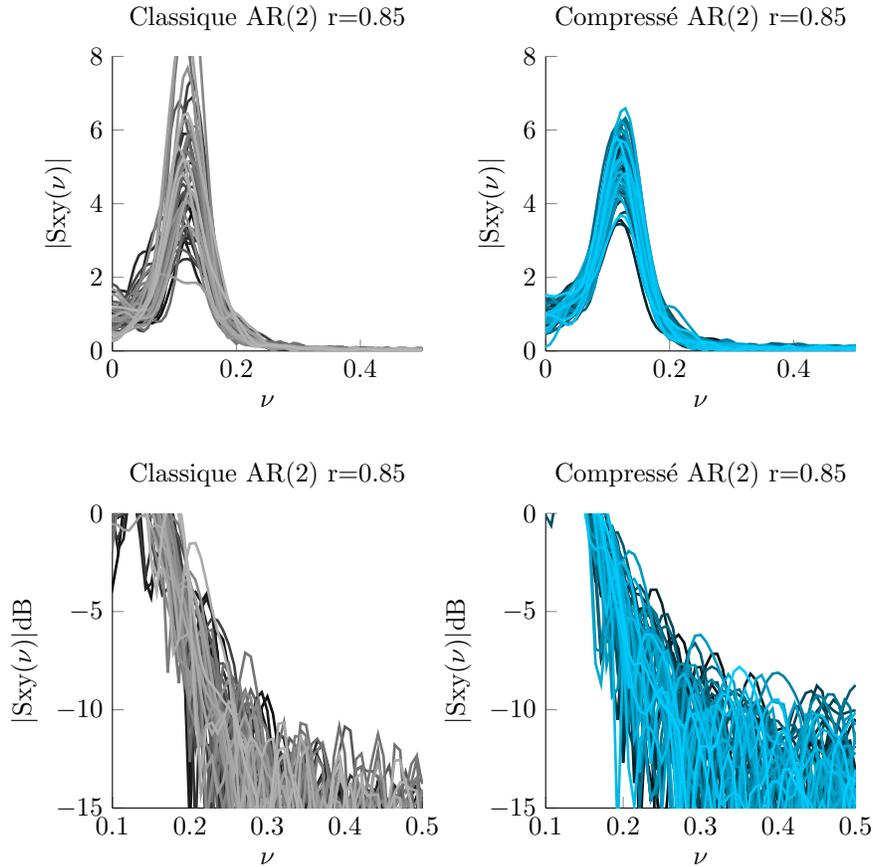


FIGURE 3.17 – Estimation spectrale pour un processus AR(2) Gaussien.  $r = 0.85$  et  $\nu = 0.125$ ,  $N = 2000$ ,  $\alpha = 10$  et  $L = 50$  avec une fenêtre de Hamming. L'estimateur compressé utilise une matrice creuse. Superposition de 50 corrélogrammes. Haut : échelle linéaire. Bas : Région à faible puissance, en échelle logarithmique. De manière globale, l'estimateur compressé s'avère plus précis que le classique.

### Conclusion

Dans ce chapitre, nous avons proposé des méthodes d'estimation qui s'appuient sur le *traitement du signal compressé* et constaté qu'il s'agit d'une méthode *aveugle* qui permet de traiter le signal sans le connaître. Cependant, comme nous l'avons observé, une méthode compressée n'est pas toujours le meilleur des choix. Il faut que le signal ait un certain degré de redondance (ce que nous avons identifié comme la *compressibilité*), plus il y a de redondance, meilleurs sont les résultats obtenus. Cette conclusion nous semble assez intuitive. Nous verrons au chapitre 6 qu'il existe des signaux réels compressibles .

A l'inverse, les résultats que nous avons observés pour l'estimation du temps de

propagation nous montrent que la compressibilité ne garantit pas de bonnes performances dans le domaine compressé.

Enfin, un autre point qui mérite d'être souligné est la très grande flexibilité quant au choix de la matrice de compression  $\Phi$  (lorsque l'on souhaite un taux de compression élevé). À ce stade, nous pourrions nous contenter d'étudier uniquement l'échantillonnage aléatoire, dont le coût calculatoire est négligeable et dont la performance est la meilleure. Pourtant, comme nous le décrirons au chapitre 4, la combinaison des échantillons d'un signal s'avère avantageuse pour la quantification.

La présentation des outils de traitement des signaux fortement quantifiés fait l'objet du chapitre suivant.

## Bibliographie

- [1] ACHLIOPTAS, D. Database-friendly random projections. In *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (2001), ACM, pp. 274–281.
- [2] ANDERSON, T. W. *The statistical analysis of time series*, vol. 19. John Wiley & Sons, 2011.
- [3] BENDAT, J. S., AND PIERSOL, A. G. *Random data : analysis and measurement procedures*, vol. 729. John Wiley & Sons, 2011.
- [4] BOX, G. E., JENKINS, G. M., REINSEL, G. C., AND LJUNG, G. M. *Time series analysis : forecasting and control*. John Wiley & Sons, 2015.
- [5] HANNAN, E. J. *Multiple time series*, vol. 38. John Wiley & Sons, 2009.
- [6] HAYES, M. H. *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.
- [7] HYVÄRINEN, A., KARHUNEN, J., AND OJA, E. *Independent component analysis*, vol. 46. John Wiley & Sons, 2004.
- [8] JENKINS, G. M., AND WATTS, D. G. *Spectral analysis*. Holden-day, 1968.
- [9] LI, P., HASTIE, T. J., AND CHURCH, K. W. Very sparse random projections. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (2006), ACM, pp. 287–296.
- [10] MANOLAKIS, D. G., INGLE, V. K., AND KOGON, S. M. *Statistical and adaptive signal processing : spectral estimation, signal modeling, adaptive filtering, and array processing*. McGraw-Hill Boston, 2000.
- [11] MARPLE, S. L., AND MARPLE, S. L. *Digital spectral analysis : with applications*, vol. 5. Prentice-Hall Englewood Cliffs, NJ, 1987.
- [12] MAX, J., DIOT, M., AND BIGRET, R. Les analyseurs de spectre à fft et les analyseurs de spectre à corrélation. *TS Traitement du signal* 3, 4-5 (1986), 241–256.
- [13] PARZEN, E. On consistent estimates of the spectrum of a stationary time series. *The Annals of Mathematical Statistics* (1957), 329–348.
- [14] PARZEN, E. Mathematical considerations in the estimation of spectra. *Technometrics* 3, 2 (1961), 167–190.
- [15] PRIESTLEY, M. B. *Spectral analysis and time series*. Academic press, 1981.
- [16] SHANMUGAN, K. S., AND BREIPOHL, A. M. *Random signals : detection, estimation, and data analysis*. Wiley, 1988.
- [17] STOICA, P., AND MOSES, R. L. *Introduction to spectral analysis*, vol. 1. Prentice hall Upper Saddle River, NJ, 1997.

# Chapitre 4

## Traitement des signaux quantifiés

### Sommaire

---

<b>4.1</b>	<b>Quantification orientée vers la tâche . . . . .</b>	<b>61</b>
<b>4.2</b>	<b>Corrélateurs fortement quantifiés . . . . .</b>	<b>62</b>
4.2.1	Corrélateurs avec des signaux à deux niveaux . . . . .	62
4.2.2	Corrélateurs hybrides . . . . .	63
4.2.3	Corrélateurs généralisés et modifiés . . . . .	65
<b>4.3</b>	<b>Acquisition compressive et quantification . . . . .</b>	<b>67</b>
<b>4.4</b>	<b>Corrélogrammes efficaces . . . . .</b>	<b>68</b>
4.4.1	Densité spectrale de puissance . . . . .	68
4.4.2	Densité spectrale d'interaction . . . . .	69
	<b>Bibliographie . . . . .</b>	<b>70</b>

---

Le traitement des signaux compressés, présenté dans le chapitre précédent, est l'un des outils permettant de diminuer dans une certaine mesure le coût lié à la transmission. Ce chapitre vise à présenter le deuxième outil : le traitement des signaux issus d'une quantification très forte, autrement dit des signaux dont les valeurs numériques sont représentées avec très peu de bits, voire un seul bit.

Procéder ainsi réduit la taille des données et donc le coût de la transmission de manière significative, mais produit une représentation grossière du signal, qui n'est évidemment pas souhaitable si c'est le signal lui-même qui nous importe. Par contre, lorsque l'on ne s'intéresse qu'à certaines propriétés du signal, cette *quantification orientée vers la tâche* s'avère très intéressante, puisqu'elle permet de réduire non seulement la taille des données, mais également la complexité des calculs et du matériel dont on a besoin pour le traitement. Dans ce chapitre on parlera des informations qui sont préservées lors d'une quantification très grossière, ainsi que des méthodes permettant de les extraire.

### 4.1 Quantification orientée vers la tâche

Dans le passé, lorsque les appareils électroniques avaient des capacités très limitées, l'étude des méthodes de traitement des signaux fortement quantifiés a permis de tirer le meilleur parti du matériel disponible et d'étendre ses capacités. Puisque nous disposons aujourd'hui d'appareils numériques très puissants, ces méthodes sont généralement négligées dans la plupart des applications du traitement du signal.

Cependant, un fort intérêt subsiste encore aujourd'hui, notamment dans les contextes sous contraintes dont le traitement distribué des signaux, qui est par nature limité en termes de communications. Des travaux se sont intéressés à l'influence de la quantification sur la précision des trois tâches canoniques en traitement du signal, évoquées précédemment : détection, estimation et classification [2] [39].

La *quantification orientée vers la tâche* dans un sens général a fait l'objet de plusieurs études, par exemple : [28]. Cf également [18] qui présente d'autres références pertinentes. C'est aussi le cas pour la quantification sur un bit, voir par exemple [32] et [33]. Elle a aussi éveillé récemment l'intérêt de thèses complètes sur le sujet, comme celles de Rodrigo Cabral [17] et Manuel Stein [35], qui ont montré la pertinence du traitement des signaux fortement quantifiés dans des contextes sous contraintes. Nous mêmes sommes persuadés que ces outils, peu répandus aujourd'hui en traitement du signal, ont encore beaucoup à offrir dans différents contextes dans le futur.

Bien que plusieurs traitements soient envisageables avec des signaux quantifiés grossièrement, notre présentation se focalise surtout sur l'estimation de la corrélation.

## 4.2 Corrélateurs fortement quantifiés

L'étude qui a conduit à des corrélateurs à forte quantification a été développée en 1943 par VanVlek : *The Spectrum of Clipped Noise* [37] (Spectre d'un bruit écrêté).<sup>1</sup> Ces travaux établissent le résultat suivant : l'écrêtage d'un signal analogique aléatoire gaussien, ce qui est équivalent à une quantification à deux niveaux pour des signaux numériques, produit un signal binaire dont la fonction d'autocorrélation, notée  $\Gamma_{xx}^*(\tau)$ , est  $2/\pi$  fois l'arcsinus de la fonction d'autocorrélation (normalisée) du signal avant écrêtage (notée  $\rho_{xx}(\tau)$ )<sup>2</sup>

$$\Gamma_{xx}^*(\tau) = \frac{2}{\pi} \arcsin\left(\frac{\Gamma_{xx}(\tau)}{\Gamma_{xx}(0)}\right) = \frac{2}{\pi} \arcsin\left(\frac{\Gamma_{xx}(\tau)}{\sigma_x^2}\right) \in [-1, 1]$$

Le but de cette étude n'était cependant pas l'estimation des fonctions de corrélation, mais la simple compréhension de l'effet produit par un écrêtage involontaire sur le spectre d'un signal.

### 4.2.1 Corrélateurs avec des signaux à deux niveaux

La relation de Van Vlek est aussi valable pour la fonction d'intercorrélation entre deux signaux aléatoires gaussiens,  $x$  et  $y$ . Elle peut s'écrire [27] :

$$\rho_{xy}(\tau) = \sin\left(\frac{\pi}{2}\Gamma_{xy}^*(\tau)\right) \quad \text{où} \quad \rho_{xy}(\tau) = \frac{\Gamma_{xy}(\tau)}{\sigma_x\sigma_y}$$

Cette écriture montre explicitement l'intérêt pratique de ce résultat : il est possible, dans le cas de deux signaux gaussiens, de retrouver leur fonction d'intercorrélation (normalisée)  $\rho_{xy}(\tau)$  à partir d'un calcul qui fait intervenir uniquement des signaux binaires, c'est à dire :  $\Gamma_{xy}^*(\tau)$ , ce qui simplifie énormément le calcul de corrélation ainsi que le matériel requis, puisque la multiplication de deux signaux binaires est très simple à mettre en œuvre. Cette méthode, connue sous le nom de *Corrélateur à coïncidence de polarité*, a été proposée en premier par Faran et Hills[16], en septembre 1952. Notons que ce calcul était conçu pour traiter des signaux analogiques. Dans ce manuscrit, nous parlerons uniquement de la version numérique, que nous appellerons *corrélateur classique sur un bit*. Voir fig 4.1.

1. Ce rapport technique [36] fut écrit pendant la deuxième guerre mondiale, mais l'accès à ce document est resté limité jusqu'en 1966, date à laquelle une version commentée a été publiée par IEEE.

2. Ce résultat est aussi connu sur le nom de loi arcsinus (*arcsin law*).

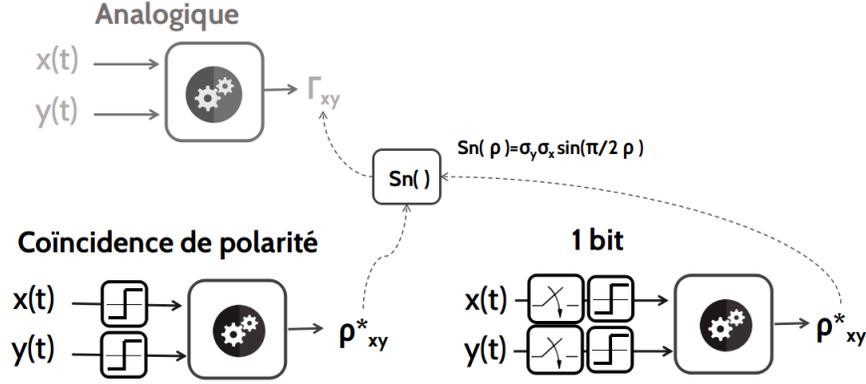


FIGURE 4.1 – Haut : Corrélateur analogique. Bas : Calcul de corrélation avec des signaux à deux niveaux. Le corrélateur à coïncidence de polarité utilise l'écrêtage. Dans le domaine numérique, le corrélateur sur 1 bit quantifie des signaux échantillonnés. Cette méthode indirecte simplifie grandement les calculs.

#### Corrélateur classique sur un bit :

$$\tilde{\mathcal{E}}_N^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{N} \langle \bar{\mathbf{x}}, \bar{\mathbf{y}}_\tau \rangle\right) = \sin\left(\frac{\pi}{2} \frac{1}{N} \bar{\mathbf{x}}^\top \bar{\mathbf{y}}_\tau\right) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N$$

avec

$$\bar{s} = \text{sign}(s)$$

$$\text{et } \mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Il peut également s'écrire :

$$\tilde{\mathcal{E}}_N^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{N} \sum_{i=1}^N \text{sign}(x_i) \text{sign}(y_{i+\tau})\right)$$

#### 4.2.2 Corrélateurs hybrides

Par ailleurs, dans un rapport datant de mars 1952, Julien Bussgang [10] aborde l'estimation de la fonction de corrélation (inter et auto) de signaux gaussiens, lorsque l'une des deux entrées du corrélateur subit une transformation non linéaire sans mémoire. Il parvient à exprimer le lien entre la fonction d'intercorrélation de ces deux signaux et celle de la fonction de corrélation avant distorsion. Il énonce alors le théorème suivant :

##### Théorème de Bussgang :

*Pour deux signaux gaussiens, la fonction de corrélation obtenue, lorsque l'un d'eux subit une distorsion  $D$  non linéaire sans mémoire en amplitude, est identique à la corrélation des signaux avant distorsion, à un facteur de proportionnalité près.*

Ce résultat, illustré dans la figure 4.2, conduit à deux implications pratiques [30] :

- Premièrement, lorsque l'un des signaux subit une distorsion involontaire à cause des composants électroniques, les calculs de corrélation ainsi obtenus peuvent être corrigés.

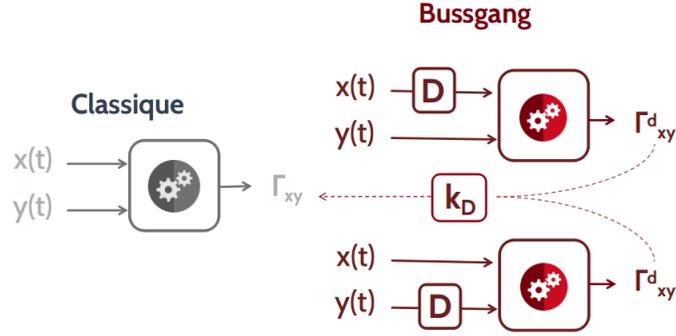


FIGURE 4.2 – Théorème de Bussgang : la sortie d'un corrélateur qui subit une distorsion non linéaire sur l'une des entrées diffère uniquement d'un facteur de proportionnalité, lorsque les signaux sont gaussiens.

- Deuxièmement (et plus important encore), il est possible d'appliquer délibérément cette distorsion, afin de simplifier le calcul de la corrélation.

Bussgang avait considéré plusieurs transformations permettant des méthodes aisées de calcul, mais elles ont été peu étudiées après lui. Des années plus tard, dans sa thèse, Albert Nuttall considère une distorsion correspondant à l'écrêtage du signal<sup>3</sup>. Cette méthode d'estimation a été très bien accueillie dans la littérature. Elle est connue sous le nom de *Corrélateur à relais* ou *Corrélateur hybride*<sup>4</sup>. La motivation pratique découlait aussi des difficultés liées à la multiplication des signaux. En effet, cette méthode de calcul élimine le besoin de multiplicateurs, tout comme pour le corrélateur à coïncidence de polarité.

Dans le cas de ce corrélateur et de signaux gaussiens  $x(t), y(t)$ , le lien entre les fonctions de corrélation avant et après écrêtage, notées  $\Gamma_{xy}$  et  $\Gamma_{xy}^h$ , s'écrit [10, p. 5][6, p. 13] :

$$\Gamma_{xy}^h(\tau) = \frac{1}{\sigma_x} \sqrt{\frac{2}{\pi}} \Gamma_{xy}(\tau) \quad \text{avec } \sigma_{xx} = \sqrt{\Gamma_{xx}(0)} \quad (4.1)$$

si  $x(t)$  est le signal que l'on écrête.

Un calcul équivalent est envisageable et intéressant pour des calculs rapides dans le domaine numérique. David Hertz propose cette approche [21] pour mesurer les fonctions d'autocorrélation. La figure 4.3 illustre les versions analogique et numérique du calcul de corrélation hybride.

Un aspect important à souligner dans les expressions précédentes concerne le terme  $\sigma_x$ , qui est a priori inconnu. Les résultats disponibles dans la littérature envisagent deux perspectives :

- D'une part, celles qui considèrent l'estimation simultanée de  $\Gamma_{xy}(\tau)$  et  $\sigma_x$ . [21].

Puisque

$$\Gamma_{xy}^h(\tau) \Gamma_{xx}^h(0) = \frac{1}{\sigma_x^2} \frac{2}{\pi} \Gamma_{xy}(\tau) \Gamma_{xx}(0) = \frac{2}{\pi} \Gamma_{xy}(\tau)$$

3. C'est vraisemblablement lui le premier à proposer cette idée en 1958 [30, p. 19]

4. Ce terme désigne le fait qu'une des voies était numérique et l'autre analogique.

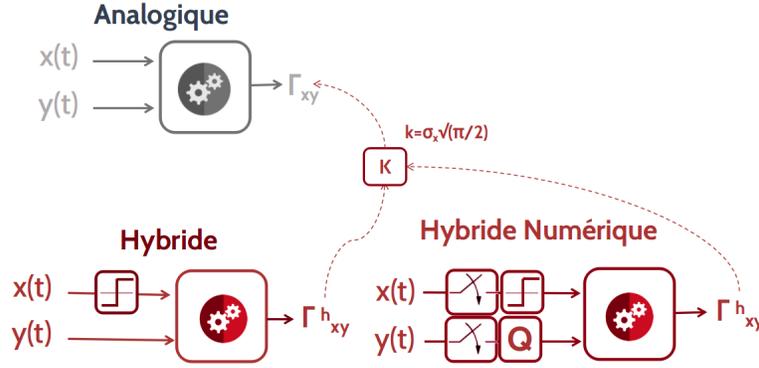


FIGURE 4.3 – Calcul de corrélation hybride. Le corrélateur hybride utilise l'écrêtage sur une des voies. Dans le domaine numérique, le corrélateur hybride numérique utilise deux niveaux de quantification pour l'une des deux voies. Cette méthode indirecte n'a pas besoin de multiplicateurs.

L'estimation avec le corrélateur hybride est dans ce cas :

$$\hat{\Gamma}_{xy}(\tau) = \frac{\pi}{2} \hat{\Gamma}_{xy}^h(\tau) \hat{\Gamma}_{xx}^h(0)$$

Jacovitti [25] a étudié les propriétés de cet estimateur pour l'autocorrélation, c'est à dire  $x = y$ . Il montre que l'estimateur hybride ainsi défini s'avère moins performant que l'estimateur à quantification fine. Ses résultats établissent que la variance de l'estimateur hybride est plus grande que celle de l'estimateur classique.

— D'autre part, celles qui considèrent que la puissance du signal,  $\sigma_x^2$ , est connue.

Les résultats montrent que, dans le cas particulier de processus gaussiens autorégressifs d'ordre 1 [23] ou 2 [24], l'estimateur hybride peut s'avérer meilleur (moins de variance) que le classique, pour les retards  $\tau$  qui conduisent aux grandes valeurs de  $\Gamma_{xy}(\tau)$ , mais moins bon (plus de variance) là où la corrélation est faible.

Par la suite, nous allons considérer uniquement le cas où la puissance du signal est connue, à savoir :

#### Corrélateur classique hybride

$$\mathcal{E}_N^h(\tau) = \frac{K_h}{N} \langle \bar{\mathbf{x}}, \mathbf{y} \rangle = \frac{K_h}{N} \bar{\mathbf{x}}^\top \mathbf{y} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^N$$

$$\text{avec } \bar{\mathbf{x}} = \text{sign}(\mathbf{x})$$

$$K_h = \sqrt{\frac{\pi}{2}} \sigma_x, \quad \mathbf{x} = [x_1, \dots, x_N]^\top, \quad \mathbf{y} = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

$$\text{où } \mathcal{E}_N^h(\tau) = \frac{K_h}{N} \sum_{i=1}^N \text{sign}(x_i) y_{i+\tau}$$

### 4.2.3 Corrélateurs généralisés et modifiés

Les corrélateurs hybride et d'un bit se voient a priori limités au cas particulier des signaux gaussiens. En effet, on ne sait pas inverser le lien entre la corrélation des signaux quantifiés et celle des signaux initiaux que sous cette hypothèse gaussienne. Plusieurs

méthodes ont été proposées dans le passé pour étendre le domaine d'application de ces corrélateurs, ou bien pour améliorer leur performance. Le lecteur intéressé peut se référer à [27] ou [14], qui offrent une vue d'ensemble sur ces méthodes.

Nous allons classifier ces techniques en deux groupes.

- Les **corrélateurs généralisés**, qui s'appuient sur une quantification à un faible nombre de bits. (Cf fig 4.4). Parmi ces méthodes on trouve le corrélateur à coïncidence de polarité généralisé, étudié par Georges Bonnet dans les années 60 [4],[5] ou bien le corrélateur hybride généralisé aussi appelé *corrélateur Stieltjes* qui a été proposé en 62 par Watts [38]. Bonnet en a étudié une variation, la *corrélateur hybride à échantillonnage* en 69 [7].

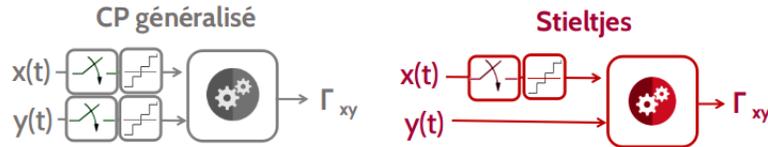


FIGURE 4.4 – Corrélateurs généralisés : une quantification plus fine permet d'améliorer la performance des corrélateurs, voire de travailler avec des signaux non gaussiens.

- Les **corrélateurs modifiés, ou à référence stochastique**<sup>5</sup> Des signaux auxiliaires aléatoires et indépendants des signaux sont ajoutés sur les voies écrêtées ou fortement quantifiées. Voir la figure 4.5. Le lecteur intéressé par ces méthodes peut consulter [3], [14] et [12].

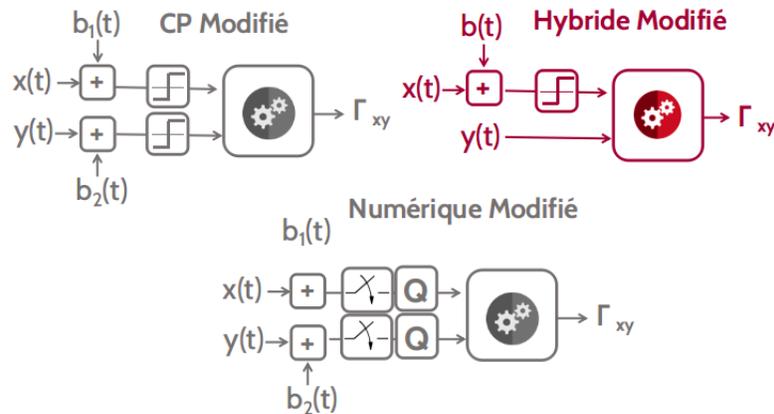


FIGURE 4.5 – Corrélateurs modifiés : ajouter un bruit auxiliaire permet de construire des corrélateurs à forte quantification, capables de traiter des signaux non-gaussiens.

<sup>5</sup>. Ce principe, aussi connu comme quantification stochastique, sort du cadre de la présente étude. Nous soulignons pourtant qu'il s'agit à notre sens d'une piste de recherche intéressante pour les problèmes d'estimation sous contraintes. Francis Castanié a étudié largement ce sujet dans les années 70 et 80. Ses travaux sont une bonne introduction au sujet : [11, 13]. Ce principe est également exploitable dans le cas de l'estimation paramétrique [15].

### 4.3 Acquisition compressive et quantification

En revenant à l'acquisition compressive, l'effet de la quantification des mesures a fait l'objet de plusieurs études. En effet, la reconstruction des signaux issus d'une acquisition compressive, elle-même quantifiée avec peu de bits, est possible [40]. C'est même le cas pour des mesures quantifiées sur un seul bit! [8, 26]. Cette approche a été nommée acquisition compressive d'un bit, (*1-bit compressive sensing*). Bien qu'intéressante au niveau pratique, on retrouve des problèmes liés à son implémentation, car la reconstruction devient plus difficile si la matrice d'acquisition utilisée est tirée selon une loi non-gaussienne [1].

Or, la quantification sur un bit préserve le produit interne et donc la corrélation. Alors, inspirés par les corrélateurs à forte quantification des années 60, nous nous proposons d'étudier dans ces travaux les estimateurs de corrélation obtenus en quantifiant sur un bit les signaux compressés. Dans notre cas, le passage par la reconstruction n'est pas envisagé. À notre connaissance, le traitement des signaux compressés et quantifiés sur un bit n'a pas été étudié.

