Radiofrequency receivers based on compressive sampling for feature extraction in cognitive radio applications

Marguerite Marnat

UNIVERSITÉ DE GRENOBLE

THÈSE
Pour obtenir le grade de

DOCTEURE DE LA COMMUNAUTÉ GRENOBLE ALPES

Spécialité: **SIGNAL IMAGE PAROLE TELECOMS**

Arêté ministériel: 25 mai 2016

Présentée par
Marguerite MARNAT

Thèse dirigée par **Olivier MICHEL** (G-INP)
et codirigée par **Laurent ROS** (G-INP)
préparée au sein du **CEA LETI**
Laboratoire Architectures Intégrées Radio-fréquences
dans l'**Ecole Doctorale Electronique, Electrotechnique, Automatique et Traitement du Signal (EEATS)**

Récepteur radiofréquence basé sur l'échantillonnage parcimonieux pour de l'extraction de caractéristiques dans les applications de radio cognitive

Radiofrequency receivers based on compressive sampling for feature extraction in cognitive radio applications

Thèse soutenue publiquement le **29 novembre 2018**
devant le jury composé de:

Madame Inbar FIJALKOW
Professeure, ENSEA ETIS, Rapporteuse
Monsieur Yves LOUËT
Professeur, Centrale SUPELEC, Rapporteur
Monsieur Dominique DALLET
Professeur, IMS, Examinateur (Président)
Monsieur Christoph STUDER
Assistant Professor, Cornell University, , Examinateur
Monsieur Olivier MICHEL
Professeur, Grenoble INP, Directeur de thèse
Monsieur Laurent ROS
Maitre de conférences, Grenoble INP, Co-Directeur de thèse
Monsieur Michaël PELISSIER
Ingénieur de recherche, CEA-LETI/LAIR, Encadrant

Acknowledgements

To begin with, I would like to express my sincere gratitude to my advisor, Michaël Pelissier, for his continuous support during my PhD study and internship. He has been a great mentor for me. I would like to thank him for his guidance, expertise and for enabling me to grow as a research scientist.

Many thanks also to my doctoral directors, Olivier Michel and Laurent Ros, for their helpful assistance. Laurent, for his patience, kindness, and excellent advices on how to write. Olivier, for his keen interest for the topic, extensive knowledge and accurate suggestions.

I thank my fellow labmates in the CEA for all the fun we have had in these last three and a half years and for the stimulating discussions. Among others, David and Raphaël for their technical support and also their cooking skills, Zoltan and Vincent for being two awesome climbing partners, Tu, Abdess and Saad for their good mood everyday, Anthony for chocolate and gold bars supply, Guillaume for still believing in a better world, all interns and in particular Etienne for the many private jokes and Florian for one of my most important discoveries: the secret falafel seller.

I would also like to thank my whole family for their unwavering supportiveness. Without them, I would not have been the person I am today and I would not have been able to achieve this. And of course I would like to thank Christian with all my heart. I dedicate this thesis to him.

Projects, snorted Gauss. Devices, plans, intrigues. Palavers with ten princes and a hundred academies, until a barometer may be put somewhere.
That is not science.
Oh, shouted Humboldt, what is science then?
Gauss sucked the pipe. A man alone at his desk, a sheet of paper in front of him, if need be a telescope, and the clear sky in front of the window. If this man does not give up before he understands. That would perhaps be science.

- Daniel Kehlmann, *Measuring the World*

Contents

Table of acronyms **1**

Notations **5**

Résumé étendu **9**

Introduction **15**

1 State-of-the-art of Compressive Sensing for Cognitive Radio applications **17**

1.1 Spectral Crowding and Cognitive Radio 18

1.2 Compressive Sensing 19

1.3 Information retrieval from the compressed samples 30

1.4 Conclusion 37

2 Compressive radiofrequency receiver architectures **39**

2.1 Framework 41

2.2 Non Uniform Sampling 47

2.3 Random Demodulation 52

2.4 Variable Rate Sampling 57

2.5 Conclusion 61

3 Code properties analysis for the implementation of a Modulated Wideband Converter **63**

3.1 From random matrices to structured acquisition 65

3.2 On the interest of CAZAC sequences for structured MWC acquisition 69

3.3 Analysis of the reconstruction performances of the sensing matrix 72

3.4 Conclusions and Perspectives 87

4 Spectral parameter estimation with compressive multiband architectures 89

4.1 Cramér-Rao bounds for spectral parametric estimation with compressive multiband architectures . . . 91

4.2 Applications to cognitive radio: definition of a framework . . . 94

4.3 Amplitude estimation . . . 95

4.4 Frequency estimation . . . 102

4.5 Influencing the Gramian . . . 104

4.6 Conclusions . . . 107

Conclusion 109

A Cramér-Rao bound 113

A.1 Fisher information matrix . . . 113

A.2 Cramér-Rao bound . . . 113

A.3 Efficient estimator . . . 114

B Computation of the MWC sensing matrix expression 115

B.1 Expression of the Fourier coefficients of the shaped code sequences . . . 115

B.2 Sensing matrix expression . . . 118

C Numerical benchmark elements between acquisition methods 119

D Exploration of time modulation in the NUWBS architectures with Zadoff-Chu codes controlled delays 121

D.1 Periodicity of the normalized sensing matrix of the NUWBS . . . 121

D.2 NUWBS architecture with Zadoff-Chu modulated wavelet stream . . . 122

E Proof of Theorem 1: Link between Fisher matrices for Compressive Sampling and Subsampling 127

F Proof of Theorem 2: Link between Fisher matrices for Subsampling and Nyquist Sampling 129

G Fisher matrix for spectral estimation of multitone signals with Nyquist Sampling **131**

Bibliography **145**

List of publications **147**

List of Figures

1.1 Spectrum access schemes and authorization regimes, inspired from [8]. 19

1.2 Matrix-wise depiction of the Compressive Sampling acquisition process. 22

1.3 Distance preservation between K-sparse vectors through Compressive Sampling. 24

1.4 Sparsity and different norm minimization, inspired from [41]. 27

1.5 Comparison of different subsampling methods in terms of information rate evolution during the acquisition process, inspired from [51]. 33

2.1 Classification of compressive radiofrequency architectures according to the way in which diversity is created. 41

2.2 Time and frequency notations. 43

2.3 Sensing matrix of a multiband architecture for two different formalisms. 45

2.4 Sketch of time sample selection in the Non Uniform Sampler. 48

2.5 Block diagram of the NUS. 48

2.6 Sensing matrix and process of the NUS (canonical formalism). 49

2.7 Non Uniform Wavelet Bandpass Sampling Block Diagram. 50

2.8 Sampling patterns of NUS, NUWS and NUWBS, inspired from [98]. 51

2.9 Non Uniform Wavelet Bandpass Sampling for a single branch, inspired from [98]. 51

2.10 Random Demodulator block diagram. 53

2.11 Action of the Random Demodulator on a pure tone, inspired from [100]. 54

2.12 Sensing process of the MWC, from 55

2.13 Mixing function $p_m(t)$. 55

2.14 Block diagram of Multirate Sampling. 58

2.15 Multirate Sampling Sensing Matrix. 58

2.16 Nyquist Folding Receiver. 59

2.17 Sensing matrix and process of the Nyquist Folding Receiver. 60

3.1 Real and imaginary part of Zadoff-Chu codes [$R = 1$, $N = 64$]. 71

3.2 Coherence $\mu(\boldsymbol{AF})$ comparison for Gold code matrices [$N = 255$]. 74

3.3 Coherence $\mu(\boldsymbol{AF})$ comparison for Zadoff-Chu codes [$N = 127$, $R \in [\![1; N]\!]$, prime N behavior]. 74

3.4 Coherence $\mu(\boldsymbol{AF})$ comparison for Zadoff-Chu circulant codes [$N = 255$, $R = 1$]. 75

3.5 Coherence $\mu(\boldsymbol{AF})$ comparison [$N = 255$, $R = 1$], from [113]. 76

3.6 Histogram of RIP-δ_K estimation [$N_s = 1000$, $N_v = 10000$, $K = 10$], from [113]. 77

3.7 Histogram of JLL-δ'_K estimation [$N_s = 1000$, $N_v = 200$, $K = 10$], from [113]. . . 78

3.8 Validation of the spectrum of the mixing codes [$K = 2, B = 5MHz$]. 81

3.9 Accuracy of the MWC with growing compression ratio. 82

3.10 Detection graph for different codes [$N = 255$, $R = 1$]. 83

3.11 Contour of the 95% confidence interval on the 80% true detection boundary of 3 previous figures, IC=Confidence interval. 84

3.12 Accuracy graph of the MWC with growing compression ratio [$N = 127$, $K = 6$, $R = 1$, 200 trials, OMP reconstruction algorithm], dashed=noiseless, plain=10dB ISNR, from [113]. 85

3.13 Contour of the 95% confidence interval on the 80% true detection boundary of noisy detection graph. 85

4.1 Disjoint subband model. 92

4.2 Variance comparison for amplitude estimation of 2 real tones. 101

4.3 Normalized off-diagonal coefficient of the Fisher matrix for frequency estimation $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ (top=overall, bottom=zoom), [$M = 10, N = 31$]. 104

4.4 Gramian of the Modulated Wideband Converter sensing matrix for different codes: Random Bernoulli (top) and Zadoff-Chu-based circulant (bottom) [N=31].106

B.1 Notations related to the mixing function $p_m(t)$. 116

B.2 Power Spectral Density of the shaped discrete Sequence $\tilde{p}_{m,[T_p]}(f)$. 117

C.1 Comparison of interferer detection performances, from [6]. 119

D.1 Definition of the different frequency aliasing zones. 122

D.2 Decomposition of the acquisition matrix of a NUWBS architecture with Zadoff-Chu modulated wavelet stream. 123

D.3 Module of the correlation $\Gamma_{n,n+L}$ between f_s-spaced columns of the ΔZAC-WBS sensing matrix, $[L = 32, R = 1,\ f_c = 4GHz,\ f_s = 125MHz]$. 125

D.4 Module of the correlation for the ΔZAC-WBS, in dBc, for varying frequency $n'\delta f$ and fixed $n\delta f = f_c$. 126

List of Tables

1.1 Definition of the coefficients necessary to ExRIP probability computation (real-valued matrices). 25

1.2 Comparisons between reconstruction algorithms, from [34], α is a non-zero constant. 28

1.3 Summary table of estimation methods from the compressed samples. 36

3.1 Experiment settings, from [123], $\boldsymbol{\Phi} = \boldsymbol{RC}$. 67

3.2 Comparison between different circulant matrices $\boldsymbol{\Phi\Psi}$ (settings described in the top) in terms of required number of measurements M (bottom), from [124] . . . 68

3.3 ExRIP [$M = 80$, $N = 511$, $K = 24$, $\delta_{2K} = \sqrt{2} - 1$, $R = 1$]. 78

3.4 Parameters of the MWC simulation. 80

3.5 Conclusion table. 87

4.1 Set-up parameters. 101

Table of acronyms

ADC	*Analog-to-Digital Converter*
AIC	*Analog-to-Information Converter*
AM	*Amplitude Modulation*
AMR	*Automatic Modulation Recognition*
ASIC	*Application-Specific Integrated Circuit*
BP	*Basis Pursuit*
BPS	*Bandpass Sampling*
BPSK	*Binary Phase Shift Keying*
CAZAC	*Constant Amplitude Zero AutoCorrelation*
CR	*Cognitive Radio*
CRB	*Cramér-Rao Bound*
CS	*Compressive Sampling/Compressive Sensing/Compressed Sensing*
CSF	*Compressive Sensing Filter*
ΔZAC-WBS	*Delay Zadoff-Chu Wavelet Bandpass Sampling*
DFT	*Discrete Fourier Transform*
DPSS	*Discrete Prolate Spheroidal Sequences*
DSA	*Dynamic Spectrum Access*
DSS	*Dynamic Spectrum Sharing*
ENOB	*Effective Number of Bits (measures the dynamic range)*
ExRIP	*Expected RIP, probability that RIP is satisfied under specific assumptions*
FM	*Frequency Modulation*
IDFT	*Inverse Discrete Fourier Transform*
i.i.d	*independent and identically distributed*
I/Q	*In-phase/Quadrature*
IoT	*Internet of Things*

ISI	*Intersymbol interference*
ISNR	*Input Signal-to-Noise Ratio*
JLL	*Johnson-Lindenstrauss Lemma*
LASSO	*Least Absolute Shrinkage and Selection Operator*
LoRa	*Long Range, a patented low-rate digital wireless communication technology for the IoT*
LTNUS	*Level Triggered Non Uniform Sampling*
ML	*Maximum Likelihood*
MRS	*Multirate Sampling*
MUSIC	*Multiple Signal Classification*
MWC	*Modulated Wideband Converter*
NP-hard	*Non-deterministic polynomial-time hardness. A decision problem H is NP-hard when for every problem L in NP, there is a polynomial-time reduction from L to H.*
NUS	*Non Uniform Sampling/Sampler*
NUWBS	*Non Uniform Wavelet Bandpass Sampling*
NYFR	*Nyquist Folding Receiver*
OMP	*Orthogonal Matching Pursuit*
PAPR	*Peak-to-Average-Power-Ratio*
PNUS	*Periodic Non Uniform Sampling/Sampler*
PSD	*Power Spectral Density*
PSK	*Phase Shift Keying*
Radar	*Radio Detection and Ranging*
QAIC	*Quadrature Analog-to-Information Converter*
RBW	*Resolution Bandwidth*
RC	*Random Convolution*
RD	*Random Demodulator*
RF	*Radiofrequency ($20kHz - 300GHz$)*
RIP	*Restricted Isometry Property*

RMPI	*Random Modulation Pre-Integrator*
RNUS	*Randomized Non Uniform Sampling*
SIR	*Signal-to-Interferer Ratio,* $10\log_{10}(\frac{Signal\ Power}{Interferer\ power})$
SNR	*Signal-to-Noise Ratio,* $10\log_{10}(\frac{Signal\ Power}{Noise\ power})$
s.t.	*such that*
StaRIP	*Statistical RIP*
UoS	*Union of Subspaces*
UWB	*Ultra Wideband (Bandwidth>*$250MHz$ *generally)*
Xampling	*Contraction of 'Compressive' and 'Sampling', framework introduced by [1]*
ZC	*Zadoff-Chu or Frank-Zadoff-Chu sequences*

Notations

Matrices and vectors will be written with uppercase and lowercase boldface letters respectively.

Notations	Signification
$.^T$	Transpose
$.^*$	Conjugate
$.^H$	Hermitian conjugate
$.^{-1}$	Inverse
$*$	Convolution
$\circ$	Hadamard product
$Tr(.)$	Trace operator
$\hat{.}$	Estimation
$\lfloor . \rfloor$	Floor operator
$\text{mod}\ (.)$	Modulo operator
$< ., . >$	Scalar product
$[\![\ldots;\ldots]\!]$	Interval of integers
$\cup$	Union
$\Re(.)$	Real part
$\Im(.)$	Imaginary part
$\lvert . \rvert$	Module
$\lvert . \rvert_p$	Norm lp, $\lvert \boldsymbol{x} \rvert_p = \sqrt[p]{\sum_n x[n]^p}$
$x(t)$	Analog time-domain input signal
$y(t)$	Analog time-domain output signal
$\boldsymbol{x}[.]$	Discrete time-domain input vector
$\tilde{x}(f)$	Fourier Transform of $x(t)$
$\tilde{\boldsymbol{x}}[.]$	Discrete Fourier Transform (DFT) of $x(t)$
$x^+(t)$	Noisy analog time-domain input signal
$x_\sharp$	Reordered NxL input signal for the compact multiband formalism
$y_\sharp$	Reordered MxL output signal for the compact multiband formalism
f_p	Width of a subband in the signal model
T_p	Period of the mixing code
f_s	2x the filter cut-off frequency
T_s	Sampling period
δf	Frequency resolution
T_{Nyq}	Time resolution of the input signal
T_{acq}	Acquisition length such that there are NL input samples
f_{Nyq}	Nyquist sampling rate of $x(t)$

T_c	Chip period of the code in the MWC
K	Degree of sparsity of $x(t)$
L	Number of bins of width δf in a subband f_p or ratio between acquisition time and period of the codes
M	Number of branches, typically
N	Number of subbands, assumed odd in the MWC
N_0	Floored half of N
Q	Length of the parameter vector
$\Gamma(.)$	Crosscorrelation function
$\Lambda(.)$	Support index indicator function
λ_q	Support index of the qth parameter
$\tilde{x}_q$	Subsignal influenced by the qth parameter
$\Delta f_{q,r}$	$-f_q + f_r$
$\Sigma f_{q,r}$	$f_q + f_r$
$\Delta\phi_{q,r}$	$-\phi_q + \phi_r$
$\Sigma\phi_{q,r}$	$\phi_q + \phi_r$
$\bar{f}$	f folded at baseband by subsampling
β_k	Support-limited function representing the kth subband
δ_K	RIP constant of sparsity K
$\eta_{1,2}$	Offset to the frequency grid of Δf, in freq. resolution < 1
$\eta'_{1,2}$	Offset to the frequency grid of Σf, in freq. resolution < 1
$\gamma_{m,n}$	Fourier coefficients of the shapped mixing codes in the MWC
μ	Coherence
μ_m	Mutual coherence
σ	Noise variance
$\boldsymbol{A}_{i,.}$	i^{th} row of $\boldsymbol{A}$
$\boldsymbol{A}_{.,i}$	i^{th} column of $\boldsymbol{A}$
$\boldsymbol{A}$	Matrix expressing, for the MWC the mixing codes in the time domain, and for the NUWBS the operations before time subsampling
$\boldsymbol{B}$	Multiband architecture matrix, compact formalism
$\check{\boldsymbol{B}}$	Multiband architecture matrix, canonical formalism
$\boldsymbol{C}$	Circulant matrix
$\boldsymbol{D}$	Diagonal matrix
$\boldsymbol{F}$	Discrete Fourier Transform matrix
$\boldsymbol{J}$	Fisher information matrix
$\boldsymbol{R}$	Matrix of Dimension reduction/Subsampling
$\boldsymbol{W}$	Wavelet frame
$\boldsymbol{Z}$	Sensing matrix applied on the sparsity domain $\boldsymbol{Z}=\boldsymbol{\Phi\Psi}$

$\boldsymbol{\Gamma}$	Covariance matrix
$\boldsymbol{\Phi}$	Acquisition/Sensing matrix describing the Compressive Sensing acquisition process on the time signal
$\boldsymbol{\Psi}$	Matrix corresponding to the basis change to the sparsity domain (Sparsifying matrix)
$\boldsymbol{\Phi\Psi}$	Acquisition/Sensing matrix describing the Compressive Sensing acquisition process on the sparse representation
$\boldsymbol{\theta}$	Vector of estimation parameter
$\boldsymbol{\Lambda}$	Support
$O(.)$	Landau notation for asymptotic evaluation
BW_{RF}	Radiofrequency bandwidth to consider in the NUWBS
@	Sampling rate indication

Résumé étendu

Ces dernières années ont été marquées par l'explosion de la quantité d'information échangée via les communications sans fil. En particulier, le volume du trafic des données mobiles a augmenté d'un facteur 18 entre 2011 et 2017 [2]. Le marché en pleine expansion de l'Internet des Objets, des objets physiques connectés tels que des appareils électroménagers intelligents, est également concerné : 500 milliards d'appareils sont attendus à l'horizon 2030 [2].

Pourtant, à un temps et un endroit donné, le spectre fréquentiel est majoritairement non occupé (le taux d'occupation estimé se situe entre 2 et 5 %, y compris dans des zones métropolitaines [3]). Mais, en raison d'une politique de régulation rigide basée sur l'attribution de licences, l'essor des communications sans fil a conduit à une situation préoccupante en termes de raréfaction de la ressource spectrale. En conséquence, le prix de l'accès au spectre augmente et des tensions apparaissent entre les différents usagers (cellulaire, gouvernemental, sécurité publique, non licencié, ...). Pour les seuls Etats-Unis, un rapport gouvernemental [4] a estimé à 1000 milliards de dollars et des millions de créations d'emplois le bénéfice potentiel si l'on parvient à relâcher la contrainte que constitue la rareté du spectre.

Ce problème appelle à une transformation du modèle de gestion de l'accès au spectre, de sorte que l'accès des utilisateurs secondaires aux bandes de fréquences non licenciées soit facilité. Pour porter ce changement, il est nécessaire de développer la Radio Cognitive (CR), des récepteurs radiofréquences (RF) conscients de l'environnement spectral et capables de s'adapter de façon intelligente en conséquence. Le récepteur intelligent doit notamment être en mesure de détecter des bandes de fréquences vides et de reconfigurer ses paramètres de façon à éviter les interférences entre utilisateurs. Cependant, l'observation efficace et à bas coût de larges bandes de fréquences constitue un défi technique majeur pour les Convertisseurs Analogiques-Numériques (CAN). Concilier à prix raisonnable les impératifs de résilience au bruit et ultra faible consommation avec des fonctions d'estimation paramétrique sur de larges bandes de fréquences est une véritable gageure.

La théorie de l'acquisition compressée (CS) introduite en 2006 par E. Candès *et al.* [5] apparaît comme un candidat prometteur pour satisfaire à ces contraintes. L'acquisition compressée est un récent changement de paradigme en termes d'acquisition de données, capable de passer outre la cadence minimale d'échantillonnage de Nyquist en considérant plutôt le concept de quantité d'information dans la base de représentation adaptée. Supposons que le signal est parcimonieux, c'est-à-dire qu'il est décrit par un nombre réduit de coefficients non nuls dans une base appropriée, le domaine spectral dans le cadre de notre étude. Alors un faible nombre de mesures incohérentes est suffisant pour reconstruire le signal originel. Le nombre d'échantillons nécessaires dépend alors de la bande de fréquence agrégée et non plus de la fréquence maximale du signal d'entrée. Pour ces raisons, il est pressenti que l'acquisition compressée va bouleverser les compromis usuels de la conception de Convertisseurs Analogique-Numérique (CAN) en matière de bande, figure de bruit et consommation. Cependant la reconstruction du signal d'origine à partir des échantillons compressés est un processus d'optimisation non-linéaire, gourmand en consommation et ainsi peu compatible

avec des solutions radio embarquées.

Pour éviter cet écueil, il est proposé de ne pas effectuer l'étape de reconstruction et d'extraire plutôt les caractéristiques d'intérêt pour les applications envisagées directement à partir des échantillons compressés. Dans ce but, il est nécessaire d'étudier à la fois les challenges de l'implémentation d'un récepteur radiofréquence compressé efficace et les spécificités de l'estimation paramétrique à partir des échantillons compressés, par opposition à l'estimation à partir de la reconstruction du signal d'origine. Dans cette thèse, nous abordons dans un premier temps les problématiques d'implémentation, qui sont considérées à travers le prisme d'une tâche de reconstruction car cet aspect est mieux compris et avancé dans la littérature que l'extraction de caractéristiques. Les spécificités de l'extraction paramétrique seront ensuite pleinement prises en compte.

Une métaphore naïve de notre approche serait la suivante. Supposons qu'on dispose d'une balance à plateau, d'un étalon de mesure et de 9 lingots d'or, dont on sait que pas plus d'un n'est truqué (c'est-à-dire plus léger). Une analogie de l'acquisition à la cadence de Nyquist serait de comparer chacun des lingots au poids étalon. Mais, deux mesures seulement (par exemple mesurer deux lots de trois lingots selon lignes et colonnes, $1+2+3$ versus $4+5+6$ et $1+4+7$ versus $2+5+8$) permettent de savoir si un lingot est truqué et si oui lequel, avec un peu de déduction. L'idée qui est exploitée dans ces travaux est similaire, dans une certaine mesure.

Ce manuscrit comporte quatre chapitres, les deux premiers chapitres dressent un panorama qui se veut presque exhaustif d'un domaine en pleine effervescence, et les deux chapitres suivants apportent des contributions originales.

Dans le premier chapitre sont introduits les enjeux de la radio cognitive pour l'accès dynamique au spectre. L'intérêt de développer des récepteurs radiofréquences capables de sonder un environnement spectral multistandard avec des contraintes aiguës de coût et de consommation est souligné. Comme solution à ces contraintes, la théorie de l'acquisition compressée est mise en avant car elle permet de passer outre l'impasse qui consiste à utiliser un CAN à haute fréquence d'échantillonnage. L'acquisition compressée dit qu'un nombre réduit de mesures est suffisant pour reconstruire le signal sous hypothèse de parcimonie du spectre du signal. Si le contenu informatif est réparti de façon adéquate entre les mesures, il devient possible de résoudre ce problème sous-déterminé. Ainsi cette approche pourrait permettre d'aller au-delà de la limite de Nyquist, en se concentrant plutôt sur la grandeur physique correspondant à la bande de fréquence agrégée.

Pour commencer, un bref historique remontant jusqu'aux années 60 est dépeint. Ensuite, les deux éléments fondamentaux qui rendent possible la reconstruction du signal d'origine en acquisition compressée sont décrits : les conditions sur la matrice qui modélise l'acquisition d'une part, et les algorithmes de reconstruction d'autre part. Tout d'abord, des métriques d'évaluation concernant la matrice d'acquisition sont présentées. En particulier la cohérence, essentielle pour limiter le nombre de mesures nécessaires, et la préservation des normes et distances par la projection (isométrie), essentielle par rapport à la résilience au bruit. Ensuite les différents algorithmes de reconstruction sont abordés, et le principe de fonctionnement des deux plus connus décrit : Orthogonal Matching Pursuit (OMP), qui est basé sur la recherche

gloutonne de la colonne du dictionnaire de représentation la plus corrélée avec les observations et Basis Pursuit, qui est basé sur la relaxation dans la formulation du problème de reconstruction de la pseudo-norme $l0$ par la norme $l1$. La comparaison montre que la magie de l'acquisition compressée a cependant un prix : Au vu de la complexité algorithmique des méthodes de reconstruction, il est choisi d'écarter l'option de la reconstruction pour une solution embarquée.

C'est la raison pour laquelle un focus est ensuite fait sur l'actuel état de l'art concernant l'extraction d'information partielle directement à partir des échantillons compressés. Tout d'abord, il convient de définir les paramètres d'intérêt pour les applications de radio cognitives et la reconfigurabilité. Ensuite, des garanties concernant les tâches d'estimation paramétrique et de classification sont données. Il apparaît que ces garanties sont essentiellement basées sur la préservation des normes et distances. Le principe d'extraction de caractéristiques avec perte d'information est également introduit. Enfin, les solutions existantes, inspirées des méthodes à la cadence de Nyquist, sont présentées et leurs limitations exposées.

De cette étude, il ressort l'existence de lacunes qui empêchent l'établissement d'une vraie méthodologie générale dans la conception d'un Convertisseur Analogique-Information (AIC) pour des tâches de détection de spectre.

Dans le chapitre 2, nous nous sommes attachés à mettre en pratique les notions abordées et à passer en revue les différentes architectures de récepteurs radiofréquence compressés. Une pléthore de convertisseurs Analogique-Information a été proposée dans le but de réaliser les promesses de l'acquisition compressée. Dans un souci de clarté, les architectures ont été regroupées suivant trois principes généraux de fonctionnement : échantillonnage non-uniforme, démodulation aléatoire et cadence d'échantillonnage variable. Les convertisseurs exploitent plusieurs idées provenant des récepteurs RF traditionnels comme l'étalement de spectre et le sous-échantillonnage, dans le but de créer de la diversité de différentes façons. Par exemple, les techniques de sous-échantillonnage aléatoire (Récepteur à repliement de Nyquist (NYFR), échantillonnage sub-Nyquist multicadence asynchrone (MASS)) s'inspirent des techniques de sous-échantillonnage usuelles (bandpass sampling) pour créer différents mélanges qui peuvent être démêlés. Aussi, les démodulateurs aléatoires (Démodulateur aléatoire (RD), Convertisseur modulé à large bande (MWC)) opèrent du codage par modulation avec un générateur de nombre (pseudo)-aléatoire à haute fréquence.

Il est cependant un peu décevant de constater que, si le signal est bien échantillonné sous la fréquence de Nyquist, de nombreuses architectures nécessitent des composants dépendant de cette fréquence. Un autre problème réside dans le manque de flexibilité de la plupart des architectures, un handicap pour des applications de radio cognitive. On compte parmi les autres points faibles les coûts de reconstruction prohibitifs, l'inévitable repliement du bruit qui dégrade le rapport signal à bruit (SNR) et une surface de silicium rendue parfois trop importante par la parallélisation de l'architecture.

Les architectures suivantes sont particulièrement encourageantes vis-à-vis des points bloquants identifiés : Le Convertisseur Analogue-Information en Quadrature (QAIC), basé sur un MWC en bande de base, fait l'objet de toutes les attentions car il parvient à améliorer le MWC en matière de consommation et de simplicité du générateur de code. Les performances en détection d'interférences présentées dans [6] éveillent particulièrement l'intérêt.
Le récepteur à repliement de Nyquist (NYFR) est une méthode astucieuse combinant le sous-

échantillonnage et la modulation par un train d'impulsions, et qui présente deux avantages : absence de composants à la fréquence de Nyquist et préservation de la structure du signal qui permet une restitution rapide de l'information.
Le sous-échantillonneur non uniforme d'ondelettes (NUWBS), qui ne nécessite pas non plus de composants à la fréquence de Nyquist, est lui aussi très prometteur grâce à une approche adaptative multi-échelles. Vu qu'il offre un haut degré de flexibilité (instant d'échantillonnage, fréquence centrale et support temporel des ondelettes), il évite les défauts classiques des architectures AIC et se présente entre autres comme le choix naturel pour des signaux d'entrée parcimonieux en temps et en fréquence.

Les codes de mélanges se distinguant comme étant l'élément structurant des architectures à démodulation aléatoire et aussi comme le point clé de leur implémentation, leurs propriétés font l'objet du Chapitre 3, dans le cadre d'un convertisseur modulé à large bande (MWC). Tout d'abord, les enjeux de la génération des codes de mélange sont esquissés. La question de la signification des notions d'aléatoire et d'universalité appliquées à l'implémentation d'architectures concrètes est posée. Les bénéfices attendus des matrices circulantes du point de vue de l'implémentation sont évoqués et leur performance en reconstruction soulignées.

Capitalisant sur les propriétés mises en lumière dans l'état de l'art poussé, la proposition d'une nouvelle matrice de code est faite, en vue de son évaluation. Les pré-requis sont d'abord définis : une famille de code particulière, retenue pour ses bonnes propriétés de corrélation, les codes Constant Amplitude Zero Autocorrelation (CAZAC) est décrite. Puis un exemple type de code CAZAC, les codes de Zadoff-Chu, est exposé plus en détails. Enfin, la nouvelle matrice de codes, circulante et à valeurs réelles, est définie à partir des codes de Zadoff-Chu par l'utilisation d'une symétrie hermitienne dans le domaine fréquentiel.

Dans une troisième partie, une analyse multi-critères originale est présentée. C'est-à-dire que les différentes familles de code et différents opérateurs de sélection de ligne parmi la matrice complète des codes sont comparés en fonction de plusieurs métriques. Tout d'abord, les propriétés mathématiques sont considérées. En premier lieu, la cohérence qui est indispensable pour limiter le nombre de branches du récepteur est calculée et comparée pour différents codes. Les propriétés d'isométrie qui sont essentielles par rapport à la résilience au bruit sont, elles, évaluées grâce à deux outils : une première estimation de la propriété d'isométrie restreinte (RIP) est menée grâce à une simulation Monte-Carlo. En outre, une comparaison de l'Expected RIP, la probabilité que la RIP soit satisfaite étant donné des hypothèses additionnelles sur le signal d'entrée, est menée. Les considérations haut niveau sont ensuite validées grâce à une plateforme de simulation MATLAB® répliquant le processus d'acquisition du convertisseur modulé à large bande (MWC), pour des scénarii sans bruit et avec bruit. Pour résumer, l'évaluation des métriques haut niveau comme la validation en simulation montrent que les codes de mélange proposés sont plus performants que ceux de l'état de l'art, notamment en matière de robustesse au bruit. Un autre élément marquant est la mise en évidence du rôle prépondérant du schéma de sélection des lignes à l'intérieur de la matrice des codes possibles.

Concernant les problématiques d'implémentation, mais pour l'architecture NUWBS, les éléments clés en vue de l'ajout éventuel d'une modulation temporelle du NUWBS par des codes de Zadoff-Chu sont également mis en évidence en annexe. Ce point fait l'objet d'un travail en cours.

Enfin, la problématique de l'estimation paramétrique à partir des échantillons compressés est abordée, avec pour objectif de quantifier la précision de l'estimation paramétrique en fonction de la matrice d'acquisition et par rapport à un échantillonnage à la fréquence de Nyquist. L'estimation paramétrique basée sur les échantillons compressés diffère fondamentalement de l'estimation paramétrique traditionnelle, dans le sens où la structure du signal a été altérée lors du processus d'acquisition. Un outil adapté à cet objectif de quantification est la borne de Cramér-Rao sur la variance d'un estimateur non-biaisé. La borne de Cramér-Rao ne dit rien sur la meilleure méthode d'estimation, mais elle permet la quantification des performances envisageables en fonction de la configuration. Elle est obtenue par la diagonale de l'inverse de la matrice de Fisher, matrice qui traduit la quantité d'information à propos du paramètre contenue dans les mesures. Cette évaluation permet ainsi au système d'être conçu par rapport à des spécifications, un point essentiel car les contraintes sur les performances ne sont pas les mêmes que pour la reconstruction.

Une forme analytique de la matrice de Fisher pour des échantillons compressés issus d'une architecture multibande est établie sous l'hypothèse d'un modèle spectral disjoint. Cette expression est donnée en fonction de la matrice de Fisher pour de l'échantillonnage à la cadence de Nyquist et du Gramien de la matrice d'acquisition d'une architecture compressée multibande, en passant par une étape intermédiaire impliquant la matrice de Fisher pour du sous-échantillonnage. Comme l'expression obtenue est adaptée à chaque terme, elle est plus précise que les résultats de l'état de l'art qui fournissent des garanties sur l'ensemble de la matrice, basées sur les propriétés statistiques. Ces nouvelles expressions soulignent le rôle clé du Gramien de la matrice d'acquisition qui traduit les contraintes géométriques en termes de produit scalaire. Cette contribution permet aussi de dissocier les effets de la compression (aspect sous-échantillonnage) de ceux de la création de diversité (aspect lié au Gramien). Les architectures multibandes sont des architectures où chaque bin fréquentiel d'une sous-bande est traité de la même manière. Comme cette définition inclut la plupart des architectures répandues (MWC, QAIC, PNUS, MRS, BPS) les résultats ont une portée assez générique.

L'influence du processus d'acquisition, notamment le couplage entre paramètres et la fuite spectrale, est illustrée par l'exemple. Dans le cadre de la radio cognitive, le cas d'application de la détection d'un interféreur en présence d'un signal utile est choisi. Tout d'abord, le cas de l'estimation de l'amplitude est posé. Les bornes théoriques sont déterminées ainsi que des variances expérimentales issues d'un estimateur de maximum de vraisemblance. Ensuite, le cas de l'estimation de fréquence est considéré. L'interprétation des formules établies n'étant pas évidente, une illustration est donnée en complément. Au vu des expressions nouvellement établies, il est souligné que les effets délétères de la présence d'un interféreur de forte puissance peuvent être contrôlés et attenués via l'ajustement des coefficients du Gramien de la matrice d'acquisition. Cela offre des opportunités intéressantes d'approches adaptées à des a priori sur la distribution du spectre. Les propriétés de cohérence et d'isométrie qui découlent du Gramien ayant été étudiées en détails dans le Chapitre 3, l'optimisation peut être menée en suivant les mêmes principes.

En revanche, ces résultats signifient aussi que, dans ce cadre, aucun paramètre n'est plus robuste vis-à-vis du processus d'acquisition compressée que les autres.

Dans ce manuscrit, un pont supplémentaire a été construit pour permettre d'enjamber le fossé entre les objets abstraits dont traite la théorie mathématique de l'acquisition compressée

et les réalités du terrain de la détection de spectre radiofréquence.

Beaucoup de questions restent encore ouvertes, comme le choix d'un estimateur paramétrique et son implémentation.

Introduction

Recent years have been marked by a steep increase in the amount of data exchanged via wireless communications. In particular the volume of mobile data traffic has increased 18-fold between 2011 and 2017 [2]. Also the rapidly expanding market of the Internet of Things (IoT), connected physical devices such as smart home appliances and other monitoring systems, is concerned: 500 billions connected devices are expected by 2030 [2].

However, at a given time and place, the frequency spectrum is usually mostly empty (an estimated 2-5% occupation, even in metropolitan areas [3]). But, because of the rigid license-based spectrum access regulation, the boom of wireless communications has lead to a worrying situation in terms of scarcity of spectral resources. As a consequence, the price of the spectrum access rises and tensions may emerge between the different users (cellular, public safety, governmental, unlicensed, ...). In the USA alone, the gain expected from achieving to relax spectrum scarcity has been estimated to more than $1 trillion benefit and millions of job creations over a decade [4].

This bottleneck calls for a model transformation in terms of spectrum access management, in order to enable better dynamic spectrum access for secondary users on unlicensed bands. To carry out this change, it is mandatory to develop Cognitive Radio (CR), radiofrequency receivers that are aware of the spectral environment and able to adjust smartly in consequence. The cognitive receiver should notably be able to detect unused frequency bandwidths and reconfigure its parameters in order to use them without creating interferences between users. The analysis of the spectral environment also enables to optimize the operating point regarding the energy consumption. However, efficiently monitoring at low-power and low-cost wide areas of spectrum is a major technical challenge for Analog-to-Digital Converters. Combining over wideband functions performing the extraction of the parameters of interest, called features, with the imperatives of robustness and ultralow-power consumption is a dare.

The Compressive Sampling (CS) theory introduced in 2006 by E. Candès *et al.* [5] is an appealing candidate to meet these constraints. Compressive Sampling is a recent paradigm shift in data acquisition, able to overcome the Nyquist rate deadlock by focusing on the concept of information amount under the adequate representation instead. If the signal is described by few non-zero coefficients in an appropriate basis, which is the spectral domain in the context of this study, the signal is said to be sparse and a small amount of incoherent samples is sufficient to recover the original information. The number of samples then depends on the gathered frequency bandwidth and not the maximal frequency of the input. Therefore the Compressive Sampling framework is foreseen to disrupt the usual trade-offs between bandwidth, noise figure and energy consumption in Analog-to-Digital-Converters conception.

However, the reconstruction of the original signal from the acquired samples is a non-linear optimization process, highly energy-consuming and hence not readily compatible with the context of embedded radio solutions. In order to avoid this pitfall, it is proposed to skip the reconstruction step and extract the features of interest for the targeted application directly from the compressed samples. But how should we design an Analog-to-Information Converter

tailored to feature extraction for Cognitive Radio ?

To answer this question, it is necessary to investigate both the implementation challenges of an efficient compressive radiofrequency receiver and the specificities of parametric estimation with compressed samples, as opposed to estimation from the reconstruction of the original signal. At first, implementation challenges will be considered through the prism of a reconstruction task, as it is better understood and advanced than for feature extraction. Specificities of estimation from the compressed samples will later be fully taken into account.

The manuscript will be organized as follows in four chapters. Chapters 1 and 2 provide a panorama which aims to be almost exhaustive of a field in booming expansion, Chapters 3 and 4 are original contributions.

In Chapter 1, the potential of Compressive Sensing for the development of Cognitive Radio is highlighted. First, it is underscored that developing further Cognitive Radio solutions is necessary to be able to relax the spectrum crowding conundrum by exploiting Dynamic Spectrum Sharing. Then, the Compressive Sensing theory, a potential response to the exposed challenges, is presented. Finally, as the reconstruction stage is discarded because computationally expensive, the stakes of information retrieval directly from compressed samples are exposed.

In the second chapter, the State-of-the-Art of existing compressive radiofrequency architectures is presented and discussed. Architectures are classified according to three main functioning principles and their drawbacks are highlighted. Among the many, two solutions rise the interest: the Modulated Wideband Converter (MWC) due to its robustness to grid mismatch and the Non Uniform Wavelet Bandpass Sampling (NUWBS) due to its flexibility.

Since the mixing codes are identified as the core and structural element of architectures based on random demodulation, their properties will be investigated further in Chapter 3, focusing on the exemple of a MWC receiver. To begin with, the potential of circulant code matrices is highlighted. An original code is proposed and a detailed evaluation benchmark with the state-of-the-art is displayed. The comparison methodology is based on both high-level metrics (coherence, isometry) and simulations on a MATLAB® platform featuring the MWC acquisition and reconstruction process.

Then, Chapter 4 focuses on the accuracy of feature extraction on the compressed samples without a reconstruction step. What differences should be considered in comparison to Analog-to-Information Converters performing reconstruction ? The Cramér-Rao lower bound on the variance of any unbiased estimator is considered in order to understand and quantify the effect of Compressive Sampling on parametric estimation guarantees. An original closed-form expression of the bound, more accurate than previous results based on global guarantees, will be established for compressive multiband architectures. The mechanisms involved are explained and the effects of bandpass sampling and diversity dissociated. A typical scenario of Cognitive Radio, detection of an interferer near useful signal, is detailed as an illustrative example.

Finally a conclusion is drawn and perspectives are given.

Note that all figures in this manuscript, except Figure C.1, have been produced or reproduced by the author. Some of them are strongly inspired by a specific source, which will be notified with the mention 'inspired from [xxx]'.

Chapter 1

State-of-the-art of Compressive Sensing for Cognitive Radio applications

Contents

1.1 Spectral Crowding and Cognitive Radio 18

1.2 Compressive Sensing 19

1.2.1 Historical development 20

1.2.2 Framework and notations 21

1.2.3 Reconstruction 22

1.2.3.1 Metrics 22

1.2.3.2 Benchmark of non-adaptive recovery algorithms 26

1.2.3.3 Adaptive approaches 28

1.3 Information retrieval from the compressed samples 30

1.3.1 Parameters of interest 30

1.3.2 Bounds and guarantees for Analog-to-Feature extraction 31

1.3.2.1 Compressive parametric estimation: Cramér-Rao bounds 31

1.3.2.2 Compressive classification 31

1.3.3 Existing solutions for compressive information retrieval 34

1.4 Conclusion 37

1.1 Spectral Crowding and Cognitive Radio

Even though at a given time and space the radiofrequency spectrum is mostly empty, spectral ressource scarcity has become an increasingly important issue in recent years due to the rigid framework of license-based spectrum access regulation. The growing amount of data exchanged in mobile services and the emergence of the Internet of Things (IoT) have provoked a drastic increase in the spectrum demand. A demand which becomes more and more difficult to meet by spectrum allocation based on an exclusive use policy.

In response to the growing pressure on the available spectrum bands, the multiple legacy regulatory organisations (the *Federal Communications Commission* in the USA, the *European Conference of Postal and Telecommunications Administrations* in Europe, the *Agence Nationale des fréquences* in France, etc.) have begun to promote the opportunity of Dynamic Spectrum Sharing (DSS) between different types of users. Dynamic Spectrum Sharing is based on Dynamic Spectrum Access (DSA), which is formally defined as 'real-time adjustement of spectrum utilization in response to changing circumstances and objectives' [7]. Typically, Secondary Users seek for opportunistic access to the spectrum, as opposed to Incombent Users that have a warrantied legal access right.

Accurate details on this legal policy evolution and the concerned frequency subbands, as well as on the current types of user accesses are provided in [8].

One famous example of DSS are the IEEE Standards *802.22* and *802.11af* which allow low-power devices in several countries (USA, UK, Canada, etc.) to operate on unused TV broadcast bands, known as TV White Spaces.

Figure 1.1 illustrates the multiple spectrum access authorization regimes currently utilized, and in particular the possibilities for unlicensed accesses.

Dynamic Spectrum Access is a consequential paradigm shift in the domain of communications, challenging in regard to the definition of new access protocols as well as associated with significant technical issues. It requires in particular ultra low-power radiofrequency receivers that are aware of the radio environment and are able to smartly adapt operational aspects. The goal is to achieve robust and enhanced performances with low-power consumption via the understanding of the radio environment and reconfigurability. Ultra low-power is all the more important because of what is called the 'Deploy and Forget' approach in the Internet of Things: a huge number of devices is settled, which should not require active maintenance such as battery substitutions from the user. In the road-map outlined in [8] with respect to Cognitive Radio development for Dynamic Spectrum Sharing, there are among others, following technical points to be dealt with:

- low-power for sustainability
- improvements in smart radio architectures to support high dynamic range for wideband operations
- reconfigurable radio hardware, interference nulling capabilities
- definition of techniques and standards for spectrum measurement

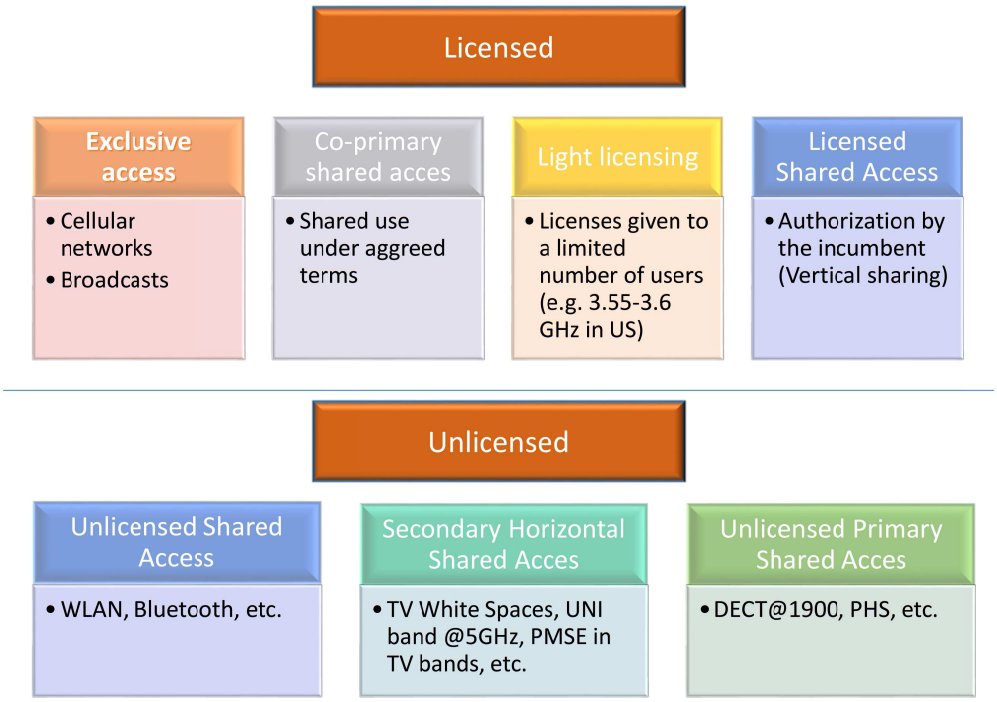


Figure 1.1: Spectrum access schemes and authorization regimes, inspired from [8].

One solution foreseen to be able to meet the exposed challenges would be to exploit the emerging theory of Compressive Sensing (CS), since it relaxes the Nyquist-rate constraint. Investigating further the opportunities offered by Compressive Sensing for Cognitive Radio applications will be the topic of this study.

Hence the concept of Compressive Sampling will be presented in the next Section. First, the roots of Compressive Sensing are sketched in Subsection 1.2.1. Then it is necessary to introduce precisely the framework and notations in Subsection 1.2.2. In following Subsection 1.2.3, the issue of the reconstruction of the original signal is investigated. This is done by discussing different evaluation metrics regarding the sensing matrix and sketching and comparing possible reconstruction algorithms. Given the computational complexity, the conducted analysis leads to the choice of exploring the potential of partial information restitution. As a consequence, the topic of information retrieval from compressed samples for cognitive radio applications is tackled in Section 1.3.

1.2 Compressive Sensing

Compressive Sensing, also known as Compressive Sampling or Compressed Sensing, represents a revolution in the way that information is extracted. This recently developed framework aims at capturing the fair amount of information through a reduced number of incoherent

measurements. It is no longer necessary to acquire data at the Nyquist-Shannon rate but rather depending on what can be seen as the 'intrinsic dimension' of the signal, which is the minimal number of parameters required to describe it in some appropriate dictionary or basis.

