



HAL
open science

Traitement parcimonieux de signaux biologiques

Paul Chollet

► **To cite this version:**

Paul Chollet. Traitement parcimonieux de signaux biologiques. Electronique. Ecole nationale supérieure Mines-Télécom Atlantique, 2017. Français. NNT : 2017IMTA0024 . tel-01797809

HAL Id: tel-01797809

<https://theses.hal.science/tel-01797809>

Submitted on 22 May 2018

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

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



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

**UNIVERSITE
BRETAGNE
LOIRE**

THÈSE / IMT Atlantique

sous le sceau de l'Université Bretagne Loire

pour obtenir le grade de

DOCTEUR D'IMT Atlantique

Spécialité : Electronique

École Doctorale Mathématiques et STIC

Présentée par

Paul Chollet

Préparée dans le département Electronique

Laboratoire Labsticc

Traitement parcimonieux de signaux biologiques

Thèse soutenue le 24 novembre 2017

devant le jury composé de :

Mustapha Nadi

Professeur des universités, Institut Jean Lamour - Vandoeuvre les Nancy / président

Noëlle Lewis

Professeur des universités, IMS - Bordeaux / rapporteur

Hervé Barthelemy

Professeur des universités, Université Sud Toulon Var / rapporteur

Patricia Desgreys

Professeur, Télécom ParisTech / examinateur

Fabrice Seguin

Maître de conférences, IMT Atlantique / examinateur

Matthieu Arzel

Maître de conférences, IMT Atlantique / examinateur

Cyril Lahuec

Maître de conférences (HDR), IMT Atlantique / directeur de thèse

Remerciements

Ce manuscrit est le résultat des travaux de recherche d'une durée de trois ans. Je tiens donc à remercier toutes les personnes sans qui ces travaux n'existeraient pas. En premier lieu, je souhaite remercier l'ensemble des membres de mon jury de thèse : Mesdames Noëlle Lewis et Patricia Desgreys ainsi que Messieurs Hervé Barthélemy, Mustapha Nadi, Cyril Lahuec, Fabrice Seguin et Matthieu Arzel. J'ai pris un réel plaisir à échanger avec vous lors de la soutenance et durant la période, trop courte, l'ayant suivie.

J'aimerais aussi présenter mes remerciements à l'ensemble des membres de PRACOM qui m'ont permis de faire ces travaux en finançant cette thèse. De plus, je remercie l'ensemble du personnel du département ELEC de l'IMT Atlantique (Télécom Bretagne pour les plus vieux) pour ces trois années très enrichissantes. Vous m'avez permis de travailler dans une ambiance des plus stimulantes et je vous souhaite plein de réussite dans vos futurs projets.

Je tiens aussi à remercier ma famille qui m'a soutenu dans les moments durs comme bons. Merci à vous d'avoir été présent durant ces moments compliqués de 2015. Un merci tout particulier à mes parents qui ont bravé le destin à plusieurs reprises afin de me rendre visite en Bretagne.

Ensuite, je souhaite dire un grand merci à tous mes collègues thésards et à la Telecom's family pour votre soutien et votre amitié (enfin j'espère). Un merci tout particulier à Franck, Paul le mou, Pierre, Valentin ainsi qu'à Erwan, Benoît et André malgré vos mensonges incessants.

Enfin, je tiens à remercier ma formidable équipe encadrante composée de Fabrice Seguin et Matthieu Arzel comme encadrants et de Cyril Lahuec comme directeur. Mille mercis pour votre support et vos précieux conseils. Si j'étais le bon candidat pour cette thèse alors je pense sincèrement que vous étiez la parfaite équipe pour ce candidat. J'ai énormément apprécié travailler avec vous et espère avoir l'occasion de le refaire dans le futur.

Résumé

Les réseaux de capteurs corporels représentent un enjeu sociétal important par la promesse d'un suivi des soins de meilleure qualité avec un coût réduit. Ils permettent en effet une surveillance des maladies chroniques depuis le domicile grâce à l'utilisation de capteurs acquérant des paramètres physiologiques. Ces derniers sont utilisés pour détecter des anomalies dès leur apparition et ainsi intervenir au plus vite. Les capteurs sont soumis à de nombreuses contraintes de fiabilité, robustesse, taille et consommation. La contrainte énergétique est d'autant plus importante que certains capteurs sont implantés sous la peau. Il est donc complexe d'en changer la source d'énergie.

Dans ce document, les différentes opérations réalisées par les réseaux de capteurs corporels sont analysées. La consommation de chacune d'elles est évaluée afin de guider les axes de recherche pour améliorer l'autonomie énergétique des capteurs. L'analyse permet de constater qu'il est préférable d'intégrer du traitement basse consommation sur les capteurs. Cela permet une utilisation parcimonieuse des organes de transmission qui présentent la moins bonne efficacité énergétique.

Un capteur pour la détection d'arythmie sur des signaux cardiaques est ensuite proposé. Il intègre un traitement du signal via l'utilisation d'un réseau de neurone à cliques. Le système proposé est simulé et offre une exactitude de classification de 95 % pour la détection de trois types d'arythmie. Le prototypage du système via la fabrication d'un circuit mixte analogique/numérique en *CMOS* 65 nm montre une consommation du capteur de l'ordre de $1,4 \mu\text{J}$. C'est 105 fois moins que l'énergie d'un capteur ne faisant que transmettre les données.

Pour réduire encore plus l'énergie, une nouvelle méthode d'acquisition est utilisée. Elle consiste à mettre en place un convertisseur analogique vers paramètre simplifiant l'architecture du capteur. Une architecture de convertisseur est proposée pour l'acquisition et le traitement de signaux cardiaques. Les simulations et les estimations laissent espérer une consommation de l'ordre de 1,18 nJ pour acquérir les paramètres tout en offrant une exactitude de classification proche de 98 %. L'ajout de cette acquisition au capteur déjà proposé permettrait de réduire la consommation d'un facteur 2 592. Cette étude permet d'ouvrir la voie vers la mise en place de capteurs très basse consommation pouvant durer toute une vie avec une simple pile.

Mots-clés : réseaux de capteurs corporels, intégration mixte analogique/numérique, réseaux de neurones à cliques, acquisition parcimonieuse.

Abstract

Body area sensor networks gained great focused through the promise of better quality and cheaper medical care system. They allow for a continuous monitoring of chronic diseases from home using sensors that acquire physiological parameters. These parameters are used to detect anomalies and treat them as soon as they arise. Sensors are under heavy constraints such as reliability, sturdiness, size and power consumption. The energy constraint is amplified by the fact that sensors can be worn under the skin. Changing the power source is thus a big challenge.

This document analyzes the operations perform by a body area sensor network. The different energy requirements are evaluated in order to choose the focus of the research to improve the battery life of the sensors. The analysis shows that integrating processing into the sensor seems the best solution. This allows for a sparse use of transmission elements that use most of the energy.

A sensor for arrhythmia detection is then proposed. It includes some signal processing through a clique-based neural network. The system simulation allows a classification between three types of arrhythmia with 95 % accuracy. The prototype, based on a 65 nm CMOS mixed signal circuit, requires only 1.4 μJ . It is 105 times less than a sensor transmitting the data. To further reduce energy consumption, a new sensing method is used. The idea is to design an analog to feature converter to simplify the sensor architecture. An analog to feature converter architecture is proposed for heart beat acquisition. Simulation and estimation show a 1.18 nJ energy requirement for parameter acquisition while offering 98 % classification accuracy. Adding the acquisition to the already proposed system yields an energy decrease by 2,592. This work leads the way to the development of low energy sensor with a lifetime battery life.

Keywords : body area sensor networks, mixed signal integration, clique-based neural networks, sparse sensing.

Table des matières

Remerciements	i
Résumé	iii
Abstract	iv
Table des Figures	ix
Liste des Tableaux	xiii
Liste des acronymes	xv
Publications associées	xvii
Introduction	1
1 Les réseaux de capteurs corporels	5
Introduction	5
1.1 Présentation des réseaux de capteurs corporels	5
1.1.1 Les réseaux de capteurs corporels	5
1.1.2 Les défis des réseaux de capteurs corporels	8
1.2 Acquisition des données : du signal analogique vers une représentation numérique .	9
1.2.1 Acquisition classique de Nyquist	9
1.2.1.1 Principe	9
1.2.1.2 Les architectures de CAN pour les signaux biologiques	12
1.2.1.2.1 Registre à approximations successives	12
1.2.1.2.2 Delta-Sigma	14
1.2.2 Acquisition compressée : échantillonnage compressif	17
1.2.2.1 Notion de parcimonie des signaux	17
1.2.2.2 Principe de l'échantillonnage compressif	18
1.2.2.3 Architectures matérielles pour les signaux biologiques	20
1.2.2.4 Algorithmes de reconstruction	22

1.3	Transmission de l'information	24
1.3.1	Caractéristiques de la transmission pour les réseaux de capteurs corporels	24
1.3.2	Les standards existants	25
1.3.3	Implantations des transmetteurs	27
1.4	Traitement de l'information : la classification supervisée	28
1.4.1	Principes	28
1.4.2	Les algorithmes de classification	31
1.4.2.1	Classifieur probabiliste	31
1.4.2.2	Arbres de décision	33
1.4.2.3	Machines à vecteurs de support	35
1.4.2.4	Réseaux de neurones	37
1.5	Discussion	39
	Conclusion	40
2	Système d'analyse d'ECG basse consommation pour l'aide au diagnostic : utilisation des réseaux à cliques	43
	Introduction	43
2.1	Les systèmes d'analyse d'ECG	44
2.1.1	Le signal ECG	44
2.1.2	Le traitement des signaux ECG pour l'arythmie	46
2.1.2.1	Le pré-traitement	47
2.1.2.2	La segmentation	48
2.1.2.3	L'extraction de paramètres	50
2.1.2.4	L'étage de classification	52
2.1.2.5	Considérations matérielles	52
2.2	Proposition d'un système de classification d'arythmie basse consommation	53
2.2.1	Présentation du système	53
2.2.1.1	Objectif du système proposé	53
2.2.1.2	Architecture du système proposé	54
2.2.2	Les réseaux à cliques	55
2.2.2.1	Structure des réseaux à cliques	55
2.2.2.2	Notion de stockage	57
2.2.2.3	Décodage	58
2.2.2.3.1	Récupération d'information	58
2.2.2.3.2	Les règles de décodage	59
2.2.3	Réseau à cliques pour la classification	61
2.2.3.1	Méthode de création des neurones	61
2.2.3.2	Résultats de classification	65
2.3	Discussion	66

Conclusion	67
3 Prototypage du système d'analyse d'ECG basse consommation	69
Introduction	69
3.1 Architecture du pré-traitement	69
3.1.1 Décomposition en ondelettes	70
3.1.1.1 Filtrage	70
3.1.1.2 Opérateur de décomposition en ondelettes	71
3.1.2 Extraction des paramètres	73
3.1.2.1 Opérateur de minimum/maximum	73
3.1.2.2 Opérateur de variance	74
3.1.3 Mise en forme	75
3.2 Conception du classifieur	76
3.2.1 Architecture mixte analogique/numérique du classifieur	76
3.2.2 Mise en œuvre du classifieur avec l'architecture proposée	79
3.2.3 Conception analogique du réseau	80
3.2.3.1 Circuit de neurone existant	80
3.2.3.2 Modification du circuit du neurone	82
3.2.3.3 Conception du réseau	84
3.2.3.4 Simulations du réseau	88
3.3 Intégration et test du réseau	90
3.3.1 Intégration du réseau	90
3.3.2 Présentation du banc de test	92
3.3.3 Test et mesures sur le prototype	94
3.3.3.1 Tests unitaires des neurones	94
3.3.3.2 Performances de classification	97
3.3.3.3 Réduction de l'énergie du classifieur.	99
Conclusion	100
4 Un nouveau système d'acquisition parcimonieux des ECG	103
Introduction	103
4.1 Le concept de convertisseur analogique vers paramètres	104
4.1.1 Principe	104
4.1.2 Architecture générale	105
4.2 Acquisition parcimonieuse de signaux ECG : conversion analogique vers paramètres	107
4.2.1 Objectifs	107
4.2.2 Échantillonnage non uniforme par ondelettes	108
4.2.3 Optimisation conjointe de la sélection des paramètres et du classifieur	112
4.2.3.1 Sélection des paramètres indépendamment de l'architecture	112
4.2.3.2 Choix de l'ondelette et validation de l'architecture	114

4.2.3.3	Algorithme de sélection adapté à l'architecture	116
4.3	Discussion sur l'implantation matérielle	117
4.3.1	Discussion sur la consommation	117
4.3.2	Synchronisation des données	119
4.3.3	Quantification des données	121
	Conclusion	122
	Conclusion et perspectives	125
	Bibliographie	127

Table des figures

1.1	Structure globale d'un <i>BASN</i> : les capteurs transmettent des données à l'agrégateur qui les traite et prévient les secours en cas de nécessité.	6
1.2	Répartition des opérations entre capteur et agrégateur pour les <i>BASN</i> détectant des anomalies. a) capteur transmetteur, b) extraction sur capteur, c) traitement sur capteur, d) compression des données.	7
1.3	Conversion d'un signal analogique en signal numérique. Le signal est d'abord échantillonné avant d'être quantifié.	9
1.4	Spectre fréquentiel d'un signal avant et après échantillonnage.	10
1.5	Un effet de repliement de spectre apparaît lorsque la fréquence d'échantillonnage est trop basse.	10
1.6	Fonction de transfert d'un CAN 3 bits. La sortie change de code lorsque l'entrée V_{en} atteint les seuils V_i . La droite représente la fonction de transfert idéale.	11
1.7	Architecture globale d'un convertisseur à approximations successives.	12
1.8	Exemple de conversion d'une entrée fixe par un convertisseur SAR. La ligne rouge représente le signal d'entrée et la noire la sortie du CNA.	13
1.9	Architecture du CAN $\Delta\Sigma$	14
1.10	Représentation fréquentielle de la sortie d'un CAN à la fréquence de Nyquist (bleu) et via sur-échantillonnage (rouge).	14
1.11	Modèle fréquentiel du CAN $\Delta\Sigma$. Le CNA 1 bit est modélisé par l'ajout d'un bruit de quantification b_q	15
1.12	Spectre fréquentiel en sortie du modulateur $\Delta\Sigma$	16
1.13	Exemple d'un signal 5-parcimonieux en fréquence.	17
1.14	Comparaison temporelle du signal d'origine et sa meilleure représentation 3-parcimonieuse.	18
1.15	Architectures pour la réalisation de convertisseur <i>A2I</i>	21
1.16	Décomposition des systèmes de communication selon le modèle OSI en sept couches.	25
1.17	Topologie de réseau en étoile. Les seules connections possibles sont entre le maître et les esclaves.	26

1.18	Schéma de principe de la classification supervisée. L'apprentissage (a) permet de fixer les paramètres du modèle grâce aux observations d'entraînement. La phase de prédiction (b) utilise le modèle pour déterminer la classe ou le label.	29
1.19	Matrice de confusion pour une classification binaire. La sensibilité et spécificité permettent d'évaluer les performances du modèle.	30
1.20	Exemple de fonction logistique permettant de transformer un nombre réel en probabilité.	32
1.21	Schéma d'un arbre de profondeur trois avec trois classes. Chaque nœud réalise deux comparaisons et en fonction du résultat pointe vers un autre nœud ou vers une feuille.	34
1.22	Illustration de l'hyperplan séparateur pour des données linéairement séparables en deux dimensions.	36
1.23	Modèle d'un perceptron à une couche cachée, n entrées, m neurones cachés et c neurones de sortie.	37
1.24	Schéma bloc d'un neurone à fonction d'activation de type sigmoïde. Les entrées pondérées e_i sont sommées et un biais b est ajouté. la fonction d'activation transforme la somme en une valeur comprise entre zéro et un.	38
1.25	Bilan énergétique des deux architectures de capteur. Les énergies minimale et maximale des éléments présentées dans le chapitre sont annotées.	39
1.26	Comparaison des budgets énergétiques des deux architectures pour a) l'application EEG et b) l'application ECG.	40
2.1	Schéma simplifié des différentes sections du cœur tiré de [Cda04].	44
2.2	Signal ECG d'un cycle cardiaque.	45
2.3	Chaîne de traitement typique pour la classification des signaux ECG.	47
2.4	Structure dyadique de filtrage pour une décomposition en ondelettes sur 3 niveaux.	48
2.5	Schéma bloc de la méthode de détection du complexe QRS de Pan et Tompkins [PT85].	49
2.6	Schéma bloc du système proposé pour la classification de signaux ECG [CCL ⁺ 16].	54
2.7	Exemple d'un ensemble de neurones pour la représentation de personnes.	56
2.8	Schéma bloc présentant les différentes fonctions d'un neurone.	56
2.9	Exemple de réseau à cliques avec stockage de trois cliques représentant trois personnes.	57
2.10	Illustration du décodage avec la règle <i>WsTA</i>	59
2.11	Illustration du décodage avec la règle <i>LsKO</i>	60
2.12	Distribution de la valeur du paramètre max(Cd1) pour chacune des classes. La plage ① représente la classe <i>LBBB</i> , et les plages ② et ③ respectivement <i>PB</i> et <i>RBBB</i>	62
2.13	Distribution de la valeur du paramètre max(Cd4). Une seule plage ① est créée du fait de la forte similitude entre <i>LBBB</i> et <i>PB</i>	63
2.14	Topologie du réseau proposé pour la classification d'ECG.	64
3.1	Schéma bloc du système complet proposé pour la classification des ECG.	70

3.2	Architecture de l'opérateur de filtrage.	71
3.3	Architecture réalisant la décomposition en ondelettes. Seuls les 3 premiers niveaux sont représentés.	72
3.4	Architecture de l'opérateur de minimum/maximum.	73
3.5	Architecture simplifiée de l'opérateur de variance. L'étape de division n'est pas effectuée pour réduire la consommation.	75
3.6	Architecture du système mixte pour le classifieur.	77
3.7	Structure de la mémoire des connexions synaptiques.	78
3.8	Circuit du neurone proposé et implanté dans [Lar15].	80
3.9	Simulation du neurone de [Lar15] pour une différence entre vainqueur et perdant de I_{unit}	81
3.10	Évolution de la résolution avec le nombre de contributions.	82
3.11	Simulation du neurone modifié pour une différence entre vainqueur et perdant de I_{unit}	83
3.12	Évolution de la résolution avec le nombre de contributions suite aux modifications.	84
3.13	Effet de l'augmentation du nombre de neurones en parallèles sur la tension V_B	85
3.14	Effet de l'augmentation du nombre de neurones sur le temps de réponse.	85
3.15	Mise en parallèle de k neurones pour former un bloc $WTA-k$	86
3.16	Circuit réalisant le WTA sur 128 neurones. Deux étages sont utilisés afin de réduire le nombre d'éléments en parallèle.	86
3.17	Architecture des interfaces numériques.	87
3.18	Signaux de contrôle pour pour une stimulation des premier et huitième neurones avec respectivement trois et une contributions.	88
3.19	Évolution des signaux dans le premier bloc $WTA-16$	89
3.20	Évolution des signaux dans le bloc $WTA-8$	89
3.21	Signaux de sortie numérique pour la stimulation considérée.	90
3.22	Layout d'un bloc $WTA-16$ et zoom sur des neurones.	91
3.23	Photo de la puce prise au microscope électronique.	91
3.24	Photo de la carte de test conçue.	92
3.25	Photo du banc de test.	93
3.26	Résultats du test des stimulations individuelles à dix contributions des neurones.	95
3.27	Réponse du réseau et des blocs $WTA-16$ lors des duels.	96
3.28	Profil de la consommation de courant durant le décodage du pire cas.	98
4.1	Exemple de spectre d'un signal à acquérir. La bande utile pour l'application considérée est B_2	104
4.2	Architecture d'un CAP telle que présentée dans [VB15].	105
4.3	Exemple de signaux alignés sur le complexe QRS issus de la base de données MIT-BIH [MIT] annotée.	107

4.4	Architecture du CAP proposé pour l'acquisition des signaux ECG.	108
4.5	Exemple des trois types d'ondelettes considérés avec un support de 180 ms et une fréquence de 22,5 Hz.	110
4.6	Évolution temporelle des 26 premières ondelettes suite à la discrétisation.	111
4.7	Exactitude de classification pour les différents types d'ondelettes considérés.	114
4.8	Exactitude de classification pour les différentes méthodes de sélection et classifieurs.	115
4.9	Exactitude de classification pour l'algorithme proposée pour une branche et sans limite de branches.	117
4.10	Architecture du convertisseur <i>A2I</i> tirée de [GAD ⁺ 14].	118
4.11	Consommation des différents éléments des trois architectures considérées : système purement transmetteur, traitement embarqué et combinaison CAP et traitement embarqué. La consommation correspond au traitement de dix secondes de signal.	119
4.12	Effet d'une mauvaise synchronisation sur les performances de classification. Les classifieurs sont entraînés avec une variance σ_a^2 et testé avec la variance σ_t^2 . Le nombre de paramètres utilisés est de 20.	120
4.13	Exactitude de classification pour différentes quantifications.	121

Liste des tableaux

1.1	Exemples de capteurs utilisés dans les <i>BASN</i> [LLW ⁺ 13].	7
1.2	Exemples de SAR utilisés pour les <i>BASN</i>	13
1.3	Exemples de convertisseurs $\Delta\Sigma$ utilisés pour les <i>BASN</i>	16
1.4	Implantations matérielles pour la reconstruction de signal.	24
1.5	Répartition des canaux de transmission pour la norme <i>IEEE 80.15.4</i>	26
1.6	Caractéristiques de transmetteurs pour les standards sans fil.	28
1.7	Implantations de classifieurs probabilistes pour les <i>BASN</i>	33
1.8	Implantations de <i>SVM</i> dans les <i>BASN</i>	37
1.9	Exemples d'utilisation de perceptrons sur des signaux biologiques.	38
2.1	Caractéristiques d'un signal ECG à 60 battements par minute pour un adulte sain [LSCCM16].	45
2.2	Synthèse des méthodes utilisées pour classifier le type d'arythmie dans les ECG.	52
2.3	Détail des plages utilisées pour caractériser les trois classes pour les 18 paramètres extraits.	63
2.4	Matrice de confusion du système.	65
2.5	Performances de classification du système proposé.	65
3.1	Coefficients des filtres pour l'ondelette <i>db2</i>	70
3.2	Matrice de confusion obtenue avec le prototype.	97
3.3	Performances de classification simulées et mesurées.	97
3.4	Comparaison entre le système proposé et un système purement transmetteur.	99
3.5	Évolution de la consommation avec le courant unitaire.	100
4.1	Répartition de l'énergie pour des circuits réalisant la classification d'ECG.	104

Liste des acronymes

A2I	Analog-to-Information
ANN	Artificial Neural Network
ASIC	Application Specific Integrated Circuit
BASN	Body Area Sensor Network
BLE	Bluetooth Low Energy
BP	Basic Pursuit
BPDN	Basic Pursuit De-Noising
CA1	Complément à un
CA2	Complément à deux
CAN	Convertisseur Analogique Numérique
CAP	Convertisseur Analogique vers Paramètres
CMOS	Complementary Metal Oxyde Semiconductor
CNA	Convertisseur Numérique Analogique
CoSaMP	Compressive Sampling Matching Pursuit
CS	Compressive Sensing
DWT	Discrete Wavelet Transform
ECG	Électrocardiogramme
EEG	Électroencephalogramme
EMG	Électromyogramme
ENOB	Effective Number Of Bits
FF	Fast Fast
FN	Faux Négatif
FOM	Factor Of Merit
FP	Faux Positif
FPGA	Field-Programmable Gate Array
HBC	Human Body Communication
HOS	High-Order Statistics
ICA	Idependant Component Analysis
IF	Importance Forêt
LBBS	Left Bundle Branch Block

LDA	Linear Discriminant Analysis
LR	Linear Regression
LSB	Least Significant Bit
LsKO	Losers-Kicked-Out
MAC	Multiplication Accumulation
MSB	Most Significant Bit
NB	Naive Bayesian
NIHT	Normalized Iterative Hard Thresholding
OMP	Orthogonal Matching Pursuit
OOB	Out-Of-Bag
OSI	Open Systems Interconnection
PB	Paced Beat
PCA	Principal Component Analysis
PISO	Parallel In Serial Out
PNN	Probabilistic Neural Network
PRD	Percentage Root-square Difference
QDA	Quadratic Discriminant Analysis
RBBB	Right Bundle Branch Block
RBF	Radial Basis Function
RF	Radio-Fréquence
RIF	Réponse Impulsionnelle Finie
RIP	Restricted Isometry Property
S/H	Sample and Hold
SAR	Successive Approximation Register
SF	Score de Fisher
SIPO	Serial In Parallel Out
SNDR	Signal over Noise and Distortion Ratio
SNR	Signal over Noise Ratio
SS	Slow Slow
SVM	Support Vector Machine
TC	Taux de Compression
TT	Typical Typical
VN	Vrai Négatif
VP	Vrai Positif
WsTA	Winners-Take-All
WTA	Winner-Takes-All

Publications associées

Conférences

P. Chollet, K. Colombier, C. Lahuec, M. Arzel and F. Seguin. Toward sub-pJ per classification in Body Area Sensor Networks. In *2016 14th IEEE International New Circuits and Systems Conference (NEWCAS)*, pages 1-4, June 2016.

P. Chollet, B. Larras, C. Lahuec, F. Seguin and M. Arzel. An ultra-low power iterative clique-based neural network integrated in 65-nm CMOS. In *2017 15th IEEE International New Circuits and Systems Conference (NEWCAS)*, pages 5-8, June 2017.

B. Larras, **P. Chollet**, C. Lahuec, M. Arzel and F. Seguin. A 65-nm CMOS 7fJ per synaptic event clique-based neural network in scalable architecture. In *2017 IEEE International Symposium on Circuits and Systems*, 2017.

P. Chollet, C. Lahuec, M. Arzel and F. Seguin. A Sub-nJ CMOS ECG Classifier for Wireless Smart Sensor. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017.

Introduction

L'évolution des technologies d'intégration des ces 50 dernières années a permis une forte amélioration de l'efficacité énergétique des circuits intégrés. L'apparition de petites unités basse consommation, dont la taille peut ne pas dépasser quelques millimètres carrés, ouvre la voie au déploiement de réseaux de capteurs sans fil. Un capteur est un élément dont la fonction est d'acquérir des données sur son environnement. Il peut ensuite directement transmettre ces données à un agrégateur ou faire un traitement de ces dernières avant la transmission. Le déploiement de nombreux capteurs forme un réseau dont le but est de surveiller l'évolution de l'environnement dans lequel ils sont placés. C'est un marché en plein essor dont la valeur est passée de 300 millions de dollars en 2010 à plus d'un milliard en 2014. On estime qu'il atteindra deux milliards de dollars d'ici les années 2020 [RSCB14]. La surveillance d'activité sismique [WALR⁺06], le contrôle du vieillissement de structures [KSC⁺08] ou la détection de tireurs embusqués [SML⁺04] sont des exemples de réseaux existants. Les domaines d'applications sont très nombreux et il existe actuellement une multitude de projets de déploiement [RSCB14].

Un domaine particulièrement intéressant est celui des réseaux de capteurs corporels. Ces derniers consistent à placer des unités d'acquisition sur une personne afin de suivre l'évolution de paramètres comme la fréquence cardiaque, la vitesse de déplacement ou la température. Leur apparition auprès du grand public se fait grâce au très fort développement des montres et bracelets connectés. Combinés avec un smartphone, ils permettent de faire du suivi d'activité physique pour les sportifs ainsi que de la surveillance des phases du sommeil.

La principale motivation pour développer les réseaux de capteurs corporels est le suivi médical. En effet, depuis 17 ans, les deux principales causes de mortalité sont les problèmes cardiaques et les accidents vasculaires cérébraux. Elles représentent un total de 15 millions de décès pour l'année 2015 [OMS]. Beaucoup d'entre eux peuvent être évités par une intervention rapide dès l'apparition des premiers symptômes. C'est l'un des objectifs des réseaux de capteurs corporels. Ils offrent un suivi complet, depuis le domicile, grâce à la surveillance de la progression de maladies. Cela permet une réduction de la fréquence des visites de contrôle et un traitement préventif plus efficace, conduisant à une amélioration de la qualité des soins ainsi qu'à une baisse des coûts. Les réseaux de capteurs corporels peuvent aussi être utilisés dans les zones ayant peu accès à des services médicaux tels que des régions très rurales ou encore des soldats et des astronautes en mission.

Les solutions techniques proposées actuellement se confrontent à de nombreux verrous, rendant complexe le déploiement des réseaux de capteurs corporels. Leur mise en place nécessite d'acquérir, de transmettre et de traiter un nombre important de données physiologiques complexes. La mesure approximative du rythme cardiaque des montres connectées ne permet pas une précision suffisante pour détecter un problème. Ce besoin de précision impose de fortes contraintes sur les capteurs. Ces derniers doivent ainsi être le plus proche possible de la source du signal à acquérir. Ils doivent donc être portés sur les vêtements ou la peau, voire implantés sous cette dernière. Ainsi, il faut des capteurs les plus petits possibles afin de ne pas causer de gêne au porteur, l'idéal étant que ce dernier l'oublie. De plus, l'autonomie est un facteur extrêmement important dans leur conception. Il est en effet très gênant pour les personnes disposant d'implants cochléaires de devoir changer les piles toutes les deux ou trois semaines voire tous les trois ou quatre jours. Cette contrainte est d'autant plus importante pour des capteurs implantés sous la peau qui ne permettent pas un changement simple de la source d'énergie.

La consommation des capteurs est principalement due à la transmission d'un nombre important de données vers un agrégateur. Ce transfert peut représenter jusqu'à 96 % du budget énergétique [VSB⁺10]. Il faut réduire la consommation de cette étape afin d'augmenter l'autonomie énergétique. Une première méthode consiste à améliorer l'efficacité énergétique de la transmission grâce à des protocoles spécifiques permettant des implantations basse consommation. Une seconde manière consiste à réduire la quantité de données à transmettre. Cela peut se faire avec des techniques de compression telles que l'échantillonnage compressif [Don06] et/ou en traitant les données sur le capteur. L'inconvénient de l'échantillonnage compressif est qu'une reconstruction coûteuse du signal doit être réalisée sur l'agrégateur. De même, les traitements sur les signaux biologiques sont parfois complexes et leur intégration sur le capteur peut conduire à une trop forte augmentation de la taille ou de la consommation. Il faut donc trouver un compromis entre traitement sur le capteur et quantité de données à transmettre afin de minimiser le budget énergétique de ce dernier.

Dans ce document, notre but est de proposer un système de traitement parcimonieux pour les signaux biologiques. Dans un premier temps, le Chapitre 1 présente les réseaux de capteurs corporels. L'objectif du chapitre est d'identifier les différentes sources de dépense énergétique. Pour cela, la structure du réseau ainsi que les différents éléments qui le composent sont présentés. Pour chaque élément de la chaîne de traitement, les principes et des implantations matérielles sont mis en avant.

Suite à cette étude, le Chapitre 2 met l'accent sur les signaux de type électrocardiogramme (ECG). Les spécificités du signal ainsi que les traitements classiques sur ce dernier sont d'abord présentés. Un système d'analyse de signaux ECG embarqué sur un capteur est ensuite proposé. Il utilise un classifieur parcimonieux reposant sur les réseaux de neurones à cliques [WBLH69][GB11]. Des simulations comportementales sont effectuées afin de valider son fonctionnement.

Le Chapitre 3 se focalise sur la réalisation d'un prototype du système précédemment proposé. Il s'appuie sur un système numérique pour le pré-traitement et sur un circuit mixte ana-

logique/numérique afin de classifier les signaux ECG. La consommation du pré-traitement est estimée avec *Synopsys Design Compiler*[®] pour la technologie *ST CMOS* 65 nm. Le circuit mixte est ensuite fabriqué avec la même technologie puis les performances de classification mesurées sont comparées à celles du modèle comportemental. Enfin, la consommation du prototype est mesurée et une comparaison avec d'autres solutions intégrées existantes est réalisée.

Le Chapitre 4 présente ensuite une méthode récente d'acquisition reposant sur la conversion analogique vers paramètres [VB15]. Elle consiste à extraire des paramètres directement sur le signal analogique. Ces paramètres sont ensuite utilisés par des classifieurs simples. Une architecture d'acquisition directe des paramètres est proposée pour les signaux ECG. Un modèle comportemental est réalisé afin de valider l'intérêt de la solution proposée. De plus, une estimation rapide de la consommation est effectuée pour évaluer les gains énergétiques. L'ensemble des travaux présentés dans le chapitre sont issus d'une collaboration de quatre mois avec l'équipe de Christoph Studer de l'université de Cornell. Cette collaboration a été possible grâce à l'obtention d'une bourse Fulbright ainsi qu'une bourse de mobilité de la part de l'Université Bretagne Loire et de Télécom Bretagne.