Voici donc les versions compressées des corrélateurs sur un bit et hybride :

**Corrélateur compressé sur 1 bit :**

$$\widetilde{C}_N^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{M} \langle \overline{\Phi \mathbf{x}}, \overline{\Phi \mathbf{y}_\tau} \rangle\right) = \sin\left(\frac{\pi}{2} \frac{1}{M} (\overline{\Phi \mathbf{x}})^\top (\overline{\Phi \mathbf{y}_\tau})\right) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

avec

$$\overline{\mathbf{s}} = \text{sign}(\mathbf{s})$$

$$\text{et } \mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

**Corrélateur compressé hybride :**

$$C_N^h(\tau) = \frac{K_h}{\sqrt{M}} \langle \overline{\Phi \mathbf{x}}, \overline{\Phi \mathbf{y}_\tau} \rangle = \frac{K_h}{\sqrt{M}} (\overline{\Phi \mathbf{x}})^\top (\overline{\Phi \mathbf{y}_\tau}) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

avec  $\overline{\mathbf{s}} = \text{sign}(\mathbf{s})$

$$K_h = \sqrt{\frac{\pi}{2}} \sigma_x, \mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Nous soulignons un aspect très intéressant concernant ces estimateurs. La quantification sur un bit est effectuée sur les signaux compressés :  $\Phi \mathbf{x}$  et  $\Phi \mathbf{y}$ , générés à partir des combinaisons linéaires d'un grand nombre d'échantillons des processus,  $\mathbf{x}$  et  $\mathbf{y}$  (supposés gaussiens). Or, pour des processus différemment distribués (non gaussiens), les matrices de compression conduisent, par construction, à *gaussianiser* la densité de probabilité des valeurs manipulées, en vertu du théorème de la limite centrale. Lors de la compression des signaux non gaussiens, on obtient donc des signaux compressés qui deviennent, eux, gaussiens et dont la distorsion engendrée par une quantification sur un bit est donnée par la relation de Van Vleck dans le cas du corrélateur compressé sur un bit et par le théorème de Bussgang pour le corrélateur compressé hybride.

À ce titre, ces corrélateurs sont plus généraux. Leur domaine d'application peut aller au delà du cas gaussien.

Leurs statistiques ainsi que la performance de ces estimateurs sont présentées dans le chapitre suivant de ce manuscrit.

## 4.4 Corrélogrammes efficaces

Portons maintenant notre attention sur l'estimation du spectre. La méthode du corrélogramme autorise n'importe quel estimateur de la fonction de corrélation. Ceci est un aspect intéressant lorsque le matériel à disposition est contraint. Il convient alors de considérer des méthodes d'estimation les plus simples possibles. En particulier, rien n'interdit de transformer le signal à étudier en vue de simplifier le matériel requis pour effectuer les calculs ultérieurs.

### 4.4.1 Densité spectrale de puissance

Vraisemblablement, cette idée date de 1962, dans une thèse qui portait sur le traitement des signaux radar pour l'exploration de Venus. Richard Goldstein devait faire face à une énorme taille de données, afin d'obtenir la précision dont son étude avait besoin. L'estimateur de densité spectrale de Goldstein, noté  $\hat{\gamma}_{xx}^G(\nu)$ , est le suivant [20, p. 20] :

$$\hat{\gamma}_{xx}^G(\nu) = \sum_{k=-L}^L \hat{\Gamma}_{xx}^G(k) e^{-j2\pi\nu k} = 1 + 2 \sum_{k=1}^L \hat{\Gamma}_{xx}^G(k) \cos(2\pi\nu k)$$

où  $\hat{\Gamma}_{xx}^G(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{N} \sum_{i=1}^N \text{sign}(x_i) \text{sign}(x_{i+\tau})\right)$

Puisque  $\hat{\Gamma}_{xx}^G(\tau)$  traite uniquement des signaux à deux niveaux, le calcul requis est significativement plus simple que l'estimation usuelle de la corrélation. Notons que l'estimateur de corrélation proposé suppose, comme nous l'avons fait dans ce manuscrit pour  $\mathcal{C}_N(\tau)$ , que l'on dispose d'une taille d'observation égale à  $N + \tau$ .

Quelques années plus tard, plusieurs travaux ont été consacrés aux propriétés asymptotiques de cet estimateur de densité spectrale, notamment Hinich[22] en 68 et McNeil [29]. Ces travaux établissent les propriétés asymptotiques de cet estimateur : Ces études, comme Parzen [31]. Elles se limitent à un cas particulier de fenêtres d'apodisation, celles du type :  $w\left(\frac{k}{L}\right)$ .

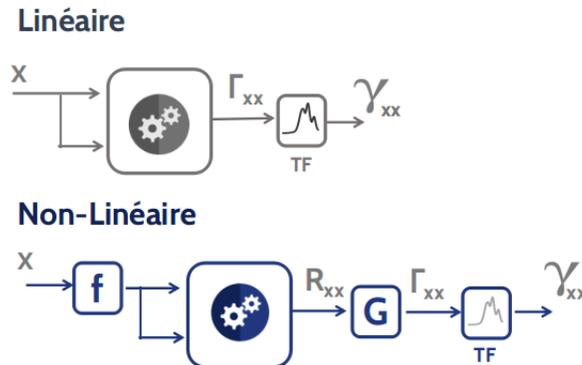


FIGURE 4.6 – Méthodes d'estimation spectrale par corrélogramme. Dans la méthode non-linéaire une fonction  $f$  transforme les signaux pour simplifier les calculs, puis une autre fonction  $G$  sert à corriger l'éventuelle distorsion dans le résultat.

De manière plus générale, le problème d'estimation de la densité spectrale de puissance à partir d'une transformation  $f$  non linéaire du signal a été étudié par Rodemich[34] et Gingras[19]. Cette idée est illustrée dans la figure 4.6.

#### 4.4.2 Densité spectrale d'interaction

Dans ce manuscrit nous nous intéressons plutôt à l'interspectre. Brillinger [9] a décrit les propriétés asymptotiques de l'estimateur non-linéaire suivant :

$$\tilde{s}_{xy}^*(\nu) = \sum_{k=-L}^L \sin \left( \frac{\pi}{2} \frac{1}{N} \sum_{i=1}^{N-|\tau|} \text{sign}(x_i) \text{sign}(y_{i+\tau}) \right) w \left( \frac{k}{N} \right) e^{-j2\pi\nu k}$$

La figure 4.7 illustre une comparaison entre cette méthode (Non-linéaire) et la méthode conventionnelle (Linéaire). Nous reviendrons sur ce résultat au chapitre 5, lorsque nous étudierons ce problème d'estimation dans un cadre distribué.

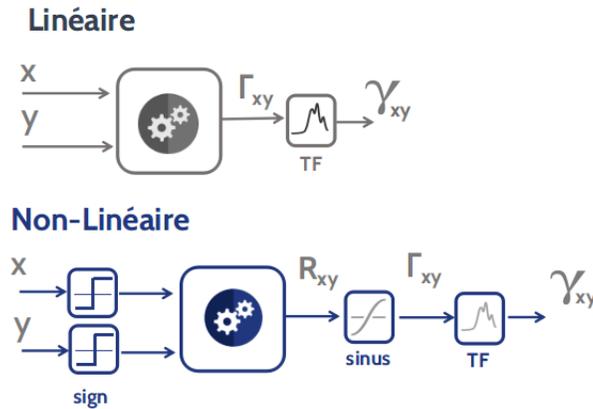


FIGURE 4.7 – Méthodes d'estimation d'interspectre par corrélogramme. Dans la méthode non-linéaire on quantifié les signaux sur 1 bit pour simplifier les calculs, puis la relation de Van-Vleck permet de corriger la distorsion dans le résultat.

### Conclusion

Comme nous l'avons vu dans ce chapitre, la faible capacité de calcul a motivé dans le passé le développement de méthodes de traitement permettant de tirer le meilleur parti des ressources de calcul disponibles. Au fil des ans, le développement de dispositifs électroniques de plus en plus puissants a inévitablement réduit ce besoin.

Comme nous l'avons souligné dans l'introduction, malgré l'énorme capacité de traitement que nous avons actuellement, la quantité de données à traiter augmente constamment, ainsi que notre désir de disposer de systèmes de mesure et de traitement omniprésents. Ces aspects font que les limites dans la capacité de traitement réapparaissent. D'où la nécessité de reconsidérer ces puissants outils pour les appliquer à la résolution des problèmes actuels et futurs.

Une des observations les plus encourageantes de nos travaux est la possibilité de faire revivre ces idées de quantification brutale, en les utilisant conjointement avec la méthode de compression que nous avons étudiée au chapitre 3.

Le chapitre 5 s'intéresse à cet aspect : le *traitement du signal compressé et fortement quantifié*.

## Bibliographie

- [1] AI, A., LAPANOWSKI, A., PLAN, Y., AND VERSHYNIN, R. One-bit compressed sensing with non-gaussian measurements. *Linear Algebra and its Applications* 441 (2014), 222–239.
- [2] AMARI, S., ET AL. Statistical inference under multiterminal data compression. *IEEE Transactions on Information Theory* 44, 6 (1998), 2300–2324.
- [3] BERNDT, H. Correlation function estimation by a polarity method using stochastic reference signals. *IEEE Transactions on Information Theory* 14, 6 (1968), 796–801.
- [4] BONNET, G. Sur la statistique de second ordre des signaux aléatoires quantifiés. *CR Acad. Sci* (1962), 825–827.
- [5] BONNET, G. Sur les corrélateurs utilisant la quantification des signaux. *École d'été OTAN-Marine nationale, Grenoble* (1964), 14–26.
- [6] BONNET, G. Transformations des signaux aléatoires à travers les systèmes non linéaires sans mémoire. *Annals of Telecommunications* 19, 9 (1964), 203–220.
- [7] BONNET, G. Corrélateurs hybrides à échantillonnage. *Annals of Telecommunications* 24, 7 (1969), 252–260.
- [8] BOUFONOS, P. T., AND BARANIUK, R. G. 1-bit compressive sensing. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on* (2008), IEEE, pp. 16–21.
- [9] BRILLINGER, D. R. Estimation of the cross-spectrum of a stationary bivariate gaussian process from its zeros. *Journal of the Royal Statistical Society. Series B (Methodological)* (1968), 145–159.
- [10] BUSSGANG, J. J. Crosscorrelation functions of amplitude-distorted gaussian signals. Tech. rep., Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.
- [11] CASTANIE, F. Signal processing by random reference quantizing. *Signal Processing* 1, 1 (1979), 27–43.
- [12] CASTANIE, F., HOFFMANN, J., AND LACAZE, B. On the performance of a random reference correlator (corresp.). *IEEE Transactions on Information Theory* 20, 2 (1974), 266–269.
- [13] CASTANIÉ, F. Quantification aléatoire. *Annals of Telecommunications* 41, 5 (1986), 235–245.
- [14] CHANG, K.-Y., AND MOORE, A. Modified digital correlator and its estimation errors. *IEEE Transactions on Information Theory* 16, 6 (1970), 699–706.
- [15] DABEER, O., AND KARNIK, A. Signal parameter estimation using 1-bit dithered quantization. *IEEE Transactions on Information Theory* 52, 12 (2006), 5389–5405.
- [16] FARAN, J., JAMES, J., HILLS, J., ET AL. Correlators for signal reception. Tech. rep., HARVARD UNIV CAMBRIDGE MA ACOUSTICS RESEARCH LAB, 1952.
- [17] FARIAS, R. C. *Estimation sous contraintes de communication : algorithmes et performances asymptotiques*. PhD thesis, Université Grenoble Alpes, 2013.
- [18] FOWLER, M. L., AND CHEN, M. Fisher-information-based data compression for estimation using two sensors. *IEEE Transactions on Aerospace and Electronic Systems* 41, 3 (2005), 1131–1137.
- [19] GINGRAS, D., AND MASRY, E. Spectral density estimation from nonlinearly observed data. *Journal of Time Series Analysis* 6, 2 (1985), 63–80.

- [20] GOLDSTEIN, R. M. *Radar exploration of Venus*. PhD thesis, California Institute of Technology, 1962.
- [21] HERTZ, D. A fast digital method of estimating the autocorrelation of a gaussian stationary process. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 30, 2 (1982), 329–329.
- [22] HINICH, M. Estimation of spectra after hard clipping of gaussian processes? *Technometrics* 9, 3 (1967), 391–400.
- [23] HUZII, M. On a simplified method of the estimation of the correlogram for a stationary gaussian process. *Annals of the Institute of Statistical Mathematics* 14, 1 (1962), 259–268.
- [24] HUZII, M. On a simplified method of the estimation of the correlogram for a stationary gaussian process, ii. In *Kodai Mathematical Seminar Reports* (1964), vol. 16, Department of Mathematics, Tokyo Institute of Technology, pp. 199–212.
- [25] JACOVITTI, G., NERI, A., AND CUSANI, R. On a fast digital method of estimating the autocorrelation of a gaussian stationary process. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32, 5 (1984), 968–976.
- [26] JACQUES, L., LASKA, J. N., BOUFONOS, P. T., AND BARANIUK, R. G. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Transactions on Information Theory* 59, 4 (2013), 2082–2102.
- [27] KORN, G. A. Hybrid-computer techniques for measuring statistics from quantized data. *Transactions of the Society for Computer Simulation* 4, 4 (1965), 229–239.
- [28] LAM, W.-M., AND REIBMAN, A. R. Design of quantizers for decentralized estimation systems. *IEEE Transactions on communications* 41, 11 (1993), 1602–1605.
- [29] MCNEIL, D. Estimating the covariance and spectral density functions from a clipped stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)* (1967), 180–195.
- [30] NUTTALL, A. H. Theory and application of the separable class of random processes. Tech. rep., Massachusetts Institute of Technology, Research Laboratory of Electronics, 1958.
- [31] PARZEN, E. On consistent estimates of the spectrum of a stationary time series. *The Annals of Mathematical Statistics* (1957), 329–348.
- [32] RIBEIRO, A., AND GIANNAKIS, G. B. Bandwidth-constrained distributed estimation for wireless sensor networks-part i : Gaussian case. *IEEE transactions on signal processing* 54, 3 (2006), 1131–1143.
- [33] RIBEIRO, A., AND GIANNAKIS, G. B. Bandwidth-constrained distributed estimation for wireless sensor networks-part ii : Unknown probability density function. *IEEE Transactions on Signal Processing* 54, 7 (2006), 2784–2796.
- [34] RODEMICH, E. R. Spectral estimates using nonlinear functions. *The Annals of Mathematical Statistics* (1966), 1237–1256.
- [35] STEIN, M. S. *Signal Parameter Estimation with 1-bit ADC : Performance Bounds, Methods and System Design*. PhD thesis, Dissertation, München, Technische Universität München, 2016, 2016.
- [36] VAN VLECK, J. H. The spectrum of clipped noise. Tech. rep., Harvard Radio Research Laboratory, July 1943.
- [37] VAN VLECK, J. H., AND MIDDLETON, D. The spectrum of clipped noise. *Proceedings of the IEEE* 54, 1 (1966), 2–19.
- [38] WATTS, D. A general theory of amplitude quantization with applications to correlation determination. *Proceedings of the IEE-Part C : Monographs* 109, 15 (1962), 209–218.

- [39] ZHANG, Z., AND BERGER, T. Estimation via compressed information. *IEEE transactions on Information theory* 34, 2 (1988), 198–211.
- [40] ZYMNIS, A., BOYD, S., AND CANDÉS, E. Compressed sensing with quantized measurements. *IEEE Signal Processing Letters* 17, 2 (2010), 149–152.

# Chapitre 5

## Estimation dans le domaine compressé et quantifié

### Sommaire

---

<b>5.1</b>	<b>Corrélateurs sur 1 bit</b>	<b>73</b>
5.1.1	Estimateur classique sur 1 bit	75
5.1.2	Estimateur compressé sur 1 bit	77
5.1.3	Compressibilité et performance	77
<b>5.2</b>	<b>Corrélateurs hybrides</b>	<b>79</b>
5.2.1	Estimateur classique hybride	80
5.2.2	Estimateur compressé hybride	81
5.2.3	Compressibilité et performance	81
<b>5.3</b>	<b>Corrélogrammes compressés et quantifiés</b>	<b>83</b>
5.3.1	Corrélogrammes sur 1 bit	84
5.3.2	Corrélogramme classique sur 1 bit	84
5.3.3	Corrélogramme compressé sur 1 bit	84
5.3.4	Compressibilité et performance	85
5.3.5	Corrélogrammes hybrides	87
	<b>Bibliographie</b>	<b>90</b>

---

Ce chapitre présente une brève étude des performances des estimateurs compressés et quantifiés que nous avons proposés précédemment, en les comparant à celles des estimateurs classiques. Nous abordons l'estimation de la fonction de corrélation, ainsi que celle de la densité spectrale de puissance d'interaction.

### 5.1 Corrélateurs sur 1 bit

Nous avons vu dans le chapitre précédent que dans le cas particulier des signaux gaussiens, il est possible de s'appuyer sur une quantification brutale des signaux (effectuée uniquement sur un bit) pour réduire considérablement la complexité de l'estimation de la fonction de corrélation. En effet, les passages par zéro de ces signaux suffisent pour reconstruire les fonctions de (inter et auto) corrélation. Dans un contexte de traitement distribué, comme celui qui nous intéresse, ces méthodes sont particulièrement attrayantes parce qu'elles permettent également de faire des économies d'énergie, en raison de la forte réduction du volume de données à transmettre.

Une première méthode visant à réduire considérablement le débit de transmission

consiste donc en la sélection de  $M$  échantillons successifs parmi  $N$  disponibles, que l'on quantifie ensuite sur un seul bit.

#### Approche classique quantifiée sur 1 bit :

Ne fournir qu'un sous-ensemble de  $M$  échantillons successifs quantifiés sur 1 bit, puis exploiter le *corrélateur quantifié* pour  $M$  échantillons dans le capteur  $y$

$$\widetilde{\mathcal{E}}_M^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{M} \bar{\mathbf{x}}^\top \bar{\mathbf{y}}_\tau\right) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^M$$

Avec

$$\mathbf{x} = [x_1, \dots, x_M]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{M+\tau}]^\top$$

où  $\bar{s} = \text{sign}(s)$

Par ailleurs, nous nous proposons d'étudier également la réduction de la taille des données par projection aléatoire, suivie là encore d'une quantification sur un bit.

#### Approche compressée quantifiée sur 1 bit :

Envoyer  $M$  mesures, obtenues par projection aléatoire, quantifiées sur un bit. Ces projections résultent de l'acquisition de  $N$  échantillons de  $\mathbf{x}$  sur le capteur  $x$ . Ensuite, on met en œuvre le *corrélateur compressé sur 1 bit* dans le capteur  $y$  :

$$\widetilde{\mathcal{C}}_N^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{M} (\Phi \mathbf{x})^\top (\Phi \mathbf{y}_\tau)\right) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

Avec

$$\mathbf{x} = [x_1, \dots, x_N]^\top \text{ et } \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

où  $\bar{s} = \text{sign}(s)$

Il faut souligner que dans le cas compressé sur 1 bit, comme on l'a fait pour l'estimateur compressé dans le chapitre 3, on suppose que le coût énergétique lié au stockage et aux calculs numériques nécessaires (réalisés sur le capteur  $y$ ) est faible par rapport au coût associé à la transmission de données.

Les deux estimateurs sur 1 bit sont illustrés sur la figure 5.1 :

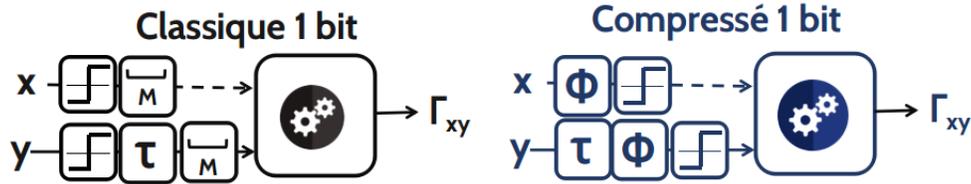


FIGURE 5.1 – Corrélateurs sur 1 bit, à gauche le classique et à droite celui que nous proposons, le compressé. Cette quantification permet une réduction supplémentaire sur la taille des données, qui peut s'avérer dans certains cas très importante. L'estimateur classique ne peut être appliqué que dans le cas de signaux gaussiens, alors que, par construction, l'estimateur compressé modifie la densité de probabilité des signaux en la faisant tendre vers la gaussienne, ce qui permet de l'employer sur presque n'importe quel type de signal.

Passons maintenant à l'étude du comportement des corrélateurs sur 1 bit. On suppose que la puissance des processus est connue et égale à 1 ( $\Gamma_{xx}(0) = 1 = \sigma_{xx}^2 = 1 = \sigma_{yy}^2 = \Gamma_{yy}(0)$ ), auquel cas  $\rho_{xy}(\tau) = \frac{\Gamma_{xy}(\tau)}{\sigma_x \sigma_y} = \Gamma_{xy}(\tau)$ .

### 5.1.1 Estimateur classique sur 1 bit

#### Biais

Nous allons d'abord introduire la notation  $\mathcal{C}_M^*(\tau)$ , pour nous référer à l'estimateur sans correction du biais :

$$\widetilde{\mathcal{C}}_M^*(\tau) = \sin\left(\frac{\pi}{2}\mathcal{C}_M^*(\tau)\right) = \sin\left(\frac{\pi}{2}\frac{1}{M}\sum_{i=1}^M \text{sign}(x_i)\text{sign}(y_{i+\tau})\right)$$

Le biais de l'estimateur sans correction, est donné par la relation de Van-Vlek :

$$\mathbf{E}[\mathcal{C}_M^*(\tau)] = \frac{2}{\pi}\arcsin(\rho_{xy}(\tau))$$

Puis,  $\mathbf{E}[\widetilde{\mathcal{C}}_M^*(\tau)]$  peut être obtenu par la méthode delta<sup>1</sup> [6], pour ce faire il est nécessaire que l'estimateur sans correction satisfasse le théorème de la limite centrale usuel :

$$\sqrt{M}\left[\mathcal{C}_M^*(\tau) - \frac{2}{\pi}\arcsin(\rho_{xy}(\tau))\right] \xrightarrow{L} \mathcal{N}(0, w(\tau))$$

où  $\xrightarrow{L}$  dénote la convergence en loi et  $w(\tau)$  la variance de  $\mathcal{C}_M^*(\tau)$

Nous admettons cette hypothèse, qui est raisonnable lorsque  $M$  est grand. En utilisant le fait que  $\cos(\arcsin(x)) = 1 - x^2$  on aboutit à :

$$\sqrt{M}\left[\widetilde{\mathcal{C}}_M^*(\tau) - \rho_{xy}(\tau)\right] \xrightarrow{L} \mathcal{N}\left(0, w(\tau)\frac{\pi^2}{4}(1 - \rho_{xy}^2(\tau))^2\right)$$

Auquel cas l'estimateur  $\widetilde{\mathcal{C}}_M^*(\tau)$  est asymptotiquement non biaisé.

#### Biais

Pour des signaux  $\mathbf{x}, \mathbf{y}_\tau$  gaussiens, cet estimateur est asymptotiquement non biaisé, c'est-à-dire que si  $M$  est grand :

$$\mathbf{E}[\widetilde{\mathcal{C}}_M^*(\tau)] = \rho_{xy}(\tau)$$

#### Variance :

Dans l'analyse précédente, il apparaît que la variance de  $\widetilde{\mathcal{C}}_M^*(\tau)$  peut se déduire de celle de  $\mathcal{C}_M^*(\tau)$ , dénotée précédemment  $w(\tau)$ . Cependant, il n'existe pas d'expression analytique pour  $w(\tau)$ . Nous en expliquons maintenant la raison. Cette variance s'écrit :

$$\begin{aligned} M^2 \mathbf{Var}\{\mathcal{C}_M^*(\tau)\} &= \mathbf{Var}\{\bar{\mathbf{x}}^\top \bar{\mathbf{y}}_\tau\} = \mathbf{E}[(\bar{\mathbf{x}}^\top \bar{\mathbf{y}}_\tau)^2] - \mathbf{E}[\bar{\mathbf{x}}^\top \bar{\mathbf{y}}_\tau]^2 \\ &= \mathbf{E}\left[\sum_{i,j} \bar{x}_i \bar{y}_{i+\tau} \bar{x}_j \bar{y}_{j+\tau}\right] - \mathbf{E}\left[\sum_{i=1}^M \bar{x}_i \bar{y}_{i+\tau}\right]^2 \end{aligned}$$

Par linéarité de l'espérance on obtient :

$$M^2 \mathbf{Var}\{\mathcal{C}_M^*(\tau)\} = \sum_{i=1}^M \mathbf{E}[\bar{x}_i^2 \bar{y}_{i+\tau}^2] + \sum_{i \neq j} \mathbf{E}[\bar{x}_i \bar{y}_{i+\tau} \bar{x}_j \bar{y}_{j+\tau}] - \left(\sum_{i=1}^M \mathbf{E}[\bar{x}_i \bar{y}_{i+\tau}]\right)^2$$

1. Cette méthode établit que pour une transformation  $g$  non linéaire de l'estimateur, on a :  $\sqrt{M}[g(\mathcal{C}_M^*(\tau)) - g(\frac{2}{\pi}\arcsin(\rho_{xy}(\tau)))] \xrightarrow{L} \mathcal{N}(0, w(\tau)[g'(\frac{2}{\pi}\arcsin(\rho_{xy}(\tau)))]^2)$   
où  $\xrightarrow{L}$  dénote la convergence en loi. Dans notre cas :  $g(x) = \sin(\frac{\pi}{2}x) \Rightarrow g'(x) = \frac{\pi}{2}\cos(\frac{\pi}{2}x)$ .

Puisque  $\bar{x}_i, \bar{y}_i = \pm 1$ ,  $\mathbf{E}[\bar{x}_i^2 \bar{y}_{i+\tau}^2]$  vaut 1, on a finalement :

$$= M + \sum_{i \neq j} \mathbf{E}[\bar{x}_i \bar{y}_{i+\tau} \bar{x}_j \bar{y}_{j+\tau}] - \left( \sum_{i=1}^M \mathbf{E}[\bar{x}_i \bar{y}_{i+\tau}] \right)^2 \quad (5.1)$$

Dans cette expression, on peut calculer uniquement  $\mathbf{E}[\bar{x}_i \bar{y}_{j+\tau}]$ . D'après la relation de Van-Vleck, ce moment vaut  $\frac{2}{\pi} \arcsin(\rho_{xy}(\tau))$ . Malheureusement, il n'existe pas de relation explicite pour le moment d'ordre 4  $\mathbf{E}[\bar{x}_i \bar{y}_{i+\tau} \bar{x}_j \bar{y}_{j+\tau}]$ , sauf dans quelques cas particuliers.

En fait, ce moment d'ordre quatre est étroitement lié à l'intégrale quadrivariable normale [4, 5], à savoir la probabilité que quatre variables aléatoires conjointement gaussiennes soient simultanément positives. Ce problème mathématique a été beaucoup étudié<sup>2</sup>, mais il ne possède de solution analytique que dans quelques cas particuliers. D'après [1], il semble qu'il n'y a pas de solution analytique au delà de l'ordre 3.

Parmi les cas qui possèdent une solution analytique figure celui des signaux échantillonnés avec une période  $T_e$  suffisamment grande pour obtenir des séquences numériques dont les échantillons successifs sont décorrélés. Autrement dit, on suppose deux processus blancs, mais corrélés entre eux :  $\mathbf{E}[x_i y_i] = \rho_{xy}$ . Gabriel [2] parvient alors au résultat suivant<sup>3</sup> :

### Biais

Pour deux processus  $x$  et  $y$  gaussiens, stationnaires, échantillonnés à une fréquence suffisamment élevée, tels que :

$$\mathbf{E}[x_i x_j] = \mathbf{E}[y_i y_j] = 0, \quad i \neq j$$

On a :

$$\mathbf{E}[\widetilde{\mathcal{C}}_M^*(\tau)] = \rho_{xy} \exp\left(-\frac{1}{M} \frac{\pi^2}{8} \left(1 - \left(\frac{2}{\pi} \arcsin(\rho_{xy})\right)^2\right)\right)$$

Pour  $M$  tendant vers l'infini on obtient :

$$\mathbf{E}[\widetilde{\mathcal{C}}_M^*(\tau)] \rightarrow \rho_{xy}$$

### Variance

Pour  $M$  grand :

$$\mathbf{Var}\{\widetilde{\mathcal{C}}_M^*(\tau)\} \approx \frac{(1 - \rho_{xy}^2) \pi^2}{M} \frac{1}{4} \left(1 - \left(\frac{2}{\pi} \arcsin(\rho_{xy})\right)^2\right) \quad (5.2)$$

L'expression précédente de la variance nous indique que la précision obtenue avec ce corrélateur dépend de la valeur à estimer. Le calcul est précis lorsque la corrélation est forte ( $|\rho_{xy}| \rightarrow 1$ ), puisque la variance s'annule. Mais l'erreur devient grande lorsque la corrélation est faible ( $|\rho_{xy}| \rightarrow 0$ ). Ce comportement est opposé à celui du corrélateur classique, pour lequel l'erreur augmente avec la valeur estimée.