If it is assumed that the signal is sparse (few non zero coefficients represent it in the said basis), it is possible to reconstruct or estimate the signal from the compressed samples through non-linear convex optimization or iterative methods. For instance, frequency tones are sparse in the Fourier domain or Radar pulse signals are sparse in the time-frequency domain.

The underlying goal is to extract and acquire only the relevant information directly at the sensing stage, in order to avoid waste of power and storage. In this way, a new constraint balance between acquisition chain and information retrieval should be figured out.

1.2.1 Historical development

For more than half a century the cornerstone of signal processing and information theory has been the Shannon-Nyquist-Whittaker theorem on the choice of the sampling frequency:

Every signal of finite energy and bandwidth W Hertz may be completely recovered from taking its samples at the rate of $2W$ per second [9]

For signals with known high frequencies, the constraints can be relaxed by using a heterodyne receiver [10] which translates the bands of interest on an intermediate frequency. But if the frequencies are unknown, this is not possible.

Premises of the present Compressive Sensing framework emerged as early as the 1960s. In 1965, the development of recovery algorithms based on $l1$-norm minimization was launched by the statement of the Logan's phenomenon:

If the product of the signal bandwidth and the measure of the support of the noise is sufficiently small, we can find a perfect decomposition of the signal into band-limited function and impulsive noise by finding the band-limited function closest to the observed signal in the $l1$ sense [11]

The relaxation in the expression of the recovery problem of the pseudo-norm $l0$ to the norm $l1$ yields, under conditions stated above, the correct original signal and simplifies computationally the resolution. Under assumption of spectral support knowledge, the Landau's theorem established in 1967 bypasses the previous Nyquist rate by considering the bandwidth occupation instead of the maximal frequency:

Every signal of finite energy may be completely recovered from its samples taken at the rate of $2W_{occupied}$ [occupied bandwidth] as soon as we have knowledge of the spectral support [12]

Other steps were made and major developments in $l1$-based reconstruction techniques occurred in the 1990s, with Matching Pursuit [13] or LASSO [14] algorithms. In 2006, E. Candès, T.

Tao, J. Romberg and D. Donoho built on those foundations to set a new groundbreaking frame for signal acquisition with a complete and formalized Compressive Sensing theory as well as optimized recovery criteria [5, 15, 16].

Compressive Sensing (CS) already proved itself very valuable in specific constraints settings: for example impossibility to take complete measurements as in imaging (e.g. medical imaging [17]), or strong asymmetry between acquisition and reconstruction costs (e.g. reconstruction delocalized from the sensor to a fusion center). In the field of radiofrequency (RF) communications, various compressive receiver architectures have been proposed since 2006, which is at the core of our topic and detailed in Chapter 2. CS is still a growing topic of interest and investigations must be carried on to identify and benchmark the potential interest for other domains and applications.

1.2.2 Framework and notations

All notations are summarized in the notation table.
As usual, $\lfloor . \rfloor$ will denote the floor operator, mod (.) the modulo operator, $.^*$ the conjugate, $.^H$ the conjugate transpose, $< ., . >$ a scalar product and $[\![\ldots ; \ldots]\!]$ an interval of integers.
The tilde $\tilde{.}$ will refer to the continuous Fourier Transform (of an analog or discrete time signal).
Indices of a vector are noted with [.] in both time and frequency domain.
The norm lp, noted $|.|_p$, is computed for a vector $\boldsymbol{x}$ as $|\boldsymbol{x}|_p = \sqrt[p]{\sum_n x[n]^p}$. If $p < 1$, the triangular inequality condition required from norms is not fulfilled, hence it is only a pseudo-norm. For $p = 0$, $l0$ matches the number of non-zero elements.

Let $\boldsymbol{x} \in \mathbb{C}^N$ be the sampled representation of a continuous time series $x(t)$. In Compressive Sensing, $\boldsymbol{x}$ is assumed to be sparse with respect to some dictionary or basis in the sense that there exists $\boldsymbol{s}$ such that $\boldsymbol{x} = \boldsymbol{\Psi s}$ and $\boldsymbol{s}$ has few non zero components ($||\boldsymbol{s}||_0 << N$). Let $\boldsymbol{\Phi} \in \mathbb{C}^{M \mathrm{x} N}$ denotes the acquisition matrix acting on the noisy input signal $\boldsymbol{x}^+ = \boldsymbol{x} + \boldsymbol{w}$. It is assumed that $M < N$.

Then the noisy signal at the output $\boldsymbol{y}^+$ is given by:

$$\boldsymbol{y}^+ = \boldsymbol{\Phi}(\boldsymbol{x} + \boldsymbol{w}) + \boldsymbol{w_{meas}} = \boldsymbol{y} + \boldsymbol{w}_{CS} + \boldsymbol{w_{meas}} \tag{1.1}$$

$$\text{with } \boldsymbol{y} = \boldsymbol{\Phi \Psi s} \tag{1.2}$$

where $\boldsymbol{w}$ and $\boldsymbol{w}_{CS} = \boldsymbol{\Phi w}$ are respectively the input noise (noise at the antenna) and the input noise after compression. $\boldsymbol{w_{meas}}$ refers to the measurement noise, i.e. the noise caused by the measurement hardware. $\boldsymbol{\Phi \Psi}$, or in some instances $\boldsymbol{\Phi}$, is called the sensing matrix.

In practice, most physical signals are not exactly sparse but considered compressible, meaning that there exists a sparse approximation of the signal. In this work focused on Cognitive Radio applications, $\boldsymbol{x}$ will be assumed to be sparse in the frequency domain. Thus $\boldsymbol{x}$ is the time domain signal sampled at Nyquist rate at the receiver input, $\boldsymbol{\Psi}$ is the Inverse Discrete Fourier Transform (IDFT) matrix $\boldsymbol{F^{-1}}$ and $\boldsymbol{s} = \tilde{\boldsymbol{x}} = DFT(\boldsymbol{x})$ is a frequency domain representation of the signal. Figure 1.2 illustrates matrix-wise Compressive Sensing acquisition and the different matrices, for a radiofrequency signal sparse in the frequency domain.

In the following subsection, it will be explained how reconstruction of the original from the compressed samples is possible.

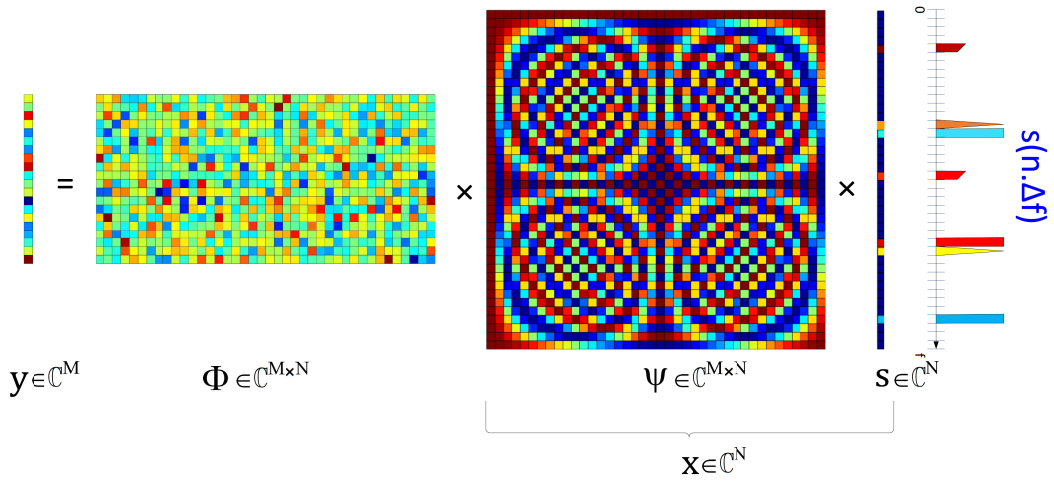


Figure 1.2: Matrix-wise depiction of the Compressive Sampling acquisition process.

1.2.3 Reconstruction

From an algebraic point of view, if N unknowns must be estimated from M linear measurements and $M < N$, the equation system is underdetermined, hence there exists an infinity of solutions. However, by adding the constraint of sparsity, it is possible with computationally-complex algorithms to completely recover the original signal, provided that the acquisition process satisfies some specific properties. These properties will be introduced in the first subsection and recovery algorithms in the second.

1.2.3.1 Metrics

In order to be able to recover the information from the compressed samples, the acquisition method must satisfy some mathematical properties, presented hereafter. Note that random matrices are generally considered because they satisfy these properties with high probability.

1.2.3.1.a Coherence

A first requirement is that each measurement contains a part of the information, so as to maximize the entropy. Coherence measures a deviation from the orthogonality condition between the columns of the sensing matrix. This quasi orthogonality ensures that the input information is spread among all the measurements. Formally, the coherence [18, 19] of the matrix $\mathbf{\Phi\Psi}$ is the largest absolute Hermitian inner product between any two different unit

normalized columns:

$$\mu((\mathbf{\Phi\Psi})) = \max_{i \neq j} \left(\frac{| < (\mathbf{\Phi\Psi})_{.,i}, (\mathbf{\Phi\Psi})_{.,j} > |}{\|(\mathbf{\Phi\Psi})_{.,i}\| \, \|(\mathbf{\Phi\Psi})_{.,j}\|} \right) \tag{1.3}$$

where $(\mathbf{\Phi\Psi})_{.,i}$ is the ith column of $(\mathbf{\Phi\Psi})$.
Coherence should be as small as possible with a lower bound given by the Welch bound [20]:

$$\sqrt{\frac{N-M}{M(N-1)}} \leq \mu \leq 1 \tag{1.4}$$

To temper the rawness of the maximum extraction operator, other types of coherence that capture a more averaged behavior of the same notion have been introduced, e.g. average or cumulative coherence [21].

1.2.3.1.b Mutual coherence

Mutual coherence [5] describes the adequacy between the acquisition process and the domain of sparsity, by measuring the correlation between rows of $\mathbf{\Phi}$ and columns $\mathbf{\Psi}$:

$$\mu_m = \sqrt{N}. \max_{i,j} | <\mathbf{\Phi}_{i,.} \, | \, \mathbf{\Psi}_{.,j}> | \tag{1.5}$$

A low value of μ_m ensures that every measurement performed by the acquistion matrix $\mathbf{\Phi}$ carries a useful amount of information of all the non-zeros of $\boldsymbol{s}$ for any location because the sampling waveforms have an extremely dense representation in $\mathbf{\Psi}$ [22]. The smaller the mutual coherence, the larger the probability that any sample provides an information about the signal. It is bounded [23] by $[1; N]$. The mutual incoherence between time and frequency e.g. is maximal.

It should be noted that in the literature the mutual coherence may sometimes also be called coherence, which can be confusing. To clarify, in our terminology, mutual coherence considers the relationship between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ whereas coherence considers the relationships between the different measurement vectors of $\mathbf{\Phi\Psi}$.

1.2.3.1.c Isometry

In order to guarantee the recovery in Compressive Sampling, approximate preservation of norms and distances is a point of major interest.

- Restricted Isometry Property (RIP)
  The RIP measures the norm deformation induced by the projection.
  A matrix $\mathbf{\Phi}$ is said to satisfy the Restricted Isometry Property (RIP) [5, 24, 19] with parameters (K, δ_K), or of order K, if:

  $$(1-\delta_K) \, \|\boldsymbol{s}\|_2^2 \leq \|\mathbf{\Phi\Psi}\boldsymbol{s}\|_2^2 \leq (1+\delta_K) \, \|\boldsymbol{s}\|_2^2 \tag{1.6}$$

  for all K-sparse vectors $\boldsymbol{s} \in \mathbb{C}^N$.
  In others words, the $l2$-norm of K-sparse vectors is preserved up to a multiplying factor

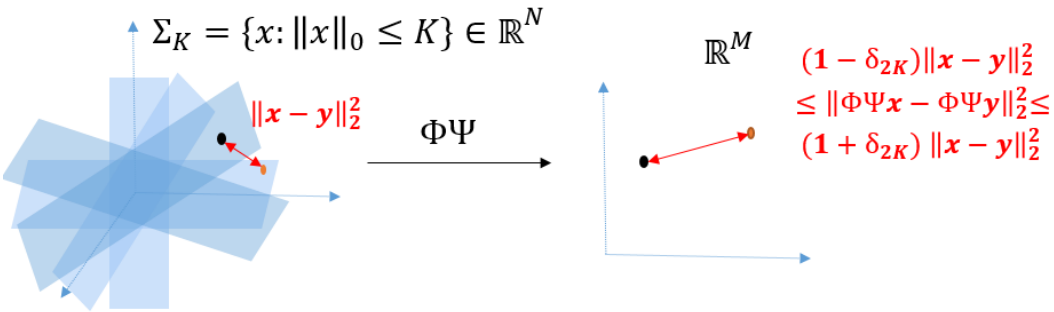


Figure 1.3: Distance preservation between K-sparse vectors through Compressive Sampling.

$(1 \pm \delta_K)$ by the projection. This leads to fundamental implications concerning robustness to noise [25, 26].
Eq. (1.6) ensures that K-sparse vectors do not lie in the nullspace of $\mathbf{\Phi\Psi}$ [27].
As noted in [28], the RIP is a convenient tool but not a necessary condition for any reconstruction algorithms. For example, a RIP-1 criteria (replacing $l2$ by $l1$ in the RIP definition) is more suitable for combinatorial recovery algorithms [29]. Furthermore, certifying the RIP is NP-hard hence it cannot be computed easily [30, 31].
If a matrix $\mathbf{\Phi}$ satisfies the RIP of order $2K$, it is easy to check [27] that eq. (1.6) is equivalent to saying that $\mathbf{\Phi}$ preserves the distance between any pair of K-sparse vectors up to the same multiplying factor $(1 \pm \delta_{2K})$, which will be referred to as Johnson-Lindenstrauss Lemma (JLL)-property in this work. As demonstrated in [32], RIP also implies preservation of angles.

- Johnson-Lindenstrauss Lemma (JLL)
  The Johnson-Lindenstrauss Lemma [33] guarantees the conservation of the pairwise distance from a cloud of points in a lower-dimensional embedding.
  In this work, a matrix $\mathbf{\Phi}$ is, perhaps abusively, said to satisfy the JLL property with parameters (K, δ_{2K}), illustrated on Figure 1.3, if:

$$(1 - \delta_{2K}) \|\mathbf{u} - \mathbf{v}\|_2^2 \leq \|\mathbf{\Phi\Psi u} - \mathbf{\Phi\Psi v}\|_2^2 \leq (1 + \delta_{2K}) \|\mathbf{u} - \mathbf{v}\|_2^2 \tag{1.7}$$

  for all K-sparse vectors $(\mathbf{u},\mathbf{v}) \in \mathbb{R}^{2xN}$.
  Hence two signals that were close in the input space will stay close in the output space and conversely, which is obviously essential with regard to classification tasks.

- Statistical Restricted Isometry Property (StaRIP)
  Introduced in [34], the Statistical RIP established itself as the non-uniform counterpart of the RIP, meaning that it is the probability that RIP is satisfied for one given K sparse vector:
  A matrix $\mathbf{\Phi}$ is said to satisfy the $(K,\ \delta,\ \epsilon)$-Statistical Restricted Isometry Property (StaRIP) if for K-sparse vectors $\boldsymbol{x}$, the inequalities (1.6) hold with probability exceeding $1 - \epsilon$ (with respect to a uniform distribution of the vectors among all sparse vectors in $\mathbb{R}^N$ with the same fixed magnitudes).

- ExRIP
  The Expected RIP (ExRIP) criterion introduced by [35] is a weaker variant of the StaRIP

where random distribution of non-zero values is additionally assumed. This assumption is reasonable and the advantage over StaRIP is that the ExRIP can be computed from three easily computable correlation-based criteria, summarized in Table 1.1.

Table 1.1: Definition of the coefficients necessary to ExRIP probability computation (real-valued matrices).

$\alpha = \frac{1}{(MN)^2}\sum_{i,k=1}^{M}(\mathbf{\Phi}_{k,.}\mathbf{\Phi}_{i,.}^H)^2$	Average correlation power of the rows
$\beta = \frac{1}{M^2N^3}\sum_{i,k=1}^{M}\lVert\mathbf{\Phi}_{k,.}\circ\mathbf{\Phi}_{i,.}\rVert$	Average cyclic auto- and cross-correlation power of the rows
$\gamma = \frac{1}{(MN)^2}\sum_{i,k=1}^{M}(\mathbf{\Phi}_{k,.}^H\mathbf{\Phi}_{i,.}^-)^2$	Average convolution power of the rows

where $\circ$ denotes circular convolution.

Then the ExRIP probability P is computed as:

$$P = 1 - \frac{(1-C_K)\rho(1+\alpha-2\beta)}{\delta_K^2} - \frac{(B_K-C_K)\rho(\gamma-\beta)+C_KM\beta-1}{\delta_K^2} \tag{1.8}$$

where $\rho = \frac{M}{M-1}$, $B_K = 1$ if non-zero entries are real-valued and $C_K = \frac{3K}{2K+K^2}$ if the entries are standard random variables.
The lower α, β, γ the higher the ExRIP probability.

1.2.3.1.d Gramian

The Gramian or Gram matrix of a matrix $\boldsymbol{M}$ is given by $\boldsymbol{M}^H\boldsymbol{M}$. Due to dimensionality reduction, the Gramian $(\mathbf{\Phi\Psi})^H\mathbf{\Phi\Psi}$ of a Compressive Sensing matrix is of maximum rank M smaller than its row length N. Since the Gramian is not full rank, some eigenvalues are zeros, which implies that neither norms nor orthogonality can be totally preserved by the projection through the sensing matrix $\mathbf{\Phi\Psi}$. The Gramian expresses this deformation of the scalar product by the acquisition process. As such it yields an accurate general overview of the geometric aspects of the projection. Note that for a unit column energy normalized sensing matrix, coherence and RIP can be deduced from it. The maximum of the Gramian without the diagonal is the coherence metric, which stems directly from both definitions, and the K-RIP property is given by extremal eigenvalues over all possible support $\mathbf{\Lambda}$ of cardinal K [19, 27], given the definition of eigenvalues.

1.2.3.1.e Combinatorial geometry

D. Donoho and J. Tanner explored high-dimensional combinatorial geometric considerations called neighboorliness which imply abrupt changes in the number of faces of convex polytopes. In [36] they propose to use a metric of distance from faces to the center on the polytope that is created by projection of the $l1$-ball with the sensing matrix. However, their geometrical conditions are restricted to $l1$ recovery, not tailored for noisy settings and less universal than RIP. In [37], these conditions are used to construct pathological vectors that will make recovery fails, in order to evaluate the RIP.

1.2.3.1.f Corresponding guarantees

Previous metrics imply various recovery guarantees, without and with additive noise, depending on the chosen reconstruction algorithm. For a Basis Pursuit reconstruction algorithm (see next Subsection), following results have been proven:

- According to [23]: If $\boldsymbol{\Phi}$ satisfies RIP with $(2K, \delta_{2K})$ and $0 < \delta_{2K} < \sqrt{2} - 1$, then any K-sparse vector $\boldsymbol{x}$ is perfectly reconstructed with overwhelming probability. Also the reconstruction error scales linearly with additive noise (C-stability), which is vital.
- According to [38]: If $K \leq 0.5(1 + \frac{1}{\mu})$ where μ is the coherence, then any K-sparse vector $\boldsymbol{x}$ is perfectly reconstructed.
- According to [23]: If M verifies $M > C\mu_m^2 K log(N)$ where C is a non-zero constant and μ_m is the mutual coherence, then the M measurements are sufficient to recover the input $\boldsymbol{x}$ with overwhelming probability. Note that this bound is relatively loose [35].

Hence the interest of a small coherence, mutual coherence and RIP constant is highlighted. For more mathematical details, a thorough overview of theoretical recovery guarantees in Compressive Sensing is for instance provided in [27].

Note that given guarantees are usually with respect to uniform recovery. As detailed in [39], the difference between uniform and non-uniform recovery is the following: the uniform recovery is defined by the high probability that after defining a matrix, every sparse vector can be reconstructed. On the other hand, non-uniform recovery is the high probability that after defining a matrix, one vector drawn among every possible can be reconstructed. In other words, it features a probability over the draw of a vector.

In the next part, algorithms able to perform signal reconstruction from compressed samples satisfying the above properties will be briefly presented.

1.2.3.2 Benchmark of non-adaptive recovery algorithms

For a signal acquired at the Nyquist rate the reconstruction of the signal is simply performed with 'sinc' interpolation, but for compressed samples, it is more complicated. Various techniques have been proposed to perform the reconstruction of the original signal from the compressed samples. The two most widely used algorithms are convex relaxation algorithms such as Basis Pursuit denoising (BPDN) and greedy algorithms such as Orthogonal Matching Pursuit (OMP) but many exists. The 6 main classes, as divided in [40], are:

- Convex relaxation algorithms
  Typical Algorithms: Basis Pursuit (BP), BP denoising (BPDN), Least absolute shrinkage and selection operator (LASSO), least angle regression (LARS).
  Principle: Recovering $\boldsymbol{x}$ from $\boldsymbol{y}$ and $\boldsymbol{\Phi}$ is an ill-posed problem since $\boldsymbol{\Phi}$ is singular.

Therefore, the solution is the sparsest $\boldsymbol{\Psi^{-1}u}$ among all possible solutions of $\boldsymbol{y} = \boldsymbol{\Phi u}$:

$$(P0) \qquad \hat{\boldsymbol{x}} = \underset{\boldsymbol{u}}{\operatorname{argmin}} ||\boldsymbol{\Psi^{-1}u}||_0 \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi u} \tag{1.9}$$

where $||.||_0$ denotes the $l0$ pseudo-norm i.e. the number of non-zero components. Because of the non convexity of $l0$, solving this problem $(P0)$ is computationally very difficult (NP-Hard, that is at least as difficult as any problem where one solution could be tested in polynomial time). The $l1$ regularization problem ($P1$ or Basis Pursuit), consists in replacing $l0$ by the norm $l1$:

$$(P1) \qquad \hat{\boldsymbol{x}} = \underset{\boldsymbol{u}}{\operatorname{argmin}} ||\boldsymbol{\Psi^{-1}u}||_1 \text{ s.t. } \boldsymbol{y} = \boldsymbol{\Phi u} \tag{1.10}$$

The Null Space Property [27] is a necessary and sufficient condition on having the same result with a relaxation of $l0$ by $l1$. Also, because $\boldsymbol{x}^+$ is generally noisy, the equality constraint is replaced by a fidelity term, leading to:

$$(P1, \lambda) \qquad \hat{\boldsymbol{x}} = \underset{\boldsymbol{u}}{\operatorname{argmin}} ||\boldsymbol{\Psi^{-1}u}||_1 + \lambda||\boldsymbol{y} - \boldsymbol{\Phi u}||_2 \tag{1.11}$$

where λ is a constant balancing the sparsity and fidelity constraints.

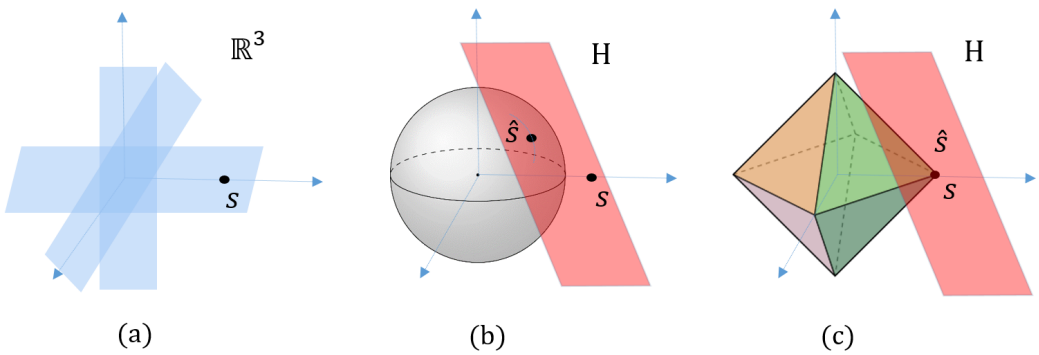


Figure 1.4: Sparsity and different norm minimization, inspired from [41].

Figure 1.4 helps to visualize on a simplified example why, in contrast to the $l2$ norm, the $l1$ norm favors sparsity and is able replace the $l0$ pseudo-norm problem with accuracy. The set of all K-sparse vectors in $\mathbb{R}^3$, depicted in (a), is a non-linear space consisting of all K-dimensional hyperplanes aligned with the coordinate axes. The hyperplane H generated by $s+Ker(\Phi\Psi)$ where $Ker(.)$ denotes the kernel, is depicted in red. The point of H with minimal $l2$-norm, $\hat{\boldsymbol{s}}$, corresponds then to the contact point between H and a dilated $l2$ ball centered at the origin. The $l2$-ball being isotropic, the probability that $\hat{\boldsymbol{s}}$ is sparse is very low, as shown in (b). On the contrary, the $l1$ ball is pointy, meaning its points are aligned along coordinate axes. Hence the first contact will appear near the coordinate axes and be sparse with high probability, as shown in (c).
Pro/cons: Too slow.

- Greedy Iterative Algorithms
  Typical Algorithms: Matching pursuit (MP), Orthogonal Matching Pursuit (OMP).

Principle: Matching Pursuit selects iteratively each column of the dictionary that correlates most with $\boldsymbol{y}$, minimizing the least square error (fidelity term). OMP [18] adds a Gram-Schmidt pre-process before MP to compensate the issues raised by the possible non-orthonormality of the dictionary.
Pro/cons: Low implementation cost, high speed of recovery, poor performances if modulations or phase noise [42].

- Iterative Thresholding Algorithms
  Typical Algorithms: Belief propagation, Expander matching pursuits.
  Principle: Iterative thresholding methods, including message passing
  Pro/cons: High computational cost for low sparsity, slow.

- Bregman Iterative Algorithms
  Principle: Iterative solving of a sequence of unconstrained problems.
  Pro/cons: Potentially unsteady but fast.

- Combinatorial/Sublinear Algorithms
  Typical Algorithms: Fourier sampling, Chaining pursuits.
  Principle: Based on group testing.

- Non Convex Minimization Algorithms
  Typical Algorithms: Focal underdetermined system solution (FOCUSS), Iteratively Reweighted Least Squares (IRLS).
  Principle: Relaxation with lp where $p \leq 1$.

A main highlight is that reconstruction with any of these algorithms will be extremely expensive from a computational point of view. This also means that real-time processing will be hard to achieve. A comparison of required number of measurements and complexity is provided as an example in Table 1.2 for a few algorithms.

Table 1.2: Comparisons between reconstruction algorithms, from [34], α is a non-zero constant.

Approaches	Number of Measurements	Complexity
Basis Pursuit	$K\log(\frac{N}{K})$	N^3
Orthogonal Matching Pursuit	$K\log^{\alpha}(N)$	$K^2\log^{\alpha}(N)$
Group testing	$K\log^{\alpha}(N)$	$K\log^{\alpha}(N)$
Greedy Expander Recovery	$K\log(\frac{N}{K})$	$N\log(\frac{N}{K})$

1.2.3.3 Adaptive approaches

Adaptive approaches have been considered to improve performances by conciliating the imperative of robustness to changing noise level and degree of sparsity K with the opportunity of savings through scalability. Scalability in our context means for instance adjusting the energy consumption by activating a given number of branches according to the level of noise.

Or in [43], adaptive thresholding and adaptive time segmentation adjust the signal detection capability of the Quadrature-Analog-to-Information Converter (QAIC, see 2.3.2) to the sparsity level of the signal. Reconfigurability, on the other hand means that at same energy consumption, the parameters of the architecture are optimized to yield better reconstruction and estimation performances.

In Bayesian adaptive methods [44, 45], the coefficients of the sensing matrix (for the Modulated Wideband Converter, MWC, see 2.3.2, that would be the mixing codes) are computed based on the available information on the current spectrum. However, achieving real-time is already an issue in Compressive Sensing, so the additional burden of computation may not be manageable or worth it [44]. Moreover such reconfigurability possibilities also implies more complicated hardware.

In his PhD work, D. Adams [46] proposes to use an MWC to perform adaptive interference canceling. Interferers are a notorious problem in Compressive Spectrum Sensing and solving this issue enhances greatly the dynamic of the system. When an interferer is detected at a given frequency, the Fourier coefficient corresponding to this frequency is set to zero in order to null his contribution in the folded spectrum. In [28], an interferer canceling method for Radar parameters extraction with the Random Modulation Pre-Integrator (RMPI), see 2.3.1, is also suggested. Once an undesirable frequency is identified, its contribution in form of the corresponding Discrete Prolate Spheroidal Sequence component is subtracted from the measurement. In [47], Ultra Wideband (UWB) signal specificities are e.g. taken into account to modify the sensing matrix energy distribution in order to mitigate the narrowband interferences.

The reconstruction of the original signal (complete reconstruction) has the drawback of being highly energy-consuming and the question must be raised whether or not the benefits at the acquisition stage preveil on the reconstruction cost in the context of this study, as it depends largely from the application. Benchmarking fairly and thoroughly Compressive Sampling and Nyquist-rate approaches for radiofrequency receivers regarding all the implementations aspects (consumption, resolution, dynamic range, die area, etc.) is a difficult task, that has been tackled but not fully completed yet. Encouraging analyses have been lead in [48], where a model-based approach showed that at lower gain requirements and low to moderate resolutions (4-6 Effective Number Of Bits), the Random Modulation Pre-Integrator (see 2.3.1) has the potential to be between 2 and 10 times more power-efficient than high-speed ADCs.

However, on the contrary, in [49] Compressive Sensing followed by reconstruction is proposed and an order of magnitude of several Watts was found (while current ultra-low power wake-up receivers achieve about $10\mu W$ consumption [50]). This seems not suitable for the purpose of Cognitive Radio with embedded solutions. That is why, in order to reduce the burden of information extraction, in this PhD manuscript the choice is made to perform partial restitution of features of interest rather than exact reconstruction of the original signal followed with parametric estimation.

An oversimplified and naive metaphor of the approach would be the following: Suppose having an equal-arm balance, a calibration weight and 9 numbered gold bars, from which it is known that no more than one is rigged, i.e. it weights less that it ought to.

A matching with the Nyquist-rate approach would be to compare each one of the gold

bars to the calibration weight. But obviously, two measurements (measuring 2 sets of 3 gold bars e.g. row- and column-wise, $1 + 2 + 3$ versus $4 + 5 + 6$ and $1 + 4 + 7$ versus $2 + 5 + 8$) would allow, with a bit of deduction, to find out if one bar is rigged and which one it is, even in the worst case. The idea that will be exploited in this work is similar, in a certain extent. This aspect will be developed in the next Section.

1.3 Information retrieval from the compressed samples

Subsection 1.2.3 has highlighted that the complete reconstruction of the signal is a non-linear optimization process, highly energy-consuming and hence barely compatible with the requirements of embedded radio solutions. Yet reconstruction is not always necessary and a promising alternative to this drawback is to perform the extraction of the sole features of interest from the compressed samples directly. The concept, called Analog-to-Feature extraction [51], consists in extracting features of interest while bypassing the Analog-to-Digital conversion (and hence the Nyquist frequency criterion). These features can be of interest for themselves or can be used further for classification purpose. Classification tasks consist in assigning the observation to a class among a set of few classes and detection tasks can be seen as a classification between two classes: presence or absence of the signal. Due to its low computational cost, this approach is relevant for various applications. The authors of [52] introduced in particular the potential for Cognitive Radios (CRs). However, the compressive acquisition has scrambled along the structure of the signal information, which renders information retrieval more difficult.

In this Section, the State-of-the-Art of information retrieval directly from the raw compressed samples is outlined. First, the parameters of interest need to be defined in Subsection 1.3.1. A second part tackles guarantees and challenges at stakes, first for estimation and then for classification. Finally, examples of existing solutions, mostly derived from solutions at the Nyquist rate, are given and their benefits and drawbacks discussed.

1.3.1 Parameters of interest

For Cognitive Radio, a certain number of features must be estimated to ensure that the radio system is able to avoid interferences with other users. Mainly, it consists in detecting if a frequency band is occupied or not. In addition to this, other applications of interest can be considered. For instance, smart RF receivers that are enhanced via reconfigurability thanks to environment sensing. That would for instance be the estimation of Peak-to-Average-Power-Ratio (PAPR), in order to adjust the dynamic of the low-rate ADC or the detection of high power interferers, in order to possibly circumvent them adaptively. Detection of interferers is all the more important in Compressive Sensing because of the noise folding phenomenon. Also environment identification (home, indoor, outdoor, car, etc.) for specific adjustments is of potential interest.

Among others, following parameters regarding useful signal or interferences are concerned:

- Central frequency, bandwidth, power level, Peak-to-Average-Power-Ratio (PAPR)
- Transmission times, direction-of-arrival and localization
- Modulation (AM/FM, size of the phase constellation, etc.) and protocol (WiFi, Bluetooth, etc.)

1.3.2 Bounds and guarantees for Analog-to-Feature extraction

First and foremost, it is necessary to evaluate which performances can be expected from an Analog-to-Feature extraction approach and in which extent they are influenced by the settings, in order to see if this strategy is viable. For estimation performances, this question will be tackled through the prism of the Cramér-Rao lower bound on the estimation variance for any unbiased estimator. For classification, several aspects will be addressed: distance preservation, feature enhancement and dictionary learning.

1.3.2.1 Compressive parametric estimation: Cramér-Rao bounds

The Cramér-Rao bound is the lower bound on the estimation variance for any unbiased estimator. It is therefore a convenient performance evaluation tool [53] to understand the mechanisms at stake in estimation from the compressed samples. The Cramér-Rao bound (CRB) is computed from the diagonal of the inverse of the Fisher information matrix, a matrix which quantifies the amount of information about the parameters that is carried by a signal observation. The reader is referred to Appendix A for more thorough details and equations.

An interesting bounding of the Fisher matrix for Compressive Sensing has been established in [54, 55], by approaching the limit in the distance preservation property ($2K$-RIP) with an infinitesimal parameter variation. Thus norm preservation is extended to the vector of partial derivative toward the estimated parameters. And since the Fisher matrix is by construction the Gramian of this partial derivative vector, the authors obtained following inequalities:

$$\frac{M}{N}(1-\delta_{2K})^2\boldsymbol{J_{Nyq}} \leq \boldsymbol{J_{CS}} \leq \frac{M}{N}(1+\delta_{2K})^2\boldsymbol{J_{Nyq}} \tag{1.12}$$

where $\boldsymbol{J_{Nyq}}$ and $\boldsymbol{J_{CS}}$ are the Fisher information matrices given the Nyquist or the compressed samples respectively and $\leq$ for matrices is understood as a positive definite difference.
$\frac{M}{N}$ corresponds to the noise folding aspect, which will be tackled in 2.1.4.
$(1 \pm \delta_{2K})^2$ corresponds to the unavoidable anisometry of the projection (RIP), which was discussed in 1.2.3.1.c. Some more specific analytics were derived for parametric model estimation with random distributions, for example the mean of a complex distribution for direction-of-arrival estimation in [56].

1.3.2.2 Compressive classification

Classification tasks output a prediction taking values in a discrete phase space. Thus it allows generally for more flexible approaches than estimation.

1.3.2.2.a Distance preservation, ellipse problem and separable classes

Distance preservation properties guarantee that classes of K-sparse signals that are separated in the N-dimensional input space stay approximately separated after projection in a M-dimensional subspace through the Analog-to-Information-Converter. For a class well-represented by a typical point, the problem has been tackled under the term of smashed filtering (cf 1.3.3). In the case where the classes are ellipsoids, this is called the Rare Eclipse problem and bounds on the sufficient number of measurements have been given for random matrices (cf [57]). But classes are in principle not separable in the input space and then they will not be in the output space either.

1.3.2.2.b Feature enhancing

Signal-agnostic approaches present a certain number of advantages. However, for a well-defined classification problem, better performances could ideally be achieved through a denoising anisotropic projection that focuses on preserving the discriminative information. It is acceptable to loose definitely information on the signal as long as this information is not discriminative. Distance preservation implies the preservation of separability but the reverse is not true. This approach, called feature enhancing [58], has for instance been successfully used for the specific task of voice activity detection [51]. In Figure 1.5, different subsampling methods are compared in terms of physical information bandwidth during the acquisition and restitution process. While Compressive Sensing and Innovation Rate Sampling allow to reconstruct all the information, Feature Extraction subsamples the signal information rate leading to information loss. The processing involved can be linear or not.
This approach is very dependent on the geometry of the problem and hence on the application.

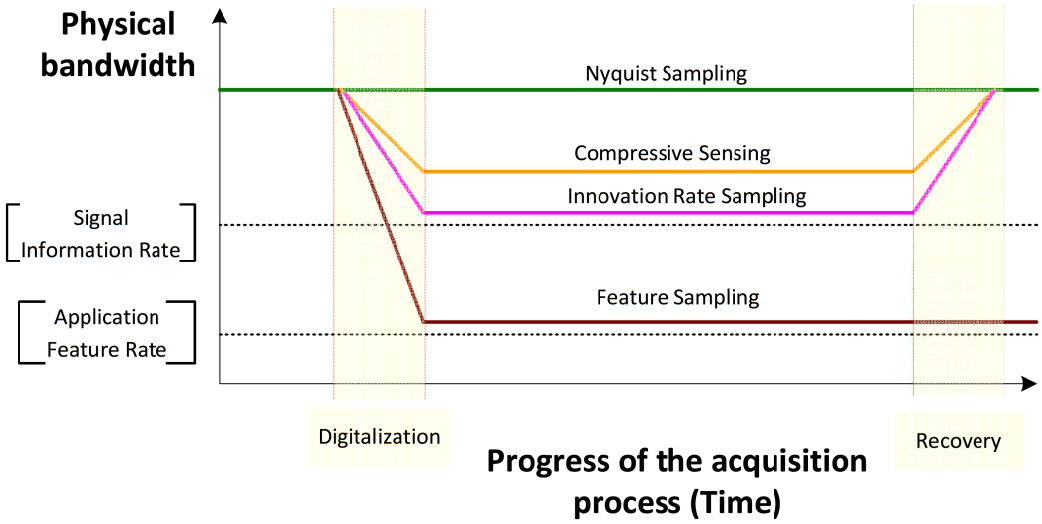


Figure 1.5: Comparison of different subsampling methods in terms of information rate evolution during the acquisition process, inspired from [51].

1.3.2.2.c Dictionary decomposition

In order to perform classification (discrete classes) or regression (continuous values), one possible technique is to use a decomposition on a dictionary.

For regression, one would for instance build a dictionary per class, with a matched filter for each type: white noise, useful signal, interferers, etc. But the problem is that orthogonality between classes is not guaranteed. For classification, the goal is to build a decomposition dictionary with a double constraint: high discriminative power (allowing accurate classification) and on which the input signal is nearly sparse. Discriminative power means that the cost function depends on the accuracy of classification on training data. Examples from the State-of-the-Art are the following:

- Sparse Representation based Classification (SRC)[59]: a shared dictionary, minimizes reconstruction error on the different possible classes.
- Fisher Discriminative Dictionary Learning (FDDL)[60]: one dictionary per class, discriminative criteria based on Fisher criterion.
- Label Consistent K-SVD (K-Singular Value Decomposition)[61]: classifier and dictionary are jointly learned with a K-SVD algorithm.
- Discriminative Bayesian dictionary learning [62]: based on the probabilities of selection of the dictionary atoms in the expansion of data from each class.

It appears that the current bounds given for estimation and classification by means of compressed samples are essentially based on distance preservation properties. These bounds enable to understand the broad lines of the problem. However, it is not entirely sufficient in order to specify precisely a compressive receiver architecture.

1.3.3 Existing solutions for compressive information retrieval

Various methods have been suggested for classification, detection and parametric estimation by means of compressed samples, often based on the adaptation of existing techniques for samples at the Nyquist rate. The most relevant strategies are mentioned non exhaustively below and summarized in a conclusion table.

- **Matching filters** propose an optimization of the architecture with respect to the expected classes. The principle is to correlate the input signal with its expected model. It is optimal in additive white Gaussian noise but requires knowledge of the signal model. The authors of [63] propose a universal CS matched filtering scheme able to detect any sparse signal, which approximates the matching signal pattern.
  In [64], an implementation of a compressive matching filter, called a 'Smashed filter', is proposed for image classification purposes.

Other exposed strategies rely on a specific estimation or classification algorithm which has been tailored to work on a reduced number of samples acquired by a compressive receiver.

- **Energy detection** is the simplest spectrum detection technique for samples at the Nyquist rate. The decision is taken based on the comparison between the amount of energy received in a frequency band and a given threshold.
  Energy detection extended to the sampled compressed observations has been formalized by [65]. However, as noted in [66] and similarly to its counterpart at the Nyquist rate, it is critically sensitive to noise.
- The principle of **cyclic detection** is to detect cyclostationarity features provoked by the modulation of the signal with sine wave carriers or repeating codes. It is more complex than energy detection but also more robust to noise.
  The authors of [66] propose a method for cyclostationarity detection from the compressed samples of a Periodic Non Uniform Sampler (PNUS) or a Modulated Wideband Converter (MWC).
  In [67], it is proposed to estimate spectrum occupancy from compressive samples through the recovery of the sparse $2D$ cyclic spectrum, the Fourier transform of the cyclic covariance with respect to time-lag and cyclic frequency index. Time-varying cross-correlations of compressive measurements are linked to cyclic statistics, which permits to solve the $2D$ cyclic spectrum via $l1$-norm minimization. Two techniques can be used to estimate the spectrum occupancy from the cyclic spectrum: a band-by-band Generalized Likelihood Ratio Test detector or fast thresholding for signals with known modulation such as BPSK.

- **MUltiple SIgnal Classification (MUSIC)** is a famous algorithm introduced by [68] which consists in estimating the frequencies contained in a signal by using an eigenspace decomposition method. The input signal is modeled as a sum of K complex exponentials and Gaussian white noise. Given the autocorrelation matrix and assuming the value of K, the space is decomposed into two subspaces, signal and noise.
  A counterpart based on compressive measurements, Compressive MUSIC, is presented in [69]. Compressive MUSIC identifies one part of the support using a conventional CS algorithm, then estimates the other part using a generalized MUSIC criterion.

- **Power Spectrum Estimation**. In [66] it is proved that the minimal sampling rate for perfect power spectrum reconstruction, i.e. with loss of the phase information, is half the rate that allows for perfect signal reconstruction. Minimal sampling rate refers to the lowest rate enabling perfect reconstruction of the power spectrum in a noiseless environment for a general sampling scheme. But the authors of [70] proved that e.g. designing the Periodic Non Uniform Sampling (PNUS, see 2.2.1) sampling matrix according to the minimal sparse ruler pattern results in a sampling rate below.

- **Automatic Modulation Recognition** Different PSK modulations have different numbers of peaks in the spectrum after raising to diverse N^{th} power. That is because the finite number of symbol phases are transformed in a constant phase through the process. These features can be used for Automatic Modulation Recognition (AMR), as well as carrier frequency and symbol rate rough estimations (except for 8PSK). The authors of [71] show that a link can be drawn between the compressed samples and the Nth power spectrum of the original signal, enabling AMR directly with compressed samples.

It would seem that the proposed solutions do not fully answer the problematic. Matched filtering requires a prior that might not be available and besides, implementing the sensing matrix solution of the problem might be expensive. Compressive MUSIC and cyclostationarity detection have a high computational cost, which should preferably be avoided. Automatic Modulation Recognition is not flexible and generic enough to be used in practice. Power Spectrum estimation has ultimately the same stakes as reconstruction, hence it does not significantly change the balance.

A quick summary table is established in Table 1.3

Table 1.3: Summary table of estimation methods from the compressed samples.

Methods	Advantages	Disadvantages
Matched filtering	Optimal performances Low computational cost	Requires prior of the primary user
Energy detection	Does not require prior Low computational cost	Poor performance for low SNR Cannot differentiate users
Cyclostationarity feature	Valid in low SNR region Robust against interference	Requires partial prior High computational cost
MUSIC	Superresolution	High computation and storage cost Knowledge of K required
Power spectrum	Less measures required	Phase information loss
AMR	Energy efficient	Limited to some modulation

1.4 Conclusion

In this chapter, the challenges at stake in the emergence of Cognitive Radio for Dynamic Spectrum Access have been introduced. The interest of developing radiofrequency receivers able to handle a multi-standard spectral environment under stringent power and cost constraints has been highlighted.

As a solution to meet these constraints, the theory of Compressive Sensing has been put forward to overcome the high-rate ADC deadlock. Compressive Sensing claims that a reduced number of incoherent measures are sufficient to recover the signal under the additional constraint of sparsity. If the information content is appropriately spread among the measurements, solving the resulting underdetermined problem becomes possible. Hence this approach would enable to bypass the Nyquist sampling rate limit by focusing instead on the gathered frequency bandwidth quantity.

To begin with, a short background historic has been given. Then the two pillars that render reconstruction of the original signal in Compressive Sensing possible have been described: conditions on the sensing matrix modeling the acquisition process on one hand and recovery algorithms on the other hand. First, evaluation metrics of the sensing matrix, in particular norm preservation and coherence which is the maximum correlation between columns, have been presented. Next, the various mechanisms of recovery algorithms has been tackled and the principle of the two most famous ones exposed: Orthogonal Matching Pursuit, based on greedy search of the column of the dictionary that correlates the most with the observation and Basis Pursuit, based on the problem relaxation to $l1$-minimization. The benchmark has pointed out that the wonders of Compressive Sensing regarding reduction of acquisition and storage costs do however come at a price. Given the computational complexity of the recovery algorithms, the option of reconstructing the original signal should preferably be discarded if an embedded solution is targeted.

Hence a focus has then been made on the current State-of-the-Art regarding partial information retrieval, limited to what is called features or characteristics, directly from compressed samples. First, parameters of interest for Cognitive Radio applications and reconfigurability have been defined. Then some guarantee bounds were given for estimation and classification tasks and it appeared that they are mostly based on distance preservation. The principle of feature extraction with loss of information was also introduced. Finally, existing solutions inspired from estimation at the Nyquist rate were presented and their limitations exposed.

From this review, it emerged that some key elements lack in order to establish a general methodology for the design of an efficient Analog-to-Feature Converter with respect to spectrum sensing tasks. To yield additional answers, two aspects must be tackled. On one hand, the implementation challenges of an efficient compressive radiofrequency receiver, whose key considerations will be outlined in Chapter 2 and explored further in Chapter 3. On the other hand, it is necessary to bring to light the fundamental specificities of parametric estimation with compressed samples, as opposed to estimation from the reconstruction of the original signal.

That is why, in the next chapter, the notions presented above are put into practice: In Chapter 2 a large variety of radiofrequency receivers based on Compressive Sampling will be presented, discussed and compared to each others. Based on the conclusions of this analysis, the role of mixing codes is investigated further in Chapter 3, focusing on one architecture, the Modulated Wideband Converter (MWC) architecture. Note that since reconstruction aspects are better understood and developed in the literature than feature extraction, reconstruction is considered as a first step to understand the challenges in Chapter 2 and 3. Specificities of parametric estimation are then explored in Chapter 4 through original contributions on the Cramér-Rao bound under specific assumptions.