En conclusion, le résumé des contributions est dressé et des perspectives de travaux sont proposées. Ces perspectives permettent d'aller plus loin dans la réalisation matérielle des solutions proposées afin d'entrevoir une réalisation complète d'un capteur intégré pour le traitement de signaux ECG ou d'autres signaux biologiques.

Chapitre 1

Les réseaux de capteurs corporels

Introduction

Les réseaux de capteurs corporels ont un potentiel d'amélioration de la qualité des soins via un suivi des paramètres physiologiques. La mise en application de tels réseaux nécessite la conception de capteurs très basse consommation afin de répondre aux contraintes d'intégration. Pour réduire la consommation de ces derniers, il convient, dans un premier temps, de les analyser et de déterminer les différentes sources de consommation énergétique. Dans ce chapitre, la structure des réseaux de capteurs corporels est d'abord présentée. Les différentes méthodes d'acquisition et de transmission classiquement utilisées sont ensuite étudiées. Enfin, les techniques de traitement utilisées pour classifier les signaux sont décrites. Pour chaque fonction du capteur, les implantations matérielles existantes dans le cadre des réseaux de capteurs corporels sont évaluées. L'accent est mis sur la consommation d'énergie afin d'identifier les éléments sur lesquels il convient d'agir pour améliorer l'autonomie des capteurs.

1.1 Présentation des réseaux de capteurs corporels

1.1.1 Les réseaux de capteurs corporels

D'après l'Organisation Mondiale de la Santé, les problèmes cardiaques et les accidents vasculaires cérébraux sont les deux premières causes de mortalité dans le monde avec un total de 15 millions de morts en 2015 [OMS]. Elles représentent 26,6% des décès. Ces deux causes sont majoritaires depuis 2010. Dans de nombreux cas, il est possible d'éviter le décès via une détection anticipée des symptômes. C'est l'un des défis que les réseaux de capteurs corporels tentent de relever via un suivi continu des paramètres physiologiques.

Un réseau de capteurs corporels (*Body Area Sensor Network, BASN*) est composé d'unités dont le rôle est d'acquérir des signaux biologiques et des informations contextuelles liées à la personne et à son environnement. Les capteurs sont des petites unités de traitement et d'acquisition très basse consommation. Ils peuvent être portés sur les vêtements ou la peau voire être implantés sous

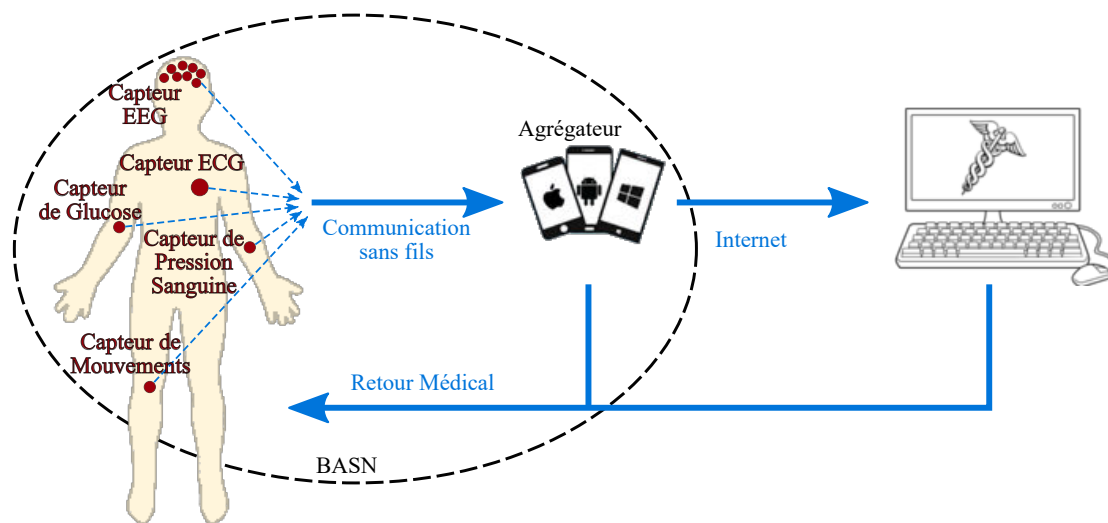


FIGURE 1.1 – Structure globale d'un *BASN* : les capteurs transmettent des données à l'agrégateur qui les traite et prévient les secours en cas de nécessité.

cette dernière afin d'être au plus proche de la source du signal. Leur petite taille et leur grande autonomie énergétique permettent un suivi à domicile à très bas coût grâce à un équipement très peu coûteux ainsi qu'à la réduction des déplacements en centre spécialisé.

Les capteurs transmettent les données acquises à un agrégateur qui les reçoit et les traite. Le réseau de capteur corporel illustré sur la Figure 1.1 utilise un smartphone en tant qu'agrégateur afin de s'intégrer au mieux à l'environnement de l'utilisateur. L'agrégateur collecte les données des capteurs et peut alors soit les relayer à un serveur médical distant, soit les traiter pour détecter une anomalie. Lorsqu'une anomalie est présente, l'agrégateur peut prévenir l'utilisateur. De plus, l'agrégateur peut contacter les services médicaux compétents afin de relayer les signaux problématiques ou prévenir les secours en cas de besoins.

Le type et le nombre de capteurs utilisés sont dépendants de l'application du *BASN*. Les capteurs de type électrocardiogramme (ECG) ou électromyogramme (EMG) ainsi que les accéléromètres et les gyroscopes sont des exemples de capteurs pouvant être utilisés. D'une manière générale, les capteurs corporels peuvent être classés en deux catégories : les capteurs continus et les discrets. Le Tableau 1.1 présente quelques capteurs utilisés pour les *BASN* [LLW⁺13] ainsi que le type de signal mesuré. La spécificité des différents signaux détermine aussi la fréquence à laquelle ces derniers sont acquis.

Dans un *BASN* ayant pour objectif de détecter des anomalies, il y a trois étapes principales réparties entre les capteurs et l'agrégateur. La première consiste à obtenir une représentation numérique du ou des signaux à analyser. Cette étape appelée acquisition est toujours réalisée par le capteur. Une fois les signaux acquis, il faut les traiter afin de détecter une éventuelle anomalie. Pour ce faire, des paramètres sont extraits puis un traitement est réalisé pour déterminer si une anomalie est présente. Ce traitement est en général une classification ou une régression. Les opérations

TABLEAU 1.1 – Exemples de capteurs utilisés dans les *BASN* [LLW⁺13].

Signal/Grandeur mesuré(e)	Capteur	Type de signal	Fréquence d'acquisition
Battement cœur	ECG	Continu	100-1000 Hz
Activité cérébrale	EEG	Continu	100-1000 Hz
Activité musculaire	EMG	Continu	100-1000 Hz
Mouvements	Accéléromètre	Continu	10-100 Hz
Mouvements	Gyroscope	Continu	10-100 Hz
Taux d'oxygénation du sang	Oxymètre	Discret	0,1-10 Hz
Température	Thermomètre	Discret	0,001-0,1 Hz
Humidité	Capteur d'humidité	Discret	0,001-0,1 Hz

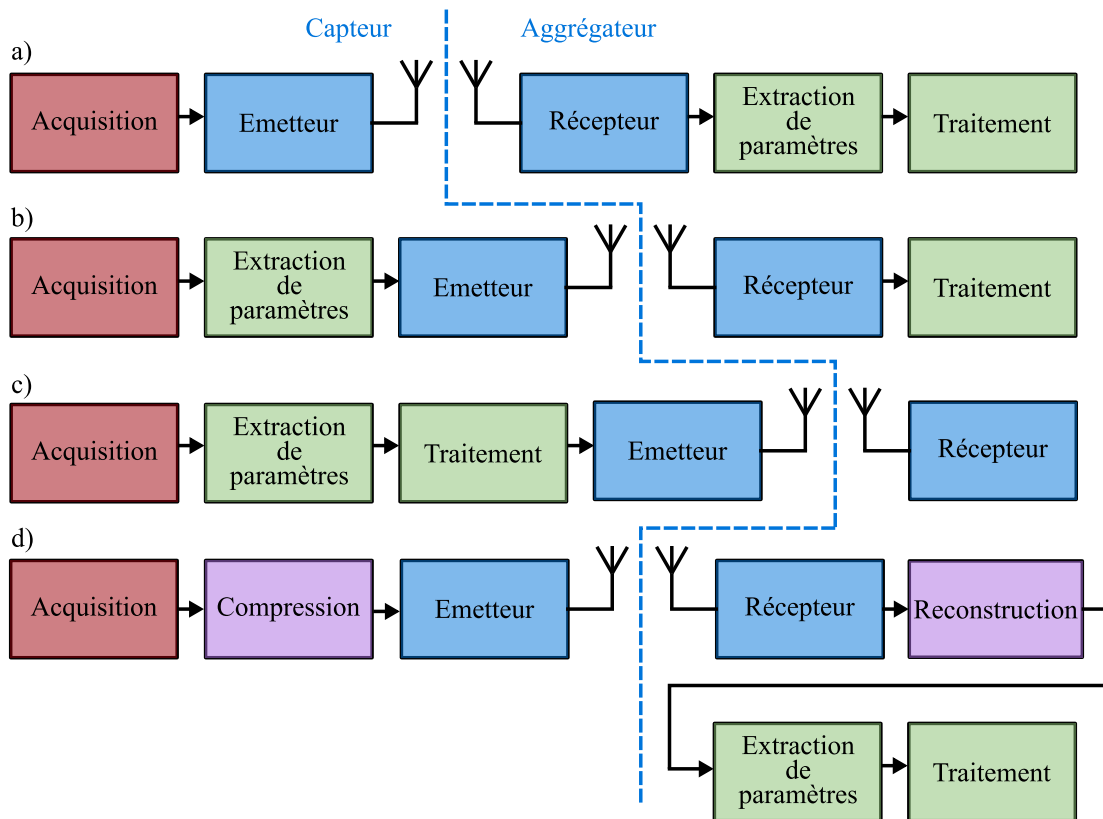


FIGURE 1.2 – Répartition des opérations entre capteur et agrégateur pour les *BASN* détectant des anomalies. a) capteur transmetteur, b) extraction sur capteur, c) traitement sur capteur, d) compression des données.

peuvent être réparties de différentes manières entre le capteur et l'agrégateur. La Figure 1.2a présente un capteur transmettant directement le signal à l'agrégateur. On parle alors de capteur purement transmetteur. Ce cas de figure est privilégié lorsque la classification nécessite l'utilisation simultanée de signaux provenant de différents capteurs.

Pour des signaux plus complexes comme les ECG ou EEG, la transmission directe du signal

à l'agrégateur est possible mais nécessite une utilisation continue de l'étage de communication. Cette utilisation intensive de la transmission a un coût non négligeable sur le budget énergétique du capteur. Dans [VSB⁺10] la transmission radio d'un signal EEG représente presque 96% du budget énergétique d'un tel capteur. Une solution pour réduire cet impact est d'utiliser un capteur plus complexe réalisant l'extraction des paramètres de manière locale (Figure 1.2b) avant de les transmettre. En allant plus loin, il est possible d'embarquer l'étape de traitement sur le capteur (Figure 1.2c). On parle alors de capteur intelligent. Ceci n'est possible que si la classification est réalisable en utilisant seulement les paramètres du capteur. De plus, il peut être complexe d'intégrer les algorithmes de classification. Enfin, une autre méthode classique pour réduire l'utilisation de l'étage radio est de mettre en place une méthode de compression (Figure 1.2d) en utilisant les caractéristiques du signal. Cette méthode nécessite alors une étape supplémentaire de reconstruction du signal sur l'agrégateur.

1.1.2 Les défis des réseaux de capteurs corporels

Les *BASN* jouent un rôle important pour la démocratisation de l'internet des objets. Si les outils de remise en forme comme les montres et bracelets connectés ont gagné leur place auprès du grand public [Coo], c'est bien le domaine de la santé qui porte les plus gros espoirs autour de cette technologie [Pri]. De nombreux défis doivent être relevés afin que les *BASN* puissent devenir des "objets" du quotidien [LIY16][JA14].

Un premier défi concerne les régulations dans le domaine médical. Il est en effet très compliqué d'obtenir les certifications nécessaires afin de pouvoir considérer un produit en tant que dispositif médical. Les démarches peuvent être très longues et onéreuses. Pour obtenir ces certifications, il est indispensable de concevoir des circuits fiables, robustes et durables afin d'assurer un bon fonctionnement.

Un deuxième verrou concerne la sécurité des données. Il est en effet impératif qu'une personne portant un *BASN* ait des garanties concernant la sécurité des données récupérées par le réseau. Il convient donc que ces données restent privées afin d'en éviter une utilisation commerciale voire malicieuse. De plus, les données ne doivent pas pouvoir être corrompues lors de la transmission au risque de voir des mauvaises prises de décisions pouvant menacer le patient. Le développement de *BASN* nécessite donc la création de protocoles de communications garantissant cette sécurité. De plus, ces protocoles doivent être intégrés sur le capteur, ajoutant de la complexité à ce dernier.

La communication entre les capteurs est l'agrégateur étant souvent sans fil, des problèmes d'interférences peuvent apparaître. L'augmentation du nombre d'objets connectés entraîne une utilisation forte du spectre radiofréquence qui devient une ressource rare. Il est donc nécessaire de mettre en place des mécanismes permettant l'interopérabilité des différents protocoles de communication.

Les défis cités nécessitent tous la mise en place de mécanismes spécifiques au sein du capteur. Ces derniers sont complexes et viennent s'ajouter à la fonction première, créant une contrainte forte sur une ressource : l'énergie. En effet, les capteurs ont vocation à être portés sur le corps voire

implantés sous la peau. Dans le cas d’une implantation, il n’est pas simple d’apporter une source externe d’énergie au capteur. Il est ainsi impératif de concevoir des circuits basse consommation afin d’avoir une autonomie énergétique la plus grande possible.

Les défis à surmonter pour démocratiser les *BASN* sont donc nombreux et en font un sujet de recherche très actif. Ils concernent à la fois des aspects communication, sécurité et circuit. Dans la suite de cette étude, l’accent est porté sur les aspects énergétiques des *BASN*. Le but est d’identifier les éléments utilisant le plus d’énergie afin de réduire le budget énergétique des capteurs. Pour cela les différentes opérations présentées dans la Figure 1.2 sont détaillées dans la suite de ce chapitre.

1.2 Acquisition des données : du signal analogique vers une représentation numérique

L’acquisition du signal est la première étape dans le processus de détection d’anomalies des *BASN*. Elle peut être faite de manière classique via un échantillonnage à la fréquence de Nyquist ou en utilisant les caractéristiques du signal pour appliquer des techniques d’échantillonnage compressif.

1.2.1 Acquisition classique de Nyquist

1.2.1.1 Principe

La méthode classique d’acquisition consiste à transformer un signal analogique en un signal numérique permettant de parfaitement le représenter. Un signal analogique est un signal continu en temps et en valeurs. À l’opposé, un signal numérique est un signal discret en temps et en valeurs. Ces valeurs sont codées de manière binaire. La conversion analogique numérique se fait donc via une étape d’échantillonnage suivie d’une quantification comme illustré en Figure 1.3.

L’échantillonnage uniforme consiste à discrétiser temporellement le signal analogique en relevant sa valeur à intervalle constant. Il est caractérisé par la fréquence d’échantillonnage F_e . Cette étape peu s’exprimer mathématiquement sous la forme :

$$v_e(k \cdot T_e) = v(t) \cdot \delta_{T_e}(t) \tag{1.1}$$

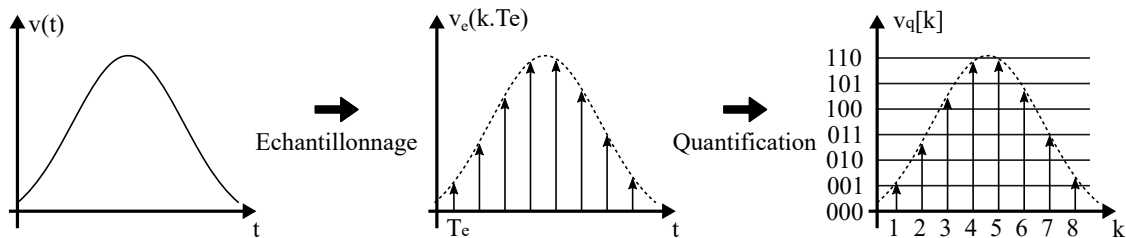


FIGURE 1.3 – Conversion d’un signal analogique en signal numérique. Le signal est d’abord échantillonné avant d’être quantifié.

où $v_e(k \cdot T_e)$ est le $k^{\text{ème}}$ élément du signal échantillonné, $v(t)$ et $\delta_{T_e}(t)$ sont respectivement le signal analogique et un peigne de Dirac de période $T_e = 1/F_e$. L'échantillonnage uniforme doit permettre la reconstruction du signal analogique à partir du signal échantillonné. Pour cela il est nécessaire d'analyser l'opération d'échantillonnage dans le domaine fréquentiel. Dans ce dernier, la multiplication est transformée en produit de convolution et un peigne de Dirac de période T_e est un peigne de Dirac d'amplitude $1/T_e$ et de période $F_e = 1/T_e$. Ainsi, dans le domaine fréquentiel, l'équation (1.1) s'écrit :

$$V_e(f) = (1/T_e) \cdot V(f) * \delta_{F_e}(f) \quad (1.2)$$

où $V_e(f)$ et $V(f)$ représentent respectivement les spectres de $v_e(t)$ et $v(t)$ et $*$ est l'opérateur du produit de convolution. L'opération de convolution est représentée graphiquement en Figure 1.4. La convolution a pour effet de recopier le spectre du signal analogique autour des fréquences multiples de F_e .

Afin de retrouver le signal le départ il suffit d'effectuer un filtrage passe-bas qui ne conserve que le signal contenu dans la bande $[-F_n; F_n]$. La fréquence F_n est appelée fréquence de Nyquist et s'exprime comme $F_n = F_e/2$. Pour assurer une bonne reconstruction du signal, il faut que la fréquence d'échantillonnage soit suffisamment grande pour éviter le recouvrement de spectre. Le recouvrement apparaît lorsque les différentes copies du spectre se superposent. Cet effet est illustré sur la Figure 1.5. Avec du recouvrement, il devient impossible de récupérer le signal de départ. La condition pour éviter les effets de recouvrement de spectre est connue sous le nom de

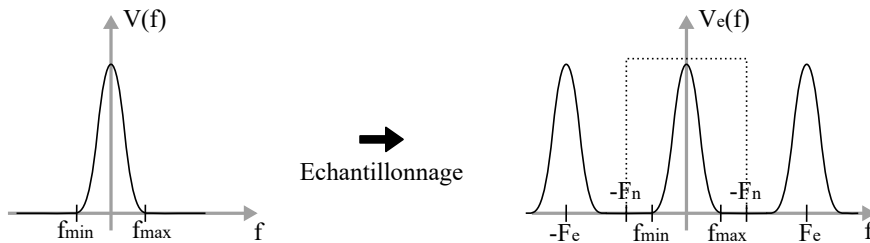


FIGURE 1.4 – Spectre fréquentiel d'un signal avant et après échantillonnage.

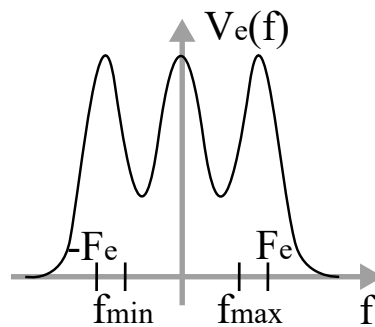


FIGURE 1.5 – Un effet de repliement de spectre apparaît lorsque la fréquence d'échantillonnage est trop basse.

théorème de Shannon et s'exprime :

$$F_n = F_e/2 \geq F_{max} = \max(|f_{min}|, |f_{max}|). \quad (1.3)$$

En pratique, il est impossible d'obtenir un filtre passe-bas assez sélectif permettant la reconstruction du signal si $F_n = F_{max}$ et une marge est donc prise.

La seconde étape pour obtenir un signal numérique est la quantification. Elle consiste à associer un code binaire à chaque échantillon du signal. Cette opération est réalisée par le convertisseur analogique numérique (CAN). Un CAN est défini par une plage d'entrée et un nombre de bits, appelé résolution, pour le code. Un convertisseur N bits pourra prendre 2^N valeurs en sortie. Avec la plage d'entrée et la résolution, il est possible de déterminer le *quantum* q . Il correspond à la plus petite variation sur le signal d'entrée entraînant un changement de code en sortie. Pour un CAN avec une plage d'entrée $[V_{min}; V_{max}]$ et une résolution de N bits, le *quantum* s'exprime :

$$q = \frac{V_{max} - V_{min}}{2^N}. \quad (1.4)$$

La fonction de transfert d'un CAN idéal 3 bits est présentée sur la Figure 1.6. La droite correspond à la droite de transfert idéale. Les tensions de basculement V_i entre codes sont définies comme :

$$V_i = \frac{q(1 + 2(i - 1))}{2}. \quad (1.5)$$

L'erreur de quantification est inhérente à toute opération de quantification. Elle s'exprime en fonction du *quantum* et correspond à la différence entre la valeur du signal quantifié et celle du signal d'entrée. Ainsi, plus le *quantum* est faible, plus le bruit de quantification est réduit.

La caractérisation d'un CAN se fait généralement en appliquant un signal sinusoïdal en entrée et en analysant le spectre du signal en sortie. Ce dernier est composé de différentes raies à la fréquence fondamentale de l'entrée ainsi qu'aux fréquences harmoniques. Il peut aussi apparaître des raies parasites ne correspondant pas à des harmoniques. Les métriques principalement utilisées

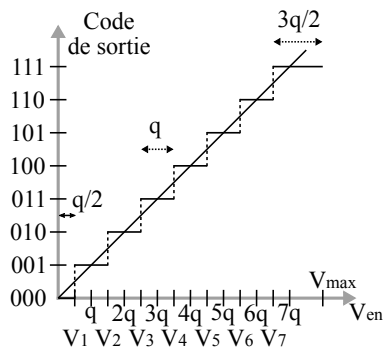


FIGURE 1.6 – Fonction de transfert d'un CAN 3 bits. La sortie change de code lorsque l'entrée V_{en} atteint les seuils V_i . La droite représente la fonction de transfert idéale.

pour évaluer les CAN sont les suivantes :

- le rapport signal sur bruit (SNR) correspond au rapport de la puissance du signal d'entrée sur la puissance du bruit, sans considérer la composante continue, les harmoniques et les fréquences parasites.
- le rapport signal sur bruit et distorsion ($SNDR$) correspond au rapport de la puissance du signal sur celle du bruit, des harmoniques et des parasites.
- le nombre de bits effectif ($ENOB$) correspond au nombre de bits nécessaire pour qu'un CAN idéal ait un SNR équivalent au SNDR du CAN caractérisé.
- le facteur de mérite (FOM) permet de comparer les implantations. Il s'exprime comme :

$$FOM = \frac{Puissance}{F_e \cdot 2^{ENOB}} \quad (1.6)$$

et permet de normaliser la consommation du CAN par rapport à la bande passante et la dynamique d'entrée. Il s'exprime en énergie par conversion et est très souvent utilisé pour comparer différentes implantations.

1.2.1.2 Les architectures de CAN pour les signaux biologiques

L'acquisition de signaux biomédicaux nécessite une fréquence de l'ordre du kHz et une précision de l'ordre de la dizaine de bits. Les architectures les plus utilisées pour les acquérir sont le registre à approximations successives et le modulateur delta-sigma. Elles sont donc présentées dans cette section.

1.2.1.2.1 Registre à approximations successives

L'architecture la plus utilisée pour l'acquisition de signaux biologiques est le registre à approximations successives (*Successive Approximation Register, SAR*) présenté en Figure 1.7. Elle est composée d'un échantillonneur (*Sample and Hold, S/H*), d'un comparateur, d'un convertisseur numérique analogique (CNA) et de logique de contrôle. Le SAR effectue une recherche binaire pour décider du code de sortie. Un algorithme d'approximations successives est utilisé afin de réduire la complexité de la recherche.

Dans un premier temps, l'échantillonneur copie la valeur V_{en} en V_1 et maintient cette dernière constante durant la conversion. Au départ, le registre est initialisé avec le bit de poids fort (*Most*

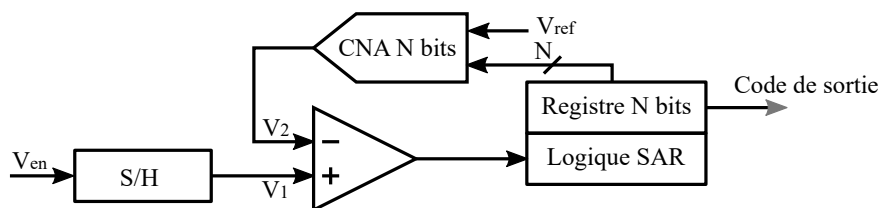


FIGURE 1.7 – Architecture globale d'un convertisseur à approximations successives.

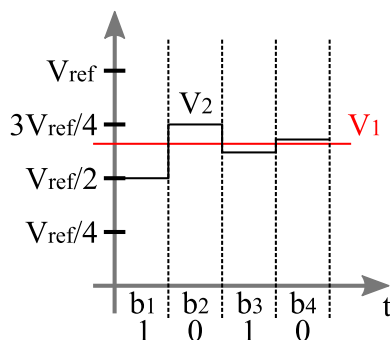


FIGURE 1.8 – Exemple de conversion d’une entrée fixe par un convertisseur SAR. La ligne rouge représente le signal d’entrée et la noire la sortie du CNA.

Significant Bit, MSB) à un et les autres à zéro. Cela équivaut à avoir en sortie du CNA une valeur de $V_{ref}/2$ pour V_2 . La valeur V_2 est ensuite comparée à V_1 afin de prendre une décision sur la valeur du bit. Si $V_2 \geq V_1$ alors le *MSB* est mis à zéro. Dans le cas contraire, il reste à un. La logique passe ensuite au bit suivant en appliquant la même méthode : le bit est d’abord forcé à un pour mettre à jour la sortie du CNA. La comparaison permet ensuite de prendre une décision sur le bit. Cette opération est répétée jusqu’au bit de poids faible (*Least Significant Bit, LSB*).

Un exemple de conversion sur quatre bits est donnée sur la Figure 1.8. La courbe noire représente la valeur V_2 en sortie du CNA et la rouge la valeur V_1 en sortie de l’échantillonneur. Le bit b_1 est le *MSB* et b_4 le *LSB*. Au départ, V_2 est fixée à $V_{ref}/2$. V_2 étant inférieure à V_1 , le bit b_1 reste à un. Le registre est ensuite forcé à la position 1100 conduisant à une valeur de V_2 de $3V_{ref}/4$. Comme V_1 est inférieure à cette valeur, le bit b_2 est donc passé à zéro et le registre passe à la position 1010. Dans cette configuration, $V_2 = 5V_{ref}/8 < V_1$ conduit à une décision de ‘1’ logique pour b_3 . Enfin le registre est mis en position 1011 et la comparaison permet de choisir la valeur logique ‘0’ pour le *LSB* car $V_2 > V_1$.

Cet exemple permet de voir qu’il faut quatre itérations pour décider de la sortie d’un code sur quatre bits. De manière générale, un *SAR* sur N bits nécessite N itérations pour convertir une nouvelle entrée. C’est beaucoup moins que les 2^N nécessaires pour effectuer une recherche exhaustive. Pour fonctionner correctement, la logique interne du *SAR* doit fonctionner à une fréquence

TABLEAU 1.2 – Exemples de SAR utilisés pour les *BASN*.

	[VC07]	[YYKY10]	[YYED ⁺ 13]	[YMPH08]
Signal	-	ECG	EEG	EEG
Technologie <i>CMOS</i> (μm)	0.18	0.18	0.18	0.5
Puissance (μW)	25	2.2	1	23
F_e (Hz)	100k	100k	32k	25k
Résolution (Bits)	12	10	10	11
ENOB	10.55	9.2	-	10.5
FOM (pJ)	0.17	0.12	-	2

au moins N fois supérieure à la fréquence d'échantillonnage. Si cela peut être une limite pour des signaux hautes fréquences pouvant nécessiter des fréquences d'échantillonnage de l'ordre du GHz, ce n'est pas du tout une contrainte pour les signaux biologiques dépassant rarement la centaine de kHz. De plus, étant donné son fonctionnement, la puissance dissipée est proportionnelle à la fréquence d'échantillonnage. Cela en fait un très bon candidat pour les *BASN* qui nécessitent des opérations à basse énergie et des fréquences d'échantillonnage faibles. Le Tableau 1.2 présente les caractéristiques de *CAN SAR* utilisés dans le cadre des *BASN*. On peut constater que la résolution est de l'ordre de 10 bits pour des fréquences d'échantillonnage maximales de 100 kHz. Les facteurs de mérite sont compris entre la centaine de fJ et quelques pJ.

1.2.1.2.2 Delta-Sigma

Le convertisseur delta-sigma ($\Delta\Sigma$) repose sur le principe de sur-échantillonnage et de filtrage de bruit. L'architecture principalement numérique (Figure 1.9) est constituée d'un modulateur $\Delta\Sigma$, d'un filtre passe-bas numérique ainsi que d'un décimateur.

Le sur-échantillonnage résulte en un étalement du bruit sur la bande passante du CAN. Cela permet d'améliorer le *SNR* par rapport aux convertisseurs échantillonnant à la fréquence de Nyquist. La figure 1.10 montre les réponses fréquentielles de deux CAN à une entrée sinusoïdale de fréquence F_s . Les CAN classiques, en bleu, échantillonnent le signal à la fréquence F_e . Le bruit de quantification étant uniquement fonction de la résolution, sa puissance sera identique quelque soit la fréquence d'échantillonnage. Le sur-échantillonnage permet de répartir cette puissance sur une

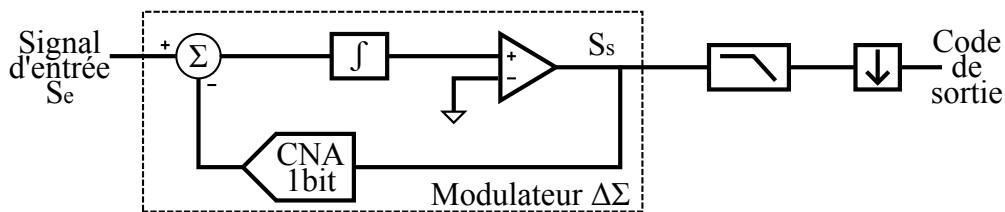


FIGURE 1.9 – Architecture du CAN $\Delta\Sigma$.

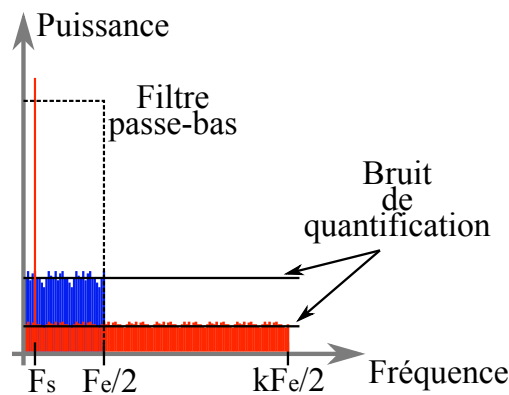


FIGURE 1.10 – Représentation fréquentielle de la sortie d'un CAN à la fréquence de Nyquist (bleu) et via sur-échantillonnage (rouge).

bande plus grande. L'utilisation d'un filtrage passe-bas en sortie permet de supprimer une partie du bruit et donc d'améliorer le SNR . L'augmentation du SNR est de 6 dB par augmentation du facteur de sur-échantillonnage k de quatre. Cela est équivalent à ajouter un bit de résolution sur un CAN sans sur-échantillonnage. Ainsi, pour un convertisseur $\Delta\Sigma$ utilisant une conversion sur un bit, un CAN N bits peut être obtenu avec un facteur de sur-échantillonnage de 4^{N-1} . Cela peut être problématique pour des signaux nécessitant une grande précision puisque pour obtenir un CAN 16 bits il faut sur-échantillonner par un facteur de plus d'un milliard.