On peut voir également que l'estimateur est consistant : il est asymptotiquement non biaisé et sa variance tend vers zéro pour  $N$  tendant vers l'infini. Notons cependant que l'hypothèse considérée par Gabriel est très contraignante, si bien que ce résultat n'est présenté ici qu'à titre indicatif. Nous reviendrons sur ce résultat par la suite.

<sup>2</sup>. Le lecteur intéressé par une vue d'ensemble du problème est invité à se référer à l'ouvrage de Tong [8, p. 188]

<sup>3</sup>. Un résultat analogue considérant une variante du corrélateur est étudié dans [7]

Faute d'une expression analytique de la variance dans le cas général, on peut l'évaluer par intégration numérique, ou à l'aide de simulations Monte Carlo. Nous allons utiliser la seconde méthode.

### 5.1.2 Estimateur compressé sur 1 bit

Rappelons que cet estimateur est une transformation non linéaire de l'estimateur sans correction du biais  $\widetilde{C}_N^*(\tau) = \sin\left(\frac{\pi}{2}C_N^*(\tau)\right)$ . D'après la relation de Van-Vleck<sup>4</sup>

$$\mathbf{E}[C_N^*(\tau)] = \frac{2}{\pi} \arcsin(\rho_{xy}(\tau))$$

Ensuite, si l'on suppose qu'il satisfait le théorème de la limite centrale habituel :

$$\sqrt{N} \left[ C_N^*(\tau) - \frac{2}{\pi} \arcsin(\rho_{xy}(\tau)) \right] \xrightarrow{L} \mathcal{N}(0, q(\tau))$$

La méthode delta établit alors que :

$$\sqrt{N} \left[ \widetilde{C}_N^*(\tau) - \rho_{xy}(\tau) \right] \xrightarrow{L} \mathcal{N} \left( 0, q(\tau) \frac{\pi^2}{4} (1 - \rho_{xy}^2(\tau))^2 \right)$$

Sous cette hypothèse l'estimateur  $\widetilde{C}_N^*(\tau)$  est asymptotiquement non biaisé.

#### Biais

Pour  $N$  grand :

$$\mathbf{E}[\widetilde{C}_N^*(\tau)] = \rho_{xy}(\tau)$$

Encore une fois, la variance de  $\widetilde{C}_N^*(\tau)$  s'écrit en fonction de celle de  $C_N^*(\tau)$ . L'évaluation de cette dernière présente les mêmes difficultés que la variance de l'estimateur classique sur un bit. Comme mentionné précédemment, la variance sera évaluée par des simulations Monte Carlo. Le paragraphe suivant a pour objectif de présenter ce résultat dans le cas très simple d'un processus autorégressif gaussien d'ordre 1 (AR(1)), dont on estime l'autocorrélation  $\Gamma_{xx}(\tau)$  en supposant  $\Gamma_{xx}(0) = 1$ .

### 5.1.3 Compressibilité et performance

#### Variance dans le cas gaussien AR(1)

Nous illustrons la variance de différents estimateurs sur la figure 5.2 : classique, compressé, classique sur 1 bit et compressé sur 1 bit. Pour obtenir ces courbes, nous avons choisi  $N = 1000$  et  $M = 100$ , soit un facteur de compression de 10, la matrice  $\Phi$  reste constante. La fonction de corrélation du processus est évaluée pour les 20 premiers retards<sup>5</sup>. La variance des différents corrélateurs est évaluée de manière empirique à partir des estimations calculées sur 1000 réalisations indépendantes des processus  $x$  et  $y$ .

$$y_t = x_t = ax_{t-1} + \sqrt{1-a^2}\epsilon_t$$

Par ailleurs, même si l'expression de la variance obtenue par Gabriel (5.2) ne correspond qu'au cas  $a = 0$  (pour lequel le processus AR(1) possède des échantillons successifs

4. Cela suppose que  $\Phi\mathbf{x}$  et  $\Phi\mathbf{y}$  sont non-seulement gaussiens mais aussi conjointement gaussiens, ce qui correspond à ce que nous avons observé expérimentalement lorsque  $N$  est grand, mais qui reste à prouver formellement.

5. Nous avons délibérément exclu le cas de  $\tau = 0$ , où une variance nulle apparaît. Ce n'est qu'un artefact, conséquence d'un calcul de variance effectué sur une valeur constante. Supposée égale à 1 :  $\Gamma_{xx}(0) = 1$

décorrélés), nous allons tracer cette approximation dans les cas étudiés, afin de montrer ce qui arrive si cette hypothèse de base n'est pas respectée.

Tout d'abord, nous retrouvons au travers de ces simulations les résultats théoriques établis au chapitre 3 : le compressé (en trait cyan) s'avère plus performant (il a une variance plus faible) que le classique (en gris) pour des processus qui ont suffisamment de redondance (Par exemple,  $a = 0.4$  et  $0.7$ ). Ensuite, on observe que ce même phénomène se produit également pour les deux estimateurs quantifiés sur un bit : compressé (représenté en bleu) et classique (en noir). Le premier est meilleur (sa variance est inférieure), à condition que le processus que l'on traite soit suffisamment compressible (par exemple  $a = 0.4, 0.7$ ). Notons que les estimateurs quantifiés sur un bit ont un comportement opposé aux estimateurs finement quantifiés. À savoir, les valeurs de corrélation élevées sont très bien estimées, puisque la variance diminue en même temps que la corrélation augmente, mais par contre, les valeurs de corrélation faibles le sont moins bien, car la variance atteint son maximum lorsque la corrélation est nulle. Cette prédiction avait été faite dans le chapitre 3 pour l'estimateur classique sur 1 bit, mais uniquement pour des séquences décorréliées en temps. Ce résultat montre que des séquences corrélées produisent le même comportement. Ces simulations nous montrent que l'estimateur compressé a lui aussi ce même comportement.

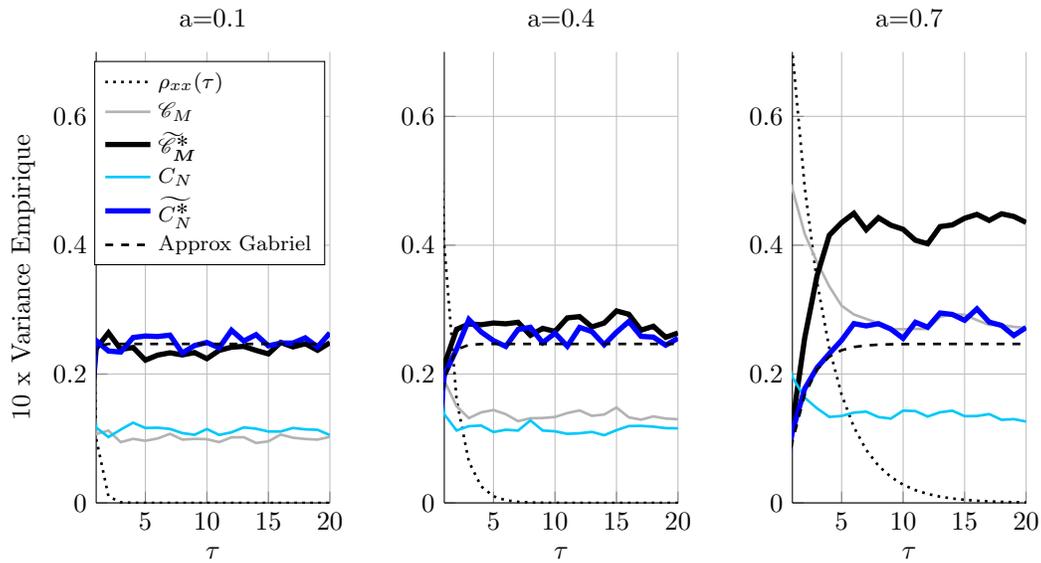


FIGURE 5.2 – Variance empirique des différents estimateurs de corrélation d'un processus AR(1), en fonction de  $\tau = 1, \dots, 20$  pour trois valeurs du coefficient de régression  $a$  : 0.1, 0.4, 0.7. Cette figure compare les estimateurs classiques à leurs versions compressées (avec une matrice de projection creuse). Les courbes correspondent à l'exploitation de 1000 réalisations des processus, pour  $N = 1000$ ,  $\widetilde{M} = 100$ , la matrice de compression est constante. Les estimateurs compressés ( $C_N$ , et  $\widetilde{C}_N^*$ ) s'avèrent meilleurs que leurs contreparties classiques ( $\mathcal{E}_M$  et  $\widetilde{\mathcal{E}}_M^*$ ) si le processus est suffisamment *compressible* ( $a = 0.4, 0.7$ ). Pour des processus très compressibles, l'estimateur compressé sur 1 bit (qui transmet donc au total  $M$  bits) est préférable au classique (qui transmet  $M$  valeurs multibit). L'approximation proposée par Gabriel ne permet pas de décrire le comportement de l'estimateur classique sur 1 bit lorsque  $a$  augmente. Ce résultat s'avère cependant proche du comportement de l'estimateur compressé.

Une autre observation qui nous semble pertinente concerne la comparaison entre l'estimateur classique (qui transmet  $M$  valeurs) et l'estimateur compressé sur un bit (qui transmet uniquement  $M$  bits). Si le processus est fortement redondant ( $a = 0.7$ ), l'estimateur compressé quantifié (en trait bleu) s'avère meilleur que le classique (en noir). Cela représente une meilleure performance, doublée d'une forte diminution du volume de données transmises.

À propos de l'approximation de la variance (équ. 5.2) de l'estimateur classique sur 1 bit, qui n'est valable que pour des séquences décorréées en temps, on voit bien ce qui arrive lorsque cette hypothèse n'est pas satisfaite : plus le processus est corrélé ( $|a|$  tend vers 1), plus l'écart entre l'approximation et la valeur obtenue expérimentalement est important. Par exemple pour  $a = 0.7$ , l'approximation s'avère très optimiste.

Chose intéressante, ces simulations nous montrent aussi que le passage par une matrice de compression d'un processus corrélé de taille  $N$ , engendre un signal compressé dont les  $M$  valeurs sont peu corrélées entre elles. En effet, la performance de l'estimateur compressé sur un bit, s'approche du résultat fourni par Gabriel, résultat valable pour  $M$  valeurs décorréées. Les valeurs des vecteurs compressés possèdent manifestement un lien moins fort que celui du processus à partir duquel ils ont été obtenus.

Exploiter la quantification sur 1 bit dans les deux capteurs n'est pas la seule solution envisageable. Il est également possible de ne quantifier qu'un seul des signaux : celui qui est transmis d'un capteur à l'autre. Le paragraphe suivant étudie deux approches hybrides qui exploitent cette idée : l'estimateur hybride et la version compressée correspondante.

## 5.2 Corrélateurs hybrides

Nous venons de voir que dans le cas particulier des signaux gaussiens, il est possible d'exploiter les corrélateurs sur 1 bit pour faire des économies énergétiques considérables en termes de données transmises. En vue de trouver un compromis entre consommation d'énergie et performance calculatoire, nous portons maintenant notre attention sur les corrélateurs hybrides, où l'on ne quantifie sur un bit que l'un des deux signaux. Clairement, le plus avantageux est de quantifier sur 1 bit les données qui sont transmises.

La première méthode étudiée visant à réduire fortement le débit de transmission est la suivante :

### Approche classique hybride :

Ne transmettre qu'un sous-ensemble de  $M$  échantillons successifs quantifiés sur 1 bit, puis exploiter le *corrélateur hybride classique* sur  $M$  échantillons dans le capteur  $y$  :

$$\mathcal{C}_M^h(\tau) = \frac{K_h}{M} \langle \bar{\mathbf{x}}, \mathbf{y}_\tau \rangle = \frac{K_h}{M} \bar{\mathbf{x}}^\top \mathbf{y}_\tau \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^M$$

Avec  $\bar{\mathbf{s}} = \text{sign}(\mathbf{s})$

$$K_h = \sqrt{\frac{\pi}{2}} \sigma_x, \quad \mathbf{x} = [x_1, \dots, x_M]^\top, \quad \mathbf{y}_\tau = [y_{t+\tau}, \dots, y_{M+\tau}]^\top$$

Nous nous proposons principalement d'étudier la version compressée de cet estimateur :

**Approche compressée hybride :**

Envoyer la version quantifiée sur 1 bit des  $M$  mesures calculées par projection aléatoire de  $N$  échantillons de  $\mathbf{x}$  dans le capteur  $x$ , puis utiliser le *corrélateur compressé hybride* dans le capteur  $y$  :

$$C_N^h(\tau) = \frac{K_h}{\sqrt{M}} \langle \bar{\Phi} \mathbf{x}, \Phi \mathbf{y}_\tau \rangle = \frac{K_h}{\sqrt{M}} (\bar{\Phi} \mathbf{x})^\top (\Phi \mathbf{y}_\tau) \quad \mathbf{x}, \mathbf{y}_\tau \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N}$$

Avec  $\bar{\mathbf{s}} = \text{sign}(\mathbf{s})$

$$K_h = \sqrt{\frac{\pi}{2}} \sigma_x, \quad \mathbf{x} = [x_1, \dots, x_N]^\top, \quad \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

Rappelons que notre étude se limite au cas où la puissance du signal  $\sigma_x^2$  est connue et égale à 1. Tout comme les corrélateurs sur 1 bit, ces deux corrélateurs permettent de réaliser des économies très importantes en termes de transmission : eux aussi ne transmettent que  $M$  bits. Les commentaires formulés précédemment pour l'estimateur compressé sur 1 bit sont tout aussi pertinents pour le corrélateur hybride compressé. Premièrement, on néglige le coût énergétique des calculs effectués sur le capteur  $y$ , devant le coût de la transmission. Deuxièmement, l'effet décrit précédemment, la *gaussianisation* obtenue lorsque l'on combine linéairement les valeurs du signal, permet d'exploiter le corrélateur hybride compressé pour des signaux non gaussiens, sachant que la distorsion engendrée est celle établie par le théorème de Bussgang dans le cas gaussien.

Ces deux nouveaux estimateurs, sont présentés sur la figure suivante :

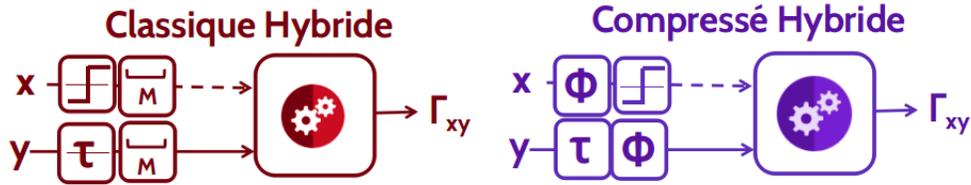


FIGURE 5.3 – Corrélateurs hybrides : à gauche le classique et à droite le compressé que nous proposons. Comme les corrélateurs précédents, ces deux estimateurs transmettent uniquement  $M$  bits. La quantification sur un bit n'est appliquée que sur le signal transmis. L'estimateur hybride classique ne peut être utilisé théoriquement que pour des signaux gaussiens. Comme l'estimateur compressé *gaussianise* les signaux, ceci permet de l'employer pour presque n'importe quel type de signal.

Étudions maintenant le comportement de ces deux corrélateurs hybrides. On suppose à nouveau que la puissance du signal  $\mathbf{x}$  est connue et vaut  $\sigma_x^2 = 1$ .

### 5.2.1 Estimateur classique hybride

Cet estimateur s'écrit :

$$\mathcal{E}_M^h(\tau) = \frac{K_h}{M} \sum_{i=1}^M \text{sign}(x_i) y_{i+\tau}$$

**Biais :**

$$\mathbf{E}[\mathcal{E}_M^h(\tau)] = \frac{K_h}{M} \sum_{i=1}^M \mathbf{E}[\text{sign}(x_i) y_{i+\tau}]$$

D'après le théorème de Bussgang, (équation 4.1) :

$$\begin{aligned}
 &= \frac{K_h}{M} \sum_{i=1}^M \frac{1}{\sigma_x} \sqrt{\frac{2}{\pi}} \rho_{xy}(\tau) \\
 &= \frac{K_h}{M} \sum_{i=1}^M \frac{\rho_{xy}(\tau)}{K_h} \\
 &= \rho_{xy}(\tau)
 \end{aligned}$$

### Biais

Cet estimateur n'est pas biaisé :

$$\mathbf{E}[\mathcal{C}_M^h(\tau)] = \rho_{xy}(\tau)$$

Le calcul de la variance a été fait par Koh : [3]. Notre étude se limite uniquement à des simulations de type Monte-Carlo.

### 5.2.2 Estimateur compressé hybride

Nous admettons à nouveau que lorsque  $N$  est grand  $\Phi \mathbf{x}$  et  $\Phi \mathbf{y}$  sont conjointement gaussiens. Auquel cas, d'après le théorème de Bussgang l'estimateur est asymptotiquement non biaisé.

### Biais

Pour  $N$  grand, l'estimateur est non biaisé.

$$\mathbf{E}[C_N^h(\tau)] = \rho_{xy}(\tau)$$

### 5.2.3 Compressibilité et performance

Notre étude sur ces corrélateurs sera réalisée uniquement au niveau qualitatif à travers des simulations Monte Carlo. Nous laissons ouverte l'analyse théorique ...

#### Variance dans le cas d'un processus gaussien AR(1)

Dans ce paragraphe, nous comparons les variances des deux corrélateurs hybrides. D'abord avec celle des deux corrélateurs numériques (quantifiés sur plusieurs bits) et ensuite avec celle des corrélateurs sur un bit. Pour ces simulations, on choisit  $N = 1000$  et  $M = 100$ , soit un taux de compression  $\alpha = 10$ . L'autocorrélation est évaluée pour les 20 premiers retards. La variance d'estimation est évaluée de manière empirique sur la base de 1000 réalisations des processus  $x$  et  $y$ , deux à deux indépendantes,  $\Phi$  reste constante.

La figure 5.4 illustre le comportement des estimateurs hybrides et numériques. Elle met en évidence quelques effets qui ont déjà été décrits pour les estimateurs présentés précédemment.

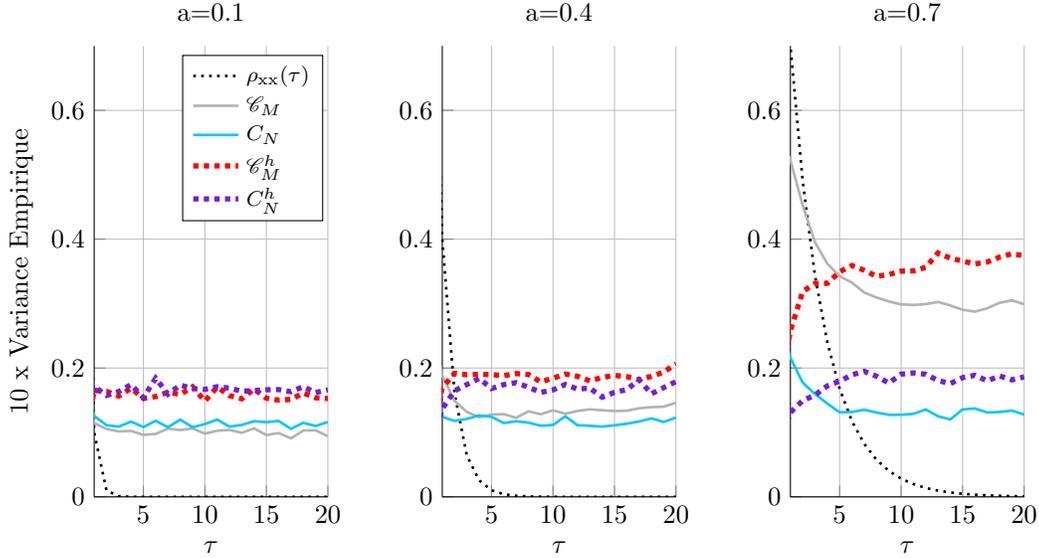


FIGURE 5.4 – Variance empirique de différents estimateurs de corrélation d’un processus AR(1) pour trois valeurs du coefficient  $a$  : 0.1, 0.4, 0.7. Cette figure compare les estimateurs hybride et classique ( $\mathcal{E}_M^h$ , et  $\mathcal{E}_M$ ) et leurs versions compressées ( $C_N^h$ , et  $C_N$ ), qui utilisent une matrice de compression creuse (de complexité  $\mathcal{O}(N)$ ). Les courbes résultent de 1000 réalisations indépendantes des processus, pour  $N = 1000, M = 100$ . Les estimateurs compressés s’avèrent meilleurs (moins de variance) que leurs contreparties classiques, si le processus est suffisamment corrélé en temps ( $a = 0.4, 0.7$ ).

D’abord, on voit que les deux estimateurs hybrides (tracés en trait pointillé) ont un comportement semblable à celui des estimateurs sur 1 bit. C’est-à-dire que ces estimateurs sont performants lorsque la corrélation est forte (parce que la variance décroît lorsque  $\rho_{xx}(\tau)$  augmente). Par contre, leur performance se dégrade lorsque la valeur de corrélation à estimer est faible (car la variance est maximale pour  $\rho_{xx}(\tau)$  tendant vers 0). Nous répétons que ce comportement est l’opposé de celui des estimateurs quantifiés finement.

Ensuite, on constate que la relation entre ces deux estimateurs est la même que celle qui existe entre les autres corrélateurs classiques et compressés : si le processus est très riche en fréquences ( $a \rightarrow 0$ ), l’estimateur hybride classique est plus performant (il a une variance inférieure). Par contre, s’il s’agit plutôt d’un signal à support fréquentiel étroit ( $a \geq 0.4$ ), l’estimateur compressé est meilleur. L’estimateur compressé hybride est donc préférable à l’hybride classique lorsque l’on traite des signaux *compressibles*.

Enfin, on peut voir que si le processus est fortement redondant ( $a = 0.7$ ), l’estimateur compressé hybride (en trait bleu pointillé) s’avère même meilleur que le classique (en trait gris continu). Dans ce cas, on obtient non seulement une meilleure performance, mais aussi une très forte réduction en termes de volume de données transmises, car le premier ne transmet que  $M$  bits, là où le deuxième envoie  $M$  échantillons quantifiés sur plusieurs bits.

Notre deuxième comparaison concerne les estimateurs à quantification grossière : hybrides et sur 1 seul bit. Elle est présentée sur la figure 5.5.

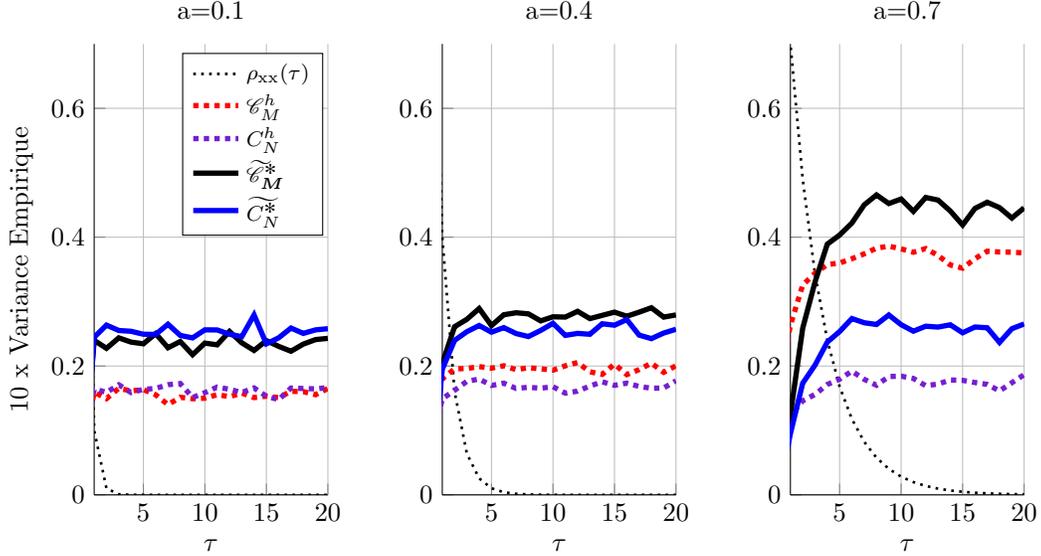


FIGURE 5.5 – Variance empirique de différents estimateurs de corrélation. Les conditions sont les mêmes que celles de la figure 5.4. Cette figure compare les estimateurs hybride et sur un bit ( $\mathcal{C}_M^h$ , et  $\widetilde{\mathcal{C}}_M^*$ ) et leurs versions compressées ( $C_N^h$ , et  $\widetilde{C}_N^*$ ). Lorsqu'il s'agit d'estimer une valeur de corrélation faible, les estimateurs hybrides ( $\mathcal{C}_M^h$ , et  $C_N^h$ ) s'avèrent meilleurs que les corrélateurs sur un bit ( $\widetilde{\mathcal{C}}_M^*$  et  $\widetilde{C}_N^*$ ). Néanmoins, cette tendance s'inverse quand la corrélation à estimer est forte.

On a vu précédemment que ces quatre estimateurs sont performants lorsque l'on cherche à estimer une valeur de corrélation forte mais que leur précision se dégrade lorsque la valeur recherchée est faible. Cette figure montre que si on quantifie sur plusieurs bits les mesures du capteur local  $y$  (celles qui ne sont pas transmises), la précision obtenue pour estimer une corrélation faible s'améliore. En effet, lorsque  $\rho_{xx}(\tau) = 0$  :  $\mathbf{Var}\{\mathcal{C}_M^h\} < \mathbf{Var}\{\widetilde{\mathcal{C}}_M^*\}$  et  $\mathbf{Var}\{C_N^h\} < \mathbf{Var}\{C_N^*\}$ . Dans ce cas, les variances des estimateurs hybrides sont inférieures à leurs contreparties sur 1 bit. Notons cependant que l'effet contraire apparaît si la valeur de corrélation à estimer est forte. La précision de l'estimation se dégrade, car lorsque  $\rho_{xx}(\tau)$  augmente :  $\mathbf{Var}\{\mathcal{C}_M^h\} > \mathbf{Var}\{\widetilde{\mathcal{C}}_M^*\}$  et  $\mathbf{Var}\{C_N^h\} > \mathbf{Var}\{\widetilde{C}_N^*\}$ . Les estimateurs hybrides s'avèrent donc préférables lorsque la corrélation entre les deux signaux traités est faible. Nous allons illustrer plus clairement cet aspect dans le chapitre suivant, consacré à l'étude de signaux réels.

À partir de ces deux nouveaux estimateurs de corrélation, on peut s'intéresser à des calculs complémentaires, tel qu'il a été fait dans le chapitre 3. Dans le paragraphe suivant nous étudions l'estimation de spectres d'interaction construits à l'aide des estimateurs de corrélation hybrides et sur 1 bit.

### 5.3 Corrélogrammes compressés et quantifiés

Voyons d'abord les estimateurs d'interspectre construits à l'aide des corrélateurs sur 1 bit. Tous s'appuient sur la même forme que le corrélogramme de Blackman-Tukey (BT), présenté au chapitre 4. Celui-ci n'est autre que la transformée de Fourier de la fonction de corrélation estimée, pondérée par une fenêtre  $w$ . Cette dernière est à support borné  $[-1, 1]$  et présente son maximum en 0 où elle vaut 1. Elle est également paire, bornée et de carré intégrable.

### 5.3.1 Corrélogrammes sur 1 bit

#### Corrélogramme classique sur 1 bit

$$\widetilde{\mathcal{F}}_M^*(\nu) = \sum_{k=-L}^L \widetilde{\mathcal{C}}_M^*(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

où  $\widetilde{\mathcal{C}}_M^*(\tau)$  est le corrélateur classique sur 1 bit.  $L$  est le retard maximum estimé.