Chapter 2

Compressive radiofrequency receiver architectures

Contents

2.1 Framework **41**
2.1.1 Considerations on the input signal model 42
2.1.2 Notations 43
2.1.3 Multiband model sensing matrices equivalence 44
2.1.4 Modelling of the input noise after compression 46
2.1.5 Calibration 46
2.1.6 Quantization 46
2.2 Non Uniform Sampling **47**
2.2.1 Non Uniform Sampler (NUS) 47
2.2.2 Non Uniform Wavelet Bandpass Sampling (NUWBS) 49
2.3 Random Demodulation **52**
2.3.1 Random Demodulator (RD) 52
2.3.2 Modulated Wideband Converter (MWC) 54
2.3.3 Compressive Sensing Filter (CSF) 56
2.4 Variable Rate Sampling **57**
2.4.1 Multirate Sampling (MRS) 57
2.4.2 Nyquist Folding Receiver (NYFR) 58
2.5 Conclusion **61**

Applied to the field of radiofrequency receivers, the Compressive Sensing theory introduced in Chapter 1 offers the promise to replace the energy-inefficient high-rate Analog-to-Digital Converters (ADCs) by architectures able to monitor large frequency bandwidths with a low sampling rate. A wide range of compressive acquisition architectures, also called Analog-to-Information Converters (AICs), have been developed for radiofrequency (RF) receivers in the last years in order to deliver this promise. Before the investigations on parametric estimation that will be presented in Chapter 4, it is vital and necessary to benchmark and fully understand the challenges of existing compressive receivers, tailored to reconstruction. Each of those AICs uses different mechanisms, has various drawbacks and dedicated applications as target. However, as they all share the purpose of reducing the sampling rate beyond the Nyquist rate, they all suffer from the phenomenon of spectrum aliasing. To solve the ambiguity and be able to discriminate different folded inputs, compressed radioreceivers need to control the spectrum folding operations by tagging the original signals in a certain way. The transformation of the signal must allow for both identification of the original frequency and reconstruction of the original signal. Various techniques have been investigated, including exploitation of time, phase, frequency or coding diversity. Among those, three main categories emerge, which are pictured in Figure 2.1:

- **Non Uniform Sampling** which is based on the incoherence between the sparse domain and the subsampling of the acquisition domain.
- **Random Demodulation** which consists in the mixing with pseudo-random sequences followed by low-pass filtering, performing a modulation/demodulation coding of the frequency bands.
- **Variable Rate Sampling** which can be viewed as generalized bandpass sampling (subsampling). It relies on the fact that if two signals are aliased for one undersampling frequency, they might not be for another.

Hence a first meaningful manner to classify architectures is by the different ways in which they create the diversity between the measurements.

The chapter is organized as follows. First of all, the framework related to compressive radiofrequency receiver architectures needs to be presented in Section 2.1. In particular the input signal models and associated notations will be introduced in Subsection 2.1.1 and 2.1.2 respectively. Then the original concept of multiband architectures is introduced in Subsection 2.1.3 in order to interpret in a generic way properties that are shared by some of the architectures. Next, the question of the modeling of the noise on the compressed data will be discussed in Subsection 2.1.4. The calibration problematic will be quickly tackled in Subsection 2.1.5, and the quantization problem mentionned in Subsection 2.1.6.

After these first considerations, the principal compressive radiofrequency receiver solutions are presented and reviewed: Non Uniform Sampling solutions are presented in Section 2.2, comprising Non Uniform Sampler (NUS) in Subsection 2.2.1 and the recent Non Uniform Wavelet Bandpass Sampling (NUWBS) in Subsection 2.2.2. Architectures based on Random Demodulation are presented in Section 2.3, including the Random Demodulator (RD) in Subsection 2.3.1, the popular Modulated Wideband Converter (MWC) in Subsection 2.3.2, and

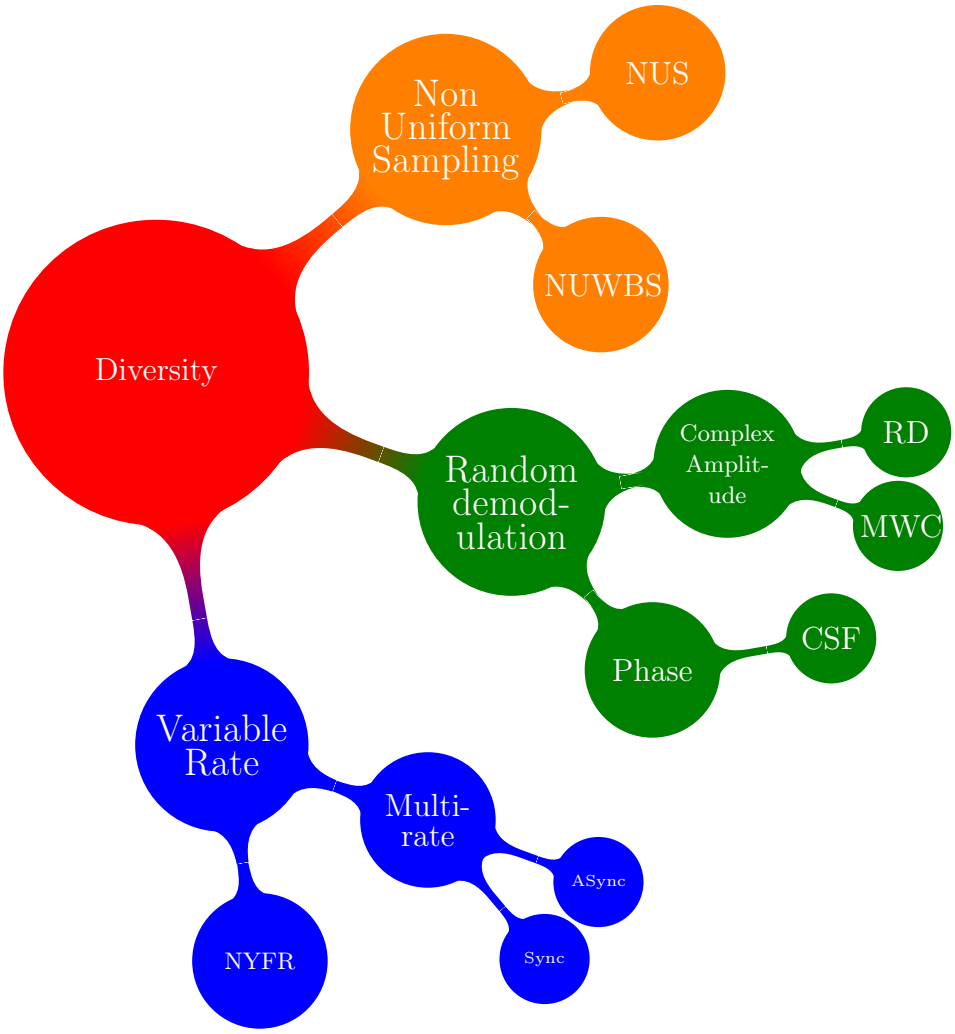


Figure 2.1: Classification of compressive radiofrequency architectures according to the way in which diversity is created.

the Compressive Sensing Filter (CSF) in Subsection 2.3.3. Variable rate sampling is tackled in Section 2.4, through the examples of synchronous and asynchronous Multirate Sampling (MRS) in Subsection 2.4.1, and the Nyquist Folding Receiver (NYFR) with his continuously variating rate in Subsection 2.4.2.

A general conclusion is then drawn in Section 2.5.

2.1 Framework

First, a framework for Analog-to-Information Converters is established. To begin with, different ways of modeling the input signal are compared and related notations in the time and frequency domains are introduced. Next, considerations on the compressed noise and then on the calibration and quantization issues are discussed.

2.1.1 Considerations on the input signal model

Compressive Sensing provides a theory for objects that are approximately sparse in a given representation space, which can be a basis or a dictionary. A compressive RF architecture specifies signal acquisition operations. These actions can be described by a sensing matrix operating on an input signal to produce output samples. Hence defining the sensing matrix (through the architecture) implicitly defines simultaneously a representation of the input (and output) signal. In order to guarantee success, the way in which the AIC represents the input signal must provide sparsity.

As mentioned earlier in Chapter 1, in the context of cognitive radio, the RF spectrum is assumed to be sparsely populated. Hence a first intuition about the input signal model might be to consider harmonic tones on a Nyquist grid as in [72]. However, alternative descriptions of the RF signal are also suitable, with benefits and drawbacks to be considered. The Union of Subspaces (UoS) model, developed within the Xampling framework by Y. Eldar [1], enables to gather under the same framework the different implied input models. In UoS, the signal belongs to a set $\cup_{\lambda\in\Lambda} A_\lambda$, where $card(\Lambda)$ and $dim(A_\lambda)$ may be finite or infinite and $\cup$ stands for union. The two most popular descriptions are the decomposition on a dictionary and the finite union of bandpass signals. Other models, see details in [1], could also be successfully taken into consideration. For example models based on Finite-Rate-of-Innovation [73], which assume a limited number of degrees of freedom per unit of time, are well-suited for the narrow-band pulses of Ultra-Wideband (UWB) [74, 75]. Manifolds [76] appear as another interesting low-dimensionnal model alternative.

One possibility is to decompose RF signals on a dictionary such as multitones (as in the Random Demodulator) or wavelets (as in the Non Uniform Wavelet Bandpass Sampling). One issue lies in the quickly prohibitory size of the dictionary in absence of additional prior on the structure of the RF signals. The other issue is basis mismatch, meaning that assumed and actual sparsity basis differ [77]. Since frequencies are rarely bin-centered on a frequency grid, they generally do not decompose themselves adequately on the grid given by the Discrete Fourier Transform matrix, a phenomenon known as spectral leakage. One notable benefit, however, is that it enables to treat a wider range of signals, for instance Radar pulses that are sparse in both time and frequency [28].

Another way to treat the input signal is as a finite union of bandpass signals, each active subband containing signal with analog representation. Union of bandpass signals model assumes that the signal is bandlimited, has support on no more than N disjoint intervals with bandwidths each smaller than a given B. This description enables to keep the analog formulation of the input signal, avoiding some pitfalls of the dictionary model [1].

In [78] a clever alternative is proposed which addresses the time/frequency limitation problem by using the Slepian basis formed by Discrete Prolate Spheroidal Sequences (DPSS) as the sparsity representation basis. Slepian bases are the solution of the spectral concentration problem, i.e. they are the time- and bandlimited representation that is the most faithful to the original signal in terms of energy. Whereas the authors of [79] tried to decompose the whole signal using one family leading to huge dimensionality, in [80] each band of the signal is described by the non-modulated Slepian families. The decomposition of multiband signals

on DPSS is very sparse, i.e. few coefficients are non-negligible. Simulation results showed an improvement between $35dB$ (noiseless) and $13dB$ (noise) in the recovery SNR with the adequate interpolation recovery algorithm, compared to the same MWC considering a sparsity in the frequency domain. The recovery SNR is defined as the ratio between the power of the original signal and of the reconstruction error.

Notations, suitable for both models, are set in the next subsection.

2.1.2 Notations

A summary of notations and symbols introduced below is given in Figure 2.2.

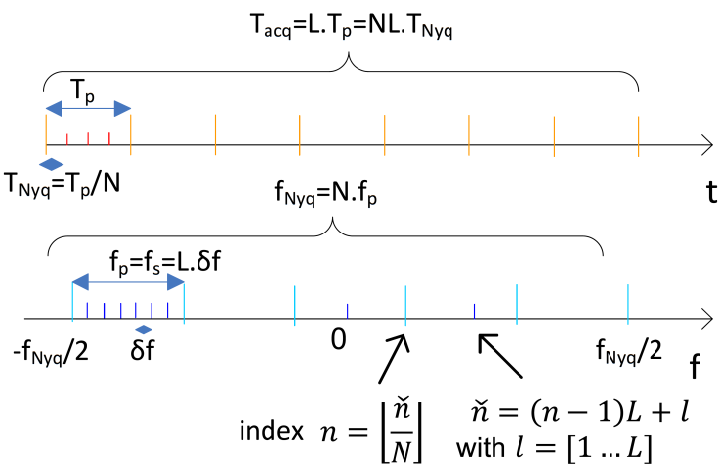


Figure 2.2: Time and frequency notations.

In this study, the number of active frequency subbands K will be assumed to satisfy $K \ll N$ and will be referred to as the sparsity degree. The frequency support of the complex-valued $\boldsymbol{x}$ is given by:

$$\boldsymbol{\Lambda} = \{n \in [\![1; N]\!] | \ \exists (n, l) \text{ satisfying } |\boldsymbol{s}[\check{n}]| > 0\}$$

$\boldsymbol{\Lambda}$ is therefore the set of indices corresponding to non-empty frequency subbands. It will be considered that the Nyquist band $[-\frac{f_{Nyq}}{2}; \frac{f_{Nyq}}{2}]$ is split into N subbands so that $f_{Nyq} = N.f_p$. The other possibility would be to consider a real-valued signal $x(t)$ with support $[0; f_{Nyq}]$.

The previously presented framework that considers input and output of respective dimensions N and M will be sometimes slightly modified for reasons that will shortly appear. Presently, the acquisition length corresponding to the time T_{acq} is set to be NL where N denotes the number of considered subbands and L is the number of bins per subband. The frequency resolution is thus given by $\delta f = \frac{f_{Nyq}}{NL}$. It will be useful to introduce the global frequency index $\check{n} = (n-1)L + l$, to index the bin l in the frequency subband n. The output

dimension is ML where M is typically, but not always, the number of channels. Similarly the index $\check{m} = (m-1)L + l$ for the lth bin in the mth branch is defined. Considering both the relationship between $\check{n}$, n, $\check{m}$, m and l and the indexation of $[-\frac{f_{Nyq}}{2}; \frac{f_{Nyq}}{2}]$ from 1 to NL, it is possible to write the discretization of the input and output signals as:

$$\tilde{x}[\check{n}] = \tilde{x}((n - \lfloor \frac{N}{2} \rfloor - \frac{3}{2})f_p + (l-1)\delta f) \tag{2.1}$$

$$\tilde{y}[\check{m}] = \tilde{y}_m((l-1)\delta f - \frac{f_p}{2}) \tag{2.2}$$

Note that depending on different versions of the architecture, and in particular a possible parallelization or serialization of the branches, the meaning of indices may differ a little.

2.1.3 Multiband model sensing matrices equivalence

In this work, the formalism of compressive multiband architecture is introduced, in order to exploit in a generic way specific properties that are shared by various architectures. Multiband architectures will hence refer to devices that process identically all frequency bins within the same subband. This is a particular subgroup within architectures defined with a union of bandpass signals model. Multiband architectures include the Periodic Non Uniform Sampler, the Modulated Wideband Converter, Random Convolution, Multirate Sampling but not the Non Uniform Wavelet Bandpass Sampling, the Nyquist Folding Receiver or the Random Demodulator. Then instead of considering the architecture-agnostic sensing matrix $\boldsymbol{\Phi\Psi}$, the acquisition process operations between the input and output spectra can be described for multiband architectures by a $M.L\text{x}N.L$ block matrix $\check{\boldsymbol{B}}$.
$\check{\boldsymbol{B}}$, illustrated in Figure 2.3a, is composed of subblocks formed by the identity matrix of size $L\text{x}L$ weighted by a coefficient $b_{m,n}$ for $(m,n) \in [\![1;M]\!]\text{x}[\![1;N]\!]$:

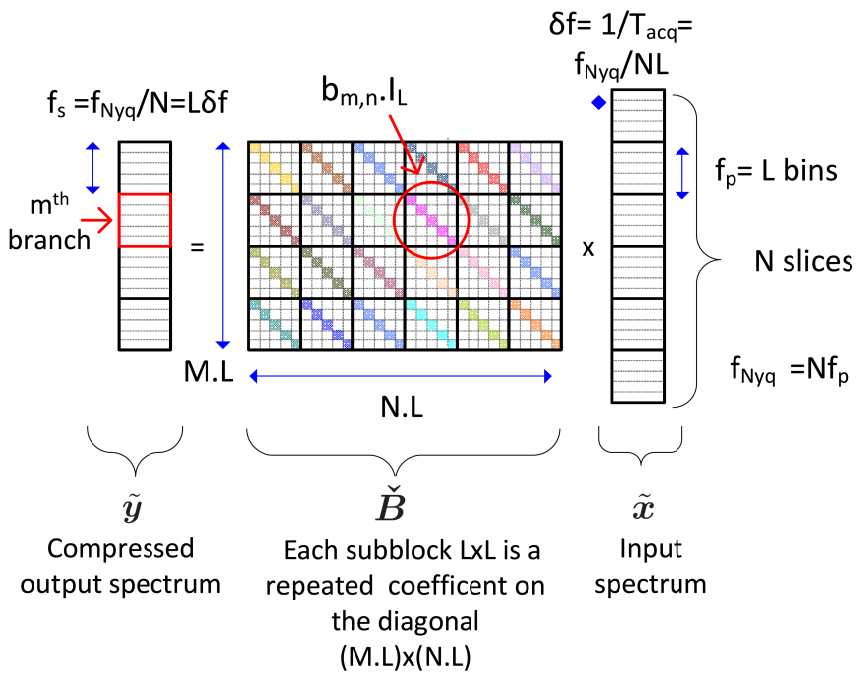
$$\tilde{\boldsymbol{y}} = \check{\boldsymbol{B}}\tilde{\boldsymbol{x}} \tag{2.3}$$

The vector $\tilde{\boldsymbol{y}}$ concatenates the spectrum components of the M acquisition channels, each having L components (bins).

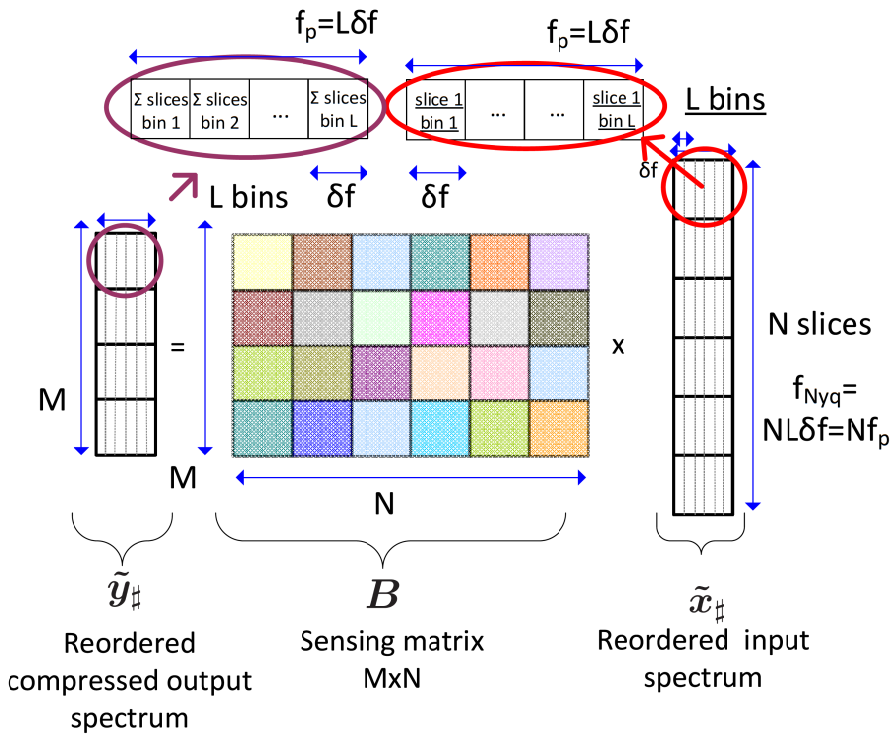
For purposes of computationally-tractable restitution, it is however better to deal with a sensing matrix whose size does not depend on the length of the acquisition. Toward this end, it is possible and preferable, as in [81], to consider a sensing matrix $\boldsymbol{B} \in \mathbb{C}^{M\text{x}N}$, compact version of $\check{\boldsymbol{B}}$, and $\tilde{\boldsymbol{x}}_\sharp$ and $\tilde{\boldsymbol{y}}_\sharp$ so that input $\tilde{\boldsymbol{x}}$ is reordered in a matrix $\tilde{\boldsymbol{x}}_\sharp$ of N rows of L bins, and output $\tilde{\boldsymbol{y}}$ is reordered in a matrix $\tilde{\boldsymbol{y}}_\sharp$ of M rows of L bins. $\boldsymbol{B}$ is illustrated in Figure 2.3b. As the considered frequency band is zero-centered, it finally yields, for odd N and $(m,n,l) \in [\![1;M]\!]\text{x}[\![1;N]\!]\text{x}[\![1;L]\!]$:

$$\tilde{x}_\sharp[n,l] = \tilde{x}((n - \lfloor \frac{N}{2} \rfloor - \frac{3}{2})f_p + (l-1)\delta f) \tag{2.4}$$

$$\tilde{y}_\sharp[m,l] = \tilde{y}_m((l-1)\delta f - \frac{f_p}{2}) \tag{2.5}$$



(a) Multiband architecture, formalism $\check{\boldsymbol{B}}$.



(b) Multiband architecture, formalism $\boldsymbol{B}$.

Figure 2.3: Sensing matrix of a multiband architecture for two different formalisms.

Then eq. (2.3) transforms into:

$$\tilde{\boldsymbol{y}}_\sharp = \boldsymbol{B}\tilde{\boldsymbol{x}}_\sharp \tag{2.6}$$

The benefit is that $\boldsymbol{B}$ is of dimension MxN whereas $\check{\boldsymbol{B}}$ was of dimension $M.L$x$N.L$. Hereinafter the canonical formalism will denote the matrix $\check{\boldsymbol{B}}$ and the multiband formalism will denote the matrix $\boldsymbol{B}$. Compressive multiband architectures will be at the core of Chapter 4.

2.1.4 Modelling of the input noise after compression

As tackled in Subsection 1.2.2, the noise on the compressed samples is composed of the measurement noise and of the projected input noise. The modeling of the input noise after compression is a key question. If the input noise $\boldsymbol{w}$ follows a normal distribution $\boldsymbol{w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_{NL})$ then the noise affecting the compressed data follows a normal distribution $\boldsymbol{\Phi w} \sim \mathcal{N}(0, \boldsymbol{\Phi w}(\boldsymbol{\Phi w})^H) = \boldsymbol{\Phi w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{\Phi\Phi}^H)$.

Assuming that $\boldsymbol{\Phi}$ satisfies properties of coherence and isometry, [82] showed through whitening the Cholesky decomposition of $\boldsymbol{\Phi}$ that assuming $\boldsymbol{\Phi w} \sim \mathcal{N}(0, \frac{N}{M}\sigma^2 \boldsymbol{I}_{ML})$ is an accurate approximation (with unitary column energy normalization of $\boldsymbol{\Phi}$). Hence the assumption that the noise after folding is still white Gaussian, with an SNR degradation of the ratio M/N, will generally be admitted, except if specified otherwise. In other words, as also demonstrated in [83], for any kind of Analog-to-Information Converter a price of a theoretical and unavoidable $10log_{10}(\frac{M}{N})$ must be paid. It is often referred to as $3dB$ SNR loss per octave of subsampling.

2.1.5 Calibration

In theory, the sensing matrix corresponding to a given Analog-to-Information-Converter is defined and known. However, in practice several hardware imperfections (non-linearities from the mixer, phase noise and jitter, non-idealities of analog components such as filters, bad synchronization, etc.) will alter the acquisition process so that the actual sensing matrix is mildly distorted [28, 84]. Some imperfections can be taken into account into a new matrix but not all. As a consequence, the theoretical sensing matrix does not correspond to the transfer function anymore so that using the theoretical matrix for recovery on a concrete prototype might fail. That is why in many cases an accurate calibration process is necessary, depending on the restitution algorithm.

2.1.6 Quantization

An implementation aspect that will not be addressed in this study but should not be forgotten is the quantization issue. The overall compression rate should logically be based on the product between the number of measurement M and the number of bits per measurement. There is a compromise to be found between the two compression phenomena, quantization on one hand and sampling below Nyquist on the other hand. This topic has been tackled e.g. in [85]

for classification tasks, in [86] for correlation estimation, in [87] with respect to the property of democratic measurements and in [88], guarantees optimization is done for a Sigma-Delta quantization scheme. Even measurements severely quantized on 1-bit [89, 90] have been proven to be potentially sufficient.

Now that the framework of compressive sensing applied to radiofrequency receivers has been defined, the different compressive receiver architectures will be presented, reviewed and compared in the next three subsections, starting with devices exploiting the principle of Non Uniform Sampling.

2.2 Non Uniform Sampling

The core concept of Non Uniform Sampling is to exploit the mutual incoherence between the representation space where the signal is sparse and the subsampled acquisition domain. The most straightforward solution is to take advantage of the maximal mutual incoherence between time and frequency [22], as in the Non Uniform Sampler (NUS).

The Non Uniform Wavelet Bandpass Sampling (NUWBS) on the other hand exploits the mutual incoherence between the frequency and time-frequency domain, making it particularly fitted to signals sparse in the latter.

2.2.1 Non Uniform Sampler (NUS)

The Non-Uniform Sampler (NUS) positions itself as the first compressive receiver. Based on ideas dating back to the 60s on spatially sparse antenna arrays subsampling [91], J. Laska, S. Kirolos et al. [92] formalize random sampling into the CS framework as early as 2006. As pictured in Figure 2.4, the principle of the NUS is to sample the incoming signal at intervals irregularly spaced in time by taking a subset of the Nyquist-regularly spaced samples (only black circles). The device, whose block diagram is presented in Figure 2.5, is composed of a Sample-and-Hold stage controlled by a Pseudo-Random Bit Sequence and of an ADC operating at the shortest sampling period between two samples. Formally, the process is described by:

$$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{x} = \boldsymbol{R}\boldsymbol{F}^H\tilde{\boldsymbol{x}} \tag{2.7}$$

where $\boldsymbol{R}$ is a $M.L$x$N.L$ matrix containing 0 and 1 that randomly subsamples the signal in the time domain and it is recalled that $\tilde{\boldsymbol{x}}$ is the Discrete Fourier transform of the signal $\boldsymbol{x}$. Use of a Non Uniform Sampler for Spectrum Sensing has been widely investigated in [93], in particular the choice of an optimal $\boldsymbol{R}$. The time-domain sensing matrix of the NUS is pictured in Figure 2.6, with a highlight on the MxN sampling pattern. The way of implementing the clocking subset patterns may rely on different strategies:

- Periodic non-uniform sampling (PNUS), also known as Multicoset sampling [94], corresponds to the periodical repetition of a stored non-regularly spaced sampling pattern. In

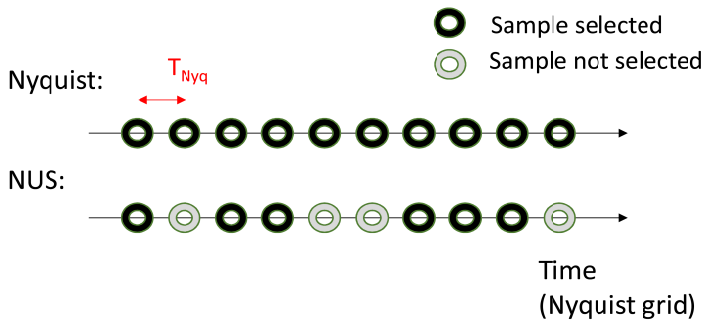


Figure 2.4: Sketch of time sample selection in the Non Uniform Sampler.

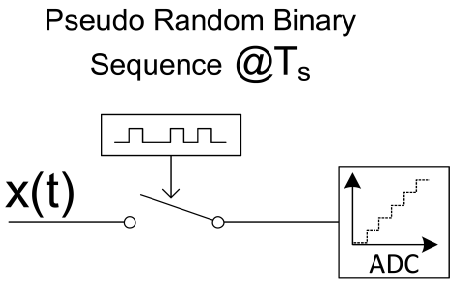


Figure 2.5: Block diagram of the NUS.

practice it consists of M parallel channels with subsampling factor N, and initial shifts p_m ($1 \leq m \leq M$, $1 \leq p_m \leq N$) determined by a pattern $\boldsymbol{P} = [p_1 — p_m]^T$. Hence the lth sample in the mth channel is given by:

$$y_m[l] = x((lN + p_m)T_{Nyq}), 1 \leq m \leq M$$

- Randomized non-uniform sampling (RNUS) exploits one on-chip randomly generated subsampling sequence for the whole set of time intervals [49] allowing for flexible under-sampling rate adjustment.
- Level-triggered non-uniform sampling (LTNUS) operates with a grid determined by the randomized (through a feedback loop) zero crossing points of the one-bit quantized representation of the input signal [95].

The post-processing is computationally relatively simple, enabling quick recovery [96]. However, taking less samples may decrease the average sampling rate but does not suffice to

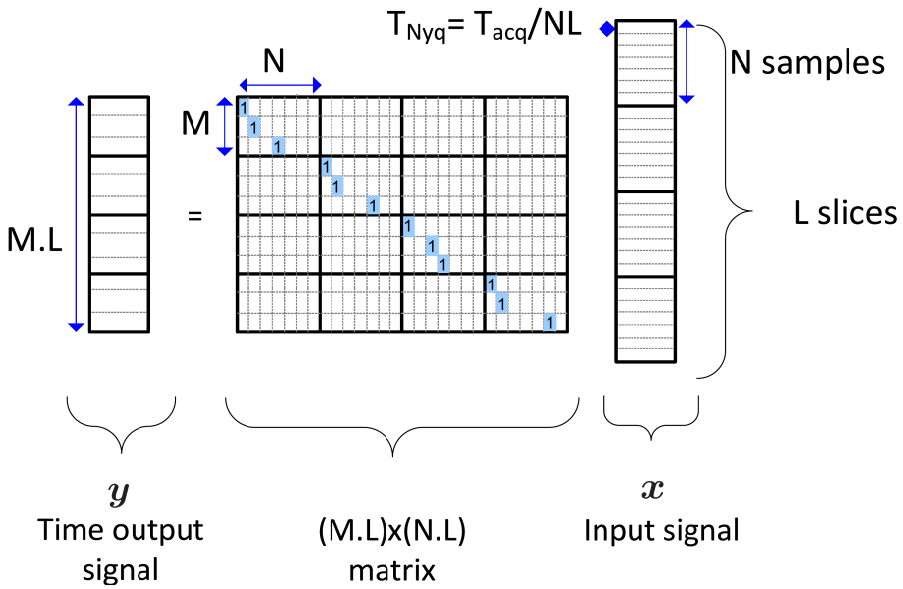


Figure 2.6: Sensing matrix and process of the NUS (canonical formalism).

alleviate completely the Nyquist rate constraint since two subsamples could be separated from just one Nyquist bin (possibly a couple of bins if the pattern is well-chosen). Also the Track-and-Hold circuitry still needs to track a Nyquist-rate varying input, even if the conversion lasts longer [81]. A major drawback is that a power-greedy Nyquist rate clock is still required to synchronize the shifters on each branch [97]. A second issue is that the NUS is extremely sensitive to clock jitter [72].

2.2.2 Non Uniform Wavelet Bandpass Sampling (NUWBS)

The principle of Non Uniform Wavelet Sampling (NUWS), introduced lately by M. Pelissier and C. Studer [98], is to perform a wavelet transform preprocessing before the non uniform sampling step to introduce more degrees of freedom and alleviate some implementation issues. Concretely, the NUWS correlates the input signal with a set of wavelets with variable time support τ and central frequency f_c, integrates over the support of each wavelet, and subsamples the resulting wavelet coefficients. The corresponding block diagram, which is similar to the NUS block diagram except for the $\boldsymbol{W}$ preprocessing block, is pictured in Figure 2.7. Hence the Dirac stream sampling scheme of the NUS is replaced by a time-subsampled wavelet comb of variable parameters, namely central frequency, sampling instant and wavelet envelop linked to its scale. Non Uniform Wavelet Bandpass Sampling (NUWBS) is a variation of the NUWS based on bandpass sampling techniques and adapted to RF multiband signals. A graphical comparison of the sampling mechanisms of NUS, NUWS and NUWBS in descending order is provided in Figure 2.8. In particular the T_s-uniform spacing of the wavelets in the NUWBS scheme is brought to light. Different colors correspond to different wavelets as the parameters

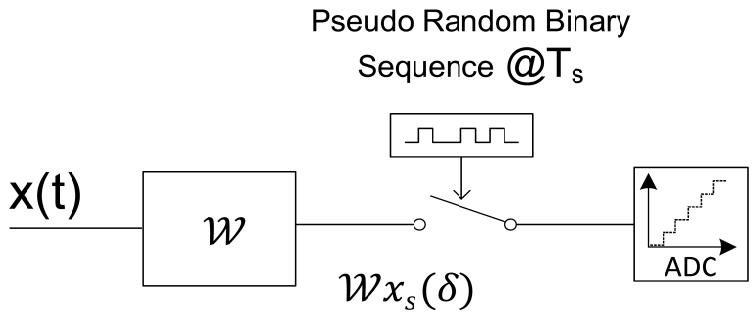


Figure 2.7: Non Uniform Wavelet Bandpass Sampling Block Diagram.

τ and f_c can be chosen different. The Gaussian envelope depicted in dark gray varies with τ. Virtual sample instants that are not selected by $\boldsymbol{R}$ are depicted with a light gray Dirac instead of a black Dirac.

Figure 2.9 represents the NUWBS operating principle in more detail. Formally the sensing process is described by:

$$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{W}^H\boldsymbol{x} = \boldsymbol{R}(\boldsymbol{F}\boldsymbol{W})^H\tilde{\boldsymbol{x}} = \boldsymbol{R}\boldsymbol{W}^H\boldsymbol{F}^{-1}\tilde{\boldsymbol{x}} \tag{2.8}$$

where $\boldsymbol{R}$ is a $M.L$xW matrix that subsamples the wavelet frame $\boldsymbol{W}^H$ of dimension Wx$N.L$, potentially over-complete since $W \geq M.L$. Hence compression is obtained by two combining effects, non uniform sampling by $\boldsymbol{R}$ of the possible time delays and folding of the spectrum by $\boldsymbol{W}^H$. The operations can be parallelized or serialized, in which case $\tilde{\boldsymbol{x}}$ is considered stationary over the different correlations. Whereas the NUS subsamples the inverse DFT matrix, eq. (2.8) highlights that the NUWS subsamples the conjugate transpose of the Discrete Fourier Transform of the wavelet frame.

From [98], the Fourier representation of the wavelet atoms $\boldsymbol{\Psi} = (\boldsymbol{F}\boldsymbol{W})^H$ for central frequency f_c and time shift δ_k with choice of a Gaussian windowing defined by its time spreading parameter τ is given by:

$$\Psi_{f_c,\delta_k}(f) = (\tau\sqrt{2\pi})^{1/2}e^{-j2\pi\delta_k f}e^{-(\pi\tau(f-f_c))^2} \tag{2.9}$$

Compared to the NUS for example, the NUWS reduces the bandwidth of the Sample-and-Hold due to mixing and low-pass processing, and avoids the use of Nyquist-rate clocks or code sequence. Compared to bandpass subsampling, the selection of adequate wavelets in the NUWBS allows to attenuate the out-of-band noise and interferers by focusing on the subbands of interest without loosing the information from the rest of the spectrum. The three offered degrees of freedom (sample time instants, wavelet bandwidth, center frequency)

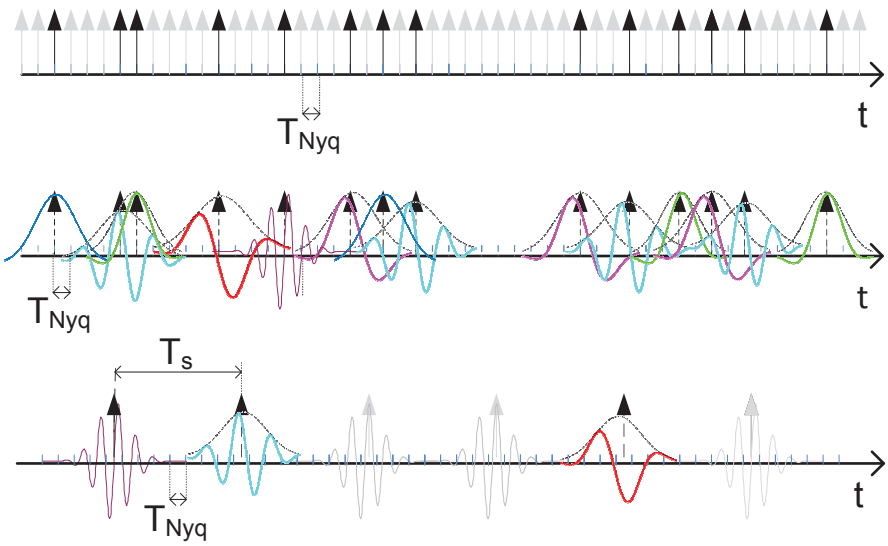


Figure 2.8: Sampling patterns of NUS, NUWS and NUWBS, inspired from [98].

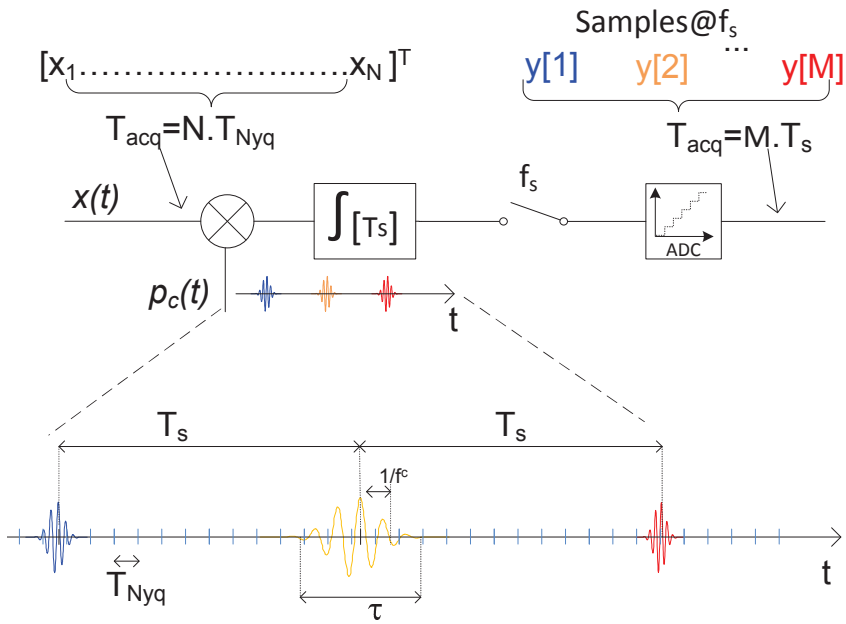


Figure 2.9: Non Uniform Wavelet Bandpass Sampling for a single branch, inspired from [98].

are particularly favorable in order to cope with the challenge of changing environments. In contrast to the other immutable architectures with e.g. a static number of channels, this flexibility enables adaptive approaches. Note that generating wavelets in the time domain is indeed very similar to generating pulses for ultra-wideband (UWB) applications. It has been shown recently that power-efficient implementation is possible [99].

2.3 Random Demodulation

The concept of Random Demodulation gathers compressive acquisition schemes inspired from Direct-Sequence Spread Spectrum (DSSS) receivers, where the input signal is modulated by a high-rate pseudo-random sequences, low-pass filtered and then demodulated during the reconstruction. The Random Modulation Pre-Integrator (RMPI) uses integration to reduce the dimension of the signal whereas the MWC uses ideal low-pass filtering. They share the same key principle but this difference implies a different modeling of the input signal. The Compressive Sensing Filter (CSF) inspired from the Random Convolution theory uses random phase modulation.

2.3.1 Random Demodulator (RD)

J. Laska, S. Kirolos et al. [100] also laid the foundation for the implementation of RF receivers based on random demodulation in 2007. In this approach, the RF signal is correlated by analog means with a Nyquist-rate sequence $p(t)$ of pseudo-random ± 1, integrated over a time window, and sampled uniformly at low rate. Thus the burden of performing high-rate sampling is alleviated by high-rate mixing. The Random Demodulator (RD) refers usually to the one-branch variant whereas the Random Modulation Pre-Integrator (RMPI) refers to the multi-branch variant. The corresponding architecture which includes a mixer, an integrator and a low-rate ADC is featured in Figure 2.10. The process in the spectral domain is pictured in Figure 2.11. It highlights that mixing a tone with a random signal amounts in the frequency domain to shifting a spectrum that is approximately white noise. As different shifts of the noise spectrum are nearly orthogonal, it is then possible to identify amplitude and frequency of few tones from low-rate samples according to the signature of the noise. Formally, the

sensing process is described in the time domain by:

$$\mathbf{\Phi} = \begin{pmatrix} \phi_1 & 0 & \dots \\ 0 & \phi_2 & \dots \\ \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} \pm1 & \pm1 & \dots & \pm1 & 0 & \dots & \dots & 0 & \vdots \\ \pm1 & \ddots & \dots & \pm1 & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ \pm1 & \pm1 & \dots & \ddots & 0 & \dots & \dots & 0 & \vdots \\ 0 & \dots & \dots & 0 & \pm1 & \pm1 & \dots & \pm1 & \vdots \\ \vdots & \ddots & & \vdots & \pm1 & \ddots & \dots & \pm1 & \vdots \\ \vdots & & \ddots & \vdots & \vdots & \dots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \pm1 & \dots & \dots & \pm1 & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{2.10}$$

where each $\mathbf{\Phi}_{l\in[\![1;L]\!]}$ is a MxN subblock containing different random ±1, drawn typically from a Bernoulli distribution.

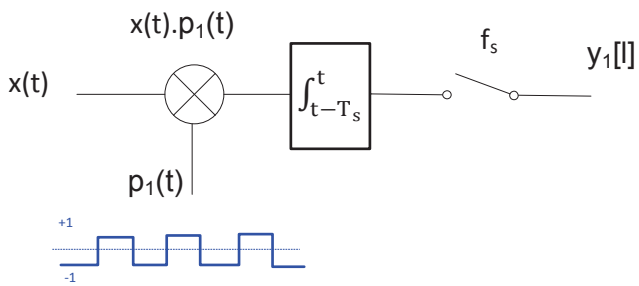


Figure 2.10: Random Demodulator block diagram.

One can note that the structure of the Random Modulation Pre-Integrator is similar to the one of a multi-user receiver in Coded-Division Multiple-Access (CDMA) communication system [101, 102], with a front end composed of a bank of matched filters (to the codes of different users), composed in each branch by despreading (multiplication by one code), integration (low pass filter), and sampling operations.

Y. Juhwan [103] proposes a $90nm$ CMOS technology implementation of the RMPI for impulse radar detection within $100MHz - 2GHz$. A benchmark with respect to Nyquist rate converters regarding jitter and aperture error and power consumption is established in [48].

Since the clock is uniform, Random Demodulators are less sensitive to timing jitter than Non Uniform Samplers [72]. However, the pseudo-noise generator must still run at the expensive Nyquist rate. And the dictionary model represents a severe limitation as it implies a high sensitivity to the grid choice and a computationally-complex reconstruction process. The multitone model for instance is not adapted to real time processing [28] because the dimension of the sensing matrix with a reasonable grid spacing is too large.

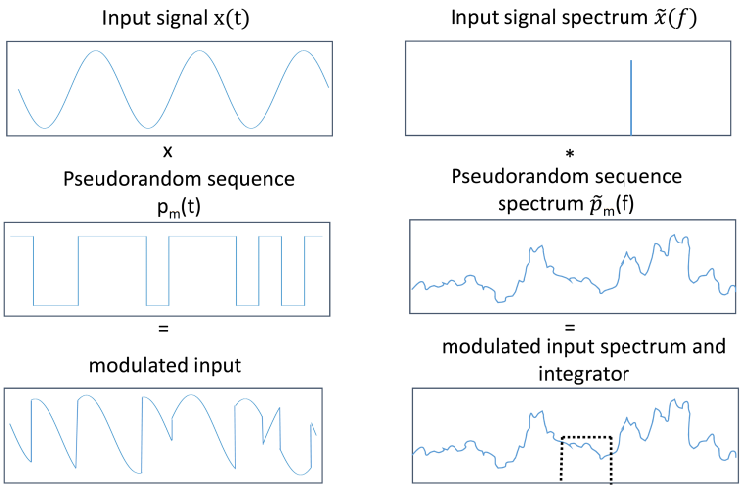


Figure 2.11: Action of the Random Demodulator on a pure tone, inspired from [100].

2.3.2 Modulated Wideband Converter (MWC)

Similarly to the RMPI, the Modulated Wideband Converter relies on analog multiplications with pulse-shaped pseudo-random code sequences, whose elements are at the Nyquist rate. One major difference is that this first step is followed by ideal low-pass filtering before uniform sampling at low rate. This results in each subband being folded at baseband with weights linked with the mixing function. As for the RMPI, performing high-rate sampling is avoided by spreading all the spectrum at baseband through mixing. The acquisition scheme of the MWC is pictured in Figure 2.12 and detailed hereafter.

In each of the M parallel branches, the K-sparse (K active frequency subbands) input signal is mixed with functions based on codes. For each branch $m \in [\![1; M]\!]$, the mixing function $p_m(t)$, pictured in Figure 2.13, is T_p-periodic and consists, within each period, in a code $\mathbf{a}_{m,.}$ of N elements, shaped with e.g. rectangular chip pulses of period T_c.

The input spectrum is thus convolved with a f_p-spaced Dirac comb so that each band is weighted by the corresponding Fourier coefficient of the code and the whole spectrum is aliased at baseband. Hence diversity is created because each linear combination of active subbands is different. The last step consists in low-pass filtering ($h(t)$ in Figure 2.12) with cut-off frequency $f_c = 1/2.T_s$ and uniform sampling at f_s. Unless stated otherwise, the canonical version where $f_s = 1/T_s = f_p$ will be considered. Because they run at the Nyquist rate, the mixing codes represent most of the energy consumption of the MWC.

Appendix B details formally the acquisition process for a Modulated Wideband Converter and demonstrates that the coefficients $b[m, n]$ correspond to the Fourier coefficients of the time domain mixing codes $p_m(t)$ in the mth branch for index $n - \lfloor \frac{N}{2} \rfloor - 1$. The result, which is established by Y. Eldar in [81], is detailed more thoroughly and with the notations introduced

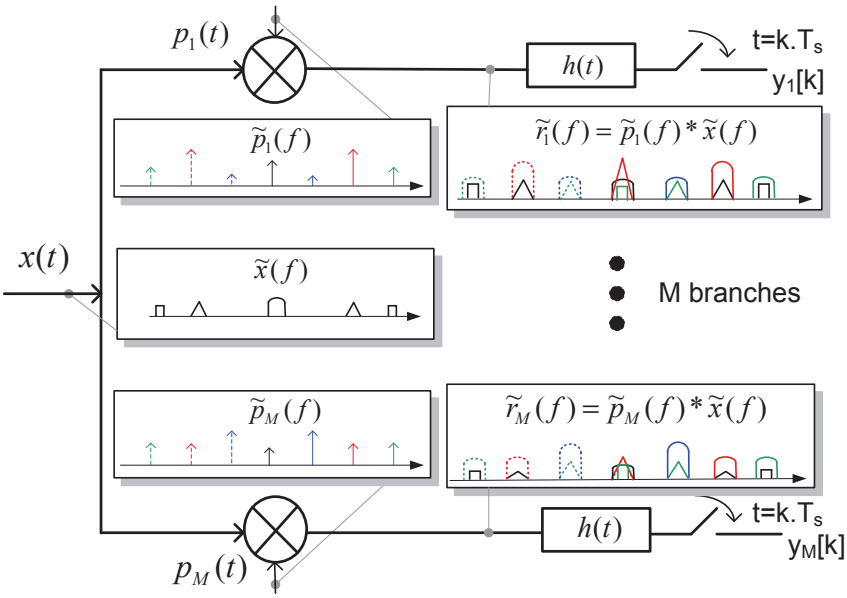


Figure 2.12: Sensing process of the MWC, from .

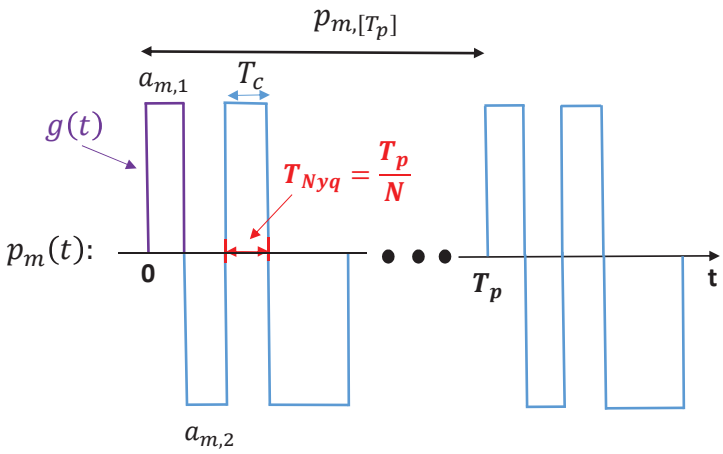


Figure 2.13: Mixing function $p_m(t)$.

in this work for sake of comprehension. Thus the MWC sensing matrix can be written, in the canonical formalism as in Figure 2.3a or in the compact multiband formalism as in Figure 2.3b with $b[m, n]$ as computed in the Appendix B.