La notion de filtrage de bruit permet d'améliorer encore plus le SNR . Le principe consiste à effectuer un filtrage du bruit de quantification afin de déplacer ce dernier dans les fréquences hautes. Cela se fait grâce au modulateur $\Delta\Sigma$ constitué d'un soustracteur, d'un intégrateur, d'un comparateur et d'un CNA 1 bit. Le CNA connecte l'entrée négative du soustracteur à une valeur V_{ref} ou $-V_{ref}$ et peut être réalisé par un simple interrupteur. La sortie du modulateur contient une densité de '1' logique proportionnelle à la valeur du signal d'entrée. Le fait d'intégrer l'erreur permet de réaliser un filtrage passe-bas bas pour le signal en même temps qu'un filtrage passe-haut pour le bruit de quantification. Cela peut s'expliquer en analysant le système d'un point de vue fréquentiel (Figure 1.11).

On a :

$$S_1 = S_e - S_s \quad (1.7)$$

$$S_2 = S_1/f \quad (1.8)$$

$$S_s = S_2 + b_q \quad (1.9)$$

où b_q représente le bruit de quantification ajouté par le comparateur. Il est alors possible d'exprimer S_s en fonction de S_i et b_q . On obtient :

$$S_s = \frac{S_i}{1+f} + \frac{q \cdot f}{1+f} \quad (1.10)$$

Cela revient bien à faire un filtrage passe-bas sur le signal et un passe-haut sur le bruit. Le spectre du signal de sortie du modulateur est représenté sur la Figure 1.12. On constate bien que la majorité du bruit a été décalée en dehors de la bande de Nyquist, permettant un meilleur SNR . Le modulateur de la Figure 1.9 est un modulateur du premier ordre et permet d'obtenir des gains de

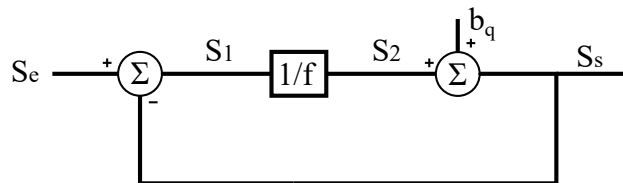


FIGURE 1.11 – Modèle fréquentiel du CAN $\Delta\Sigma$. Le CNA 1 bit est modélisé par l'ajout d'un bruit de quantification b_q .

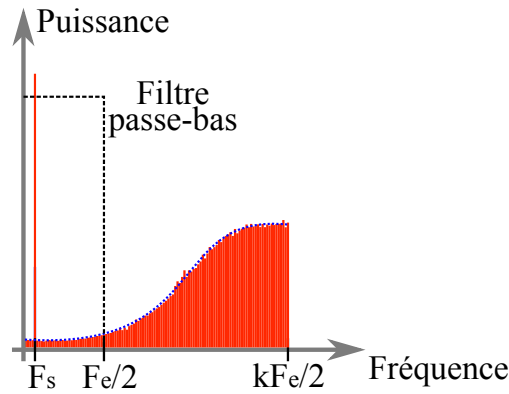


FIGURE 1.12 – Spectre fréquentiel en sortie du modulateur $\Delta\Sigma$.

9 dB par facteur deux de sur-échantillonnage. Il est possible d'augmenter l'ordre du modulateur en ajoutant des intégrateurs afin d'avoir un filtrage du bruit plus sélectif. Ainsi, les gains sont de 15 dB et 21 dB par doublement de la fréquence de sur-échantillonnage pour les ordres deux et trois respectivement. Cela se traduit par des facteurs de sur-échantillonnage de 1024, 64 et 20 pour obtenir des CAN 16 bits avec des modulateurs $\Delta\Sigma$ respectivement d'ordre un, deux et trois.

La sortie du modulateur $\Delta\Sigma$ est une succession de bits à la fréquence de sur-échantillonnage F_{se} . Le filtre passe-bas est utilisé pour réduire la quantité de bruit. Il produit des échantillons à la fréquence F_{se} k fois supérieure à la fréquence d'échantillonnage. Le sous-échantillonnage sert alors à réduire le débit des données la fréquence F_e .

Les CAN $\Delta\Sigma$ sont très utilisés dans les *BASN* car ils offrent une flexibilité en termes de bande passante et de précision. Cela permet de s'adapter à différents types de signaux. Des exemples de CAN $\Delta\Sigma$ utilisés dans le cadre des *BASN* sont présentés dans le Tableau 1.3. On constate que ces CAN ont des fréquences d'échantillonnage plus limitées que les *SAR* et présentent un *FOM* plus élevé. Cette différence peut s'expliquer par une architecture plus simple du *SAR* face au $\Delta\Sigma$. En effet, le *SAR* nécessite seulement un comparateur, un CNA et de la logique. De son côté, le $\Delta\Sigma$ utilise aussi un comparateur mais à cela s'ajoute un, voire plusieurs, intégrateurs, un soustracteur, un filtre numérique et un sous-échantillonneur. De plus, le convertisseur $\Delta\Sigma$ fonctionne en permanence à une fréquence élevée tandis que le *SAR* n'a besoin que d'un nombre réduit d'itérations pour

TABLEAU 1.3 – Exemples de convertisseurs $\Delta\Sigma$ utilisés pour les *BASN*.

	[CC10]	[GRR13]	[MMCT09]
Signal	biosignal	biosignal	EEG
Technologie <i>CMOS</i> (μm)	0.5	0.15	0.5
Puissance (μW)	20	96	76
F_e (Hz)	128-2.048k	8k	500
Résolution (Bits)	8-12	12	10
ENOB	9.5	9.7	8.8
FOM (pJ)	54	14.7	341

fournir un code de sortie. Cela lui permet de travailler avec des fréquences élevées le temps d'une conversion puis de couper l'alimentation pour économiser de l'énergie lors des phases d'inactivité.

La section suivante présente une méthode d'acquisition récemment développée prenant en compte la structure du signal. Cette méthode peut permettre de réduire la fréquence d'acquisition des signaux biologique afin d'améliorer l'efficacité énergétique des capteurs et donc leur autonomie.

1.2.2 Acquisition compressée : échantillonnage compressif

1.2.2.1 Notion de parcimonie des signaux

L'échantillonnage compressif (*Compressed Sensing, CS*) s'appuie sur la parcimonie des signaux. Un signal est parcimonieux lorsque sa représentation dans une base est composée majoritairement de zéros. Le nombre K d'éléments non nuls dans cette base est le degré de parcimonie du signal. On parle alors d'un signal K -parcimonieux. La base permettant une telle représentation est appelée base parcimonieuse. Il existe de nombreuses bases dans lesquelles un signal peut être parcimonieux. Les plus courantes sont les bases temporelle, fréquentielle, d'ondelettes discrètes ou de cosinus discret.

Un signal $x(t)$, 5-parcimonieux en fréquence, est illustré en Figure 1.13a. Ce dernier est une somme de cinq sinusoïdes à des fréquences différentes. L'observation de la représentation temporelle ne permet pas de voir que le signal est parcimonieux. La transformée de Fourier $X(f)$ du signal de la Figure 1.13b permet de constater que sa représentation fréquentielle ne contient que cinq coefficients non nuls. Le signal est donc bien 5-parcimonieux. D'une manière générale, il est possible d'approcher n'importe quel signal par sa meilleure représentation K -parcimonieuse. Pour ce faire, seul les K plus grands coefficients du signal sont conservés, les autres sont forcés à zéro. La meilleure

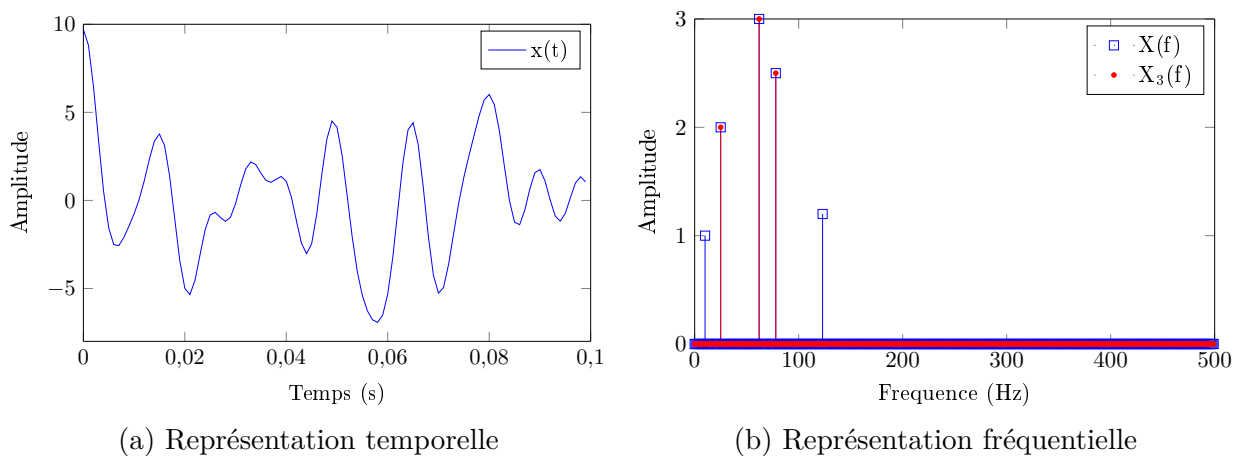


FIGURE 1.13 – Exemple d'un signal 5-parcimonieux en fréquence.

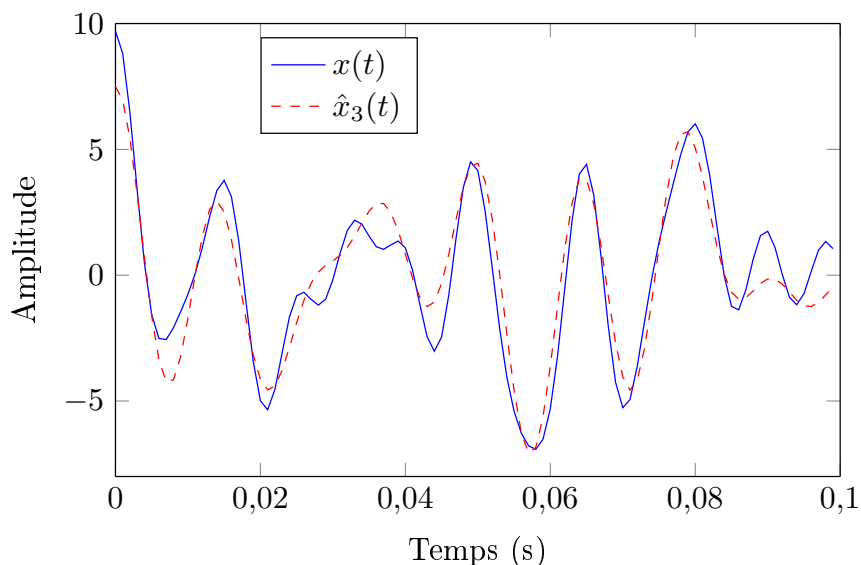


FIGURE 1.14 – Comparaison temporelle du signal d’origine et sa meilleure représentation 3-parcimonieuse.

représentation 3-parcimonieuse de x en fréquence, notée $X_3(f)$, est illustrée en Figure 1.13b.

Ce type d’approximation des signaux est utilisé afin de pouvoir utiliser les méthodes d’échantillonnage compressif sur des signaux réels. En effet, les signaux réels ne sont, en pratique, jamais parfaitement parcimonieux du fait qu’ils sont bruités. On utilise alors une approximation parcimonieuse de ces derniers. L’approximation parcimonieuse se traduit par une déformation du signal dans la base d’origine. La Figure 1.14 montre le signal d’origine $x(t)$ et la représentation temporelle $\hat{x}_3(t)$ de $X_3(f)$. On constate bien l’apparition d’une erreur. Il y a ainsi un compromis à faire entre degré de parcimonie de l’approximation et fidélité au signal d’origine. Plus le degré de parcimonie diminue, plus l’erreur augmente.

1.2.2.2 Principe de l’échantillonnage compressif

L’échantillonnage compressif a été présenté pour la première fois en 2006 par Donoho et Candes [CRT06][CT06][Don06]. C’est une nouvelle technique d’acquisition qui vise à réduire le nombre d’échantillons à prendre afin de représenter un signal. Elle permet, grâce à une connaissance *a priori* des propriétés d’un signal, d’acquérir ce dernier à des fréquences d’échantillonnages inférieures à celle prescrite par le théorème de Shannon-Nyquist. Le *CS* s’appuie sur la connaissance de la parcimonie des signaux à acquérir.

Soit \mathbf{x} un signal K -parcimonieux composé de N éléments. D’un point de vue mathématique, la représentation temporelle de \mathbf{x} peut s’exprimer comme :

$$\mathbf{x} = \Psi \cdot \boldsymbol{\alpha} \quad (1.11)$$

avec α la représentation de \mathbf{x} dans la base parcimonieuse et Ψ la matrice de changement de base. Le vecteur α est lui aussi de taille N mais avec seulement K coefficients non nuls. La matrice de changement de base Ψ est de taille $N \times N$. Avec le CS, le signal à acquérir n'est plus échantillonné de manière uniforme dans le temps comme cela est fait classiquement par un CAN. Au lieu de cela, $M \ll N$ mesures linéaires y_i sur le signal sont effectuées. Ces mesures peuvent s'exprimer sous la forme

$$y_i = \phi_i^T \cdot \mathbf{x} \quad \forall i = 1, 2, \dots, M, \quad (1.12)$$

avec ϕ_i le vecteur de mesure i et T l'opérateur de transposition matricielle. L'acquisition des M mesures peut s'écrire sous forme matricielle comme :

$$\mathbf{y} = \Phi \cdot \mathbf{x} = \Phi \cdot \Psi \cdot \alpha \quad (1.13)$$

avec \mathbf{y} le vecteur des M mesures et Φ la matrice de mesure de taille $M \times N$ composée des vecteurs ϕ_i^T comme ligne. Afin de traiter le signal par la suite, il est nécessaire de pouvoir reconstruire le signal \mathbf{x} à partir des mesures \mathbf{y} . C'est le rôle de l'algorithme de reconstruction qui sera présenté dans la Section 1.2.2.4.

Pour que la reconstruction du signal soit possible, il y a deux conditions à respecter : le signal doit être K -parcimonieux et la matrice $\Phi \cdot \Psi$ doit respecter la propriété d'isométrie restreinte (*restricted isometry property, RIP*) [CRT06][CT06] :

$$(1 - \delta_K) \cdot \|\mathbf{p}_k\|_2^2 \leq \|\Phi \cdot \Psi \cdot \mathbf{p}_k\|_2^2 \leq (1 + \delta_K) \cdot \|\mathbf{p}_k\|_2^2 \quad (1.14)$$

pour tous vecteurs K -parcimonieux \mathbf{p}_k . δ_K s'appelle la constante d'isométrie de la matrice $\Phi \cdot \Psi$ et est compris en zéro et un. La notation $\|\cdot\|_2$ correspond à la norme deux. Cette propriété devant être valide pour n'importe quel vecteur K -parcimonieux, il est impossible de la vérifier en pratique. À la place, une mesure de cohérence entre les matrices Φ et Ψ est utilisée. Elle s'exprime comme [CW08] :

$$\mu(\Phi, \Psi) = \sqrt{N} \cdot \max_{1 \leq j, k \leq n} |\langle \phi_k, \psi_j \rangle| \quad (1.15)$$

où $\langle \cdot, \cdot \rangle$ est l'opérateur de produit scalaire. La cohérence quantifie la similitude entre les deux matrices. Elle mesure la plus grande corrélation existante entre les lignes des deux matrices. Elle doit être la plus proche possible de sa borne inférieure valant un.

Bien que les conditions pour utiliser le CS soient complexes à vérifier, il existe des familles de matrices qui ont une très grande chance de respecter la condition *RIP* si :

$$M \geq C \cdot K \cdot \log(N/K). \quad (1.16)$$

avec C une constante empirique en fonction de la famille de matrice [CW08]. Parmi ces familles de matrice on trouve les matrices aléatoires à entrées indépendantes et identiquement distribuées (iid) selon une loi normale de moyenne nulle et variance $1/M$. Les matrices iid selon une distribution de

Bernoulli symétrique entre $1/\sqrt{M}$ et $-1/\sqrt{M}$ avec une probabilité de $1/2$ sont aussi fonctionnelles. Les métriques utilisées pour caractériser le CS sont les suivantes :

— Le taux de compression TC défini par :

$$TC = \frac{b_x - b_y}{b_x} \cdot 100 \quad (1.17)$$

avec b_x et b_y les nombres de bits utilisés pour respectivement représenter le signal d'origine et le signal compressé.

— Le *percentage root-mean-square difference* (PRD) permet d'évaluer la différence entre le signal reconstruit $\tilde{\mathbf{x}}$ et sa version originale \mathbf{x} et s'écrit

$$PRD = \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2} \cdot 100 \quad (1.18)$$

— Le rapport signal sur bruit SNR permet une mesure en dB de la qualité de reconstruction du signal et s'exprime :

$$SNR = -20 \cdot \log_{10}(0,01 \cdot PRD) \quad (1.19)$$

Le CS est un principe très intéressant pour les *BASN* puisqu'il permet de réduire la quantité de données à transmettre du capteur vers l'agrégateur. La mise en pratique des techniques de CS peut être faite de différentes manières au niveau du capteur. Elle peut être faite de manière totalement numérique suite à une acquisition classique du signal. La manière la plus intéressante est son intégration dans le CAN afin de faire une acquisition directe des échantillons. Cette méthode offre de grandes perspectives puisque qu'elle n'ajoute pas de traitement numérique pour le CS et réduit le nombre d'échantillons transmis. Les systèmes visant à acquérir des échantillons compressés via l'ajout d'un codeur CS sont appelés des convertisseurs "analogique vers information" *A2I* (*Analog-to-information*). La partie suivante présente des implantations existantes du CS dans une optique de *BASN*.

1.2.2.3 Architectures matérielles pour les signaux biologiques

La mise en place du CS est réalisée via la conception d'un convertisseur *A2I*. Ce dernier est composé de deux éléments : un CAN pour transformer la grandeur analogique en signal numérique et un codeur CS qui effectue l'opération de création des échantillons compressés. L'objectif de ce codeur est de réaliser la multiplication par la matrice Φ . La Figure 1.15 présente les trois principales architectures pour réaliser un convertisseur *A2I*.

Le démodulateur aléatoire de la Figure 1.15a est la première implantation à avoir été mise en place [LKD⁺07]. La multiplication avec la matrice est réalisée avec un mélangeur et un intégrateur dans le domaine analogique. Un échantillonneur permet alors de sélectionner les instant d'échantillonnage du CAN. Une telle architecture à une branche limite le type de matrices pouvant être utilisées. En effet, il n'est pas possible de réaliser plusieurs mélanges simultanément. Ainsi, chaque

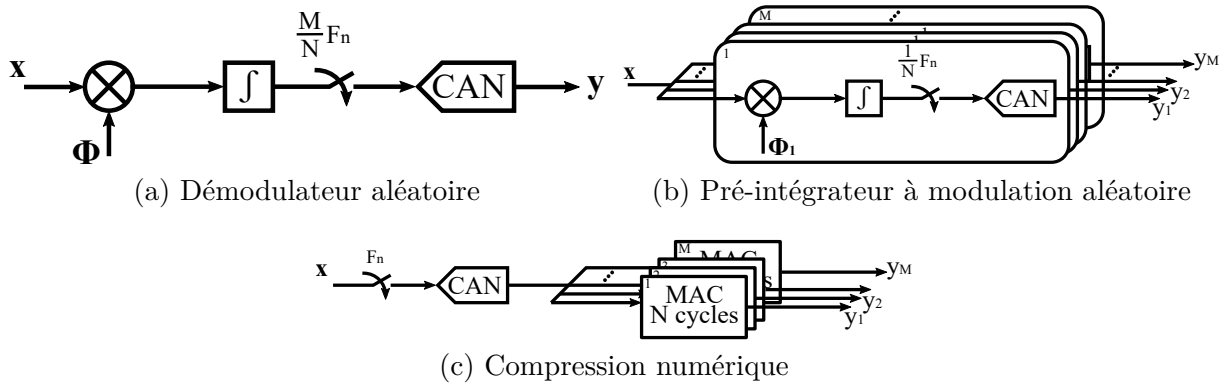


FIGURE 1.15 – Architectures pour la réalisation de convertisseur $A2I$.

ligne de la matrice Φ est composée d'un bloc de N/M éléments non nuls, le reste étant nul. Les blocs non nuls de chaque ligne doivent être distincts pour assurer le fonctionnement avec une seule branche.

Cette limitation est levée avec l'architecture du pré-intégrateur à modulation aléatoire de la Figure 1.15b. Elle utilise plusieurs branches d'acquisition, chacune effectuant les opérations d'une ligne de la matrice Φ . Cela permet d'avoir des mesures utilisant les mêmes échantillons d'entrée, ce qui n'est pas possible avec le démodulateur aléatoire.

Enfin, la compression de type CS de la Figure 1.15c consiste à faire une acquisition classique du signal. Le signal passe ensuite dans M branches d'unité de multiplication accumulation (MAC) afin de générer les échantillons compressés.

Le démodulateur aléatoire proposé dans [LKD⁺07] permet l'utilisation de matrice Φ de type Bernoulli. L'architecture est simulée pour une technologie $CMOS$ $0,13 \mu m$. Elle permet l'acquisition d'un signal synthétique 3-parcimonieux en fréquence avec une fréquence d'échantillonnage six fois inférieure à celle de Nyquist. Les mesures [RLN⁺08] sur le même système montrent des SNR de reconstruction de 62, 59 et 56 dB pour des fréquences d'échantillonnage de respectivement $1/2$, $1/4$ et $1/8$ de la fréquence de Nyquist.

Le pré-intégrateur à modulation aléatoire est implanté dans [GAD⁺14] pour une application sur des signaux ECG. Il est constitué de 64 branches avec chacune un CAN SAR dix bits. Les matrices pouvant être générées sont de type Bernoulli ou uniforme. Le circuit est réalisé avec un process $CMOS$ $0,13 \mu m$ et mesure $2 \times 3 \text{ mm}^2$. La consommation totale des 64 branches est de $1,8 \mu W$ pour une bande passante de 1 kHz et un ENOB de 6.5 bits. Cela revient à un FOM de 9,9 pJ/conversion. Les taux de compression atteints sont de 95% pour des signaux synthétiques 2-parcimonieux. Pour des signaux ECG avec comme base parcimonieuse des ondelettes ou le temps, les taux de compression sont respectivement de 83% et 93,7%. La consommation est de l'ordre de grandeur de celle des architectures SAR existantes rendant la méthode extrêmement intéressante pour les $BASN$.

La compression par CS est mise en place dans [CCS12] avec une technologie $CMOS$ 90 nm. La matrice de mesure utilisée est une matrice dont les éléments sont générés uniformément dans $\{-1, 1\}$. Le codeur CS consomme $1,9 \mu W$ pour une alimentation de 0,6 V et une surface de $200 \times 450 \mu m$. Le taux de compression obtenu atteint 90% pour des signaux EEG.

La compression CS est aussi utilisée dans [MKAV11] sur la plate-forme Shimmer [BGM⁺10] dédiée à la recherche biomédicale. La plate-forme permet l'acquisition, le traitement et la transmission des échantillons. L'application mise en place permet l'utilisation de matrices aléatoires iid uniformes sur $\{-1, 1\}$ ou gaussiennes. Pour un objectif de PRD entre 9 et 2%, les taux de compression obtenus sont de 71% pour un prolongement de la durée de vie de la plate-forme de 9,7%. Ce faible gain d'autonomie est dû au fait que 84 % de l'énergie de la plate-forme est utilisée pour l'échantillonnage et le système d'exploitation.

L'échantillonnage compressif permet de réduire la quantité de données à transmettre tout en n'augmentant pas sensiblement la consommation énergétique du capteur. Il semble donc qu'il est un excellent candidat pour une application sur des bio-signaux qui sont souvent parcimonieux. Si l'acquisition compressée est un bon moyen de réduire la consommation du capteur, il faut aussi prendre en compte la reconstruction du signal sur l'agrégateur. La partie suivante explique brièvement le principe des principaux algorithmes de reconstruction du signal. Une discussion synthétique aborde par la suite la complexité de ces algorithmes ainsi que les conséquences pour les $BASN$.

1.2.2.4 Algorithmes de reconstruction

Comme montré dans la Partie 1.2.2.2, le CS effectue une transformation sur le signal $\mathbf{x} \in \mathbb{R}^N$ via l'opération représentée par la matrice $\Phi \in \mathbb{R}^{M \times N}$. De plus, on a $\mathbf{x} = \Psi \cdot \alpha$, avec α un signal K -parcimonieux de \mathbb{R}^N . Il est possible d'écrire :

$$\mathbf{y} = \mathbf{A} \cdot \alpha = \Phi \cdot \Psi \cdot \alpha. \quad (1.20)$$

L'algorithme de reconstruction a pour objectif de retrouver α parcimonieux à partir de \mathbf{y} et \mathbf{A} . Ce problème est un système linéaire sous-déterminé de M équations à N inconnues avec $M \ll N$. Le problème à résoudre est le suivant :

$$\tilde{\alpha} = \arg \min_{\hat{\alpha} \in \mathbb{R}^N} \|\hat{\alpha}\|_0 \quad \text{sous la contrainte} \quad \mathbf{A} \cdot \hat{\alpha} = \mathbf{y} \quad (1.21)$$

où $\|\cdot\|_0$ représente la norme ℓ_0 qui correspond au nombre d'éléments non nuls du vecteur. Si la matrice \mathbf{A} respecte la contrainte RIP alors il est montré dans [CT06] que la solution au problème (1.21) est atteinte pour $\tilde{\alpha} = \alpha$. Ce problème est NP-complet et ne peut pas être résolu en pratique. Les nombreuses recherches concernant le CS ont cependant permis de simplifier le problème en remplaçant la norme ℓ_0 par la norme ℓ_1 . Il est démontré dans [CRT06] que si le nombre

de mesures M est suffisamment grand alors la solution du problème

$$\tilde{\alpha} = \arg \min_{\hat{\alpha} \in \mathbb{R}^N} \|\hat{\alpha}\|_1 \quad \text{sous la contrainte} \quad \mathbf{A} \cdot \hat{\alpha} = \mathbf{y} \quad (1.22)$$

est $\tilde{\alpha} = \alpha$ avec une très forte probabilité.

À partir de cette définition du problème, il est possible de définir de nombreux algorithmes permettant de le résoudre. Ces algorithmes se décomposent en deux principales familles. Les algorithmes d'optimisation convexe reformulent le problème sous forme d'un problème convexe qui est ensuite résolu. Ces algorithmes permettent d'obtenir une solution exacte mais sont complexes d'un point de vue calcul. Les deux principaux algorithmes d'optimisation convexe utilisés pour reconstruire le signal sont la *Basic Pursuit (BP)* et la *Basic pursuit de-noising (BPDN)*. Ils sont présentés dans [CDS01] et s'expriment sous les formes :

$$\text{BP :} \quad \tilde{\alpha} = \arg \min_{\hat{\alpha} \in \mathbb{R}^N} \|\mathbf{y} - \Phi \cdot \hat{\alpha}\|_1 \quad (1.23)$$

$$\text{BPDN :} \quad \tilde{\alpha} = \arg \min_{\hat{\alpha} \in \mathbb{R}^N} \|\mathbf{y} - \Phi \cdot \hat{\alpha}\|_2^2 + \lambda \cdot \|\hat{\alpha}\|. \quad (1.24)$$

La *BP* cherche une solution ayant la plus petite norme ℓ_1 . La *BPDN* ajoute un paramètre de régularisation afin de régler l'équilibre entre la recherche d'une solution parcimonieuse et la minimisation du bruit. Ces deux algorithmes ont une complexité de calcul de l'ordre de $\mathcal{O}(N^3)$ avec N la taille du signal de départ.

La deuxième famille est la famille des algorithmes itératifs. Ces algorithmes gloutons cherchent un extremum local à chaque itération en espérant que la solution finale sera optimale. Cela permet une réduction de la complexité mais ne garantit pas d'obtenir la meilleure solution. Les 3 algorithmes principaux utilisant cette méthode sont l'*Orthogonal Matching Pursuit (OMP)* [TG07], le *Compressive Sampling Matching Pursuit (CoSaMP)* [NT09] et le *Normalized Iterative Hard Thresholding (NIHT)* [BD10]. Les complexités de ces algorithmes sont de l'ordre de $\mathcal{O}(K \cdot M \cdot N)$ pour l'*OMP* et $\mathcal{O}(M \cdot N)$ pour le *CoSaMP* et le *NIHT*.

Les différents algorithmes sont évalués pour des applications sur des signaux ECG et EMG dans [DAGA12]. Les taux de compression utilisés sont de 75% et différents degrés de parcimonie sont utilisés. Les temps d'exécution sur une machine exécutant *Matlab*[®] ainsi que les SNR de reconstruction obtenus pour différents niveaux de parcimonie sont présentés. Une conclusion de cette étude est que l'algorithme le plus performant est variable selon l'application. De plus, les temps d'exécution sont dans la majorité des cas supérieurs à la seconde. Avec les machines modernes, cela équivaut à un nombre de cycles de calcul sur processeur généraliste de l'ordre du milliard. Cela représente un coût trop important pour espérer une intégration basse consommation. D'une manière générale, l'*OMP* est l'algorithme qui s'exécute le plus rapidement.

Ces algorithmes de reconstruction du signal sont donc complexes et ne sont pas bien adaptés pour une implantation matérielle. La majorité des implémentations est faite de manière logicielle.

TABLEAU 1.4 – Implantations matérielles pour la reconstruction de signal.

	[CTH16]	[JH16]	[RM16]
Application	ECG	-	ECG/EMG/EEG
Technologie <i>CMOS</i> (<i>nm</i>)	90	90	40
Puissance (<i>mW</i>)	11.7	82.5 (simu post-layout)	8.6-78
(<i>M</i> , <i>N</i>)	(64, 128)	(64, 256)	(512, 1024)*
Parcimonie <i>K</i>	40, 44, 48	12	jusqu'à 192
Débit (k échantillons/s)	228	4204	12-237
Efficacité énergétique (nJ/échantillon)	51	20	32-389

* Valeurs maximales pour *M*, nombre de mesures, et *N*, nombre initial d'échantillons.

Il existe quelques implantations matérielles pour l'*OMP* dont les principales caractéristiques sont présentées dans le Tableau 1.4. La consommation de l'ordre du nJ est très supérieure à celle de l'acquisition de l'ordre du pJ. De plus, les implantations de [CTH16] et [JH16] ne permettent de résoudre que des problèmes de tailles fixe pour *N* et *M*.

Si les techniques d'échantillonnage compressif permettent de réduire la quantité de données à acquérir et à transmettre, elles entraînent un surcoût lors de la réception des échantillons. La mise en place d'une reconstruction de signal robuste nécessite la connaissance de la base parcimonieuse du signal ainsi que la mise en place d'algorithmes complexes. Son utilisation dans les *BASN* peut être intéressante si la puissance de calcul de l'agrégateur n'est pas limitante et que les restrictions énergétiques le permettent. Les gains en énergie pour le capteur sont dépendants de la technologie utilisée pour transmettre les données. La partie suivante présente brièvement les bases des transmissions numériques ainsi que les protocoles existants permettant la mise en place de *BASN*.

1.3 Transmission de l'information

1.3.1 Caractéristiques de la transmission pour les réseaux de capteurs corporels

L'ensemble des capteurs et l'agrégateur du *BASN* forment un réseau de communication. Les contraintes d'un tel réseau pour des applications médicales sont nombreuses. Il est indispensable que les données médicales soient transmises et reçues sans erreurs et dans un temps réduit. De plus, la question de la sécurisation de ces données est très importante afin que les données ne puissent pas être interceptées. Le réseau étant localisé autour de la personne, l'interopérabilité avec d'autres *BASN* ou avec les infrastructures environnantes doit être possible. Enfin, la technique utilisée pour transmettre l'information doit être adaptée à l'environnement du réseau. La personne étant au cœur du réseau, ce réseau est mobile et les conditions environnementales sont susceptibles de beaucoup varier.

Afin de répondre à ces contraintes, le modèle OSI (Figure 1.16) permet de décomposer la communication en couches ayant chacune un objectif propre. Les couches quatre à sept sont dites

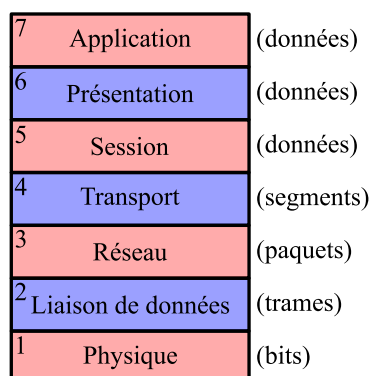


FIGURE 1.16 – Décomposition des systèmes de communication selon le modèle OSI en sept couches.

hautes et permettent de déterminer les données à transmettre. Elles s’occupent notamment de mettre en forme les données sous un format spécifique et de les chiffrer. Elles gèrent aussi la vérification des erreurs dans les messages transmis ainsi que l’optimisation des ressources réseaux. Ces couches sont généralement implémentées de façon logicielle.