Nous nous proposons de comparer cet estimateur à celui obtenu en remplaçant l'estimateur classique de corrélation par l'estimateur compressé sur 1 bit :

#### Corrélogramme compressé sur 1 bit

$$\widetilde{S}_N^*(\nu) = \sum_{k=-L}^L \widetilde{C}_N^*(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

où  $\widetilde{C}_N^*(k)$  est le corrélateur compressé sur 1 bit.

### 5.3.2 Corrélogramme classique sur 1 bit

D'après les résultats présentés en annexe les propriétés de cet estimateur sont les suivantes :

#### Covariance

$$\begin{aligned} \lim_{M \rightarrow +\infty} \frac{M}{L_M} \mathbf{Cov}\{\widetilde{\mathcal{F}}_M^*(\nu_1), \widetilde{\mathcal{F}}_M^*(\nu_2)\} = \\ = \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ \left[\frac{\pi}{2}\right]^2 \psi_{xx}(\nu)\psi_{yy}(\nu) \times \left[\int_{-1}^1 w^2(x) dx\right] & \text{si } \nu_1 = \nu_2 \neq 0, \frac{1}{2} \\ \left[\frac{\pi}{2}\right]^2 (|\psi_{xy}(\nu)|^2 + \psi_{xx}(\nu)\psi_{yy}(\nu)) \times \left[\int_{-1}^1 w^2(x) dx\right] & \text{si } \nu_1 = -\nu_2 = 0, \frac{1}{2} \end{cases} \end{aligned}$$

L'écriture  $\psi(\nu)$  dénote les densités des signaux quantifiés sur 1 bit.

#### Variance

$$\begin{aligned} \lim_{M \rightarrow +\infty} \frac{M}{L_M} \mathbf{Var}\{\widetilde{\mathcal{F}}_M^*(\nu)\} = \\ \begin{cases} \left[\frac{\pi}{2}\right]^2 \psi_{xx}(\nu)\psi_{yy}(\nu) \times \left[\int_{-1}^1 w^2(x) dx\right] & \text{si } \nu \neq 0, \frac{1}{2} \\ \left[\frac{\pi}{2}\right]^2 (|\psi_{xy}(\nu)|^2 + \psi_{xx}(\nu)\psi_{yy}(\nu)) \times \left[\int_{-1}^1 w^2(x) dx\right] & \text{si } \nu = 0, \frac{1}{2} \end{cases} \end{aligned}$$

### 5.3.3 Corrélogramme compressé sur 1 bit

D'après le calcul présenté en annexe on a :

## Covariance

$$\lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\widetilde{S}_N^*(\nu_1), \widetilde{S}_N^*(\nu_2)\} \approx$$

$$\left[\frac{\pi}{2}\right]^2 \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N^*(\nu_1), \mathcal{S}_N^*(\nu_2)\}$$

$$+ \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ \left[\frac{\pi}{2}\right]^2 (C_{4,\varphi} + N/M)\psi_{yy}(\nu) \left[\int_{-1}^1 w^2(x)dx\right] & \text{si } \nu_1 = \nu_2 = \nu \end{cases}$$

## Variance

$$\lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Var}\{S_N(\nu)\} \approx$$

$$\left[\frac{\pi}{2}\right]^2 \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Var}\{\mathcal{S}_N^*(\nu)\} + \left[\frac{\pi}{2}\right]^2 (C_{4,\varphi} + N/M)\psi_{yy}(\nu) \left[\int_{-1}^1 w^2(x)dx\right]$$

D'après ce résultat, le comportement du corrélogramme compressé sur 1 bit s'écrit sous la même forme que celui du corrélogramme compressé étudié au chapitre 3. Deux différences importantes apparaissent : la première est le facteur  $\left[\frac{\pi}{2}\right]^2$ . La deuxième est que le comportement n'est plus fonction du spectre des signaux, mais des propriétés spectrales des signaux quantifiés sur 1 bit.

Nous allons maintenant étudier le comportement des corrélogrammes à forte quantification au moyen de simulations Monte Carlo. Comme nous l'avons fait dans le chapitre 4 pour les corrélogrammes, nous étudierons le cas particulier des processus AR(1) et AR(2) gaussiens.

### 5.3.4 Compressibilité et performance

Pour comprendre l'influence de la nature du signal dans la performance, nous considérons des signaux dont la *compressibilité* est variable. Dans le cas du processus AR(1) on choisit à nouveau les cas  $a = 0.1, 0.4$  et  $0.7$ . Pour le processus d'ordre 2, on considère des processus avec pôles complexes conjugués :  $p_{1,2} = re^{\pm j\theta}$ . L'angle  $\theta = 0.125$  reste constant et on teste différentes valeurs de  $r$ ,  $0.3, 0.6$  et  $0.8$ .<sup>6</sup> Pour évaluer la variance de ces corrélogrammes, nous avons utilisé 1000 simulations indépendantes pour lesquelles  $N = 1000$  et  $M = 100$ , soit un taux de compression  $\alpha = 10$ . La densité spectrale de puissance est calculée à l'aide d'une transformée de Fourier rapide appliquée à la corrélation sur 256 points : les  $2L + 1 = 41$  retards effectifs, complétés par 215 zéros pour lisser la courbe. Autrement dit, la fonction de pondération  $w$  est une fenêtre rectangulaire de 41 points de largeur.

La figure 5.6 illustre les variances empiriques des corrélogrammes sur 1 bit classique ( $\mathcal{S}_M^*$ ) et compressé ( $\widetilde{S}_N^*$ ). À titre de comparaison, nous avons également tracé le comportement des corrélogrammes classique ( $\mathcal{S}_M$ ) et compressé ( $S_N$ ), étudiés dans le chapitre 3, ainsi que les variances asymptotiques théoriques de tous ces estimateurs.

6. Nous rappelons que pour  $r \rightarrow 0$ , le processus obtenu est un bruit blanc (un signal qui n'est pas compressible). Par contre, si  $r \rightarrow 1$ , on obtient plutôt un signal sinusoïdal, de fréquence  $\theta$  (un signal fortement compressible).

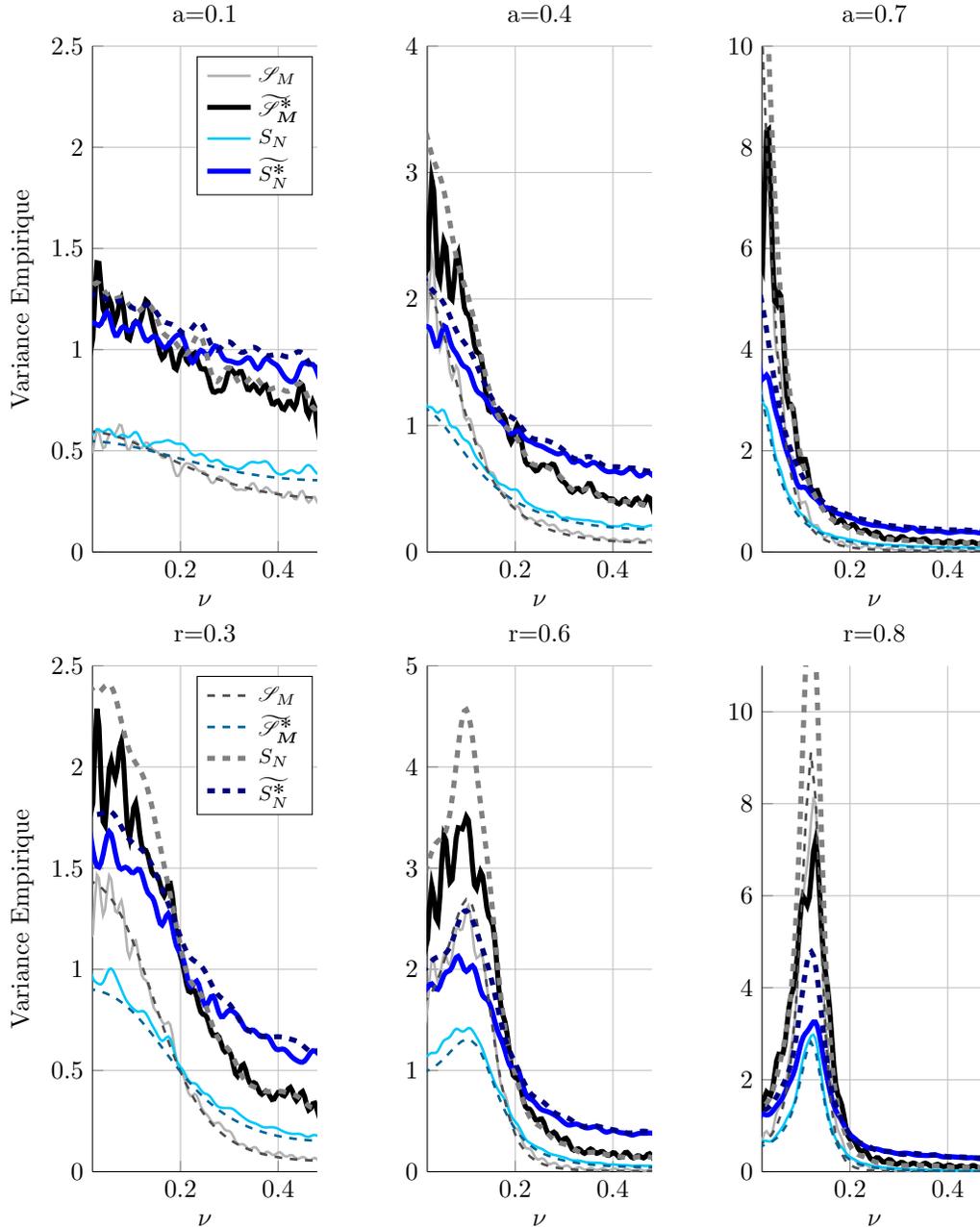


FIGURE 5.6 – Variance empirique d’estimation spectrale pour différents estimateurs (trait continu) et variance asymptotique théorique (trait pointillé). Haut : processus AR(1), Bas : AR(2). Les courbes émanent de la simulation de 1000 réalisations indépendantes. On prend  $N = 1000$ ,  $\alpha = 10$ , une fenêtre de pondération rectangulaire de demi-largeur  $L = 20$  et une transformée de Fourier calculée sur 256 points. Les corrélogrammes sur un bit (classique et compressé) ont une relation analogue à celle retrouvée pour les corrélogrammes classique et compressé. La version compressée de l’estimateur peut s’avérer plus performante que le classique. Ceci est vrai uniquement pour des processus *compressibles* et uniquement dans la région du spectre où le processus concentre son énergie. La quantification sur un bit dégrade la précision car la variance des corrélogrammes sur un bit (en trait gras) est supérieure à celle des estimateurs quantifiés finement (en trait fin).

Sur la figure 5.6 on peut voir que les résultats asymptotiques théoriques<sup>7</sup>, établis pour les variances dans ce chapitre et le 3, concordent assez bien avec les résultats obtenus par simulation. Les estimateurs compressés sont en effet plus performants que les classiques lorsque le processus est suffisamment redondant ( $a \geq 0.4$  et  $r \geq 0.3$ ). Cette affirmation n'est pourtant valable que dans la région du spectre qui concentre la puissance du signal. Ailleurs, dans la région du spectre où la puissance du signal est faible, ce sont les estimateurs classiques qui s'avèrent meilleurs. Notons que quantifier les signaux sur un bit dégrade la performance d'estimation.

Passons maintenant au paragraphe suivant, qui montre que les méthodes hybrides parviennent à atténuer ce problème.

### 5.3.5 Corrélogrammes hybrides

L'étude qui suit concerne les estimateurs hybrides. À notre connaissance, aucun de ces estimateurs n'a été étudié de manière analytique. Faute d'une analyse théorique, nous nous appuyons de nouveau sur des simulations de type Monte Carlo, réalisées sous les mêmes conditions que celles du paragraphe précédent.

Voici le corrélogramme classique hybride :

#### Corrélogramme classique hybride

$$\mathcal{S}_M^h(\nu) = \sum_{k=-L}^L \mathcal{C}_M^h(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

où  $\mathcal{C}_M^h(\tau)$  est le corrélateur classique hybride.

Et sa version compressée :

#### Corrélogramme compressé hybride

$$S_N^h(\nu) = \sum_{k=-L}^L C_N^h(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

$C_N^h(\tau)$  désignant le corrélateur compressé hybride.

Les signaux simulés sont à nouveau des processus autorégressifs, d'ordre 1 avec  $a = 0.1, 0.4$  et  $0.7$ , ou d'ordre 2 avec  $\theta = 0.125$  et  $r = 0.3, 0.6$  ou  $0.8$ . Le nombre de mesures disponibles est  $N = 1000$  ramené à  $M = 100$  échantillons, soit un taux de compression  $\alpha = 10$ . Chaque densité spectrale de puissance est une transformée de Fourier rapide, appliquée à la corrélation estimée sur  $2L + 1 = 41$  retards, complétés par 215 zéros. Les variances sont calculées de manière empirique, à partir de 1000 réalisations indépendantes.

7. Dans les cas des estimateurs quantifiés sur 1 bit les propriétés asymptotiques s'écrivent en fonction de la densité spectrale du signal quantifié. Faute d'une expression analytique,  $\psi_{xx}(\nu)$  a été estimée à partir du signal quantifié sur 1 bit.

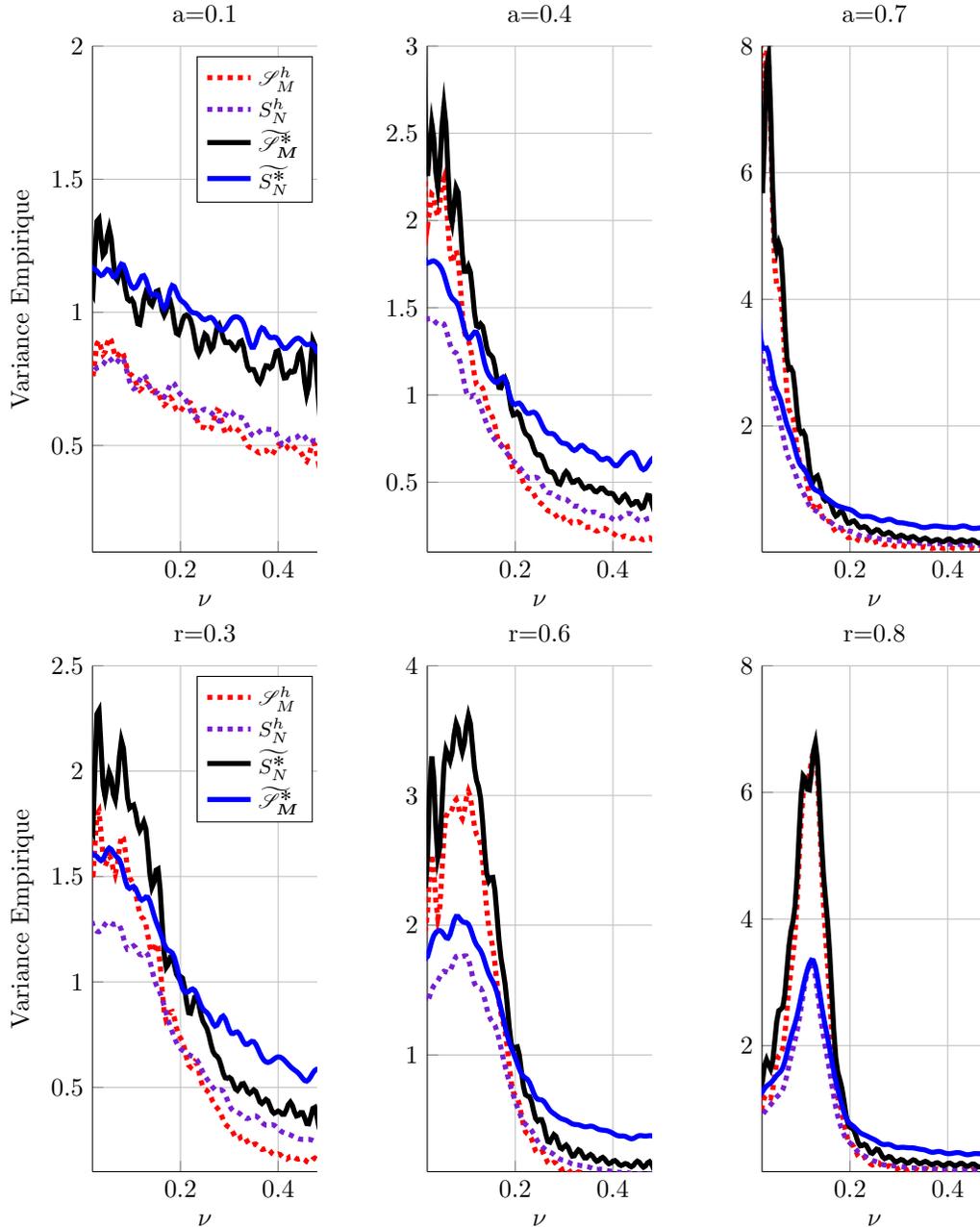


FIGURE 5.7 – Variance empirique d’estimation spectrale pour différents estimateurs. Haut : processus AR(1), Bas : AR(2). Les courbes découlent de 1000 réalisations indépendantes, obtenues sous les mêmes conditions que celles de la figure 5.6. Les corrélogrammes hybrides (classique et compressé) ont aussi un lien entre eux, analogue à celui retrouvé pour les corrélogrammes numériques et les corrélogrammes sur un bit. L’estimateur compressé hybride est plus performant que le classique pour des processus *compressibles*, mais seulement dans la zone du spectre qui concentre la puissance du signal. Les méthodes hybrides permettent d’améliorer la précision par rapport aux méthodes sur 1 bit. En effet, la variance des corrélogrammes hybrides (en trait pointillé) est inférieure à celle des estimateurs sur un bit (en trait mince).

La figure 5.7 illustre le comportement de la variance des corrélogrammes hybrides (classique et compressé), ainsi que celui des corrélogrammes sur un bit étudiés dans le paragraphe précédent. Dans le cas des estimateurs hybrides de densité spectrale, on retrouve le phénomène déjà décrit : l'estimateur compressé s'avère meilleur que sa contrepartie classique pour des processus dont le support fréquentiel est étroit (des signaux *compressibles*). On voit aussi que les méthodes hybrides permettent de gérer le compromis entre précision et volume de données à transmettre.

Pour conclure ce chapitre, nous comparons la performance des estimateurs hybrides qui transmettent  $M$  bits, avec celle des estimateurs numériques qui transmettent  $M$  valeurs quantifiées sur plusieurs bits.

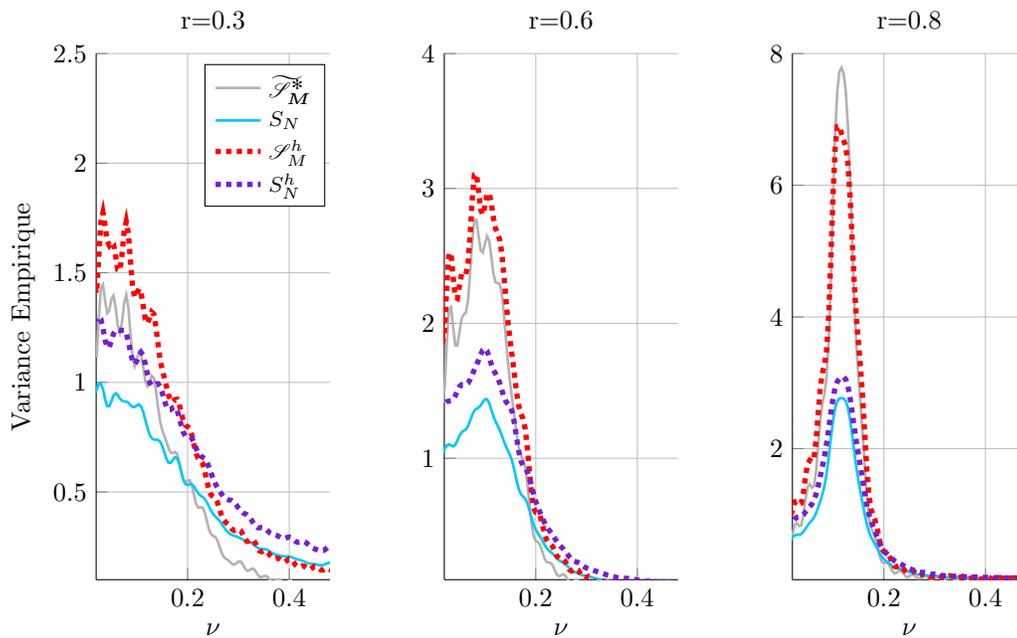


FIGURE 5.8 – Variance empirique d'estimation spectrale pour les corrélogrammes hybrides et numériques, dans leurs versions classique et compressée. Processus AR(2) sous les mêmes conditions que la figure 5.7. Le corrélogramme hybride compressé offre un bon compromis entre performance d'estimation et transmission de données. Bien qu'il transmette uniquement  $M$  bits, sa performance est comparable à celle de l'estimateur compressé qui transmet  $M$  valeurs quantifiées sur plusieurs bits.

La figure 5.8 nous montre à quel point les méthodes hybrides sont avantageuses. Si le processus est très compressible, la variance de l'estimateur hybride compressé est comparable à celle de l'estimateur compressé, même si le premier transmet significativement moins de données que le deuxième.

## Conclusion

Dans ce cinquième chapitre, nous avons montré que la compression peut être combinée avec la quantification brutale pour traiter les signaux avec une méthode simple qui permet de réduire encore plus les communications.

Les deux nouveaux estimateurs compressés discutés dans ce chapitre (hybride et sur un bit) présentent un comportement analogue au corrélateur compressé, discuté au chapitre 3. Encore une fois, ces méthodes ne sont préférables à leurs homologues

classiques que si l'on travaille avec des signaux compressibles.

Compte tenu de la difficulté de l'étude analytique, notre étude repose essentiellement sur des simulations Monte Carlo. Ces simulations ont rempli deux objectifs : elles ont permis de comprendre l'effet de la quantification dans les différentes tâches d'estimation et aussi de valider certains résultats théoriques asymptotiques, en particulier ceux qui concernent l'analyse spectrale.

Comme on pouvait s'y attendre, l'exploitation de la quantification se traduit par une perte de performance, qui se manifeste le plus clairement dans l'estimation spectrale. Selon nos résultats, les méthodes hybrides offrent le meilleur compromis entre transmission de données et qualité d'estimation. Nous sommes certains que l'étude théorique de leurs propriétés mérite d'être poursuivie.

Pour terminer ce manuscrit, nous allons maintenant appliquer toutes les méthodes étudiées à des données réelles issues de l'application qui nous intéresse, l'étude de l'état physique des bâtiments.

## Bibliographie

- [1] DAVID, F. A note on the evaluation of the multivariate normal integral. *Biometrika* 40, 3-4 (1953), 458–459.
- [2] GABRIEL, K. Comparison of three correlation coefficient estimators for gaussian stationary processes. *IEEE transactions on acoustics, speech, and signal processing* 31, 4 (1983), 1023–1025.
- [3] KOH, T., AND POWERS, E. Efficient methods of estimate correlation functions of gaussian processes and their performance analysis. *IEEE transactions on acoustics, speech, and signal processing* 33, 4 (1985), 1032–1035.
- [4] MCFADDEN, J. The fourth product moment of infinitely clipped noise. *IRE Transactions on Information Theory* 4, 4 (1958), 159–162.
- [5] MCFADDEN, J. Two expansions for the quadrivariate normal integral. *Biometrika* 47, 3/4 (1960), 325–333.
- [6] OEHLERT, G. W. A note on the delta method. *The American Statistician* 46, 1 (1992), 27–29.
- [7] RUCHKIN, D. Error of correlation coefficient estimates from polarity coincidences (corresp.). *IEEE Transactions on Information Theory* 11, 2 (1965), 296–297.
- [8] TONG, Y. *Multivariate normal distribution*. Springer-Verlag, 1990.

# Chapitre 6

## Application à des signaux réels

### Sommaire

---

<b>6.1 Structure A (Signaux modérément corrélés)</b>	<b>91</b>
6.1.1 Estimation de $\Gamma_{xy}(\tau)$	93
6.1.2 Estimation de $\Gamma_{xz}(\tau)$	95
<b>6.2 Structure L (Signaux faiblement corrélés)</b>	<b>98</b>
6.2.1 Estimation de $\Gamma_{xy}(\tau)$	99
6.2.2 Estimation de $\gamma_{xy}(\nu)$	101

---

Ce dernier chapitre s'intéresse à la performance des méthodes d'estimation étudiées dans ce manuscrit, appliquées à des données réelles. Les signaux traités correspondent à des vibrations d'édifices en service, enregistrées à l'aide d'un accéléromètre COLIBRYS SF3000L ([www.colibrys.com](http://www.colibrys.com)). Les signaux dont nous disposons<sup>1</sup> ont été enregistrés dans deux bâtiments différents, que nous appellerons **Structure A** et **Structure L**. Quelques informations les concernant seront fournies par la suite.

Cette étude considère l'estimation de l'intercorrélation et de l'interspectre des signaux, à partir desquels on peut déduire des indicateurs qui présentent un intérêt, dans le but de surveiller l'évolution de l'état de la structure, par exemple. Notre but est donc d'évaluer la pertinence des estimateurs compressés dans ce contexte.

Malheureusement, nous ne disposons pas de vérité terrain pour ces signaux réels. Faute de mieux, nous admettons que les résultats obtenus en analysant l'intégralité de ces données constituent l'objectif à atteindre, ce qui nous permet d'évaluer des biais et des variances empiriques lorsque nous n'en traitons que des sous-ensembles de taille réduite. Pour simplifier la présentation, nous omettons tout commentaire sur les unités utilisées : le temps, la fréquence et l'amplitude seront exprimés en unités arbitraires.

### 6.1 Structure A (Signaux modérément corrélés)

Cette première structure est le bâtiment de quinze étages qui apparaît sur la figure 6.1. L'instrument de mesure a été placé au sol du treizième étage. On dispose de trois signaux différents, dénotés  $x$ ,  $y$  et  $z$ , qui décrivent le mouvement que subit le bâtiment selon ses trois axes. Bien que ces mesures proviennent du même lieu, nous allons supposer qu'elles *ont été effectuées à des endroits différents du bâtiment*, par des capteurs différents, auquel cas les contraintes énergétiques dues à la transmission doivent être prises en compte.

---

<sup>1</sup>. Nous remercions Mikael Carmona et le CEA / LETI Grenoble d'avoir partagé ces données avec nous.



FIGURE 6.1 – Structure A : bâtiment résidentiel en service, doté de 15 étages. Les signaux proviennent du sol du treizième étage.

La figure 6.2 montre l’allure des signaux, ainsi que la densité spectrale de puissance et le spectrogramme de chaque signal, estimés avec l’ensemble des données disponibles (voir les détails sur la figure). D’après sa représentation temps fréquence, on voit que pour la fenêtre d’observation, il est raisonnable de considérer ces signaux comme stationnaires, car le contenu fréquentiel reste presque constant au cours du temps.