As stated earlier, Random Demodulators differ from the MWC regarding the expected signal model, which is for the latter a sum of analog subbands. This corresponds in terms of block diagram to the fact that the MWC uses an ideal low pass filter with a finite frequency response, localized in frequency but not in time. Whereas Random Demodulators use an ideal

integrator with a finite impulse response, localized in time but with infinite spectral support. This difference in representation models leads to better performances of the MWC regarding robustness to model mismatch, hardware complexity and computational loads [1] .

Note also that non-ideal time-domain properties have no effect on the MWC as long as spectral properties are preserved [1]. Also the ideal low-pass filter, which would be difficult to implement, can be replaced by any filter fulfilling the Nyquist intersymbol interference (ISI) criterion for symbol frequency f_p [104]. However, compared to other receivers, an MWC would yield poor performance if inputs are time-dependent like Radar pulses [28] because a static spectrum needs to be assumed on many time samples. A calibration process based on injecting consecutive sinusoidal inputs at incrementing rates is described in [84].

The Quadrature-Analog-to-Information-Converter (QAIC) system proposed by [105] is a variation of the MWC that performs frequency down-conversion of the real-valued signal into a baseband complex-valued signal (in phase and quadrature low-pass components) before the analog mixing with the pseudo-random codes. Reducing the analyzed bandwidth allows to use code sequences with shorter length and lower frequency and is often justified for cognitive radio applications, where the entire spectrum does usually not need to be scanned. Energy consumption has been estimated [6] to be potentially an order of magnitude lower than the MWC, since the code generation is the most power-greedy part. Of course frequency downconversion implies in return the issue of I/Q mismatch. Encouraging performances in interferer detection were presented in [6]: the QAIC detects interferers in the spectrum between 2.7GHz and 3.7GHz with a 4.4μs sensing time for a 20MHz resolution bandwidth and a consumption of $81mW$. The interested reader is referred to Appendix C for a more detailed numerical benchmark.

2.3.3 Compressive Sensing Filter (CSF)

The Random Convolution theory proposed by J. Romberg [106] and its application to Compressive Sensing Filters (CSFs) equates to a random dephasing of the frequency subbands through convolution with a filter impulse response. It is shown in [106] that such an approach is universal, which means that it can tackle signals that are sparse in any domain. More notably, it is particularly suited for signals sparse in the time domain. Preservation of the time structure requires precise relationships between the phases of the Fourier coefficients, which will be destroyed by the random filtering. Hence it is ensured that the information is not concentrated anymore in the measurement (time) domain. The acquisition process can formally be described by following equation:

$$\boldsymbol{y} = \frac{1}{\sqrt{N}} \boldsymbol{R} \boldsymbol{F}^{-1} \boldsymbol{D} \boldsymbol{F} \boldsymbol{x} \tag{2.11}$$

where $\boldsymbol{D}$ is a $N.L$x$N.L$ diagonal filtering matrix that must satisfy additional conditions (unit magnitude coefficients drawn from a uniform distribution and conjugation between two halves) and $\boldsymbol{R}$ is a $M.L$x$N.L$ subsampling matrix. In [107], $\boldsymbol{R}$ subsamples with a regular spacing the time domain, while in the more general framework of [106], $\boldsymbol{R}$ selects a subset of M rows out of N. The foreseen implementations with random phase filters as in [107]

did however not meet success for wideband applications because affecting only phase over a wide bandwidth is horrendous from an implementation point of view, as magnitude must be preserved over all frequencies.

Note a certain similarity between Random Convolution and a Modulated Wideband Converter with mixing codes that have random phases but constant magnitudes, a device which would then also operate a simple dephasing. However, a notable difference lays in the reduction of the signal dimension, which is done by subsampling the time representation for Random Convolution and by filtering the spectral representation for the MWC.

2.4 Variable Rate Sampling

It is first recalled that Bandpass sampling [108] consists in subsampling under the Nyquist rate with a carefully chosen frequency f_s. The output spectrum is given by:

$$\tilde{y}(f) = \sum_{k=-\infty}^{+\infty} \tilde{x}(f - kf_s) \tag{2.12}$$

Variable Rate Sampling corresponds to Bandpass Sampling, carried out with different sampling frequencies f_s. It relies on the fact that with different sampling rates, the spectrum folds in a different manner, hence aliases happen differently. In the case of Multirate Sampling (MRS), a predetermined set of sampling frequencies is used in each of the parallel branches. But also the Nyquist Folding Receiver (NYFR) can be considered to be a variable rate sampling architecture, with a rate that is varied continuously.

2.4.1 Multirate Sampling (MRS)

Multirate Sampling is based on the fact that since each branch provokes frequency aliases between different frequencies, it is possible to locate bands by comparing the spectra. Figure 2.14 pictures the block diagram of Multirate Sampling, which consists of M branches with bandpass sampling at different frequencies. Among Multirate Sampling, a distinction can be made between what is called asynchronous and synchronous multi-rate sampling.

- Synchronous Multirate Sampling [109, 110], represented in Figure 2.15, implies that the sampling starts simultaneously in all channels.
- In Asynchronous Multirate Sampling, the target application is energy detection so the phase information is not needed [97]. Hence synchronization between the channels is not required, therefore hardware design constraints are relaxed and the calibration effort is reduced.

The choice of an optimal set of undersampling frequencies, based on coprime integers, is detailed in [110, 111]. The main advantage of this approach is that it does not need a Nyquist

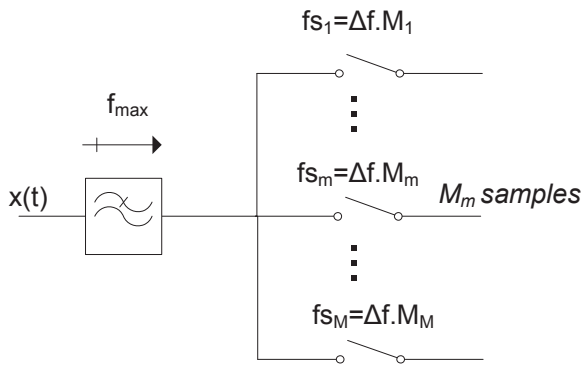


Figure 2.14: Block diagram of Multirate Sampling.

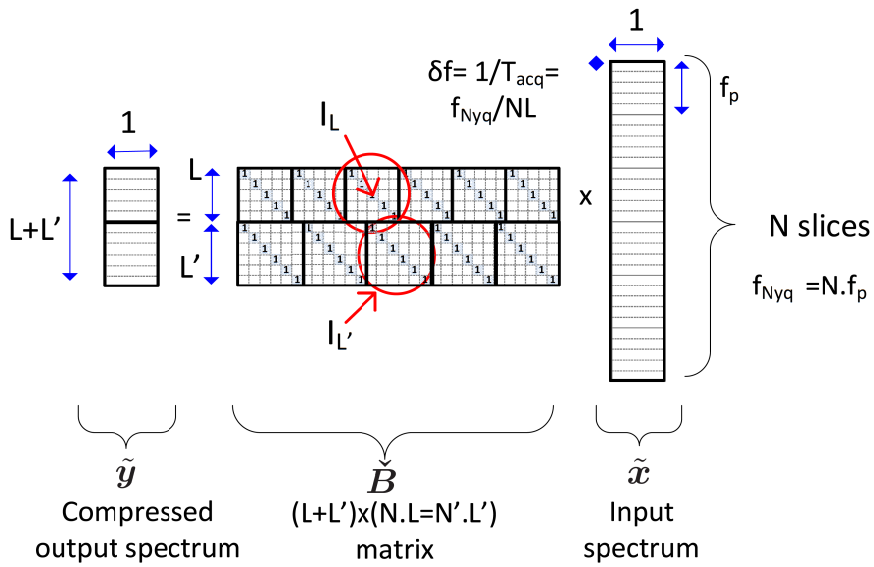


Figure 2.15: Multirate Sampling Sensing Matrix.

rate clock. Another asset is its robustness when the signal is not very sparse (>50% subband occupation typically). Compared to the NUS, less branches are needed and the processing time is also relatively short. However, there remains the need of a Nyquist-bandwidth front-end and the issue of high sensitivity to jitter.

2.4.2 Nyquist Folding Receiver (NYFR)

The Nyquist Folding Receiver (NYFR) [42] performs variable rate chirp sampling so as to modulate the input spectrum in frequency and bandwidth. As such, it could also be seen as hybrid between (structured) Non Uniform Sampling and Random Demodulation (but with a

sub-Nyquist pulse rate) approaches. As pictured in Figure 2.16, the folding is achieved by undersampling the input spectrum according to a stream of time-modulated short pulses and then interpolating via continuous-time low-pass filtering.

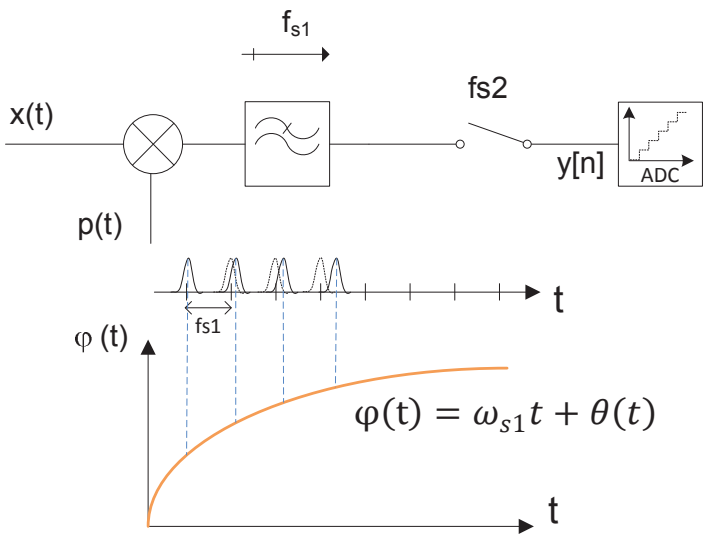


Figure 2.16: Nyquist Folding Receiver.

The sampling operations mark the original spectrum bands with frequency-dependent signatures in the following manner: The sampling rate is on average f_{s1} but varies sinusoidally between two values $f_{s1} + \Delta$ and $f_{s1} - \Delta$, with a given modulation index MI according to a narrowband clock modulation $\Theta(t)$. The aliased signal will change frequency over time since the sampling rate is changing over time. The modulation scale factor MI is an integer corresponding bijectively to the fold (or alias) number, i.e. the corresponding support element. As a result, signals at the output of the interpolation filter are modulated by $MI.\Theta(t)$.
For instance if, as in [42], a sinusoidal input is considered:

$$x(t) = cos(\omega_c t + \Psi(t))$$

where $\Psi(t)$ is the information content of the input signal. Then the output is given by:

$$y(t) = cos(\omega_f \pm \beta\Psi(t) - MI.\Phi(t))$$

where:

- $\omega_f = |\omega_c - \omega_{s1} k_h|$ gives the folded spectrum orientation (positive or reversed)
- $k_h = \lfloor \frac{\omega_c}{\omega_{s1}} \rfloor$ is the sampling harmonic
- $MI = \beta.k_h$ is the induced modulation factor

The acquisition matrix corresponding to these operations is pictured in Figure 2.17.

A major asset of the Nyquist Folding Receiver lies in its ability to sample high analog

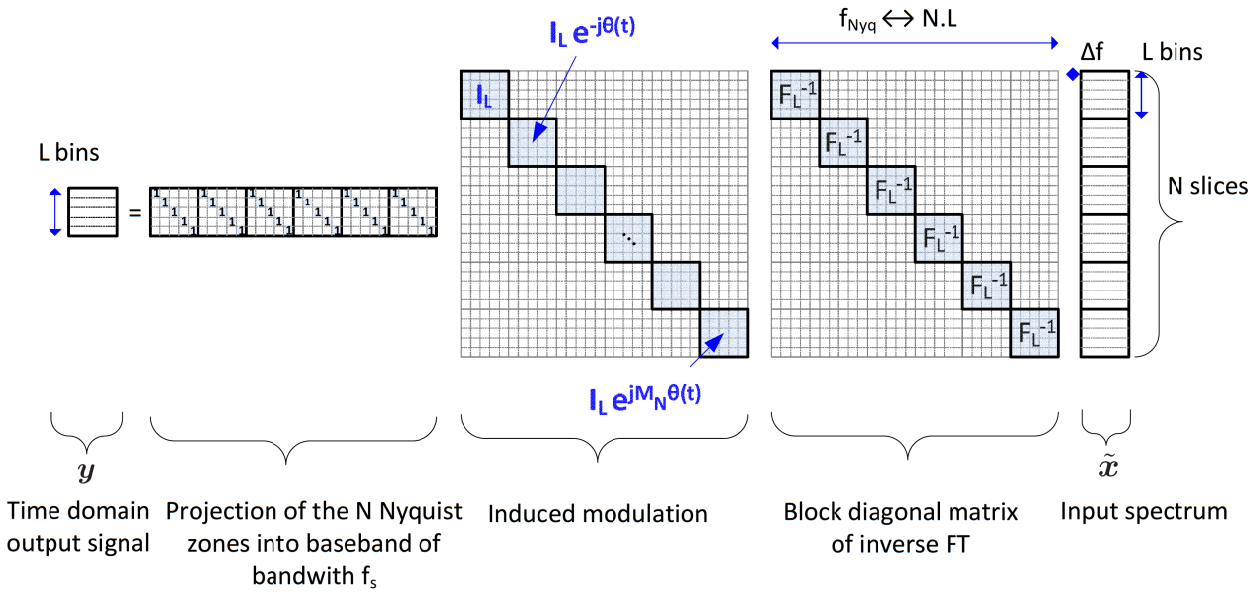


Figure 2.17: Sensing matrix and process of the Nyquist Folding Receiver.

input frequencies without the need of any high-speed components operating at the Nyquist rate for the maximum analog input frequency. Another perk is the possibility to use fast and power-efficient recovery techniques instead of traditional algorithms, since the NYFR preserves essentially the signal structure. The solution is robust to mismatch with the frequency basis as the samples do not need to correspond to a grid defined by the Nyquist rate. As magnitude is preserved, it is also less sensitive to noise folding than other architectures such as the MWC, because there is no subband where the power is relatively amplified. However, on the other hand, a stumbling point lies in the high sensitivity to noise at the zero-crossing risings (jitter).

2.5 Conclusion

A large variety of Analog-to-Information-Converters has been put forward in order to make the promises of Compressive Sensing to overcome the Nyquist rate come true. In addition to the main sparsity concept, AICs exploit several ideas originating from traditional RF receivers such as Spread Spectrum and Bandpass Sampling, in order to create diversity in miscellaneous ways. For instance, random bandpass sampling techniques (Nyquist Folding Receiver, Multi-rate Asynchronous Sub-Nyquist Sampling) take inspiration from bandpass sampling to create different frequency aliases that can be disentangled. Also random pre-integrators (Random Demodulator, Modulated Wideband Converter) perform coding by the modulation with a high-rate random number generator.

It appears though that even if the signal is sampled under the Nyquist rate, many architectures still require Nyquist-rate components, which is disappointing. Another issue seems to be the lack of versatility of most architectures, which is inconvenient for cognitive radio applications. Prohibitive reconstruction costs, unavoidable noise folding that degrades the SNR and sometimes large silicon area due to parallelization of the architecture are other topics of concern. Those rising technologies are still prototypes.

The following architectures are particularly encouraging with respect to the stumbling points that have been raised:

The Quadrature-Analog-to-Information-Converter architecture (QAIC), based on a band-passed MWC, focuses the spotlights as it enhances further the popular field-proven MWC regarding to both power and code sequence generator simplicity. Performances presented in [6] regarding the common issue of interferer detection in particular generate much interest.

The NYFR is a clever method combining Compressive Sampling, Bandpass Sampling, and pulse train modulation that has the two assets of no Nyquist-rate components and signal structure preservation for quicker restitution.

The Non Uniform Wavelet Bandpass Sampler, which does not need any Nyquist-rate component either, is also very promising thanks to an adaptive multiscale approach. As it offers a high degree of flexibility and adaptation, it overcomes some notorious drawbacks of AIC architectures and presents itself as a natural choice for signals sparse in both time and frequency in particular.

The presented receivers were conceived with the purpose of reconstruction of the original signal from the compressed samples but they can be equally considered for partial information retrieval purposes. In Chapter 3, some aspects of the presented receiver architectures will be investigated further. A structural principle found across many receivers inspired by random demodulation is the pseudo-random mixing sequences, and it is the principal bottleneck of their implementation. Hence the choice of adequate codes is a topic that needs to be tackled further in order to be able to yield answers to the implementation challenges. The Modulated Wideband Converter architecture will generally be taken as baseline in this work, due to its qualities of robustness to basis mismatch and compact sensing matrix. That is why properties of the mixing codes will be studied and discussed for the particular instance of a Modulated Wideband Converter, and different codes benchmarked.

Chapter 3

Code properties analysis for the implementation of a Modulated Wideband Converter

Contents

3.1 From random matrices to structured acquisition 65

3.1.1 Limitation of random matrices 65

3.1.2 Circulant matrix structures 66

3.1.2.1 Motivation and definition 66

3.1.2.2 Typical performances 67

3.2 On the interest of CAZAC sequences for structured MWC acquisition 69

3.2.1 Constant Amplitude Zero Autocorrelation (CAZAC) sequences 69

3.2.2 One example: Zadoff-Chu codes 70

3.2.3 Proposition of an original mixing code sequence 70

3.3 Analysis of the reconstruction performances of the sensing matrix . . 72

3.3.1 Coherence 72

3.3.2 Isometry 76

3.3.2.1 Statistical estimation 76

3.3.2.2 Isometry under additional assumptions: Expected RIP 78

3.3.3 Empirical validation of the reconstruction performances by means of simulation platform 79

3.3.3.1 Simulation framework 79

3.3.3.2 Noiseless scenario 81

3.3.3.3 Noisy scenario 84

3.4 Conclusions and Perspectives 87

A known difficulty of the MWC is choosing appropriate periodic functions $p_n(t)$ so that their Fourier coefficients fulfill CS requirements.

- S. Stein, O. Yair, D. Cohen, Y. Eldar, *CaSCADE: Compressed Carrier and DOA Estimation [112]*

It was demonstrated that the main performance bottleneck of sub-Nyquist systems proposed to date are the pseudo-random bit sequence generators employed in the analog frontend.

- T. Haque, R. T. Yazicigil, K. J.-L. Pan, J. Wright, P. R. Kinget, *Theory and Design of a Quadrature Analog-to-Information Converter for Energy-Efficient Wideband Spectrum Sensing [105]*

In Chapter 2, different architectures of radiofrequency receivers based on Compressive Sampling have been presented and discussed. The Modulated Wideband Converter (MWC) and by extension its downconverted to baseband version, the Quadrature-Analog-to-Information Converter (QAIC), notably stood out. This is due in particular to robustness to basis mismatch and reduced dimension of the recovery problem, qualities that are inherited from the compact multiband formalism [1]. As the cornerstone of architectures based on random demodulation is clearly the mixing codes regarding both implementation, power consumption and recovery performances, it is necessary to investigate this crucial topic further, which is at the core of this chapter. This work has resulted in the conference articles [113] and [114].

In the first section, system aspects regarding number of branches of the MWC architecture and generation of the code sequences will be tackled. The potential benefits of circulant matrices from an implementation point of view will be highlighted and their performances in signal reconstruction will be outlined.

Our contributions begin in the second section, where innovative codes which capitalize on specific properties brought to light are introduced. A particular code family, selected because of its interesting correlation properties, Constant Amplitude Zero Autocorrelation (CAZAC) sequences will be featured. Then a specific instance, the Zadoff-Chu codes, will be more precisely investigated. Based on this, a proposition of a new code matrix is formulated, in anticipation of its evaluation.

In the third section, an original multifaceted performance analysis is presented. That is, different code families with different row selection matrices $\boldsymbol{R}$ are compared based on multiple criteria. In the first place, mathematical properties are addressed. To begin with, the coherence, which is vital to limit the number of measurements and hence of parallel branches, is computed and compared for different codes. Properties of norm and distance preservation, crucial to ensure robustness to noise, are evaluated through two different tools. A first estimation of the Restricted Isometry Property (RIP) is conducted via a statistical Monte-Carlo simulation. Additionally, a comparison is lead based on the Expected RIP which

is the probability that RIP is satisfied under furtherl prior on the input. High-level insights are validated by means of a MATLAB® simulation platform featuring the Modulated Wideband Converter acquisition and reconstruction process, in both noiseless and noisy settings.

Finally, a conclusion is drawn from the presented results.

3.1 From random matrices to structured acquisition

In order to implement the Modulated Wideband Converter (MWC), the degree of parallelization i.e. the number of branches M, see Figure 2.12, should be kept small. However, less measurements means that the manageable sparsity level K is limited, which limits the range of application. One way to overcome this limitation is to optimize as much as possible the information captured in each measurement by choosing very carefully the best sensing matrix $\mathbf{\Phi}$ while keeping in mind implementation problems. Other complementary solutions are mentioned in the literature, including serialized [115] or collapsed architectures [81, 116] but it is not sufficient.

Typically, following matrices are employed: the authors in [81] use random Bernoulli codes, the inconvenient being that it requires $M.N$ flip-flops for their generation. In [35], a shift register with a pattern chosen according to a criterion of high ExRIP value (cf 1.2.3.1.c) is implemented. The authors in [117] use a circulant matrix generated from a maximal or Legendre sequence. On the other hand, the RMPI developed by Northrop Grumman [28] or the QAIC [6] use Gold sequences, which only exists for $M = 2^k - 1$ where k is an integer corresponding to the number of flip-flops [118]. Since it is typically used in Spread Spectrum communications techniques and benchmarked in [119], Hadamard sequences are mentioned but never chosen because not efficient for information retrieval.

3.1.1 Limitation of random matrices

Random matrices whose entries are independent and identically distributed (i.i.d.) satisfy the Restricted Isometry Property (RIP) with high probability if the entries are chosen according to a Gaussian, Bernoulli, or more generally any sub-Gaussian distribution [5, 23]. Therefore in theoretical representation of Compressive Sampling, a typical way to generate $\mathbf{\Phi}$ consists in sampling i.i.d entries from a symmetric Bernoulli distribution taking values $\pm\frac{1}{\sqrt{M}}$ with equal probability, so as to generate M vectors.

However, in practical applications the actual hardware of the sensing devices, for which some examples have been given in the previous chapter, must be taken into account. It means that the sensing matrix exhibits an underlying structure, a notion referred to as Structured Compressed Sensing [120]. Also, only one instance of a random matrix can be implemented at once. In other words, the object of interest in this study is a deterministic matrix structured around one element (or more) which should provide 'random-like' properties. In the case of receivers based on random demodulation, that would be the mixing code sequences. These 'random-like' properties, that are required to guarantee recovery, can be the consequence of the element being one draw from a random distribution. But they can also derived from the

deterministic definition of an expression adequately chosen. Since generating true randomness is, to say the least, not an easy task in hardware, the latter approach is promising for practical purposes. The randomness aspect is exploited fruitfully in other application contexts, e.g. in compressive imagers [121] or optical computing [122], however in compressive radiofrequency receivers, it is perhaps not the most suitable option.

Another potentially misleading theoretical concept would be universality, meaning that the recovery guarantees given by a sensing matrix must hold independently from the basis or dictionary $\mathbf{\Psi}$ in which the signal is sparse. This is another reason why matrices with entries drawn from a standard random distribution, which ensures this property, are often used as a reference in the mathematical framework of CS. However, imposing such constraint to a concrete architecture is not mandatory. In this study it is known that the signals are sparse in the frequency domain, so it is possible to explicitly take $\mathbf{\Psi}$, the IDFT, into account in the construction of $\mathbf{\Phi}$. Matrices that perform poorly in a very general frame could perform very well in a specific setting, in this study a signal sparse in frequency or time.

Therefore the goal is to find practical matrices that are incoherent with bases such as time or frequency domain and not necessarily with any basis.

3.1.2 Circulant matrix structures

3.1.2.1 Motivation and definition

Besides the performance of the sensing matrix with respect to information retrieval, the hardware cost of code generation must be considered. Toward this end, circulant matrices which only require to generate one code stand out. Additionally, they exist for every number of row $M \leq N$ and the structure in the matrix should improve the speed of recovery computations. For example, in [123], specific variations in relaxation reconstruction algorithms are proposed in order to seize benefit of the structure of the circulant sensing matrix, enabling much faster processing of the reconstruction.

The possible use of a matrix whose M' first sequences were shifted to create the $(M - M')$ last rows was quickly mentioned in [81] regarding the MWC. It is reported that there was no empirical performance degradation for $M' = M/5$. Therefore circulant matrices will be investigated in the following.

As represented in eq. (3.1), circulant matrices, noted with $\boldsymbol{C}$, can be defined by a shift of the N elements of the first row $\mathbf{c}$ in the time domain or by the diagonal $\boldsymbol{\sigma}_{k\in\{1,\ldots,N\}}$ in the frequency domain [123].

$$\mathbf{C} = \begin{pmatrix} \overbrace{\boxed{c_N \quad c_{N-1} \quad \cdots \quad c_1}}^{\text{temporal code in the first branch}} \\ c_1 \quad c_N \quad \cdots \quad c_2 \\ \ddots \quad \ddots \quad \ddots \quad \vdots \\ c_{N-1} \quad c_{N-2} \quad \cdots \quad c_N \end{pmatrix} = \mathbf{F}^{-1} \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_N \end{pmatrix} \mathbf{F} \tag{3.1}$$

where $\mathbf{F}$ is the Discrete Fourier Transform (DFT) matrix. This is simply because the appli-

cation of a circulant matrix $\boldsymbol{C}$ to an input vector $\boldsymbol{x}$ corresponds to the discrete convolution of $\boldsymbol{x}$ with $\boldsymbol{c}$, which amounts in the frequency domain to a multiplication with the Fourier Transform of $\boldsymbol{c}$.

3.1.2.2 Typical performances

As stated in [120], standard techniques used in proving CS guarantees cannot be employed for a subsampled circulant matrix, since entries of the matrix are not independent. However, different methods still allow to yield guarantees. In the following, state-of-the-art proofs of the good performances of circulant matrices for Compressive Sampling are sketched, starting with the most empirical ones.

A first contribution are the simulations presented in [123]. Performances of circulant matrices defined by one random sequence are compared toward those of matrices where all elements are randomly drawn. The sensing matrix is given by $\boldsymbol{\Phi} = \boldsymbol{RC}$, where $\boldsymbol{R}$ is a row selection matrix which can be random in different extents, as detailed in Table 3.1, last column. $\boldsymbol{\Psi}$ is the identity matrix or the Discrete Cosine Transform (DCT) matrix. Input signals are K-sparse and follow various distributions e.g. independent and identically distributed (i.i.d.) Gaussian values with randomly distributed support.

Table 3.1: Experiment settings, from [123], $\boldsymbol{\Phi} = \boldsymbol{RC}$.

$\boldsymbol{C}$-type				$\boldsymbol{R}$-type	
$\boldsymbol{R1}$	Circulant	1st row i.i.d. Gaussian	real	$\boldsymbol{P1}$	First m rows
$\boldsymbol{R2}$	1st row given	1st row i.i.d. Bernoulli			
$\boldsymbol{D1}$	Circulant	d_i random phase, $\lvert d_i \rvert = 1$		$\boldsymbol{P2}$	Equally spaced m rows
$\boldsymbol{D2}$	$\boldsymbol{C} = \boldsymbol{F}^*\boldsymbol{DF}$,	d_i random phase			
$\boldsymbol{D3}$	$\boldsymbol{D} = diag(d)$	$\boldsymbol{D1}$ without conjugate symmetry	complex	$\boldsymbol{P3}$	Randomly spaced m rows
$\boldsymbol{I1}$	i.i.d. random	i.i.d. Gaussian	real		
$\boldsymbol{I2}$	Non Circulant	i.i.d. Bernoulli			

Whether for random distributions ($\boldsymbol{I1}$, $\boldsymbol{I2}$), for circulant matrices with random diagonal elements ($\boldsymbol{D2}$) or for others described in Table 3.1, it appears that the chosen evaluation metric, the average frequency of a relative reconstruction error lower than 10^{-3}, is nearly identical for all signals. Therefore it is empirically proved that matrices tailored for signals sparse in specific bases can perform as well as their totally random counterpart, even with a non-random subsampling operator (such as $\boldsymbol{P1}$ or $\boldsymbol{P3}$). This good performance is explained by the incoherence between circulant matrices and the time domain. $\boldsymbol{D1}$ and $\boldsymbol{D3}$, which matches in fact the modeling of a Compressive Sensing Filter (CSF) (2.3.3), are moreover incoherent with any orthonormal sparsifying basis with high probability [106].

As well, in [117], simulation results show that circulant matrices based on Legendre sequences perform better than Bernoulli matrices in support recovery tasks.

More analytical considerations are exposed in [124]. If $\boldsymbol{C}$ is a unitary deterministic circulant matrix of dimension NxN with mutual coherence $\mu_m(\boldsymbol{CF}) = O(1)$, sampled ran-

domly by $\boldsymbol{R}$ of dimension MxN, then uniform recovery (cf. 1.2.3.1.f p.26) is achieved with $M \geq O(K(logN)^4)$ and non-uniform recovery with $M \geq O(KlogN)$ measurements. Uniform recovery conditions are satisfactory from a mathematical point of view but non-uniform recovery is more adapted for the purposes of this work, which is more concerned with statistical success.

As it is not easy to ensure a unitary $\boldsymbol{C}$ using the time domain approach, the authors of [124] propose to use the frequency-domain approach, i.e. generating the matrices from the chosen sequence $\boldsymbol{\sigma}$. Yet $\boldsymbol{C}$ is a unitary matrix if and only if $\boldsymbol{\sigma}$ is a unimodular sequence, i.e. the module of all coefficients is 1. It is then pointed out that for signals sparse in time and frequency, a randomly sampled deterministic sensing matrix generated in the frequency domain by a unimodular sequence $\boldsymbol{\sigma}$ has mutual coherence bounded by $O(1)$ if $\boldsymbol{\sigma}$ has unimodular Inverse Discrete Fourier Transform (IDFT), which is notably the case if $\boldsymbol{\sigma}$ is a perfect (or almost perfect in a certain extent) sequence. The definition of perfect sequences is that they have ideal correlation properties i.e. the cyclic autocorrelation function $\Gamma_{xx}(\tau)$ for a sequence x of size N and a delay τ is a centered Dirac (Zero AutoCorrelation):

$$\Gamma_{xx}(\tau) = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x^*[n+\tau] = \delta(\tau), \tau \in \mathbb{Z}$$

Theoretical guarantees regarding the required number of samples M to satisfy the Restricted Isometry Property (RIP) and high probability of non-uniform recovery are compared for the codes analyzed in [124] against earlier state-of-the-art in Table 3.2.

Table 3.2: Comparison between different circulant matrices $\boldsymbol{\Phi\Psi}$, from [124].
Note that a typing mistake happened in the original table for non-uniform recovery in random convolution.

Sensing matrix $\boldsymbol{\Phi\Psi}$	Random Convolution [106]	Partial Circulant matrix [125, 126]	Coherence bounded circulant unitary matrices [124]
Filter coefficients	Random	Random	Deterministic
Subsampling matrix $\boldsymbol{R}$	Random	Deterministic	Random
Sparsifying Transform $\boldsymbol{\Psi}$	Arbitrary	I_N	F^{-1}
Restricted Isometry Prop.	$M \geq O(K(logN)^5)$	$M \geq O(K(logN)^4)$	$M \geq (K(logN)^4)$
Non-uniform Recovery	$M \geq O(K(logN)^2)$	$M \geq O(KlogN)$	$M \geq O(KlogN)$

An improvement compared to random convolution of the number of required measurements M to satisfy RIP and non-uniform guarantees appears for the sensing matrices in the last two columns of the Table, partial circulant matrices and coherence bounded circulant unitary matrices. This improvement is explained by the fact that the mutual coherence of the random filter is only bounded by $O(\sqrt{logN})$ for an arbitrary sparsifying transform $\boldsymbol{\Psi}$. Whereas for a specific sparsifying matrix, identity $\boldsymbol{I}$ or inverse Discrete Fourier Transform $\boldsymbol{F^{-1}}$, it is known that the proposed deterministic sequences have mutual coherence $O(1)$.

In [127], the interest of circulant matrices built on perfect sequences is developed further. It is noted that perfect sequences have a constant discrete periodic spectrum according to the Wiener-Khinchin theorem. Hence their eigenvalues, which are given by the DFT of the first row, are unimodular, and the Power Spectral Density (PSD) is preserved, a quality which provides robustness to noise.

The Subsampled Circulant Matrix based Analogue Compressed Sensing (SCM-ACS) architecture proposed in [78, 80] uses randomly sampled complex circulant Zadoff-Chu codes, which are perfect unimodular sequences. One very promising result of recovery simulation is shown ([80], Fig. 4.6) in the study, however, a detailed high level metric analysis is lacking. The authors only argue that the RIP constant (see Sect. 1.2.3.1) of the sensing matrix will be the same as the RIP constant from the row selection matrix $\boldsymbol{R}$, since the matrix consisting of Zadoff-Chu codes is unitary. This is an interesting remark but not a decisive argument.

Considering the promising nature of the exposed results, properties of Constant Amplitude Zero Autocorrelation (CAZAC) sequences, and in particular the Zadoff-Chu (ZC) sequences, will be investigated in details in the following. Based on these considerations, a new MWC sensing matrix $\boldsymbol{B}$ combining the properties sketched above and an idea originating from Random Convolution is then proposed. Afterwards, a thorough benchmark of the novel sensing matrix is conducted with respect to the state-of-the-art in order to evaluate the respective recovery performances. Considering its thoroughness, this benchmark methodology is another contribution of this work,.

3.2 On the interest of CAZAC sequences for structured MWC acquisition

A specific code family, Constant Amplitude Zero Autocorrelation (CAZAC) sequences, will be introduced, and one specific instance of them, Zadoff-Chu codes, investigated further. Based on this, a proposition of innovative codes, which have the benefit of being real-valued, is made.

3.2.1 Constant Amplitude Zero Autocorrelation (CAZAC) sequences

CAZAC sequences have two distinctive features:

- unimodularity, i.e. unit magnitude (Constant Amplitude) for all elements
- ideal cyclic autocorrelation function $\Gamma_{xx}(\tau)$, a property also called being a perfect sequence:

$$\Gamma_{xx}(\tau) = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x^*[n+\tau] = \delta(\tau), \tau \in \mathbb{Z}$$

These properties are interesting because no real-valued binary perfect sequence is known [124] except the sequence $[1; 1; 1; -1]$ for $N = 4$, and therefore codes usually proposed for the MWC so far are suboptimal.

For circulant codes, only one instance is considered to generate the matrix, so the key point will be the cyclic autocorrelation rather than cyclic cross-correlations.

Also the Fourier Transform of an unimodular perfect sequences is another unimodular

perfect sequence [128], which implies that the Power Spectral Density (PSD) of the sequence is constant along the frequency. Note that an expression for the Fourier Transform of generalized chirp-like sequences is established in [129], which could be useful to study properties for wider class of codes.

3.2.2 One example: Zadoff-Chu codes

Coming from the field of Spectrum Spreading techniques, one interesting instance of such unimodular perfect codes are Zadoff-Chu codes (ZC) [130]. Zadoff-Chu, or Frank-Zadoff-Chu, sequences are complex-valued and constant envelope codes known for long time in Radar [131] and more recently for synchronization in Long-Term Evolution (LTE) mobile communications systems due to their perfect cyclic autocorrelation function. They are also used to build chirp-based orthogonal modulation dictionaries for energy-efficient long range transmission as in LoRa (**Lo**ng **Ra**nge) communications [132] or Turbo-Zadoff-Chu systems [133].
Zadoff-Chu codes are defined by:

$$ZC_R[k] = e^{-j\pi Rk(k+1)/N}, \text{ N odd} \tag{3.2}$$

$$ZC_R[k] = e^{-j\pi Rk(k)/N}, \text{ N even} \tag{3.3}$$

$$\text{for } k \in [\![0; N-1]\!], R \text{ prime to } N$$

where R is called the index of the sequence. For illustration purposes, an example of Zadoff-Chu sequences highlighting real and imaginary parts is depicted in Figure 3.1 for $N = 64$. There are as many sequences of length N as number of possible R prime to N, $R \leq N$. If N is prime, a square matrix can be generated.

Additionally, they have constant cyclic cross-correlation under certain conditions. In particular $|R_{xy}(\tau)| = \frac{1}{\sqrt{N}}$ for prime N [134].

Note also that with prime N, the Discrete Fourier Transform of a Zadoff-Chu sequence is another time-scaled Zadoff-Chu sequence [135]:

$$\tilde{ZC}_R[k] = ZC_R^*[R^{-1}k]ZC_R[0]$$

In the particular case ($R = 1$, N even), a simple expression of the DFT is derived in [136]:

$$\tilde{ZC}_R[k] = ZC_R[k]exp(\frac{-j\pi}{4})$$

3.2.3 Proposition of an original mixing code sequence

In [106], it is pointed out that if a simple Hermitian symmetry condition is respected on the elements $\sigma_{k\in\{1,\ldots,N\}}$ of a circulant matrix, it is possible to generate a real-valued circulant matrix, which is of main interest compared to [117] for instance. For even N, the condition is

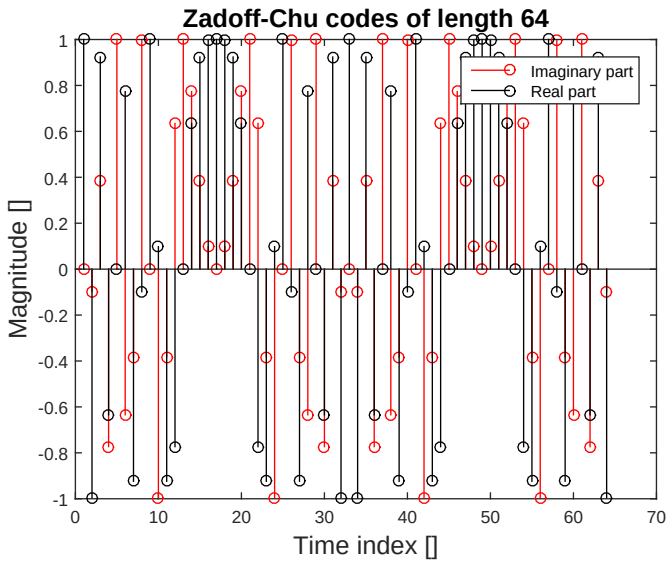


Figure 3.1: Real and imaginary part of Zadoff-Chu codes [$R = 1$, $N = 64$].

given by:

$$\begin{aligned} & k \in \{1; \frac{N}{2}+1\} && \sigma_k = \pm 1 \text{ with equal proportion} \\ & \text{for } 2 \leq k < \frac{N}{2}+1 && \sigma_k \text{ s.t. } |\sigma_k| = 1 \\ & \text{for } \frac{N}{2}+2 < k \leq N && \sigma_k = \sigma^*_{N-k+2} \end{aligned} \tag{3.4}$$

These real-valued circulant matrices can be viewed as a specific case of the Random Convolution introduced by [106] where phases of $\boldsymbol{\sigma_k}$ are supposed to be uniformly distributed. Note that preservation of the power in the frequency domain through multiplication with a unimodular sequence implies, as in the Nyquist Folding Receiver for example, a better robustness to noise because the power of no frequency subband is amplified compared to the others.

Based on previous elements, a new, real-valued, matrix which will be referred to as 'ZC circ real' and such that $\boldsymbol{A} = \boldsymbol{R}.\boldsymbol{ZC_{circ\,real}}$ is proposed in this work. That is, $\boldsymbol{A}$, the MxN code expression in the time domain, is the subsampling by $\boldsymbol{R}$ of the proposed $\boldsymbol{ZC_{circ\,real}}$. This real-valued circulant matrix is obtained as the time domain expression of a diagonal matrix in the frequency domain, diagonal which is defined by the symmetry condition (3.4) imposed on a Zadoff-Chu sequence of length $\lfloor N/2 \rfloor$, as pictured in eq. (3.5):

$$\boldsymbol{ZC_{circ\,real}} = \mathbf{F}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & ZC_R[2] & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & ZC^*_R[2] \end{pmatrix} \mathbf{F} \tag{3.5}$$

Note that despite the similarity of performing pseudo-random dephasing, this is intrinsically different from the Random Convolution (see 2.3.3), as dimension reduction is performed in the frequency domain for a Modulated Wideband Converter through bandpass whereas the

dimension reduction in the Random Convolution is performed by subsampling the time dimension.

The matrix referred to as 'ZC circ tmp' on the other hand consists in the circulant matrix defined by one complex-valued Zadoff-Chu sequence as the first row in the time domain.

In the next Section, proposed codes are compared against codes from the state-of-the-art according to various criteria.

3.3 Analysis of the reconstruction performances of the sensing matrix

Different codes will be evaluated based on a thorough comparison methodology including both theoretical and practical metrics. First, the high-level aspects of coherence and isometry properties of the sensing matrix $\boldsymbol{\Phi\Psi}$ are investigated. Then performances are confirmed by means of a simulation platform mimicking the MWC acquisition and reconstruction process.

3.3.1 Coherence

This analysis begins with the comparison of the coherence of matrices $\boldsymbol{\Phi\Psi}$ typically benchmarked in classical Compressive Sampling (Bernoulli [81], Gold [6]) and more advanced technics (a circulant Zadoff-Chu code in the time domain 'ZC circ tmp' [78, 80], our original proposition 'ZC circ real'). Coherence is essential because, as seen on p.26, it is proportionally related to the degree of sparsity K that can be handled and hence with the number of required branches M.

Comparisons will be carried out for a specific architecture, the Modulated Wideband Converter (MWC), and therefore the sensing matrix to be analyzed is not solely the temporal code pattern $\boldsymbol{A}$. Indeed, as detailed in Appendix B, the sensing matrix for the MWC framework is given by $\boldsymbol{B} = \boldsymbol{A}\bar{\boldsymbol{F}}\boldsymbol{D}$ where $\bar{\boldsymbol{F}}$ is a reordered Discrete Fourier Transform matrix, $\boldsymbol{A}$ of dimension MxN matches the code in the time domain and $\boldsymbol{D}$ is a diagonal matrix accounting for the decay of the code's Fourier transform at high frequencies. Therefore the analysis must be lead on $\boldsymbol{B} = \boldsymbol{A}\bar{\boldsymbol{F}}\boldsymbol{D}$. In [35] and [117], it is pointed out that if a random $\boldsymbol{A}$ is considered, $\boldsymbol{D}$ and the permutation matrix of $\bar{\boldsymbol{F}}$ can be ignored in the analysis because the input signal $\boldsymbol{z}$ or $\boldsymbol{Dz}$ would have the same sparsity. In fact $\bar{\boldsymbol{F}}$ cannot be ignored with deterministic codes but is equivalently replaced by $\boldsymbol{F}$, as sorting colmuns does not matter. And $\boldsymbol{D}$ can be ignored for the coherence because of the column normalization taking place which cancels the effect of $\boldsymbol{D}$. Hence for the computation of the coherence, $\boldsymbol{AF}$ will be considered.

The choice of a code family is an important part, but the choice of $\boldsymbol{R}$, the subsampling of the square matrix must not be neglected either. The question of whether it is possible to improve performances by a specific and non naive row selection within the square sensing matrix will also be investigated.

- Publications [5, 16, 33, 123, 137] often stay on the theoretical level and use random selectors, typically Gaussian or Bernoulli with zero mean. Those approaches will be pictured with a uniform distribution (noted 'rand').
- Naive selectors are taking the first lines (noted 'fl') or performing regularly spaced subsampling (noted 'sub').
- Another idea in this study was to test a sparse ruler row selector on circulant matrices. A set $\mathbf{\Omega} \subset \{0, \ldots, L-1\}$ is a length $(L-1)$ sparse ruler [138] if for every $l \in \{0, .., L-1\}$, there exists at least one pair of elements (k, k') in $\mathbf{\Omega}$ satisfying $k - k' = l$. Thus a sparse ruler guarantees that all distances are represented as a difference between two elements of the selected set. Therefore it allows to capture all dephasings through a dimensionally reduced matrix and this might perform better. Wichmann rulers [139] are conjectured to be optimal rulers and are easy to generate so they were chosen (noted 'Wh.'). However, they do not exist for all (M, N) value, only very few values can be plotted for small values of M/N.
- Exploiting the asset that coherence is easy to compute, a method for choosing a matrix with very good coherence with the help of random selectors is also proposed. 1000 selectors are generated and for each value of M, the matrix with best coherence (noted 'stat') and its coherence value are extracted. Since it is much easier to generate circulant matrices than regular ones, they should be preferred if it is shown that they also perform better.

The coherence is computed for different values of M. The optimal Welch lower bound (cf 1.2.3.1) is drawn in dashed blue. Note that for clarity, in all figures the color represents the code nature and the pattern represents the row selector $\boldsymbol{R}$.

Gold codes

The study begins with Gold codes (brown), which are currently the state-of-the-art. Coherences curves are presented in Figure 3.2

All curves for sensing matrices based on Gold codes have approximately the same profile, the selection process does not seem to be a highly relevant element regarding the coherence.

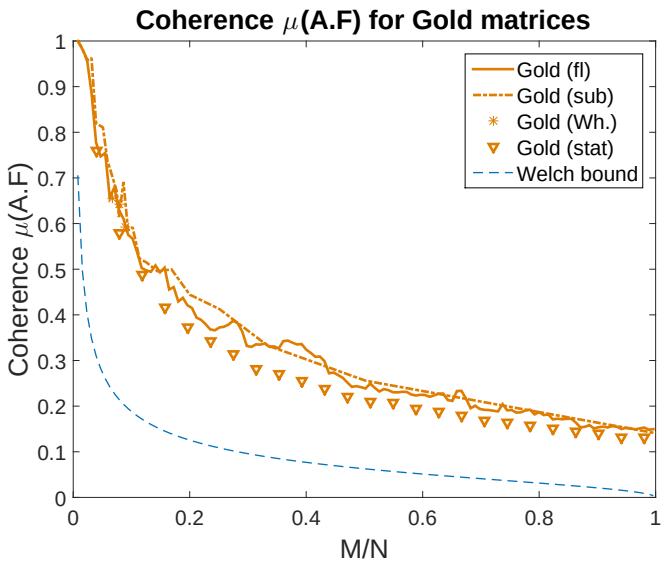


Figure 3.2: Coherence $\mu(\boldsymbol{AF})$ comparison for Gold code matrices $[N = 255]$.

Zadoff-Chu codes

The study is extended to Zadoff-Chu codes (dark blue)(ZC) in Figure 3.3.

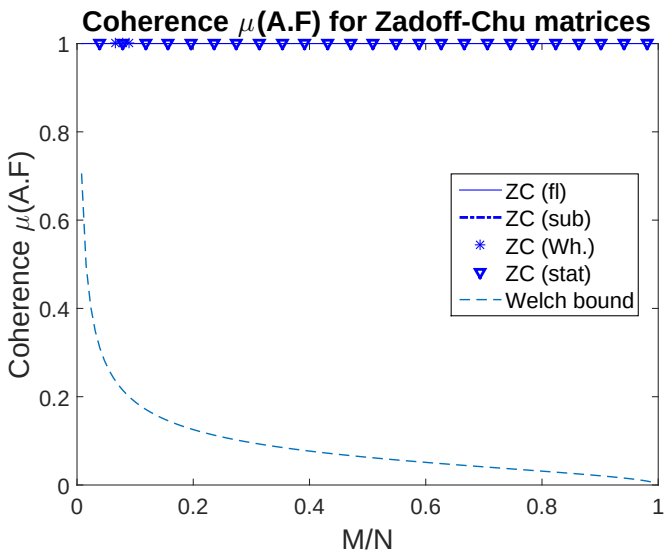


Figure 3.3: Coherence $\mu(\boldsymbol{AF})$ comparison for Zadoff-Chu codes $[N = 127,\ R \in [\![1; N]\!]$, prime N behavior$]$.