Les couches basses (une à trois) sont généralement réalisées de manière matérielle. La couche réseau est chargée de déterminer le parcours d’un paquet dans le réseau afin d’acheminer ce dernier d’un point A à un point B. La couche liaison de données permet de gérer la communication directe entre deux éléments communicants. C’est elle qui gère le format des trames envoyées pour établir une connexion entre deux machines. Elle permet aussi de définir les règles d’accession au canal de communication. Enfin, la couche physique est chargée de la transmission physique des bits sous forme numérique ou analogique. On y trouve les spécifications de l’interface RF, dans le cadre des transmissions sans fils, ainsi que la modulation des données.

Dans le cadre de cette étude axée sur la consommation énergétique, seules les couches une et deux sont considérées. Ce sont en effet les couches qui vont déterminer le coût énergétique de la transmission d’un bit. Les couches supérieures auront comme influence de modifier la quantité de données transmises. La métrique qui sera utilisée pour comparer les différentes méthodes de transmission est l’énergie nécessaire pour transmettre un bit.

1.3.2 Les standards existants

La mise en place de standards a pour but d’uniformiser les règles de conception des éléments communicants. Cela permet d’assurer que des capteurs produits par des entreprises différentes pourront communiquer entre eux. Les standards définissent un ensemble de critères qu’un produit doit respecter afin d’assurer sa compatibilité avec d’autres produits respectant le même standard.

Dans le cadre des *BASN*, il y a trois principaux standards : *Bluetooth Low Energy (BLE)*, *IEEE 802.15.4* et *IEEE 802.15.6*. Chacun d’eux permet une communication sans fils via la transmission radio-fréquence. Le *BLE* est apparu avec la version 4.0 du *Bluetooth Core Specification* de 2010.

Cette spécification en est à la version 5.0 depuis décembre 2016 [BLE]. Le *BLE* est une adaptation du *Bluetooth* classique pour s'adapter aux unités très faible puissance ayant peu de données à transmettre. Il permet la communication sur des distances inférieures à 50 mètres pour un débit de 1 Mbit/s. La distance maximale est atteinte pour une communication en espace libre sans obstacles. Son fonctionnement en mode maître/esclave limite son utilisation à une topologie en étoile. Dans une topologie en étoile (Figure 1.17), chaque nœud esclave ne peut communiquer directement qu'avec le maître. Avec le *BLE*, le nombre d'esclaves maximum pour un maître n'est pas spécifié. Au niveau physique, le *BLE* utilise la bande de fréquence libre ISM (Industriel, Scientifique, Médical) à 2,4 GHz. La bande est divisée en 40 canaux physiques séparés de 2 MHz et la technique d'étalement par saut de fréquences est utilisée.

La norme *IEEE 802.15.4* [IEE16] spécifie les couches une et deux pour les *Low Rate Wireless Personal Area network*. L'objectif est d'offrir un cadre permettant la mise en place simple de réseaux à faible portée pour des unités ayant peu de données à communiquer. La portée de ces réseaux est de l'ordre de 10 m. Les fréquences de transmission sont réparties sur différentes bandes comme illustré dans le Tableau 1.5. L'avantage du standard *IEEE 802.15.4* est de permettre une topologie en étoile ainsi qu'une topologie *mesh*. Dans une topologie *mesh*, chaque élément peut être connecté à un autre élément. Un nœud spécifique du réseau sert de coordinateur afin de gérer l'ajout ou la suppression de nœuds. Pour les *BASN*, c'est l'agrégateur qui sert de coordinateur permettant d'avoir un réseau qui s'organise de manière autonome. C'est sur cette norme que s'appuie le *ZigBee*[Zig].

Le standard *IEEE 802.15.6* [IEE12] est le standard spécialisé dans les communications destinées aux réseaux corporels. Il permet des communications ultra basse consommation pour des portées très faibles ne dépassant pas les deux mètres. Pour cela, trois couches physiques sont spécifiées :

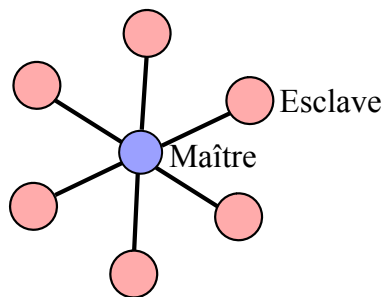


FIGURE 1.17 – Topologie de réseau en étoile. Les seules connections possibles sont entre le maître et les esclaves.

TABLEAU 1.5 – Répartition des canaux de transmission pour la norme *IEEE 80.15.4*.

Bande	Région	Nombre de canaux	Débit
868 MHz	Europe	1	20 kb/s
915 MHz	Amérique	10	40 kb/s
2.4 GHz	Monde	16	250 kb/s

la bande étroite, l'ultra-large bande et l'utilisation du corps comme canal de communication. La communication à bande étroite peut se faire sur sept bandes de fréquences différentes comprenant la bande à 2.4 GHz. Une des bandes intéressantes est celle située entre 402 et 405 MHz. Elle est en effet dédiée à la communication à l'intérieur du corps car moins atténuée par ce dernier. Les débits atteignables varient de 57,5 à 971,4 kb/s en fonction de la modulation.

Les communications ultra-large bande sont la deuxième manière de transmettre l'information. Elles se font sur les bandes de fréquences 3,2-4,7 GHz et 6,2-10,3 GHz avec des largeurs de canal de 499 MHz. La communication ultra-large bande permet la communication via la transmission de signaux très brefs, permettant ainsi de laisser l'étage RF désactivé une majorité du temps. Les débits spécifiés varient de 250 kb/s à 15,6 Mb/s [IEE12].

Enfin la dernière méthode de transmission repose sur la communication par champ électrique. Elle utilise les propriétés diélectriques du corps pour s'en servir comme canal de communication, d'où son nom *Human Body Communication (HBC)*. Pour faire cela les fréquences utilisées sont situées autour de 21 MHz, c'est à ces fréquences que les couches supérieures du corps se comportent de manière la plus proche d'un conducteur. Les débits fluctuent entre 164 et 1 312 kb/s [IEE12].

Les différents standards permettent de placer un cadre pour la conception de capteurs pour les *BASN* afin d'assurer une compatibilité de ces derniers entre eux. Cette nécessité de compatibilité impose des contraintes lors de la communication pouvant conduire à une perte de l'efficacité énergétique. Les différents axes de recherche concernant la communication pour les *BASN* tentent de proposer de nouvelles méthodes de transmission et d'accès au canal tout en restant compatible avec les standards. Afin de déterminer le standard le plus intéressant, il faut comparer ces derniers d'un point de vue énergétique. La partie suivante présente des implantations existantes pour transmettre l'information au cœur des *BASN*.

1.3.3 Implantations des transmetteurs

Les implantations présentées dans cette partie se focalisent sur le coût de transmission d'un bit au niveau du capteur. Ce choix s'explique par le fait que le capteur est un élément difficile à recharger, il convient donc de maximiser les efforts pour améliorer l'efficacité énergétique de ce dernier.

Le Tableau 1.6 présente les performances de différents émetteurs existants. Certains comme [WDD⁺13] et [ZJW⁺13] sont configurables afin de fonctionner selon plusieurs standards. On constate que malgré l'utilisation de technologies plus énergivores, les solutions utilisant la *HBC* présentent les meilleures efficacités énergétiques. Le principal désavantage de cette méthode de transmission est qu'elle impose que l'agrégateur soit un élément en contact avec la peau.

Le transmetteur proposé dans [BVK⁺15] permet de dépasser les débits du standard *802.15.6* pour atteindre 4,5 Mb/s. Cela lui permet d'avoir une efficacité énergétique de 0,51 nJ/b.

L'efficacité énergétique est obtenue en considérant la transmission au débit maximum. Ces débits sont très fortement supérieurs à ce qui est nécessaire pour une application biomédicale. En

TABLEAU 1.6 – Caractéristiques de transmetteurs pour les standards sans fil.

Référence	Standard	Bande	Technologie CMOS (nm)	Débit	Efficacité (nJ/b)
[KFC ⁺ 17]	<i>BLE</i>	2,4 GHz	28	1 Mb/s	3,7
[OST ⁺ 16]	<i>BLE</i>	2,4 GHz	65	1 Mb/s	5,5
[WDD ⁺ 13]	<i>BLE</i>	2,4 GHz	130	1 Mb/s	14,5
[WDD ⁺ 13]	<i>802.15.6</i>	2,4 GHz	130	971,4 kb/s	12,5
[BVK ⁺ 15]	<i>802.15.6</i>	400 MHz	40	187,5 kb/s	12,1
[ZJW ⁺ 13]	<i>802.15.6</i>	400 MHz	180	187,5 kb/s	70
[ZJW ⁺ 13]	<i>802.15.4</i>	2,4 GHz	180	250 kb/s	72
[WFKA16]	<i>802.15.6 (HBC)</i>	10-60 MHz	<i>FPGA</i>	1,25 Mb/s	3,4
[BSL ⁺ 12]	<i>HBC</i>	40-120 MHz	180	10 Mb/s	0,38

considérant un signal ECG échantillonné à 360 Hz sur 11 bits, le débit de données à transmettre est de 3,96 kb/s. Ainsi, l'efficacité énergétique par bit utile est réduite d'autant plus que le rapport entre débit d'émission et débit utile est grand. Pour améliorer l'efficacité, il est possible d'envisager de stocker les données avant transmission. Cela nécessite alors la mise en place d'une mémoire qui consommera aussi de l'énergie.

Toutes ces implantations permettent donc de constater que l'émission d'un bit utile via l'air ou le corps est coûteux. Comparée aux CAN qui acquièrent un échantillon de dix bits pour un coût de quelques dizaines de pJ, la transmission de ces mêmes dix bits a un coût de l'ordre de la dizaine voire centaine de nJ. Ce rapport d'un ordre de grandeur montre que, pour augmenter l'autonomie d'un capteur, il est plus intéressant de focaliser les efforts sur la diminution du budget RF. La partie suivante décrit le traitement effectué sur les données dans le cadre des *BASN* afin de comparer sa consommation à celle de l'acquisition et de la transmission. Elle ouvre une réflexion sur un usage réduit du frontal RF.

1.4 Traitement de l'information : la classification supervisée

L'objectif principal des *BASN* est de détecter la présence d'une anomalie. Il est aussi possible de déterminer le type de l'anomalie suite à sa détection. Cette étape est, en général, réalisée via une classification. Grâce aux nombreuses bases de données médicales commentées par des médecins spécialisés, il est possible d'utiliser des méthodes supervisées afin de maximiser les performances. Cette section présente d'abord le principe de la classification supervisée. Dans un second temps, les méthodes couramment utilisées dans le cadre des *BASN* sont analysées.

1.4.1 Principes

La classification consiste à regrouper des observations ayant des caractéristiques communes. Les groupes formés sont appelés classes ou labels. L'objectif d'un processus de classification est

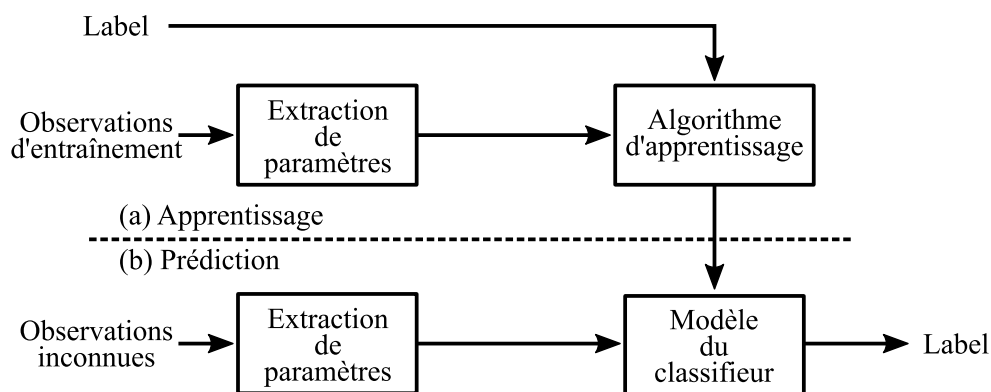


FIGURE 1.18 – Schéma de principe de la classification supervisée. L'apprentissage (a) permet de fixer les paramètres du modèle grâce aux observations d'entraînement. La phase de prédiction (b) utilise le modèle pour déterminer la classe ou le label.

d'assigner une classe à une nouvelle observation inconnue. Le nombre de classes est dépendant de l'application. Lorsqu'il n'y a que deux classes, la classification est dite binaire et peut s'apparenter à de la détection.

La classification se fait généralement en deux phases suivant le schéma de la Figure 1.18. La phase d'apprentissage ou d'entraînement a pour but d'inférer le modèle à partir de données d'entraînement. Pour ce faire, des paramètres sont extraits des observations afin de réduire la dimension du signal. Dans le cas des *BASN*, les observations sont typiquement les signaux temporels issus de l'acquisition par le capteur. Les paramètres extraits sont très dépendants de l'application de classification et du signal et peuvent avoir une très forte influence sur les performances. Pour les signaux biologiques, des exemples de paramètres utilisés incluent l'énergie dans différentes bandes de fréquences, des mesures temporelles entre des événements, des mesures statistiques ou sur une transformation du signal. Il n'est donc pas possible de déterminer la consommation énergétique de l'extraction de paramètres dans un cas général. Pour l'étudier il est nécessaire de fixer un cadre spécifique sur le type de signal considéré ainsi que l'application de classification.

Après l'extraction des paramètres, l'algorithme d'apprentissage les analyse afin de configurer le modèle du classifieur. Cela peut se faire manière supervisée ou non-supervisée. Dans le cas non-supervisé, l'algorithme ne connaît pas les classes, ni leur nombre. Le but de l'algorithme est alors de trouver des similitudes entre les différentes observations pour les regrouper en classes. Avec l'apprentissage supervisé, chaque observation d'entraînement est marquée avec un label. L'algorithme tente alors d'ajuster le modèle afin que la décision de la classe de chaque observation corresponde à celle donnée en entrée. Cela permet de guider l'algorithme qui sait en avance la réponse correcte et peut ainsi modifier les paramètres du modèle pour obtenir la bonne sortie. L'étape d'apprentissage est le plus souvent réalisée hors ligne.

L'étape de prédiction consiste à extraire un vecteur de paramètres d'une nouvelle observation et de déterminer sa classe d'appartenance. Les paramètres extraits sont les mêmes que lors de l'entraînement. Cette étape est aussi appelée étape de généralisation puisqu'elle permet d'évaluer

la capacité du modèle à classifier des observations n'ayant aucun lien avec l'apprentissage.

Afin d'éviter des biais, les observations sont généralement séparées en trois ensembles de données. Un premier ensemble est celui d'apprentissage. Il est utilisé pour faire l'ajustement du modèle. Les algorithmes d'apprentissage étant souvent itératifs, il est nécessaire d'avoir une condition d'arrêt. Cette condition est déterminée par les performances de classification sur un deuxième ensemble de données appelé ensemble de validation. Les ensembles d'apprentissage et de validation sont généralement regroupés en un seul ensemble d'entraînement. Enfin, l'ensemble de test permet d'évaluer le pouvoir de généralisation du classifieur. C'est un jeu de données indépendant de ceux d'apprentissage et de validation. Cela permet de s'assurer que l'évaluation des performances est effectuée sur des données qui n'ont pas servi à entraîner le modèle.

Les performances de classification sont évaluées avec différentes métriques s'appuyant sur les notions de faux positif (FP), faux négatif (FN), vrai positif (VP) et vrai négatif (VN). La matrice de confusion de la Figure 1.19 illustre, pour une classification binaire, les quatre cas possibles de décision. Avec une classification binaire, les deux classes peuvent être vues comme présence ou non d'une pathologie par exemple. La matrice de confusion décrit aussi deux métriques très utilisées pour évaluer un classifieur. La sensibilité évalue la capacité à correctement prédire la présence pathologie lorsque l'observation présente une pathologie. Une prédiction négative avec un classifieur ayant une forte sensibilité permet de valider l'absence de pathologie. La spécificité évalue la capacité à correctement prédire l'absence de pathologie quand l'observation ne présente pas de pathologie. Avec une forte spécificité, une prédiction positive permet de valider la présence de pathologie. Il est donc indispensable d'utiliser ces deux métriques afin de bien évaluer un classifieur. En effet, la seule connaissance de la sensibilité ne permet pas de juger puisqu'un système prédisant tout le temps la présence de pathologie aura une sensibilité de 100 % alors qu'il est évident qu'un tel classifieur est mauvais. Une troisième métrique utile est l'exactitude qui est définie comme :

$$Exactitude = \frac{VP + VN}{VP + VN + FP + FN}. \quad (1.25)$$

Pour les applications médicales devant détecter une anomalie, une autre métrique utilisée est le taux de fausse alarme exprimé en nombre par heure.

Nombre d'observations VP+VN+FP+FN	Prédiction positive	Prédiction négative	
Observation vraie	Vrai Positif VP	Faux Négatif FN	Sensibilité VP/(VP+FN)
Observation fausse	Faux Positif FP	Vrai Négatif VN	Spécificité VN/(VN+FP)

FIGURE 1.19 – Matrice de confusion pour une classification binaire. La sensibilité et spécificité permettent d'évaluer les performances du modèle.

Toutes les métriques présentées dans le cadre d'une classification binaire sont aussi valables pour le cas multi-classes. Dans une classification multi-classes, les métriques sont calculées pour chacune des classes en se ramenant à un cas binaire. Pour se faire, chaque classe est opposée à l'union de toutes les autres classes. Il est aussi possible de faire une moyenne pour chaque métrique en pondérant par le taux d'apparition de chaque classe dans les observations de test.

Dans la suite de cette partie, différents modèles de classifieur sont présentés afin d'évaluer leur complexité d'implantation et leur performances. Pour les classifieurs ayant été implantés dans le cadre des *BASN*, la consommation est donnée afin de comparer les différentes méthodes. L'objectif de cette étude étant d'évaluer la consommation des différents éléments de la chaîne d'acquisition et de traitement des *BASN*, les algorithmes d'apprentissage ne seront pas détaillés. Il n'est, en effet, pas nécessaire d'implanter ces algorithmes puisqu'ils sont réalisés hors-ligne lors de la conception du modèle.

1.4.2 Les algorithmes de classification

1.4.2.1 Classifieur probabiliste

Les premiers classifieurs à avoir fait leur apparition sont les classifieurs probabilistes. Ils s'appuient sur des calculs de probabilités et font des suppositions sur la distribution des données de manière à prédire la classe d'une nouvelle entrée. Considérons un problème de classification simple avec c classes. Soit $\mathbf{x} \in \mathbb{R}^n$ un vecteur de n paramètres réels et $y \in \{1, 2, \dots, c\}$ la classe à laquelle il appartient. Un classifieur probabiliste cherche à calculer la probabilité conditionnelle $P(y = k|\mathbf{x})$ que l'observation appartienne à la classe k . Le choix de la classe est ensuite fait en sélectionnant la valeur de k offrant la plus forte probabilité.

Un premier type de classifieurs probabilistes s'appuie sur la méthode du maximum de vraisemblance. En utilisant la loi de Bayes, la probabilité conditionnelle peut s'écrire :

$$P(y = k|\mathbf{x}) = \frac{P(\mathbf{x}|y = k) \cdot P(y = k)}{P(\mathbf{x})}. \quad (1.26)$$

Afin de pouvoir utiliser ce modèle pour classifier il est nécessaire de connaître la probabilité d'apparition de la classe k ainsi que la densité de probabilité de chaque classe. La probabilité d'avoir l'observation \mathbf{x} n'est pas importante puisque cette dernière est une constante ne dépendant pas de k . Si le jeu de données utilisé pour le test est représentatif de la réalité, $P(y = k)$ peut être obtenue en prenant le taux d'apparition de la classe k dans ce dernier. La décision de la classe se fait alors en résolvant le problème :

$$\hat{y} = \arg \min_y P(y) \cdot P(\mathbf{x}|y) \quad (1.27)$$

Il est ensuite nécessaire de faire un choix sur la distribution de chaque classe. Les classifieurs Bayésien naïfs (NB) font la supposition que les éléments de \mathbf{x} sont indépendants. Il est alors possible

de réécrire (1.27) comme :

$$\hat{y} = \arg \min_y P(y) \cdot \prod_{i=1}^n P(x_i|y) \quad (1.28)$$

Dans le cadre des *BASN* manipulant des données continues, un choix classique de distribution est une gaussienne avec $P(x_i|y) \sim \mathcal{N}(\mu_{y,i}, \sigma_{y,i})$. L'inconvénient des classifieurs Bayésiens naïfs est qu'ils reposent sur une supposition forte d'indépendance des paramètres.

Pour palier ce problème il est possible d'utiliser la classification par analyse discriminante linéaire (*LDA*). Avec cette méthode, les classes sont supposées suivre une distribution gaussienne multivariée de la forme :

$$P(\mathbf{x}|y) \sim \mathcal{N}(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}). \quad (1.29)$$

Il est à noter que la matrice de covariance $\boldsymbol{\Sigma}$ est constante et ne dépend pas de la classe. Cette méthode permet de créer des séparations linéaires dans l'espace des paramètres. Lorsque les observations ne sont pas linéairement séparables, il est nécessaire de créer un modèle plus complexe. Pour ce faire, il est possible d'utiliser l'analyse discriminante quadratique (*QDA*). Le principe est identique sauf que les matrices de covariances des différentes classes ne sont plus identiques et sont notées $\boldsymbol{\Sigma}_y$. L'algorithme d'apprentissage pour ces trois classifieurs a pour but de déterminer les paramètres des différentes distributions à partir du jeu de données d'entraînement.

Il existe un deuxième type de classifieur qui ne s'appuie ni sur la loi de Bayes, ni sur la supposition d'une distribution pour chaque classe. Le classifieur par régression logistique (*LR*) est un classifieur qui utilise une fonction logistique pour prendre une décision. Une fonction logistique s'exprime de la forme

$$h(x) = \frac{1}{1 + e^{-ax}} \quad (1.30)$$

avec a un paramètre du modèle à définir lors de l'apprentissage. Une telle fonction est représentée sur la figure 1.20. Elle a une forme de *S* et est comprise entre zéro et un. Elle permet de transformer un réel en une valeur interprétable comme une probabilité.

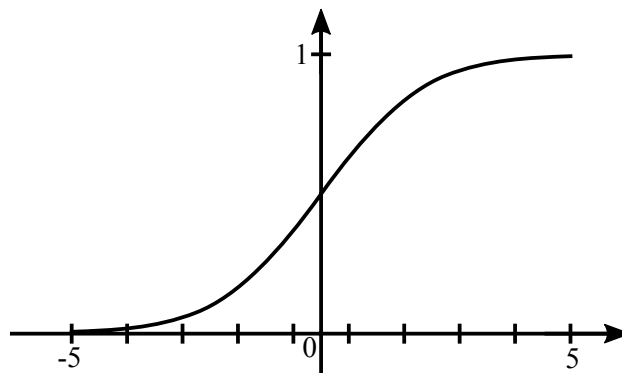


FIGURE 1.20 – Exemple de fonction logistique permettant de transformer un nombre réel en probabilité.

TABLEAU 1.7 – Implantations de classifieurs probabilistes pour les *BASN*.

	[HHC ⁺ 14]	[PSS ⁺ 15] ¹	[CMM ⁺ 13] ¹	[BTS ⁺ 16] ¹
Classification	Crise cardiaque	Épilepsie	ECG anormal	Arythmie
Classifieur	QDA	LR	LDA	NB
Technologie <i>CMOS</i> (<i>nm</i>)	90	65	130	65
Nombre de paramètres	44	110	8	7
Sensibilité	0.99	0.952	0.865	0.842
Spécificité	-	-	0.942	0.87
Énergie par classification	24.7 μJ^2	19 nJ ³	826 nJ ³	2.1 μJ^2

¹ Performances obtenues en simulation post-synthèse et consommation estimée post-layout.

² Énergie estimée à partir la puissance en considérant un rythme de 80 classification/min

³ Consommation dynamique uniquement.

Avec la *LR*, la probabilité d'appartenance à la classe k connaissant l'observation est exprimée comme :

$$P(y = k|\mathbf{x}) = h(\boldsymbol{\theta}_y^T \mathbf{x}) \quad (1.31)$$

où $\boldsymbol{\theta}_y$ est un vecteur de taille n permettant de réaliser une combinaison linéaire des éléments x_i . Avec cette méthode, la classe ayant la plus grande probabilité est choisie. De même qu'avec les classifieurs précédents, l'algorithme d'apprentissage a pour rôle de déterminer les $\boldsymbol{\theta}_y$ à partir des observations d'entraînement.

Les classifieurs du Tableau 1.7 ont été implantés dans le cadre des *BASN*. L'utilisation de techniques calculatoires permet de simplifier les circuits en évitant les calculs de fonctions exponentielles. Cela permet d'obtenir des classifieurs ne demandant que des additions et des multiplications.

Les temps de traitement pour [HHC⁺14] et [BTS⁺16] n'étant pas précisés, l'énergie par classification est estimée à partir de la puissance et en considérant un rythme de 80 classifications par minute. De plus, cette énergie par classification inclut l'énergie pour extraire les paramètres, ce qui n'est pas le cas pour [PSS⁺15] et [CMM⁺13]. Ces derniers nécessitent peu d'énergie pour effectuer une classification. Dans le cas de [PSS⁺15] la classification équivaut à la transmission d'une dizaine de bits.

1.4.2.2 Arbres de décision

Les arbres de décision sont une famille de classifieurs qui prédisent une classe par descente d'un arbre. Comme illustré en Figure 1.21, un arbre est composé de nœuds et de feuilles. Chaque nœud compare la valeur d'un paramètre d'entrée avec un ou deux seuils déterminés lors de l'apprentissage. Chaque nœud possède deux sorties qui indiquent le nœud suivant en fonction du résultat de la comparaison. Les nœuds finaux sont appelés feuilles et sont associés à une classe. La prise de décision d'un arbre est réalisée en descendant ce dernier jusqu'à atteindre une feuille. La profondeur d'un arbre est définie comme le nombre maximum de nœuds à traverser avant d'arriver sur une

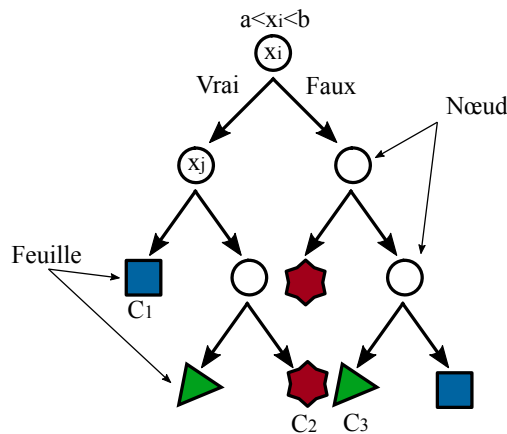


FIGURE 1.21 – Schéma d’un arbre de profondeur trois avec trois classes. Chaque nœud réalise deux comparaisons et en fonction du résultat pointe vers un autre nœud ou vers une feuille.

feuille. L’arbre présenté sur la Figure 1.21 a une profondeur de trois.

L’utilisation d’un unique arbre comme classifieur peut conduire à des problèmes comme le surapprentissage (*overfitting*) ou la création d’arbres trop profonds. Le surapprentissage apparaît lorsque le classifieur s’appuie trop sur les données d’entraînement lors de l’apprentissage. Cela signifie que le classifieur, en plus de mémoriser les spécificités des classes, mémorise le bruit qui est spécifique aux données d’entraînement. Cela se traduit par une chute des performances de classification lors du test du classifieur. Pour limiter cet effet, il est possible d’utiliser des méthodes d’ensemble dont le but est de créer plusieurs arbres. Chaque arbre est entraîné avec un sous-ensemble des données d’entraînement afin de répartir les effets du surapprentissage sur les différents arbres. La décision finale est ensuite réalisée en effectuant un vote majoritaire des décisions des arbres.

Parmi ces méthodes on trouve l’ensachage (*bagging*) et le *boosting*. En considérant un jeu de données d’apprentissage de N observations, l’ensachage consiste à faire un tirage aléatoire uniforme avec remplacement de N éléments parmi l’ensemble de départ. Avec un tel tirage, en moyenne, 37 % des éléments ne sont pas sélectionnés. En faisant un nouveau tirage pour chaque nouvel arbre, il est possible de créer un ensemble d’arbres pour réaliser la classification. La méthode du *boosting* est similaire à l’ensachage mais garde en mémoire les performances des arbres déjà construits. Au lieu de faire un tirage uniforme, le *boosting* favorise les observations n’ayant pas été correctement classifiées par les arbres existants. La décision finale est ensuite réalisée par vote majoritaire en pondérant la prédiction des arbres par leur exactitude de classification sur les données de validation.

Les forêts aléatoires sont une autre méthode d’ensemble. Elles s’appuient sur l’ensachage mais limitent aussi le choix lors de la création d’un nœud. La recherche du paramètre à utiliser pour un nœud n’est pas effectuée sur l’ensemble des paramètres. Elle est limitée en général à \sqrt{N} paramètres qui sont tirés aléatoirement et uniformément parmi la liste des N paramètres. Un nouveau tirage est effectué pour chaque nœud.

La complexité d'implantation d'un arbre est principalement due à la mémorisation. Il est en effet nécessaire de stocker cinq éléments par nœud : le paramètre sur lequel faire la comparaison, les deux valeurs des seuils et les deux nœuds suivants. D'un point de vue calculs, les arbres sont très simples puisqu'ils ne reposent que sur des comparaisons et ne nécessitent pas de calculs complexes.

Un classifieur utilisant l'algorithme *AdaBoost* est implanté dans [HHL⁺16]. *AdaBoost* est un algorithme reposant sur la méthode du *boosting*. Le circuit proposé détecte les crises d'épilepsie sur un signal EEG à 16 canaux. Le classifieur utilise 16 paramètres et est constitué de 50 arbres d'une profondeur de deux. La sensibilité de la classification est de 93,8 % pour un taux de 0,16 fausse alarme par heure. La consommation évaluée en simulation post-layout avec une technologie *CMOS* 65 nm est de 390 nJ/classification. Elle comprend l'extraction des paramètres et la classification.

Dans [IJRV17], un détecteur des étapes du sommeil est réalisé en utilisant un signal EEG sur un canal. 17 paramètres sont extraits avant d'être classifiés. Le système utilise 15 classifieurs différents composé d'un arbre chacun. En fonction de l'étape de sommeil courante, seul un sous-ensemble de ces classifieurs est activé. La méthode proposée présente une exactitude de 98,72 %. Le circuit fabriqué en technologie *CMOS* 0.18 μm nécessite 700 nJ pour effectuer une classification.

1.4.2.3 Machines à vecteurs de support

Les machines à vecteurs de support (*SVM*) ou séparateurs à vaste marge sont un classifieur binaire. L'objectif est de trouver un hyperplan séparateur des deux classes dans l'espace formé par les paramètres. L'ensemble des points \mathbf{p} formant l'hyperplan suivent l'équation :

$$\langle \mathbf{w}, \mathbf{p} \rangle + b = 0 \quad (1.32)$$

où \mathbf{w} est le vecteur normal à l'hyperplan et $b/\|\mathbf{w}\|$ est l'ordonnée à l'origine selon la direction de \mathbf{w} . Cet hyperplan séparant les données, le signe de $\langle \mathbf{w}, \mathbf{x} \rangle + b$ avec \mathbf{x} une nouvelle observation, sera positif pour les observations d'une classe et négatif pour l'autre.