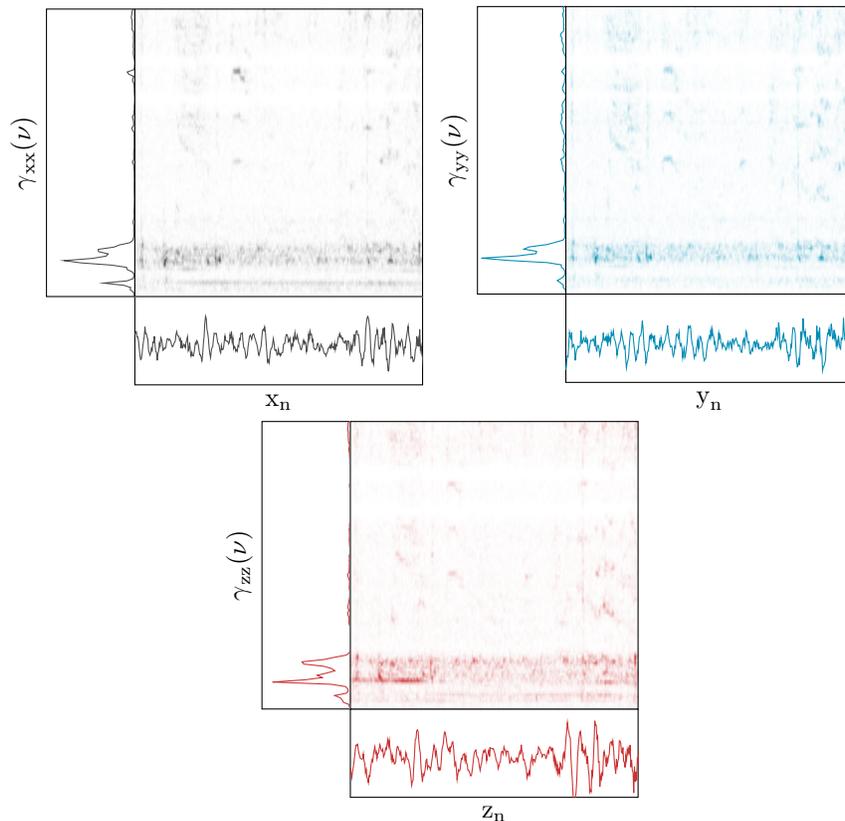


FIGURE 6.2 – Spectrogrammes des trois signaux étudiés (fenêtre de pondération rectangulaire de taille 500 et FFT calculée sur 512 échantillons). Plus l’amplitude est élevée, plus la couleur est sombre, les spectrogrammes montrent que la stationnarité est une hypothèse raisonnable pour chacun des signaux. Leur densité spectrale estimée respective est présentée à gauche de chaque spectrogramme. On constate qu’il s’agit de signaux à support fréquentiel étroit, signaux a priori *compressibles*. Toutes les quantités sont en unités arbitraires.

On voit que tous les signaux sont *compressibles* : comme le montrent leurs densités spectrales, le support fréquentiel des signaux est étroit. C'est le genre de signaux sur lequel les estimateurs compressés sont sensés être performants.

Dans nos études, les signaux  $x, y$  et  $z$  sont constitués de 12000 échantillons. Ils sont découpés en 12 blocs de taille  $N = 1000$ . Nous avons choisi un taux de compression  $\alpha = 10$ , soit un nombre de projections  $M = 100$ . L'évaluation de la performance d'un estimateur  $\hat{\Gamma}(\tau)$  est faite à partir de l'erreur quadratique moyenne :  $\mathbf{E}[(\hat{\Gamma}(\tau) - \Gamma(\tau))^2]$ , somme de la variance et du carré du biais  $\mathbf{E}[\hat{\Gamma}(\tau) - \Gamma(\tau)]$ . Ces deux quantités sont estimées de manière empirique, à partir des résultats obtenus sur chacun des 12 blocs consécutifs. Intéressons nous d'abord uniquement aux signaux  $x$  et  $y$ , qui sont les vibrations horizontales enregistrées par l'accéléromètre.

### 6.1.1 Estimation de $\Gamma_{xy}(\tau)$

Rappelons que dans le cas des estimateurs à forte quantification, nous avons supposé que la puissance du signal était connue. Pour satisfaire cette hypothèse et simplifier l'analyse, nous avons normalisé les signaux pour obtenir  $\sigma_x^2 = \sigma_y^2 = 1$  (ou encore  $\Gamma_{xx}(0) = \Gamma_{yy}(0) = 1$ ), auquel cas estimer les quantités normalisées et non normalisées est équivalent.

La figure 6.3 représente les fonctions d'autocorrélation puis les densités spectrales des deux signaux, estimées avec l'ensemble des données disponibles (12000 échantillons) ainsi que l'intercorrélation (la quantité à estimer) et la densité spectrale d'interaction (uniquement son module). Faute de connaître  $\Gamma_{xx}(\tau)$ , cette fonction est remplacée par l'estimation que nous pouvons en faire à partir de l'ensemble de données à disposition.

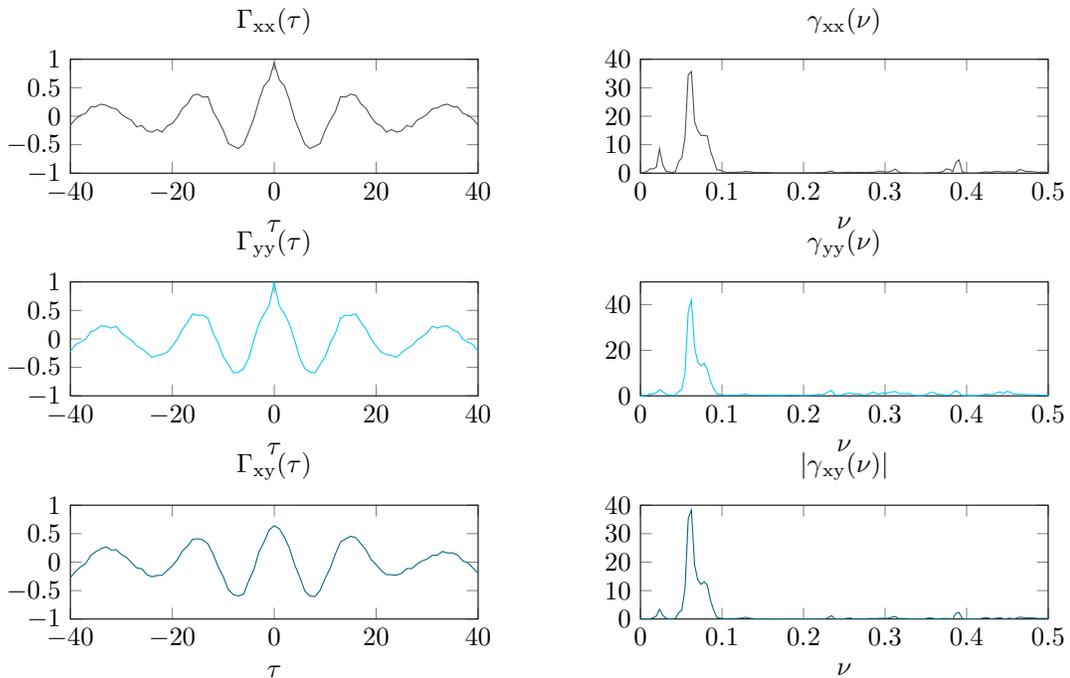


FIGURE 6.3 – Gauche : fonctions d'autocorrélation et d'intercorrélation des signaux. Droite : Densités spectrales de puissance des deux signaux et module de l'interspectre. Ces quantités sont estimées avec l'ensemble des données disponibles. Les signaux à traiter sont des processus très corrélés en temps. Leur spectre est donc à support étroit.

La figure 6.4 résume les performances obtenues pour chacun des estimateurs d'intercorrrelation. Elle présente les estimateurs déjà connus dans la littérature (classique, classique d'un bit et classique hybride), comparés aux versions compressées correspondantes que nous proposons dans ce manuscrit. Observons d'abord, que le biais estimé de manière empirique, diffère de zéro. Cela n'a rien d'étonnant car il est difficile d'avoir une bonne estimation avec uniquement 12 blocs. La valeur des biais est pourtant faible et du même ordre de grandeur pour tous les estimateurs considérés.

Pour juger la performance des estimateurs, observons plutôt l'erreur quadratique moyenne d'estimation (EQM) (qui prend en compte le biais et la variance). Ces performances concordent bien avec les prédictions effectuées précédemment, théoriquement ou avec des données synthétiques. En effet, les trois estimateurs compressés proposés sont plus performants que leurs homologues classiques (l'erreur quadratique moyenne obtenue est inférieure). Les méthodes compressées s'avèrent donc effectivement préférables aux méthodes classiques, sous réserve que le signal traité soit *compressible*.

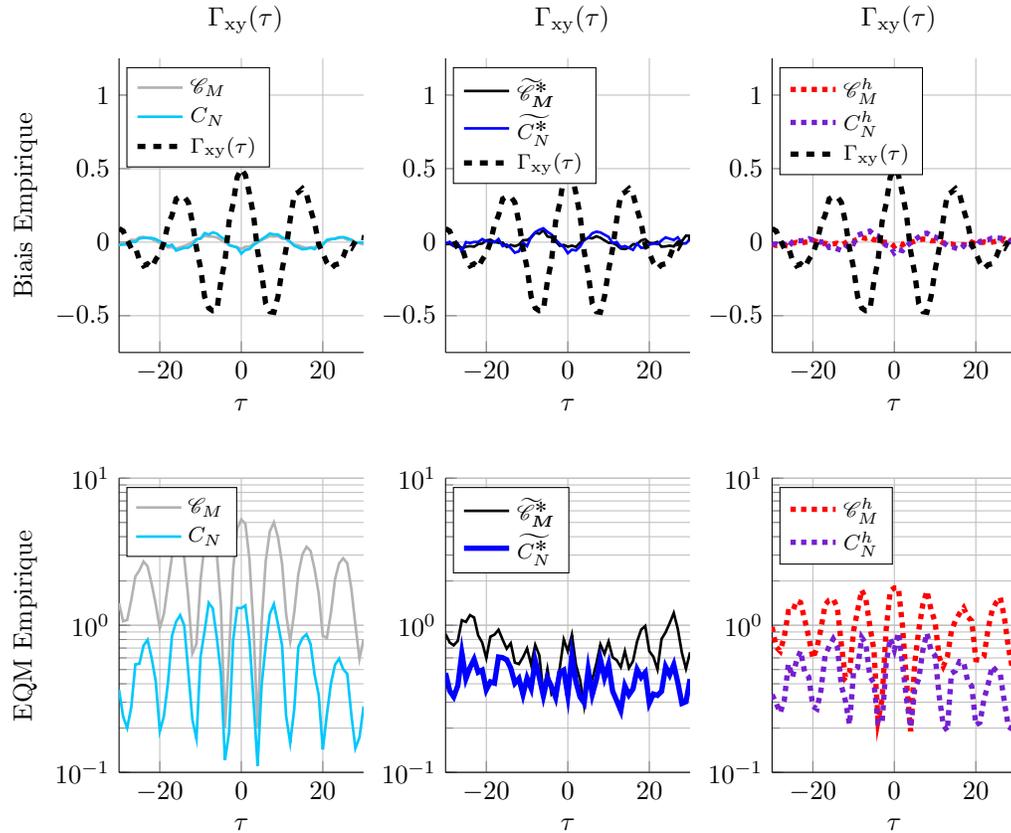


FIGURE 6.4 – Performance d'estimation. Signaux  $x$  et  $y$  de la structure A. Moyenne sur 12 blocs, avec  $N = 1000$  et  $M = 100$ . Haut : biais empirique. Bas : erreur quadratique moyenne. Trois types d'estimateurs sont considérés : Gauche : numériques, Centre : sur un bit, Droite : hybrides. Les estimateurs compressés :  $C_N$ ,  $\widehat{C}_N^*$ , et  $C_N^h$  présentent une erreur inférieure à leur équivalent classique. Ils sont donc préférables pour ce type de signal.

Pour certains retards, la fonction d'intercorrrelation présente des valeurs modérées ( $\geq 0.3$ ). En ces points, le corrélateur classique s'avère le pire des estimateurs, à savoir celui qui présente l'EQM la plus grande. Il s'agit d'un cas paradoxal puisque une

transmission de données plus économique ( $M$  bits) devient plus performante qu'une transmission plus chère ( $M$  échantillons multibit). Cela n'est cependant pas toujours le cas. La fonction d'intercorrrelation oscille, si bien que certains points approchent ou même atteignent zéro. En ces points, les estimateurs à forte quantification affichent de moins bonnes performances et c'est plutôt l'estimateur compressé (dont les échantillons sont quantifiés sur plusieurs bits) qui présente la meilleure performance. Nous rappelons la remarque effectuée au chapitre 5 : *les corrélateurs sur un bit et hybrides constituent de bons estimateurs si la valeur de corrélation à estimer est forte, mais de mauvais estimateurs lorsque la corrélation recherchée est faible*, ce qui explique ce comportement.

Afin de quantifier le gain en précision avec les méthodes compressées, on trace maintenant sur la figure 6.5 le rapport entre les EQM obtenues avec les méthodes classiques et celles obtenues avec leurs équivalents compressés. Ces trois courbes mettent en évidence que les estimateurs compressés sont tous meilleurs que leur équivalent classique. Le plus remarquable est le cas de l'estimateur compressé : il est en moyenne quatre fois plus précis que le classique. Les gains observés dans les autres cas sont plus faibles mais restent non négligeables. L'estimateur compressé hybride augmente la précision environ trois fois par rapport au classique hybride, tandis que le corrélateur compressé sur 1 bit est deux fois plus précis que le classique sur 1 bit.

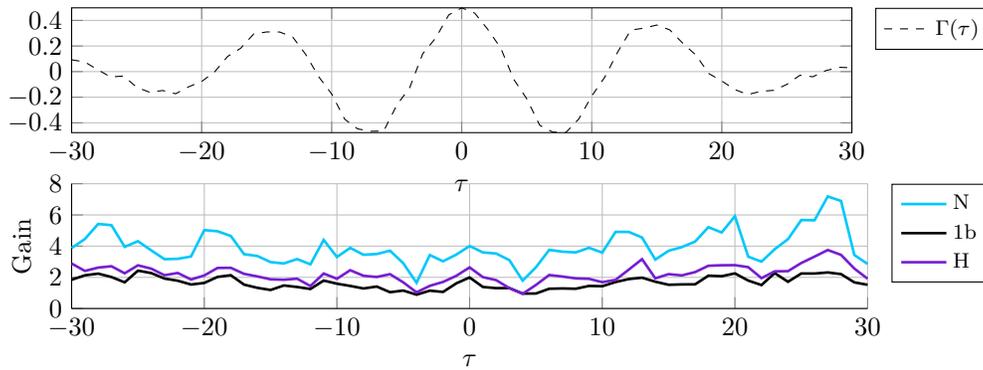


FIGURE 6.5 – Haut : fonction d'intercorrrelation à estimer. Bas : rapport des erreurs quadratiques moyennes, en fonction de  $\tau$ , sous la forme classique/compressé. (N : Numériques 1b : Sur 1 bit, H :Hybrides). Les estimateurs compressés sont meilleurs que leur équivalent classique. Pour les signaux étudiés, la précision augmente d'un facteur 4 avec l'estimateur compressé, 3 avec le compressé hybride et 2 avec le compressé sur 1 bit.

Passons maintenant à l'étude des signaux  $x$  et  $z$ .

### 6.1.2 Estimation de $\Gamma_{xz}(\tau)$

Le signal  $x$  déjà étudié est maintenant traité avec le signal  $z$ . La figure 6.6 montre les autocorrélations des deux signaux, ainsi que l'intercorrrelation et le module de l'interspectre. Ces courbes ont été calculées avec l'ensemble des données disponibles. Pour évaluer les différents estimateurs, on travaille sous les mêmes conditions que lors des simulations précédentes : 12000 échantillons, découpés en 12 blocs de  $N = 1000$  points.

A nouveau, pour juger la performance, nous étudions le biais et l'erreur quadratique moyenne, calculés avec les 12 blocs disponibles des signaux. Les performances obtenues pour l'estimation de l'intercorrrelation sont présentées dans la figure 6.7.

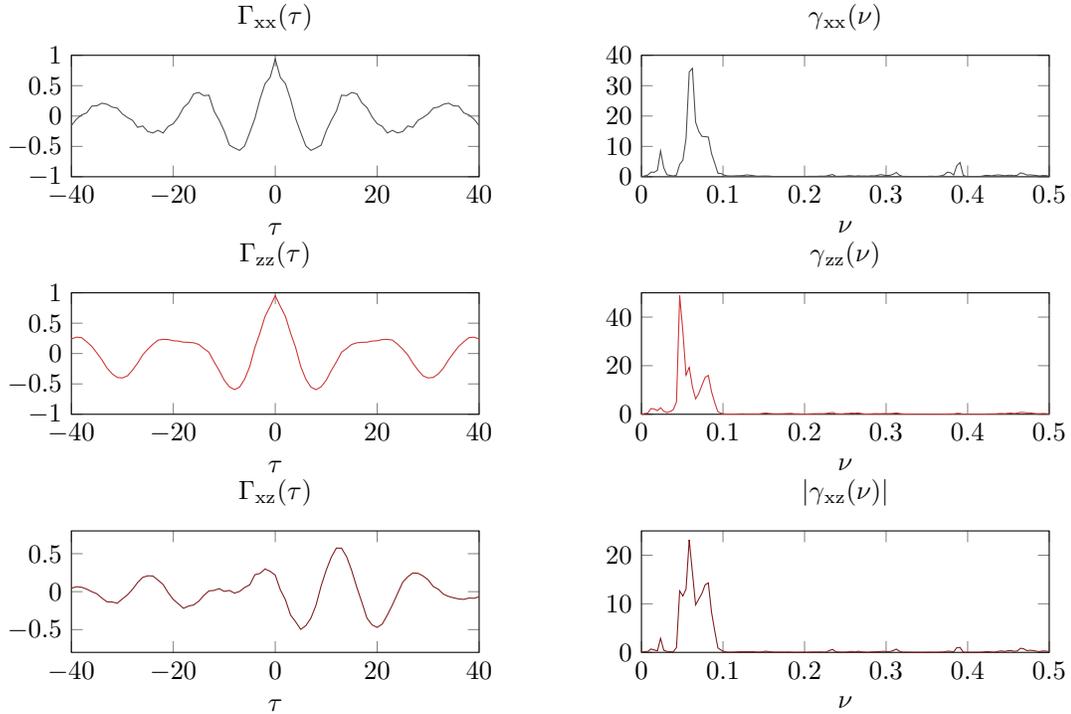


FIGURE 6.6 – Gauche : fonctions d'autocorrélation des signaux et leur intercorrélations. Droite : densités spectrales de  $x$ ,  $z$  et module de l'interspectre. Les deux signaux à traiter sont très corrélés en temps. Leur spectre étant à support borné, ils sont a priori *compressibles*. Ces courbes ont été évaluées à partir de l'ensemble des données disponibles, c'est-à-dire les 12000 échantillons des processus.

Portons d'abord notre attention sur les courbes séparément. Chaque estimateur classique est comparé à sa version compressée. On retrouve le même comportement que celui observé pour les signaux  $x$  et  $y$  : le biais d'estimation est du même ordre de grandeur (mais sa valeur est faible). Pour l'EQM, les versions compressées des corrélateurs produisent une erreur plus faible que leurs homologues classiques, quelle que soit la valeur de  $\tau$ .

Dans cette nouvelle fonction de corrélation à estimer, on a deux régions dont les caractéristiques sont distinctes. Dans la partie gauche ( $\tau < -10$ ), la corrélation est faible, tandis que dans la partie droite ( $\tau > -10$ ) la corrélation est modérée. Dans la partie gauche, les estimateurs sur 1 bit et hybride se retrouvent dans des conditions défavorables, au profit de l'estimateur compressé ( $C_N$ ) qui quantifie ses échantillons sur plusieurs bits. C'est lui qui affiche l'erreur la plus faible, suivi du corrélateur compressé hybride ( $C_N^h$ ) puis du compressé sur un bit ( $\widetilde{C}_N^*$ ). Les corrélateurs classiques ont tous les trois des performances inférieures. Pour cette gamme de retards, la réduction en complexité de la méthode d'estimation se paye par une perte en performance. Dans la partie droite les conditions s'améliorent pour les estimateurs à forte quantification. Ces estimateurs montrent des comportements comparables aux estimateurs qui quantifient sur plusieurs bits.

Enfin, la figure 6.8 montre le gain en précision obtenu avec les différentes méthodes. On a tracé le rapport entre l'EQM obtenue avec chaque estimateur classique divisée par l'EQM de l'estimateur compressé équivalent. Ce gain est différent selon la complexité des méthodes : 4 fois plus de précision dans le meilleur des cas (l'estimateur compressé par rapport au classique) et un facteur d'environ 2 pour l'amélioration la plus modeste

(l'estimateur compressé d'un bit par rapport au classique d'un bit).

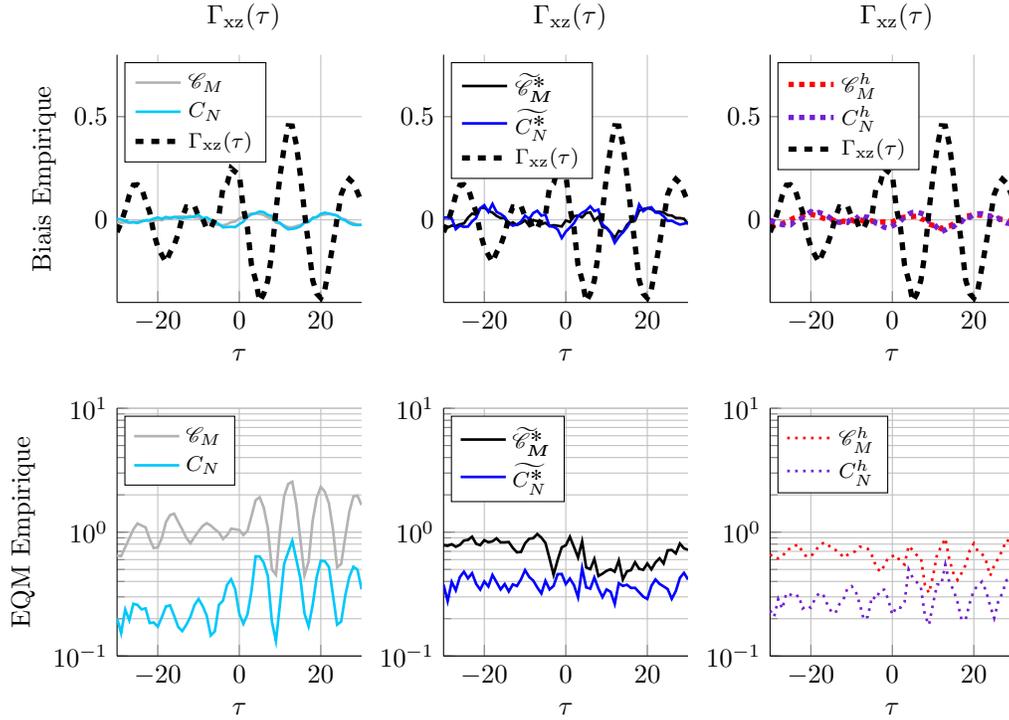


FIGURE 6.7 – Performance d'estimation de l'intercorrélation des signaux  $x$  et  $z$  de la structure A. L'estimation est évaluée sur 100 réalisations,  $N = 1000$ ,  $M = 100$ . Haut : moyenne des réalisations. Bas : erreur quadratique moyenne. Trois types d'estimateurs sont considérés : Gauche : numériques, Centre : d'un bit, Droite : hybrides. Ces quantités sont calculées à partir des 12 blocs du signal dont nous disposons. Les estimateurs compressés :  $C_N$ ,  $C_N^*$ ,  $C_N^h$  présentent une erreur inférieure à leur équivalent classique. Nos estimateurs sont donc préférables aux classiques pour ce type de signaux.

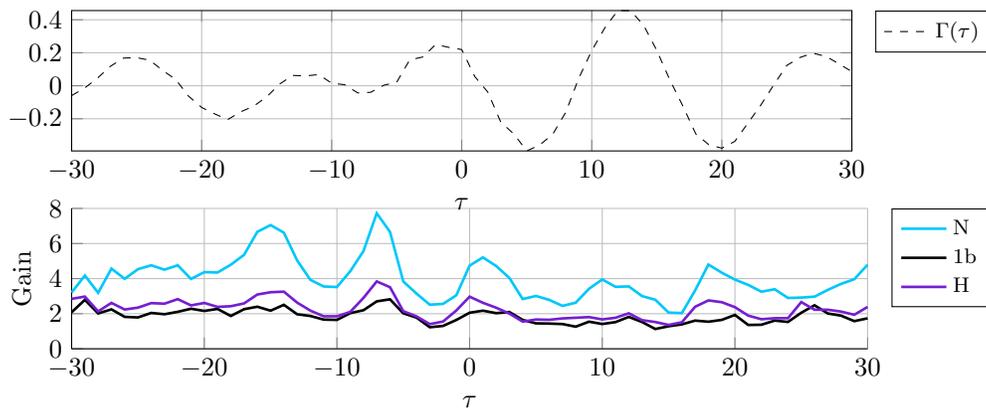


FIGURE 6.8 – Haut : fonction d'intercorrélation à estimer. Bas : rapport des erreurs quadratiques moyennes, en fonction de  $\tau$ , sous la forme classique/compressé. Les estimateurs compressés sont meilleurs que leurs équivalents classiques. La précision augmente d'un facteur 4 avec l'estimateur compressé, 3 avec le compressé hybride et 2 avec le compressé d'un bit.

Passons maintenant au deuxième jeu de données dont nous disposons, les signaux

provenant de la structure L.

## 6.2 Structure L (Signaux faiblement corrélés)

Notre deuxième exemple, la **Structure L** présentée dans la figure 6.9, est un pont qui supporte les rails d'un funiculaire utilisé comme salle blanche mobile à Grenoble. Dans ce cas, les signaux ont été enregistrés à la surface de l'une des poutres transversales en béton qui relient les rails.

On dispose uniquement de deux signaux (utiles)  $x$  et  $y$ , qui correspondent aux vibrations que subit la structure étudiée dans le plan parallèle au sol à un endroit donné. Comme on l'a fait avec le premier jeu de données, nous exploitons ces signaux comme s'ils provenaient d'endroits (et donc de capteurs) différents, situation dans laquelle on retrouve les contraintes énergétiques liées à la transmission des informations. Notre but est à nouveau d'étudier le gain en performance obtenu avec les méthodes compressées, par rapport aux méthodes dites *classiques*.



FIGURE 6.9 – Structure L : funiculaire utilisé comme salle blanche mobile. Les signaux vibratoires proviennent de l'une des poutres transversales en béton qui relient les rails.

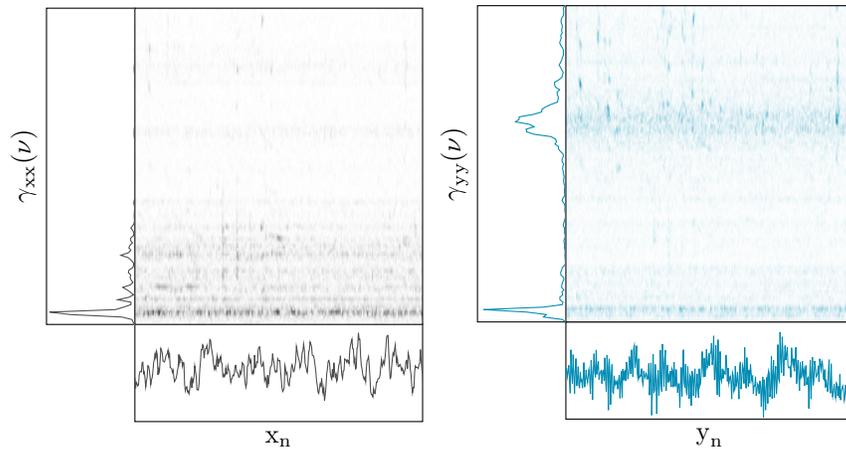


FIGURE 6.10 – Signaux étudiés : spectrogrammes (fenêtre de pondération rectangulaire de taille 500 et une FFT calculée sur 512 échantillons). Plus l'amplitude est élevée, plus la couleur est sombre. La stationnarité est une hypothèse raisonnable. Les densités spectrales estimées apparaissent à gauche de chaque spectrogramme. On constate que l'on traite des signaux *compressibles*.

Dans la figure 6.10 on montre les caractéristiques des signaux étudiés : l'allure, ainsi

que la densité spectrale de puissance et le spectrogramme de chaque signal, estimés avec l'ensemble des données disponibles (voir les détails dans la figure). Dans le spectrogramme on voit que, pour la fenêtre d'observation, il est raisonnable de considérer ces signaux comme stationnaires. Les supports fréquentiels étant étroits, il s'agit bien de signaux *compressibles*.