Non-circulant Zadoff-Chu matrices exhibit the worst possible performance, meeting the $\mu = 1$ limit. The explanation lies in the symmetry appearing on the antidiagonal (for sequence indices k and $k' = N - k$) of the Gramian of the matrix featuring the Zadoff-Chu code 'ZC circ tmp'.

The coherence of circulant Zadoff-Chu matrices (green) is studied as well, either for a Zadoff-Chu code in the time domain 'ZC circ tmp' or for Zadoff-Chu-based real-valued matrices 'ZC circ real', pictured in Figure 3.4. In constrast to Zadoff-Chu codes with different sequence

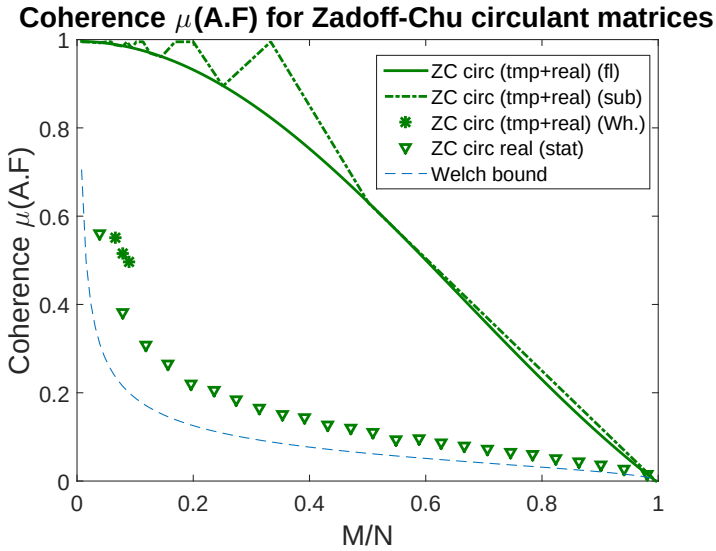


Figure 3.4: Coherence $\mu(\boldsymbol{AF})$ comparison for Zadoff-Chu circulant codes $[N = 255,\ R = 1]$.

indices R, diagonal and antidiagonal coefficients of the Gramian for circulant codes 'ZC circ tmp' and 'ZC circ real' are not equal, so the coherence is lower than 1. The second observation in Figure 3.4 is that the proposed Zadoff-Chu-based real-valued matrices perform exactly the same as complex-valued circulant matrices based on a Zadoff-Chu code in the time domain, hence the notation 'tmp+real' in the legend because the curves are superimposed. This validates that good circular correlation properties of the Zadoff-Chu sequences are preserved by the symmetry condition on the diagonal.

Then it appears that circulant Zadoff-Chu matrices selected through a Wichmann ruler outperform Gold codes. However, the choice of the size is limited and this adds a significant constraint. Zadoff-Chu circulant matrices selected with the best random selector, whether for a pattern defined in time ('ZC circ tmp') or as the half diagonal in the frequency domain ('ZC circ real'), yield similar values. They have significantly lower coherence than Gold codes and also better than the Wichmann ruler, which is very promising.

Summary

Intuitions based on analytical considerations established in the previous part are therefore confirmed by the summary graph pictured in Figure 3.5.

Zadoff-Chu-based matrices confirm that they have better coherence than matrices from the state-of-the-art. Statistical selection proved to be the best method, capitalizing on the advantage that the coherence is quickly computable. If these results are confirmed by the other metrics of interest, a good strategy would be to pick selection schemes with good coherence properties before implementation. The outcomes are encouraging because they confirm the

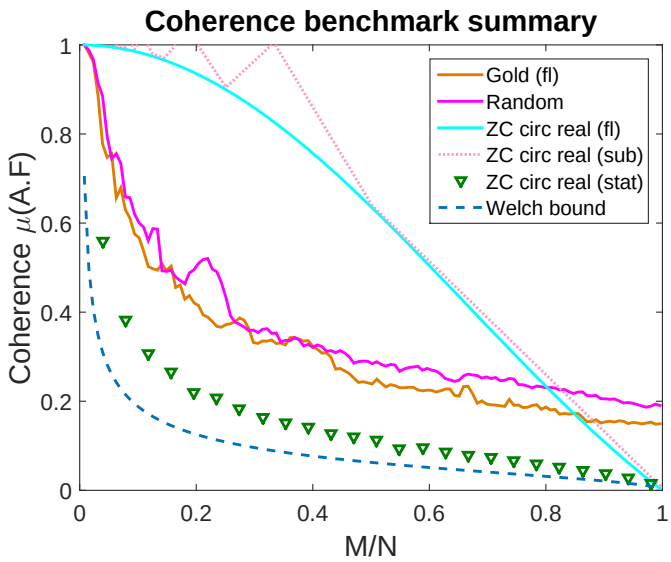


Figure 3.5: Coherence $\mu(\boldsymbol{AF})$ comparison [$N = 255$, $R = 1$], from [113].

quality of circulant matrices that have the advantage of being easy to compute, store and manipulate.

After the coherence, another vital high-level property of the sensing matrix, preservation of norms and distances, will be evaluated in the next subsection.

3.3.2 Isometry

As evoked in 1.2.3.1, preservation of norm and distances (isometry) plays a major role in Compressive Sensing and in particular with respect to the noise robustness aspect [25, 26]. However, it has been mentioned that the Restricted Isometry Property (RIP) constant is NP-hard to compute. Hence two strategies are developed in this work to approach it: first a statistical evaluation of δ_K, and second, computing a theoretical guarantee under additional assumptions, the Expected RIP.

3.3.2.1 Statistical estimation

The empirical Restricted Isometry Property (RIP) constant δ_K which describes norm preservation of K-sparse vectors and the empirical Johnson-Lindenstrauss Lemma (JLL) constant δ'_K which describes distance preservation between K-sparse vectors are evaluated through statistical estimation. The JLL constant of order K corresponds to the RIP constant of order $2K$, hence δ'_K might also be noted and computed as δ_{2K} but in this section it will be computed as the constant governing the preservation of distances. Such statistical simulations are known to miss specific pathological cases [37] and be overoptimistic. However, it could be expected that the maximum δ_K and δ'_K encountered in practice will almost always be close

to the maximum δ_{N_v} estimated over N_v vectors.

For the RIP constant, 10^7 test vectors are generated, subdivided into $N_s = 1000$ subsets of $N_v = 10000$ vectors. Each vector of length 127 has $K = 10$ non-zero values, uniformly distributed on the support and with values uniformly distributed on $[-0.5; 0.5]$. Then vectors are projected with a sensing matrix of dimension 50×127 to study the variations of the norm of K-sparse vectors through projection.

For the JLL constant, $N_s = 1000$ subsets of $N_v = 200$ vectors ($N_s.\frac{N_v(N_v-1)}{2} = 19.900.000$ distances) are similarly generated to study the variations of the norm of distances between K-sparse vectors through projection.

To get an estimation of the variation of the estimation for different input vector sets, histograms are established, depicting the values taken by δ_K in Figure 3.6 for RIP and by δ'_K in Figure 3.7 for JLL for the subsets N_s. It is reminded that the overall Restricted Isometry Property (RIP) constant δ_K is given by the worst value over all input vectors from all subsets. Following quantities are reported:

- the Restricted Isometry Property (RIP) constant δ_{N_v,N_s} and JLL constant δ'_{N_v,N_s} per subset
- the average $E_{N_s}(\delta_{Nv}) = \frac{1}{N_s}\sum_{N_s}\delta_{N_v,N_s}$ of the estimation of δ_K and δ'_K over the different N_s subsets
- the standard deviation $\sigma_{N_s}(\delta_{Nv}) = \sqrt{\frac{1}{N_s}\sum_{N_s}(\delta_{N_v,N_s} - E_{N_s}(\delta_{Nv}))^2}$ of the estimation of δ_K and δ'_K over the different N_s subsets

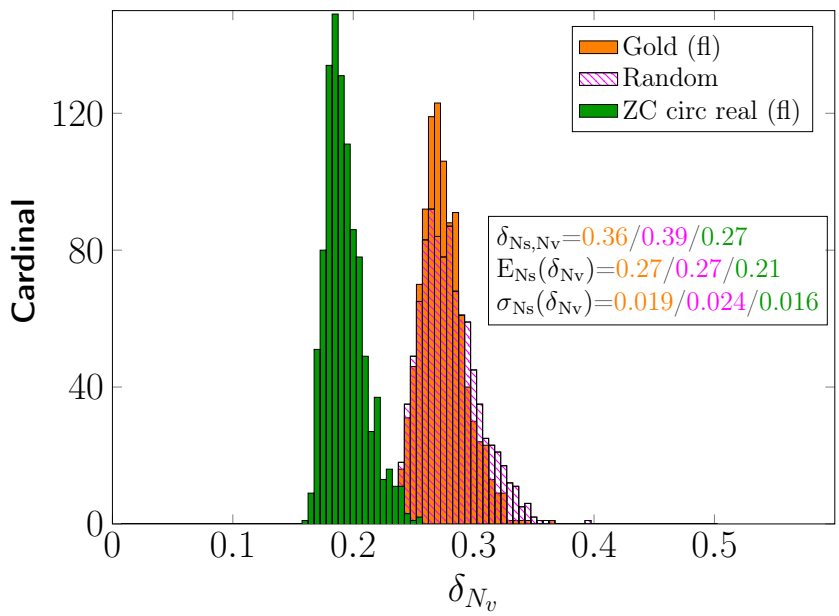


Figure 3.6: Histogram of RIP-δ_K estimation [$N_s = 1000$, $N_v = 10000$, $K = 10$], from [113].

The standard deviations being very small compared to the mean values, the estimation methodology is validated. It appears in Figure 3.6 that RIP-δ_K is nearly 1.5 times smaller

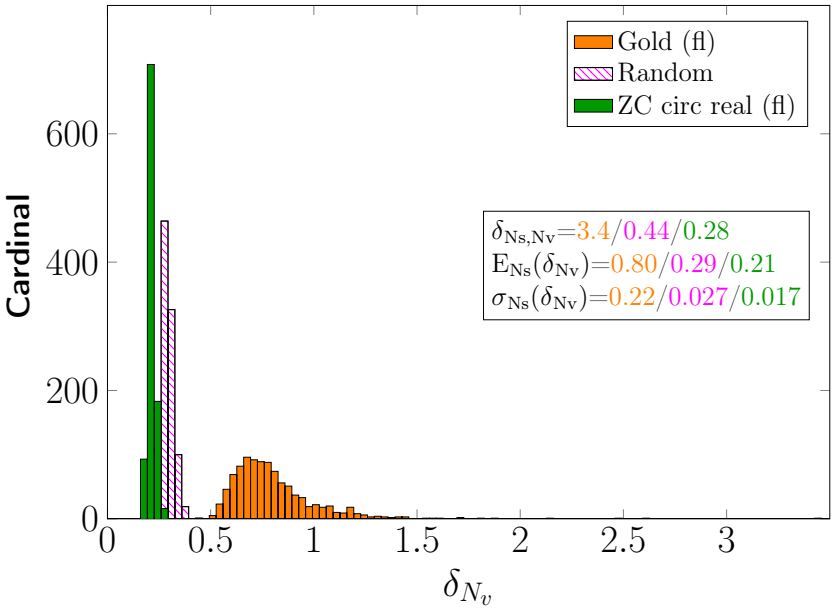


Figure 3.7: Histogram of JLL-δ'_K estimation [$N_s = 1000$, $N_v = 200$, $K = 10$], from [113].

for real-valued circulant sensing matrices based on Zadoff-Chu codes than Gold and Random codes (0.27 instead of 0.36 and 0.39). It appears in Figure 3.7 that JLL-δ'_K is more than 1.5 smaller for circulant sensing matrices based on Zadoff-Chu codes than Random codes and more than 12 times smaller than Gold codes (0.28 instead of 0.44 and 3.4). These results mean that one can reasonably rely on a small norm and distance preservation bound for circulant Zadoff-Chu-based matrices, in comparison to the codes from the state-of-the-art.

3.3.2.2 Isometry under additional assumptions: Expected RIP

The Expected RIP (ExRIP) criterion introduced in Subsection 1.2.3.1 expresses the probability P that a matrix satisfies the Restricted Isometry Property (RIP), assuming an uniform distribution of the support and random distribution of non-zero values. To the best of the author's knowledge, ExRIP guarantees are only established for real-valued codes, hence complex-valued Zadoff-Chu circulant codes can not be compared. Table 3.3 shows a comparison of ExRIP probability P computed for various codes ("This work"), benchmarked with results from [35]. Note that the value $\delta_{2K} = \sqrt{2} - 1$ is presumably chosen because it corresponds to the value guaranteeing exactitude of the convex relaxation of $l0$ to $l1$ [24].

Table 3.3: ExRIP [$M = 80$, $N = 511$, $K = 24$, $\delta_{2K} = \sqrt{2} - 1$, $R = 1$].

Code		ZC circ real (fl)	ZC circ real (stat)	Random	Gold (fl)	Hadamard
P	This work	0.9498	0.9511	0.9270	0.9405	0
	[35]	–	–	0.927	0.939	0

The benchmark for random, Gold and Hadamard codes is consistent. Note that parameters of [35] are more detailed in the corresponding technical report [119], but not completely, which explains the thin discrepancy for Gold codes with different settings.

Thus three remarks: Hadamard codes are absolutely not recommended, the poor probability of satisfying the RIP is caused by high values of the coefficients β and γ (see Table 1.1) due to many peaks in their crosscorrelation function. Zadoff-Chu-based circulant codes perform better than all other analyzed codes, e.g. the probability $(1 - P)$ that RIP is not met is decreased by almost a third compared to random codes. Also, for Zadoff-Chu-based circulant codes, a statistical best coherence row selector ('stat') is slightly more effective than first row selection ('fl'). This performance can be related to their known good correlations properties.

3.3.3 Empirical validation of the reconstruction performances by means of simulation platform

The aforementioned metrics (coherence, RIP, ExRIP) are not entirely sufficient to conclude on overall restitution performances as they only shed light on specific theoretical aspects of the acquisition process. Hence it is proposed to confirm or not the good properties outlined in the theoretical framework by means of simulation of the Modulated Wideband Converter (MWC) architecture. A MATLAB® platform has been built, inspired from the code available in [116, 140]. The purpose is to validate the choice of the mixing sequences based on practical evaluation metrics through Monte-Carlo simulations.

3.3.3.1 Simulation framework

The simulation platform that is used enables to picture the functioning of the MWC architecture more precisely than just using the sensing matrix: in particular, it accounts for the specificity of finite unions of bandpass signals model, which enables to describe the process of acquisition on the analog representation of the frequency subbands.

To represent numerically an analog hardware, especially the non-bandlimited signal after analog mixing, the time grid is oversampled by a factor *Rep*. In order to mimic the analog filtering and sampling steps, an oversampled digital Finite Impulse Response (FIR) filter followed by decimation at the adequate factor and removal of the delay is used. However, more advanced realism such as non-ideality of the low-pass filters, non-linearity of the mixer or clock jitter is not tackled. Parameters are set as defined in Table 3.4.

Similarly to [116, 117, 78] and others, the input signal is given by:

$$x(t) = \sum_{k=1}^{K} \sqrt{E_k B} sinc(B(t - \tau_k)) cos(2\pi f_k (t - \tau_k)) + w(t) \tag{3.6}$$

where τ_k and f_k are chosen uniformly at random and other parameters according to Table 3.4. That is, each active band is modeled as a 'sinc' in time domain/window in frequency domain,

Table 3.4: Parameters of the MWC simulation.

	Name	Unit	Specification	Default value	Variable
Architecture	N	[]	Number of bands before downconversion =Code length	127	Settings Settings $2^k - 1$
	f_p	Hz	Width of a band in the signal model = 2x filter cut-off frequency	$78.74MHz$	Settings
	M	[]	Number of channels	50	Settings
Signal	K	[]	Degree of sparsity	6	Settings
	B	Hz	Maximal bandwidth of the input signal	$78MHz$	Settings $B \leq fp$
	ISNR	dB	Global SNR	NS	Settings
	E_k	J	Energy in the active band k		Settings = 1
	f_k	Hz	Centralized frequency of the band k		Variable, random
	τ_k	Hz	Dephasing of the kth band		Variable, random
	f_{Nyq}	Hz	2x the maximal frequency of the signal model	$1GHz$	$L.f_p$
	L	[]	Signal length coefficient	128	Settings
	Rep	[]	Oversampling factor	10	Settings
Reconstruction: Orthogonal Matching pursuit (OMP)					

centered around the randomly-drawn frequency f_k. w(t) is white Gaussian noise, scaled such that the test signal has the desired Signal-to-Noise Ratio (SNR).

Let us note with $\mathbf{\Lambda}$ the true support and with $\hat{\mathbf{\Lambda}}$ the identified support. As in the original platform [140], a variation in the reconstruction process is employed for denoising: In the Continuous-to-Finite (CTF) reconstruction block, thresholding is applied on negligible eigenvalues of the frame constructed from the samples, in order to remove at best the noise space. Hence success is declared for $\mathbf{\Lambda} \subset \hat{\mathbf{\Lambda}}$ and linear independence of the columns $\hat{\mathbf{\Lambda}}$ of the sensing matrix instead of the more straightforword condition $\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}$.

The platform has been validated step by step concerning various spectral aspects:

- First, that the frequency are translated as expected
- Then that the intermediate spectral aspects of the coding sequences match the expectations.
- Last, that the spectrum components present at baseband have the amplitudes corresponding to those of the Dirac they were expected to be coming from.

  In Figure 3.8, it is first verified that the expression of the spectrum of the sensing matrix matches the analytical expectations, in particular the effect of windowing ('Windowed Repeated Code Spectrum') and artificial upsampling ('Non Repeated Upsampled Code Spectrum'). The magnitude of the 13th coefficient of the spectrum of the shaped mixing codes is $8.6dB$. Then the low-passed spectrum of the compressed samples is measured

and indeed a magnitude at the active frequency $Magnitude(f = f_i[f_p]) = 8.6dB$ is found (the energy of the signal spectrum has been normalized at the earlier stages).

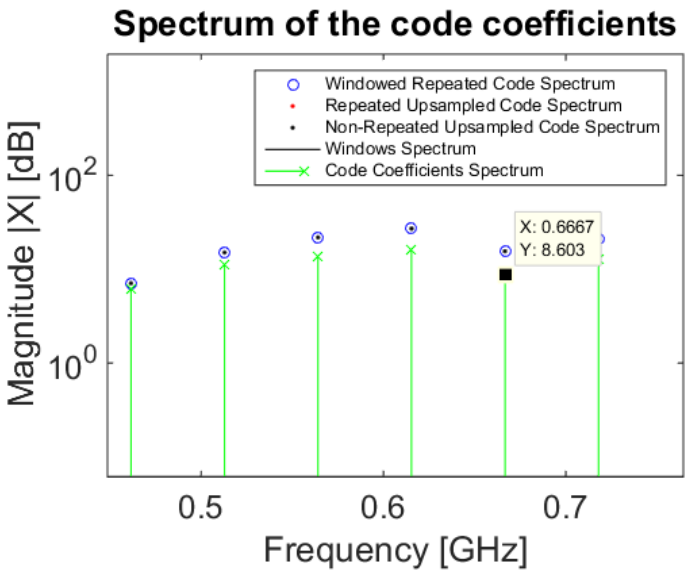


Figure 3.8: Validation of the spectrum of the mixing codes $[K = 2, B = 5MHz]$.

Evaluations of the support identification performances will be carried out, first in noiseless then in noisy settings.

3.3.3.2 Noiseless scenario

Accuracy graph

The transition graph which pictures the percentage of successful support recovery for varying relative number of measurement M/N (x-axis) and relative sparsity degree K/N (y-axis) is a metric frequently used because, for geometrical reasons developed in [36], the transition graph shows two distinctive phases: a success phase, where exact reconstruction typically occurs, and a failure phase, where exact reconstruction typically fails. Results strongly depend on the chosen reconstruction algorithm [141].

In the first place, for faster comparison accuraccy graph which shows the evolution of accuracy depending of the compression ratio are established for the Orthogonal Matching Pursuit (OMP) algorithm. An accuracy graph is a cut at constant K of a transition graph. The accuracy, on the y-axis, is given by the fraction of the trials with successful support identification. An accuracy graph is represented in Figure 3.9 for a noiseless environment.

Gold codes, Zadoff-Chu-based circulant codes with statistical selection and Random Bernoulli codes show similar good accuracy performance. Additionally, Zadoff-Chu-based codes (stat) show a success rate increasing already for higher compression ratio (0.075 instead of 0.1). For Zadoff-Chu-based circulant codes with first rows selection, which it is reminded

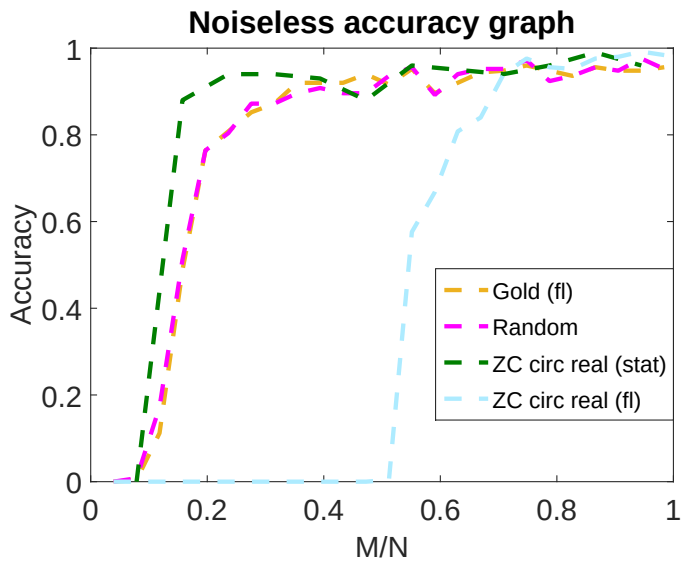


Figure 3.9: Accuracy of the MWC with growing compression ratio [$N = 127$, $K = 6$, $R = 1$, 200 trials, OMP reconstruction algorithm].

had lower coherence than state-of the-art and 'ZC circ real (stat)', the success rate increases only at small compression ratio (> 0.5). So it seems, as expected, that the required number of measurements is affected by coherence properties, which for Zadoff-Chu are dependent on the selector and can be better than the state-of-the-art.

Detection graph

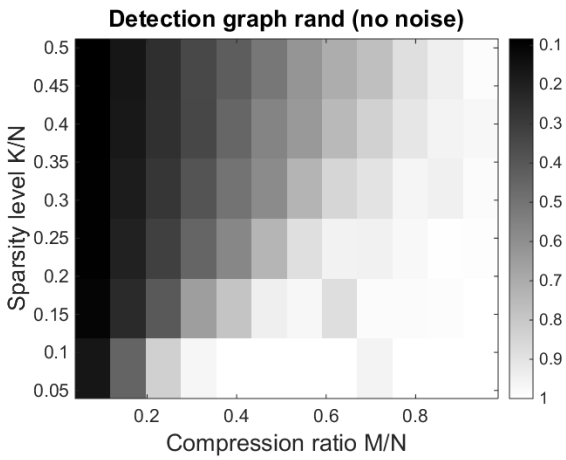
The last step of the reconstruction process is the operation of pseudo-inverting the restriction of the sensing matrix to the support. Hence transition graphs are essential for reconstruction tasks because, if there is a single error in the support estimation, it will yield results that are completely off the mark. However, if the aim is to perform detection, estimation or classification, it is rather desirable to know how often errors appear overall.

Hence the concept of detection graph, which represents the percentage of support elements successfully identified for a given sparsity level K and number of branches M, will be introduced. The true detection rate is the probability that a bin of the support is identified correctly, $\Pr(f \in \hat{\mathbf{\Lambda}} | f \in \mathbf{\Lambda})$. More formally, the true detection rate is given by:

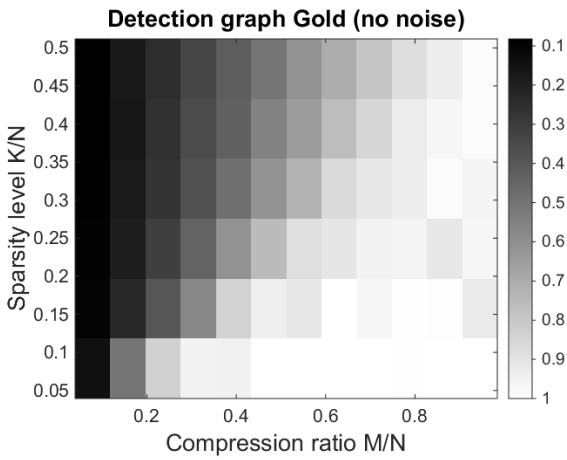
$$TD = \frac{card(\mathbf{\Lambda} \cap \hat{\mathbf{\Lambda}})}{card(\mathbf{\Lambda})} \tag{3.7}$$

The chance of true detection rises at lower sparsity levels. Indeed if $\mathbf{\Lambda}$ is estimated uniformly at random, the probability for one element of $\hat{\mathbf{\Lambda}}$ to belong effectively to $\mathbf{\Lambda}$ is K/N. At $K/N > 0.5$ for example, it is obvious that the intersection of both ensembles can not be empty and thus there is some success independently of the algorithm. This effect does not transpire with transition graphs which requires a perfect support identification to consider a success.

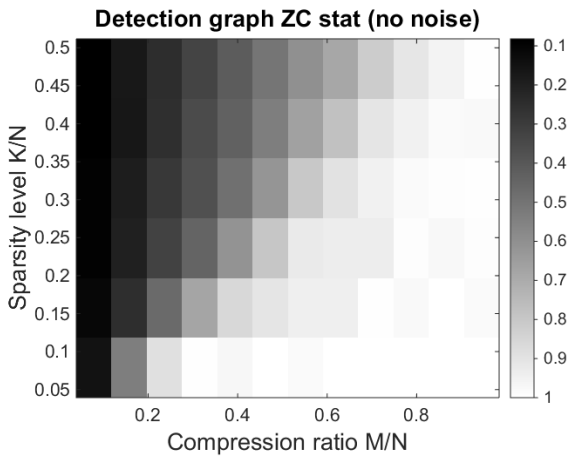
Figure 3.10 pictures the detection graph for circulant matrices based on Bernoulli, Gold and real-valued Zadoff-Chu codes defined on the frequency domain for 250 averaging trials.



(a) Bernoulli



(b) Gold



(c) Real-valued Zadoff-Chu-based circulant

Figure 3.10: Detection graph for different codes [$N = 255$, $R = 1$].

The transition between favorable (light) and insufficient (dark) settings appears relatively clearly on all three figures. The comparison between the graphs shows that with increasing sparsity degree K, Zadoff-Chu-based circulant codes keep a slight advantage over Bernoulli codes.

To interpret more precisely the gray levels, it is looked at the frontier (TD=80%). In order to establish if this difference is significant, the bound of the 95% confidence interval (i.e. there is 95% chance that the true value lies between the bounds under assumption of a normal distribution of the estimations) is plotted in Figure 3.11. At 95%, it is given by:

$$\text{Confidence interval}_{95\%} = [\frac{\hat{x} - (1.96\sigma)}{\sqrt{N}}; \frac{\hat{x} + (1.96\sigma)}{\sqrt{N}}] \tag{3.8}$$

where σ is the standard deviation of the estimator.

The transition appears quite clearly for all three codes and the measurement inaccuracies

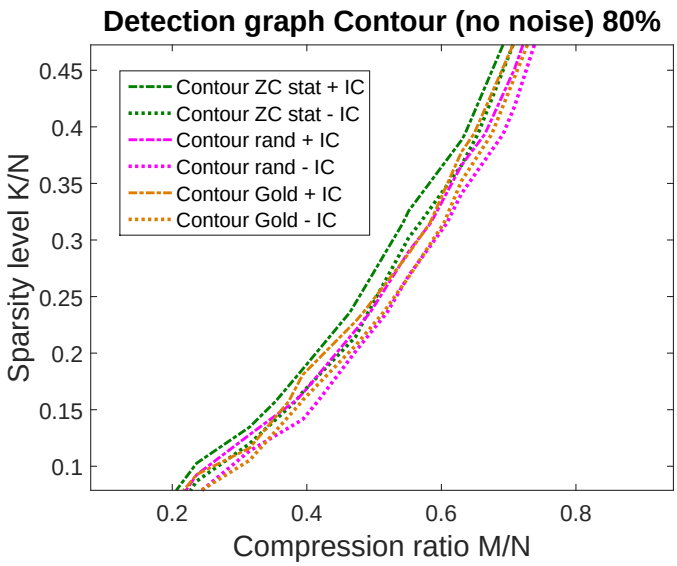


Figure 3.11: Contour of the 95% confidence interval on the 80% true detection boundary of 3 previous figures, IC=Confidence interval.

are closely bounded. The zone of detection $< 20\%$, under the curves, remains quite identical for all codes. Still, note that there is a slight advantage to 'ZC circ real (stat)'.

Now, the more realistic scenario of a noisy environment will be considered.

3.3.3.3 Noisy scenario

Accuracy graph

The accuracy graph is represented in Figure 3.12 for a noisy environment: $10dB$ Input Signal to Noise Ratio (ISNR) as defined in [83].

In the noisy context, Gold codes performance collapses (at $M/N = 0.5$, 70% loss) whereas for Zadoff-Chu-based codes it degrades less (plateau at 0.85) than random codes (plateau at 0.7). Performance in a noisy context fits therefore our analysis of isometry properties: Zadoff-Chu-based codes are more resilient to noise than other analyzed codes.

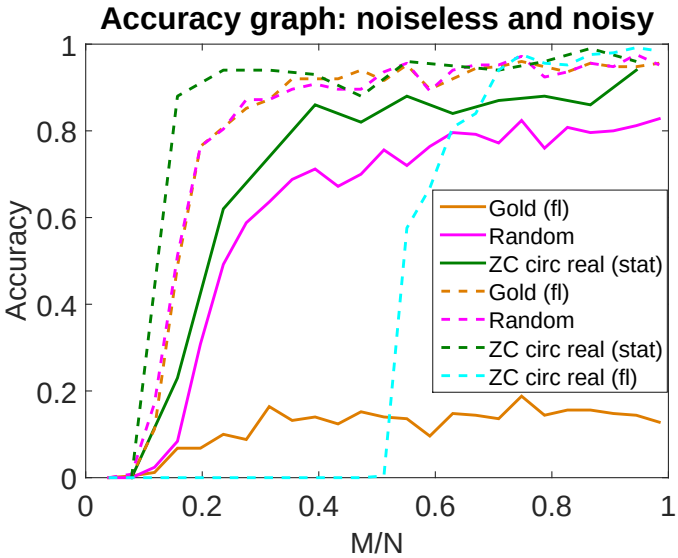


Figure 3.12: Accuracy graph of the MWC with growing compression ratio [$N = 127$, $K = 6$, $R = 1$, 200 trials, OMP reconstruction algorithm], dashed=noiseless, plain=10dB ISNR, from [113].

Detection graph

Adding noise has the effect of making the detection graph darker and the transition phenomenon more blurry hence the contour graph at 80% is directly shown in Figure 3.13

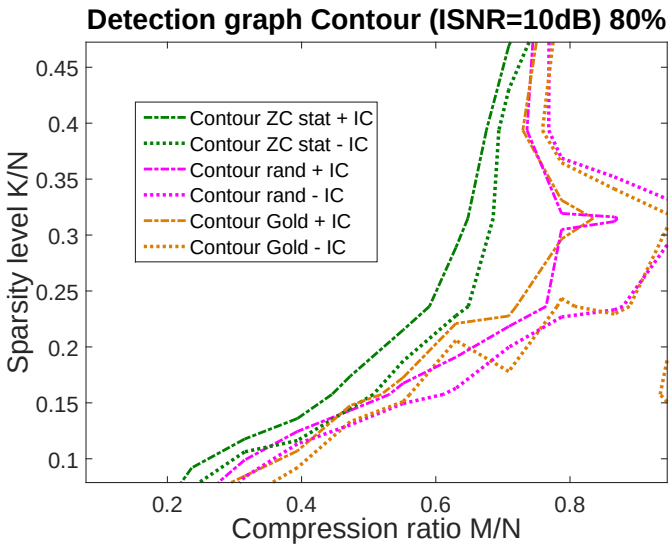


Figure 3.13: Contour of the 95% confidence interval on the 80% true detection boundary of noisy detection graph.

The zone defined by true detection over 80% is wider for Zadoff-Chu-based codes than for random and Gold codes, even if the worst estimation error is assumed. It shows that Zadoff-

Chu codes have significantly better support identification performance than the two other codes in this settings, and therefore that they have better resilience to noise. In contrast to the noiseless setting (Figure 3.11) where it was difficult to conclude, the difference is statistically significant.

Less clear is, however, why for a few settings with very high sparsity level and very high compression ratio, the chances of true detection of one support element are better. One lead is that with growing sparsity degree K, the chance of finding randomly one element of the support also grows (K/N).

3.4 Conclusions and Perspectives

First, a quick state-of-the art of mixing code sequences that are used in the Modulated Wideband Converter was sketched. The meaning of notions such as randomness and universality for matrices implemented as a RF receiver in hardware was discussed. In the light of these reflexions, the interest of using circulant matrices regarding both performances and ease of implementation was underscored, based on the literature. Then a proposition of a new code matrix was made: this real-valued circulant matrix is determined by its diagonal in the frequency domain where a symmetry condition is used on a half-length Zadoff-Chu sequence. A thorough benchmark of the respective interest of our proposition and matrices from the state-of-the-art was conducted, based on both high-level metrics such as coherence and isometric properties and on a simulation platform mimicking the acquisition and restitution process of a MWC receiver. The different elements brought to light are summarized in Table 3.5.

Table 3.5: Conclusion table.

Metric	Goal	Results
Coherence	Minimal number of measurements or lower sparsity level handleable	.Row selection is decisive for circulant codes .Circulant codes with statistical selection perform better than Bernoulli and Gold codes
ExRIP (Theoretical)	Isometry for noise robustness	ZC circulant codes are the best, they are better with 'rand' or 'stat' selector
Estimated RIP (Practical)	Isometry for noise robustness	Reliably smaller constants for ZC circulant codes
Accuracy	Validation of the concrete recovery performance	
.Noiseless settings		No significant difference slight advantage to ZC circ stat
.Noisy settings	Noise resilience	ZC circ stat best, Gold → strong degradation
Detection graph	for partial restitution Confirm results at lower sparsity	
Noiseless settings		No significant difference
Noisy settings	Noise resilience	ZC circulant stat best

For circulant codes, it is shown that the coherence is essentially determined by the row selector $\boldsymbol{R}$ and not by the nature of the code ('ZC circ tmp' or 'ZC circ real'). Selection by taking the best coherence out of matrices from a given type yields the best coherence result overall, better than the state-of-the art. The selection by a sparse ruler which performs worse than Zadoff-Chu codes but also better than state-of-the-art, might be an alternative. However it is not recommended, as it is more size-constrained.

Regarding estimations of the isometry properties, conclusions emerge as follows. Concerning the Expected RIP criteria, Zadoff-Chu-based circulant real-valued codes perform better than Gold codes or any other considered codes. Concerning statistical RIP and JLL estima-

tion, Zadoff-Chu circulant codes, especially real-valued ones, also perform better than Gold codes or any other considered codes. Additionally, for accuracy at sparsity $K = 6$, it appears that Zadoff-Chu-based circulant codes are much more resilient to noise. These results are then confirmed and extended by detection graphs for varying sparsity levels, regarding performances in both noiseless and noisy environment.

Some questions remain open:

1. Implementation aspect: How to implement these codes easily? What is the best index R of Zadoff-Chu sequence for the circulant matrix ? And notably, is it possible to optimize the use/generation of complex codes in the QAIC ?

2. Extension of the results to feature extraction: Are these conclusions also valid for estimation and classification tasks ?

To the light of the review lead in Chapter 2, the NUWBS solution showed promising results, with respect to implementation considerations [98] [99]. To tackle the critical points 1. and 2., we suggest in Appendix D to analyze in details the benefits of ZC sequences mapped to the principle of a NUWBS solution. Some implementation considerations have also beeen analyzed, which are not reported in this manuscript for the sake of intellectual property concern. This developments lead to a patent deposit [142]. The purpose is to examine further the potential of flexibility of the NUWBS.

As mentioned earlier, reconstruction of the original signal is too energy-consuming for embedded radio solutions, and hence in this work feature extraction directly from the compressed samples is considered. Since the principles and struggles governing the conception of compressive receiver architectures are better understood, by now the second aspect of this work, compressive spectral parametric estimation without reconstruction, should be addressed. This will be done in the next chapter where new theoretical bounds on the estimation error are established. In order to depict the influence of Compressive Sensing on spectral parametric estimation, compressive multiband architectures will be considered, with the particular example of the MWC in mind. More precisely, a closed-form expression of the Cramér-Rao bounds on the variance of any unbiased estimator for multiband architectures with a disjoint spectral subband model will be demonstrated. These results will allow to distinguish the effect of subsampling and diversity on the coupling between parameters. It will then be illustrated on a typical scenario of Cognitive Radio, interferer detection.

Chapter 4

Spectral parameter estimation with compressive multiband architectures

Contents

4.1 Cramér-Rao bounds for spectral parametric estimation with compressive multiband architectures . . . 91

4.1.1 Original theorems: Fisher information matrices for different sampling schemes under assumption of a disjoint subband spectral model . . . 91

4.1.2 Discussions . . . 93

4.2 Applications to cognitive radio: definition of a framework . . . 94

4.3 Amplitude estimation . . . 95

4.3.1 Expressions of the Fisher information matrices . . . 96

4.3.1.1 Nyquist Sampling . . . 96

4.3.1.2 Bandpass Sampling . . . 97

4.3.1.3 Multiband Compressive Sampling . . . 98

4.3.2 Interpretation and discussions . . . 99

4.3.3 Maximum Likelihood Estimator (MLE) . . . 99

4.3.3.1 Theoretical expressions . . . 99

4.3.3.2 Experimental variance benchmark . . . 101

4.4 Frequency estimation . . . 102

4.4.1 Expressions of the Fisher information matrices . . . 102

4.4.1.1 Nyquist Sampling . . . 102

4.4.1.2 Bandpass Sampling . . . 102

4.4.1.3 Multiband Compressive Sampling . . . 103

4.4.2 Illustration and interpretation of the analytical expressions . . . 103

4.5 Influencing the Gramian . . . 104

4.5.1 Interest of exploiting non-uniform spectrum distribution assumptions . . . 105

4.5.2 Illustrative example . . . 105

4.6 Conclusions . . . 107

As outlined in Chapter 1, the reconstruction of the original signal is considered to be computationally too expensive with regard to the stringent power constraints of embedded radio solutions. This observation has lead to the proposition of performing feature extraction directly from the samples acquired with a compressive radiofrequency architecture, an approach whose implications in terms of performances and design were only briefly tackled in previous Chapters. That is why, in this Chapter, specificities of parametric estimation from the compressed samples, as opposed to parametric estimation on the signal reconstructed from the compressive samples, are investigated.

As far as spectral information is concerned, the estimation of the parameters governing the underlying signal model is mostly based on Nyquist rate sampling since folding must be avoided to recover the information. Parametric estimation directly from the compressed samples, on the other hand, differs strongly from traditional parametric estimation in the sense that the structure of the information has been scrambled along during the acquisition process. When reconstruction is performed after the Compressive Sensing acquisition, an additional assumption on the sparsity of the input enables to recover the information. However, extracting information about features of interest directly from the compressed samples is another challenge, still in exploration.

The goal of this chapter is to quantify the accuracy of parametric estimation depending on the sensing matrix and compared to sampling at the Nyquist rate. A convenient performance assessment tool toward this end, introduced in Section 1.3 and presented in Appendix A, is the Cramér-Rao lower bound (CRB) on the variance of any unbiased estimator, which is computed from the diagonal of the inverse of the Fisher information matrix. Although the Cramér-Rao bound says nothing about the most suitable estimation method, it allows to evaluate in which extent settings impact the estimation performances. This evaluation will in return enables a feature extraction system to be designed according to needs. It is notably essential because constraints on performances appearing in partial restitution of the information are different from those in better-known reconstruction tasks.

The State-of-the-Art of parametric estimation from compressed samples has been detailed in Section 1.3.2.1. In [54], a parametric estimation based on compressed samples is developed, with a specific focus on frequency estimation. Bounds are derived and asymptotic results given as a function of the statistical properties of the sensing matrix:

$$\frac{M}{N}(1-\delta_{2K})^2\boldsymbol{J_{Nyq}} \leq \boldsymbol{J_{CS}} \leq \frac{M}{N}(1+\delta_{2K})^2\boldsymbol{J_{Nyq}} \tag{4.1}$$

where $\boldsymbol{J_{Nyq}}$ and $\boldsymbol{J_{CS}}$ are the Fisher matrices for Nyquist and Compressive Sampling respectively, it is recalled that δ_{2K} is the Restricted Isometry Constant for $2K$-sparse vectors, $\leq$ should be understood as an order relationship between matrices, and $\frac{M}{N}$ is the consequence of compression ratio.

Yet in practice, sensing matrices, whether randomly generated from a distribution or structured, are deterministic. That is why, in order to be able to specify a compressive architecture performing feature extraction, a more thorough framework is advisable. The purpose of following original developments is two-fold: First, to establish asymptotic constraints on each parameter to be estimated, instead of global guarantees with the same scaling factor for all coefficients of the Fisher matrix. Second, to provide bounds straightforwardly computable

from the coefficients of the sensing matrix instead of statistical guarantees, which depends on the RIP constant for instance. Potentially, this analysis could also enable to provide advices on an adequate choice of parameters of interest.

In Section 4.1.1, an original result yielding a closed-form expression of the Fisher information matrix for samples from a compressive multiband architecture is determined under assumption of a spectral disjoint subband model, i.e. assuming that one parameter influences no more than one frequency subband. The expression is given as a function of the Fisher information matrix for Nyquist sampling and of the Gramian of the sensing matrix of a compressive multiband architecture, a result which is more accurate than global guarantees. An intermediate step in function of the Fisher information matrix for Bandpass Sampling is used for sake of clarity. As introduced in Subsection 2.1.3, multiband architectures are receivers where each bin in the same frequency subband is projected identically (multiplied with the same coefficient). Since that includes most popular compressive receivers such as the Modulated Wideband Converter (MWC), the Quadrature Analog to Information Converter (QAIC), the Periodic Non Uniform Sampling (PNUS), Multirate Sampling or Bandpass Sampling, results have a relatively broad application scope.

The targeted application field being Cognitive Radio, simple typical scenarios of detecting interferers near an useful signal are then taken as illustrative examples. Working scenarios with dual-tone signals are first defined in Section 4.2. In Section 4.3, the topic of amplitude estimation for known frequencies is considered. Theoretical bounds are compared as well as variances of Maximum-Likelihood estimators used in simulations. In Section 4.4, frequency estimation with known amplitudes is considered. As the interpretation of the derived formulas is not straightforward, the equations are graphically interpreted. This interpretation then suggests a manner, presented in Section 4.5, of tailoring the sensing matrix to priors on the spectrum so as to tame the influence of strong interferers. Finally, implications of these results are discussed and conclusions are drawn in Section 4.6.

The work presented in this Chapter has lead to the submission of the journal article [143].

4.1 Cramér-Rao bounds for spectral parametric estimation with compressive multiband architectures

To begin with, two original theorems will be stated and then their significance and key elements will be discussed. The reader is kindly asked to refer to Subsection 2.1.3 regarding further properties of multiband architectures and to Subsection 2.1.2 regarding more exhaustive notations, in particular time and frequency indices that are illustrated in Figure 2.2.

4.1.1 Original theorems: Fisher information matrices for different sampling schemes under assumption of a disjoint subband spectral model

Consider a Qx1 parameter vector $\boldsymbol{\theta}$ to be estimated and a K-sparse signal x, meaning that it contains K active subbands. If there is a unique unknown parameter in each active subband,

then $K = Q$. In general $K < Q$. The vector $\mathbf{\Lambda}$ denotes the support and $\Lambda(f)$ denotes the support index of frequency f.

Hypothesis 1 (Disjoint subband model)**.** *Suppose the existence of an underlying disjoint signal model: one parameter influences no more than one subband signal. An illustration example is given in Figure 4.1. More formally, this assumption translates into the existence of a function g that maps each parameter θ_q with its unique corresponding support element $g(\theta_q) = \lambda_q \in \mathbf{\Lambda}$ such that:*
$\forall q \in [\![1; Q]\!]$ if $\lambda_q \neq \Lambda(f)$ then

$$\frac{\partial \tilde{x}(f)}{\partial \theta_q} = 0 \tag{4.2}$$

where it is recalled that $\tilde{x}(f)$ is the continuous Fourier Transform of the input signal.
Tones satisfy this model and so do modulated signals whose channel bandwidths belong to a unique frequency subband of the architecture. For sake of conciseness $\tilde{x}_q$ will denote abusively the subsignal influenced by the qth parameter and central frequencies will be noted f_k for $k \in [\![1; K]\!]$.

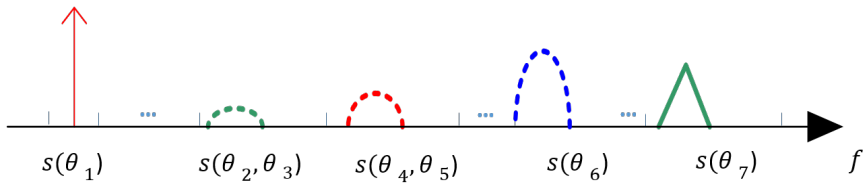


Figure 4.1: Disjoint subband model.

Theorem 1 (Compressive Multiband architectures versus Bandpass Sampling)**.** *Assuming the above model and a real-valued signal x, the Fisher information matrix for compressive multiband architectures $\boldsymbol{J_{MB}}$ can be expressed as a function of the Fisher information matrix for Bandpass sampling $\boldsymbol{J_{BP}}$ in the following manner:*

$$\boldsymbol{J_{MB}} = \Re[\boldsymbol{B}_{\mathbf{\Lambda}}^H \boldsymbol{B}_{\mathbf{\Lambda}}] \circ \boldsymbol{J_{BP}} \tag{4.3}$$

where $\boldsymbol{B}_{\mathbf{\Lambda}}$ is the restriction of the sensing matrix $\boldsymbol{B}$ of a compressive multiband architecture to the support $\mathbf{\Lambda}$, with column energy normalization to M, $\circ$ is the Hadamard (element-wise) product and $\Re[.]$ the real part.
If the signal x is complex-valued, then eq. (4.3) transforms into:

$$\boldsymbol{J_{MB}} = \Re[\boldsymbol{B}_{\mathbf{\Lambda}}^H \boldsymbol{B}_{\mathbf{\Lambda}}] \circ \boldsymbol{J_{BP}} + \Im[\boldsymbol{B}_{\mathbf{\Lambda}}^H \boldsymbol{B}_{\mathbf{\Lambda}}] \circ \check{\boldsymbol{J}}_{\boldsymbol{BP}} \tag{4.4}$$

where $\Im[.]$ denotes the imaginary part. $\check{J}_{BP}[q, r]$ corresponds to the $(q, r)^{th}$ coefficient of the Fisher matrix for Bandpass Sampling in a virtual scenario where the initial dephasing between components $\tilde{x}_q$ and $\tilde{x}_r$ is increased by $\frac{\pi}{2}$.

The interpretation for complex-valued $x(t)$, which would corresponds to the baseband signal in a complex architecture, is less intuitive due to the effect of a complex Gramian on a complex signal which burdens the notations. Note, however, that the behaviour regarding parameters influence and order of magnitude is similar as in the case where $x(t)$ is real-valued.

The Cramér-Rao bound is then obtained through eq. (A.5).

Refer to Appendix E for proof.

Hypothesis 2 (Bandlimited discrete signal)**.** *Suppose also that the sampled subsignals are bandlimited:*

$$\forall q \in [\![1;Q]\!] \;\; \tilde{x}_q[\check{n}] = 0 \; for \; \check{n} \notin [\![\check{n}_0(q) - \frac{L}{2}; \check{n}_0(q) + \frac{L}{2}]\!] \tag{4.5}$$

where $\check{n}_0(q)$ corresponds to the central bin of $\tilde{x}_q$ and L to the number of frequency bins per subband.