Si les données sont linéairement séparables, il peut exister une infinité d'hyperplans séparateurs. L'objectif de l'algorithme d'apprentissage des *SVM* est de trouver l'hyperplan offrant la plus grande marge. La Figure 1.22 illustre la notion de marge dans un cas simple de classification avec deux paramètres. On appelle vecteurs de support l'ensemble des vecteurs de chaque classe les plus proches de l'hyperplan. Ces derniers sont entourés en noir sur la Figure 1.22. La distance d entre une observation \mathbf{x} et l'hyperplan est donnée par la formule :

$$d = \frac{|\langle \mathbf{w}, \mathbf{x} \rangle + b|}{\|\mathbf{w}\|}. \quad (1.33)$$

La marge est la distance entre les deux plans portés par les vecteurs de support de chaque classe et parallèle à l'hyperplan séparateur. Il est possible de considérer que ces hyperplans suivent les équation $\langle \mathbf{w}, \mathbf{x} \rangle + b = \pm 1$. Dans ce cas, la marge m s'écrit comme $m = 2/\|\mathbf{w}\|$. L'algorithme d'apprentissage doit donc utiliser l'ensemble des N observations d'apprentissage pour résoudre le

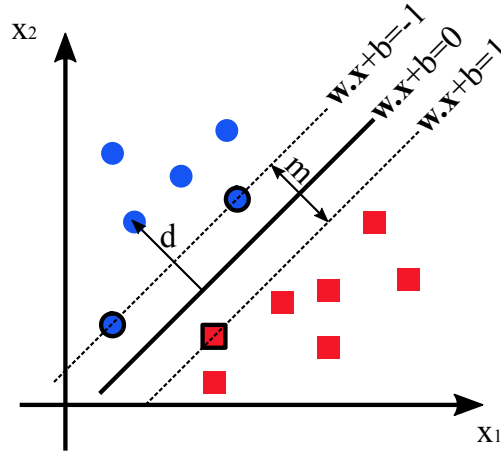


FIGURE 1.22 – Illustration de l’hyperplan séparateur pour des données linéairement séparables en deux dimensions.

problème suivant :

$$(\hat{\mathbf{w}}, \hat{b}) = \arg \min_{\mathbf{w}, b} \|\mathbf{w}\| \quad \text{sous la contrainte} \quad y_i \cdot (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 \quad \forall i \in \{1; N\} \quad (1.34)$$

où $y_i = \pm 1$ en fonction de la classe de l’observation i . Lorsque les données ne pas parfaitement séparables, il est possible de faire intervenir une fonction de régularisation afin que le problème d’optimisation minimise à la fois la norme de \mathbf{w} et le nombre d’observations mal classées.

Si les données ne sont pas linéairement séparables, il est possible d’utiliser une transformation dans un espace de plus grande dimension où elles le sont. La classification et l’apprentissage des *SVM* reposent principalement sur le calcul du produit scalaire entre deux vecteurs. Soit la transformation $\Phi : \mathbf{x} \mapsto \Phi(\mathbf{x})$ permettant le changement de dimension. Le produit scalaire dans ce nouvel espace devient alors :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle. \quad (1.35)$$

la fonction $K(\mathbf{x}_i, \mathbf{x}_j)$ est appelée fonction noyau. Pour apprendre et utiliser les *SVM* il n’est pas nécessaire d’effectuer la transformation, seule la connaissance du noyau est utile. Les noyaux classiquement utilisés sont les suivants :

- linéaire : $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$
- polynomiale de degré p : $K(\mathbf{x}_i, \mathbf{x}_j) = (1 + \langle \mathbf{x}_i, \mathbf{x}_j \rangle)^p$
- gaussien ou fonction à base radiale (*RBF*) : $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$

Les *SVM* sont beaucoup utilisées car elles permettent d’obtenir de bons résultats de classification tout en garantissant une implantation relativement peu complexe. Le Tableau 1.8 présente des implantations de type *Application Specific Integrated Circuit (ASIC)* de *SVM* dans le cadre des *BASN*.

Dans [LV13], l’implantation proposée permet de réaliser des noyaux de type linéaire, polynomial et gaussien. De plus, le système est testé pour des applications de détection d’épilepsie et de clas-

TABLEAU 1.8 – Implantations de *SVM* dans les *BASN*.

	[LV13]	[YYED ⁺ 13]	[ZAY16]	[AY16]
Application	Arythmie	Épilepsie	Épilepsie	Épilepsie
Technologie <i>CMOS</i> (μm)	0.13	0.18	0.18	0.18
Nombre de paramètres	21	8	112	56
Noyau	Gaussien	Linéaire	Linéaire	Gaussien
Énergie par classification	127 μJ	2.03 μJ	1.85 μJ	1.83 μJ
Sensibilité	0.85 ~ 0.97	0.827	0.951	0.951
Spécificité	1	0.955	0.962	0.991

sification d'arythmie. La classification d'arythmie est effectuée de manière spécifique pour chaque patient. C'est ce qui explique la plage de sensibilité dans le tableau. Dans ce même document, une comparaison est effectuée entre la réalisation d'une *SVM* gaussienne sur processeur et avec l'implantation proposée. Des réductions d'énergie de facteurs de 62.4 et 144.7 sont respectivement obtenus pour la détection d'épilepsie et d'arythmie.

1.4.2.4 Réseaux de neurones

Les réseaux de neurones artificiels (*Artificial Neural networks, ANN*) ont été mis en place dans le but d'imiter la structure interne du cerveau humain. Ils sont composés de neurones connectés entre eux selon différentes organisations en fonction du type de réseau. Un type d'*ANN* classiquement utilisé pour faire de la classification est le perceptron multi-couches. Le modèle d'un perceptron est présenté en Figure 1.23. Il est composé d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une couche de sortie. La couche d'entrée contient autant de neurones qu'il y a d'éléments dans le vecteur de paramètres. Les neurones d'entrées sont connectés à la première couche cachée via des connexions unidirectionnelles pondérées par les poids $w_{1,i,j}$. Chaque neurone de la couche

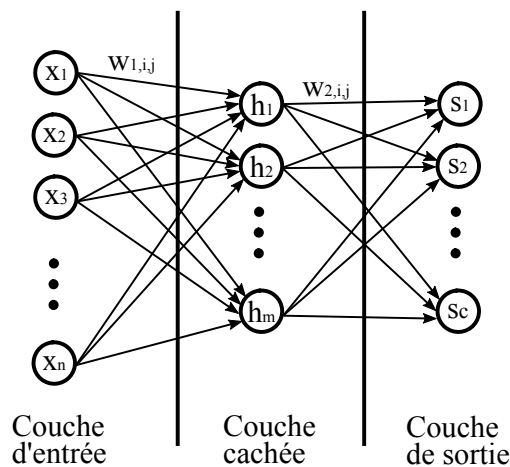


FIGURE 1.23 – Modèle d'un perceptron à une couche cachée, n entrées, m neurones cachés et c neurones de sortie.

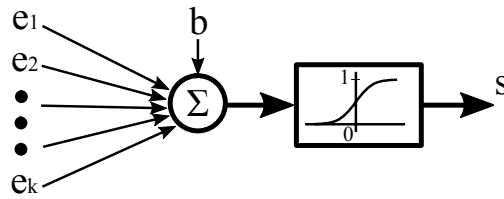


FIGURE 1.24 – Schéma bloc d'un neurone à fonction d'activation de type sigmoïde. Les entrées pondérées e_i sont sommées et un biais b est ajouté. la fonction d'activation transforme la somme en une valeur comprise entre zéro et un.

cachée (Figure 1.24) effectue la somme de ses entrées et ajoute un biais. Cette somme est ensuite transformée via la fonction d'activation. La fonction d'activation est, en général, une fonction de type sigmoïde afin d'obtenir un classifieur non linéaire. Ces neurones sont ensuite connectés à la couche cachée suivante ou à la couche de sortie. Les connexions sont toujours pondérées et unidirectionnelles d'une couche vers la suivante. Les neurones de la couche de sortie peuvent utiliser des fonctions d'activation sigmoïde ou linéaire. Le nombre de neurones dans la couche cachée est en général égal au nombre de classes.

Avec un tel modèle, le rôle de l'algorithme d'apprentissage est de déterminer la valeur des poids et des biais entre chaque couche. Pour entraîner le réseau de la Figure 1.23, il y a $n \cdot m + m \cdot c$ poids et $m + c$ biais à déterminer. Un des inconvénients de l'entraînement de ces réseaux est que le nombre de couches cachées ainsi que le nombre de neurones les formant sont des méta-paramètres qui doivent être connus. C'est aussi le cas pour le type de fonction d'activation utilisé dans les couches cachées et la couche de sortie. Cela conduit généralement à l'entraînement d'un grand nombre de réseaux avec différentes combinaisons de méta-paramètres. De plus, la complexité de l'apprentissage augmente linéairement avec le nombre de neurones dans les couches cachées et quadratiquement avec le nombre de couches cachées.

L'utilisation des perceptrons est fréquente pour résoudre des problèmes de classification liés aux signaux biologiques. Il n'existe cependant pas encore, à la connaissance de l'auteur, d'implantation matérielle de ces derniers dans le cadre des *BASN*. Le Tableau 1.9 présente des réseaux de neurones utilisés pour classer des signaux biologiques.

TABLEAU 1.9 – Exemples d'utilisation de perceptrons sur des signaux biologiques.

	[ZWF14]	[MCXF16]	[PPG16]	[KK13]
Type de signal	EEG, EMG	EEG	ECG	ECG
Application	Fatigue	Dépression	Apnée sommeil	Arythmie
Nombre de paramètres	4	800	20	25
Nombre de couches cachées	2	1	1	2
Nombre de neurones cachées	9/9	10	8	4/4
Nombre de classes	4	2	2	2
Exactitude	0.965-0.995	0.9	0.821	0.95

Afin d'évaluer la consommation d'un perceptron un cas similaire issu du domaine de la classification de gaz polluants [ZAAB16] est utilisé. L'implantation sur un *Field-Programmable Gate Array FPGA* d'un réseau avec 12 neurones d'entrée, trois neurones cachés et un neurone de sortie consomme 959 nJ/classification. Cette consommation ne concerne que la partie classification et non l'extraction des paramètres. Le système atteint une exactitude de 96.8%.

L'analyse des différentes opérations à réaliser dans les *BASN* permet d'orienter les choix pour réduire la consommation. Le capteur étant l'élément le plus contraint, les efforts sont concentrés sur ce dernier. La section suivante discute de la stratégie choisie pour augmenter l'autonomie du capteur.

1.5 Discussion

L'étude des *BASN* menée dans ce chapitre permet de faire un bilan de la consommation énergétique des différentes opérations. Ce bilan est résumé sur la Figure 1.25 pour les capteurs transmetteur et intelligent. On constate très rapidement que l'émission d'un bit est l'étape la plus coûteuse. La consommation des étapes d'extraction et de traitement est très variable selon les applications.

La comparaison entre les deux types de capteurs est effectuée pour deux applications liées aux *BASN*. La première est une application de détection de crise d'épilepsie sur un signal EEG. La seconde consiste à détecter des arythmies dans les signaux ECG. Pour l'application EEG, huit canaux sont échantillonnés sur dix bits à 4 kHz. Le traitement considéré est celui de [YYED⁺13] qui travaille sur deux secondes de signal. Suite au traitement, seul le résultat de la classification binaire sur un bit est transmis. Pour l'application ECG, 800 échantillons de 10 bits sont acquis à 1 kHz. Le classifieur de [CMM⁺13] permet de traiter ces échantillons pour détecter la présence d'arythmie. La Figure 1.26 présente les bilans énergétiques des deux capteurs pour chacune des deux applications. Le CAN considéré est celui de [YYKY10] qui consomme 0,12 pJ/échantillons avec un *ENOB* de dix bits. Le frontal RF utilisé est tiré de [KFC⁺17] qui utilise le standard *BLE* et nécessite 3,7 nJ par bit transmis.

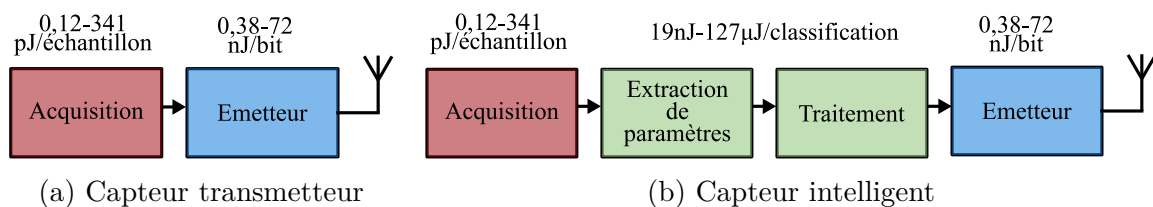


FIGURE 1.25 – Bilan énergétique des deux architectures de capteur. Les énergies minimale et maximale des éléments présentées dans le chapitre sont annotées.

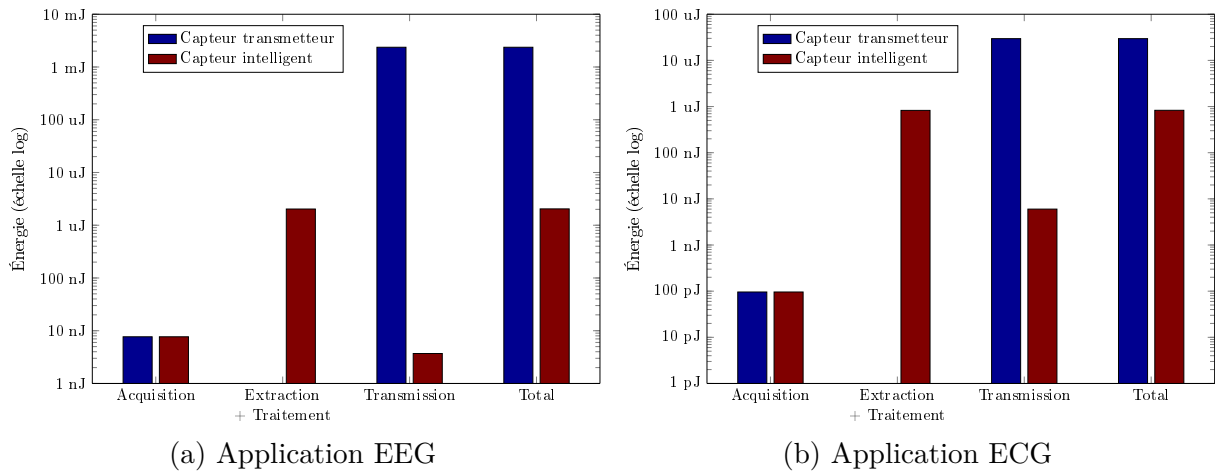


FIGURE 1.26 – Comparaison des budgets énergétiques des deux architectures pour a) l’application EEG et b) l’application ECG.

Pour les deux applications, on constate que l’énergie est dominée par le traitement pour le capteur intelligent et par la transmission pour le capteur transmetteur. Les gains d’énergie obtenus suite à l’intégration du traitement sont de 36 et 1160 pour respectivement l’ECG et l’EEG. Plus la quantité de données à transmettre est réduite plus les gains possibles sont élevés.

La piste étudiée pour augmenter l’autonomie des capteurs est donc de limiter au maximum l’utilisation de l’émetteur RF. L’application considérée est la détection d’arythmie dans les ECG. Cette détection permet l’identification de maladies cardiovasculaires qui est un enjeu majeur pour la société. De plus, c’est un sujet qui est beaucoup étudié et qui offre une littérature riche et variée.

Conclusion

Dans ce chapitre, une revue des différents éléments constituant un *BASN* a été réalisée avec un accent sur le capteur. L’étage RF étant la partie la plus gourmande en énergie, il semble logique d’étudier comment réduire son impact pour améliorer la durée de vie des capteurs. Mise à part l’amélioration des protocoles de communication et de leur implantation, la réduction de la quantité de données à transmettre permet une utilisation sporadique de l’étage RF. Pour ce faire l’utilisation de techniques d’échantillonnage compressif permet d’obtenir des gains énergétiques pour le capteur via une compression des données lors de l’acquisition. Ces techniques se heurtent cependant à une limitation importante : le signal transmis ne peut pas être directement exploité et doit d’abord être reconstruit via des algorithmes complexes à implanter. Si cette solution permet de réduire la consommation du capteur, la complexité rajoutée du côté traitement peut résulter en une augmentation du budget énergétique global. L’intégration du traitement sur le capteur pour ne transmettre que le résultat du traitement semble être la meilleure solution pour réduire l’utilisation de l’étage RF. Cela nécessite un travail sur l’algorithme de classification afin d’implanter ce dernier

de manière efficace.

Dans la suite de ce document, les signaux ECG seront considérés comme cas d'application. Dans un premier temps, un système de classification des ECG propice à l'intégration est proposé. La conception matérielle de ce dernier est ensuite réalisée afin de valider l'intérêt de la solution ainsi que son aspect basse consommation. Enfin, une nouvelle méthode d'acquisition est évaluée afin de réduire encore plus la consommation du système proposé.

Chapitre 2

Système d'analyse d'ECG basse consommation pour l'aide au diagnostic : utilisation des réseaux à cliques

Introduction

Dans ce chapitre, un système d'analyse de signaux ECG est proposé. L'objectif est de concevoir une chaîne de traitement afin d'améliorer l'autonomie énergétique d'un capteur ECG en transmettant uniquement un résultat de classification via un traitement *in situ*. Pour se faire, il est donc nécessaire d'utiliser des traitements basse consommation et intégrables. Dans un premier temps, les spécificités du signal ECG sont présentées ainsi que les méthodes usuelles pour le traiter. Il est possible d'effectuer de nombreux traitements différents sur le signal. Dans le cadre des *BASN*, l'objectif du traitement est la détection d'un comportement anormal. C'est le cas des arythmies qui correspondent à un battement irrégulier du cœur. Elles sont dans la majorité des cas bénignes, mais peuvent, si elles sont prolongées, provoquer un malaise voire la mort du patient. De plus, les personnes souffrant de maladies cardiovasculaires sont particulièrement sujettes aux arythmies. Ainsi l'étude menée dans ce chapitre se focalise sur la classification des arythmies afin de permettre une surveillance de ces dernières. Le système proposé est ensuite détaillé, il utilise des paramètres extraits du signal numérisé afin de classer ce dernier en utilisant des réseaux de neurones à cliques. Les réseaux de neurones à cliques peuvent être considérés comme des mémoires auto-associatives dont le fonctionnement est explicite. De plus, leur utilisation en tant que classifieur est évaluée et appliquée aux signaux de la base de données d'arythmie MIT-BIH [MIT]. Enfin, une estimation de la consommation du système s'appuyant sur une technologie *CMOS* 65 nm montre l'intérêt du système proposé.

2.1 Les systèmes d'analyse d'ECG

Cette section présente les caractéristiques du signal ECG ainsi qu'une chaîne de traitement classique de ce dernier. Le but n'est pas de faire une liste exhaustive des méthodes existantes mais de présenter les plus utilisées. Cette étude permet de motiver les choix fait pour la mise en place du système de la Section 2.2.1.

2.1.1 Le signal ECG

Le cœur est en charge de la circulation du sang dans l'ensemble du corps. Un schéma simplifié de ce dernier est présenté en Figure 2.1. Il est principalement composé de deux oreillettes qui reçoivent le sang venant des organes et de deux ventricules qui renvoient le sang vers ces derniers. Les valvules sont des clapets anti-retour assurant le passage du sang dans un sens unique : des veines vers les artères en passant par les oreillettes puis les ventricules. Le cycle cardiaque est le suivant :

1. remplissage des oreillettes,
2. contraction des oreillettes pour remplir les ventricules,
3. décontraction des oreillettes,
4. contraction des ventricules pour renvoyer le sang dans le système,
5. décontraction des ventricules.

Ce cycle est commandé en continu par le système nerveux. Une manière simple d'observer ce cycle est par la mesure de la variation de potentiel électrique à la surface de la peau. C'est ce que l'on appelle l'électrocardiographie.

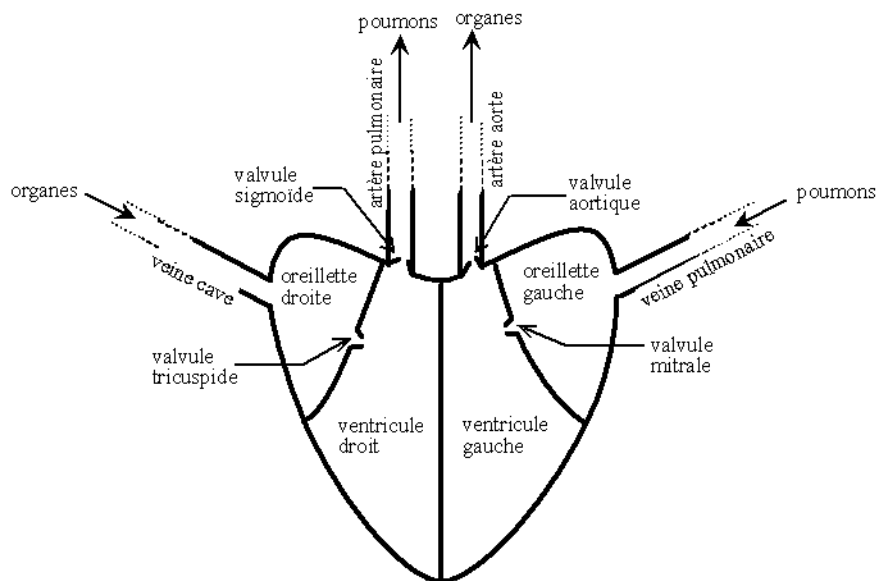


FIGURE 2.1 – Schéma simplifié des différentes sections du cœur tiré de [Cda04].

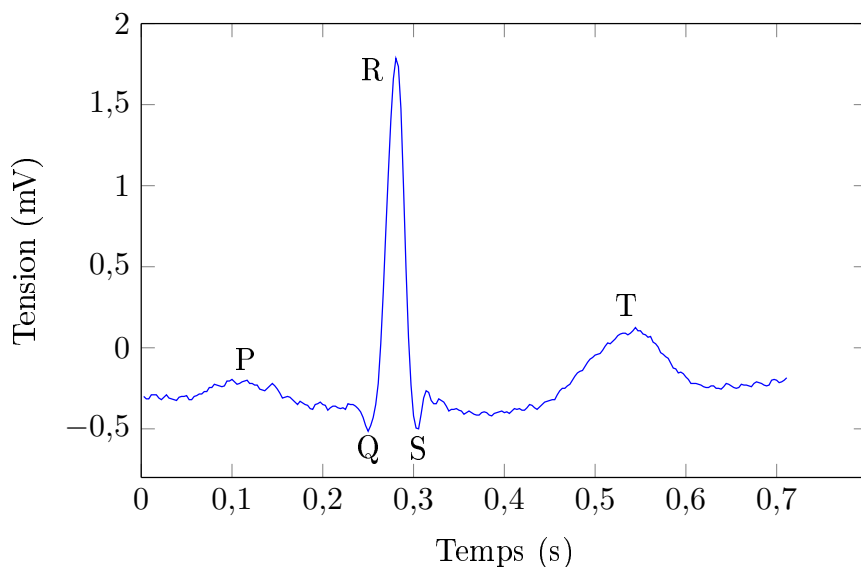


FIGURE 2.2 – Signal ECG d’un cycle cardiaque.

Le signal ECG est le signal obtenu par cette mesure. Un exemple d’un tel signal est présenté sur la Figure 2.2. On peut y voir ce que l’on appelle un complexe PQRST qui représente les différentes étapes du cycle cardiaque :

- l’onde P correspond à la dépolarisation des oreillettes issue de leur contraction,
- le complexe QRS englobe la repolarisation des oreillettes et la dépolarisation des ventricules issue de leur contraction,
- l’onde T est la conséquence de la repolarisation des ventricules.

Les caractéristiques typiques d’un rythme cardiaque normal à 60 battements par minute sont données dans le Tableau 2.1 [LSCCM16].

La mesure est effectuée en plaçant des électrodes sur la peau du patient. Cette mesure est donc soumise à des bruits de différentes natures qui peuvent perturber l’analyse du signal. Le réseau électrique à 50 Hz ou 60 Hz, les mouvements du patient (activité respiratoire, déplacement) ou

TABLEAU 2.1 – Caractéristiques d’un signal ECG à 60 battements par minute pour un adulte sain [LSCCM16].

Caractéristique	Valeur	Variation
Largeur P	110 ms	± 20 ms
Intervalle PR	160 ms	± 40 ms
Largeur QRS	100 ms	± 20 ms
Intervalle QT	400 ms	± 40 ms
Amplitude de P	0.115 mV	± 0.05 mV
Amplitude de QRS	1.5 mV	± 0.5 mV
Amplitude de T	0.3 mV	± 0.2 mV

encore l'usure et le déplacement des électrodes sont par exemples à l'origine de ces bruits. Certains de ces bruits sont connus à l'avance comme par exemple celui issu du réseau électrique ou la ligne de base. Cette dernière est un décalage variable de l'amplitude du signal dans le temps. C'est un bruit qui est très souvent présent et dont la fréquence est de l'ordre de 0,4 Hz. Si certains de ces bruits sont prévisibles ce n'est pas le cas de la majorité d'entre eux qui sont dépendants de nombreux facteurs.

C'est le résultat de cette mesure bruitée qui doit être traité afin de détecter une anomalie. Les anomalies considérées dans ce document sont les arythmies. Elles correspondent à un problème de rythme au niveau du cœur. Une base de données très souvent utilisée afin d'évaluer les traitements sur les ECG est la base MIT-BIH [MIT]. Elle contient 30 minutes d'enregistrement pour 47 patients présentant des problèmes d'arythmie cardiaque. Chaque complexe QRS est annoté avec le type du battement. De même, le point de départ des différentes ondes est aussi indiqué. Ces annotations ont été effectuées de manière indépendante par différents cardiologues afin d'assurer leur exactitude. Les principaux types de battements et leur signification sont les suivants :

- normal : battement normal,
- bloc de branche gauche : blocage de la conductivité du signal pour le ventricule gauche causant un retard de sa contraction,
- bloc de branche droit : identique au bloc de branche gauche mais pour le ventricule droit,
- extrasystole auriculaire : contraction prématurée des oreillettes,
- extrasystole ventriculaire : contraction prématurée des ventricules,
- battement électro-entraîné : battement induit par un stimulateur cardiaque.

La section suivante présente la chaîne de traitement classique du signal ECG afin de classifier le type d'arythmie présente.

2.1.2 Le traitement des signaux ECG pour l'arythmie

La chaîne de traitement présentée en Figure 2.3 utilise les échantillons produits par un frontal d'acquisition contenant un CAN. Le signal ECG n'a subi aucun traitement préalable si ce n'est une amplification linéaire afin de permettre sa numérisation. Dans un premier temps, le pré-traitement consiste à supprimer un maximum de bruit. Cela améliore la qualité du signal afin d'augmenter les performances de classification. Le signal est ensuite segmenté en blocs de données de taille fixe afin de constituer un vecteur d'observation de taille finie. Des paramètres permettant de décrire le signal sont ensuite extraits de chaque vecteur d'observation. Enfin ces paramètres sont utilisés par le classifieur qui décide de la classe de signal présent : signal normal ou type d'arythmie présente.

Comme précisé dans [LSCCM16], les étapes de pré-traitement et de segmentation sont intimement liées aux paramètres à extraire. En effet, si les paramètres extraits sont de type temporel, il est alors indispensable d'avoir une segmentation permettant la détection du complexe QRS et du pic R de manière précise. La suite de cette section se focalise sur les méthodes les plus utilisées pour chaque étape du traitement.

2.1.2.1 Le pré-traitement

Le pré-traitement a pour but de filtrer le bruit sur le signal ECG. Une méthode très simple et efficace est l'utilisation de filtres numériques à réponse impulsionnelle finie (RIF). Un filtre RIF d'ordre N est définie par $N + 1$ coefficients. Chaque nouvelle sortie du filtre est une somme pondérée des N entrées précédentes tel que :

$$y[n] = \sum_{k=0}^{N-1} c_k \cdot x[n - k], \quad (2.1)$$

où $y[n]$ et $x[n]$ sont respectivement la sortie et l'entrée à l'instant n et les c_i sont les coefficients du filtre. Ces filtres sont simples à réaliser et il est possible d'en synthétiser afin d'éliminer une bande de fréquence particulière.

Une méthode très utilisée [ZDL⁺14][MZM⁺11][COR04] consiste à utiliser un filtre qui remplace la valeur d'un échantillon par la médiane des valeurs des échantillons qui l'entourent. Un tel filtre est appelé un filtre médian. La méthode de filtrage utilise un premier filtre médian de largeur 200 ms suivi d'un second de 600 ms. Le filtre de 200 ms est utilisé pour supprimer le complexe QRS et l'onde P alors que le filtre de 600 ms supprime l'onde T. Le signal ainsi obtenu est la ligne base qui est ensuite soustraite du signal d'origine. Le signal corrigé est ensuite filtré par un passe-bas d'ordre 11 avec une fréquence de coupure à -3 dB de 35 Hz. Cela permet de supprimer le bruit d'alimentation à 50 Hz ou 60 Hz ainsi que tout autre bruit de fréquence supérieure à 35 Hz. Le soucis de cette méthode réside dans l'utilisation d'un filtre médian. Un tel filtre nécessite le stockage d'un grand nombre d'échantillons et le classement de ces derniers afin de déterminer la médiane. Il est donc nécessaire d'implanter un algorithme de tri en matériel.

Une autre méthode de filtrage présentée dans [ST06] et utilisée dans [ESH⁺16][MAM13] consiste à utiliser une décomposition en ondelettes discrète (*Discrete Wavelet Transform, DWT*) de type Daubechies 6 (*db6*) sur neuf niveaux. C'est le débruitage par ondelettes. Il est effectuée sur un signal issu de la base MIT-BIH échantillonné à 360 Hz. La *DWT* est un outil d'analyse multi-résolutions permettant d'obtenir des informations à la fois temporelles et fréquentielles. Elle est réalisée via des filtres effectuant des décompositions successives en deux sous-bandes couvrant le spectre du signal. Les filtres divisant la bande passante des signaux par deux, ceux-ci sont sous-échantillonnés par deux afin d'éliminer les échantillons inutiles. Une telle décomposition est dite dyadique. La structure de décomposition dyadique, sur trois niveaux, d'un signal s quelconque est illustrée en Figure 2.4. Les coefficients issus du filtre passe-haut sont les coefficients de détail (Cd) et ceux du passe-bas les coefficients d'approximation (Ca). La structure duale permet de

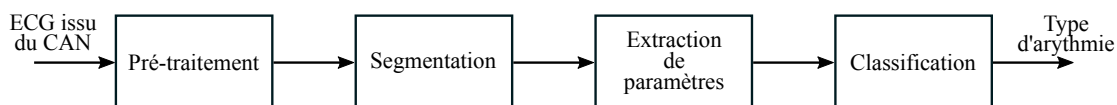


FIGURE 2.3 – Chaîne de traitement typique pour la classification des signaux ECG.

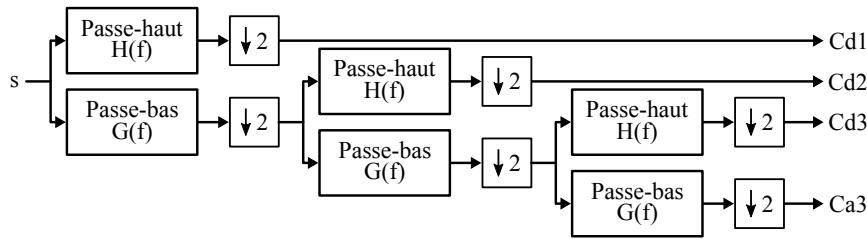


FIGURE 2.4 – Structure dyadique de filtrage pour une décomposition en ondelettes sur 3 niveaux.

reconstruire le signal d'origine à partir des coefficients. Ces derniers sont sur-échantillonnés par un facteur deux, puis les filtres de reconstruction sont utilisés. Ainsi pour chaque type ondelette, quatre filtres sont associés : deux pour la décomposition et deux pour la reconstruction. Pour les ondelettes de type Daubechies dbn le nombre de coefficients de chacun des filtres est de $2 \cdot n$.

Le débruitage par ondelettes de [ST06] consiste à forcer à zéro tous les coefficients des niveaux de décomposition un, deux et neuf. En effet, ces niveaux de décomposition ne comportent pas d'information utile sur le signal. Ce dernier étant échantillonné à 360 Hz, les niveaux de décomposition un et deux contiennent l'information fréquentielle contenue dans les bandes 90-180 Hz et 45-90 Hz. Les supprimer permet de filtrer le bruit d'alimentation à 50 Hz ou 60 Hz ainsi que d'autres artefacts à haute fréquence. Comme la bande passante du signal est divisée par deux à chaque niveau, le signal issu du dernier passe-bas contient l'information entre 0 Hz et $180/2^9 = 0,3516$ Hz. La seule partie du signal à cette faible fréquence est la ligne de base. Il est donc intéressant de la supprimer. Suite à la mise à zéro des coefficients, les filtres de reconstruction sont utilisés afin d'obtenir le signal débruité. L'inconvénient de cette méthode est qu'elle nécessite de nombreux filtres. Il faut en effet 18 filtres pour la décomposition et autant pour la reconstruction. Cela implique donc l'implantation de 36 filtres d'ordre 11 sur le capteur ce qui peut être limitant pour une implantation matérielle.