Dans nos études les signaux  $x$  et  $y$  comportent 35000 échantillons, que l'on a coupé en 35 blocs de longueur  $N = 1000$ . On utilise  $M = 100$ , soit un taux de compression  $\alpha = 10$ . L'évaluation de la performance de chaque estimateur  $\hat{\Gamma}(\tau)$  est faite à partir du biais  $\mathbf{E}[\hat{\Gamma}(\tau)] - \Gamma(\tau)$  et de l'erreur quadratique moyenne,  $\mathbf{E}[(\hat{\Gamma}(\tau) - \Gamma(\tau))^2]$ , estimés de manière empirique à partir des 35 différents blocs.

### 6.2.1 Estimation de $\Gamma_{xy}(\tau)$

Les propriétés à l'ordre 2 de ces signaux sont illustrées par la figure 6.11.

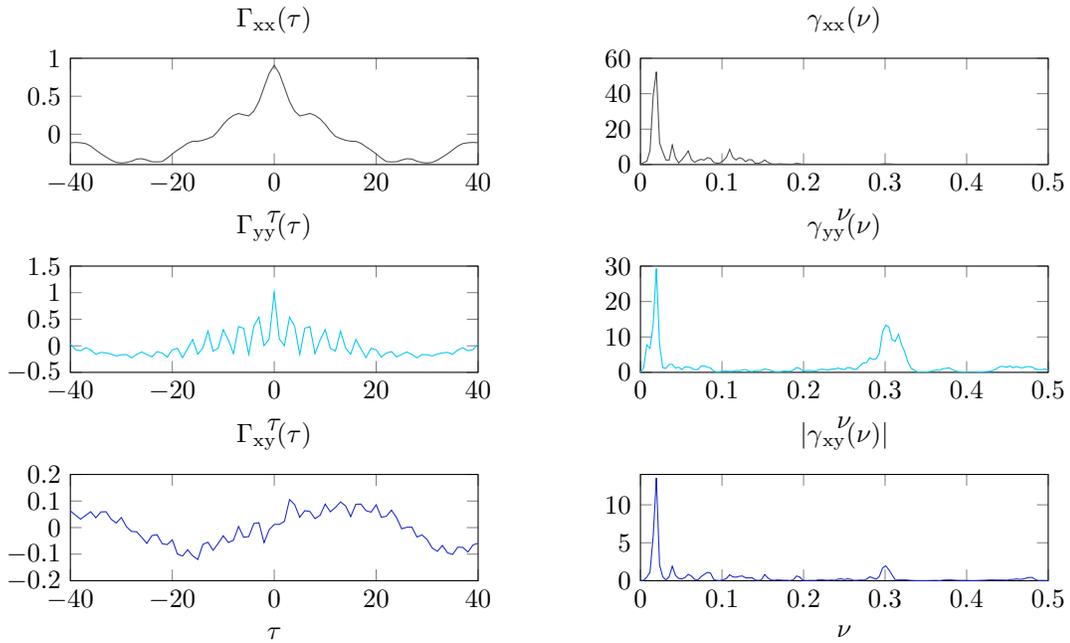


FIGURE 6.11 – Gauche : fonctions d'autocorrélation et d'intercorrélations des signaux. La corrélation entre  $x$  et  $y$  est faible. Droite : densités spectrales des deux signaux et module de l'interspectre. Ces quantités ont été calculées à partir de l'ensemble des 35000 échantillons disponibles. Les signaux à traiter sont *compressibles* car leur spectre est à support restreint.

Dans les fonctions d'autocorrélation, on peut voir la normalisation que nous avons effectuée pour simplifier notre étude. Les signaux sont de puissance unitaire car  $\Gamma_{xx}(0) = \sigma_x^2 = \Gamma_{yy}(0) = \sigma_y^2 = 1$ . Puis dans la fonction d'intercorrélations, on voit que contrairement aux cas précédents, cette fois-ci les signaux sont peu corrélés, ce qui implique des conditions défavorables pour les estimateurs à forte quantification (hybrides et sur 1 bit).

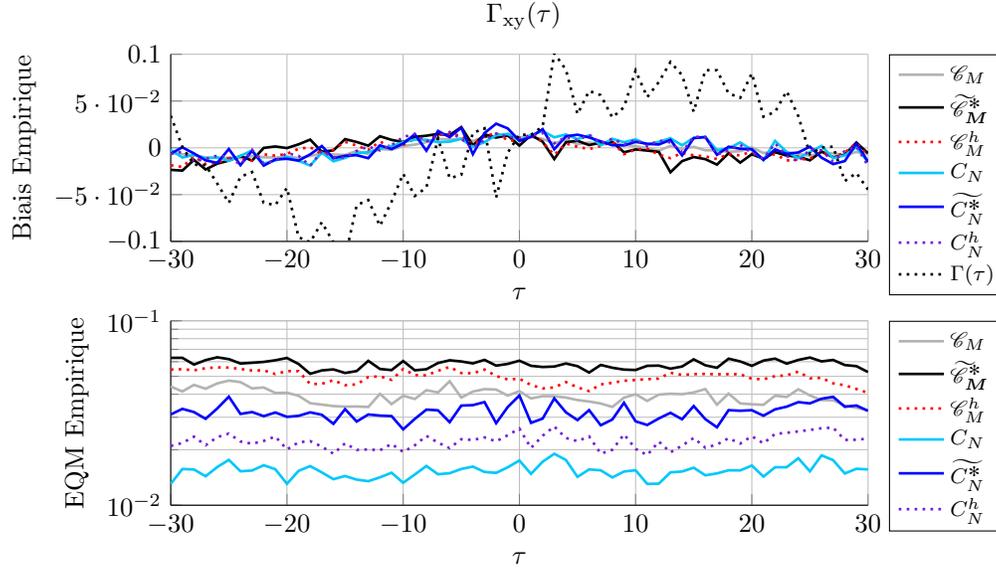


FIGURE 6.12 – Performance d’estimation pour les signaux  $x$  et  $y$  de la structure L. L’estimation est réalisée 35 fois, à partir de  $N = 1000$  mesures compressées en  $M = 100$  points. Haut : moyenne des réalisations. Bas : erreur quadratique moyenne. Les estimateurs compressés :  $C_N$ ,  $\widetilde{C}_N^*$ ,  $C_N^h$  présentent une erreur inférieure à leur équivalent classique. Lorsque la valeur de corrélation recherchée est faible, les méthodes à forte quantification (hybrides et d’un bit) voient leur performance se dégrader par rapport aux méthodes numériques (quantification multibit).

Les résultats obtenus par simulation pour l’estimation de corrélation sont tracés dans la figure 6.12. Elle nous montre que tous les estimateurs étudiés estiment l’allure de la courbe d’intercorrélation avec un certain biais. Théoriquement il devrait être nul, mais cela n’est pas le cas. L’estimation n’est effectuée que sur 35 blocs. Comme le biais est du même ordre de grandeur pour tous les estimateurs, la performance est plutôt évaluée selon l’erreur quadratique moyenne. Celle ci apparaît en bas de la figure. Toutes ces erreurs semblent invariantes selon le retard  $\tau$ . En effet, la valeur de corrélation que l’on cherche à estimer est faible dans tous les cas ( $\rho(\tau)$  tend vers 0). Les six estimateurs étudiés ont des performances distinctes, de sorte que nous pouvons les classer en deux paquets.

Un premier groupe est formé par les corrélateurs classiques ( $\widetilde{C}_M^*$ ,  $C_M^h$ ,  $C_M$ ) qui présentent les pires performances (l’EQM la plus grande). Le plus précis est l’estimateur classique et le moins performant est le classique d’un bit. On constate à nouveau que lorsque la valeur à estimer est faible, simplifier les méthodes d’estimation se traduit par une perte en performance : l’estimateur le plus économique du groupe est aussi le moins performant.

Le deuxième groupe est constitué des trois estimateurs compressés ( $C_N$ ,  $\widetilde{C}_N^*$ ,  $C_N^h$ ). Il est clair que chaque estimateur compressé est préférable à son équivalent classique, mais ce qui est le plus intéressant, c’est que les méthodes compressées quantifiées sur 1 bit, qui permettent d’économiser énormément en termes de transmission, s’avèrent plus performantes que l’estimateur classique qui quantifie ses valeurs transmises sur plusieurs bits et consomme plus d’énergie dans la transmission. Enfin, nous soulignons, d’après les résultats obtenus, que le prix à payer pour simplifier la méthode d’estimation est une perte en performance.

Pour terminer cette étude sur l'estimation de la corrélation, nous illustrons plus précisément le gain en performance associé aux méthodes compressées, c'est-à-dire l'EQM des estimateurs classiques divisée par l'EQM de leurs équivalents compressés en fonction de  $\tau$ . Ces rapports sont illustrés dans la figure 6.13. En effet, les méthodes compressées offrent environ deux fois plus de précision que les méthodes classiques pour l'estimation de corrélation sur ce type de signal.

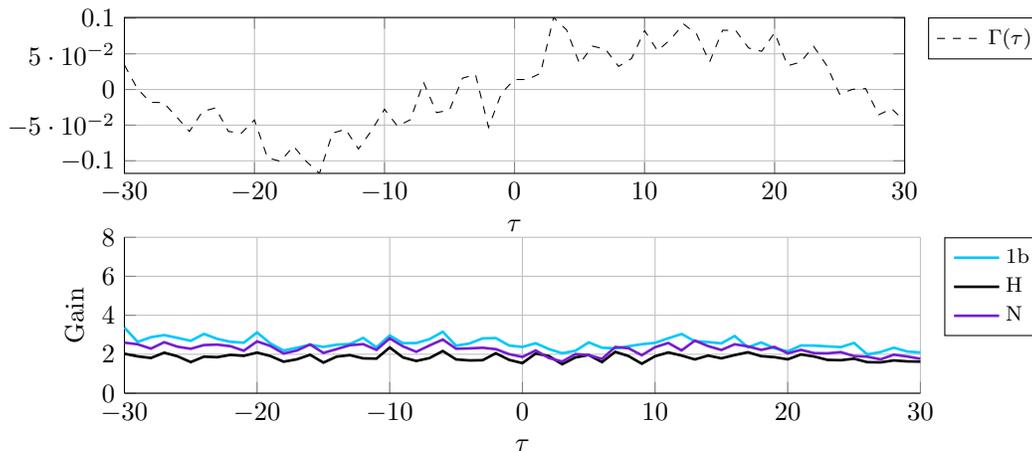


FIGURE 6.13 – Haut : fonction d'intercorrélation à estimer. Bas : rapport des erreurs quadratiques moyennes en fonction de  $\tau$ , sous la forme classique/compressé. Les estimateurs compressés sont meilleurs que leur équivalent classique. La précision augmente d'un facteur 2 avec tous les estimateurs compressés.

Notre dernière étude concerne l'estimation de l'interspectre à partir des fonctions de corrélation que l'on vient d'estimer.

### 6.2.2 Estimation de $\gamma_{xy}(\nu)$

Comme on l'a fait précédemment, nous allons observer uniquement le module de l'interspectre. Notre comparaison concerne les 3 corrélogrammes classiques et leurs équivalents compressés. Les conditions de simulation sont les mêmes que pour l'estimation de corrélation : on exploite  $L = 30$  retards avec une pondération rectangulaire pour calculer la transformée de Fourier sur 256 échantillons.

Les biais, ainsi que les erreurs quadratiques moyennes d'estimation du module de l'interspectre, sont tracés sur la figure 6.14. Ce jeu de données illustre assez bien le comportement des estimateurs compressés de densité spectrale. Notons d'abord que l'on retrouve une composante à basse fréquence et une autre à haute fréquence d'amplitude plus faible. Pour la première composante, les méthodes compressées produisent toutes une erreur plus faible que leur contrepartie classique. L'effet contraire se produit cependant sur la deuxième composante. À cet endroit les méthodes classiques sont plus performantes. D'après cet exemple, on constate que les corrélateurs compressés sont préférables si le signal traité est *compressible* et si la valeur spectrale recherchée est suffisamment loin de zéro.

Enfin, la figure 6.15 montre le gain en précision pour le module de l'interspectre. Cette fois-ci le gain en précision des méthodes compressées est assez important : la précision augmente 6 fois dans la partie qui contient le plus de puissance.

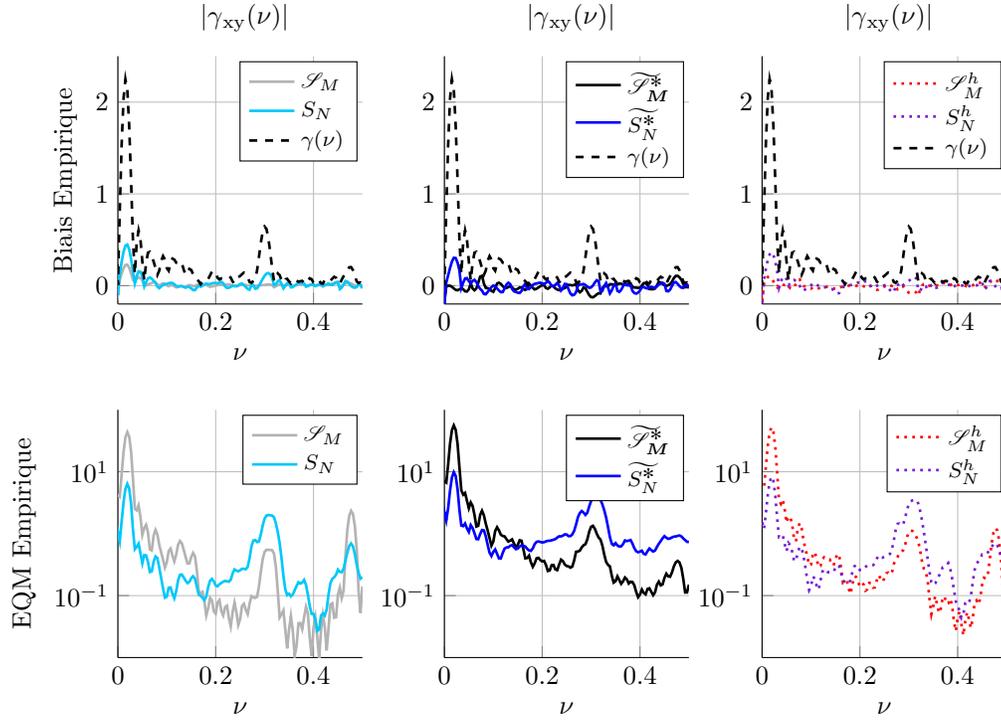


FIGURE 6.14 – Biais et EQM d'estimation du module de l'interspectre des signaux de la structure L. Moyenne empirique sur 35 blocs de taille  $N = 1000$ ,  $M = 100$ , pour  $L = 30$  et une transformée de Fourier calculée sur 256 échantillons. Haut : Biais. Bas : EQM. Les estimateurs compressés :  $S_N$ ,  $\widehat{S}_N^*$ ,  $S_N^h$  ont une erreur inférieure à leur équivalent classique pour ce type de signaux, mais uniquement lorsque la valeur du spectre est importante.

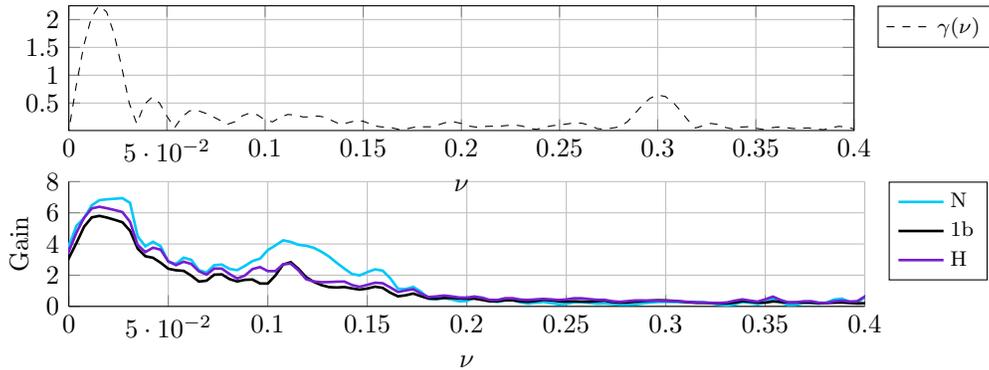


FIGURE 6.15 – Haut : module de l'interspectre à estimer. Bas : rapport des erreurs quadratiques moyennes, en fonction de  $\tau$ , sous la forme classique/compressé. Les estimateurs compressés sont meilleurs que leur équivalent classique. La précision augmente d'un facteur 6 avec les estimateurs compressés dans la région où le processus concentre sa puissance.

## Conclusion

Pour conclure ce paragraphe nous soulignons que les prédictions effectuées dans les chapitres précédents, qu'il s'agisse de résultats théoriques ou bien de simulations sur des données synthétiques, concordent bien avec les résultats sur les données réelles. Les

méthodes compressées que nous avons proposées peuvent s'avérer plus performantes que les méthodes classiques, sous les conditions que nous avons déjà décrites. Bien que l'on puisse exploiter n'importe quel signal, les méthodes compressées n'ont d'intérêt pratique que si les signaux traités sont *compressibles*.

L'affirmation précédente est aussi valable pour les estimateurs compressés et quantifiés sur 1 bit, qui offrent une réduction supplémentaire assez importante en termes de données transmises.

Une étude sur des données réelles enregistrées sous conditions plus rigoureuses et mieux adaptées aux traitements que l'on envisage mérite d'être poursuivie, par exemple avec des capteurs placés à des endroits différents, ou bien avec des jeux de données plus longs, afin d'améliorer la précision sur la performance des estimateurs.



# Conclusions et perspectives

Cette thèse s'inscrit dans la continuité des travaux menés au sein de l'équipe CICS au GIPSA-lab consacrés aux problèmes d'estimation à l'aide d'un réseau de capteurs. La contrainte majeure du contexte : la limitation des ressources énergétiques des capteurs a déjà motivé l'étude de deux outils particuliers : la quantification brutale [5, 8] et l'acquisition compressive [2]. Dans ce travail, nous cherchons à utiliser conjointement les deux techniques pour des problèmes d'estimation : corrélation, temps de propagation et interspectre.

Ceci nous a conduit à aborder les questions suivantes.

*Dans quelles conditions un traitement du signal compressé est-il pertinent ? Quel est l'impact de la compression sur la performance ? Est-il possible faire cohabiter compression et quantification grossière ? Comment le réaliser ?*

Nos réponses ne constituent que les premiers pas vers la pleine compréhension de ces outils de traitement.

## Principales contributions

### Traitement des signaux aléatoires

Nous avons repris les idées de Davenport [3] : traiter les signaux dans le domaine compressé et ainsi éliminer la phase de reconstruction et étendu ces outils au cas des processus aléatoires, ce qui nous a conduits à définir intuitivement la notion de *compressibilité* qui est analogue à la parcimonie dans des signaux déterministes.

### Étude généraliste sur matrice de compression $\Phi$

Contrairement aux travaux présentes la littérature, qui étudient seulement des cas particuliers, nous n'avons pas restreint le choix de la matrice. Ceci nous a conduit à évoquer le compromis entre complexité et précision de calcul. Nous avons également considéré le cas de l'échantillonnage aléatoire, modélisé comme une matrice aléatoire dont les lignes ne sont pas indépendantes. Nos résultats montrent que contrairement à ce que suggère l'intuition, le type de matrice de compression n'a que peu d'influence sur la performance.

### Description asymptotique des propriétés

Puisque l'analyse des propriétés des estimateurs à nombre d'échantillons fini est très lourde, nous avons eu recours à l'analyse asymptotique [9] qui permet de décrire la qualité d'estimation lorsque la taille des données traitées est très grande. Nos simulations et résultats expérimentaux montrent la pertinence de ces études asymptotiques pour évaluer la performance des traitements dans le domaine compressé.

Pour conclure ce manuscrit, nous réfléchissons sur les problèmes qui restent ouverts et sur de possibles directions pour la recherche future.

## Perspectives

### Concernant la performance des estimateurs

En ce qui concerne l'estimation spectrale, la description que nous avons donnée est incomplète. L'étape suivante consisterait à évaluer la qualité de l'estimation du module et de la phase de la densité spectrale d'interaction, ainsi que la cohérence, comme le font [1, 6] dans le cas classique.

Dans cette thèse, les estimateurs les plus intéressants sont malheureusement ceux pour lesquels les résultats théoriques font défaut. Comme nous l'avons constaté, les estimateurs hybrides offrent le meilleur compromis entre performance et économie d'énergie. L'étude théorique de leurs propriétés est sans doute l'aspect qui mérite la plus d'attention dans des travaux ultérieurs.

### Concernant la compression

Un aspect fondamental concerne le raffinement de la notion de *compressibilité* que nous avons introduite dans les cas des processus autorégressifs gaussiens. La théorie de l'information offre des indicateurs du caractère aléatoire d'un processus stochastique, comme le coefficient de planitude, aussi appelé entropie de Wiener (et ses généralisations)[4]. Nous sommes persuadés que l'inclusion de ces outils permettrait de comprendre plus précisément dans quelles conditions un traitement du signal compressé reste pertinent.

Deux effets secondaires de la compression par projection aléatoire que nous n'avons pas abordés sont. Premièrement le fait que  $\Phi \mathbf{x}$  subisse un étalement de spectre, tel qu'il est exploité en télécommunications, ce qui pourrait offrir une éventuelle robustesse contre les interférences en transmissions. Deuxièmement, la sécurité apportée (l'exploitation de données nécessite la connaissance exacte de la matrice de compression).

Un troisième aspect mériterait d'être exploré : l'exploitation de matrices structures (déterministes ou pseudo-aléatoires) afin de bénéficier de une complexité réduite ou d'algorithmes rapides dans le traitement. Non seulement le coût calculatoire pourrait diminuer, mais la performance pourrait aussi s'améliorer. Des travaux récents montrent que certains de ces matrices s'avèrent plus performantes que les constructions aléatoires, par exemple pour préserver la distance euclidienne[7].

### Concernant la quantification

Bien que nous ayons indiqué la capacité de compression aléatoire pour *gaussianiser* les signaux, nous n'avons pas étudié ce phénomène, comprendre dans quelles conditions il se produit permettrait d'établir plus précisément sous quelles conditions on peut compresser et quantifier pour une matrice donnée.

Enfin, le dernier aspect concerne l'étude de la quantification stochastique que nous avons exclue de cette étude. Nous pensons que dans certains cas cet outil peut être exploité conjointement, par exemple avec un échantillonnage aléatoire pour les tâches d'estimation. Nous répétons que la quantification stochastique est un autre outil qui généralise la quantification forte au-delà du cas gaussien.

Pour conclure, nous revenons sur la question fondamentale que nous avons évoquée au début de ce manuscrit :

*Comment peut-on traiter les signaux lorsque les exigences dépassent les moyens disponibles ?*

Nous espérons avoir persuadé le lecteur que dans la mesure où les limitations en termes de capacité de traitement doivent être prises en compte, le traitement du signal compressé et la quantification sur 1 bit sont deux réponses tout à fait pertinentes dont l'étude mérite d'être poursuivie.

## Bibliographie

- [1] BENDAT, J. S., AND PIERSOL, A. G. *Random data : analysis and measurement procedures*, vol. 729. John Wiley & Sons, 2011.
- [2] COPPA, B. *Sur quelques applications du codage parcimonieux et sa mise en oeuvre*. PhD thesis, Université Grenoble Alpes, 2013.
- [3] DAVENPORT, M. A., BOUFONOS, P. T., WAKIN, M. B., AND BARANIUK, R. G. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing* 4, 2 (2010), 445–460.
- [4] DUBNOV, S. Generalization of spectral flatness measure for non-gaussian linear processes. *IEEE Signal Processing Letters* 11, 8 (2004), 698–701.
- [5] FARIAS, R. C. *Estimation sous contraintes de communication : algorithmes et performances asymptotiques*. PhD thesis, Université Grenoble Alpes, 2013.
- [6] JENKINS, G. M., AND WATTS, D. G. *Spectral analysis*. Holden-day, 1968.
- [7] MARNAT, M., PELISSIER, M., MICHEL, O., AND ROS, L. Code properties analysis for the implementation of a modulated wideband converter.
- [8] VINCENT, R. *Identification passive en acoustique : estimateurs et applications au SHM*. PhD thesis, Université Grenoble Alpes, 2016.
- [9] WU, W. B. Asymptotic theory for stationary processes. *Stat. Interface* 4, 2 (2011), 207–226.



# Annexes

## Statistiques du corrélateur compressé

Rappelons la définition de cet estimateur :

$$C_N(\tau) = (\Phi \mathbf{x})^\top (\Phi \mathbf{y}_\tau)$$

$$\text{où } \mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$$

### Biais

Puisque les signaux et la matrice de compression sont aléatoires, d'après le théorème de l'espérance totale on a :

$$\mathbf{E}[C_N(\tau)] = \mathbf{E}_{\mathbf{x}\mathbf{y}_\tau}[\mathbf{E}_\Phi[C_N(\tau)|\mathbf{x}\mathbf{y}_\tau]]$$

Pour calculer l'espérance conditionnelle on écrit :

$$C_N(\tau) = \sum_{i=1}^M \left( \sum_{\alpha} \varphi_{i\alpha} x_{\alpha} \right) \left( \sum_{\beta} \varphi_{i\beta} y_{\beta+\tau} \right)$$

$$C_N(\tau) = \sum_{i=1}^M Q_i^\top \quad \text{où } Q_i^\top = \sum_{\alpha, \beta} x_{\alpha} y_{\beta+\tau} \varphi_{i\alpha} \varphi_{i\beta}$$

Puis :

$$\begin{aligned} \mathbf{E}[Q_i|\mathbf{x}\mathbf{y}_\tau] &= \mathbf{E} \left[ \sum_{\alpha, \beta} x_{\alpha} y_{\beta+\tau} \varphi_{i\alpha} \varphi_{i\beta} \middle| \mathbf{x}\mathbf{y}_\tau \right] \\ &= \mathbf{E}[\varphi^2] \mathbf{x}^T \mathbf{y}_\tau \end{aligned}$$

puisque les éléments  $\varphi$  de la matrice sont i.i.d. et de moyenne nulle. Alors

$$\mathbf{E}[C_N(\tau)|\mathbf{x}\mathbf{y}_\tau] = M \mathbf{E}[\varphi^2] \mathbf{x}^T \mathbf{y}_\tau$$

et finalement :

$$\begin{aligned} \mathbf{E}[C_N(\tau)] &= \mathbf{E}_{\mathbf{x}\mathbf{y}_\tau}[\mathbf{E}[C_N(\tau)|\mathbf{x}\mathbf{y}_\tau]] \\ &= M \mathbf{E}[\varphi^2] \mathbf{E}[\mathbf{x}^T \mathbf{y}_\tau] \\ &= M \mathbf{E}[\varphi^2] N \Gamma_{xy}(\tau) \end{aligned}$$

Pour que l'estimateur soit non biaisé, il suffit donc de fixer

$$\mathbf{E}[\varphi^2] = \frac{1}{MN}$$

## Covariance

Les signaux et la matrice étant aléatoires, ce résultat est obtenu à partir de la formule de décomposition de la covariance :

$$\mathbf{Cov}\{C_N(k), C_N(\ell)\} = \mathbf{E}_{x\mathbf{y}_k\mathbf{y}_\ell}[\mathbf{Cov}_{\Phi}\{C_N(k), C_N(\ell)|x\mathbf{y}_k\mathbf{y}_\ell\}] + \mathbf{Cov}_{x\mathbf{y}_k\mathbf{y}_\ell}\{\mathbf{E}_{\Phi}[C_N(k)|x\mathbf{y}_k], \mathbf{E}_{\Phi}[C_N(\ell)|x\mathbf{y}_\ell]\}$$

$$\text{Où } \mathbf{x} = [x_1, \dots, x_N]^\top, \mathbf{y}_k = [y_{1+k}, \dots, y_{N+k}]^\top, \mathbf{y}_\ell = [y_{1+\ell}, \dots, y_{N+\ell}]^\top.$$

Premier terme,  $\mathbf{E}_{x\mathbf{y}_k\mathbf{y}_\ell}[\mathbf{Cov}_{\Phi}\{C_N(k), C_N(\ell)|x\mathbf{y}_k\mathbf{y}_\ell\}]$ .