Due to finite duration of the acquisition and therefore spectral leakage, this hypothesis is an approximation. Note that Hyp. 2 deals, through the decomposition of the signal on the frequency grid, with bandlimited discretized representations. Whereas Hyp. 1, in contrast, was dealing with bandlimited analog representations and hence did not need to neglect spectral leakage on remote frequencies.

Theorem 2 (Bandpass Sampling versus Nyquist Sampling)**.** *Under the above assumptions, Hypothesis 1 on a disjoint subband model and Hypothesis 2 on a bandlimited discrete representation, the $(q,r)^{th}$ element of $\boldsymbol{J_{BP}}$ for $(q,r) \in [\![1;Q]\!]^2$ is given as a function of the $(q,r)^{th}$ element of the Fisher information matrix for Nyquist rate sampling $\boldsymbol{J_{Nyq}}$:*

$$J_{BP}[q,r](\Delta f_{q,r}) = \frac{1}{N}(J_{Nyq}[q,r](\Delta f_{q,r}).H_{LP}(\Delta f_{q,r})) * \sum_{k=-\infty}^{\infty} \delta(\Delta f_{q,r} - k f_p) \tag{4.6}$$

where $\Delta f_{q,r}$ is the distance in the frequency domain between $\tilde{\boldsymbol{x}}_q$ and $\tilde{\boldsymbol{x}}_r$, $H_{LP}(f)$ is the transfer function of an ideal low-pass filter with cut-off frequency $f_c = f_p/2$ and gain 1, $*$ is the convolution operator and $\delta(.)$ is the Dirac distribution. Note that the term $\sum_{k=-\infty}^{\infty} \delta(\Delta f_{q,r} - k f_p)$ corresponds to a f_p-spaced Dirac comb distribution, used as periodisation operator.

In other words, eq. (4.6) means that the Fisher information matrix for Bandpass Sampling as a function of the frequency distance between subsignals $\tilde{\boldsymbol{x}}_q$ and $\tilde{\boldsymbol{x}}_r$, $J_{BP}[q,r](\Delta f_{q,r})$, is f_p-periodic and matches with the Fisher information matrix for Nyquist sampling $J_{Nyq}[q,r]$ at baseband, up to a N factor.

Refer to Appendix F for proof.

4.1.2 Discussions

The reference benchmark is the inequality (1.12) derived in [54] and discussed in 1.3.2.1:

$$\frac{M}{N}(1-\delta_{2K})^2 \boldsymbol{J_{Nyq}} \leq \boldsymbol{J_{CS}} \leq \frac{M}{N}(1+\delta_{2K})^2 \boldsymbol{J_{Nyq}} \tag{4.7}$$

where it is recalled that δ_{2K} is the Restricted Isometry Constant for $2K$-sparse vectors and $\frac{M}{N}$ is the consequence of dimension reduction.

In this work, an additional hypothesis on a compressive multiband architecture and a disjoint subband model is made in the context of Spectrum Sensing to obtain eq. (4.3)

and (4.4). This assumption enables to describe the effect of the acquisition matrix on each frequency subband and thus enables to determine each term thanks to the deterministic nature of the sensing matrix. It results in a better accuracy than the state-of-the-art that considers statistical guarantees on the globality of the Fisher matrix.

Also, the relationships established in eq. (4.4) and eq. (4.6) between Fisher matrices for compressive multiband, subsampling and Nyquist sampling allow to distinguish the influence on the estimation bound of the effect of subsampling from the effect of diversity creation between channels.

Regarding Theorem 1, it is not surprising to see the Gramian of the sensing matrix restricted to the support, $\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}$, appearing for Theorem 1 in eq. (4.3) and (4.4), since the Gramian also holds a crucial role in signal reconstruction. Recall that coherence and isometric properties analyzed in Chapter 3 can be derived from the Gramian. This is because the Gramian defines the deformation of the scalar product by the dimension-reducing acquisition process and hence expresses the geometric constraints.

One fundamental but thin subtlety allowing the establishment of the first theorem lies in the fact that the sensing matrix of compressive multiband architectures can be expressed in the 'compact formalism' (see 2.1.3). Hence all frequencies of an analog subband are projected identically, even if they do not lie on the Nyquist grid with frequency resolution δf, which means that leakages provoked by the discretization of the analog disjoint model will depend on the coefficient of the sensing matrix $\boldsymbol{B}$ given by the original support index (cf proof).

Regarding Theorem 2, the f_p-periodicity of the Fisher information matrix for Bandpass Sampling $J_{BP}[q,r](\Delta f_{q,r})$ due to the signal folding is relatively obvious and the scaling factor is related to noise folding (cf proof). The link between $\boldsymbol{J_{BP}}$ and $\boldsymbol{J_{Nyq}}$ is drawn under a small approximation on band-limitation of the discrete representation (Hyp.2) an approximation that is more accurate as the acquisition time grows, because potential spectral leakage from components that do not lie on the Nyquist grid with finite resolution δf is reduced.

Note that in contrast to the usual unit column energy normalization of the sensing matrix $\boldsymbol{B}$ which matches an equal measurement power budget allocation per bin, it is chosen here to use the convention of a column energy normalization of the sensing matrix to M for sake of compactness and so that the scaling effects appears more clearly.

In the following section, these new theorems will be illustrated through a practical application case in order to clarify and exemplify their implications.

4.2 Applications to cognitive radio: definition of a framework

To illustrate the interest of Theorems 1 and 2 for cognitive radio applications, it will be focused on the canonical application case of amplitude and frequency estimation of two tones. This admittedly simplified scenario mimics the detection of an interferer in presence of useful signal with background noise, a common issue for wideband compressive receivers (see e.g.[6,

46]). Thus it allows to illustrate the challenges at stake. Ideally, one would like to be able to handle interferers whose power is up to $80dB$ above the signal of interest and located in adjacent channels, which should be in a different frequency subband from the point of view of the compressive receiver.

The noisy time domain input signal sampled at rate f_{Nyq} is given at the sampling instant $\check{n}T_{Nyq}$ for $\check{n} \in [\![1; NL]\!]$ by:

$$x[\check{n}]^+ = A_1 cos(2\pi f_1 \check{n} + \phi_1) + A_2 cos(2\pi f_2 \check{n} + \phi_2) + w[\check{n}] \tag{4.8}$$

where $(f_1,f_2) \in [-0.5; 0.5]^2$ are the tone frequencies normalized by f_{Nyq}, A_1 and A_2 the amplitudes and $w \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_{NL})$.
$\boldsymbol{x}^+$ can be decomposed in the following manner:

$$\boldsymbol{x}^+ = \boldsymbol{E}\boldsymbol{a} + \boldsymbol{w} \tag{4.9}$$

with

$$\boldsymbol{E} = \begin{pmatrix} cos(2\pi f_1) & cos(2\pi f_2) \\ \vdots & \vdots \\ cos(2\pi f_1 \check{n}) & cos(2\pi f_2 \check{n}) \\ \vdots & \vdots \\ cos(2\pi f_1 NL) & cos(2\pi f_2 NL) \end{pmatrix}, \tag{4.10}$$

where $\boldsymbol{E} \in \mathbb{R}^{NLx2}$ and $\boldsymbol{a}^T = (A_1, A_2)$.

For sake of conciness of the equations, following notations will be used:
$\Delta f = f_2 - f_1$, $\Delta\phi = \phi_2 - \phi_1$, $\Sigma f = f_2 + f_1$ and $\Sigma\phi = \phi_2 + \phi_1$.
For illustrations, the Modulated Wideband Converter architecture is chosen but the proposed Theorem 1 is valid for all multiband architectures (Bandpass, MWC, QAIC, PNUS, MRS); results can be easily extended by replacing the corresponding sensing matrix $\boldsymbol{B}$, see section 2.1.3.

The exhaustive Fisher matrix for Nyquist sampling of K real tones which is used as benchmark throughout the following is derived in Appendix G. For sake of clarity, non-realistic disjoint estimation with 2 real tones is presented, however the study of joint estimation or a multitone input signal would be methodologically similar. Note that the benchmark is logically established for equal acquisition times T_{acq} between architectures, and not equal number of samples, as scanning time is also a constraint for our targeted application.

4.3 Amplitude estimation

In this scenario, it is considered that the normalized frequencies of the useful signal f_1 and of a possible interferer f_2 are known, and that the parameters to be estimated are $\boldsymbol{\theta} = (A_1, A_2)$.

First, the expressions of Fisher matrices for three sampling schemes (Nyquist rate, Bandpass and compressive multiband sampling) are established and then compared to each others.

To supplement these results with another point of view, a typical unbiased estimator, the Maximum Likelihood estimator, is considered. Analytical expressions are established and benchmarked with experimental values.

4.3.1 Expressions of the Fisher information matrices

4.3.1.1 Nyquist Sampling

According to the definition (A.5), for $(q,r) \in [\![1;2]\!]^2$, the $(q,r)^{th}$ coefficient of the Fisher information matrix for Nyquist Sampling is expressed by:

$$J_{Nyq}[q,r] = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{NL} \frac{\partial y[\check{n}]}{\partial A_q} \frac{\partial y[\check{n}]}{\partial A_r} \tag{4.11}$$

For Nyquist sampling, it is immediate that $y[\check{n}] = x[\check{n}]$ for $\check{n} \in [\![1;NL]\!]$. Given the real-valued signal model (G.1), the partial derivative with respect to amplitude of the signal sampled at the Nyquist rate is given by:

$$\frac{\partial x[\check{n}]}{\partial A_i} = cos(2\pi f_i \check{n} + \phi_i), \text{ for } i = 1,2. \tag{4.12}$$

Easy but long computations detailed in Appendix G, which consider the real part of the sum of geometrical sequence terms with common ratio $e^{2j\pi f}$ and use half angle factorization, then yields:

$$\boldsymbol{J_{Nyq}} \simeq \tfrac{NL}{2\sigma^2} \begin{pmatrix} 1 & J_{.,Nyq}[1,2] \\ J_{.,Nyq}[1,2] & 1 \end{pmatrix} \tag{4.13}$$

where $J_{.,Nyq}[1,2]$ is given by:

$$J_{.,Nyq}[1,2] = J_{1/2,Nyq}[1,2](\Delta f, \Delta\phi) + J_{1/2,Nyq}[1,2](\Sigma f, \Sigma\phi) \tag{4.14}$$

with:

$$\begin{aligned} J_{1/2,Nyq}[1,2](f,\phi) &= \tfrac{sin(2\pi NLf)}{NLsin(2\pi f)} cos(\phi + \pi f(NL+1)) \\ &\simeq sinc(\pi NLf) cos(\phi + \pi f(NL+1)) \end{aligned} \tag{4.15}$$

where $sinc(x)$ stands for $\frac{sin(x)}{x}$.

The lower the cross-term coefficient of the Fisher matrix $J_{Nyq}[1,2]$, the lower the determinant $\Delta = \frac{NL}{2\sigma^2}[(1 - J_{.,Nyq}[1,2]^2)]$ of $\boldsymbol{J}_{.,Nyq}$ and the better the estimation.

When $J_{.,Nyq}[1,2] = 0$, the matrix is diagonal so that not knowing one amplitude does not degrade the estimation of the other. In other words, if $J_{.,Nyq}[1,2]$ is non-zero, the presence of one signal makes the estimation of the parameters of the other more difficult; an interaction that is slightly different from the correlation notion (one frequency cannot be expressed as a function of the other).

When $f_1 = f_2$ and $\phi_1 = \phi_2 \mod \pi$, or $f_1 = -f_2$ and $\phi_1 = -\phi_2 \mod \pi$, then $J_{Nyq}[1,2] = J_{Nyq}[1,1]$, all terms of the Fisher matrix are equal. Because foldings are indiscernible in this situation, the matrix is singular, meaning that amplitude estimation by an unbiased estimator is not possible.

Consider first, as often in the literature, the frequency on-grid case, meaning that the frequencies Σf and Δf lie exactly on the Nyquist grid with frequency resolution δf:

$$\exists (k,k') \in \mathbb{N}^{*2} \text{ such that } \Delta f.f_{Nyq} = k\delta f \text{ and } \Sigma f.f_{Nyq} = k'\delta f. \tag{4.16}$$

where $f_{Nyq} = NL.\delta f$. Then $sin(2\pi NL\Delta f) = sin(2\pi k) = 0$ and $sin(2\pi NL\Sigma f) = sin(2\pi k') = 0$ hence $J_{.}[1,2] = 0$ and $\boldsymbol{J_{Nyq}} \simeq \frac{NL}{2\sigma^2}.\boldsymbol{I}_2$. There is no spectral leakage if $f_2 - f_1$ and $f_2 + f_1$ are multiples of the frequency resolution given by the inverse of the acquisition time $\frac{1}{T_{acq}}$. Furthermore, it is important to take note that the Cramér-Rao bound and hence the accuracy of the estimation depends only on the number of samples NL.

However, in practice the probability that tones lay exactly on the grid is equal to zero. Consider now the more realistic off-grid case meaning that there exists $0 < \eta, \eta' < 1$ such that:

$$\begin{aligned}\Delta f.f_{Nyq} &= \lfloor \frac{(f_2 - f_1).f_{Nyq}}{\delta f} \rfloor \delta f + \eta \delta f \\ \Sigma f.f_{Nyq} &= \lfloor \frac{(f_1 + f_2).f_{Nyq}}{\delta f} \rfloor \delta f + \eta' \delta f\end{aligned}$$

where $\eta\delta f$ and $\eta'\delta f$ stands for the distance of $(f_2 - f_1)f_{Nyq}$ and $(f_1 + f_2)f_{Nyq}$ respectively to the Nyquist grid with frequency resolution δf. Then eq. (4.15) is transformed into:

$$\begin{aligned}J_{.,Nyq}[1,2] = \frac{1}{NL}[&\frac{sin(\pi(k+\eta))}{sin(\pi\Delta f)} cos(\Delta\phi + \pi\Delta f(NL+1)) \\ &+ \frac{sin(\pi(k+\eta'))}{sin(\pi\Sigma f)} cos(\Sigma\phi + \pi\Sigma f(NL+1))]\end{aligned} \tag{4.17}$$

Hence for large number of samples NL, $J_{.,Nyq}[1,2] = O(sinc(\pi NL\Delta f))$ approaches 0.
This matches the fact that the spectral leakage created by the mismatch with the frequency grid approaches 0 as the acquisition time tends to infinity (hence the signal model converges to a frequency-continuous model).

4.3.1.2 Bandpass Sampling

For Bandpass Sampling at normalized subsampling frequency $f_p = \frac{1}{N}$, let us define the frequencies folded at baseband with the notation $\bar{.}$:

$$\begin{aligned}\bar{f_1} &= f_1 - f_p \lfloor \frac{f_1}{f_p} \rfloor, \\ \bar{f_2} &= f_2 - f_p \lfloor \frac{f_2}{f_p} \rfloor \text{ and} \\ \Delta\bar{f} &= \bar{f_1} - \bar{f_2}.\end{aligned} \tag{4.18}$$

From the expression of the signal in (G.1), the partial derivative with respect to amplitude of the subsampled signal can be derived as:

$$\frac{\partial y[l]_{BP}}{\partial A_i} = cos(2\pi \bar{f}_i N l + \phi_i), \text{ for } i = 1, 2. \tag{4.19}$$

Hence from the definition formula (A.6), the Fisher information matrix for Bandpass Sampling is yielded by:

$$\boldsymbol{J_{BP}} = \frac{L}{2\sigma^2} \begin{pmatrix} 1 & J_{.,BP}[1,2] \\ J_{.,BP}[1,2] & 1 \end{pmatrix} \tag{4.20}$$

where the element in the first row and first column of $\boldsymbol{J_{BP}}$ is given by:

$$J_{BP}[1,1] = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{L} (\frac{1}{2} + \frac{cos(4\pi \check{n} \bar{f}_1 + 2\bar{\phi}_1)}{2}) \simeq \frac{L}{2\sigma^2} \tag{4.21}$$

and:

$$J_{.,BP}[1,2] = \frac{1}{L} \sum_{l=1}^{L} [cos(2\pi \Delta \bar{f} N l + \Delta\phi) + cos(2\pi \Sigma \bar{f} N l + \Sigma\phi)] \tag{4.22}$$

Similarly to the establishment of eq. (G.10) in Appendix G, by considering the real part of the sum of geometrical sequence terms with common ratio $e^{2j\pi f}$ and factorizing through the half angle, it is established that:

$$J_{.,BP}[1,2] = J_{1/2,BP}[1,2](N\Delta\bar{f}, \Delta\phi) + J_{1/2,BP}[1,2](N\Sigma\bar{f}, \Sigma\phi) \tag{4.23}$$

where:

$$J_{1/2,BP}[1,2](f,\phi) = \frac{sin(\pi L f)}{L sin(\pi f)} cos(\phi + \pi f(L+1)) \simeq sinc(\pi L f) cos(\phi + \pi f(L+1)) \tag{4.24}$$

which will be discussed shortly after.

4.3.1.3 Multiband Compressive Sampling

From the application of Theorem 1 and eq. (4.4), the Fisher information matrix for a compressive multiband architecture is derived as a function of the Fisher information for bandpass sampling $\boldsymbol{J_{BP}}$:

$$\boldsymbol{J_{MB}} = \Re[\boldsymbol{B_\Lambda^H B_\Lambda}] \circ \boldsymbol{J_{BP}} \tag{4.25}$$

where it is recalled that $\boldsymbol{B_\Lambda}$ is the restriction to the active frequency support of the columns of $\boldsymbol{B}$, the sensing matrix of a compressive multiband architecture expressed in the 'compact formalism'. $\boldsymbol{B_\Lambda^H B_\Lambda}$ is called its Gramian. Previous results on $\boldsymbol{J_{BP}}$ and $\boldsymbol{J_{MB}}$, eq. (4.23) and (4.25), will be discussed below and compared to $\boldsymbol{J_{Nyq}}$. Note that coefficients energy normalization implies that diagonal elements of the Gramian are equal to M.

4.3.2 Interpretation and discussions

Consider first the on-grid frequency case, i.e. the frequencies $\Delta\bar{f}$ and $\Sigma\bar{f}$ lie on the Nyquist grid:

$$\exists(k,k') \in \mathbb{N}^{*2} \text{ such that } \Delta\bar{f}.f_{Nyq} = k\delta\bar{f} \text{ and } \Sigma\bar{f}.f_{Nyq} = k'\delta\bar{f} \tag{4.26}$$

As was obtained for Nyquist samples, $J_{BP}[1,2] = J_{BP}[2,1] = 0$ and $J_{MB}[2,1] = J_{MB}[1,2] = 0$. For all sampling methods, the values taken by $J[1,1]$ and $J[2,2]$ are proportional to the number of samples, which is L for bandpass sampling; that is, N times less than Nyquist sampling, for the same acquisition time. Hence when the two tones do not interfere with each others (hence the Fisher matrix is diagonal), the accuracies of Bandpass and multiband compressive amplitude estimation with an unbiased estimator compared to estimation from the Nyquist samples are exactly characterized by the respective compression ratios $1/N$ and M/N.

Consider now the off-grid frequency case, i.e. frequencies $\Delta\bar{f}$ and $\Sigma\bar{f}$ do not lie on the Nyquist grid. The interference phenomenon between the two tones under compressive sampling acquisition is then characterized by the product of two terms. The first term, $J_{BP}[1,2]$, depends on the distance toward the frequency grid of the sum and difference of the tone frequencies. The second term is an element of the Gramian sensing matrix, $B_\Lambda^H B_\Lambda[1,2]$.

For $\Phi_1 = \Phi_2$, one may notice that $f_1 = f_2$ for Nyquist and Multiband Compressive Sampling and $f_1 = f_2 \mod (f_p)$ for Bandpass Sampling makes the estimation problem singular because of the mismatch between the two tones model and the single tone observation.

Here, note that Hyp. 2 on bandlimited spectral representations of the sampled tones needs not be assumed. Hence obtained results are slightly more accurate than the application of Theorem 2 because the effect of spectral leakage is not approximated.

4.3.3 Maximum Likelihood Estimator (MLE)

Another perspective on the topic is to consider the Maximum Likelihood estimator (MLE), which is an unbiased estimator and hence lower-bounded by the Cramér-Rao bound. In this part, theoretical expressions of the Maximum-Likelihood estimator and its variance are established for Nyquist rate and compressive multiband sampling. Then an experimental benchmark illustrates and validates previous equations.

4.3.3.1 Theoretical expressions

Under noise Gaussianity assumption, it is well-known that the Maximum-Likelihood (ML) estimator in the non-compressed case is given from eq. (4.9) by [53]:

$$\hat{\boldsymbol{a}}_{\boldsymbol{Nyq}} = (\boldsymbol{E}^{\boldsymbol{H}}\boldsymbol{E})^{-1}\boldsymbol{E}^{\boldsymbol{H}}\boldsymbol{x}^{+} \tag{4.27}$$

with $\boldsymbol{E}$ as in (4.11) that matches the expression of a real-valued dual tones signal.
Also the Maximum-Likelihood covariance matrix is given by:

$$\boldsymbol{Cov}(\hat{\boldsymbol{a}}_{\boldsymbol{Nyq}}) = \sigma^2(\boldsymbol{E}^H\boldsymbol{E})^{-1} \simeq \frac{2\sigma^2}{NL} \tag{4.28}$$

For a compressive multiband architecture, it is possible to write the output as $\boldsymbol{y} = \boldsymbol{Ha}$ where

$$\boldsymbol{H} = \begin{pmatrix} \begin{pmatrix} cos(2\pi\bar{f}_1N+\phi_1) & cos(2\pi\bar{f}_2N+\phi_2) \\ \vdots & \vdots \\ cos(2\pi\bar{f}_1NL+\phi_1) & cos(2\pi\bar{f}_2NL+\phi_2) \end{pmatrix} \begin{pmatrix} b[1,1] & 0 \\ 0 & b[1,2] \end{pmatrix} \\ - - - - - - - - - - - - - - - - \\ \vdots \\ - - - - - - - - - - - - - - - - \\ \begin{pmatrix} cos(2\pi\bar{f}_1N+\phi_1) & cos(2\pi\bar{f}_2N+\phi_2) \\ \vdots & \vdots \\ cos(2\pi\bar{f}_1NL+\phi_1) & cos(2\pi\bar{f}_2NL+\phi_2) \end{pmatrix} \begin{pmatrix} b[M,1] & 0 \\ 0 & b[M,2] \end{pmatrix} \end{pmatrix} \tag{4.29}$$

is a matrix of dimension MLx2, and $b[i,j]$ is the $(i,j)^{th}$ coefficient of the sensing matrix of the compressive multiband architecture $\boldsymbol{B}$.

If the compressed noise is assumed to be white Gaussian, the Maximum-Likelihood estimator coincides with the best linear estimator and is given by [53]:

$$\hat{\boldsymbol{a}}_{\boldsymbol{MB}} = (\boldsymbol{H}^{\boldsymbol{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\boldsymbol{H}}\boldsymbol{y}^{+} \tag{4.30}$$

and the covariance matrix of the estimation is given by:

$$\boldsymbol{Cov}(\hat{\boldsymbol{a}}_{\boldsymbol{MB}}) = \sigma^2(\boldsymbol{H}^H\boldsymbol{H})^{-1} \tag{4.31}$$

The more the coherence and isometric properties of the sensing matrix are degraded, the more the approximation made in Subsection 2.1.4 of Chapter 2 on a diagonal noise covariance matrix $\boldsymbol{\Gamma}$ becomes rough. Without this approximation, the covariance matrix given in (4.31) would become:

$$\boldsymbol{Cov}(\hat{\boldsymbol{a}}_{\boldsymbol{MB}}) = (\boldsymbol{H}^H\boldsymbol{\Gamma}^{-1}\boldsymbol{H})^{-1} \tag{4.32}$$

It is recalled that as the number of samples tends to infinity, the covariance of any Maximum-Likelihood estimator converges to the corresponding Cramér-Rao bound.
For multiband architectures, cross-terms of $\boldsymbol{J}$ are negligible toward diagonal terms as the number of samples tends to infinity, approaching the on-grid case:

$$\begin{aligned} \boldsymbol{Cov}(\hat{\boldsymbol{a}}_{\boldsymbol{MB}}) &\approx \frac{2\sigma^2}{ML}\begin{pmatrix} \frac{1}{(\boldsymbol{B}_\Lambda^H\boldsymbol{B}_\Lambda)_{1,1}} & 0 \\ 0 & \frac{1}{(\boldsymbol{B}_\Lambda^H\boldsymbol{B}_\Lambda)_{2,2}} \end{pmatrix} \\ &\approx \frac{N}{M}\boldsymbol{Cov}(\hat{\boldsymbol{a}}_{\boldsymbol{Nyq}}) \end{aligned} \tag{4.33}$$

As expected, the compression ratio $\frac{M}{N}$ appears as scaling between the variances of Nyquist sampling and compressive multiband sampling.

4.3.3.2 Experimental variance benchmark

In the Subsection above, the expressions of the theoretical variance lower bounds for a Maximum-Likelihood estimator (MLE) has been established for Nyquist rate and compressive multiband sampling. To complement this benchmark, variances obtained by simulation means with a MLE are compared to these analytical expressions.

A simplified version of the MWC platform is used to compute the compressed samples. Set-up parameters are reported in Table 4.1.

Table 4.1: Set-up parameters.

Parameters	Number of trials	N	M	L
Default value	1000	127	50	120

The empirical variance of the MLE obtained for both the Nyquist (blue) and Modulated Wideband Converter (red) architecture with random Bernoulli codes (stars, from (4.27) and (4.30)) is shown in Figure 4.2. It is compared to the expected variance of the MLE, from (4.31) and (4.32) (triangle). Results are also compared to the Cramér-Rao bound from Theorem 2 (line).

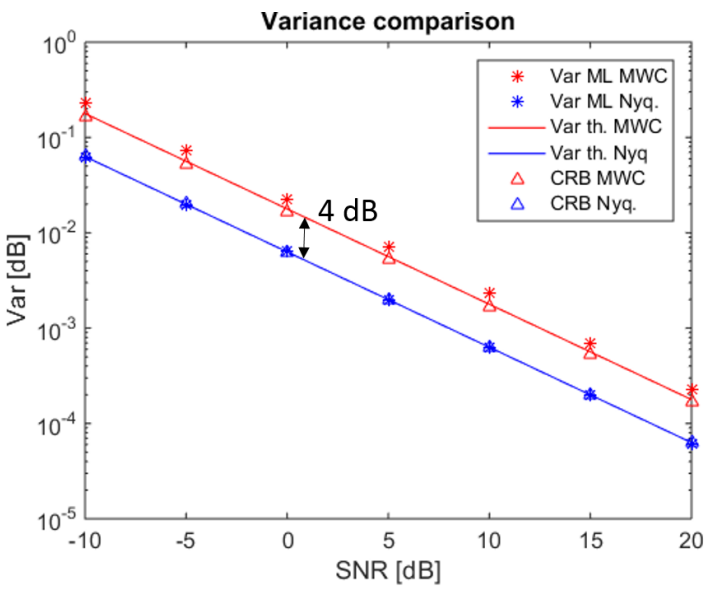


Figure 4.2: Variance comparison for amplitude estimation of 2 real tones.
[$M = 50$, $N = 127$, 1000 trials, $f_1 = 1GHz$, $f_2 = 3GHz$, $f_{Nyq} = 10GHz$, $A1 = A2 = 1$].

The variance obtained by simulation means with a Maximum Likelihood estimator indeed converges quickly to the Cramér-Rao bound, as is expected from the theory (cf Appendix G). Note also the expected $10\log_{10}(\frac{M}{N}) = -4.04dB$ loss between the acquisition methods, which

matches with the Signal-to-Noise degradation by the compressive sensing projection. Hence the simulation results validate with an illustration the bound established analytically.

In the next Subsection, the study will be lead on another parameter of interest: the example of frequency estimation will be tackled, with a focus on the role of the Gramian.

4.4 Frequency estimation

The second scenario mimics the detection of an interferer at any frequency with known amplitude and phase in presence of background noise. The parameters to be estimated are the normalized frequencies of the two tones $\boldsymbol{\theta} = (f_1, f_2)$.

In this Section, the expressions of Fisher matrices for three sampling schemes (Nyquist rate, Bandpass and compressive multiband) are established. A graphical illustration is provided to interpret the equations.

4.4.1 Expressions of the Fisher information matrices

4.4.1.1 Nyquist Sampling

According to the definition given in eq. (A.5), the input model given in eq. (G.1) and computations that are developed in Appendix G, following expression of the Fisher information matrix given Nyquist samples is established:

$$\boldsymbol{J_{Nyq}} \simeq \frac{2\pi^2}{\sigma^2} \begin{pmatrix} A_1^2 S_2(NL) & J_{.}[1,2] \\ J_{.}[1,2] & A_2^2 S_2(NL) \end{pmatrix}$$

where $S_2(NL) = \sum_{\check{n}=1}^{NL} \check{n}^2 = \frac{NL(NL+1)(NL+2)}{6}$ is the sum of squares, and from (G.12):

$$J_{.,Nyq}[1,2] = A_1 A_2 \Re[e^{j\Delta\phi_{1,2}} \chi_2(\Delta f, NL) - e^{j\Sigma\phi_{1,2}} \chi_2(\Sigma f, NL)] \tag{4.34}$$

where $\chi_2(f, NL) = \sum_{\check{n}=1}^{NL} \check{n}^2 e^{j2\pi f \check{n}}$.
Hence the Cramér-Rao bound for frequency estimation is of the order of $O(\frac{\sigma^2}{(NL)^3})$.

4.4.1.2 Bandpass Sampling

The derivative of a subsampled dual-tone signal with respect to frequency is given by:

$$\frac{\partial y[l]_{BP}}{\partial f_i} = -2\pi N A_i l sin(2\pi \bar{f}_i N l + \phi_i), \text{ for } i = 1, 2. \tag{4.35}$$

where it is recalled that $\bar{f}_i$ is the normalized subsampling frequency as defined in eq. (4.18). Therefore, from the definition formula (A.5), and similarly to the establishment of eq. (G.12), the Fisher information matrix given bandpass samples is given by:

$$\boldsymbol{J_{BP}} \simeq \frac{2\pi^2}{\sigma^2} \begin{pmatrix} N^2 A_1^2 S_2(L) & J_{.,BP}[1,2] \\ J_{.,BP}[1,2] & N^2 A_2^2 S_2(L) \end{pmatrix}$$

where:

$$\begin{aligned} J_{.,BP}[1,2] &= A_1 A_2 \Re[\textstyle\sum_{l=1}^{L} l^2 e^{2j\pi\Delta\bar{f}Nl+\Delta\phi}] \\ &= A_1 A_2 \Re[e^{j\Delta\bar{\phi}_{1,2}}\chi_2(\Delta\bar{f}N, L) - e^{j\Sigma\bar{\phi}_{1,2}}\chi_2(\Sigma\bar{f}N, L)] \end{aligned} \tag{4.36}$$

Hence the Cramér-Rao bound for frequency estimation which was of order of $O(\frac{\sigma^2}{(NL)^3})$ for Nyquist Sampling is modified in of order of $O(\frac{\sigma^2}{N^2L^3})$ for Bandpass Sampling.

4.4.1.3 Multiband Compressive Sampling

From applying Theorem 1, the Fisher information matrix for a compressive multiband architecture is derived as a function of the Fisher information for bandpass sampling $\boldsymbol{J_{BP}}$:

$$\boldsymbol{J_{MB}} = \Re[\boldsymbol{B_\Lambda^H B_\Lambda}] \circ \boldsymbol{J_{BP}} \tag{4.37}$$

Due to scaling of the Gramian, the variance of frequency estimation with an unbiased estimator is modified in of order of $O(\frac{\sigma^2}{(NL)^2ML})$ for Compressive Sensing, which matches the results obtained in [55].

As the behaviour of eq. (4.36) and (4.34) is not straightforward to analyze, these expressions will be studied in simulation.

4.4.2 Illustration and interpretation of the analytical expressions

For these three acquisition methods (Nyquist sampling, bandpass sampling and compressive multiband sampling) the normalized coefficient $J_{norm}[1,2] = |\frac{J[1,2]}{J[1,1]}|$, which translates the ambiguity of the estimation due to coupling between parameters, is pictured in Figure 4.3. It is reminded that if it is equal to 0, there is no degradation of the estimation of one amplitude due to not knowing the other amplitude, and if it is equal to 1, it is impossible to raise the ambiguity between the two signals and no estimation can be made. The x-axis corresponds to the relative normalized frequency $\frac{f_2}{f_{Nyq}}$, and the term $\frac{f_1}{f_{Nyq}} = 0.22581$ is arbitrarily fixed. The codes of the acquisition matrix in the time domain are randomly generated from a Bernoulli distribution.

It appears that the value of $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ is for Bandpass Sampling (blue) the f_p-periodic repetition of the value of $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ in $[-f_p/2; +f_p/2]$ for a Nyquist approach (black). For multiband compressive sensing (red), $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ is additionally weighted in each frequency subband. This weight, corresponding to $\Re[B_\Lambda^H B_\Lambda[\lambda_1, \lambda_2]]$, is given by the real part of the cross-correlation between the two support columns of the sensing matrix of indices $\lambda_1 = 7$ and variable λ_2. It would be 1 for Bandpass Sampling.

Hence two frequencies f_1 and f_2 belonging to the same frequency subband fold similarly (with the same coefficient). In this setting, the Cramér-Rao bound for Multiband CS is equal to the Cramér-Rao bound for Bandpass, up to a factor M. However, for two frequencies belonging to different subbands, there is an attenuation coefficient (< 1) controlled by the

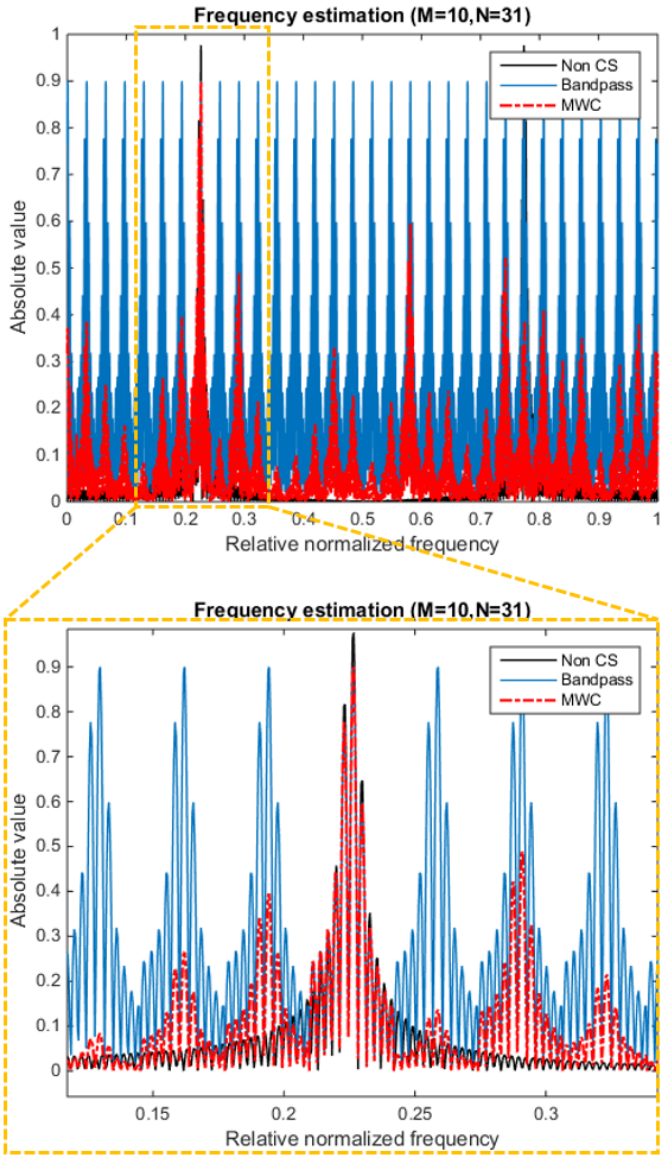


Figure 4.3: Normalized off-diagonal coefficient of the Fisher matrix for frequency estimation $J_{norm}[1,2](\frac{f_2}{f_{Nyq}})$ (top=overall, bottom=zoom), $[M = 10, N = 31]$.

sensing matrix.

Note that the figure is not symmetric which is due to the influence of the second term depending on $\Sigma\bar{f}$, $\Re[-e^{j\Sigma\bar{\phi}}\chi_2(\Sigma\bar{f})]$ in eq. (4.36).

4.5 Influencing the Gramian

In this section, opportunities of adaptive approaches that ensue directly from the analysis which has been conducted above on the parametric estimation accuracy are sketched. First,

the working principle, based on the exploitation of non uniform priors on the spectrum distribution, is described. Secondly, this principle is pictured on the Gramian of the sensing matrix $\boldsymbol{B}$.

4.5.1 Interest of exploiting non-uniform spectrum distribution assumptions

For generality purpose, the sensing matrix $\boldsymbol{\Phi\Psi}$, noted $\boldsymbol{B}$ for compressive multiband architectures, is often supposed to be random. However, the sensing matrix depends on both the structure of the architecture and a pseudo-random element (mixing codes, sample selection patterns, etc.). This opens the opportunity to design codes to improve the overall performances of the system. For example, it can be used to take into account a non uniform input spectrum distribution.

In an adaptive Bayesian approach [144], it is required to recompute the sensing matrix, which is extremely expensive, especially if the spectrum is changing quickly.

In comparison, an approach where an adapted sensing matrix is determined once based on a non-uniform spectrum distribution criterion, may raise interest. This is similar to the concept of rakeness presented in [145]. It must be ensured that signals and interferers are well discriminated and hence that the corresponding columns of interest of $\boldsymbol{B}$ are weakly correlated. On the other hand, columns of $\boldsymbol{B}$ corresponding to frequency ranges where no useful signal is expected are allowed to be far from orthogonal to each others. Very concretely, for the MWC, one frequency subband and its matching column in $\boldsymbol{B}$ corresponds on each of the M branches to one coefficient of the Fourier transform of the shaped codes.

Note that rakeness is against the very notion of CS uniform guarantees for which it is guaranteed that any K-sparse vectors can be recovered given sufficient measurement. Conversely, adaptive schemes would rather try to maximize the detection probability over the non-detection probability given a non-random distribution of the sparse input signal. In this sense, one walks a few steps away from the agnostic random compressive sensing to move toward tailored dimensionality reduction techniques.

4.5.2 Illustrative example

Element $[i, j]$ of the Gramian $\boldsymbol{B}^{\boldsymbol{H}}\boldsymbol{B}$ is the cross-correlation between columns i and j of the matrix. A simple example on how to influence the Gramian is given hereafter to yield a brief insight into the possibilities.

It is shown in Appendix B that the MWC sensing matrix may be expressed as $\boldsymbol{B} = \boldsymbol{A}\tilde{\boldsymbol{F}}\boldsymbol{D}$ where $\boldsymbol{A}$ is the time expression of the code, $\tilde{\boldsymbol{F}}$ is a reordered subset of the DFT matrix, $\boldsymbol{D}$ is a diagonal matrix accounting for pulse shaping. Figure 4.4 presents the Gramian for a Modulated Wideband Converter, and two different codes studied in Chapter 3 for reconstruction.

"Random" corresponds to a Bernoulli distribution, "ZC circ" corresponds to Zadoff-Chu-based circulant codes introduced in Chapter 3 which have good cross-correlation properties. The weights of the coefficients for multiband CS architectures (red dots on the red curve) in

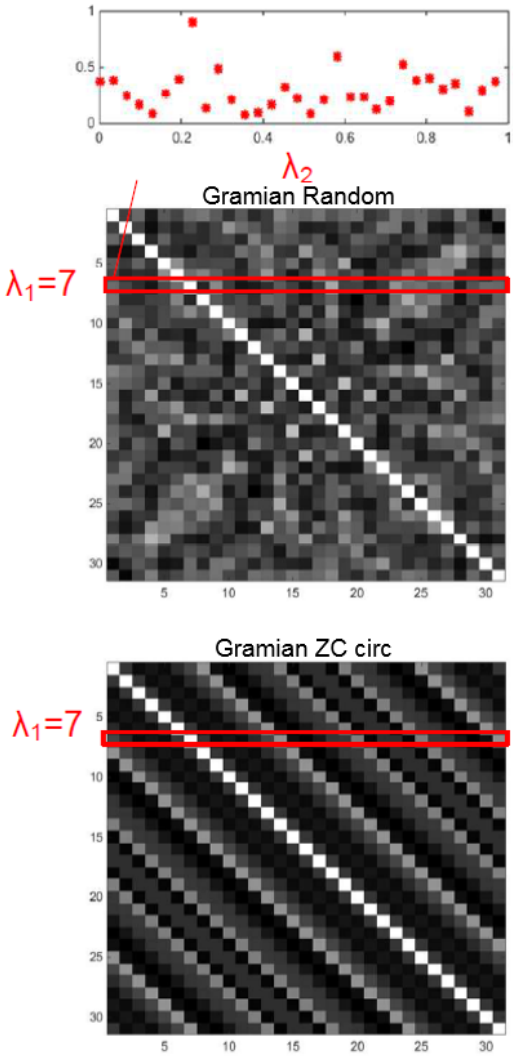


Figure 4.4: Gramian of the Modulated Wideband Converter sensing matrix for different codes: Random Bernoulli (top) and Zadoff-Chu-based circulant (bottom) [N=31].

Figure 4.3 are given by the values on the 7th row (because $\mathbf{\Lambda}(f_1) = 7$) of the Gramian in Figure 4.4, values which are reproduced at the top. For the two different structures of codes, the Gramian presents two specific patterns. For "Random" codes, values of the off-diagonal coefficients appear uniformly distributed whereas the circulant character of code (bottom) implies equal-valued diagonal lines. Hence by changing the shift of the circulant matrix, it is possible to choose the circular permutation with the most favorable coefficient between λ_1 and λ_2. More sophisticated tunings, by optimizing the choice of M rows among N for example, are also possible. Note that it is shown in Chapter 3 that the coherence (maximal value of the off-diagonal coefficients) of the second code ('ZC circ') is also lower. Thus it is possible to adapt the structure of the mixing code in the MWC, and the choice of the pseudo-random element in the general case, in order to promote non-uniformly distributed performance according to the frequency subband index.

Combined with the presented theorems, this perspective shows that it is possible to control the Cramér-Rao bound and hence the frequency estimation precision thanks to the cross-correlation properties of the acquisition matrix. With the random codes for the MWC given in example, one reads at the 4th row and 7th column of the Gramian, $\Re[\boldsymbol{B}_{\boldsymbol{\Lambda}}^H \boldsymbol{B}_{\boldsymbol{\Lambda}}[4,7]] = 0.9$. Thus frequency estimation of a single tone signal is as accurate if there is an interferer in the 4th subband with $-80dB$ Signal-to-Interference ratio (SIR), as if the interferer is in the 7th subband (same as the signal, $\boldsymbol{\Lambda}(f_1) = 7$), same bin, with $-70dB$ SIR. Given the orders of magnitude obtained, it is possible to handle even strong blockers, if the sensing matrix is adequately chosen.

4.6 Conclusions

In the state-of-the-art, the Cramér-Rao bound for parametric estimation from compressed samples had so far only been given as a guarantee on the globality of the Fisher matrix [54], based on a statistical property, the Restricted Isometry Property. In this Chapter, two theorems have been presented and demonstrated, which offer a bound on each term of the Fisher matrix from samples acquired with a multiband compressive radiofrequency receiver. As it is adapted to each term, it is more accurate than [54]. This is achieved thanks to the deterministic nature of the sensing matrix in practice and the assumption of a specific signal model, where each parameter influences no more than one signal subband. These expressions underline the key role of the Gramian of the sensing matrix $\boldsymbol{B}$. It is interesting to note that since reconstruction performance depends on coherence and norm preservation, metrics which are closely linked to the Gramian, a good matrix for feature extraction is a good matrix for reconstruction and vice versa. This is not surprising as, in this educational scenario, the signal is accurately described by the set of parameters. Also this contribution enables to consider separately the two fundamental aspects of compressive architectures and their influence: bandpass compression on one hand, and diversity creation on the other.

From the application examples, following main points emerged: For amplitude estimation, the on-grid case yields a diagonal Fisher matrix, and Cramér-Rao bounds are exactly described by the inverse of the number of samples, hence by the compression ratio. Otherwise performance depends on the crossterms of $\boldsymbol{J}$, and therefore on correlations between the columns of the sensing matrix that belong to the support. Since the precision of frequency and amplitude estimation can be controlled for each pair of subbands through the cross-correlation properties of the sensing matrix, a simple way is pointed out to optimize performances based on priors on the input spectrum distribution.

It is highlighted that choices on the structure of a pseudo-random $\boldsymbol{\Phi}$ may favor non uniform performances. Hence it is sketched within a computationally-lighter framework than Bayesian approaches how to optimize the performances given prior on the parameter distribution. In Chapter 3, the question of the optimal choice of code parameters was raised. An interesting lead would be to optimize the parameters of the code (for example the row selection process) in order to enhance the Cramér-Rao bounds between specific portions of the spectrum. Note that in this setting, an average or cumulative coherence might be more relevant than usual

coherence.

Compared to the interference cancellation method proposed in [146], the adaptive approach suggested here is more flexible and simpler as it is not necessary to generate new codes at every interferer detection.

The proposed framework is relatively general as it concerns all types of parameters and all compressive multiband architectures, whether for real- or complex-valued signals (after frequency downconversion). Also the effect of spectral leakage, which is rarely tackled even with Nyquist rate samples, is discussed. The limitation could rather be found in the signal model, and the assumption of bandlimited representation of the samples of the signal (Hyp. 2).

Cramér-Rao bounds expression for other wideband RF receivers and wider class of signals ought to be exposed as well. Also the Cramér-Rao bound is given for unbiased estimators, so it might be of interest to look into other bounds, more accurate at low SNR than the Cramér-Rao bound to picture the threshold effect, for example the Barankin bound [147] or the Ziv-Zakai bound [148].

In the chosen framework, no parameter is better than others regarding the robustness to dimension reduction. It means that, given a particular sensing matrix, $\boldsymbol{B}$ for compressive multiband acquistion, there is no set of parameters that is more resilient to the detrimental effects of compressive sensing. Hence, there is no preference inherent to the sensing matrix that governs the choice of a set of parameters with whom to possibly perform classification afterwards.

The most interesting question to investigate remains whether (biased) estimators with lower variance can be built and how.

Conclusion and perspectives

Conclusion

In this work, the potential of radiofrequency receivers based on Compressive Sampling for Cognitive Radio applications has been investigated.
Through this manuscript, another small bridge has been built to help overcome the gap between the appealing abstract objects and concepts of the mathematical theory of Compressive Sensing and realities and stakes of Radiofrequency Spectrum Sensing.

First and foremost, Chapter 1 started by introducing the urge to develop smart and adaptive radiofrequency receivers in order to meet the technical challenges of low-power Dynamic Spectrum Sharing. Then the Compressive Sampling theory has been presented as a potential answer to this daunting task. Applied to Spectrum Sensing, Compressive Sampling allows to go substantially lower than the Nyquist acquisition rate and hence modifies the usual trade-offs, provided that the spectrum is sparse and the architecture is modeled by a matrix that satisfies some mathematical properties.

However, if acquisition is facilitated, the reconstruction of the original signal, on the other hand, requires an amount of energy that is not readily compatible with an embedded solution. Hence it has been proposed in this work to perform the extraction of the features of interest directly from the compressed samples. Theoretical guarantees as well as existing solutions and their limitations have been sketched.

The second Chapter has focused on the State-of-the-Art of compressive radiofrequency receivers. Architectures have been gathered according to three main functioning principles: Non Uniform Sampling, Random Demodulation and Variable Rate Sampling. However, it has been highlighted that most of them lack versatility and still require Nyquist-rate components.

Two receivers stood out: first, the Modulated Wideband Converter (MWC) which is based on the mixing with pseudorandom periodic sequences and showed an interesting potential for interferer detection among other things. Second, the Non Uniform Wavelet Bandpass Sampling (NUWBS), a variant of Non Uniform Sampling applied on wavelet coefficients instead of time samples which provides three appealing degrees of freedom and reconfigurability.