2.1.2.2 La segmentation

La segmentation est la deuxième étape de la chaîne présentée en Figure 2.3. Son objectif est de créer un vecteur d'observation de taille finie sur lequel des paramètres seront extraits. La segmentation consiste à passer d'un traitement en ligne à un traitement par bloc. Pour ce faire il est nécessaire de déterminer la taille du bloc en nombre d'échantillons ainsi que l'échantillon de départ pour chaque bloc. Dans la plupart des cas [PS84] [KHO02] [OAT14], la segmentation du signal ECG consiste à prendre un nombre fixe d'échantillons avant et après le complexe QRS. Cela revient à prendre un cycle cardiaque. Pour effectuer cette segmentation, il est donc nécessaire de détecter la position du complexe QRS. Cette étape de détection du complexe est très importante puisque c'est sur elle que repose l'étape suivante d'extraction de paramètres. En effet, si un complexe est manqué alors aucun paramètre ne sera extrait et la classification ne pourra pas avoir lieu. De même, un faux positif entraînera une extraction de paramètres erronés causant une erreur de classification par la suite.

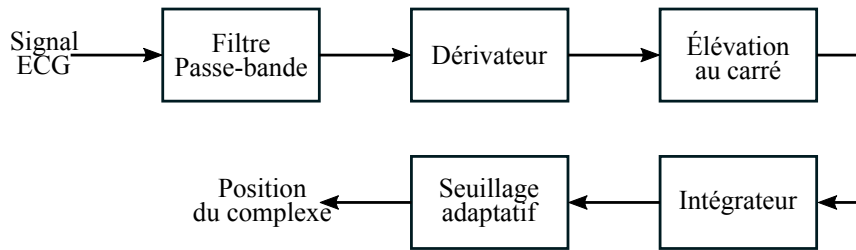


FIGURE 2.5 – Schéma bloc de la méthode de détection du complexe QRS de Pan et Tompkins [PT85].

Il existe deux principales techniques pour détecter le complexe QRS : les méthodes travaillant dans le domaine temporel et celle utilisant une analyse temps fréquence du signal. La méthode la plus utilisée parmi les méthodes temporelles est celle proposée par Pan et Tompkins [PT85] et illustrée en Figure 2.5. Elle est implémentée de base dans *Matlab*[®] afin d'en permettre son déploiement simple et rapide. La méthode consiste à filtrer le signal afin d'en extraire le complexe QRS. Le premier filtre est un passe-bande filtrant toutes les fréquences en dehors de la plage 5-15 Hz. Le signal est ensuite dérivé avec un dérivateur d'ordre trois avant d'être élevé au carré. L'élévation au carré permet à la fois d'avoir des valeurs positives et d'obtenir une amplification non linéaire. Enfin, le signal est intégré avec un fenêtrage glissant de largeur 150 ms. Pour obtenir la position des pics, un seuil est appliqué au signal. L'algorithme mis en place [PT85] utilise une mise à jour du seuil afin de s'adapter à l'évolution du signal. Cette méthode a été évaluée sur la base de données MIT-BIH. Une comparaison est effectuée entre la sortie du détecteur et l'annotation du pic R fournie avec la base de données. Sur l'ensemble de la base, 0,437 % des battements détectés sont des faux positifs et 0,239 % ne sont pas détectés. Cela se traduit par une sensibilité de 99,76 %.

Les auteurs de [RRDLL14] et [ZL14] ont également proposé une architecture de détection temporelle du complexe QRS. Elle utilise un CAN de type *Level crossing*. Au lieu de produire des échantillons de manière uniforme, ce type de convertisseur produit un échantillon dès que le signal d'entrée franchit certains seuils. Cette méthode d'échantillonnage permet de réduire le nombre d'échantillons acquis tout en effectuant un filtrage du bruit. En effet, les bruits de faible amplitude ne franchiront pas le seuil et donc ne déclencheront pas l'acquisition d'un nouvel échantillon. Cette architecture est bien adaptée pour les signaux parcimonieux dans le temps, c'est à dire souvent constant et dont la valeur ne change que sur de petits intervalles, comme l'ECG. De plus, selon le seuil franchi, il est possible de savoir si le signal est en train de croître ou de décroître. C'est cette information qui est utilisée pour détecter une forte augmentation ou diminution du signal. Ensuite l'utilisation d'un seuillage adaptatif sur cette augmentation/diminution permet la décision de la position du complexe QRS. La sensibilité obtenue dans [RRDLL14] est de 98,89 % et celle de [ZL14] atteint 97,76 %. Ces méthodes réduisent la complexité matérielle du fait d'un nombre réduit d'échantillons à traiter et d'un traitement plus simple à base de seuils. Ainsi la solution proposée dans [ZL14] consomme 34 nW pour une implantation en *CMOS* 0,13 μm . Le désavantage est que, suite à la détection du complexe, le signal est toujours représenté via un échantillonnage

non-uniforme. Il est donc nécessaire, soit de modifier les traitements suivants afin de les adapter, soit de transformer le signal pour avoir une représentation uniforme.

La deuxième famille de méthodes de détection du complexe QRS consiste à utiliser une analyse temps fréquence du signal. Cela se fait principalement par l'utilisation de la transformée en ondelettes. Lorsque les coefficients sont élevés sur plusieurs niveaux simultanément, une pente est détectée. Des seuils sont alors utilisés afin de déterminer si la pente détectée correspond à celle du complexe QRS. Une implantation matérielle d'un tel détecteur est mise en place dans [KKH⁺14]. Les expérimentations sur puce ont permis d'obtenir une sensibilité de 100 % pour un *SNR* de -18 dB. Les données de tests sont produites en ajoutant un bruit sur un signal ECG non bruité. La consommation de la partie détection du circuit, fabriqué en technologie CMOS 0,18 μm , est de 43 μW . Bien que la technologie soit plus ancienne, la consommation est 100 à 1 000 fois supérieure que celle d'une méthode temporelle.

D'autres méthodes de détection sont aussi mentionnées dans [KHO02] et [OAT14] comme l'utilisation de réseaux de neurones, de filtres adaptés ou encore d'algorithmes génétiques. Parmi tous les détecteurs présentés dans [KHO02], plus de 70 % atteignent au minimum une sensibilité de 99 % et seule une solution présente une sensibilité inférieure à 95 %.

2.1.2.3 L'extraction de paramètres

L'extraction de paramètres a pour but d'analyser le signal afin d'obtenir des informations utilisables par un classifieur. Cette section présente les paramètres les plus souvent utilisés. Il n'est pas possible d'évaluer directement la qualité d'un paramètre puisque celle-ci est aussi dépendante du classifieur. Les performances liées aux différents paramètres seront donc traitées dans la section suivante sur les classifieurs.

Il existe trois familles de paramètres afin de classifier les signaux ECG : les paramètres temporels, morphologiques et spectraux. Les paramètres temporels sont des mesures de temps comme par exemple ceux du Tableau 2.1. Il est aussi possible d'utiliser l'intervalle entre deux complexe QRS successifs appelé intervalle RR. Dans [COR04], la moyenne globale ainsi que la moyenne locale sur dix battements de l'intervalle RR sont utilisées.

Les paramètres morphologiques sont les paramètres représentant la structure du signal. Ils consistent généralement à prendre directement les échantillons de différentes sections du signal ECG. Afin que le vecteur de paramètres ne soit pas trop grand, tous les échantillons ne sont pas utilisés. Seul un sous-ensemble d'échantillons sont conservés pour les différentes sections du signal. Ainsi les auteurs de [COR04] proposent plusieurs jeux de paramètres morphologiques comme par exemple la prise de dix échantillons uniformément répartis entre le début et la fin du complexe QRS. Il est aussi possible d'utiliser des mesures statistiques sur le signal comme la moyenne, la variance ou des statistiques utilisant des puissances du signal supérieures à deux. Ces mesures sont appelées *High-order statistics (HOS)* et permettent d'obtenir des informations sur les non-linéarités du signal. Ces paramètres sont complexes à obtenir puisqu'ils nécessitent le calcul du signal à des puissances supérieures à deux. De tels paramètres sont utilisés dans le cadre de l'étude

menée dans [ESH⁺16].

Enfin la famille des paramètres spectraux consiste à effectuer une transformation en ondelettes sur le signal afin d'obtenir des informations sur le spectre du signal. La famille d'ondelette ainsi que le nombre de niveaux de décomposition diffèrent selon les publications et peut jouer un rôle important. Ainsi, dans [MAM13] l'ondelette de Meyer *dmey* est utilisée sur quatre niveaux. Dans [MZM⁺11], une décomposition sur cinq niveaux avec une ondelette de type spline quadratique est utilisée. L'ondelette de Daubechies 2 *db2* sur quatre niveaux est utilisée dans [Gb05]. De la même manière que pour les paramètres temporels, la décomposition en ondelettes produit un grand nombre d'échantillons. Afin de réduire ce nombre des statistiques sur chacun des niveaux de décomposition sont utilisées. Des statistiques souvent utilisées sont le minimum, le maximum, la différence ou le rapport entre minimum et maximum ou encore la variance.

Pour les paramètres morphologiques et spectraux, il est aussi possible d'utiliser des techniques de réduction de dimension comme l'analyse linéaire discriminante (*LDA*), l'analyse en composantes principales (*Principal Component analysis PCA*) ou l'analyse en composantes indépendantes (*Independent Component analysis ICA*). Leur objectif est de projeter le signal d'entrée sur une base de dimension plus petite ayant des propriétés particulières. Les méthodes *LDA* et *PCA* sont linéaires, alors que l'*ICA* ne l'est pas. La *PCA* est une méthode non supervisée qui vise à maximiser la variance des données. L'objectif de l'algorithme est de trouver des combinaisons linéaires orthogonales afin de maximiser la variance des données dans le nouvel espace. La projection des données sur ce nouvel espace permet de réduire le nombre d'éléments à fournir au classifieur tout en minimisant la corrélation entre les données. La *LDA* est une méthode supervisée, c'est-à-dire que pour chaque observation, la classe associée est connue. Grâce à cela, à la manière de la *PCA*, la *LDA* cherche des combinaisons linéaires des éléments d'entrée afin de maximiser la séparation entre les classes. La *LDA* pour la réduction de dimension est similaire à la *LDA* présentée dans le Chapitre 1 qui est utilisée pour la classification. Enfin l'*ICA* tente de décomposer le signal d'entrée sur une base dont les vecteurs sont des composants indépendants du signal. L'algorithme considère le signal observé comme est une combinaison linéaire d'éléments indépendants. Il tente alors de déterminer à la fois quels sont ces composants ainsi que les coefficients des combinaisons linéaires. Ces méthodes sont appliquées sur le jeu de données d'entraînement afin de déterminer les combinaisons linéaires à appliquer aux signaux lors des tests.

Comme mentionné dans la section précédente, la plupart des ces paramètres nécessitent une bonne détection du complexe QRS afin d'être extraits. Si les différents éléments du signal ne sont pas bien identifiés, les paramètres temporels et morphologiques ne seront pas bons. Les paramètres issus des statistiques sur la décomposition en ondelettes sont moins impactés par une imprécision dans la détection du pic R. En effet, même si la détection du pic R est décalée de quelques échantillons cela ne changera pas la valeur du minimum et du maximum.

2.1.2.4 L'étage de classification

Le classifieur est l'élément qui, à partir des paramètres extraits, décide la classe présente à l'entrée du système. Il existe de nombreuses manières d'effectuer la classification. Les principaux classifieurs utilisés ont été présentés dans le Chapitre 1. Parmi ceux-ci on trouve les classifieurs probabilistes avec les classifieurs Bayésien naïfs (NB), la *LDA*, l'analyse discriminante quadratique (*QDA*) et de la régression logistique (LR). On peut aussi trouver des arbres de décision, des machines à vecteurs de supports *SVM*, des réseaux de neurones artificiels (*ANN*) ou des réseaux de neurones probabilistes (*PNN*).

Le Tableau 2.2 présente l'exactitude de classification des arythmies pour les références citées. Une étude présentant un plus grand nombre de méthodes est disponible dans [LSCCM16].

Toutes ces méthodes offrent de très bonnes performances de classification mais sont réalisées de manière logicielle. La complexité matérielle n'y est pas considérée. Or dans le cadre de l'étude menée dans ce document, le fait de pouvoir intégrer le traitement à basse consommation est primordial. En effet, l'intégration du traitement sur le capteur est intéressante uniquement si il y a un gain en terme de consommation. Afin de prendre cela en compte, la section suivante discute de l'aspect implantation matérielle de la chaîne de traitement.

2.1.2.5 Considérations matérielles

Les méthodes de traitements présentées dans la section précédente offrent des exactitudes de classification pouvant atteindre 99,7 %. Le problème est que ces méthodes sont purement logicielles. L'objectif étant la mise en place d'un capteur intelligent pouvant traiter le signal *in situ*, il est impératif d'avoir un traitement basse consommation. Afin de pouvoir fixer des objectif en termes

TABLEAU 2.2 – Synthèse des méthodes utilisées pour classifier le type d'arythmie dans les ECG.

Ref	Nombre de classes	Paramètres	Classifieur	Exactitude
[ESH ⁺ 16]	5	(DWT+PCA)+(HOS+ICA)	SVM(RBF)	98,91 %
[WCHY13]	8	LDA	PNN	98,26 %
		PCA	PNN	99,6 %
		LDA+PCA	PNN	99,71 %
[MAM13]	7	DWT+PCA	PNN	99 %
		DWT+LDA	ANN	98,59 %
		DWT+ICA	PNN	99,28 %
[COR04]	5	RR+morphologique	LDA	83 %
[MZM ⁺ 11]	4	RR+morphologique		
		+DWT (statistiques)	ANN	89 %
[Gb05]	4	DWT (statistiques)	ANN	96,9 %

DWT : transformation en ondelettes discrète, PCA : analyse en composantes principales, HOS : statistique d'ordre supérieur, ICA : analyse en composantes indépendantes, SVM : machine à vecteurs de support, LDA : analyse linéaire discriminante, PNN : réseau de neurones probabilistes, ANN : réseau de neurones artificiels, RR : intervalle R-R.

de performances de classification pour le système, des circuits intégrés réalisant la classification d'arythmie sont utilisés comme référence.

Dans [BMD⁺14], un système numérique multi-cœurs est mis en place afin de débruiter et traiter le signal. La classification mise en place est une classification binaire pour détecter la fibrillation atriale. C'est un type d'arythmie durant laquelle les oreillettes se contractent de manière désordonnée. Le système est testé avec la base de donnée MIT-BIH AF [MM92] qui est spécialisée dans ce type d'arythmie. L'exactitude de classification n'est pas donnée mais la spécificité et la sensibilité sont respectivement de 93 et 96 %. Le système consomme environ 137 μW . En considérant un rythme cardiaque de 80 battements par minute cela équivaut à une énergie de 103 μJ par classification en prenant en compte la chaîne complète.

Les auteurs de [CMM⁺13] utilisent des paramètres spectraux et un classifieur de type *LDA* afin de classifier des signaux ECG. La classification considérée est aussi une classification binaire qui détecte si le signal est normal ou pas. Le circuit proposé est un circuit numérique utilisant une technologie *CMOS* 0,13 μm . Les simulations post-synthèse ont montré les performances de classification suivante : spécificité de 94,2 %, sensibilité de 86,5 % et exactitude de 90,4 %. De plus, l'estimation via *Synopsys Design Compiler*[®] de la consommation dynamique pour l'ensemble du traitement est de 182,94 nW. Le traitement d'un battement étant effectué en 4,516 secondes, le système requière 826 nJ par classification. Il est à noter que cette énergie est uniquement l'énergie dynamique et ne prend pas en compte la consommation statique qui n'est pas donnée.

Enfin, une architecture autour d'un processeur basse énergie combiné à un accélérateur matériel pour réaliser une *SVM* est proposée dans [LV13]. Le système effectue une extraction de 21 paramètres dont trois temporels et 18 morphologiques. Le circuit est implanté en utilisant une technologie *CMOS* basse énergie en 0,13 μm . Les tests du circuit mettent en évidence une exactitude de classification moyenne autour de 92 %. L'énergie pour classifier un vecteur de paramètres est de 124 μJ .

Ces trois exemples d'implantation montrent que les solutions matérielles ont des performances réduite d'environ 7 % par rapport aux solutions logicielles. Afin de se placer dans le contexte des systèmes de traitement intégrés, le système proposé dans la section suivante devra faire au moins aussi bien que les systèmes matériels. Cela correspond à une exactitude moyenne d'environ 92 %.

2.2 Proposition d'un système de classification d'arythmie basse consommation

2.2.1 Présentation du système

2.2.1.1 Objectif du système proposé

Dans cette section, un système de traitement des signaux ECG pour la classification d'arythmie est proposée. L'objectif est de permettre un traitement basse consommation pouvant être effectué sur le capteur. La consommation totale du système doit être suffisamment basse pour que l'énergie

nécessaire au traitement et à la transmission du résultat soit inférieure à celle de la transmission directe du signal.

Le signal d'entrée considéré est issu de la base de données d'arythmie MIT-BIH [MIT]. Il est échantillonné à une fréquence de 360 Hz et encodé sur 11 bits pour une plage d'entrées comprises entre -5 et 5 mV. L'objectif du système est d'être capable de traiter le signal issu de la base de donnée afin de détecter différents types de battements arythmiques. Cette détection d'arythmie permet un suivi des patients atteints de maladies cardiovasculaires. Pour des patients sains, un tel système peut permettre de détecter l'apparition d'un problème cardiaque ou encore de détecter un pic de stress. Le contexte de ce système étant son intégration dans un capteur au sein d'un *BASN*, il devra avoir une faible consommation énergétique. De même, les performances de classification devront être au niveau, voire supérieures, de celles d'autres systèmes sur puce. Comme montré dans la section précédente, cela revient à une exactitude de classification d'au moins 92 %.

2.2.1.2 Architecture du système proposé

Le système proposé reprend l'architecture présentée en Figure 2.3. Les principales différences reposent dans les choix de la segmentation et du classifieur.

Afin de réduire la consommation énergétique, des blocs de données de dix secondes sont considérés. Avec une telle segmentation, il n'est pas nécessaire de détecter le complexe QRS, conduisant à une simplification de l'architecture matérielle. De plus, dans la majorité des systèmes, le débruitage est utilisé pour améliorer la détection du complexe QRS. Le système proposé n'ayant pas besoin de détection du complexe, le filtrage du bruit n'est pas réalisé. L'architecture du système complet est présentée sur la Figure 2.6.

Un bloc de données de dix secondes de signal ECG correspond à 3600 échantillons temporels. Ce signal est d'abord transformé via une décomposition en ondelettes de type Daubechies 2 (*db2*). Le type d'ondelette est choisi pour son utilisation classique dans les système d'analyse d'ECG [Gb05] et pour sa simplicité. En effet, les filtres numériques pour effectuer la décomposition ne nécessitent que quatre coefficients. Cela permet une implantation matérielle simple. La décomposition est effectuée sur cinq niveaux, produisant cinq jeux de coefficients de détail (Cd_i) et un jeu de coefficients d'approximation (Ca_5). Afin d'avoir suffisamment de coefficients la décomposition s'arrête au cinquième niveau. Le signal d'entrée étant de taille 3600, les signaux Cd_5 et Ca_5 ne

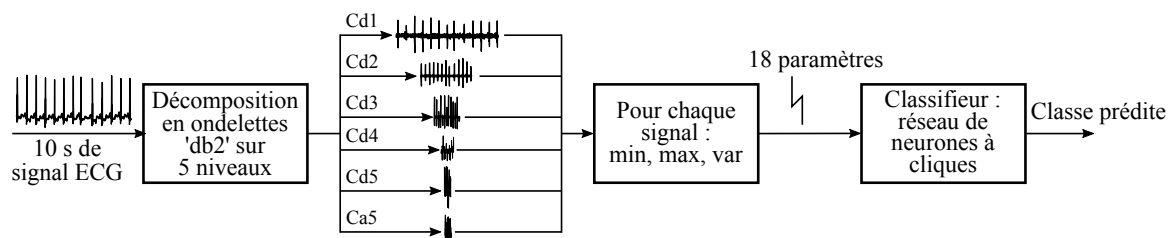


FIGURE 2.6 – Schéma bloc du système proposé pour la classification de signaux ECG [CCL⁺16].

contiennent que 112 coefficients. De plus, les coefficients d'approximation Ca5 contiennent l'information contenue dans la bande 0-5,6 Hz qui est principalement composée de bruit. Pour réduire le nombre de paramètres à fournir au classifieur, le maximum, le minimum et la variance sont extraits de chacun des six jeux de coefficients. Cela conduit à la création d'un vecteur de paramètres de 18 éléments.

Pour effectuer la classification, l'utilisation des réseaux à cliques est proposée. Les réseaux à cliques sont un type de réseaux de neurones artificiels parcimonieux. L'information est portée par l'interconnexion de certains neurones via des connexions binaires. Cela s'oppose aux réseaux de type perceptron ou récurrent dans lesquels toutes les connexions sont réalisées et l'information est portée par des poids synaptiques. Une récente implantation mixte analogique/numérique [Lar15][LLSA16] a montré une consommation très faible de l'ordre de 68 fJ par neurone. Les réseaux à cliques sont décrits dans la section suivante. Ce travail est la première utilisation d'un tel type de réseau pour effectuer une opération de classification.

2.2.2 Les réseaux à cliques

Cette section présente un type particulier de réseau de neurones artificiels : le réseau à cliques. Le réseau à cliques a dans un premier temps été présenté comme une mémoire auto-associative dans les réseaux de Willshaw et Palm [WBLH69][Pal13]. Une mémoire auto-associative est capable de retrouver une information stockée à partir d'une version incomplète ou erronée de cette même information. Une extension des réseaux à cliques est proposée dans [GB11]. Elle ajoute une organisation des nœuds en *cluster*. Ces clusters ne sont pas utilisés dans le cadre de cette application mais ils peuvent améliorer les performances des réseaux à cliques dans des cas précis. Cette section décrit d'abord l'organisation des éléments dans le réseau. Les notions de stockage de l'information puis les procédures de décodage sont ensuite explicitées.

2.2.2.1 Structure des réseaux à cliques

Un réseau à cliques est composé de nœuds appelés *neurones* connectés entre eux par des connexions binaires. Dans la suite du document le nombre de neurones dans un réseau est noté N_N . Chaque neurone représente une information spécifique. Les informations portées par les différents neurones peuvent être de même nature ou de nature totalement différente. L'exemple de la Figure 2.7 montre un ensemble de 19 neurones portant différentes informations. Chaque neurone de la catégorie age représente une plage d'âge allant de très âgé (TA) à très jeune (TJ). De même pour les différents neurones des catégories poids (très lourd à très léger) et taille (très grand à très petit). Enfin les quatre neurones A1 à A4 caractérisent des attributs d'une personne comme par exemple le fait d'être sportif, musicien, amateur de cuisine ou de peinture. L'association de ces informations de base permet de représenter des notions plus complexe comme par exemple un type de personne. Cela correspond à stocker de l'information dans le réseau. Ce principe est expliqué dans la Section 2.2.2.2.

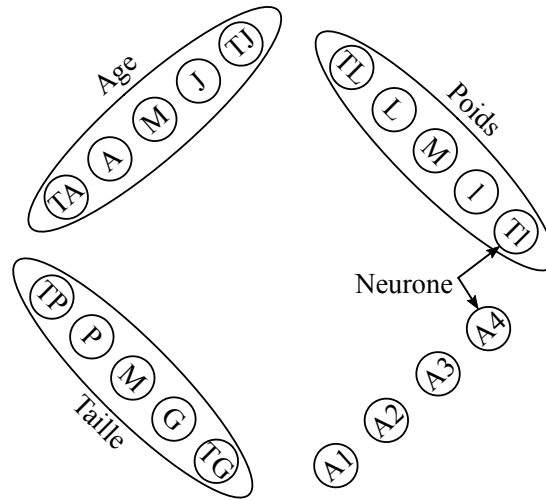


FIGURE 2.7 – Exemple d’un ensemble de neurones pour la représentation de personnes.

Dans les réseaux à cliques, les différents neurones sont connectés les uns aux autres via des connexions binaires. Le réseau n’est pas totalement interconnecté, seul un sous-ensemble des connexions possibles est réalisé. Le choix des connexions existantes est effectué durant la phase de stockage du réseau décrite en Section 2.2.2.2. Lorsque qu’elles existent, les connexions sont binaires, c’est à dire qu’elles ne sont pas pondérées comme dans d’autres modèles de réseaux de neurones artificiels. Cela se traduit par un échange d’informations binaires entre les différents neurones, simplifiant les opérations du neurone.

Le modèle d’un neurone est donné sur la Figure 2.8. Il est constitué de n entrées, appelées *synapses*, provenant de la sortie d’autres neurones et d’une entrée de stimulation extérieure. La première fonction d’un neurone est de calculer son *nombre de contributions actives*, ν_i , correspondant à la somme de ses entrées :

$$\forall i, 1 \leq i \leq N_N : \nu_i = \sum_{j=1}^n w_{i,j} + w_{i,ext}, \quad (2.2)$$

où $w_{i,j}$ est la valeur binaire de l’entrée j du neurone i et $w_{i,ext}$ celle de la stimulation externe. Ce nombre de contributions actives est ensuite comparé à celui des autres neurones afin de vérifier s’il respecte la règle d’activation. Si elle est respectée, alors la sortie S_i est fixée à un et le neurone est

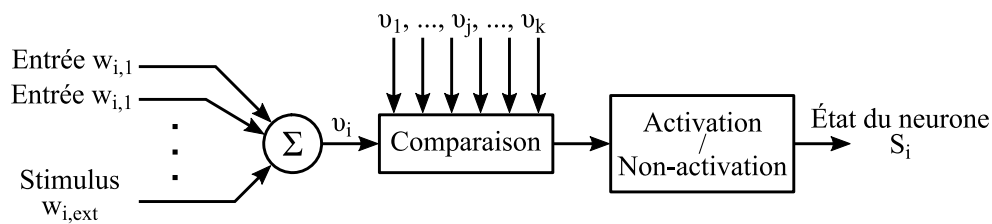


FIGURE 2.8 – Schéma bloc présentant les différentes fonctions d’un neurone.

dit *actif*. Dans le cas contraire la sortie vaut zéro et le neurone reste *inactif* ou est *désactivé* s'il était préalablement actif.

La règle d'activation met en compétition la totalité des neurones du réseau. Elle peut être de type *k-Winners-Take-All* (*k-WsTA* ou *WsTA*) ou *Losers-Kicked-Out* (*LsKO*). Ces règles sont décrites dans la Section 2.2.2.3.2.

2.2.2.2 Notion de stockage

Comme précisé dans la section précédente, il est possible de stocker une information générale à partir des informations élémentaires que représentent les neurones. L'ensemble des informations générales stockées dans le réseau est appelé *dictionnaire*. La phase de stockage dans le réseau est réalisée hors-ligne, c'est-à-dire qu'elle est effectuée avant de faire fonctionner le réseau. Il est donc nécessaire d'avoir une connaissance *a priori* des informations à stocker. Cette phase de stockage consiste à réaliser des connexions entre neurones et peut être comparé à la phase d'apprentissage dans les autres types de réseaux. Pour ce faire, l'information générale est décomposée en informations élémentaires correspondant à celles des neurones du réseau. Le stockage consiste ensuite à venir interconnecter ensemble tous les neurones formant cette information. Le sous-graphe complet réalisé est appelé une *clique*. La Figure 2.9 s'appuie sur les neurones de la Figure 2.7 en y ajoutant le stockage de trois cliques. Les trois cliques permettent de représenter trois informations qui dans ce cas correspondent à trois personnes différentes :

- une personne d'âge moyen, de très petite taille et plutôt lourde avec les attributs A3 et A4 en rouge,
- une personne très âgée, petite, de poids moyen présentant les attributs A2 et A3 en bleu,
- une personne jeune, très grande et très lourde avec les attributs A1 et A2 en gris.

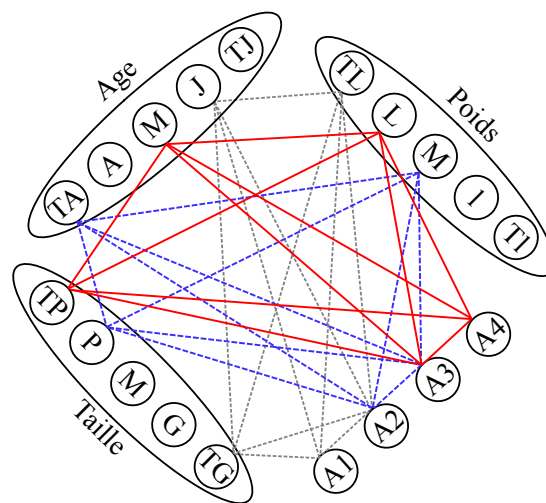


FIGURE 2.9 – Exemple de réseau à cliques avec stockage de trois cliques représentant trois personnes.

Le stockage de plusieurs cliques peut entraîner la création d'une connexion partagée. Dans ce cas, cette dernière reste unique et n'est pas doublée.

La taille c d'une clique est définie comme le nombre de neurones qu'elle connecte. Il est possible d'avoir des cliques de tailles différentes au sein d'un même réseau. Dans ce cas, il n'est plus possible de d'utiliser la règle de décodage globale k -*WsTA* qui nécessite que toutes les cliques aient la même taille.

2.2.2.3 Décodage

2.2.2.3.1 Récupération d'information

Le décodage est l'étape consistant à utiliser le réseau pour retrouver une information stockée. Le réseau fonctionne comme une mémoire auto-associative. Cela signifie que le réseau retrouve une information de son dictionnaire à partir d'une version bruitée ou incomplète de cette même information. Cela est possible de part l'interconnexion entre les neurones d'une clique qui crée une redondance entre les éléments qui la composent.

Le processus de décodage débute par une stimulation extérieure des neurones du réseau. Pour ce faire l'information disponible en entrée est décomposée en informations élémentaires correspondant aux neurones. Elle est ensuite appliquée via les entrées de stimulation des neurones. Les neurones appliquent ensuite les règles de décodage afin de décider de leur état. Les états sont ensuite propagés aux autres neurones via les connexions. Cette séquence de règles d'activation et de propagation est répétée N_{it} fois, jusqu'à ce que les états des neurones convergent. La valeur empirique maximale de N_{it} déterminée suite aux expérimentations menées sur les réseaux à cliques est de quatre. Une fois la convergence atteinte, l'ensemble des neurones actifs correspond à la réponse du réseau à la stimulation.

La procédure de récupération de l'information est présentée dans l'Algorithme 2.1. Sa complexité est principalement due à la règle d'activation et aux nombres d'éléments impliqués dans

Stocker le dictionnaire dans le réseau.

Tant que en fonctionnement

Stimuler les neurones avec l'information d'entrée.

Pour tout itération $k = 1$ à N_{it}

Pour tout neurone $i = 1$ à N_N

 Calculer et stocker ν_i .

Fin

 Appliquer la règle d'activation entre tous les neurones.

 Mettre à jour la sortie des neurones.

 Propager l'état des neurones.

 Mettre à jour les entrées des neurones.

Fin

Fin

Algorithme 2.1 : Procédure d'une itération de décodage

cette règle. Les deux règles possibles sont explicités dans la section suivante.

2.2.2.3.2 Les règles de décodage

Les règles d'activation ont pour but de déterminer si un neurone sera activé ou non en fonction de son nombre de contributions actives et de celui des autres neurones. La règle *k-Winners-Take-All* (*WsTA*) cherche à activer les neurones avec le plus de contributions. Elle peut s'exprimer de la manière suivante :

$$\forall i, 1 \leq i \leq N_N : S_i = \begin{cases} 1 & \text{si } \nu_i \geq \theta_k, \\ 0 & \text{sinon} \end{cases} \quad (2.3)$$

où θ_k est le k^{ieme} élément de la liste des ν_i classés par ordre décroissant. Cette règle garantie que au moins k neurones seront activés dans le réseau. Afin de retrouver des cliques de taille c , le paramètre k est fixé à la valeur c . L'utilisation de cette règle nécessite donc un réseau avec des cliques ayant toute la même taille c connue.

La Figure 2.10 illustre le fonctionnement du décodage avec la règle du *WsTA*. Suite à la stimulation (neurones grisés), la liste des ν_i classés dans l'ordre décroissant est la suivante : [4; 3; 3; 3; 3; 3; 2; 2; 2; 2; 1; 1; 1; 1; 0; ...; 0]. La taille des cliques étant de cinq, la cinquième valeur de cette liste est utilisée afin de déterminer le seuil d'activation θ_k . Dans ce cas, le seuil vaut trois. Ainsi tous les neurones ayant au moins trois contributions actives sont activés. La deuxième itération commence donc avec six neurones activés. Cette fois, la liste des contributions est : [5; 5; 5; 5; 5; 2; 2; 2; 2; 0; ...; 0] et le seuil devient cinq. À la fin de cette itération seul les cinq neurones de la clique rouge sont activés. Le décodage est donc terminé et le réseau a correctement convergé vers la clique la plus proche de la stimulation initiale.