On écrit à nouveau :

$$C_N(\tau) = \sum_{i=1}^M Q_i^\tau \quad \text{avec} \quad Q_i^\tau = \sum_{\alpha, \beta} x_\alpha y_{\beta+\tau} \varphi_{i\alpha} \varphi_{i\beta}$$

La covariance conditionnelle entre  $Q_i^k$  et  $Q_i^\ell$  s'écrit :

$$\mathbf{Cov}_{\Phi}[Q_i^k, Q_i^\ell | x\mathbf{y}_k\mathbf{y}_\ell] = \sum_{\alpha, \beta, \delta, \gamma} x_\alpha y_{\beta+k} x_\delta y_{\gamma+\ell} \mathbf{Cov}[\varphi_{i\alpha} \varphi_{i\beta}, \varphi_{i\gamma} \varphi_{i\delta}]$$

Puisque les  $\varphi_{iz}$  sont centrés et deux à deux indépendants

$$\begin{aligned} \mathbf{Cov}[\varphi_{i\alpha} \varphi_{i\beta}, \varphi_{i\gamma} \varphi_{i\delta}] &= \mathbf{E}[\varphi_{i\alpha} \varphi_{i\beta} \varphi_{i\gamma} \varphi_{i\delta}] - \mathbf{E}[\varphi_{i\alpha} \varphi_{i\beta}] \mathbf{E}[\varphi_{i\gamma} \varphi_{i\delta}] \\ &= \begin{cases} 0 & \alpha \neq \beta \neq \gamma \neq \delta \\ \mathbf{E}^2[\varphi^2] & \alpha = \gamma \neq \beta = \delta \text{ ou } \alpha = \delta \neq \beta = \gamma \\ \mathbf{E}^2[\varphi^2] - \mathbf{E}^2[\varphi^2] = 0 & \alpha = \beta \neq \gamma = \delta \\ 0 & \text{trois indices égaux, mais pas quatre} \\ \mathbf{E}[\varphi^4] - \mathbf{E}^2[\varphi^2] & \alpha = \beta = \gamma = \delta \end{cases} \end{aligned}$$

On obtient donc :

$$\mathbf{Cov}_{\Phi}[Q_i^k, Q_i^\ell | x\mathbf{y}_k\mathbf{y}_\ell] = (\mathbf{E}[\varphi^4] - \mathbf{E}^2[\varphi^2]) \sum_{\alpha} x_\alpha^2 y_{\alpha+k} y_{\alpha+\ell} + \mathbf{E}^2[\varphi^2] \sum_{\alpha \neq \beta} (x_\alpha^2 y_{\beta+k} y_{\beta+\ell} + x_\alpha y_{\alpha+k} x_\beta y_{\beta+\ell})$$

En écrivant l'expression précédente comme suit,

$$\begin{aligned} &= (\mathbf{E}[\varphi^4] - 3\mathbf{E}^2[\varphi^2]) \sum_{\alpha} x_\alpha^2 y_{\alpha+k} y_{\alpha+\ell} + 2\mathbf{E}^2[\varphi^2] \sum_{\alpha} x_\alpha^2 y_{\alpha+k} y_{\alpha+\ell} \\ &\quad + \mathbf{E}^2[\varphi^2] \sum_{\alpha \neq \beta} (x_\alpha^2 y_{\beta+k} y_{\beta+\ell} + x_\alpha y_{\alpha+k} x_\beta y_{\beta+\ell}) \end{aligned}$$

on fait apparaître le kurtosis normalisé,  $\kappa[\varphi] = \frac{\mathbf{E}[\varphi^4] - 3\mathbf{E}^2[\varphi^2]}{\mathbf{E}^2[\varphi^2]}$ , on a donc :

$$= \kappa[\varphi] \mathbf{E}^2[\varphi^2] \sum_{\alpha} x_\alpha^2 y_{\alpha+k} y_{\alpha+\ell} + \mathbf{E}^2[\varphi^2] \sum_{\alpha, \beta} (x_\alpha^2 y_{\beta+k} y_{\beta+\ell} + x_\alpha y_{\alpha+k} x_\beta y_{\beta+\ell})$$

Ou bien,

$$= \mathbf{E}^2[\varphi^2] \left( \kappa[\varphi] \sum_{\alpha} x_\alpha^2 y_{\alpha+k} y_{\alpha+\ell} + \|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_\ell) + (\mathbf{x}^T \mathbf{y}_k) (\mathbf{x}^T \mathbf{y}_\ell) \right)$$

D'autre part,

$$\begin{aligned} \mathbf{Cov}_{\Phi}\{C_N(k), C_N(\ell)|\mathbf{x}\mathbf{y}_k\mathbf{y}_\ell\} &= \mathbf{Cov}_{\Phi}\left\{\sum_{i=1}^M Q_i^k, \sum_{j=1}^M Q_j^\ell \middle| \mathbf{x}\mathbf{y}_k\mathbf{y}_\ell\right\} \\ &= \sum_{i=1}^M \sum_{j=1}^M \mathbf{Cov}_{\Phi}\left\{Q_i^k, Q_j^\ell \middle| \mathbf{x}\mathbf{y}_k\mathbf{y}_\ell\right\} \end{aligned}$$

Les éléments de  $\Phi$  étant deux à deux indépendants,  $Q_i^k$  et  $Q_j^\ell$  le sont aussi pour  $j \neq i$  conditionnellement à  $\mathbf{x}, \mathbf{y}_k$ , et  $\mathbf{y}_\ell$ . Les termes qui restent correspondent au seul cas  $j = i$

$$\begin{aligned} &= \sum_{i=1}^M \mathbf{Cov}_{\Phi}[Q_i^k, Q_i^\ell | \mathbf{x}\mathbf{y}_k\mathbf{y}_\ell] \\ &= M\mathbf{E}^2[\varphi^2] \left( \kappa[\varphi] \sum_{\alpha} x_{\alpha}^2 y_{\alpha+k} y_{\alpha+\ell} + \|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_\ell) + (\mathbf{x}^T \mathbf{y}_k)(\mathbf{x}^T \mathbf{y}_\ell) \right) \end{aligned}$$

soit, en introduisant la condition de non biais  $\mathbf{E}[\varphi^2] = \frac{1}{MN}$  :

$$= \frac{1}{MN^2} \left( \kappa[\varphi] \sum_{\alpha} x_{\alpha}^2 y_{\alpha+k} y_{\alpha+\ell} + \|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_\ell) + (\mathbf{x}^T \mathbf{y}_k)(\mathbf{x}^T \mathbf{y}_\ell) \right)$$

Finalement,

$$\begin{aligned} \mathbf{E}_{\mathbf{x}\mathbf{y}_k\mathbf{y}_\ell}[\mathbf{Cov}_{\Phi}\{C_N(k), C_N(\ell)|\mathbf{x}\mathbf{y}_k\mathbf{y}_\ell\}] &= \\ \frac{1}{MN^2} \left( \kappa[\varphi] \sum_{\alpha} \mathbf{E}[x_{\alpha}^2 y_{\alpha+k} y_{\alpha+\ell}] + \mathbf{E}[\|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_\ell)] + \mathbf{E}[(\mathbf{x}^T \mathbf{y}_k)(\mathbf{x}^T \mathbf{y}_\ell)] \right) \end{aligned}$$

Dans cette expression :

$$\begin{aligned} &\sum_{\alpha} \mathbf{E}[x_{\alpha}^2 y_{\alpha+k} y_{\alpha+\ell}] \\ &= N(C_{xyxy}(k, 0, \ell) + 2\Gamma_{xy}(k)\Gamma_{xy}(\ell) + \Gamma_{xx}(0)\Gamma_{yy}(k - \ell)) \end{aligned}$$

Ensuite :

$$\begin{aligned} &\mathbf{E}[\|\mathbf{x}\|_2^2 (\mathbf{y}_k^T \mathbf{y}_\ell)] = \sum_{\alpha, \beta} \mathbf{E}[x_{\alpha}^2 y_{\beta+k} y_{\beta+\ell}] \\ &= \sum_{\alpha, \beta} (C_{xyxy}(\alpha - \beta + k, 0, \alpha - \beta + \ell) + 2\Gamma_{xy}(\alpha - \beta + k)\Gamma_{xy}(\alpha - \beta + \ell) + \Gamma_{xx}(0)\Gamma_{yy}(k - \ell)) \\ &= \sum_{r=-N}^N (N - |r|)(C_{xyxy}(r + k, 0, r + \ell) + 2\Gamma_{xy}(r + k)\Gamma_{xy}(r + \ell)) + N^2\Gamma_{xx}(0)\Gamma_{yy}(k - \ell) \end{aligned}$$

Et :

$$\begin{aligned} \mathbf{E}[(\mathbf{x}^T \mathbf{y}_k)(\mathbf{x}^T \mathbf{y}_\ell)] &= \mathbf{Cov}\{\mathbf{x}^T \mathbf{y}_k, \mathbf{x}^T \mathbf{y}_\ell\} + \mathbf{E}[(\mathbf{x}^T \mathbf{y}_k)] \mathbf{E}[(\mathbf{x}^T \mathbf{y}_\ell)] \\ &= N^2 \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} + N^2 \Gamma_{xy}(k)\Gamma_{xy}(\ell) \end{aligned}$$

Deuxième terme,  $\mathbf{Cov}_{\mathbf{x}\mathbf{y}_k\mathbf{y}_\ell}\{\mathbf{E}_{\Phi}[C_N(k)|\mathbf{x}\mathbf{y}_k], \mathbf{E}_{\Phi}[C_N(\ell)|\mathbf{x}\mathbf{y}_\ell]\}$ .

Lors du calcul du biais, on a déjà établi

$$\mathbf{E}_{\Phi}[C_N(k)|\mathbf{x}\mathbf{y}_\tau] = M\mathbf{E}[\varphi^2] \mathbf{x}^T \mathbf{y}_k = \frac{1}{N} \mathbf{x}^T \mathbf{y}_k$$

Finalement,  $\mathbf{Cov}_{xy_k y_\ell} \{ \mathbf{E}_\Phi [C_N(k) | \mathbf{x} \mathbf{y}_k], \mathbf{E}_\Phi [C_N(\ell) | \mathbf{x} \mathbf{y}_\ell] \}$  s'écrit :

$$\mathbf{Cov}_{xy_k y_\ell} \left\{ \frac{1}{N} \mathbf{x}^T \mathbf{y}_k, \frac{1}{N} \mathbf{x}^T \mathbf{y}_\ell \right\} = \mathbf{Cov} \{ \mathcal{C}_N(k), \mathcal{C}_N(\ell) \}$$

On retrouve tout simplement la covariance de l'estimateur classique, lorsqu'il manipule  $N$  points.

## Échantillonnage aléatoire

Soient  $\mathbf{x} = [x_1, \dots, x_N]^\top$  et  $\mathbf{y}_\tau = [y_{1+\tau}, \dots, y_{N+\tau}]^\top$  les données mesurées par chaque capteur. On écrit  $\mathbf{s} = \mathbf{x} \circ \mathbf{y}_\tau$  le produit, terme à terme, des deux vecteurs. On prélève uniformément sans remise  $M$  éléments parmi  $\{1, \dots, N\}$ . Soit  $E_M$  le sous-ensemble obtenu. Alors,  $i \in E_M$  avec probabilité  $\pi_1 = \frac{M}{N}$ . Pour  $i \neq j, (i, j) \in E_M$  avec probabilité  $\pi_2 = \frac{M(M-1)}{N(N-1)}$ . Maintenant, nous pouvons écrire l'estimateur par échantillonnage aléatoire, sans remise, de la fonction d'intercorrélacion comme

$$\mathring{C}_N(\tau) = \frac{1}{M} \sum_{i \in E_M} s_i = \frac{1}{M} \sum_{i=1}^N s_i \epsilon_i$$

Ici  $\epsilon_i, i = \{1, \dots, N\}$  est une suite de variables de Bernoulli de paramètre  $\pi_1$ . Ces variables sont corrélées :  $\mathbf{E}[\epsilon_i \epsilon_j] = \pi_2$ , mais on les suppose indépendantes des composantes de  $\mathbf{s}$ .

## Biais

L'estimateur compressé par échantillonnage aléatoire est non biaisé. En effet

$$\begin{aligned} \mathbf{E}\{\mathring{C}_N\} &= M^{-1} \sum_{i=1}^N \mathbf{E}_s [s_i \mathbf{E}_\epsilon [\epsilon_i | s_i]] = M^{-1} \sum_{i=1}^N \mathbf{E}_s [s_i \pi_1] \\ &= M^{-1} \frac{M}{N} \sum_{i=1}^N \mathbf{E}[s_i] = \frac{1}{N} \mathbf{E}[x_i y_{i+\tau}] = \Gamma_{xy}(\tau) \end{aligned}$$

## Covariance

Soient les vecteurs  $\mathbf{x} = [x_1, \dots, x_N]^\top$ ,  $\mathbf{y}_k = [y_{1+k}, \dots, y_{N+k}]^\top$ ,  $\mathbf{z} = \mathbf{x} \circ \mathbf{y}_k$  et  $\mathbf{w} = \mathbf{x} \circ \mathbf{y}_\ell$

La covariance est dérivée à partir de la formule de décomposition, c'est-à-dire :

$$\mathbf{Cov} \{ \mathring{C}_N(k), \mathring{C}_N(\ell) \} = \mathbf{E} \mathbf{x} \mathbf{y} [\mathbf{Cov}_\epsilon \{ \mathring{C}_N(k), \mathring{C}_N(\ell) | \mathbf{x} \mathbf{y} \}] + \mathbf{Cov} \mathbf{x} \mathbf{y} \{ \mathbf{E}_\epsilon [\mathring{C}_N(k) | \mathbf{x} \mathbf{y}], \mathbf{E}_\epsilon [\mathring{C}_N(\ell) | \mathbf{x} \mathbf{y}] \}$$

Calculons d'abord la covariance conditionnelle :  $\mathbf{Cov}_\epsilon \{ \mathring{C}_N(k), \mathring{C}_N(\ell) | \mathbf{x} \mathbf{y} \}$

$$\begin{aligned} &= \mathbf{Cov}_\epsilon \left\{ M^{-1} \sum_i z_i \epsilon_i, M^{-1} \sum_j w_j \epsilon_j \middle| \mathbf{x} \mathbf{y} \right\} \\ &= M^{-2} \sum_{j,i} z_i w_j \mathbf{Cov} \{ \epsilon_i, \epsilon_j \} \end{aligned}$$

Le premier terme,  $\mathbf{E} \mathbf{x} \mathbf{y} [\mathbf{Cov}_\epsilon \{ \mathring{C}_N(k), \mathring{C}_N(\ell) | \mathbf{x} \mathbf{y} \}]$ , vaut alors :

$$= M^{-2} \sum_{j,i} \mathbf{E} \mathbf{x} \mathbf{y} [z_i w_j \mathbf{Cov} \{ \epsilon_i, \epsilon_j \}]$$

$$= \frac{1}{M^2} \sum_i \pi_1(1 - \pi_1) \mathbf{E}[z_i w_i] + \frac{1}{M^2} \sum_{i \neq j} (\pi_2 - \pi_1^2) \mathbf{E}[z_i w_j]$$

En remplaçant  $\pi_1$  par  $M/N = 1/\alpha$  et  $\pi_2$  par  $\frac{M(M-1)}{N(N-1)} = \frac{N-\alpha}{(N-1)\alpha^2}$ , et en utilisant le taux de compression,  $\alpha = N/M$ , on obtient :

$$\begin{aligned} &= \frac{\alpha - 1}{N^2} \sum_i \mathbf{E}[z_i w_i] + \frac{1}{N^2} \left( \frac{N - \alpha}{N - 1} - 1 \right) \sum_{i \neq j} \mathbf{E}[z_i w_j] \\ &= \frac{\alpha - 1}{N^2} \left( \sum_i \mathbf{E}[z_i w_i] - \frac{1}{N - 1} \sum_{i \neq j} \mathbf{E}[z_i w_j] \right) \\ &= \frac{\alpha - 1}{N^2} \left( \frac{N}{N - 1} \sum_i \mathbf{E}[z_i w_i] - \frac{1}{N - 1} \sum_{ij} \mathbf{E}[z_i w_j] \right) \\ &= \frac{\alpha - 1}{N - 1} \left( \frac{1}{N} \sum_i \mathbf{E}[z_i w_i] - (\mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} + \mathbf{E}[z_i] \mathbf{E}[w_j]) \right) \\ &= \frac{\alpha - 1}{N - 1} (C_{xyxy}(k, 0, \ell) + \Gamma_{xx}(0) \Gamma_{yy}(k - \ell) + 2\Gamma_{xy}(k) \Gamma_{xy}(\ell) - \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} - \Gamma_{xy}(k) \Gamma_{xy}(\ell)) \\ &= \frac{\alpha - 1}{N - 1} (C_{xyxy}(k, 0, \ell) + \Gamma_{xx}(0) \Gamma_{yy}(k - \ell) + \Gamma_{xy}(k) \Gamma_{xy}(\ell) - \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\}) \end{aligned}$$

Ensuite, pour *le second terme*, on établit facilement que

$$\mathbf{Cov}_{\mathbf{xy}}\{\mathbf{E}_\epsilon[\hat{\mathcal{C}}_N(k)|\mathbf{xy}], \mathbf{E}_\epsilon[\hat{\mathcal{C}}_N(\ell)|\mathbf{xy}]\} = \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\}$$

puisque l'espérance conditionnelle vaut

$$\mathbf{E}_\epsilon[\hat{\mathcal{C}}_N(k)|\mathbf{xz}] = \mathbf{E}[M^{-1} \sum_{i=1}^N z_i \epsilon_i | \mathbf{xz}] = N^{-1} \sum_{i=1}^N z_i = N^{-1} \sum_{i=1}^N x_i y_{i+k}$$

et de la même manière

$$\mathbf{E}_\epsilon[\hat{\mathcal{C}}_N(\ell)|\mathbf{w}] = N^{-1} \sum_{i=1}^N x_i y_{i+\ell}$$

## Covariance asymptotique du corrélogramme classique

Rappelons d'abord sa définition :

$$\mathcal{S}_N(\nu) = \sum_{k=-L}^L \mathcal{C}_N(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

$$\text{avec } \mathcal{C}_N(\tau) = \frac{1}{N} \sum_{i=1}^N x_i y_{i+\tau}$$

$N$  est le nombre d'échantillons disponibles

$L$  est le retard maximum estimé

$w(x)$  est une fenêtre de pondération, symétrique, et normalisée ( $w(0)=1$ )

Des résultats existent dans la littérature [1, 4], mais ceux-ci considèrent l'estimateur biaisé standard de corrélation :

$$c_{xy}(\tau) = \frac{1}{N} \sum_{i=1}^{N-|\tau|} x_i y_{i+\tau}$$

Nous adaptions ici la démonstration d'Anderson [1, p. 527] au cas du corrélateur classique.

On s'intéresse au comportement asymptotique de la covariance, c'est à dire  $N \rightarrow \infty$ , et  $L \rightarrow \infty$ . Ces deux quantités ne sont pas liées entre elles. Cependant, la covariance évolue en  $\mathcal{O}(\frac{L}{N})$ , pour que l'estimateur soit consistant, il faut  $\frac{L}{N} \rightarrow 0$ . d'où l'écriture  $L_N$ .

$$\lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\}$$

D'abord,  $\frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\}$  s'écrit :

$$= \sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \mathbf{Cov}\{\mathcal{C}_M(k), \mathcal{C}_M(\ell)\}$$

Où  $\mathbf{Cov}\{\mathcal{C}_M(k), \mathcal{C}_M(\ell)\}^2$  est constitué de 3 termes.

$$\frac{1}{N} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(k, r, \ell + r) + \Gamma_{xx}(r)\Gamma_{yy}(r + \ell - k) + \Gamma_{xy}(k - r)\Gamma_{xy}(\ell + r)\}$$

$$\frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} = \Sigma_1 + \Sigma_2 + \Sigma_3$$

On commence par le terme  $\Sigma_3$ .

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \left( \frac{1}{N} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{ \Gamma_{xy}(r - k)\Gamma_{xy}(r + \ell) \} \right)$$

Avec le changement de variable  $u = r - k$  et  $v = r + \ell$ , on peut écrire :

$$\frac{1}{L_N} \sum_{u, v = -(L_N + N)}^{L_N + N} \sum_{r = \max[u - L_N, v - L_N, -N]}^{\min[u + L_N, v + L_N, N]} \left(1 - \frac{|r|}{N}\right) \left[ w\left(\frac{r - u}{L_N}\right) w\left(\frac{v - r}{L_N}\right) \right] \times e^{-2\pi j[(r - u)\nu_1 - (v - r)\nu_2]} \Gamma_{xy}(u)\Gamma_{xy}(v) \quad (6.1)$$

Puisque  $\Gamma_{xy}(\tau) \rightarrow 0$  rapidement lorsque  $\tau \rightarrow \pm\infty$ , l'expression (6.1) est bien approchée par une série tronquée constituée de uniquement  $2m + 1$  termes ( $m \leq L_N$ ), c'est-à-dire :

$$\frac{1}{L_N} \sum_{u, v = -m}^m \sum_{r = \max[u - L_N, v - L_N, -N]}^{\min[u + L_N, v + L_N, N]} \left(1 - \frac{|r|}{N}\right) \left[ w\left(\frac{r - u}{L_N}\right) w\left(\frac{v - r}{L_N}\right) \right] \times e^{-2\pi j[(r - u)\nu_1 - (v - r)\nu_2]} \Gamma_{xy}(v)\Gamma_{xy}(u) \quad (6.2)$$

Suivant Anderson, si  $w(x)$  est continue,  $-1 \leq x \leq 1$ , et symétrique alors pour  $|u| < m$ ,  $|v| < m$ ,  $|r| < m + L_N$  et  $L_N$  suffisamment grand<sup>3</sup> :

$$\left| w\left(\frac{r - u}{L_N}\right) w\left(\frac{v - r}{L_N}\right) - w^2\left(\frac{r}{L_N}\right) \right| < \epsilon$$

2. Dans le calcul d'Anderson qui considère  $\mathbf{Cov}\{c_{xy}(k), c_{xy}(\ell)\}$ , au lieu du terme  $\left(1 - \frac{|r|}{N}\right)$  on retrouve une fonction  $\phi(r; k, \ell)$  qui compte le nombre d'éléments non-nuls dans la somme, celle-ci dépend aussi de  $k$  et  $\ell$ .

3. Cela permet d'approcher  $w\left(\frac{r - u}{L_N}\right) w\left(\frac{v - r}{L_N}\right)$  par  $w^2\left(\frac{r}{L_N}\right)$

Auquel cas, on peut approcher (6.2) par :

$$\frac{1}{L_N} \sum_{u,v=-m}^m \sum_{r=\max[u-L_N, v-L_N, -N]}^{\min[u+L_N, v+L_N, N]} \left(1 - \frac{|r|}{N}\right) w^2\left(\frac{r}{L_N}\right) \times e^{-2\pi j[(r-u)\nu_1 - (v-r)\nu_2]} \Gamma_{xy}(v) \Gamma_{xy}(u) \quad (6.3)$$

Ensuite, si  $N$  est suffisamment large  $\left(1 - \frac{|r|}{N}\right) \rightarrow 1$ , et donc (6.3) est bien approchée par :

$$\frac{1}{L_N} \sum_{u,v=-m}^m \sum_{r=-L_N}^{L_N} w^2\left(\frac{r}{L_N}\right) e^{-2\pi j[(r-u)\nu_1 - (v-r)\nu_2]} \Gamma_{xy}(v) \Gamma_{xy}(u)$$

ou bien :

$$\left[ \sum_{r=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{r}{L_N}\right) e^{-2\pi j r[\nu_1 + \nu_2]} \right] \left[ \sum_{u=-m}^m e^{2\pi j u \nu_1} \Gamma_{xy}(u) \right] \left[ \sum_{v=-m}^m e^{2\pi j v \nu_2} \Gamma_{xy}(v) \right]$$

Il existe deux cas, d'abord si  $\nu_1 + \nu_2 \neq 0$  ou  $\pm 1$ , d'après Anderson :

$$\lim_{N \rightarrow +\infty} \sum_{r=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{r}{L_N}\right) e^{-2\pi j r[\nu_1 + \nu_2]} = 0$$

Auquel cas la contribution de ce terme est nulle.

Ensuite si  $\nu_1 + \nu_2 = 0$ , ou  $\pm 1$ . Alors  $\nu_1 = -\nu_2 = \nu$ , et donc :

$$\lim_{N \rightarrow +\infty} \sum_{r=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{r}{L_N}\right) = \int_{-1}^1 w^2(x) dx$$

Pour  $m$  suffisamment large :

$$\begin{aligned} \sum_{u=-m}^m e^{2\pi j u \nu} \Gamma_{xy}(u) &= \sum_{p=-m}^m e^{2\pi j p \nu} \Gamma_{xy}(-p) \rightarrow \overline{\gamma_{xy}}(\nu) \\ \sum_{v=-m}^m e^{-2\pi j v \nu} \Gamma_{xy}(v) &\rightarrow \gamma_{xy}(\nu) \end{aligned}$$

Enfin, on aboutit à :

$$\lim_{N \rightarrow +\infty} \Sigma_3 = \begin{cases} \left[ \int_{-1}^1 w^2(x) dx \right] |\gamma_{xy}(\nu)|^2 & \nu_1 = -\nu_2 = 0, \pm 1 \\ 0 & \text{sinon} \end{cases}$$

Ensuite, le terme  $\Sigma_2$ , s'écrit :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \left( \frac{1}{N} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \Gamma_{xx}(r) \Gamma_{yy}(r + \ell - k) \right)$$

Avec le changement de variable  $r = u, v = r + \ell - k$  et  $s = u - k$  on peut écrire :

$$\frac{1}{L_N} \sum_{u=-(N-1)}^{N-1} \sum_{v=u-2L_N}^{u+2L_N} \sum_{s=\max(u,v)+L_N}^{\min(u,v)+L_N} \left(1 - \frac{|u|}{N}\right) \left[ w\left(\frac{u-s}{L_N}\right) w\left(\frac{v-s}{L_N}\right) \right] \times e^{-2\pi j[(u-s)\nu_1 - (v-s)\nu_2]} \Gamma_{xx}(u) \Gamma_{yy}(v) \quad (6.4)$$

Comme  $\Gamma_{xx}(\tau) \rightarrow 0$  et  $\Gamma_{yy}(\tau) \rightarrow 0$  lorsque  $\tau \rightarrow \pm\infty$ , l'expression (6.4) est bien approchée par uniquement  $2m + 1$  termes. où  $m \leq L_N$  :

$$\frac{1}{L_N} \sum_{u,v=-m}^m \sum_{s=\max(u,v)+L_N}^{\min(u,v)+L_N} \left(1 - \frac{|u|}{N}\right) \left[ w\left(\frac{u-s}{L_N}\right) w\left(\frac{v-s}{L_N}\right) \right] \times e^{-2\pi j[(u-s)\nu_1 - (v-s)\nu_2]} \Gamma_{xx}(u) \Gamma_{yy}(v) \quad (6.5)$$

D'après Anderson, pour  $L_N$  large  $\left[ w\left(\frac{u-s}{L_N}\right) w\left(\frac{v-s}{L_N}\right) \right]$  peut être approchée par  $w^2\left(\frac{-s}{L_N}\right) = w^2\left(\frac{s}{L_N}\right)$

A nouveau, pour  $N$  grand,  $\left(1 - \frac{|u|}{N}\right) \rightarrow 1$ , ce qui permet d'approcher (6.5) par :

$$\frac{1}{L_N} \sum_{u,v=-m}^m \sum_{s=-L_N}^{L_N} w^2\left(\frac{s}{L_N}\right) e^{-2\pi j[(u-s)\nu_1 - (v-s)\nu_2]} \Gamma_{xx}(u) \Gamma_{yy}(v)$$

ou bien :

$$\left[ \sum_{s=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{s}{L_N}\right) e^{2\pi j[s(\nu_1 - \nu_2)]} \right] \left[ \sum_{u=-m}^m e^{-2\pi j u \nu_1} \Gamma_{xx}(u) \right] \left[ \sum_{v=-m}^m e^{2\pi j v \nu_2} \Gamma_{yy}(v) \right]$$

Il existent aussi deux cas,  $\nu_1 - \nu_2 \neq 0$  ou  $\pm 1$ , d'après Anderson :

$$\lim_{N \rightarrow +\infty} \sum_{s=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{s}{L_N}\right) e^{2\pi j s [\nu_1 - \nu_2]} = 0$$

dont la contribution est nulle.