Chapter 3 has hence focused on the core and bottleneck element of all architectures based on random demodulation, the mixing codes. The Modulated Wideband Converter is chosen as a concrete example. First, the meaning of the notions of randomness and universality for practical matrices implementing a RF receiver in hardware was discussed. Also the potential of circulant matrices has been highlighted, based on the literature. Then a proposition was made of an original code sensing matrix, to be tested afterwards against the codes used in the state-of-the-art: circulant and defined by its diagonal in the frequency domain, which is a half-length Zadoff-Chu code on which hermitian symmetry is applied.

In a second part, a systematic study of the influence on various evaluation criteria of dif-

ferent choices of code families and of row selection, including our original proposition, was carried out. First, mathematical properties of the sensing matrix have been addressed. The coherence, which is vital to limit the number of branches, was investigated. Then an approach that focuses on evaluating norm and distance preservation, which is essential to noise robustness, was developed. The relevance of previous high-level metrics was validated by means of a simulation platform reproducing the acquisition and reconstruction process of a MWC, in both noiseless and noisy settings. From this study, it appeared that the proposed codes had better coherence and isometric properties and also provided better recovery performances than codes used in the state-of-the-art, especially in noisy environment. Another highlight was the importance of appropriate row selection. Altogether a methodology for a thorough comparison between usual compressive sensing matrices and new proposals was delivered.

We also paved the way for practical implementation of Zadoff-Chu codes in the NUWBS architecture.

In order to design a compressive receiver performing feature extraction, specificities of compressive parametric estimation must be clarified. That is why, in Chapter 4, the question of the accuracy of spectral parametric estimation directly from the compressed samples has been tackled. An analysis based on the Cramér-Rao lower bound on the variance of any unbiased estimator has been carried out. A new closed-form expression of the Fisher information matrix for samples from a compressive multiband architecture has been established: it is equal to the element-wise multiplication between the real part of the Gramian of the sensing matrix and the Fisher information matrix for bandpass sampling. Then it is shown how the Fisher information matrix for bandpass sampling can be expressed with respect to the Fisher information matrix for Nyquist sampling, under a small approximation. In contrast to the existing state-of-the-art, where bounds had so far only been given statistically on the Fisher matrix as a whole, these results are deterministically given for each coefficient and hence each parameter, a bounding which provides more accuracy. The established theorems also shed light on the mechanisms at stake in compressive estimation, in particular dissociate the effect of the two cornerstones of compressive sampling: compression on one hand and diversity creation on the other hand. In the process, it is highlighted that properties required from the sensing matrix regarding feature extraction are similar to those required for signal reconstruction, as it depends on the Gramian of the sensing matrix and hence on coherence and isometric properties.

In addition to the closed-form equation of the theorems, interferer detection, a common issue in Cognitive Radio, has been tackled through the examples of amplitude and frequency estimation of multitone signals. In particular the effects of spectral leakage, that might be very detrimental in presence of a strong interferer due to folding, were discussed. Based upon the newly established expressions of the Cramér-Rao bounds, it was then highlighted that the effects of this spectral leakage can be controlled by the coefficients of the Gramian of the sensing matrix. By adjusting the coefficients of the Gramian matrix, it is hence possible to control the precision of parametric estimation, offering interesting opportunities for adaptation to specific spectrum priors. As properties of coherence and isometry had been studied in details in Chapter 3, the optimization of the accuracy of parametric estimation can be done following the same guidelines.

Perspectives

As attested by the abundance of references in the recent literature, the level of current expectations regarding performances offered by compressive RF receivers is high. Some answers have been given, but so many questions remain open.

The future of Analog-to-Information Converter for Cognitive Radio will probably consists in a further hybridization between two elements: on one hand, the original theory of Compressive Sampling with universal statistical guarantees and on the other hand, additional application-related priors. For instance compressive modulation recognition [71] seems a promising technique. More generally, improvements through the proposition of innovative feature extraction algorithms from the compressed samples are expected. Also taking into consideration additional structure on top of sparsity, for example through manifold-like structures [54] or in the direction of the Finite-Rate of Innovation framework.

It would also be very interesting to analyze the possibilities of feature extraction with the NUWBS architecture in the light of Mallat's work on the scattering transform and the preservation of invariants [149, 150, 151]. Ideally, simple but adequately chosen embedded processing would allow to highlight the inherent structure of the information, which would make restitution of the features of interest possible from few samples.

Quantization is of course another vital aspect with respect to efficient implementation that was not adressed in this study, and quantization on few bits seems to entail an interesting potential.

On a different level, Cooperative Spectrum Sensing has not been tackled in this manuscript as it adds many other considerations but it is currently a major field of interest with respect to the Internet of Things.

Appendix A

Cramér-Rao bound

A.1 Fisher information matrix

Let us note with $\hat{.}$ an estimator. The probability density function $p(\boldsymbol{y};\boldsymbol{\theta})$ of a signal $\boldsymbol{y}$ as a function of a parameter vector $\boldsymbol{\theta}$ can be interpreted as a likelihood function. The 'sharper' the probability density function, the more accurate the parametric estimation can be. In fact the amount of information on parameters which is carried by a signal observation can be measured by the Fisher information matrix, noted $\boldsymbol{J}$. $\boldsymbol{J}$ is computed as the averaged second derivative of the log-likelihood function or as the covariance of the derivative of the log-likelihood:

$$\boldsymbol{J} = -E[\frac{\partial^2 ln(p(\boldsymbol{y};\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}^2}] = E[\frac{\partial ln(p(\boldsymbol{y};\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}}\frac{\partial ln(p(\boldsymbol{y};\theta))^H}{\partial \boldsymbol{\theta}}] \tag{A.1}$$

A.2 Cramér-Rao bound

More precisely, let us note with $\boldsymbol{\Gamma} \in \mathbb{R}^{Q\text{x}Q}$ the covariance matrix of an unbiased estimator of a Qx1 parameter vector $\boldsymbol{\theta}$. Then under mild conditions [53], $\boldsymbol{\Gamma}$ is lower bounded, in the sense of positive definiteness, by the Cramér-Rao lower bound (CRB) [53]:

$$\boldsymbol{\Gamma} \geq CRB(\theta) = \boldsymbol{J}(\boldsymbol{\theta})^{-1} \tag{A.2}$$

$$\boldsymbol{\Gamma} - \boldsymbol{J}(\boldsymbol{\theta})^{-1} \geq 0 \tag{A.3}$$

Therefore the unbiased estimation of θ_q, the qth parameter of the vector $\boldsymbol{\theta}$ ($q \in [\![1;Q]\!]$), is lower bounded by [53]:

$$var(\hat{\theta}_q) \geq CRB(\hat{\theta}_q) = (\boldsymbol{J}(\boldsymbol{\theta})^{-1})[q,q] \geq (\boldsymbol{J}(\boldsymbol{\theta})[q,q])^{-1} \tag{A.4}$$

where the two right hand terms are equal if and only if $\boldsymbol{J}$ is diagonal.
If $J[q,r] \neq 0$ for $q \neq r$, the additional parameter interferes, degrading performances. This coupling interaction is different from the correlation notion.
In this thesis, we are mainly concernend by Gaussian vectors, for which the computation of the CRB (A.2) can be done by a more specific form. Indeed, for a complex vector $y^+[\check{n};\theta] = y[\check{n};\theta] + w[\check{n}]$ where $\boldsymbol{w} \sim \mathcal{N}(0,\sigma^2\boldsymbol{I}_{NL})$ is a circular white complex Gaussian noise [53], for $(q,r) \in [\![1;Q]\!]^2$:

$$J[q,r] = \frac{2}{\sigma^2}\Re[\sum_{\check{n}=1}^{NL}\frac{\partial y^*[\check{n};\boldsymbol{\theta}]}{\partial \theta_q}\frac{\partial y[\check{n};\boldsymbol{\theta}]}{\partial \theta_r}] \tag{A.5}$$

where $\Re[.]$ denotes the real part. For a real vector $\boldsymbol{y}^+$:

$$J[q,r] = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{NL} \frac{\partial y[\check{n};\boldsymbol{\theta}]}{\partial \theta_q} \frac{\partial y[\check{n};\boldsymbol{\theta}]}{\partial \theta_r} \tag{A.6}$$

A.3 Efficient estimator

An estimator is efficient if and only if the probability density function comes from a family of exponentials (Darmois theorem). An estimator is called asymptotically efficient if it achieved the Cramér-Rao bound when the number of samples N approaches infinity. It is the case of Maximum Likelihood estimator under mild regularity conditions. An unbiased estimator does not always exist and it is not always the best estimator in the sense of minimal square error. The best unbiased estimator in sense of Minimum Square Error (MSE) is called the Minimum-Variance unbiased estimator (MVUE) and can be built from any unbiased estimator using the Rao-Blackwell theorem. There exists some other bounds, more accurate at low SNR than the Cramér-Rao bound in the picturing of the threshold effect, the Barankin bound [147] or the Ziv-Zakai bound [148].

Appendix B

Computation of the MWC sensing matrix expression

First, the expression of the Fourier coefficients of the shaped code sequences is established. Then it is proven that the sensing matrix $\boldsymbol{B}$ of the Modulated Wideband Converter used in eq. (2.6) is indeed given by these coefficients.

B.1 Expression of the Fourier coefficients of the shaped code sequences

For sake of simplicity, the case where N is odd will be considered. As the mixing code in the mth branch of the MWC receiver is assumet to be Tp-periodic, it admits following Fourier decomposition:

$$p_m(t) = \sum_{n'=-\infty}^{+\infty} \gamma_{m,n'} e^{2i\pi t \frac{n'}{T_p}} \tag{B.1}$$

Its restriction to a period T_p is given by:

$$p_{m,[T_p]}(t) = \sum_{q=0}^{N-1} \alpha_m(q) g(t - qT_c) \tag{B.2}$$

where $\boldsymbol{\alpha_m}$ corresponds to the N code elements and g is a shaping function, by default a unit-amplitude rectangular window of length T_c. Notations are illustrated in Fig. B.1. Thus the Fourier Transform of the signal after mixing is given by the sum of weighted subbands:

$$\begin{aligned} FT[x(t)p_m(t)] = \tilde{x}(f) * \tilde{p}_m(f) = & \textstyle\sum_{n'=-\infty}^{+\infty} \gamma_{m,n'} \tilde{x}(f - \frac{n'}{T_p}) \\ = & \textstyle\sum_{n'=-\infty}^{+\infty} \gamma_{m,-n'} \tilde{x}(f + \frac{n'}{T_p}) \end{aligned} \tag{B.3}$$

Yet the Fourier coefficients are given by:

$$\begin{aligned} \gamma_{m,n'} = \frac{1}{T_p} \int_0^{T_p} p_m(t) e^{-j(\frac{2\pi}{T_p})n't} dt = \frac{1}{T_p} \int_{-\infty}^{\infty} p_{m,[T_p]}(t) e^{-j(\frac{2\pi}{T_p})n't} dt \\ = \frac{1}{T_p} \tilde{p}_{m,[T_p]}(f = \frac{n'}{T_p}) \end{aligned} \tag{B.4}$$

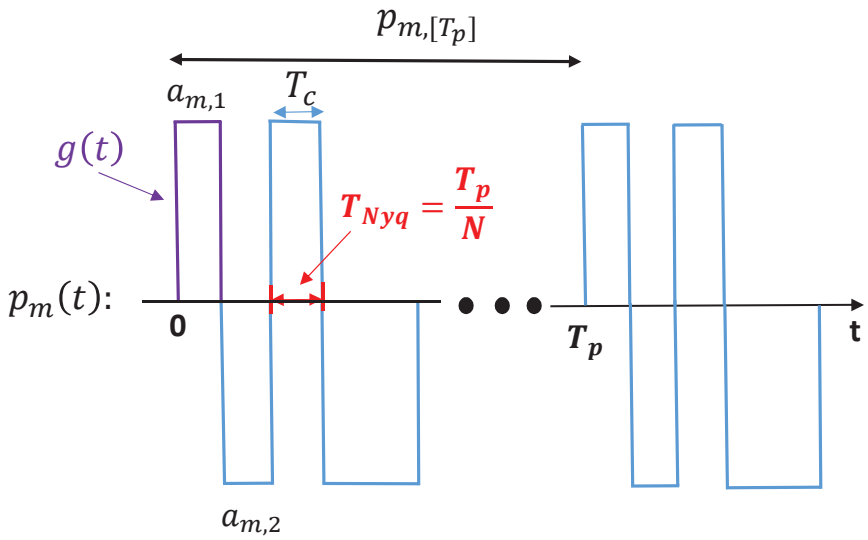


Figure B.1: Notations related to the mixing function $p_m(t)$.

where $\tilde{p}_{m,[T_p]}(f)$ is the Fourier transform of the finite duration analog code $p_{m,[T_p]}(t)$ which can be rewritten in the time domain as a convolution product:

$$p_{m,[T_p]}(t) = (\sum_{q=0}^{N-1} \alpha_m(q).\delta(t - qT_c)) * g(t) \tag{B.5}$$

Hence in the frequency domain, the term in eq. (B.4) is given by:

$$\tilde{p}_{m,[T_p]}(\frac{n'}{T_p}) = \tilde{a}_m(\frac{n'}{T_p}).\tilde{g}(\frac{n'}{T_p}) \tag{B.6}$$

where

$$\tilde{a}_m(f) = \sum_{q=0}^{N-1} \alpha_m(q) e^{-j2\pi q T_c f} \tag{B.7}$$

is the analog Fourier Transform of the Discrete Sequence and $\tilde{g}(f)$ is the Fourier transform of $g(t)$. Fig. B.2 illustrates previous formula.

Note that the effect of shaping is seldom, if ever, mentioned in the literature but it is clear from the above equations that shaping implies that the compressive sampling acquistion grants less measurement power to higher frequencies. As a consequence, higher frequencies have lower probability to be recovered, i.e. recovery guarantees are not uniform.

Finally, by replacing (B.5) in (B.4) for rectangular window shaping, following expression of the Fourier coefficients is obtained:

$$\gamma_{m,n'} = \frac{1}{T_p}\tilde{g}[\frac{n'}{T_p}]DFT(\alpha_m)[n'] = \frac{\tilde{a}_m[n' f_p]}{T_p} T_c \text{sinc}(\frac{\pi n'}{N}) e^{-j\frac{\pi n'}{N}} \tag{B.8}$$

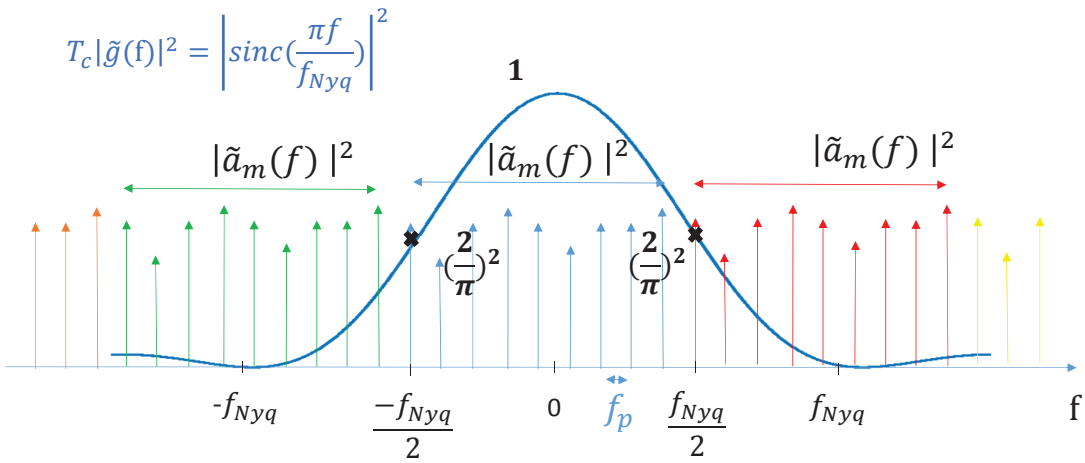


Figure B.2: Power Spectral Density of the shaped discrete Sequence $\tilde{p}_{m,[T_p]}(f)$.

For sake of clarity, let us note with $\boldsymbol{\sigma}$ the Discrete Fourier Transform (DFT) of the vector containing the code sequence $\boldsymbol{a_m}$:

$$\boldsymbol{\sigma}_{[\![1;N]\!],m} = \boldsymbol{F}\boldsymbol{a}_{[\![1;N]\!],m} \tag{B.9}$$

or with reindexing:

$$\boldsymbol{\sigma}_{[-N_0;N_0],m} = \bar{\boldsymbol{F}}\boldsymbol{a}_{[\![1;N]\!],m} \tag{B.10}$$

where $\boldsymbol{F}$ is the Discrete Fourier Transform matrix of dimension $N = 2N_0 + 1$ and $\bar{\boldsymbol{F}}$ is a flipped row version of $\boldsymbol{F}$ so that $\bar{\boldsymbol{F}}_{m,n} = e^{-j\frac{2\pi(N-m)n}{N}}$.
Let us define the diagonal matrix of dimension N, for $n \in [\![1;N]\!]$:

$$\boldsymbol{D} = diag(\frac{1}{N}\mathrm{sinc}(\frac{\pi n}{N})e^{-j\frac{\pi n}{N}}) \tag{B.11}$$

Then it is possible to rewrite as a matrix expression eq. (B.8):

$$\boldsymbol{\Gamma}_-^T = \boldsymbol{D}\boldsymbol{\sigma} = \boldsymbol{D}\bar{\boldsymbol{F}}\boldsymbol{A}^T \tag{B.12}$$

where $\boldsymbol{A}$ and $\boldsymbol{\Gamma}_-^T$ correspond to NxM matrices containing the coefficients $a_{m,n}$ and $\gamma_{-n',m}$ respectively. Hence after transpose:

$$\boldsymbol{\Gamma}_- = \boldsymbol{A}\bar{\boldsymbol{F}}\boldsymbol{D} \tag{B.13}$$

because $\bar{\boldsymbol{F}}^T = \bar{\boldsymbol{F}}$ and $\boldsymbol{D}^T = \boldsymbol{D}$.

B.2 Sensing matrix expression

The input signal $\tilde{x}(f)$ belongs to a union of bandpass signals hence it can be rewritten as a sum of shifted support limited functions:

$$\tilde{x}(f) = \sum_{k=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \tilde{\beta}_k(f - kf_p) \tag{B.14}$$

where $\tilde{\beta}_k(f) = 0$ if $f \notin [-\frac{f_p}{2}; \frac{f_p}{2}]$.
Hence the last term of eq. (B.3) can be rewritten:

$$FT[x(t)p_m(t)] = \sum_{k=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_k(f - (k - n')f_p) \tag{B.15}$$

After filtering, only $k = n'$ is kept at baseband since $\tilde{\beta_k}$ are support limited:

$$\tilde{y}_m(f) = FT[x(t)p_m(t)]_{BB} = \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_{n'}(f) \tag{B.16}$$

For a given acquisition time, suppose that the spectrum does not have very quick fluctuations, we can decompose the continuous $\tilde{\beta_k}$ on the l discrete frequency bins by averaging on each one for $k \in [\![-\lfloor N/2 \rfloor; \lfloor N/2 \rfloor]\!]$:

$$\tilde{\beta}_k(f) = \sum_{l=1}^{L} \tilde{\beta}_{k,l}(f - (l-1)\delta f) \tag{B.17}$$

Hence eq. (B.16) can be rewritten as:

$$\tilde{y}_m(f) = \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta_{n'}}(f) = \sum_{l=1}^{L} \sum_{n'=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \gamma_{m,-n'} \tilde{\beta}_{n',l}(f) \tag{B.18}$$

where $\tilde{\beta}_{n',l}(f)$ represents the $\tilde{n}^{th} = (n-1)L + l^{th}$ bin of the input spectrum and $\tilde{\beta}_{n'}(f)$ represents the concatenation of the l bins of the nth subband.
This is because n' spans $[\![-\lfloor N/2 \rfloor; \lfloor N/2 \rfloor]\!]$ but index of the sensing matrix n spans $[\![1; N]\!]$.
Eq. (B.18) highlights perfectly the equivalence between the canonical model and the multiband compact formalism described in Subsection 2.1.3. In other words:

$$\begin{aligned} \tilde{y}_m[l] &= \tilde{y}_m(f = (l-1)\delta f - \tfrac{f_p}{2}) \\ &= \textstyle\sum_{n=1}^{N} b_{m,n} \tilde{x}((l-1)\delta f + ((n - \lfloor \frac{N}{2} \rfloor - \frac{3}{2}).f_p) \end{aligned} \tag{B.19}$$

for $l \in [\![1; L]\!]$. It also establishes that the acquisition matrix $\check{\boldsymbol{B}}$ in eq. (2.3) is indeed given by coefficients $b_{m,n} = \gamma_{m,-(n-\lfloor \frac{N}{2} \rfloor - 1)}$ with $\gamma_{m,n}$ defined by eq. (B.1) as the Fourier coefficients of the shaped mixing codes. Hence in combination with (B.13) it can be concluded that:

$$\boldsymbol{B} = \boldsymbol{A}\bar{\boldsymbol{F}}\boldsymbol{D} \tag{B.20}$$

■

Appendix C

Numerical benchmark elements between acquisition methods

To underpin the potential of the Quadrature-to-Analog-Information Converter (QAIC) presented in 2.3.2, a comparison benchmark of interferer detection performances is reported.

TABLE III
Comparison of the Measured Performance of the QAIC Interferer Detector to the State of the Art in Integrated Spectrum Sensors and Scanners

		[5, 6, 36]	[9]	[10]	[11]	[14, 37]	[23]	[24]	This Work
Architecture		Cross-Correlation SA	Dual Up/ Down Conversion SA	Multi-Resolution Spectrum Sensing	Tunable RF Filter & Dynamic-Range-Scalable Energy Detector	Digital & Analog Multi-Band Sensing (Hybrid SA)	Random-Modulation Pre-Integrator CS	Parallel Path Time Segmented CS	**CS with Quadrature Analog-to-Information Converter**
System Config.		Cross Correlation					Low pass CS	Low pass CS	**Band pass CS**
Application		Spectrum Analysis	Spectrum Analysis	Spectrum Analysis	Spectrum Analysis	Spectrum Analysis	Sub Nyquist Receiver	Sub Nyquist Receiver	**Sub Nyquist Interferer Detector**
CMOS Tech.	**[nm]**	65	130	180	90	40	90	90	**65**
Die Area	**[mm^2]**	< 0.26 (active area)	14.43	11.52	2.3	5	8.85	0.93	**1.96**
Supply Voltage	**[V]**	1.2	1.8	1.8	1.2	1.1	1.5 (A), 1.2 (D)	N/R	**1.1**
Frequency Range		50MHz-1.5GHz	100Hz-6GHz	600-603MHz[1]	30-334MHz (LPF) 334-800MHz (BPF) 800MHz-2.4GHz (Bypass)	500MHz-2.5GHz[2]	100MHz-2GHz	5-500MHz	**2.7-3.7GHz**
Number of Branches		4 = 2 x I/Q	2 = 1 x I/Q	2 = 1 x I/Q	2 = 1 x I/Q	2 = 1 x I/Q	8	8	**16 = 8 x I/Q**
RBW Options	**[MHz]**	1	0.4-11 (step 0.5)	0.194	0.2-30	1-40	N/R	5	**10, 20[3]**
PN Seq. Clock Rate	**[GHz]**	N/A	N/A	N/A	N/A	N/A	4	1.25	**1.26**
Sensitivity	**[dBm]**	-81 / -109[4,5]	-82[6]	-74	-83	N/R	-64[7]	N/R	**-72[8]**
Power	**[mW]**	166[9,10]	678.6[9]	180	30-44	33-99	506.4[9,11]	55[9,12]	**81[9,13]**

N/A = Not Applicable; N/R = Not Reported

[1] Measurements are shown for 600–603 MHz signals; [2] Demo frequency range is reported; [3] RBW of 20 MHz is the nominal setting for length 63; Fig. 15 shows RBW of 10 MHz (length 127) and RBW of 20 MHz (length 63); [4] Calculated the −81 dBm from the −141 dBm/Hz DANL for a 1 MHz RBW and calculated the −109 dBm from the −169 dBm/Hz DANL for a 1 MHz RBW; [5] After cross-correlation, the sensitivity is −109 dBm for a 1 MHz RBW; [6] Maximum gain setting; [7] For a single tone detection; [8] For a single band detection; [9] ADC off-chip; [10] Estimated DSP power consumption is 25 mW; [11] Without output buffers; [12] Estimated ADC power is less than 1 mW (neglected) and estimated PLL power is 1mW; [13] 400 fF on-chip capacitors are included to emulate the ADC load.

Figure C.1: Comparison of interferer detection performances, from [6].

Appendix D

Exploration of time modulation in the NUWBS architectures with Zadoff-Chu codes controlled delays

In this appendix, an investigation of the use of Zadoff-Chu codes for time modulation in the Non Uniform Wavelet Bandpass Sampling (NUWBS) scheme is proposed. The NUWBS has demarcated itself as a promising candidate from the overall comparison in Chapter 2 due to its high degree of flexibility and versatility, a valuable asset for Cognitive Radio applications.

In [99], a hardware-in-the-loop proof of the intraband reconstruction capabilities of the NUWBS for fixed central frequency f_c and time support τ has been demonstrated, based on a previous Application-Specific Integrated Circuit (ASIC) developed in our CEA-LETI laboratory. To go further, the exploration of the potential of the NUWBS architecture for Spectrum Sensing is proposed through time modulation with code sequences. It is chosen not to reveal in details the implementation aspect in this manuscript.

The core idea is to use a T_{acq}-periodic Zadoff-Chu sequence, with rapidity rate of f_p matching the wavelet rate and sampling rate f_s ($f_p = f_s$). For each wavelet, the phase of the corresponding Zadoff-Chu element is used to command an additional delay on the starting time of each wavelet. As a first step to simplify the problem, it will be considered that there is no flexibility on time supports τ and central frequencies f_c, which will be the same for all wavelets. A potential further improvement in the ability of the acquisition sensing matrix to generate diversity is foreseen.

D.1 Periodicity of the normalized sensing matrix of the NUWBS

The different aliasing zones of the input spectrum are sketched in Figure D.1. Note as BW_{RF} the width of the occupied RF bandwidth and f_i the frequency of another active signal with whom aliasing could happen. "Intraband" refers to the frequency range $[f_i - f_s; f_i + f_s]$. The term "interband" corresponds to the physical quantity of the signal frequency bandwidth, BW_{RF} and denotes the frequency range $[f_i - BW_{RF}; f_i + BW_{RF}]$. "Out-of-band" indicates frequencies further away than the "interband" range but belonging to the Nyquist band.

In this study, it is desired to consider the acquisition process independently of the beneficial windowing effect of the envelope. Hence it is appropriate to introduce the matrix $\boldsymbol{A_n} \in \mathbb{C}^{LxN}$,

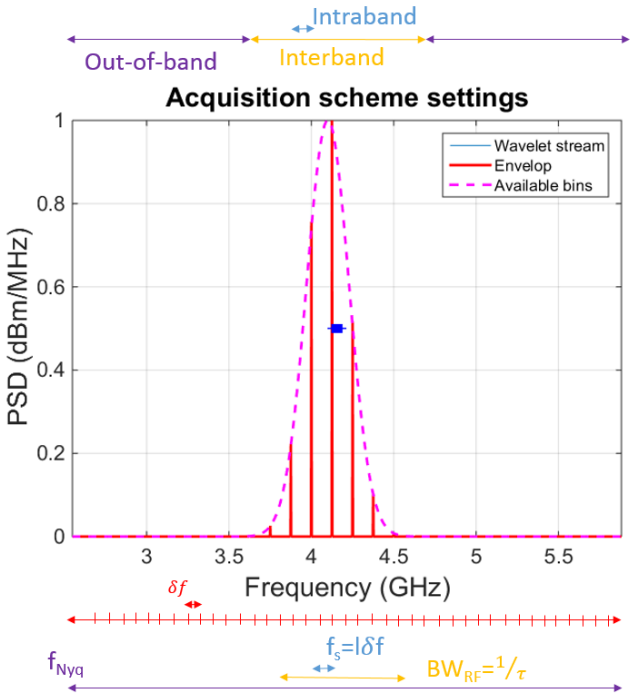


Figure D.1: Definition of the different frequency aliasing zones.

whose elements have the same phase as those from $\boldsymbol{A}$ but their module is normalized to 1. From eq. (2.9), it appears that columns of $\boldsymbol{A_n}$ are f_p-periodic in frequency.

The periodicity appears because the subsampling period is chosen below the Nyquist rate. This well-known aliasing phenomenon is similar to what happens in classical bandpass sampling. Unlike the former, the NUWBS benefits from the pulse envelope which acts as a selective band-pass filter. However, a potential point of interest would be to know whether the discrimination capabilities guaranteed by the decreasing of the Gaussian envelope can be further reinforced for signals spaced from f_s, for which the corresponding normalized columns of the sensing matrix are identical. Hence, in the following, time modulation is proposed to temper the periodicity.

D.2 NUWBS architecture with Zadoff-Chu modulated wavelet stream

In the proposed system, a delay is added with respect to the expected periodic starting date of each wavelet of the NUWBS. This new architecture will be referred to as ΔZAC-WBS.
For sake of simplicity, a single branch architecture is considered.
Consider $\boldsymbol{Z_\Delta} \in \mathbb{C}^{LxN}$, the sensing matrix corresponding to the acquisition with a ΔZAC-WBS. As pictured on Figure D.2, the application of the delays can be expressed as a convolution product in the time domain with a matrix $\boldsymbol{C} \in \mathbb{C}^{LxN}$ such that the sensing matrix $\boldsymbol{Z_\Delta}$ is

given by:

$$\boldsymbol{Z_\Delta} = \boldsymbol{R}\boldsymbol{A_\Delta} = \boldsymbol{R}\boldsymbol{C} \circ \boldsymbol{A} = \boldsymbol{R}\boldsymbol{C} \circ \boldsymbol{W^H}\boldsymbol{F^{-1}} \tag{D.1}$$

where it is recalled that $\boldsymbol{R}$ is a matrix selecting a subset of rows, $\boldsymbol{A} = \boldsymbol{W^H}\boldsymbol{F^{-1}}$ and $\boldsymbol{C}$ representing the delays is defined as follows:

$$C[k,n] = e^{-j2\pi\epsilon_k n\delta f} \tag{D.2}$$

with the delay on the kth wavelet, noted ϵ_k, that can be linked to a phase ϕ_k through the relationship $\epsilon_k = \frac{\phi_k}{2\pi f_c}$.

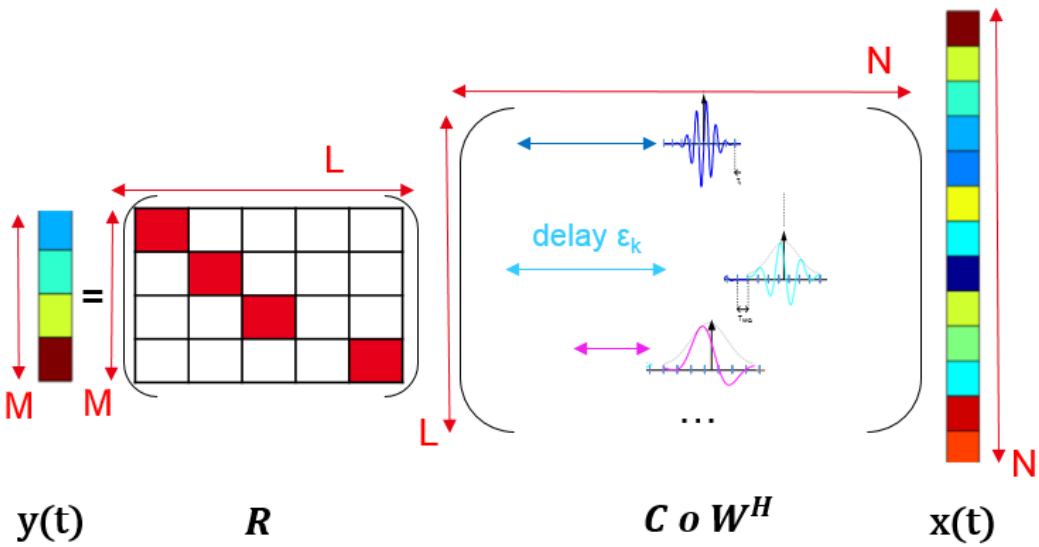


Figure D.2: Decomposition of the acquisition matrix of a NUWBS architecture with Zadoff-Chu modulated wavelet stream.

The expression of the kth delayed measurement vector in the time domain is given by [98]:

$$\Psi_{\Delta,\delta_k}(t) = \Psi_{\delta_k}(t-\epsilon_k) = \frac{2^{1/4}}{\sqrt{\tau}(\pi)^{1/4}} e^{-j2\pi f_c(t-\epsilon_k-\delta_k)} e^{\frac{-(t-\epsilon_k-\delta_k)}{\tau}} \tag{D.3}$$

Hence the Fourier Transform of eq. (D.3) is given by [98]:

$$\check{\Psi}_{\Delta,\delta_k}(f) = FT[\Psi_{\delta_k}(t)].e^{-j2\pi\epsilon_k f} = (\tau\sqrt{2\pi})^{1/2} e^{-j2\pi\delta_k f} e^{-(\pi\tau(f-f_c))^2}.e^{-\frac{\phi_k}{f_c}f} \tag{D.4}$$

Replacing δ_k as in eq. (2.9), and replacing ϕ_k by assuming that it is chosen to be the phase of a Zadoff-Chu sequence according to eq. (3.4) so that $\epsilon_k = \frac{Rk^2}{2f_cL}$ yields:

$$a_\Delta[k,n] = (\tau\sqrt{2\pi})^{1/2} e^{-j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})n\delta f} e^{-(\pi\tau(n\delta f-f_c))^2} \tag{D.5}$$

where R is the index of the Zadoff-Chu sequence and δf corresponds to the frequency resolution so that $f = n\delta f$.

For sake of dissociating the effect of windowing, the matrix $\boldsymbol{A_{n\Delta}} \in \mathbb{C}^{LxN}$, whose elements

have the same phase as those from $\boldsymbol{A_\Delta}$ but the module is normalized to 1 is also introduced. The expression of its elements is then:

$$a_{n\Delta}[k,n] = e^{-j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})n\delta f} \tag{D.6}$$

By considering this equation, it appears that the application of a delay amounts in fact to defining a new equivalent sampling grid with sampling instant defined by $(\delta_k + \epsilon_k).T_p$. That is, it is possible to rewrite (D.2) as:

$$a_{n\Delta}[k,n] = \Psi_{n\Delta,\delta_k}(n\delta f) = e^{-j2\pi(\delta_k+\epsilon_k)n\delta f} \tag{D.7}$$

This point of view reminds for instance of chirp sampling [152, 153], a method consisting in sampling at instants given by the zero-crossings of a chirp. A parallel could perhaps also be drawn with the choice of the optimal sampling pattern in the Non Uniform Sampling (NUS), a problem tackled for instance by S. Traoré in [93].
Then for indices $(n,n') \in \mathbb{Z}^2$, the correlation $\Gamma_{n,n'}$ between two columns of the normalized sensing matrix $\boldsymbol{A_{n\Delta}}$ is given by:

$$\begin{aligned}\Gamma_{n,n'} &= \sum_{k=1}^{L} e^{j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})n\delta f} e^{-j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})n'\delta f} \\ &= \sum_{k=1}^{L} e^{-j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})(n'-n)\delta f}\end{aligned} \tag{D.8}$$

If a distance $f_s = L.\delta f$ is considered, which corresponds to an index $n' = n + L$, previous equation becomes:

$$\Gamma_{n,n+L} = \sum_{k=1}^{L} e^{-j2\pi(\frac{k}{f_s}+\frac{Rk^2}{2f_cL})f_s} = \sum_{k=1}^{L} e^{-j2\pi\frac{Rk^2}{2f_c}\delta f} \tag{D.9}$$

This expression is smaller than L in general. A given delay entails a different phase depending on frequency, thus creating some diversity. In particular for R and k odd and f_c multiple of δf, a first insight highlights that the fraction $\frac{Rk^2\delta f}{2f_cL}$ can not be an integer since Rk^2 is odd and $2f_cL$ contains a factor 2. Hence Γ is not f_s- periodic anymore.

Previous equation is then analyzed more thoroughly for varying parameters R and f_c, and a fixed distance f_s (i.e. $n' - n = L$), in Figure D.3. For an appropriate choice of $(R, \frac{f_c}{f_s})$ the cross-correlation $\Gamma_{n,n+L}$ in the ΔZAC-WBS is much smaller (i.e. better) than $L = 32$, the worst case of the NUWBS. For the optimal parameter set $[R = 55,\ f_c = 6.f_s]$, a gain of $-24.41dB$ is achieved. For $f_c = 4GHz = 32.f_s$ which corresponds to the available hardware used in [99], a gain of $-15.68dB$ is achieved for optimal $R = 62$. This shows that introducing pseudo-random delays at each wavelet may indeed allow to prevent further interband aliasing through f_s-periodicity.

The f_s-periodicity, which is characterized by the analysis for $n' = n + L$, was the most obvious issue. Varying n' allows to go further in evaluating the amount of information brought globally by the acquisition process. It is similar to the analysis of coherence, which would then be the maximum correlation over all possible columns.

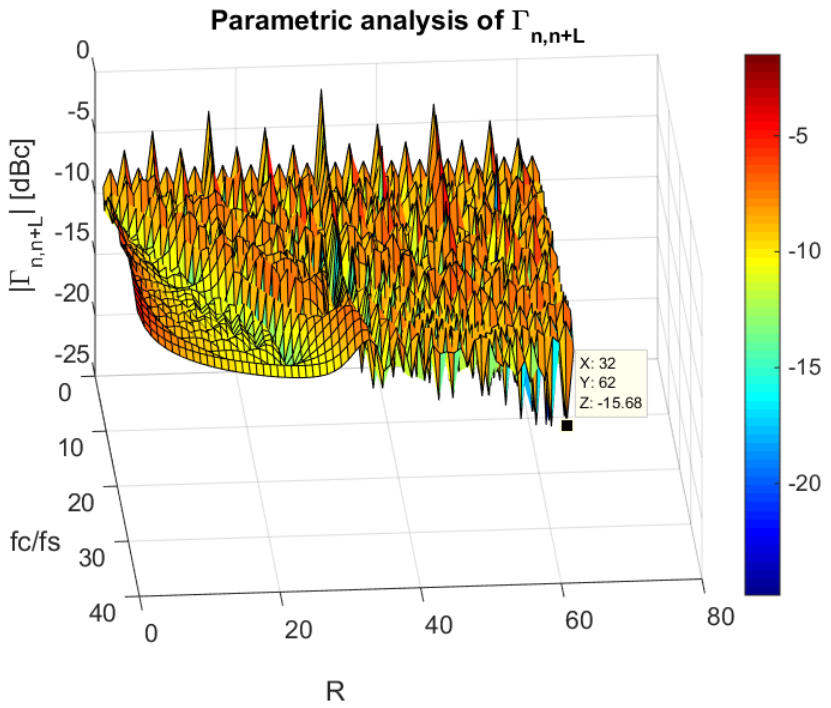


Figure D.3: Module of the correlation $\Gamma_{n,n+L}$ between f_s-spaced columns of the ΔZAC-WBS sensing matrix, $[L = 32, R = 1,\ f_c = 4GHz,\ f_s = 125MHz]$.

Parameters are fixed according to the hardware constraint $f_c = 4GHz$ and with the optimal R found in the previous parametric analysis: $R = 62$, $f_s = 125MHz$ and $f_{Nyq} = 32GHz$. In Figure D.4, the module of the crosscorrelation is pictured for the NUWBS and the ΔZAC-WBS respectively, in dBc. The reference column is set to $f = f_c$ but note that since eq. (D.9) points out that the result only depends on the distance $n - n'$, any reference point could have been chosen. For the NUWBS (red dashed), the module of the correlation is a f_s-spaced Dirac comb distribution. However, for the ΔZAC-WBS (blue), it appears that the f_s-periodicity has been broken: For $f = f_c \pm f_s$, the amplitudes of the intercorrelation peak with f_c is at $-15.68dBc$. Hence the proposed method achieves the targeted purpose: taming the periodicity ofthe measurement. In inter-peak zones, coefficients are not zeros anymore hence orthogonality within a subband is not ensured anymore, but this does not necessarily mean that the signals cannot be discrimated.

The use of Zadoff-Chu codes for time modulation in the NUWBS has been sketched. A correlation analysis showed that the f_s-periodicity of the normalized sensing matrix is indeed broken. The interest of this compromise on the overall performances of the RF receiver will be investigated further, notably through a practical performance benchmark with the NUWBS simulation platform. Also, further interpretation of the spectrum of the measurement vectors with respect to properties of Zadoff-Chu codes could shed additional light on the acquisition process.

These elements have been studied further, leading to the patent deposit [142].

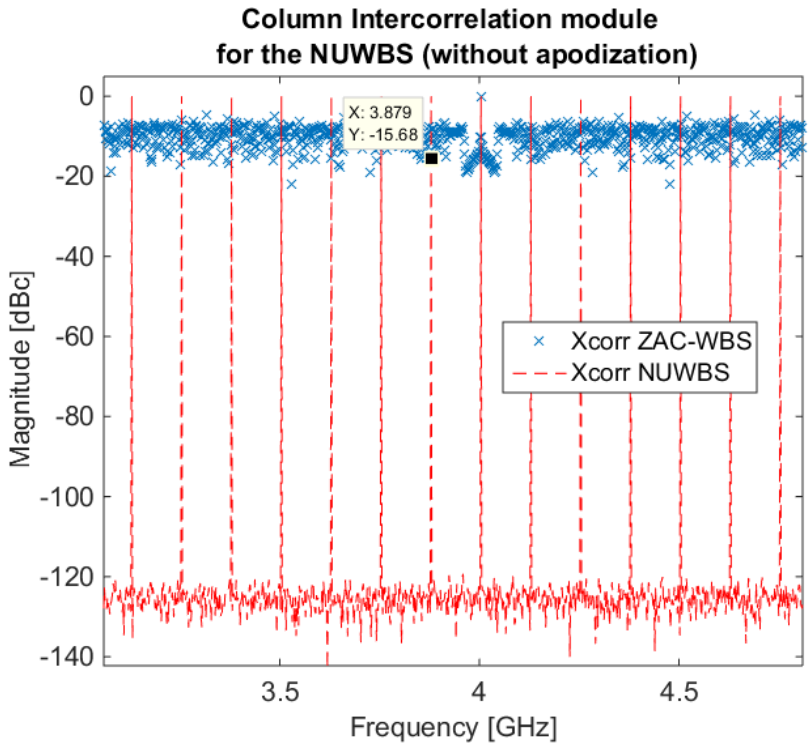


Figure D.4: Module of the correlation for the ΔZAC-WBS, in dBc, for varying frequency $n'\delta f$ and fixed $n\delta f = f_c$.

Appendix E

Proof of Theorem 1: Link between Fisher matrices for Compressive Sampling and Subsampling

---

First, (A.5) is derived, then (A.6) as a simplification. According to (A.5),(1.1) the $(q,r)^{th}$ element of the Fisher matrix from the compressed samples of a multiband architecture (MB) $\boldsymbol{J_{MB}}$ is given by:

$$J_{MB,[q,r]} = \frac{2}{\sigma_{CS}^2}\Re[\sum_{\check{m}=1}^{ML}\frac{\partial y^*[\check{m}]}{\partial\theta_q}\frac{\partial y[\check{m}]}{\partial\theta_r}] \tag{E.1}$$

where the compressed noise satisfies $\sigma_{CS}^2 = \sigma^2$ since $\boldsymbol{B}$ has column energy normalization to M. Due to orthonormality and independence toward parameters of the Discrete Fourier Transform matrix $\boldsymbol{F}$, it is also true in the spectral domain:

$$J_{MB,[q,r]} = \frac{2M}{N\sigma^2}\Re[\sum_{\check{m}=1}^{ML}\frac{\partial\tilde{y}^*[\check{m}]_{MB}}{\partial\theta_q}\frac{\partial\tilde{y}[\check{m}]_{MB}}{\partial\theta_r}] \tag{E.2}$$

For Bandpass sampling:

$$\tilde{y}_{BP}(f) = \sum_{k=-\infty}^{+\infty}\tilde{x}(f-kf_p) \tag{E.3}$$

The $(q,r)^{th}$ element of the Fisher matrix is given by:

$$J_{BP}[q,r] = \frac{2}{N\sigma^2}\Re[\sum_{l=1}^{L}\frac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}] \tag{E.4}$$

where from the discretization of (E.3) the sample derivatives are given by:

$$\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_q} = \sum_{n=1}^{N}\frac{\partial\tilde{x}[L(n-1)+l]}{\partial\theta_q} = \frac{\partial\tilde{x}_q[L(\lambda_q-1)+l]}{\partial\theta_q} \tag{E.5}$$

where the second equality is due to the subband model (Hyp.1). Whereas for Compressive Multiband Sensing, the samples derivatives are given, from (B.19), by:

$$\begin{aligned}\frac{\partial\tilde{y}_m[l]_{MB}}{\partial\theta_q} &= \sum_{n=1}^{N} b_{m,n}\frac{\partial\tilde{x}[L(n-1)+l]}{\partial\theta_q}\\ &= b_{m,\lambda_q}\frac{\partial\tilde{x}_q[L(\lambda_q-1)+l]}{\partial\theta_q} = b_{m,\lambda_q}\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_q}\end{aligned} \tag{E.6}$$

where the second equality is due to (Hyp.1) and the third by recognizing (E.5). Hence:

$$\sum_{m=1}^{M}\sum_{l=1}^{L}\frac{\partial\tilde{y}_m^*[l]_{MB}}{\partial\theta_q}\frac{\partial\tilde{y}_m[l]_{MB}}{\partial\theta_r}=\sum_{m=1}^{M}b_{m,\lambda_q}^*b_{m,\lambda_r}\sum_{l=1}^{L}\frac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r} \tag{E.7}$$

Yet:

$$\Re[b_{m,\lambda_q}^*b_{m,\lambda_r}\tfrac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\tfrac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}]=$$
$$\Re[b_{m,\lambda_q}^*b_{m,\lambda_r}]\Re[\tfrac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\tfrac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}]-\Im[b_{m,\lambda_q}^*b_{m,\lambda_r}]\Im[\tfrac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\tfrac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}] \tag{E.8}$$

One may recognize that:

$$\Im[\frac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}]=-\Re[i\frac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}\frac{\partial\tilde{y}[l]_{BP}}{\partial\theta_r}]. \tag{E.9}$$

Since $i\frac{\partial\tilde{y}^*[l]_{BP}}{\partial\theta_q}=\frac{\partial e^{j\frac{\pi}{2}}\tilde{y}^*[l]_{BP}}{\partial\theta_q}$, it therefore corresponds to a simple additional dephasing of $\frac{\pi}{2}$ between the two subsignals. So according to (E.2) and (E.4), it is possible to rewrite matrix-wise for $(q,r)\in[\![1;Q]\!]^2$:

$$\boldsymbol{J_{MB}}=\Re[\boldsymbol{B_\Lambda^H}\boldsymbol{B_\Lambda}]\circ\boldsymbol{J_{BP}}+\Im[\boldsymbol{B_\Lambda^H}\boldsymbol{B_\Lambda}]\circ\boldsymbol{\check{J}_{BP}} \tag{E.10}$$

where $\check{J}_{BP}[q,r]$ corresponds to the $(q,r)^{th}$ coefficient of the Fisher matrix for Bandpass Sampling in a virtual scenario where the initial dephasing between components x_{λ_q} and x_{λ_r} is increased by $\frac{\pi}{2}$.