La règle du *WsTA* effectue une comparaison sur la totalité du réseau et nécessite la réalisation

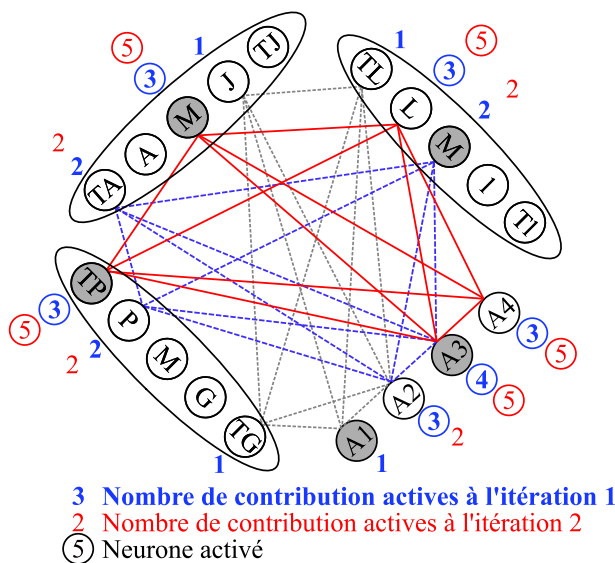


FIGURE 2.10 – Illustration du décodage avec la règle *WsTA*.

d'une opération de tri. Afin de simplifier les opérations, il est possible de réaliser cette règle en appliquant successivement une règle de *Winner-Takes-All* (*WTA*). Cette règle est équivalente à celle du *WsTA* pour $k = 1$. L'avantage de cette règle est qu'elle nécessite simplement un opérateur de maximum. En appliquant le *WTA* jusqu'à avoir au moins k neurones d'activés, une règle de k -*WsTA* est réalisée.

La règle des *Losers-Kicked-Out* est différente des règles cherchant des maxima. Cette règle s'effectue en trois phases et est illustré en Figure 2.11. La première phase est appliquée uniquement sur les neurones stimulés. Les neurones ayant le moins de contributions sont désactivés et les contributions de l'ensemble des neurones conservés sont recalculées. Cette étape est répétée tant que les ν_i des neurones encore actifs ne sont pas tous égaux. La valeur des ν_i à la fin de cette phase est enregistrée dans la variable θ . Dans l'exemple de la Figure 2.11, θ vaut trois. La deuxième phase consiste à propager la sortie des neurones gardés actifs à l'ensemble du réseau. L'activation des neurones dans tout le réseau se fait alors selon la règle suivante :

$$\forall i, 1 \leq i \leq N_N : S_i = \begin{cases} 1 & \text{si } \nu_i \geq \theta, \\ 0 & \text{sinon.} \end{cases} \quad (2.4)$$

A la fin de cette phase, une nouvelle phase d'élimination successive est effectuée. La règle se termine lorsque tous les neurones actifs possèdent le même nombre de contributions. Dans l'exemple, la phase trois est limitée à une seule itération puisque tous les neurones actifs ont le même nombre de contributions actives.

Cette règle ne nécessite pas de connaître la taille des cliques. Elle autorise donc l'utilisation de

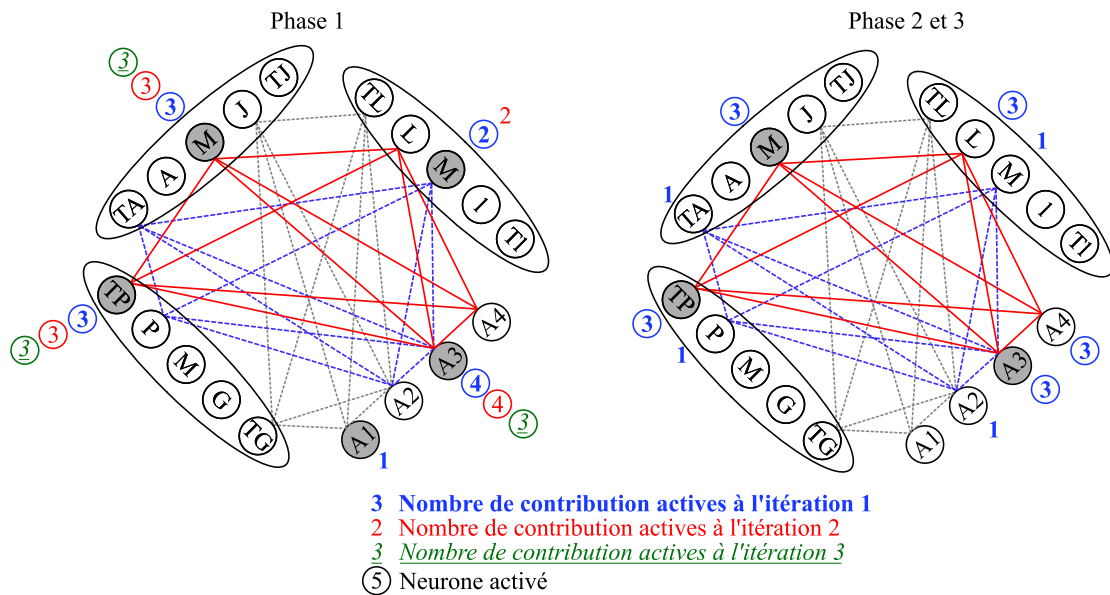


FIGURE 2.11 – Illustration du décodage avec la règle *LsKO*.

cliques de tailles différentes. Elle est aussi plus complexe que les règles précédentes du fait de son fonctionnement en trois phases nécessitant de multiples désactivations de neurones. Il est possible de désactiver plusieurs neurones à la fois afin d'accélérer le processus. Cela a pour conséquence de complexifier les opérations puisqu'un tri devient alors nécessaire afin de désactiver les k plus faible neurones. De plus, l'ensemble des neurones impliqués dans la règle est différent entre les phases. Cela peut complexifier l'architecture matérielle qui doit pouvoir sélectionner un sous-ensemble des neurones pour appliquer la règle. Cette règle est donc à privilégier dans une situation où la taille des cliques est variable ou inconnue.

2.2.3 Réseau à cliques pour la classification

2.2.3.1 Méthode de création des neurones

Le réseau à cliques de Willshaw-Palm présenté dans la Section 2.2.2 est utilisé afin de classifier les signaux ECG. Pour ce faire il est nécessaire de déterminer la topologie du réseau ainsi que le dictionnaire. L'objectif étant de détecter des anomalies dans le signal d'entrée, chaque type de battement sera stocké dans le réseau sous la forme d'une clique. Afin de pouvoir utiliser la règle de décodage du *WsTA*, les cliques sont conçues pour avoir la même taille. Suite à l'étape d'extraction de paramètres de la Figure 2.6, un vecteur de 18 paramètres réels est produit. Les neurones d'un réseau à cliques représentant une information spécifique binaire, il est nécessaire de définir une transformation pour passer du vecteur de 18 paramètres vers des informations unitaires. De la même manière que dans l'exemple de la Figure 2.7, il est possible de créer, pour chacun des 18 paramètres, des neurones représentant une plage de valeurs.

Ainsi, pour représenter les anomalies dans le réseau, chaque neurone correspond à une plage de valeurs pour un paramètre spécifique. Cette étude étant la première utilisation des réseaux à cliques en tant que classifieur, un cas simple est mis en place. Ainsi seuls trois types de battements sont considérés : le bloc de branche gauche (*Left Bundle Branch Block, LBBB*), le bloc de branche droit (*Right Bundle Branch Block, RBBB*) et le battement électro-entraîné (*Paced Beat, PB*). Ces battements ont été sélectionnés car ils offrent une bonne séparabilité des paramètres pour la création des neurones.

Dans un premier temps, l'ensemble de la base de données est segmentée afin d'obtenir des blocs de dix secondes de signal sans superposition. Cela signifie qu'il n'y a pas d'échantillons communs entre deux blocs de données consécutifs. Chaque bloc contient environ une dizaine de complexes QRS. Les annotations étant réalisées complexe par complexe il est nécessaire de décider de la classe pour un bloc. Pour ce faire chaque bloc présentant au moins une apparition d'une des trois anomalies est marqué comme présentant l'anomalie. Tous les autres blocs sont écartés puisqu'ils ne sont pas intéressants pour l'étude considérée. Le résultat de cette étape est un ensemble de 1572 blocs dont 486 *LBBB*, 523 *RBBB* et 563 *PB*. La décomposition en ondelettes est ensuite appliquée afin d'obtenir un jeu de données annotées de 1572 vecteurs de 18 paramètres.

Afin de créer le réseau de neurones, il est nécessaire de disposer d'une connaissance *a priori*.

Pour ce faire, une partie des données de la base est utilisée en tant que connaissance préalable. Cela équivaut à prendre un ensemble de données pour l'apprentissage pour les classifieurs classiques de type *SVM* ou perceptron. Le reste des données sera utilisées pour évaluer le pouvoir de généralisation du classifieur. Ainsi, pour chaque classe, 60 % des données sont utilisés en tant que connaissance *a priori*. Cette séparation conduit à l'utilisation de 291, 313 et 337 vecteurs pour respectivement les classes *LBBB*, *RBBB* et *PB*. Ces vecteurs sont utilisés pour construire le réseau à cliques. La mise en place des plages de valeurs est réalisée en analysant la distribution de chacun des 18 paramètres pour les trois classes. L'objectif est de créer des plages permettant de discriminer les trois classes pour chaque paramètres.

La Figure 2.12 montre la distribution pour le paramètre du maximum des coefficients $Cd1$ noté $\max(Cd1)$. Le nombre d'éléments pour chaque classe étant différent, le taux d'apparition est utilisé au profit du nombre d'occurrences. Dans cet exemple, trois plages de valeurs sont créées afin de discriminer les classes. Si les classes *LBBB* et *PB* sont presque totalement contenues dans les plages ① et ③, ce n'est pas le cas pour *RBBB*. Cette dernière est séparée entre les plages ① et ② qui contiennent respectivement 28.7 et 65 % des apparitions. L'effet de cette séparation est que 28.7 % des cas *RBBB* seront attribués à la mauvaise classe pour ce paramètre. Cela sera corrigé par la capacité du réseau à retrouver la bonne clique même si des erreurs sont insérées lors de la stimulation.

La Figure 2.13 présente l'exemple du paramètre $\max(Cd4)$ pour lequel il n'est pas possible de créer trois plages. En effet, les classes *LBBB* et *PB* possèdent des distributions très proches. Elles ne peuvent alors pas être discriminées via le paramètre considéré et seule la plage ① est créée pour représenter *RBBB*.

L'étape de création des plages est répétée pour les 18 paramètres et conduit à la décomposition

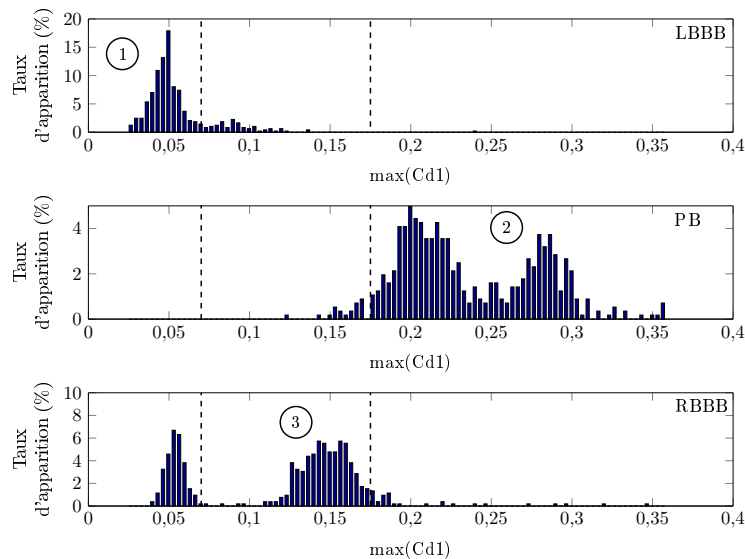


FIGURE 2.12 – Distribution de la valeur du paramètre $\max(Cd1)$ pour chacune des classes. La plage ① représente la classe *LBBB*, et les plages ② et ③ respectivement *PB* et *RBBB*.

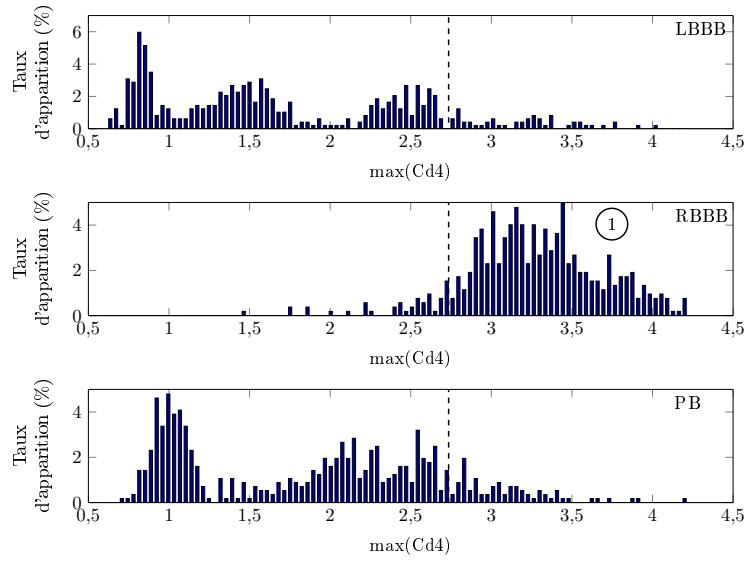


FIGURE 2.13 – Distribution de la valeur du paramètre $\max(\text{Cd}4)$. Une seule plage (1) est créée du fait de la forte similitude entre *LBBB* et *PB*.

TABLEAU 2.3 – Détail des plages utilisées pour caractériser les trois classes pour les 18 paramètres extraits.

Paramètre (x)	LBBB	RBBB	PB
$\max(\text{Cd}1)$	$x < 0,07$	$0,07 \leq x < 0,18$	$x \geq 0,18$
$\max(\text{Cd}2)$	$x < 0,33$	-	$x \geq 0,5$
$\max(\text{Cd}3)$	$x < 0,91$	$x \geq 1,64$	$0,91 \leq x < 1,64$
$\max(\text{Cd}4)$	-	$x \geq 2,74$	-
$\max(\text{Cd}5)$	-	-	$x \leq 1,63$
$\max(\text{Ca}5)$	-	-	-
$\min(\text{Cd}1)$	-	-	$x \leq -0,09$
$\min(\text{Cd}2)$	$x \geq -0,31$	-	$x < -0,31$
$\min(\text{Cd}3)$	$x \geq -0,68$	-	-
$\min(\text{Cd}4)$	-	$x \leq -1,67$	-
$\min(\text{Cd}5)$	-	-	-
$\min(\text{Ca}5)$	-	-	-
$\text{var}(\text{Cd}1)$	$x < 0,015$	-	$x \geq 0,015$
$\text{var}(\text{Cd}2)$	$x < 0,038$	-	$x \geq 0,056$
$\text{var}(\text{Cd}3)$	$x < 0,153$	$x \geq 0,153$	-
$\text{var}(\text{Cd}4)$	-	$x \geq 0,4$	-
$\text{var}(\text{Cd}5)$	-	-	-
$\text{var}(\text{Ca}5)$	-	-	-

présentée dans le Tableau 2.3. Cette séparation est effectuée manuellement et permet d'associer une classe à chaque neurone. Les cliques sont alors réalisées en connectant ensemble les neurones représentant la même anomalie. La décomposition proposée dans le Tableau 2.3 ne permet pas

d'avoir des cliques de tailles identiques. En effet, la classe *RBBB* n'utilise que six plages alors que les deux autres en utilisent huit. Pour pallier ce problème, il est possible de doubler des neurones qui sont seuls sur un paramètre. De plus, pour cinq paramètres, il n'a pas été possible de créer de plages pour discriminer les trois classes considérées. Ces derniers ne sont pas utilisés.

Le réseau ainsi obtenu est composé de 24 neurones et est illustré en Figure 2.14a. Il n'utilise que 13 des 18 paramètres car cinq ne sont pas suffisamment discriminants entre les classes. Chaque forme représente un paramètre et les numéros à l'intérieur identifient les différentes plages du paramètre. Afin que la clique *RBBB* ait une taille de huit, les neurones correspondant aux paramètres $\max(Cd4)$ et $\min(Cd4)$ ont été doublés. Ce choix est justifié par le fait que ce sont les deux plages qui permettent de mieux discriminer la classe *RBBB* des deux autres.

Le réseau proposé ne permet pas une lecture directe de la sortie. Pour obtenir la décision du réseau il est nécessaire de comparer l'ensemble des neurones activés avec les cliques. Cela demande des étapes de calcul en plus. La solution pour obtenir une réponse directe du réseau est l'utilisation du réseau de sortie de la Figure 2.14b. Il est constitué de trois neurones représentant chacun une classe. Chaque neurone du réseau cœur constituant une clique est connecté au neurone de sortie représentant cette dernière. Les connexions sont unidirectionnelles du réseau cœur vers le réseau de sortie. Le réseau de sortie ainsi formé peut être vu comme un réseau avec trois cliques de taille unitaire et dont la stimulation est l'état des neurones du réseau cœur. Avec une telle topologie, la sortie du réseau est un message de trois bits qui peut être simplement encodé sur deux bits pour la transmission à l'agrégateur.

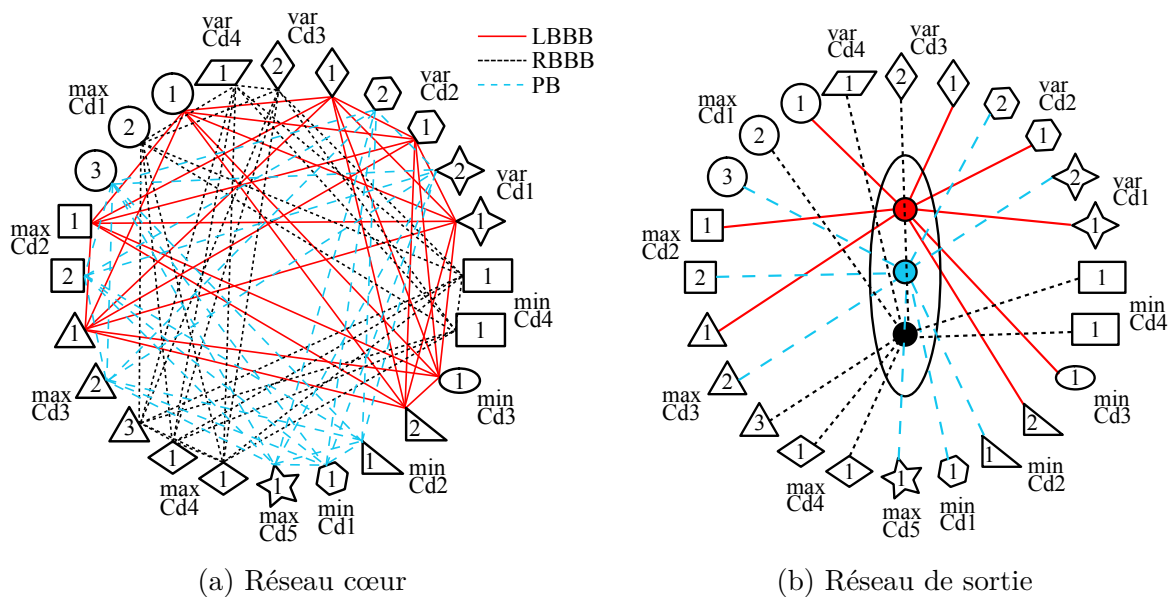


FIGURE 2.14 – Topologie du réseau proposé pour la classification d'ECG.

2.2.3.2 Résultats de classification

Le système proposé est simulé via un modèle *Matlab*[®] afin d'évaluer les performances de classification. Le test est effectué sur l'ensemble de la base de données en incluant les données ayant servi pour créer la connaissance *a priori*. Le nombre total d'éléments de chaque classe est de 486, 523 et 563 pour respectivement *LBBB*, *RBBB* et *PB*. Le modèle réalise d'abord la décomposition en ondelettes puis en extrait les paramètres. La stimulation du réseau est ensuite produite en créant un message de 24 bits avec chaque bit correspondant à un neurone. Si le paramètre est dans la plage du neurone alors le bit correspondant du message est mis à un, sinon il reste à zéro. La règle de décodage utilisée est celle du *WsTA* pour le réseau cœur et le *WTA* pour le réseau de sortie.

La matrice de confusion du Tableau 2.4 présente les résultats de classification obtenus. Chaque colonne représente la classe prédite par le système. Les lignes correspondent à la classe présente en entrée. La somme de tous les éléments de cette matrice est 1588. Cela ne correspond pas au nombre d'éléments testés qui est de 1572. Cette différence s'explique par le fait que certaines stimulations conduisent à une égalité entre deux classes. Cette matrice de confusion permet de constater que le système tend à favoriser la classe *RBBB*. Cela peut s'expliquer par le doublement de deux neurones pour cette classe. Si une observation de type *LBBB* stimule un des neurones doublés alors cela renforce doublement la clique *RBBB* qui est alors plus facilement dominante dans le réseau.

Les performances de classification du Tableau 2.5 sont calculées à partir de la matrice de confusion. Les performances moyennes sont calculées via une moyenne pondérée des classes par leur taux d'apparition dans l'ensemble testé. Ainsi le poids de la classe *LBBB* est $486/1572 \approx 0,309$, de même ceux des classes *RBBB* et *PB* sont respectivement 0,333 et 0,358. Ces performances de classification sont similaires à celles obtenues par les implantations présentées en Section 2.1.2.5. L'objectif de performances est donc bien atteint. Ces résultats ont fait l'objet d'une publication lors de la conférence *New Circuits and Systems (NEWCAS) 2016* [CCL⁺16].

TABLEAU 2.4 – Matrice de confusion du système.

		Classe prédite		
		<i>LBBB</i>	<i>RBBB</i>	<i>PB</i>
Classe d'entrée	<i>LBBB</i>	437	50	4
	<i>RBBB</i>	13	493	24
	<i>PB</i>	0	27	540

TABLEAU 2.5 – Performances de classification du système proposé.

	Sensibilité	Spécificité	Exactitude
<i>LBBB</i>	89%	98,8%	95,8%
<i>RBBB</i>	93%	92,7%	92,8%
<i>PB</i>	95,2%	97,2%	96,5%
Moyenne	92,6%	96,2%	95%

2.3 Discussion

Bien que le système proposé dans ce chapitre ait de bonnes performances de classification, ce dernier n'est capable d'identifier que trois cas de battements anormaux. Il est tout à fait possible d'imaginer vouloir ajouter de nouvelles classes à l'application. Pour ce faire, il est alors nécessaire de répéter la totalité du processus de création des neurones et des connexions. Ce dernier étant effectué à la main, cela demande beaucoup de temps et ne garanti pas une solution optimale lors de la mise en place des plages de valeurs. Cette forte limitation peut rendre l'utilisation de tels réseaux complexes pour des classifieurs avec de nombreuses classes ou de nombreux paramètres. Une solution envisagée afin de pallier cette limitation est de mettre en place un algorithme de création du réseau dont le but est de maximiser les performances tout en minimisant le nombre de neurones et la taille des cliques. De plus, l'ajout de nouvelles classes se traduit par l'augmentation du nombre de neurones, ainsi que la taille des cliques. Il sera donc intéressant d'étudier le comportement de l'implantation matérielle du classifieur avec l'augmentation du nombre de neurones et de la taille des cliques.

Le système proposé dans ce chapitre permet d'obtenir des performances de classification au niveau de l'état de l'art des solutions matérielles existantes. Afin de présenter un réel intérêt dans le cadre des *BSAN*, le système doit aussi présenter une faible consommation énergétique. L'objectif étant l'augmentation de l'autonomie du capteur, il est nécessaire que la consommation du système proposé soit inférieure à celle d'un capteur simple ne faisant aucun traitement. Pour se faire on considère un capteur faisant simplement l'acquisition et la transmission directe des données comme présenté en Figure 1.2a. Le convertisseur considéré est le CAN SAR sur dix bits de [YYKY10] présentant une consommation de 0,12 pJ/bit. L'émetteur de type *BLE* est présenté dans [KFC⁺17] et consomme 3,7 nJ/bit transmis. Avec un tel capteur, le traitement de dix secondes de signal équivaut à l'acquisition et à l'émission de 39 600 bits. On considère des signaux similaires à ceux de la base de données utilisée, soit un signal à 360 Hz sur 11 bits. L'énergie nécessaire pour l'acquisition et la transmission des dix secondes est d'environ 147 μ J. Pour être intéressant, le système proposé doit donc avoir une consommation inférieure à 147 μ J. Dans le système proposé, la fonction la plus complexe est celle de la décomposition en ondelettes. Une implantation *ASIC* récente proposée dans [WZL⁺15] présente un opérateur effectuant une *DWT* de type Daubechies 4 à trois niveaux sur 256 points avec des données de taille 16 bits pour une énergie de 192 nJ. Bien que la transformation du système proposé soit plus simple, cette dernière est effectuée sur 3 600 points au lieu de 256. Il est donc possible d'estimer la consommation de la *DWT* du système comme étant de l'ordre du μ J. Cette valeur est très inférieure à la limite de 147 μ J et permet de justifier que le système proposé peut être basse consommation. Le chapitre suivant présente le prototype du système afin de vérifier que ce dernier présente bien une faible consommation.

Conclusion

Dans ce chapitre, les méthodes traditionnelles de traitement pour l'analyse de signaux ECG sont présentées. Leurs très bonnes performances sont mitigées par une grande complexité pour l'implantation matérielle. Cette complexité rend l'utilisation des meilleures méthodes impossible pour la mise en place de capteurs intelligents qui nécessitent un traitement très basse consommation. Un système axé sur une décomposition en ondelettes et utilisant les réseaux de neurones à cliques est proposé pour classifier des arythmies sur les signaux ECG. L'exactitude de classification moyenne obtenue de 95 % est supérieure à celle obtenue par des systèmes d'analyses intégrés existant [BMD⁺14] [CMM⁺13] [LV13]. De plus, la simplicité des traitements proposés rendent une implantation *ASIC* basse consommation possible. Bien que le système ne considère que trois classes, il reste possible d'imaginer d'autres topologies de réseaux de neurones à cliques afin d'offrir un pouvoir de classification plus fort. Le système proposé permet d'avoir un capteur capable de fournir une aide pour des patients ayant besoin d'un suivi cardiaque continu sans devoir se déplacer régulièrement chez le médecin.

Bien que les fonctions mise en place dans le système soit simple, la faible consommation du système n'est pas évidente. Il convient donc d'évaluer la consommation énergétique des différents éléments du système. De plus, il peut être intéressant de voir comment cette consommation évolue avec la complexification du classifieur via l'ajout de nouvelles classes. C'est l'objet du chapitre suivant qui établit une architecture pour la mise en place d'un prototype du système proposé. La consommation du système est alors estimée à partir du prototype afin d'évaluer l'aspect basse consommation de la solution.

Chapitre 3

Prototypage du système d'analyse d'ECG basse consommation

Introduction

Ce chapitre présente la conception du système d'analyse proposé précédent. L'objectif premier est de mettre en place une architecture matérielle complète du système afin de valider son aspect basse consommation. Dans un premier temps, une architecture numérique pour le pré-traitement est proposée. Sa consommation est déterminée à partir de l'estimation de la consommation d'éléments de base fournit par *Synopsys Design Compiler*[®] pour la technologie *ST CMOS 65 nm*. Dans un second temps, une architecture s'appuyant autour d'un circuit mixte analogique/numérique et d'un *FPGA* est mise en place afin d'implanter le classifieur. L'utilisation d'un système mixte est justifiée par la faible consommation obtenue dans [Lar15] par rapport à une implantation numérique. Le circuit conçu est fabriqué et testé afin d'obtenir des mesures précises de la consommation du classifieur et de valider son fonctionnement. Enfin, la validité du système proposé en tant que capteur intelligent est discuté.

3.1 Architecture du pré-traitement

Dans cette section, une architecture des différents éléments de la chaîne est proposée dans le but d'estimer sa consommation. L'architecture globale de traitement est présentée sur la Figure 3.1. Le signal d'entrée, issu de la base de données MIT-BIH, est codé sur 11 bits et est échantillonné à 360 Hz. Suite à la décomposition en ondelettes, les paramètres sont extraits. Ces derniers sont ensuite mis en forme afin de pouvoir être utilisés par le classifieur à base de réseau à cliques. Cette architecture est pipelinée et vise un fonctionnement à la fréquence de d'échantillonnage de 360 Hz. Cela permet un traitement au fil de l'eau des données afin d'éviter de devoir stocker les données et, ainsi, réduire la consommation. Les architectures proposées dans cette section visent à avoir une estimation de la consommation. L'estimation globale est réalisée à partir de celle des éléments

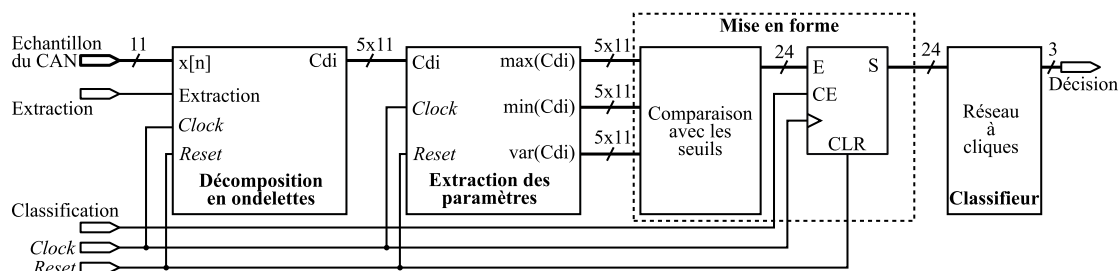


FIGURE 3.1 – Schéma bloc du système complet proposé pour la classification des ECG.

de base suivant : registre 11 bits, additionneur 11 bits et multiplieurs 11 et 22 bits. Ces derniers sont alimentés à la tension nominale de 1 V.

3.1.1 Décomposition en ondelettes

3.1.1.1 Filtrage

L'opération de décomposition en ondelettes discrète est réalisée via une structure de filtrage dyadique. Les coefficients des deux filtres pour l'ondelette *db2* utilisée sont présentés dans le Tableau 3.1. L'architecture utilisé pour réaliser un filtre est la forme directe de la Figure 3.2. Elle est constituée de quatre registres, quatre multiplieurs et trois additionneurs. Les données d'entrée sont considérées être sur 11 bits avec une représentation du type complément à un (CA1). C'est à dire que le premier bit représente le signe et les dix suivants l'amplitude.

Grâce à la représentation CA1 des données l'opération de multiplication peut être réalisée simplement avec des additions et des décalages. En effet, la multiplication, en valeur absolue, peut être décomposée comme une somme de multiplications élémentaires. Par exemple la multiplication $x \cdot 3$ peut s'écrire $x \cdot 1 + x \cdot 2$. La multiplication par une puissance de deux étant un décalage sur la représentation binaire, la multiplication peut être réalisée très simplement. Le nombre d'additions nécessaires correspond donc au nombre de uns dans la représentation binaire des coefficients. Pour obtenir le signe, il suffit alors de comparer les bits de signe.

Pour estimer la consommation du filtre, les coefficients des filtres sont donc représentés de manière binaire sur 11 bits en CA1. Avec une telle représentation des coefficients, il faut 19 additions et décalages par filtre. Afin de ne pas perdre en précision, ces additions sont réalisées sur 22 bits puis la sortie est tronquée sur 11. Le signal est ensuite converti en complément à deux (CA2)

TABEAU 3.1 – Coefficients des filtres pour l'ondelette *db2*.

Passe-haut	Passe-bas
$h_0 = -0,48296$	$g_0 = -0,12941$
$h_1 = 0,83652$	$g_1 = 0,22414$
$h_2 = -0,22414$	$g_2 = 0,83652$
$h_3 = -0,12941$	$g_3 = 0,48296$

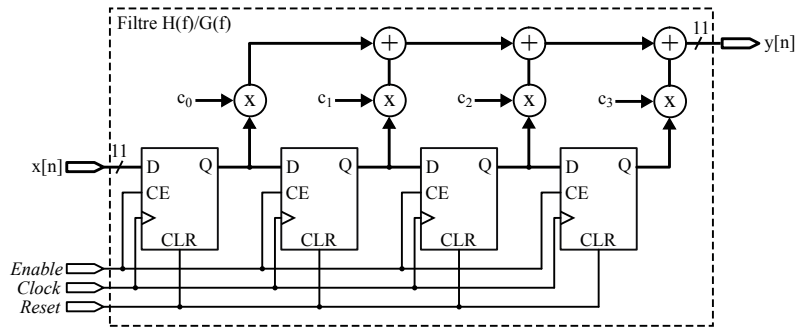


FIGURE 3.2 – Architecture de l'opérateur de filtrage.

afin de sommer facilement les résultats des multiplications. La conversion en CA2 consiste en une inversion de tous les bits puis à l'ajout de un. Elle n'est nécessaire que pour les nombres négatifs. Comme il n'est pas possible de savoir à l'avance les signes de sortie des multiplications, on considère que cette conversion est tout le temps effectuée. De plus, sa consommation est estimée être de l'ordre de celle d'un additionneur 11 bits. La consommation réelle sera inférieure mais cela n'est pas problématique puisque l'estimation réalisée est une estimation pessimiste.