Ensuite pour si  $\nu_1 - \nu_2 = 0$ , ou  $\pm 1$ . Alors  $\nu_1 = \nu_2 = \nu$ , et donc :

$$\lim_{N \rightarrow +\infty} \sum_{s=-L_N}^{L_N} \frac{1}{L_N} w^2\left(\frac{s}{L_N}\right) = \int_{-1}^1 w^2(x) dx$$

Pour  $m$  suffisamment large :

$$\sum_{u=-m}^m e^{-2\pi j u \nu} \Gamma_{xx}(u) \rightarrow \gamma_{xx}(\nu)$$

$$\sum_{v=-m}^m e^{2\pi j v \nu} \Gamma_{yy}(v) = \sum_{q=-m}^m e^{-2\pi j q \nu} \Gamma_{yy}(q) \rightarrow \gamma_{yy}(\nu)$$

Enfin, on aboutit à :

$$\lim_{N \rightarrow +\infty} \Sigma_2 = \begin{cases} \left[ \int_{-1}^1 w^2(x) dx \right] |\gamma_{xx}(\nu) \gamma_{yy}(\nu)| & \nu_1 = \nu_2 \\ 0 & \text{sinon} \end{cases}$$

Finalement, le Terme  $\Sigma_1$  :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L}\right) w\left(\frac{\ell}{L}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \left( \frac{1}{N} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(k, r, \ell + r)\} \right)$$

est inférieure en valeur absolue à :

$$\frac{1}{L_N} \sup_{-1 \leq x \leq 1} w^2(x) \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} |C_{xyxy}(r, s, t)| = \mathcal{O}\left(\frac{1}{L_N}\right)$$

Puisque la somme triple est finie, la contribution de ce terme est aussi nulle.

$$\lim_{N \rightarrow +\infty} \Sigma_1 = 0$$

## Covariance asymptotique du corrélogramme compressé

Nous rappelons sa définition :

$$S_N(\nu) = \sum_{k=-L}^L C_N(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

$$\begin{aligned} \text{avec } C_N(\tau) &= \langle \Phi \mathbf{x}, \Phi \mathbf{y}_\tau \rangle = (\Phi \mathbf{x})^\top (\Phi \mathbf{y}_\tau) \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^N, \Phi \in \mathbb{R}^{M \times N} \\ \mathbf{x} &= [x_1, \dots, x_N]^\top \text{ et } \mathbf{y} = [y_{1+\tau}, \dots, y_{N+\tau}]^\top \end{aligned}$$

$N$  est le nombre d'échantillons que l'on compresse

$M$  est le nombre de valeurs numériques après compression.

$L$  est le retard maximum estimé

$w(x)$  est une fenêtre de pondération, symétrique, et normalisée ( $w(0)=1$ )

On s'intéresse au comportement asymptotique :  $N, M, L \rightarrow \infty$ . Il faut à nouveau  $\frac{L}{N} \rightarrow 0$  pour obtenir un estimateur consistant. On considère un taux de compression fixe, c'est-à-dire,  $\frac{N}{M} = \alpha$ .

Pour évaluer :

$$\lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{S_N(\nu_1), S_N(\nu_2)\}$$

On écrit d'abord  $\frac{N}{L_N} \mathbf{Cov}\{S_N(\nu_1), S_N(\nu_2)\}$  :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L}\right) w\left(\frac{\ell}{L}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \mathbf{Cov}\{C_N(k), C_N(\ell)\}$$

Qui dépend des covariances du corrélateur compressé :

$$\begin{aligned} \mathbf{Cov}\{C_N(k), C_N(\ell)\} &= \left(1 + \frac{1}{M}\right) \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\} \\ &+ \frac{\kappa[\varphi]}{MN} (C_{xyxy}(k, 0, \ell) + 2\Gamma_{xy}(k)\Gamma_{xy}(\ell) + \Gamma_{xx}(0)\Gamma_{yy}(k - \ell)) \\ &+ \frac{1}{M} (\Gamma_{xx}(0)\Gamma_{yy}(k - \ell) + \Gamma_{xy}(k)\Gamma_{xy}(\ell)) \\ &+ \frac{1}{M} \left( \frac{1}{N} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(r + k, 0, r + \ell) + 2\Gamma_{xy}(\ell + r)\Gamma_{xy}(r + k)\} \right) \end{aligned}$$

$$= \Sigma'_1 + \Sigma'_2 + \Sigma'_3 + \Sigma'_4 + \Sigma'_5 + \Sigma'_6 + \Sigma'_7 + \Sigma'_8$$

Le premier terme,  $\Sigma'_1$ , est connu,  $(1 + \frac{1}{M})$  fois la covariance du corrélateur classique calculé sur  $N$  points, sa contribution est donc :

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L} \left(1 + \frac{1}{M}\right) \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} &= \lim_{N \rightarrow +\infty} \left(1 + \frac{1}{M}\right) \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \\ &= \lim_{N \rightarrow +\infty} \frac{N}{L} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \end{aligned}$$

Ensuite, les termes  $\Sigma'_4$  et  $\Sigma'_5$  sont similaires, ils s'écrivent :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times C_0 \frac{N}{L_N} \Gamma_{xx}(0) \Gamma_{yy}(\ell - k)$$

$C_0$  vaut  $\frac{\kappa[\varphi]}{NM}$  pour  $\Sigma'_4$  et  $\frac{1}{M}$  pour  $\Sigma'_5$

Avec le changement de variable  $z = \ell - k$ , on peut écrire :

$$C_0 N \sum_{z=-2L_N}^{2L_N} \Gamma_{xx}(0) \Gamma_{yy}(z) e^{-2\pi j z \nu_1} \times \left[ \frac{1}{L_N} \sum_{\ell=-L_N}^{L_N} w\left(\frac{\ell+z}{L_N}\right) w\left(\frac{\ell}{L_N}\right) e^{-2\pi j \ell(\nu_1 - \nu_2)} \right]$$

Notons que  $\Gamma_{yy}(\tau) \rightarrow 0$  rapidement lorsque  $\tau \rightarrow \pm\infty$ , on peut donc approcher cette expression par une série tronquée constituée de  $2m + 1$  termes. où  $m \leq L_N$  :

$$C_0 N \sum_{z=-m}^m \Gamma_{xx}(0) \Gamma_{yy}(z) e^{-2\pi j z \nu_1} \times \left[ \frac{1}{L_N} \sum_{\ell=-L_N}^{L_N} w\left(\frac{\ell+z}{L_N}\right) w\left(\frac{\ell}{L_N}\right) e^{-2\pi j \ell(\nu_1 - \nu_2)} \right] \quad (6.6)$$

Suivant Anderson, on approche  $\left[ w\left(\frac{\ell+z}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right]$  par  $w^2\left(\frac{\ell}{L_N}\right)$  ce qui est valable pour  $L_N$  large, (6.6) est donc approchée par :

$$C_0 N \left[ \sum_{z=-m}^m \Gamma_{xx}(0) \Gamma_{yy}(z) e^{-2\pi j z \nu_1} \right] \left[ \frac{1}{L_N} \sum_{\ell=-L_N}^{L_N} w^2\left(\frac{\ell}{L_N}\right) e^{-2\pi j \ell(\nu_1 - \nu_2)} \right] \quad (6.7)$$

La limite de (6.7) lorsque  $N \rightarrow \infty$ , est le produit des limites des trois facteurs, à condition que ces limites existent. Pour le premier on a :

$$\lim_{N \rightarrow +\infty} C_0 N = \begin{cases} \lim_{N \rightarrow +\infty} \frac{\kappa[\varphi]}{M} = C_{4,\varphi} & \text{pour } \Sigma'_4 \\ \lim_{N \rightarrow +\infty} \frac{N}{M} = \alpha & \text{pour } \Sigma'_5 \end{cases}$$

Le deuxième est ne dépend pas de  $N$ , mais pour  $m$  suffisamment large :

$$\sum_{z=-m}^m e^{-2\pi j z \nu} \Gamma_{xx}(0) \Gamma_{yy}(z) \rightarrow \sigma_x^2 \gamma_{yy}(\nu)$$

Pour le troisième, il faut considérer deux cas. D'abord si  $\nu_1 - \nu_2 \neq 0$  ou  $\pm 1$ , d'après Anderson :

$$\lim_{N \rightarrow +\infty} \sum_{s=-L_N}^{L_N} \frac{1}{L_N} w^2 \left( \frac{\ell}{L_N} \right) e^{2\pi j z [\nu_1 - \nu_2]} = 0$$

La contribution de  $\Sigma'_4$  et  $\Sigma'_5$  est donc nulle dans ce cas.

Par contre, si  $\nu_1 - \nu_2 = 0$ , ou  $\pm 1$ . Alors  $\nu_1 = \nu_2 = \nu$ , et donc :

$$\lim_{N \rightarrow +\infty} \sum_{\ell=-L_N}^{L_N} \frac{1}{L_N} w^2 \left( \frac{\ell}{L_N} \right) = \int_{-1}^1 w^2(x) dx$$

Avec les trois limites on aboutit à :

$$\lim_{N \rightarrow +\infty} \Sigma'_4 = \begin{cases} \left[ \int_{-1}^1 w^2(x) dx \right] C_{4,\varphi} \sigma_x^2 \gamma_{yy}(\nu) & \nu_1 = \nu_2 \\ 0 & \text{sinon} \end{cases}$$

$$\lim_{N \rightarrow +\infty} \Sigma'_5 = \begin{cases} \left[ \int_{-1}^1 w^2(x) dx \right] \alpha \sigma_x^2 \gamma_{yy}(\nu) & \nu_1 = \nu_2 \\ 0 & \text{sinon} \end{cases}$$

Par la suite nous montrons que tous les autres termes ont une contribution asymptotique nulle.

On commence par,  $\Sigma'_3$  et  $\Sigma'_6$ , ces termes s'écrivent :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w \left( \frac{k}{L_N} \right) w \left( \frac{\ell}{L_N} \right) \right] e^{-2\pi j (k\nu_1 - \ell\nu_2)} \times C_0 \frac{N}{L_N} \Gamma_{xy}(k) \Gamma_{xy}(\ell) \quad (6.8)$$

$C_0$  vaut  $\frac{\kappa[\varphi]}{NM}$  pour  $\Sigma'_3$  et  $\frac{1}{M}$  pour  $\Sigma'_6$

Dans les deux cas, la limite existe :

$$\lim_{N \rightarrow +\infty} C_0 N = \begin{cases} \lim_{N \rightarrow +\infty} \frac{\kappa[\varphi]}{M} = C_{4,\varphi} & \text{pour } \Sigma'_3 \\ \lim_{N \rightarrow +\infty} \frac{N}{M} = \alpha & \text{pour } \Sigma'_6 \end{cases}$$

$$:= C_0^\infty$$

|(6.8)| est alors dominé par :

$$\frac{1}{L_N} C_0^\infty \sup_{-1 \leq x \leq 1} w^2(x) \sum_{k=-\infty}^{\infty} |\Gamma_{xy}(k)| \sum_{\ell=-\infty}^{\infty} |\Gamma_{xy}(k)|$$

Puisque  $\sum_{\tau=-\infty}^{\infty} |\Gamma_{xy}(\tau)| < \infty$ , la contribution de  $\Sigma'_3$  et  $\Sigma'_6$ , est de l'ordre de  $\mathcal{O}\left(\frac{1}{L_N}\right)$ , ils convergent vers 0 lorsque  $N \rightarrow \infty$ .

Puis, on considère les termes qui dépendent du cumulants d'ordre 4 :  $\Sigma'_2$  et  $\Sigma'_7$

Le terme  $\Sigma'_2$ , est :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w \left( \frac{k}{L_N} \right) w \left( \frac{\ell}{L_N} \right) \right] e^{-2\pi j (k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \left( \frac{\kappa[\varphi]}{MN} \{C_{xyxy}(k, 0, \ell)\} \right)$$

$|\Sigma'_2|$  est donc dominé par :

$$\frac{\kappa[\varphi]}{ML_N} \sup_{-1 \leq x \leq 1} w^2(x) \sum_{k, \ell = -\infty}^{\infty} |C_{xyxy}(k, 0, \ell)| = \mathcal{O}\left(\frac{1}{ML}\right)$$

Qui converge vers 0, lorsque  $N \rightarrow \infty$ .

Car

$$\sum_{a, b, c = -\infty}^{\infty} |C_{xyxy}(a, b, c)| < \infty \implies \sum_{k, \ell = -\infty}^{\infty} |C_{xyxy}(k, 0, \ell)| < \infty$$

Ensuite, le terme  $\Sigma'_7$ ,

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \left( \frac{1}{NM} \sum_{r=-M}^M \left(1 - \frac{|r|}{N}\right) \{C_{xyxy}(k, r, \ell + r)\} \right)$$

est inférieur en valeur absolue à :

$$\frac{1}{ML_N} \sup_{-1 \leq x \leq 1} w^2(x) \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} |C_{xyxy}(r, s, t)| = \mathcal{O}\left(\frac{1}{ML_N}\right)$$

car la somme triple est finie et donc  $\Sigma'_7 \rightarrow 0$  lorsque  $N \rightarrow \infty$ .

Finalement,  $\Sigma'_8$  :

$$\begin{aligned} & \sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{M}{L_N} \left( \frac{1}{MN} \sum_{r=-N}^N \left(1 - \frac{|r|}{N}\right) \Gamma_{xy}(r + \ell) \Gamma_{xy}(r + k) \right) \\ &= \frac{1}{N} \Sigma_3 \text{ (du corrélateur classique), et donc :} \end{aligned}$$

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{N}{L} \Sigma'_8 &= \lim_{N \rightarrow +\infty} \left( \frac{1}{N} \right) \lim_{N \rightarrow \infty} \frac{N}{L} \Sigma_3 \\ \lim_{N \rightarrow \infty} \frac{N}{L} \Sigma_3 &= \begin{cases} \frac{1}{N} \left[ \int_{-1}^1 w^2(x) dx \right] |\gamma_{xy}(\nu)|^2 & \nu_1 = -\nu_2 = 0, \pm 1 \\ 0 & \nu_1 \neq \nu_2 \end{cases} \end{aligned}$$

Elle est de l'ordre de  $\mathcal{O}\left(\frac{1}{N}\right)$  et converge vers 0 lorsque  $N \rightarrow \infty$ .

### Cas échantillonnage aléatoire

Pour calculer :

$$\lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\hat{S}(\nu_1), \hat{S}(\nu_2)\}$$

On écrit d'abord  $\frac{N}{L_N} \mathbf{Cov}\{\hat{S}(\nu_1), \hat{S}(\nu_2)\}$  :

$$\sum_{k=-L_N}^{L_N} \sum_{\ell=-L_N}^{L_N} \left[ w\left(\frac{k}{L_N}\right) w\left(\frac{\ell}{L_N}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{N}{L_N} \mathbf{Cov}\{\hat{C}_N(k), \hat{C}_N(\ell)\}$$

Rappelons que :

$$\mathbf{Cov}\{\hat{C}_N(k), \hat{C}_N(\ell)\} = \left(1 - \frac{\alpha-1}{N-1}\right) \mathbf{Cov}\{\mathcal{C}_N(k), \mathcal{C}_N(\ell)\}$$

$$+ \frac{\alpha-1}{N-1} (C_{xyxy}(k, 0, \ell) + \Gamma_{xx}(0)\Gamma_{yy}(k-\ell) + \Gamma_{xy}(\ell)\Gamma_{xy}(k))$$

$$= \Sigma_1'' + \Sigma_2'' + \Sigma_3'' + \Sigma_4''$$

A nouveau,  $\Sigma_1''$ , est connu :  $\left(1 - \frac{\alpha-1}{N-1}\right)$  fois la covariance du corrélateur classique calculé sur  $N$  points, sa contribution est donc :

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L_N} \left(1 - \frac{\alpha-1}{N-1}\right) \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} &= \lim_{N \rightarrow +\infty} \left(1 - \frac{\alpha-1}{N-1}\right) \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \\ &= \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N(\nu_1), \mathcal{S}_N(\nu_2)\} \end{aligned}$$

Le terme  $\Sigma_3''$ , est analogue à  $\Sigma_4'$  et  $\Sigma_5'$ , suivant la même procédure on arrive à :

$$\lim_{N \rightarrow +\infty} \Sigma_3'' = \begin{cases} \left[ \int_{-1}^1 w^2(x) dx \right] (\alpha-1) \sigma_x^2 \gamma_{yy}(\nu) & \nu_1 = \nu_2 \\ 0 & \text{sinon} \end{cases}$$

Le terme  $\Sigma_2''$  est analogue à  $\Sigma_2'$  sa contribution est donc aussi nulle. C'est aussi le cas de  $\Sigma_4''$  qui est équivalent à  $\Sigma_3'$  et  $\Sigma_6'$ .

## Covariance asymptotique des corrélogrammes sur 1 bit

### Corrélogramme classique sur 1 bit :

Ce résultat découle directement du théorème 1 de Brillinger [2], qui établie les propriétés asymptotiques de l'estimateur d'interspectre suivant :

$$\tilde{s}^*(\nu) = \sum_{k=-L}^L \tilde{c}^*(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k}$$

$$\text{Ici } \tilde{s}^*(\tau) = \sin\left(\frac{\pi}{2} \frac{1}{N} \sum_{i=1}^{N-|\tau|} \text{sign}(x_i) \text{sign}(y_{i+\tau})\right)$$

Ce résultat, s'exprime en fonction de l'estimateur :

$$s^*(\nu) = \sum_{k=-L}^L c^*(k) w\left(\frac{k}{L}\right) e^{-j2\pi\nu k} \quad (6.9)$$

$$\text{où } c^*(\tau) = \frac{1}{N} \sum_{i=1}^{N-|\tau|} \text{sign}(x_i) \text{sign}(y_{i+\tau})$$

Notons que  $\tilde{c}^*(\tau)$  et  $c^*(\tau)$  sont respectivement des estimateurs de corrélation avec et sans la correction du biais. (6.9) est le corrélogramme usuel dont les propriétés sont connues. Ses propriétés asymptotiques s'écrivent en fonction des densités spectrales des signaux quantifiés sur 1 bit, qui seront notés  $\psi_{xx}(\nu)$ ,  $\psi_{yy}(\nu)$ ,  $\psi_{xy}(\nu)$ .

Dans son théorème 1, Brillinger établie que :

$$\mathbf{Cov}\{\tilde{s}^*(\nu_1), \tilde{s}^*(\nu_2)\} = \left[\frac{\pi}{2}\right]^2 \mathbf{Cov}\{s^*(\nu_1), s^*(\nu_2)\} + \mathcal{O}\left(\frac{L}{N}\right)$$

L'estimateur de corrélation considéré dans ce résultat est le biaisé *standard*.

Dans la preuve du théorème 1, l'équation (6.2) est constitué de 5 termes, les deux premiers établissent le résultat, les 3 derniers (3),(4) et (5) convergent vers zero.

Une preuve pour l'estimateur classique s'écrit sous la même forme et diffère uniquement dans le terme (4) qui est absent, le résultat pour l'estimateur classique est alors :

$$\mathbf{Cov} \left\{ \widetilde{\mathcal{I}}_N^*(\nu_1), \widetilde{\mathcal{I}}_N^*(\nu_2) \right\} = \left[ \frac{\pi}{2} \right]^2 \mathbf{Cov} \left\{ \mathcal{I}_N^*(\nu_1), \mathcal{I}_N^*(\nu_2) \right\} + \mathcal{O} \left( \frac{L}{N} \right)$$

Puis, on s'intéresse au comportement asymptotique, c'est à dire  $N \rightarrow \infty$ , et  $L \rightarrow \infty$  et pour que l'estimateur soit consistant  $\frac{L}{N} \rightarrow 0$ .

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov} \left\{ \widetilde{\mathcal{I}}_N^*(\nu_1), \widetilde{\mathcal{I}}_N^*(\nu_2) \right\} &= \left[ \frac{\pi}{2} \right]^2 \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov} \left\{ \mathcal{I}_N^*(\nu_1), \mathcal{I}_N^*(\nu_2) \right\} \\ &= \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ \left[ \frac{\pi}{2} \right]^2 \psi_{xx}(\nu) \psi_{yy}(\nu) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 \neq 0, \frac{1}{2} \\ \left[ \frac{\pi}{2} \right]^2 (|\psi_{xy}(\nu)|^2 + \psi_{xx}(\nu) \psi_{yy}(\nu)) \times \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = -\nu_2 = 0, \frac{1}{2} \end{cases} \end{aligned}$$

Nous rappelons que l'écriture  $\phi(\nu)$  se réfère aux densités des signaux quantifiés sur 1 bit.

### Corrélogramme compressé sur 1 bit :

Nous suivons ici les résultats de Hinich, McNeil et Gingras [5, 6, 3] sur l'estimation de densité spectrale de puissance et de Brillinger [2] sur l'interspectre. Ces travaux exploitent un développement en série de Taylor pour étudier les propriétés des estimateurs.

Nous allons nous contenter d'une approximation l'ordre 1.<sup>4</sup>

$$G(x) \approx G(a) + G'(a)(x - a)$$

L'idée des études évoquées est d'approcher une transformation non-linéaire  $G$  de l'estimateur  $C_N^*(\tau)$  : autour de sa moyenne  $\rho_{xy}^*(\tau) = \frac{2}{\pi} \text{asin}(\rho_{xy}(\tau))$  :

$$G(C_N^*(\tau)) \approx G(\rho_{xy}^*(\tau)) + G'(\rho_{xy}^*(\tau)) \{C_N^*(\tau) - \rho_{xy}^*(\tau)\}$$

Plus précisément,  $G(x) = \sin\left(\frac{\pi}{2}x\right) \implies G'(x) = \frac{\pi}{2} \cos\left(\frac{\pi}{2}x\right)$ .

Cette approximation s'écrit donc :

$$\sin\left(\frac{\pi}{2}C_N^*(\tau)\right) \approx \rho_{xy}(\tau) + \frac{\pi}{2} \cos\left(\frac{\pi}{2}\rho_{xy}^*(\tau)\right) \{C_N^*(\tau) - \rho_{xy}^*(\tau)\}$$

Ou bien

$$\widetilde{C}_N^*(\tau) - \rho_{xy}(\tau) \approx \left(\frac{\pi}{2} \sqrt{1 - \rho_{xy}^2(\tau)}\right) \{C_N^*(\tau) - \rho_{xy}^*(\tau)\}$$

D'autre part, la covariance que l'on cherche à évaluer, celle de  $\widetilde{C}_N^*$  s'écrit :

$$\mathbf{Cov} \left\{ \widetilde{C}_N^*(k), \widetilde{C}_N^*(\ell) \right\} = \mathbf{E} \left[ \left( \widetilde{C}_N^*(k) - \rho_{xy}(k) \right) \left( \widetilde{C}_N^*(\ell) - \rho_{xy}(\ell) \right) \right]$$

4. L'égalité exacte reste à prouver, elle semble possible d'après les résultats cités

En remplaçant les deux facteurs à l'intérieur de l'espérance par leurs approximations on obtient :

$$\begin{aligned} \mathbf{Cov}\{\widetilde{C}_N^*(k), \widetilde{C}_N^*(\ell)\} &\approx \mathbf{E} \left[ \left( \frac{\pi}{2} a(k) \{C_N^*(k) - \rho_{xy}^*(k)\} \right) \left( \frac{\pi}{2} a(\ell) \{C_N^*(\ell) - \rho_{xy}^*(\ell)\} \right) \right] \\ &\quad \text{avec } a(\tau) = \sqrt{1 - \rho_{xy}^2(\tau)} \\ &\approx \left[ \frac{\pi}{2} \right]^2 a(k)a(\ell) \mathbf{E} \left[ (\{C_N^*(k) - \rho_{xy}^*(k)\}) (\{C_N^*(\ell) - \rho_{xy}^*(\ell)\}) \right] \\ &\approx \left[ \frac{\pi}{2} \right]^2 a(k)a(\ell) \mathbf{Cov}\{C_N^*(k), C_N^*(\ell)\} \end{aligned}$$

Enfin, la covariance de l'interspectre s'écrit :

$$\begin{aligned} \frac{M}{L} \mathbf{Cov}\{\widetilde{S}_N^*(\nu_1), \widetilde{S}_N^*(\nu_2)\} &= \\ &\sum_{k=-M}^M \sum_{\ell=-M}^M \left[ w\left(\frac{k}{L}\right) w\left(\frac{\ell}{L}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{M}{L} \mathbf{Cov}\{\widetilde{C}_N^*(k), \widetilde{C}_N^*(\ell)\} \end{aligned}$$

Et avec l'approximation :

$$\approx \sum_{k=-L}^L \sum_{\ell=-L}^L \left[ w\left(\frac{k}{L}\right) w\left(\frac{\ell}{L}\right) \right] e^{-2\pi j(k\nu_1 - \ell\nu_2)} \times \frac{M}{L} \frac{\pi^2}{4} a(k)a(\ell) \mathbf{Cov}\{C_N^*(k), C_N^*(\ell)\}$$

Puis, on s'intéresse au comportement asymptotique :  $N, M, L \rightarrow \infty$ . Il faut à nouveau  $\frac{L}{N} \rightarrow 0$  pour obtenir un estimateur consistant. On considère un taux de compression fixe, c'est-à-dire,  $\frac{N}{M} = \alpha$ . D'après Hinich [5, p. 396], les fonctions  $a(\tau)$  n'ont pas d'influence dans le résultat asymptotique car  $a(\tau) = 1$  pour  $\tau$  grand, auquel cas on a :

$$\lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\widetilde{S}_N^*(\nu_1), \widetilde{S}_N^*(\nu_2)\} \approx \left[ \frac{\pi}{2} \right]^2 \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{S_N^*(\nu_1), S_N^*(\nu_2)\}$$

C'est à dire, le corrélogramme compressé appliqué aux signaux quantifiés, et donc.

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\widetilde{S}_N^*(\nu_1), \widetilde{S}_N^*(\nu_2)\} &\approx \left[ \frac{\pi}{2} \right]^2 \lim_{N \rightarrow +\infty} \frac{N}{L_N} \mathbf{Cov}\{\mathcal{S}_N^*(\nu_1), \mathcal{S}_N^*(\nu_2)\} \\ &+ \begin{cases} 0 & \text{si } \nu_1 \neq \nu_2 \\ \left[ \frac{\pi}{2} \right]^2 (C_{4,\varphi} + \alpha) \psi_{yy}(\nu) \left[ \int_{-1}^1 w^2(x) dx \right] & \text{si } \nu_1 = \nu_2 = \nu \end{cases} \end{aligned}$$

## Bibliographie

- [1] ANDERSON, T. W. *The statistical analysis of time series*, vol. 19. John Wiley & Sons, 2011.
- [2] BRILLINGER, D. R. Estimation of the cross-spectrum of a stationary bivariate gaussian process from its zeros. *Journal of the Royal Statistical Society. Series B (Methodological)* (1968), 145–159.
- [3] GINGRAS, D., AND MASRY, E. Spectral density estimation from nonlinearly observed data. *Journal of Time Series Analysis* 6, 2 (1985), 63–80.
- [4] HANNAN, E. J. *Multiple time series*, vol. 38. John Wiley & Sons, 2009.
- [5] HINICH, M. Estimation of spectra after hard clipping of gaussian processes? *Technometrics* 9, 3 (1967), 391–400.
- [6] MCNEIL, D. Estimating the covariance and spectral density functions from a clipped stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)* (1967), 180–195.