Similarly but more simply in the real case, by using eq. (A.6) $\boldsymbol{J_{MB}}$ can be rewritten matrix-wise for $(q,r)\in[\![1;Q]\!]^2$ into:

$$\boldsymbol{J_{MB}}=\Re[\boldsymbol{B_\Lambda^H}\boldsymbol{B_\Lambda}]\circ\boldsymbol{J_{BP}} \tag{E.11}$$

■

Appendix F

Proof of Theorem 2: Link between Fisher matrices for Subsampling and Nyquist Sampling

---

We will first note that $J_{BP}[q,r](\Delta f_{q,r})$ is f_p-periodic and then demonstrate that $J_{BP}[q,r](\Delta f_{q,r}) = \frac{1}{N} J_{Nyq}[q,r](\Delta f_{q,r})$ for $\Delta f_{q,r} \in [-\frac{f_s}{2}; \frac{f_s}{2}]$.
First, notice from (E.3) that $\tilde{\boldsymbol{y}}$ is f_p-periodic, hence it follows from $(E.4)$ that $J_{BP}[q,r](\Delta f_{q,r})$ is f_p-periodic.
Concerning the second point, the Fisher information matrix for a Nyquist-rate acquisition is given by:

$$J_{Nyq}[q,r] = \frac{1}{\sigma^2} \sum_{l=1}^{L} \sum_{n=1}^{N} \frac{\partial \tilde{x}^*[(n-1)L+l]}{\partial \theta_q} \frac{\partial \tilde{x}[(n-1)L+l]}{\partial \theta_r} \tag{F.1}$$

According to Hyp. 2, for a given $l \in [\![1; L]\!]$:

$$\frac{\partial \tilde{x}_q[(n-1)L+l]}{\partial \theta_q}, \forall q \in [\![1; Q]\!] \tag{F.2}$$

has among $n \in [\![1; N]\!]$ at most one non-zero term, for $n = \lambda_q$.
Therefore for $\Delta f_{q,r} \in [-\frac{f_p}{2}; \frac{f_p}{2}]$ it is possible to simplify the sum on n leading to:

$$J_{Nyq}[q,r] = \frac{1}{\sigma^2} \sum_{l=1}^{L} \frac{\partial \tilde{x}_q^*[(\lambda_q-1)L+l]}{\partial \theta_q} \frac{\partial \tilde{x}_r[(\lambda_r-1)L+l]}{\partial \theta_r} \tag{F.3}$$

where λ_q and λ_r are equal, or consecutive indices (in case of leakage from the adjacent subband). Hence by recognizing (E.4)(E.5), we have: $J_{Nyq}[q,r] = N.J_{BP}[q,r]$ for $\Delta f_{q,r} \in [-\frac{f_p}{2}; \frac{f_p}{2}]$. Combining both proofs, the theorem is proven:

$$J_{BP}[q,r](\Delta f_{q,r}) = \frac{1}{N}(J_{Nyq}[q,r](\Delta f_{q,r}).H_{LP}(\Delta f_{q,r})) * \sum_{k=-\infty}^{\infty} \delta(\Delta f_{q,r} - k f_p) \tag{F.4}$$

■

Appendix G

Fisher matrix for spectral estimation of multitone signals with Nyquist Sampling

Consider a K-tone signal, generalization of eq. (G.1).

$$x[\check{n}]^{+} = \sum_{k=1}^{K} A_k cos(2\pi f_k \check{n} + \phi_k) + w[\check{n}] \tag{G.1}$$

The parameter vector is $\boldsymbol{\theta} = [A_1...A_K, \Phi_1...\Phi_K, f_1...f_K]$. For Nyquist sampling, $y[\check{n}] = x[\check{n}]$ for $\check{n} \in [\![1; NL]\!]$. Hence the coefficients of the Fisher matrix $\boldsymbol{J_{Nyq}}$ are given by:

$$J_{Nyq}[q,r] = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{NL} \frac{\partial y[\check{n}]}{\partial \theta_q} \frac{\partial y[\check{n}]}{\partial \theta_r} = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{NL} \frac{\partial x[\check{n}]}{\partial \theta_q} \frac{\partial x[\check{n}]}{\partial \theta_r} \tag{G.2}$$

The partial derivatives of the measured signal model with respect to the considered parameter are given by, for $k \in [\![1; K]\!]$:

$$\frac{\partial \tilde{x}[\check{n}]}{\partial A_k} = cos(2\pi f_k \check{n} + \phi_k) \tag{G.3}$$

$$\frac{\partial \tilde{x}[\check{n}]}{\partial \phi_k} = -A_q sin(2\pi f_k \check{n} + \phi_k) \tag{G.4}$$

$$\frac{\partial \tilde{x}[\check{n}]}{\partial f_k} = -A_q 2\pi \check{n} sin(2\pi f_k \check{n} + \phi_k) \tag{G.5}$$

Then coefficients of the Fisher matrix $\boldsymbol{J_{Nyq}}$ for Nyquist sampling are given as follows. For same parameters, $(k, k') \in [\![1; K]\!]^2$:

$$J[A_k, A_{k'}] = \frac{1}{\sigma^2} \sum_{\check{n}=1}^{NL} cos(2\pi f_k \check{n} + \phi_k) cos(2\pi f_{k'} \check{n} + \phi_{k'}) \tag{G.6}$$

Recall that $cos(a)cos(b) = \frac{1}{2}(cos(a+b) + cos(a-b))$ for (a,b) $\in \mathbb{R}$. Hence:

$$J[A_k, A_{k'}] = \frac{1}{2\sigma^2} \sum_{\check{n}=1}^{NL} cos(2\pi(\Delta f)_{k,k'} \check{n} + (\Delta\phi)_{k,k'}) + cos(2\pi(\Sigma f)_{k,k'} \check{n} + (\Sigma\phi)_{k,k'}) \tag{G.7}$$

where the notations $(\Delta f)_{k,k'} = f'_k - f_k$ and $(\Sigma f)_{k,k'} = f'_k + f_k$ are used for sake of brevity. Consider then the real part of a sum of geometrical terms with common ratio $e^{2j\pi f}$ given by ($f \neq 0$):

$$\sum_{\check{n}=1}^{NL} e^{2j\pi f\check{n}+\Delta\phi} = e^{\Delta\phi+2j\pi f}\frac{1-e^{2j\pi NLf}}{1-e^{2j\pi f}} \tag{G.8}$$

Then from half angle factorization (for $f \neq \{0; 0.5\}$):

$$e^{\Delta\phi+2j\pi f}\frac{1-e^{2j\pi NLf}}{1-e^{2j\pi f}} = e^{\Delta\phi+j\pi f(NL+1)}\frac{sin(2\pi NLf)}{sin(2\pi f)} \tag{G.9}$$

Taking the real part in eq. (G.9) to replace in eq. (G.7) yields:

$$\begin{aligned} J[A_k, A_{k'}] = & \tfrac{1}{2\sigma^2}[\tfrac{sin(NL(\Delta f)_{k,k'})}{sin((\Delta f)_{k,k'})}cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) + \\ & \tfrac{sin(NL(\Sigma f)_{k,k'})}{sin((\Sigma f)_{k,k'})}cos(\Sigma\Phi_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))] \\ \simeq & \tfrac{NL}{2\sigma^2}[sinc(NL(\Delta f)_{k,k'})cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) + \\ & sinc(NL(\Sigma f)_{k,k'})cos(\Sigma\Phi_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))] \end{aligned} \tag{G.10}$$

Similarly by using $sin(a)sin(b) = \frac{1}{2}(-cos(a+b) + cos(a-b))$ for (a,b) $\in \mathbb{R}$, the partial derivatives with respect to phase and frequency are given by:

$$\begin{aligned} J[\Phi_k, \Phi_{k'}] \simeq & \tfrac{A_k A_{k'} NL}{2\sigma^2}[sinc(NL(\Delta f)_{k,k'})cos((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) - \\ & sinc(NL(\Sigma f)_{k,k'})cos((\Sigma\Phi)_{k,k'}) + \pi(\Sigma f)_{k,k'}(NL+1))] \end{aligned} \tag{G.11}$$

$$J[f_k, f_{k'}] = \tfrac{2\pi^2}{\sigma^2}A_k A_{k'}\Re[e^{j(\Delta\Phi)_{k,k'}}\chi_2(\Delta f, NL) - e^{j(\Sigma\Phi)_{k,k'}}\chi_2(\Sigma f, NL)]$$

where $\chi_2(f, NL) = \sum_{\check{n}=1}^{NL}\check{n}^2 e^{j2\pi f\check{n}}$. For diagonal terms and $k = k' \in [\![1; K]\!]$, previous equations simplify to:

$$J[A_k, A_k] = \sum_{\check{n}=1}^{NL}(\frac{1}{2} + \frac{cos(4\pi\check{n}f_k + 2\phi_k)}{2}) \simeq \frac{NL}{2\sigma^2} \tag{G.12}$$

$$J[\Phi_k, \Phi_k] = A_k^2\sum_{\check{n}=1}^{NL}(\frac{1}{2} - \frac{cos(4\pi\check{n}f_k + 2\phi_k)}{2}) \simeq \frac{NLA_q^2}{2\sigma^2} \tag{G.13}$$

$$\begin{aligned} J[f_k, f_k] \simeq & \frac{2\pi^2 A_k^2}{\sigma^2}\sum_{\check{n}=1}^{NL}\check{n}^2 \\ = & \frac{2\pi^2 A_k^2}{\sigma^2}\frac{NL(NL+1)(NL+2)}{6} \end{aligned} \tag{G.14}$$

Other terms of the Fisher information matrix are given as follows.

$$J[A_k, \Phi_{k'}] = -A_{k'}\frac{NL}{\sigma^2}\sum_{\check{n}=1}^{NL} sin(2\pi f_k\check{n} + \phi_k)cos(2\pi f_{k'}\check{n} + \phi_{k'}) \tag{G.15}$$

Using $sin(a)cos(b) = \frac{1}{2}(sin(a+b) + sin(a-b))$ for (a,b) $\in \mathbb{R}$ yields:

$$J[A_k, \Phi_{k'}] \simeq \frac{-A_{k'}NL}{2\sigma^2}[sinc(NL(\Delta f)_{k,k'})sin((\Delta\Phi)_{k,k'} + \pi(\Delta f)_{k,k'}(NL+1)) + sinc(NL(\Sigma f)_{k,k'})sin((\Sigma\Phi)_{k,k'} + \pi(\Sigma f)_{k,k'}(NL+1))] \quad \text{(G.16)}$$

$$J[A_k, f_{k'}] = \frac{-2\pi^2}{\sigma^2}A_{k'}\Im[e^{j(\Delta\Phi)_{k,k'}}\chi_2(\Delta f, NL) + e^{j(\Sigma\Phi)_{k,k'}}\chi_2(\Sigma f, NL)] \quad \text{(G.17)}$$

$$J[\Phi_k, f_{k'}] = \frac{2\pi}{\sigma^2}A_kA_{k'}\Re[e^{j(\Delta\Phi)_{k,k'}}\chi_1(\Delta f, NL) - e^{j(\Sigma\Phi)_{k,k'}}\chi_1(\Sigma f, NL)] \quad \text{(G.18)}$$

where $\chi_1(f, NL) = \sum_{\check{n}=1}^{NL} \check{n}e^{j2\pi f\check{n}}$.
For $k = k'$:

$$J[A_k, \Phi_k] \simeq \frac{NL}{2\sigma^2}sinc(2NLf_k)sin(2\Phi_k + 2\pi f_k(NL+1))] \quad \text{(G.19)}$$

$$J[A_k, f_k] = \frac{2\pi^2}{\sigma^2}A_k\Im[e^{j2\Phi_k}\chi_2(2f_k, NL)] \quad \text{(G.20)}$$

$$J[\Phi_k, f_k] = \frac{2\pi A_k^2}{\sigma^2}\frac{NL(NL+1)}{2} \quad \text{(G.21)}$$

Note that eq. (G.12-G.14) and (G.19-G.21) match the reference benchmark of eq. (3.41) established in [53]. It is possible to compute further χ_1 and χ_2, either by simple but tedious calculations consisting in deriving twice $\sum_{n=1}^{N} e^{j2\pi fn}$ or by recognizing the Fourier Transform of a ramp function and a square function.

■

Bibliography

[1] M. Mishali, Y. C. Eldar, and A. J. Elron. "Xampling: Signal Acquisition and Processing in Union of Subspaces." In: *Signal Processing, IEEE Transactions on* 59.10 (2011), pp. 4719–4734 (cit. on pp. 3, 42, 56, 64).

[2] CISCO. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper." In: CISCO, 2017 (cit. on pp. 9, 15).

[3] H. Hassanieh et al. "GHz-wide sensing and decoding using the sparse Fourier transform." In: *IEEE INFOCOM- IEEE Conference on Computer Communications*. 2014, pp. 2256–2264 (cit. on pp. 9, 15).

[4] Executive Office of the US President's Council of Advisors on Science and Technology. *Report to the president: Realizing the full potential of government-held spectrum to spur economic growth*. US Government Document. 2012 (cit. on pp. 9, 15).

[5] E. J. Candès et al. "Compressive Sampling." In: *Proc. of the International Congress of Mathematicians* (2006) (cit. on pp. 9, 15, 21, 23, 65, 73).

[6] R. T. Yazicigil et al. "Wideband Rapid Interferer Detector Exploiting Compressed Sampling With a Quadrature Analog-to-Information Converter." In: *Solid-State Circuits, IEEE Journal of* PP.99 (2015), pp. 1–18 (cit. on pp. 11, 56, 61, 65, 72, 94, 119).

[7] International Symposium on Advanced Radio Technologies ISART. "Regulatory Framework(s) for Facilitating New Spectrum Sharing Schemes Session IV: Validating and Regulating New Sharing Schemes." In: 2012 (cit. on p. 18).

[8] S. Bhattarai et al. "An Overview of Dynamic Spectrum Sharing: Ongoing Initiatives, Challenges, and a Roadmap for Future Research." In: *Cognitive Communications and Networking, IEEE Transactions on* 2.2 (2016), pp. 110–128 (cit. on pp. 18, 19).

[9] C. E. Shannon. "Communication In The Presence Of Noise." In: *Proceedings of the IEEE* 86.2 (1949), pp. 447–457 (cit. on p. 20).

[10] J. Hogan. "Developments of the Heterodyne Receiver." In: *Proceedings of the IRE* 3.3 (1915), pp. 249–260 (cit. on p. 20).

[11] B. F. Logan. "Properties of high pass signal." PhD thesis. 1965 (cit. on p. 20).

[12] H. J. Landau. "Sampling, data transmission, and the Nyquist rate." In: *Proceedings of the IEEE* 55.10 (1967), pp. 1701–1706 (cit. on p. 20).

[13] S. G. Mallat and Z. Zhang. "Matching pursuits with time-frequency dictionaries." In: *Signal Processing, IEEE Transactions on* 41.12 (1993), pp. 3397–3415 (cit. on p. 20).

[14] R. Tibshirani. "Regression shrinkage and selection via the Lasso." In: *Journal of the Royal Statistical Societies B* 58.1 (1996), pp. 267–288 (cit. on p. 20).

[15] E. J. Candès, J. Romberg, and T. Tao. "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information." In: *Information Theory, IEEE Transactions on* 52.2 (2006), pp. 489–509 (cit. on p. 21).

[16] D. L. Donoho. "Compressed sensing." In: *Information Theory, IEEE Transactions on* 52.4 (2006), pp. 1289–1306 (cit. on pp. 21, 73).

[17] G. Wang, Y. Bresler, and V. Ntziachristos. "Guest Editorial Compressive Sensing for Biomedical Imaging." In: *Medical Imaging, IEEE Transactions on* 30.5 (2011), pp. 1013–1016 (cit. on p. 21).

[18] J. A. Tropp and A. C. Gilbert. "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit." In: *Information Theory, IEEE Transactions on* 53.12 (2007), pp. 4655–4666 (cit. on pp. 22, 28).

[19] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing.* Springer, 2013 (cit. on pp. 22, 23, 25).

[20] L. Welch. "Lower bounds on the maximum cross correlation of signals." In: *Information Theory, IEEE Transactions on* 20.3 (1974), pp. 397–399 (cit. on p. 23).

[21] M. Golbabaee and P. Vandergheynst. "Average case analysis of sparse recovery with thresholding: New bounds based on average dictionary coherence." In: *2008 Acoustics Speech and Signal Processing (ICASSP), IEEE International Conference on.* 2008, pp. 3877–3880 (cit. on p. 23).

[22] E. J. Candès and M. Wakin. *"People Hearing Without Listening:" An Introduction To Compressive Sampling.* 2008 (cit. on pp. 23, 47).

[23] E. J. Candès and M. B. Wakin. "An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition." In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 21–30 (cit. on pp. 23, 26, 65).

[24] E. J. Candès. "The restricted isometry property and its implications for compressed sensing." In: *Comptes Rendus Mathematiques* 346.9 (2008), pp. 589–592 (cit. on pp. 23, 78).

[25] Y. C. Eldar and M. Mishali. "Robust Recovery of Signals From a Structured Union of Subspaces." In: *Information Theory, IEEE Transactions on* 55.11 (2009), pp. 5302–5316 (cit. on pp. 24, 76).

[26] R. Baraniuk et al. "An introduction to compressive sensing." In: *Connexions e-textbook* (2011) (cit. on pp. 24, 76).

[27] M. A. Davenport et al. "Introduction in Compressed Sensing. Ed. Yonina C. Eldar and Gitta Kutyniok." In: 1st ed. Cambridge University Press, 2012. Chap. pp. 269-304 (cit. on pp. 24–27).

[28] S. Becker. "Practical Compressed Sensing: modern data acquisition and signal processing." PhD thesis. 2011 (cit. on pp. 24, 29, 42, 46, 53, 56, 65).

[29] R. Berinde et al. "Combining geometry and combinatorics: A unified approach to sparse signal recovery." In: *2008 46th Annual Allerton Conference on Communication, Control, and Computing.* 2008, pp. 798–805 (cit. on p. 24).

[30] A. S. Bandeira et al. "Certifying the Restricted Isometry Property is Hard." In: *Information Theory, IEEE Transactions on* 59.6 (2013), pp. 3448–3450 (cit. on p. 24).

[31] A. M. Tillmann and M. E. Pfetsch. "The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing." In: *Information Theory, IEEE Transactions on* 60.2 (2014), pp. 1248–1259 (cit. on p. 24).

[32] J. Haupt and R. Nowak. *A Generalized Restricted Isometry Property.* Report. 2007 (cit. on p. 24).

[33] R. G. Baraniuk et al. "A Simple Proof of the Restricted Isometry Property for Random Matrices." In: *Constructive Approximation* 28.3 (2008), pp. 253–263 (cit. on pp. 24, 73).

[34] R. Calderbank, S. Howard, and S. Jafarpour. "Construction of a Large Class of Deterministic Sensing Matrices That Satisfy a Statistical Isometry Property." In: *IEEE Journal of Selected Topics in Signal Processing* 4.2 (2010), pp. 358–374 (cit. on pp. 24, 28).

[35] M. Mishali and Y. C. Eldar. "Expected RIP: Conditioning of the Modulated Wideband Converter." In: *Information Theory Workshop. ITW 2009. IEEE.* 2009, pp. 343–347 (cit. on pp. 24, 26, 65, 72, 78, 79).

[36] D. L. Donoho and J. Tanner. "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing." In: *Philosophical Transactions of the Royal Society A* 367.1906 (2009), p. 4273 (cit. on pp. 25, 81).

[37] C. Dossal, G. Peyré, and J. Fadili. "A numerical exploration of compressed sampling recovery." In: *Linear Algebra and its Applications* 432.7 (2010), pp. 1663–1679 (cit. on pp. 25, 76).

[38] R. Gribonval and M. Nielsen. "Sparse representations in unions of bases." In: *Information Theory, IEEE Transactions on* 49.12 (2003), pp. 3320–3325 (cit. on p. 26).

[39] H. Rauhut and U. Ayaz. "Nonuniform sparse recovery with subgaussian matrices." In: *Electronic Transactions on Numerical Analysis.* 41 (2014), pp. 167–178 (cit. on p. 26).

[40] S. Qaisar et al. "Compressive sensing: From theory to applications, a survey." In: *Communications and Networks, Journal of* 15.5 (2013), pp. 443–456 (cit. on p. 26).

[41] R. G. Baraniuk. "Compressive Sensing." In: *IEEE Signal Processing Magazine* 24.4 (2007), pp. 118–121 (cit. on p. 27).

[42] R. Maleh et al. "Analog-to-Information and the Nyquist Folding Receiver." In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 2.3 (2012), pp. 564–578 (cit. on pp. 28, 58, 59).

[43] R. T. Yazicigil et al. "How to Make Analog-to-Information Converters Work in Dynamic Spectrum Environments With Changing Sparsity Conditions." In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* PP.99 (2017), pp. 1–10 (cit. on p. 29).

[44] J. Haupt and R. Nowak. "Adaptive sensing for sparse recovery Compressed Sensing in Compressed Sensing. Ed. Yonina C. Eldar and Gitta Kutyniok." In: 1st ed. Cambridge University Press, 2012. Chap. pp. 269-304 (cit. on p. 29).

[45] M. L. Malloy and R. D. Nowak. "Near-Optimal Adaptive Compressed Sensing." In: *Information Theory, IEEE Transactions on* 60.7 (2014), pp. 4001–4012 (cit. on p. 29).

[46] D. Adams. "A Practical Implementation of the Modulated Wideband Converter Compressive Sensing Receiver Architecture." PhD thesis. 2016 (cit. on pp. 29, 95).

[47] Y. Shuai, D. Kai, and W. Deqiang. "Compressed sensing based UWB receiver scheme with NBI mitigation." In: *Wireless Communications & Signal Processing (WCSP), 2015 International Conference on.* 2015, pp. 1–5 (cit. on p. 29).

[48] O. Abari et al. "Why Analog-to-Information Converters Suffer in High-Bandwidth Sparse Signal Applications." In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 60.9 (2013), pp. 2273–2284 (cit. on pp. 29, 53).

[49] D. E. Bellasi et al. "VLSI Design of a Monolithic Compressive-Sensing Wideband Analog-to-Information Converter." In: *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on* 3.4 (2013), pp. 552–565 (cit. on pp. 29, 48).

[50] H. Milosiu and F. Oehler. "Sub-10 microW CMOS wake-up receiver IP for green SoC designs." In: *SOC Conference (SOCC), IEEE 26th International.* 2013, pp. 88–91 (cit. on p. 29).

[51] K. M. H. Badami et al. "A 90 nm CMOS, 6 microW Power-Proportional Acoustic Sensing Frontend for Voice Activity Detection." In: *Solid-State Circuits, IEEE Journal of* 51.1 (2016), pp. 291–302 (cit. on pp. 30, 32, 33).

[52] S. K. Sharma et al. "Application of Compressive Sensing in Cognitive Radio Communications: A Survey." In: *IEEE Communications Surveys and Tutorials* 18 (2016), pp. 1838–1860 (cit. on p. 30).

[53] S. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* 1993 (cit. on pp. 31, 99, 100, 113, 133).

[54] D. Ramasamy, S. Venkateswaran, and U. Madhow. "Compressive estimation in AWGN: General observations and a case study." In: *Proc. Signals, Systems and Computers (ASILOMAR).* 2012, pp. 953–957 (cit. on pp. 31, 90, 93, 107, 111).

[55] D. Ramasamy, S. Venkateswaran, and U. Madhow. "Compressive Parameter Estimation in AWGN." In: *Signal Processing, IEEE Transactions on* 62.8 (2014), pp. 2012–2027 (cit. on pp. 31, 103).

[56] P. Pakrooh et al. "Analysis of Fisher Information and the Cramer-Rao Bound for Nonlinear Parameter Estimation After Random Compression." In: *Signal Processing, IEEE Transactions on* 63.23 (2015), pp. 6423–6428 (cit. on p. 31).

[57] V. Cambareri, C. Xu, and L. Jacques. "The rare eclipse problem on tiles: Quantised embeddings of disjoint convex sets." In: *Sampling Theory and Applications (SampTA), 2017 International Conference on.* 2017, pp. 241–245 (cit. on p. 32).

[58] M. Verhelst and A. Bahai. "Where Analog Meets Digital: Analog-to-Information Conversion and Beyond." In: *Solid-State Circuits Magazine, IEEE* 7.3 (2015), pp. 67–80 (cit. on p. 32).

[59] R. Rubinstein, A. M. Bruckstein, and M. Elad. "Dictionaries for Sparse Representation Modeling." In: *Proceedings of the IEEE* 98.6 (2010), pp. 1045–1057 (cit. on p. 33).

[60] M. Yang et al. "Fisher Discrimination Dictionary Learning for sparse representation." In: *IEEE International Conference on Computer Vision (ICCV)* (2011), p. 543 (cit. on p. 33).

[61] Z. Jiang, Z. Lin, and L. S. Davis. "Label Consistent K-SVD: Learning a Discriminative Dictionary for Recognition." In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.11 (2013), pp. 2651–2664 (cit. on p. 33).

[62] N. Akhtar, F. Shafait, and A. Mian. "Discriminative Bayesian Dictionary Learning for Classification." In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 38.12 (2016), pp. 2374–2388 (cit. on p. 33).

[63] J. Haupt and R. Nowak. "Compressive Sampling for Signal Detection." In: *Acoustics, Speech and Signal Processing, 2007, IEEE International Conference on.* Vol. 3. 2007, pp. III–1509–III–1512 (cit. on p. 34).

[64] A. Amaravati et al. "A 130nm 165 nJ/frame Compressed-Domain Smashed-Filter-Based Mixed-Signal Classifier for "In-Sensor" Analytics in Smart Cameras." In: *Circuits and Systems II: Express Briefs, IEEE Transactions on* 65.3 (2018), pp. 296–300 (cit. on p. 34).

[65] M. A. Davenport et al. "Signal Processing With Compressive Measurements." In: *IEEE Journal of Selected Topics in Signal Processing* 4.2 (2010), pp. 445–460 (cit. on p. 34).

[66] D. Cohen and Y. C. Eldar. "Sub-Nyquist Cyclostationary Detection for Cognitive Radio." In: *Signal Processing, IEEE Transactions on* 65.11 (2017), pp. 3004–3019 (cit. on pp. 34, 35).

[67] T. Zhi, Y. Tafesse, and B. M. Sadler. "Cyclic Feature Detection With Sub-Nyquist Sampling for Wideband Spectrum Sensing." In: *Selected Topics in Signal Processing, IEEE Journal of* 6.1 (2012), pp. 58–69 (cit. on p. 34).

[68] R. O. Schmidt. "Multiple emitter location and signal parameter estimation." In: *Antennas and Propagation, IEEE Transactions on* 34.3 (1986), pp. 276–280 (cit. on p. 35).

[69] K. Jong Min, L. Ok Kyun, and Y. Jong Chul. "Compressive MUSIC: Revisiting the Link Between Compressive Sensing and Array Signal Processing." In: *Information Theory, IEEE Transactions on* 58.1 (2012), pp. 278–301 (cit. on p. 35).

[70] D. D. Ariananda and G. Leus. "Compressive Wideband Power Spectrum Estimation." In: *Signal Processing, IEEE Transactions on* 60.9 (2012), pp. 4775–4789 (cit. on p. 35).

[71] X. Zhengli et al. "Automatic Modulation Recognition of PSK signals using nonuniform compressive samples based on high order statistics." In: *Communication Problem-Solving (ICCP), IEEE International Conference on.* 2014, pp. 611–614 (cit. on pp. 35, 111).

[72] J. A. Tropp et al. "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals." In: *Information Theory, IEEE Transactions on* 56.1 (2010), pp. 520–544 (cit. on pp. 42, 49, 53).

[73] M. Vetterli, P. Marziliano, and T. Blu. "Sampling signals with finite rate of innovation." In: *Signal Processing, IEEE Transactions on* 50.6 (2002), pp. 1417–1428 (cit. on p. 42).

[74] S. Gishkori and G. Leus. "Compressive Sampling Based Energy Detection of Ultra-Wideband Pulse Position Modulation." In: *Signal Processing, IEEE Transactions on* 61.15 (2013), pp. 3866–3879 (cit. on p. 42).

[75] E. Matusiak and Y. C. Eldar. "Sub-Nyquist sampling of short pulses." In: *2011 Acoustics Speech and Signal Processing (ICASSP), IEEE International Conference on.* 2011, pp. 3944–3947 (cit. on p. 42).

[76] R. G. Baraniuk and M. B. Wakin. "Random Projections of Smooth Manifolds." In: *Foundations of Computational Mathematics* 9.1 (2009), pp. 51–77 (cit. on p. 42).

[77] Y. Chi et al. "Sensitivity to Basis Mismatch in Compressed Sensing." In: *Signal Processing, IEEE Transactions on* 59.5 (2011), pp. 2182–2195 (cit. on p. 42).

[78] X. Yang et al. "Subsampled circulant matrix based analogue compressed sensing." In: *Electronics Letters* 48.13 (2012), pp. 767–768 (cit. on pp. 42, 69, 72, 79).

[79] M. A. Davenport and M. B. Wakin. "Compressive sensing of analog signals using discrete prolate spheroidal sequences." In: *Applied and Computational Harmonic Analysis* 33 (2012), pp. 438–472 (cit. on p. 42).

[80] X. Yang. "Research on compressed sensing and its applications in wireless communications." PhD thesis. 2015 (cit. on pp. 42, 69, 72).

[81] M. Mishali and Y. C. Eldar. "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals." In: *IEEE Journal of Selected Topics in Signal Processing* 4 (2010), pp. 375–391 (cit. on pp. 44, 49, 54, 65, 66, 72).

[82] E. Arias-Castro and Y. C. Eldar. "Noise Folding in Compressed Sensing." In: *IEEE Signal Processing Letters* 18.8 (2011), pp. 478–481 (cit. on p. 46).

[83] M. A. Davenport et al. "The Pros and Cons of Compressive Sensing for Wideband Signal Acquisition: Noise Folding versus Dynamic Range." In: *Signal Processing, IEEE Transactions on* 60.9 (2012), pp. 4628–4642 (cit. on pp. 46, 84).

[84] E. Israeli et al. "Hardware Calibration of the Modulated Wideband Converter." In: *IEEE Global Communications Conference.* 2014, pp. 948–953 (cit. on pp. 46, 56).

[85] P. T. Boufounos and H. Mansour. "Universal embeddings for kernel machine classification." In: *Sampling Theory and Applications (SampTA), 2015 International Conference on.* 2015, pp. 307–311 (cit. on p. 46).

[86] A. Zebadua et al. "Compressed and quantized correlation estimators." In: *Signal Processing, IEEE Transactions on* PP.99 (2016), pp. 1–1 (cit. on p. 47).

[87] J. N. Laska et al. "Democracy in action: Quantization, saturation, and compressive sensing." In: *Applied and Computational Harmonic Analysis* 31.3 (2011), pp. 429–443 (cit. on p. 47).

[88] C. S. Güntürk et al. "Sobolev Duals for Random Frames and $\Sigma\Delta$ Quantization of Compressed Sensing Measurements." In: *Foundations of Computational Mathematics* 13.1 (2013), pp. 1–36 (cit. on p. 47).

[89] A. Zymnis, S. Boyd, and E. Candes. "Compressed Sensing With Quantized Measurements." In: *Signal Processing Letters, IEEE* 17.2 (2010), pp. 149–152 (cit. on p. 47).

[90] L. Jacques et al. "Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors." In: *Information Theory, IEEE Transactions on* 59.4 (2013), pp. 2082–2102 (cit. on p. 47).

[91] D. King, R. Packard, and R. Thomas. "Unequally-spaced, broad-band antenna arrays." In: *IRE Transactions on Antennas and Propagation* 8.4 (1960), pp. 380–384 (cit. on p. 47).

[92] J. Laska et al. "Random Sampling for Analog-to-Information Conversion of Wideband Signals." In: *IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*. 2006, pp. 119–122 (cit. on p. 47).

[93] S. Traoré. "Contribution à l'étude de l'échantillonnage non uniforme dans le domaine de la radio intelligente." PhD thesis. 2015 (cit. on pp. 47, 124).

[94] M. Mishali and Y. C. Eldar. "Blind Multiband Signal Reconstruction: Compressed Sensing for Analog Signals." In: *Signal Processing, IEEE Transactions on* 57.3 (2009), pp. 993–1009 (cit. on p. 47).

[95] A. A. Lazar and L. T. Toth. "Perfect recovery and sensitivity analysis of time encoded bandlimited signals." In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 51.10 (2004), pp. 2060–2073 (cit. on p. 48).

[96] M. Wakin et al. "A Nonuniform Sampler for Wideband Spectrally-Sparse Environments." In: *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on* 2.3 (2012), pp. 516–529 (cit. on p. 48).

[97] L. Bai and S. Roy. "Compressive Spectrum Sensing Using a Bandpass Sampling Architecture." In: *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on* 2.3 (2012), pp. 433–442 (cit. on pp. 49, 57).

[98] M. Pelissier and C. Studer. "Non-Uniform Wavelet Sampling for RF Analog-to-Information Conversion." In: *Circuits and Systems–I: Regular Papers, IEEE Transactions on* (2017) (cit. on pp. 49–51, 88, 123).

[99] M. Pelissier et al. "Hardware platform of Analog-to-Information converter using Non Uniform Wavelet Bandpass Sampling for RF signal activity detection." In: *IEEE International Symposium on Circuits and Systems (ISCAS)*. 2018, pp. 1–5 (cit. on pp. 52, 88, 121, 124).

[100] J. N. Laska et al. "Theory and Implementation of an Analog-to-Information Converter using Random Demodulation." In: *Proc. IEEE International Symposium on Circuits and Systems*. 2007, pp. 1959–1962 (cit. on pp. 52, 54).

[101] L. Ros, G. Jourdain, and M. Arndt. "Interpretations and performances of linear reception in downlink TD-CDMA and multi-sensor extension." In: *Annals of Telecommunications* 56.5 (2001), pp. 275–290 (cit. on p. 53).

[102] L. Ros. "Multi-sensor reception for radio-mobile terminal in a Code Division Multiple Access system. Application to the TDD mode of UMTS." PhD thesis. 2001 (cit. on p. 53).

[103] Y. Juhwan. "Compressed sensing receivers: theory, design, and performance limits." PhD Thesis. 2012 (cit. on p. 53).

[104] C. Yilun et al. "Modulated Wideband Converter with non-ideal lowpass filters." In: *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on.* 2010, pp. 3630–3633 (cit. on p. 56).

[105] T. Haque et al. "Theory and Design of a Quadrature Analog-to-Information Converter for Energy-Efficient Wideband Spectrum Sensing." In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 62.2 (2015), pp. 527–535 (cit. on pp. 56, 64).

[106] J. Romberg. "Compressive sensing by random convolution." In: *SIAM Journal on Imaging Sciences* 2.4 (2009), pp. 1098–1128 (cit. on pp. 56, 67, 68, 70, 71).

[107] J. A. Tropp. "Random Filters for Compressive Sampling." In: *Information Sciences and Systems, 40th Annual Conference on.* 2006, pp. 216–217 (cit. on p. 56).

[108] R. G. Vaughan, N. L. Scott, and D. R. White. "The Theory of Bandpass Sampling." In: *Signal Processing, IEEE Transactions on* 39.9 (1991), pp. 1973–1984 (cit. on p. 57).

[109] M. Fleyer et al. "Multirate Synchronous Sampling of Sparse Multiband Signals." In: *Signal Processing, IEEE Transactions on* 58.3 (2010), pp. 1144–1156 (cit. on p. 57).

[110] M. E. Dominguez-Jimenez and N. Gonzalez-Prelcic. "Analysis and design of multirate synchronous sampling schemes for sparse multiband signals." In: *European Signal Processing Conference (EUSIPCO), Proc. of the 20th.* 2012, pp. 1184–1188 (cit. on p. 57).

[111] S. Hongjian et al. "Wideband Spectrum Sensing With Sub-Nyquist Sampling in Cognitive Radios." In: *Signal Processing, IEEE Transactions on* 60.11 (2012), pp. 6068–6073 (cit. on p. 57).

[112] S. Stein et al. "CaSCADE: Compressed Carrier and DOA Estimation." In: *Signal Processing, IEEE Transactions on* PP.99 (2017), pp. 1–1 (cit. on p. 64).

[113] M. Marnat et al. "Code Properties Analysis for the Implementation of a Modulated Wideband Converter." In: *European Signal Processing Conference (EUSIPCO), Proceedings of the 25th.* 2017 (cit. on pp. 64, 76–78, 85).

[114] M. Marnat et al. "Analyse comparative des codes de mélange du convertisseur comprimé MWC." In: *GRETSI.* 2017 (cit. on p. 64).

[115] O. Taheri and S. A. Vorobyov. "Segmented Compressed Sampling for Analog-to-Information Conversion: Method and Performance Analysis." In: *Signal Processing, IEEE Transactions on* 59.2 (2011), pp. 554–572 (cit. on p. 65).

[116] M. Mishali et al. "Xampling: Analog to digital at sub-Nyquist rates." In: *Circuits, Devices & Systems, IET* 5.1 (2010), pp. 8–20 (cit. on pp. 65, 79).

[117] L. Gan and H. Wang. "Deterministic Binary Sequences for Modulated Wideband Converter." In: *Sampling Theory and Applications (SampTA), 2013 International Conference on.* 2013 (cit. on pp. 65, 67, 70, 72, 79).

[118] R. Gold. "Optimal binary sequences for spread spectrum multiplexing (Corresp.)" In: *Information Theory, IEEE Transactions on* 13.4 (1967), pp. 619–621 (cit. on p. 65).

[119] M. Mishali and Y. C. Eldar. "Conditioning of the Modulated Wideband Converter." In: *CCIT Technion Report* (2009) (cit. on pp. 65, 79).

[120] M. F. Duarte and Y. C. Eldar. "Structured Compressed Sensing: From Theory to Applications." In: *Signal Processing, IEEE Transactions on* 59.9 (2011), pp. 4053–4085 (cit. on pp. 65, 67).

[121] W. Guicquero. "Exploring Information Retrieval using Image Sparse Representations: From Circuits Designs and Acquisition Processes to specific Reconstruction Algorithms." PhD Thesis. 2014 (cit. on p. 66).

[122] A. Saade et al. "Random projections through multiple optical scattering: Approximating Kernels at the speed of light." In: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016, pp. 6215–6219 (cit. on p. 66).

[123] W. Yin et al. "Practical Compressive Sensing with Toeplitz and Circulant Matrices." In: *2010 Visual Communications and Image Processing*. Vol. 7744. 2010, 77440K–77440K–10 (cit. on pp. 66, 67, 73).

[124] K. Li, G. Lu, and L. Cong. "Convolutional Compressed Sensing Using Deterministic Sequences." In: *Signal Processing, IEEE Transactions on* 61.3 (2013), pp. 740–752 (cit. on pp. 67–69).

[125] F. Krahmer, S. Mendelson, and H. Rauhut. "Suprema of chaos processes and the restricted isometry property." In: *Communications on Pure and Applied Mathematics* 67.11 (2014), pp. 1877–1904 (cit. on p. 68).

[126] H. Rauhut, J. Romberg, and J. A. Tropp. "Restricted isometries for partial random circulant matrices." In: *Applied and Computational Harmonic Analysis* 32.2 (2012), pp. 242–254 (cit. on p. 68).

[127] J. P. Elsner et al. "Compressed spectrum estimation for cognitive radios." In: *Proc. 19th Virginia Tech. Symp. Wireless Commun.* 2009 (cit. on p. 68).

[128] E. M. Gabidulin and V. V. Shorin. "Unimodular perfect sequences of length p^s." In: *Information Theory, IEEE Transactions on* 51.3 (2005), pp. 1163–1166 (cit. on p. 70).

[129] A. K. Brodzik. "On the Fourier Transform of Finite Chirps." In: *IEEE Signal Processing Letters* 13.9 (2006), pp. 541–544 (cit. on p. 70).

[130] D. Chu. "Polyphase codes with good periodic correlation properties (Corresp.)" In: *Information Theory, IEEE Transactions on* 18.4 (1972), pp. 531–532 (cit. on p. 70).

[131] B. M. Popovic. "Generalized chirp-like polyphase sequences with optimum correlation properties." In: *Information Theory, IEEE Transactions on* 38.4 (1992), pp. 1406–1409 (cit. on p. 70).

[132] Semtec. *Lora Modulation Theory*. Report. 2014 (cit. on p. 70).

[133] Y. Roth et al. "Contender waveforms for Low-Power Wide-Area networks in a scheduled 4G OFDM framework." In: *EURASIP Journal on Advances in Signal Processing* 2018.1 (2018), p. 43 (cit. on p. 70).

[134] K. Jae Won et al. "Generalized Cross-Correlation Properties of Chu Sequences." In: *Information Theory, IEEE Transactions on* 58.1 (2012), pp. 438–444 (cit. on p. 70).

[135] B. M. Popovic. "Efficient DFT of Zadoff-Chu sequences." In: *Electronics Letters* 46.7 (2010), pp. 502–503 (cit. on p. 70).

[136] C. P. Li and W. C. Huang. "A Constructive Representation for the Fourier Dual of the Zadoff-Chu Sequences." In: *Information Theory, IEEE Transactions on* 53.11 (2007), pp. 4221–4224 (cit. on p. 70).

[137] M. Rudelson and R. Vershynin. "On sparse reconstruction from Fourier and Gaussian measurements." In: *Communications on Pure and Applied Mathematics* 61.8 (2008), pp. 1025–1045 (cit. on p. 73).

[138] D. Romero et al. "Compressive Covariance Sensing: Structure-based compressive sensing beyond sparsity." In: *Signal Processing Magazine, IEEE* 33.1 (2016), pp. 78–93 (cit. on p. 73).

[139] B. Wichmann. "A note on restricted difference bases." In: *Journal of the London Mathematical Society* s1-38.1 (1963), pp. 465–466 (cit. on p. 73).

[140] M. Mishali and Y. C. Eldar. "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals." In: *Selected Topics in Signal Processing, IEEE Journal of* 4.2 (2010), pp. 375–391 (cit. on pp. 79, 80).

[141] D. L. Donoho, A. Maleki, and A. Montanari. "Message-passing algorithms for compressed sensing." In: *Proc. of the National Academy of Sciences* 106.45 (2009), pp. 18914–18919 (cit. on p. 81).

[142] M. Marnat and M. Pelissier. "Méthode d'échantillonnage passe-bande par ondelettes modulées en position." In: *Patent EN 1900181* (cit. on pp. 88, 125).

[143] M. Marnat et al. "Cramér-Rao bounds for spectral parameter estimation with compressive multi-band architectures." In: *submitted in IEEE Transactions on Signal Processing.* under revision (cit. on p. 91).

[144] J. Shihao, X. Ya, and L. Carin. "Bayesian Compressive Sensing." In: *Signal Processing, IEEE Transactions on* 56.6 (2008), pp. 2346–2356 (cit. on p. 105).

[145] M. Mangia, R. Rovatti, and G. Setti. "Rakeness in the Design of Analog-to-Information Conversion of Sparse and Localized Signals." In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 59.5 (2012), pp. 1001–1014 (cit. on p. 105).

[146] D. Adams, Y. Eldar, and B. Murmann. "A mixer frontend for a four-channel Modulated Wideband Converter with 62 dB blocker rejection." In: *IEEE Radio Frequency Integrated Circuits Symposium (RFIC).* 2016, pp. 286–289 (cit. on p. 108).

[147] H. L. Van Trees and K. L. Bell. "Barankin Bounds on Parameter Estimation." In: *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking.* Wiley-IEEE Press, 2007, p. 951 (cit. on pp. 108, 114).

[148] S. Bellini and G. Tartara. "Bounds on Error in Signal Parameter Estimation." In: *Communications, IEEE Transactions on* 22.3 (1974), pp. 340–342 (cit. on pp. 108, 114).

[149] S. Mallat. "Understanding Deep Convolutional Networks." In: *Philosophical Transactions of the Royal Society A* (2018) (cit. on p. 111).

[150] J. Andén and S. Mallat. "Deep Scattering Spectrum." In: *Signal Processing, IEEE Transactions on* 62.16 (2014), pp. 4114–4128 (cit. on p. 111).

[151] R. Bammer and M. Dörfler. "Invariance and stability of Gabor scattering for music signals." In: *Sampling Theory and Applications (SampTA), 2017 International Conference on.* 2017, pp. 299–302 (cit. on p. 111).

[152] R. Calderbank. "Fast Reconstruction Algorithms for Deterministic Sensing Matrices and Applications." In: *Program in Applied and Computational Mathematics Princeton University.* 2009 (cit. on p. 124).

[153] L. Applebaum et al. "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery." In: *Applied and Computational Harmonic Analysis* 26.2 (2009), pp. 283–290 (cit. on p. 124).

List of publications

[99] M. Pelissier et al. "Hardware platform of Analog-to-Information converter using Non Uniform Wavelet Bandpass Sampling for RF signal activity detection." In: *IEEE International Symposium on Circuits and Systems (ISCAS)*. 2018, pp. 1–5 (cit. on pp. 52, 88, 121, 124).

[113] M. Marnat et al. "Code Properties Analysis for the Implementation of a Modulated Wideband Converter." In: *European Signal Processing Conference (EUSIPCO), Proceedings of the 25th*. 2017 (cit. on pp. 64, 76–78, 85).

[114] M. Marnat et al. "Analyse comparative des codes de mélange du convertisseur comprimé MWC." In: *GRETSI*. 2017 (cit. on p. 64).

[142] M. Marnat and M. Pelissier. "Méthode d'échantillonnage passe-bande par ondelettes modulées en position." In: *Patent EN 1900181* (cit. on pp. 88, 125).

[143] M. Marnat et al. "Cramér-Rao bounds for spectral parameter estimation with compressive multi-band architectures." In: *submitted in IEEE Transactions on Signal Processing*. under revision (cit. on p. 91).

Résumé – Cette thèse traite de la conception de récepteurs radiofréquences basés sur l'acquisition compressée pour de l'estimation paramétrique en radio cognitive. L'acquisition compressée est un changement du paradigme de la conversion analogique-numérique qui permet de s'affranchir de la fréquence d'échantillonnage de Nyquist sous hypothèse d'occupation parcimonieuse du spectre. Dans ces travaux, les estimations sont effectuées sur les échantillons compressés vu le coût prohibitif de la reconstruction du signal d'entrée. Après avoir dressé un état de l'art de l'acquisition compressée pour la radio cognitive et passé en revue différentes architectures de récepteurs, la première contribution concerne l'étude des codes de mélange pour une architecture particulière, le convertisseur modulé à large bande (MWC). Une analyse haut niveau des propriétés de la matrice d'acquisition, à savoir la cohérence pour réduire le nombre de mesures et l'isométrie pour la robustesse au bruit, est menée puis validée par une plateforme de simulation. Ensuite, l'estimation paramétrique à partir des échantillons compressés est abordée à travers la borne de Cramér-Rao sur la variance d'un estimateur non biaisé. Une forme analytique de la borne est établie sous certaines hypothèses et permet de dissocier les effets de la compression et de la création de diversité. L'influence du processus d'acquisition, notamment le couplage entre paramètres et la fuite spectrale, est illustrée par l'exemple.
Mots clés : Acquisition compressée, Acquisition comprimée, Détection de Spectre, Estimation Paramétrique, Radio Cognitive, MWC, NUWBS, Internet des Objets, Récepteurs radio

Abstract – This work deals with the topic of radiofrequency receivers based on Compressive Sampling for feature extraction in Cognitive Radio. Compressive Sampling is a paradigm shift in analog to digital conversion that bypasses the Nyquist sampling frequency under assumption of spectral sparsity of the signal. In this work, estimations are carried out on the compressed samples due to the prohibitive cost of signal reconstruction. After a state-of-the-art on Compressive Sampling in Cognitive Radio and a discussion on different compressive receiver architectures, our first contribution is a study of the mixing codes of a particular receiver, the Modulated Wideband Converter. A high-level analysis on properties of the sensing matrix, coherence to reduce the number of measurement and isometry for noise robustness, is carried out and validated by a simulation platform. Then, parametric estimation based on compressed samples is tackled through the prism of the Cramér-Rao lower bound on unbiased estimators. A closed form expression of the bound is established under certain assumptions and enables to dissociate the effects of compression and diversity creation. The influence of Compressive Sampling on estimation bounds, in particular coupling between parameters and spectral leakage, is illustrated by the example.
Keywords: Compressive Sensing, Compressive Sampling, Spectrum Sensing, Parametric Estimation, Cognitive Radio, Modulated Wideband Converter, Non Uniform Wavelet Bandpass Sampling, IoT, Radiofrequency receiver

CEA Grenoble, DRT/LETI/DACLE/SCCI/LAIR
17 rue des martyrs
38054 Grenoble cedex 9, France