Avec toutes ces considérations, il est possible d'estimer la consommation d'un filtre. Elle correspond à la consommation de quatre registres, 23 additionneurs 22 bits et trois additionneurs 11 bits. Cela correspond à une consommation statique de $P_s \text{ filtre} = 9,25 \text{ nW}$ et une consommation dynamique de $E_{dl \text{ filtre}} = 1,12 \text{ pJ/cycle}$ et $E_{dr \text{ filtre}} = 0,3 \text{ pJ/cycle}$ pour respectivement la logique et les registres.

3.1.1.2 Opérateur de décomposition en ondelettes

L'architecture réalisant la décomposition en ondelettes, présenté en Figure 3.3, est constituée de filtres, de registres et de bascules D. Seuls les trois premiers des cinq niveaux de décomposition sont représentés afin de garder la lisibilité de la figure. Suite à l'étude menée dans la Section 2.2.3.1, les coefficients d'approximation du cinquième étage ne sont jamais utilisés. Il n'est donc pas nécessaire d'effectuer le filtrage passe-bas sur le dernier niveau. C'est ce qui est illustré sur la Figure 3.3 avec trois niveaux.

Les blocs de filtrages ne contenant pas de registres en sortie, il est nécessaire d'en ajouter après les filtres passe-haut. Le sous-échantillonnage par deux est réalisé par l'utilisation d'une bascule D dont la sortie complémentée est rebouclée sur l'entrée. La bascule est commandée par l'horloge du système et effectue donc une division de l'horloge par deux. Sa sortie est utilisée comme signal *Enable* pour les registres des niveaux suivants. Cette sortie est aussi utilisée comme horloge pour la bascule D suivante, réalisant ainsi une division par quatre de la fréquence de départ. En répétant, cette structure il est possible de réaliser les sous-échantillonnages par deux successifs. Pour une décomposition en cinq niveaux, il y aura donc cinq bascules. La consommation globale de ces bascules est comptée comme celle d'un registre 11 bits car c'est l'élément dont elles se rapprochent le plus.

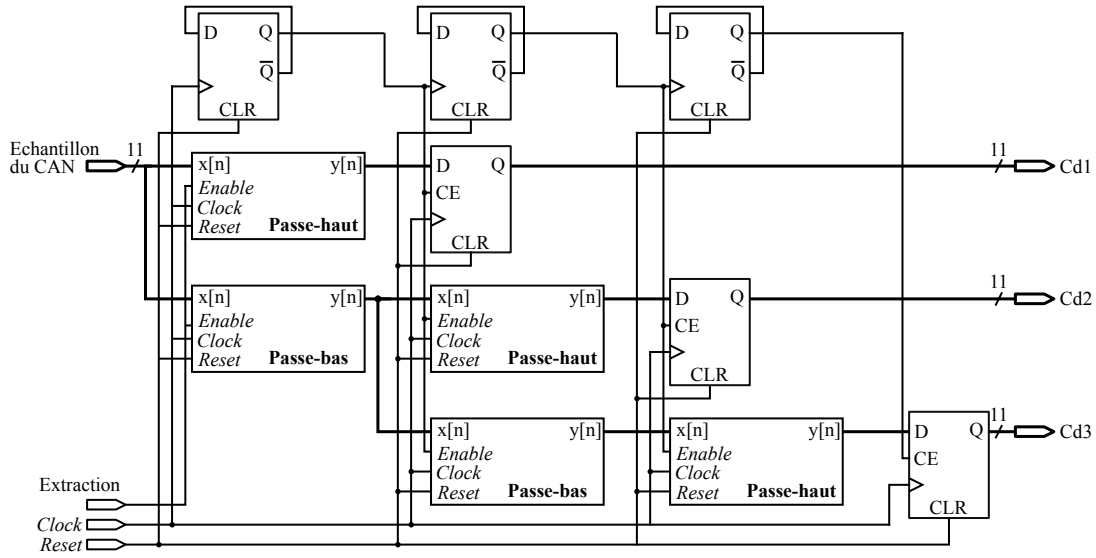


FIGURE 3.3 – Architecture réalisant la décomposition en ondelettes. Seuls les 3 premiers niveaux sont représentés.

Pour estimer la consommation de la décomposition en ondelettes, il est à noter que l'activité moyenne de la logique suite à une opération de filtrage est divisée par deux. Ainsi la consommation dynamique de la logique diminue avec les étages. Concernant les registres, les estimations obtenues ne différencient pas la consommation issue de l'activité de l'entrée et celle de l'arbre d'horloge la consommation dynamique de ces derniers est considérée maximale. Bien que cela ne reflète pas réalité, cela permet de s'assurer de ne pas sous-estimer la consommation. Ainsi la consommation statique de l'opérateur de décomposition en ondelettes est de :

$$\begin{aligned}
 P_{s\text{ ondelettes}} &= 9 \cdot P_{s\text{ filtre}} + 6 \cdot P_{s\text{ reg}} \\
 &= 84,74 \text{ nW},
 \end{aligned}
 \tag{3.1}$$

avec $P_{s\text{ reg}}$ la puissance statique d'un registre 11 bits. Les six registres supplémentaires correspondent aux cinq registres de sortie et à celui utilisé pour les bascules D. De même, la consommation dynamique des registres vaut :

$$\begin{aligned}
 E_{dr\text{ ondelettes}} &= 9 \cdot E_{dr\text{ filtre}} + 6 \cdot E_{dr} \\
 &= 3,13 \text{ pJ/cycle},
 \end{aligned}
 \tag{3.2}$$

avec E_{dr} la consommation dynamique d'un registre. Enfin la consommation dynamique de la

logique en tenant compte de la réduction d'activité est donnée par :

$$\begin{aligned}
 E_{dl\ ondelettes} &= E_{dl\ filtre} \cdot \left(\sum_{i=1}^4 2 \cdot A_{f_i} + A_{f_5} \right) & (3.3) \\
 &= E_{dl\ filtre} \cdot (2 \cdot (1 + 0,5 + 0,25 + 0,125) + 0,0625) \\
 &= 4,9\ pJ/cycle,
 \end{aligned}$$

avec A_{f_i} l'activité des filtres de l'étage i . L'énergie totale nécessaire pour effectuer la décomposition en ondelettes des dix secondes de signal à 360 Hz est donc :

$$\begin{aligned}
 E_{ondelettes} &= P_{s\ ondelettes} \cdot 10 + (E_{dr\ ondelettes} + E_{dl\ ondelettes}) \cdot 3600 & (3.4) \\
 &= 876\ nJ.
 \end{aligned}$$

3.1.2 Extraction des paramètres

3.1.2.1 Opérateur de minimum/maximum

L'extraction des paramètres consiste à calculer le maximum, le minimum et la variance des coefficients en sortie de chacun des niveaux de décomposition. L'opérateur de calcul du minimum ou du maximum est présenté sur la Figure 3.4. Il reçoit le coefficient issu de la décomposition et le compare avec la valeur courante du minimum ou maximum. Si besoin le résultat est mis à jour dans le registre, sinon il reste inchangé. La comparaison est effectuée en regardant le signe de la différence. Les données étant représentées en CA2, l'opération nécessite un changement de signe ainsi qu'une addition. Comme auparavant, le changement de signe est considéré consommer autant qu'un additionneur. L'opérateur de minimum/maximum nécessite donc un registre et deux additionneurs. Ainsi les différentes consommations pour un opérateur sont données par :

$$P_{s\ minmax} = 615\ pW \quad (3.5)$$

$$E_{dr\ minmax} = 75,68\ fJ/cycle \quad (3.6)$$

$$E_{dl\ minmax} = 19,7\ fJ/cycle \quad (3.7)$$

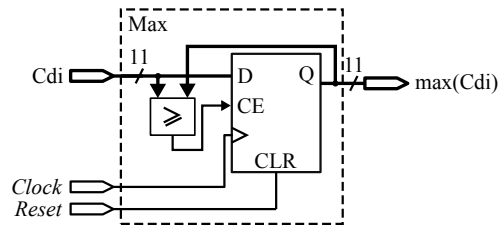


FIGURE 3.4 – Architecture de l'opérateur de minimum/maximum.

avec $P_{s\ minmax}$, $E_{dr\ minmax}$ et $E_{dl\ minmax}$ respectivement la puissances statique, l'énergie dynamique du registre et l'énergie dynamique de la logique d'un opérateur. Le paramètre Ca5 n'étant pas calculé, il y a donc un total de dix opérateurs : cinq minima et cinq maxima. La consommation statique ainsi que la consommation dynamique des registres sont donc multipliées par dix. En prenant en compte la réduction de l'activité de la logique, sa consommation devient :

$$\begin{aligned} E_{dl\ minmax\ tot} &= E_{dl\ minmax} \cdot \left(\sum_{i=1}^5 2 \cdot A_{li} \right) \\ &= E_{dl\ minmax} \cdot (2 \cdot (0,5 + 0,25 + 0,125 + 0,0625 + 0,03125)) \\ &= 38,1\ fJ/cycle, \end{aligned} \quad (3.8)$$

avec A_{li} l'activité de la logique de l'étage i .

La consommation totale des opérateurs de minimum/maximum pendant les dix secondes de traitement est donc :

$$\begin{aligned} E_{minmax\ tot} &= P_{s\ minmax} \cdot 100 + 3600 \cdot (E_{dr\ minmax} \cdot 10 + E_{dl\ minmax\ tot}) \\ &= 64,4\ nJ. \end{aligned} \quad (3.9)$$

3.1.2.2 Opérateur de variance

La dernière opération d'extraction de paramètres est le calcul de la variance. Afin de pouvoir exécuter un calcul pipeliné il est nécessaire d'utiliser la formulation de la variance suivante :

$$var(x) = E(x^2) - E(x)^2, \quad (3.10)$$

où $E(x)$ est l'espérance de x . Ce calcul requiert l'utilisation d'un opérateur de carré et d'un opérateur de division. Afin de simplifier l'architecture, l'étape de division n'est pas réalisée. La variance étant ensuite utilisée pour être comparée à un seuil fixé à l'avance, il est possible d'adapter ce dernier. Cela est réalisable car la valeur par laquelle il faut diviser est fixe et connue. L'architecture résultante de cette simplification est présentée sur la Figure 3.5. La branche supérieure effectue le carré de la somme tandis que la branche inférieure calcule la somme des carrés. Durant les calculs intermédiaires, le nombre de bits est augmenté afin d'assurer qu'il n'y ait pas de perte de précision. Le résultat final est ensuite tronqué afin de revenir sur une dynamique de 11 bits. De même qu'auparavant, l'opération de soustraction est réalisée via un changement de signe et une addition. Le changement de signe est de nouveau estimé comme une addition.

Le matériel nécessaire pour réaliser le calcul est de trois registres de tailles 11, 22 et 44 bits, trois additionneurs 44 bits et un additionneur 22 bits ainsi que deux multiplieurs avec des sorties

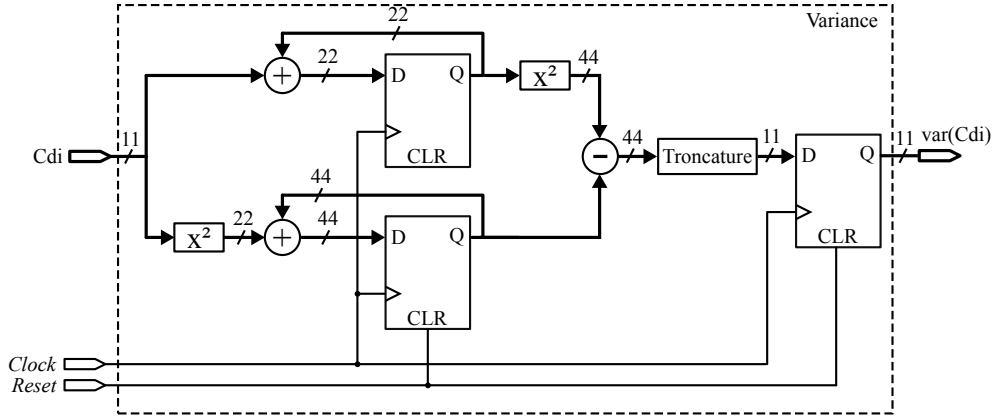


FIGURE 3.5 – Architecture simplifiée de l'opérateur de variance. L'étape de division n'est pas effectuée pour réduire la consommation.

de 22 et 44 bits. L'estimation de consommation d'un opérateur de variance est donc :

$$P_{s\ var} = 7,92\ nW \quad (3.11)$$

$$E_{dr\ var} = 530\ fJ/cycle \quad (3.12)$$

$$E_{dl\ var} = 656\ fJ/cycle, \quad (3.13)$$

où $P_{s\ var}$, $E_{dr\ var}$ et $E_{dl\ var}$ sont respectivement la consommation statique de l'ensemble, l'énergie dynamique des registres et de la logique. Le système étant limité à cinq étage, il y a un total de cinq opérateurs de variance. La consommation totale des cinq opérateurs est donc donnée par :

$$E_{var\ tot} = P_{s\ var} \cdot 50 + 3600 \cdot (E_{dr\ var} \cdot 5 + E_{dl\ var\ tot}) \quad (3.14)$$

$$= 408\ nJ, \quad (3.15)$$

avec

$$E_{dl\ var\ tot} = E_{dl\ var} \cdot \left(\sum_{i=1}^5 A_{l\ i} \right) \quad (3.16)$$

$$= E_{dl\ var} \cdot (0,5 + 0,25 + 0,125 + 0,0625 + 0,03125)$$

$$= 636\ fJ/cycle.$$

3.1.3 Mise en forme

La dernière étape avant le classifieur est celle de la mise en forme du signal. L'étape comparaison avec les seuils transforme les 15 paramètres en un message de 24 bits. D'après le Tableau 2.3, 17 comparaisons différentes sont nécessaires. Elles peuvent être réalisées simplement avec 17 comparateurs. L'estimation de cet élément est donc faite en considérant 17 additionneurs. De plus,

un registre de sortie est ajouté afin d'isoler le réseau de la partie extraction de paramètres. La consommation de ce registre 24 bit est estimée en considérant deux registres 11 bits. Ainsi la consommation pour effectuer la transformation est de $E_{tr\ tot} = 37\ nJ$.

Toutes ces estimations donnent une consommation du pré-traitement de 1385 nJ. Cela est très faible comparé aux 238 μJ nécessaires pour la transmission directe des données. Le dernier bloc du système est la classification via le réseau à cliques. La section suivante décrit l'architecture du système mixte réalisant le classifieur.

3.2 Conception du classifieur

L'objectif premier des travaux présentés dans cette section est de réaliser le classifieur du Chapitre 2. La conception doit permettre de reproduire les résultats de classification simulés tout en offrant une faible consommation. Afin de se placer dans un contexte général, le choix est fait de mettre en place une architecture flexible pour le classifieur. Cela permet à la fois de réaliser la topologie de réseau proposée dans le Chapitre 2 ainsi que d'autres topologies de réseaux à cliques. Il sera donc possible de tester d'autres réseaux suite à l'ajout de nouvelles classes pour la classification de signaux ECG ou pour d'autres applications utilisant ce type de réseau. La partie analogique du circuit reprend les travaux présentés dans [Lar15] en apportant des modifications afin de pallier les limitations. La réalisation d'un deuxième circuit implantant ce type de réseau de neurones offre l'opportunité d'améliorer sa conception ainsi que d'effectuer de nouvelles mesures.

3.2.1 Architecture mixte analogique/numérique du classifieur

La contrainte principale concernant l'architecture à mettre en place est la flexibilité de la structure et des connexions. Cela nécessite, en effet, d'être en mesure de réaliser toutes les connexions possibles entre les neurones. Il n'est pas envisageable de créer un ensemble de neurones totalement interconnectés car les connexions métalliques prendraient une place trop importante. En effet la surface nécessaire pour créer les connexions augmente avec le carré du nombre de neurones [Lar15]. De plus, rendre ces connexions paramétrables nécessite l'ajout d'un contrôle sur chacune d'elle, augmentant d'autant plus la surface. Afin de limiter la surface des connexions, [Lar15] propose une architecture reposant sur la réutilisation matérielle. Une version plus complète et plus détaillée de cette dernière est proposée sur la Figure 3.6.

Elle utilise un circuit *ASIC* contenant 128 neurones de 31 synapses chacun. L'utilisation de 31 synapses par neurone limite le réseau à un dictionnaire conduisant, au maximum, à 31 connexions entrantes par neurone. Le réseau à tester pour le système proposé étant de seulement 27 neurones avec huit connexions par neurones, il peut être réalisé avec l'architecture proposée. Un tel choix pour l'*ASIC* permet une réutilisation de ce dernier pour tester des classifieurs plus complexes suite à l'ajout de classes ou dans le cadre d'autres applications.

Sur l'*ASIC*, chaque synapse est contrôlée par une entrée numérique déterminant son état. L'*ASIC* effectue ensuite une opération de *WTA* sur l'ensemble des neurones. Ainsi les neurones avec

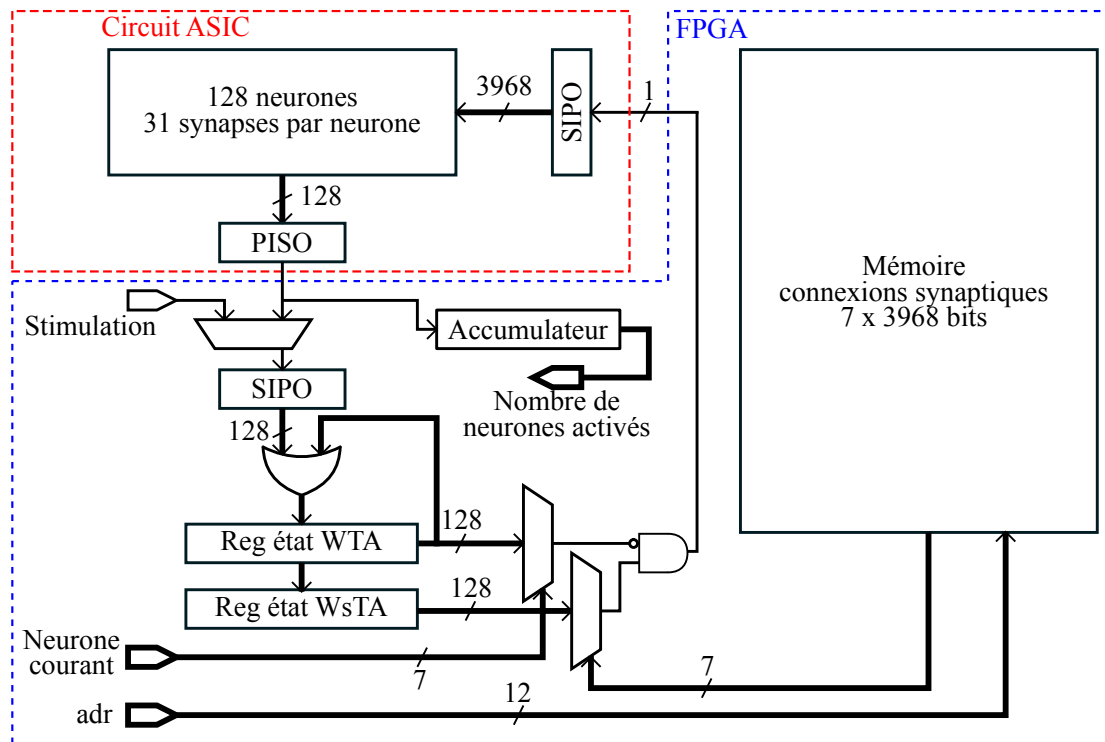


FIGURE 3.6 – Architecture du système mixte pour le classifieur.

le maximum de contributions actives sont activés. Pour réaliser une itération de décodage *WsTA* il est nécessaire d'effectuer plusieurs fois la règle de *WTA* en mémorisant les neurones activés. Pour se faire, un *FPGA* est utilisé afin d'effectuer la gestion des différentes itérations de règles *WTA*. Afin de réduire le nombre d'entrées/sorties, des registres à entrée série et sorties parallèles (*Serial In Parallel Out*, *SIPO*) ainsi que des registres à entrées parallèles et sortie série (*Parallel In Serial Out*, *PISO*) sont utilisés sur l'*ASIC*. L'utilisation du *FPGA* apporte une flexibilité dans le contrôle de l'*ASIC*. De plus, l'utilisation d'un *FPGA* permet une reconfiguration afin d'adapter la taille des différents éléments de l'architecture au réseau à implémenter.

Sur la Figure 3.6, les dimensions de la mémoire et des registres sont données pour le cas le plus gros. Le rôle du *FPGA* est de générer les entrées de l'*ASIC* lors des différentes étapes du décodage. Pour cela une mémoire et deux registres sont utilisés. Les registres servent à stocker l'état des neurones et la mémoire stocke les connexions réalisées dans le réseau. La structure de la mémoire des connexions synaptiques est présentée en Figure 3.7. Chaque ligne de cette mémoire correspond à une synapse du réseau en commençant par la première synapse du premier neurone (S_1, N_1). Les lignes suivantes représentent les autres synapses du premier neurone puis celles des différents neurones du réseau. Cela se traduit par une mémoire avec une profondeur de $31 \cdot 128 = 3968$ mots.

Chaque mot de la mémoire contient l'origine de la connexion sur la synapse. Cette connexion est unique et provient d'un neurone du réseau, elle peut donc être codée sur $\log_2(128) = 7$ bits. Ainsi chaque ligne de la mémoire contient l'indice du neurone à l'origine de la connexion.

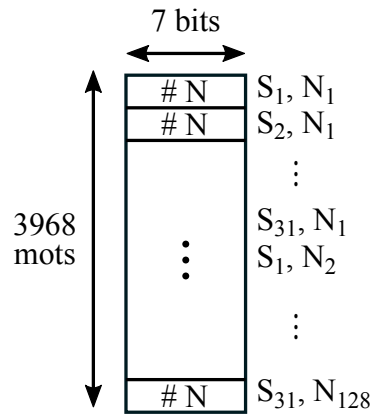


FIGURE 3.7 – Structure de la mémoire des connexions synaptiques.

Les registres se chargent de stocker l'état des neurones du réseau. Ils ont donc une taille maximale de 128. Ils sont au nombre de deux afin de stocker l'état des neurones entre les itérations de décodage *WsTA* mais aussi durant les opérations d'une itération.

Le fonctionnement du *FPGA* pour réaliser un décodage est le suivant. La stimulation est fournie au système via un message série par l'entrée Stimulation. Durant le chargement, les registres d'état ainsi que l'accumulateur sont remis à zéro. Le message d'entrée de 128 bits est ensuite stocké dans le registre état *WsTA* en passant par le registre d'état *WTA*. Le registre d'état *WTA* est alors forcé et à zéro et l'itération de *WTA* peut alors commencer. Elle est effectuée comme décrit dans l'algorithme 3.1.

L'état de chaque synapse de chaque neurone est calculé en regardant l'état du neurone à l'origine de la connexion. C'est ce qui est fait par le multiplexeur 128 vers un à la sortie de *Reg état WsTA*. Pour réaliser la règle *WsTA*, il est nécessaire de faire des règles *WTA* successives en éliminant les neurones activés lors des itérations précédentes. C'est ce qui est réalisé par la porte OU entre la sortie du réseau et *Reg état WTA* ainsi que le multiplexeur 128 vers un à sa sortie. Ainsi entre

```

Mettre à zéro adr
Pour tout neurone  $i = 1$  à 128
  Neurone courant =  $i$ 
  Pour tout synapse  $j = 1$  à 31
     $adr = adr + 1$ 
    Lire dans la mémoire des connexions
    Stocker l'état de la synapse
  Fin
Fin
Stimuler l'ASIC
Compter le nombre de neurones activés
Mettre à jour Reg état WTA

```

Algorithme 3.1 : Procédure d'une itération de *WTA*.

chaque itération *WTA*, les neurones activés sont stockés. Lors du calcul de l'état d'une synapse, si le neurone est déjà activé, cette dernière sera forcée à zéro par la porte ET. Cette séquence est répétée jusqu'à ce que le nombre de neurones activés en sortie de l'accumulateur soit égal à la taille des cliques. Dans ce cas, le contenu de *Reg état WTA* est copié dans *Reg état WsTA* et une nouvelle itération de *WsTA* peut commencer.

La section suivante présente comment réaliser le réseau à cliques proposé au Chapitre 2 avec cette architecture.

3.2.2 Mise en œuvre du classifieur avec l'architecture proposée

Le réseau à mettre en place est composé de deux sous-réseaux : le cœur et la sortie. Ils sont réalisés de manière successive et deux mémoires de connexions sont mises en place. Le réseau cœur est un réseau utilisant 24 neurones et stockant trois cliques de taille huit. Le nombre maximum de connexions entrantes par neurones est de huit. Ainsi il est possible de réduire la taille des registres d'état à 24 et celle de la mémoire des connexions M_1 à une largeur de $\lceil \log_2(24) \rceil = 5$ et une profondeur de $24 \cdot 8 = 192$ bits. Cela revient à une mémoire de 120 octets.

Le réseau de sortie est un réseau très simple de trois neurones. Les seules connexions dans ce réseau sont celles allant des neurones du réseau cœur vers le réseau de sortie. Ainsi, chaque neurone du réseau de sortie reçoit huit connexions du réseau cœur. Cela se traduit par une mémoire des connexions M_2 de profondeur 24 et de largeur cinq bits pour un total de 15 octets. L'état du réseau de sortie dépend seulement de l'état du réseau cœur. Il n'est donc pas nécessaire d'ajouter un registre d'état. La procédure pour réaliser un décodage est présentée dans l'algorithme 3.2. La sortie peut alors être obtenue soit directement à la sortie série de l'*ASIC* soit dans le registre d'état *WTA*.

L'architecture de contrôle sur *FGPA* offre donc la possibilité de mettre en place le classifieur pour les signaux ECG. La flexibilité via l'utilisation d'une mémoire permet la réalisation d'autres

```

Initialiser avec la stimulation externe
Pour toute itération  $it = 1$  à 3
    Mettre Reg état WTA à zéro
    Tant que Nombre de neurones activés < 8
        Faire décodage WTA avec  $M_1$ 
    Fin
    Remettre à zéro l'accumulateur
    Copier Reg état WTA dans Reg état WsTA
Fin
Copier Reg état WTA dans Reg état WsTA
Mettre Reg état WTA à zéro
Faire décodage WTA avec  $M_2$ 
Fin

```

Algorithme 3.2 : Procédure de décodage complet d'une stimulation.

topologies de réseaux à cliques. Afin de fonctionner correctement, il faut que l'ASIC soit capable de réaliser la règle de *WTA*. La section suivante présente la conception de ce dernier.

3.2.3 Conception analogique du réseau

La partie analogique du réseau est un circuit de 128 neurones de 31 synapses chacun effectuant une règle de *WTA*. Pour concevoir un tel circuit, les travaux de [Lar15] sont utilisés comme base. Une implantation de neurone pour les réseaux à cliques y est proposée. Les tests réalisés sur le circuit [CLL⁺17] [LCL⁺17] ont permis d'en identifier les limites. Les travaux de conception consistent à modifier le circuit existant afin de pouvoir réaliser la règle *WTA* avec 128 neurones. Les modifications apportées au neurone sont d'abord présentées puis l'architecture pour l'implantation de 128 neurones est exposée.

3.2.3.1 Circuit de neurone existant

Les fonctions réalisées par le neurone sont présentées dans la Figure 2.8. Elles sont au nombre de trois : somme des contributions binaires, comparaison des sommes et décision de l'activation. Le circuit de la Figure 3.8 est proposé par [LLAS13] afin de réaliser ces fonctions.

Les entrées et la sortie des neurones sont des signaux numériques. À l'intérieur du neurone, le fonctionnement se fait en mode courant. Les entrées numériques sont transformées en courant unitaire via le banc de synapses. Chaque synapse est une source de courant réalisée par le miroir ($M_5 ; M_6$) dont la sortie est activée ou non via M_7 . Le transistor d'entrée M_5 du miroir est commun à toutes les synapses. Le transistor M_7 est commandé par le signal d'entrée numérique. Le courant en sortie de chaque synapse est donc soit nul, soit I_{unit} . Dans [Lar15] le courant I_{unit} est fixé à 300 nA. Du fait d'un fonctionnement en mode courant, l'opération de somme est réalisée par un simple nœud au point *A*. Ce courant est ensuite recopié dans le circuit *WTA* via le miroir ($M_3 ; M_4$).

Le circuit *WTA* est composé du transistor M_1 et de M_2 monté en diode. La tension de sortie de cet élément est la tension V_B . Tous les neurones sont placés en parallèle et sont connectés

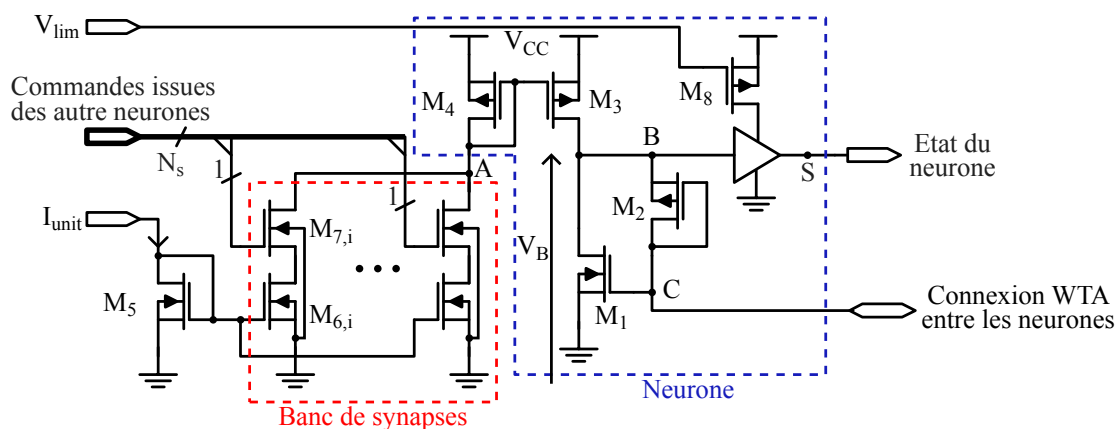


FIGURE 3.8 – Circuit du neurone proposé et implanté dans [Lar15].

ensemble au point C . Tous les transistors M_1 partagent donc la même tension grille-source V_{GS1} . Le transistor M_1 de la branche avec le plus grand courant fonctionne en régime saturé et impose sa tension V_{GS1} . Les autres transistors M_1 étant parcourus par un courant plus faible seront dans un régime non saturé ou en plus faible saturation. De plus, seule la diode M_2 du neurone vainqueur sera passante, les autres seront bloquées.

En appliquant un seuil sur la valeur de la tension V_B , il est possible de déterminer le neurone avec le plus de contributions en entrée. La comparaison est réalisée avec un buffer limité en courant. Cette limitation modifie sa tension de basculement. Le contrôle du courant dans le buffer via la tension de grille V_{lim} du transistor M_8 permet de régler la valeur du seuil. Dans le circuit de [Lar15], ce seuil est de 400 mV. La tension au point S correspond à la sortie numérique du neurone.

Le circuit [Lar15] est une preuve de concept de la réalisation des réseaux à cliques en mixte. Pour cette preuve de concept, le nombre de neurones mis en parallèle est de six, et chaque neurone possède 13 synapses. Le réseau intégré n'est pas flexible, les connexions sont directement réalisées via des pistes métalliques. L'objectif du circuit étant de réaliser une opération de WTA , le circuit doit être capable de différencier deux neurones avec une contribution d'écart. Cela se traduit par une différence de courant de I_{unit} entre deux neurones. Cette différence doit être détectée quel que soit le nombre de contributions du neurone vainqueur.

La résolution est définie comme la différence minimale en nombre de contributions que le circuit peut correctement détecter. Elle s'exprime en fonction du nombre de contributions du neurone vainqueur. Pour l'obtenir, les tensions V_B de deux neurones sont simulées pour avec une différence de une contribution entre les neurones. La Figure 3.9 présente cette évolution pour le circuit proposé dans [Lar15]. Chaque réseau de courbes présente les conditions de simulations suivantes :

- TT : process typique, température de 27°C , et alimentation de 1 V.

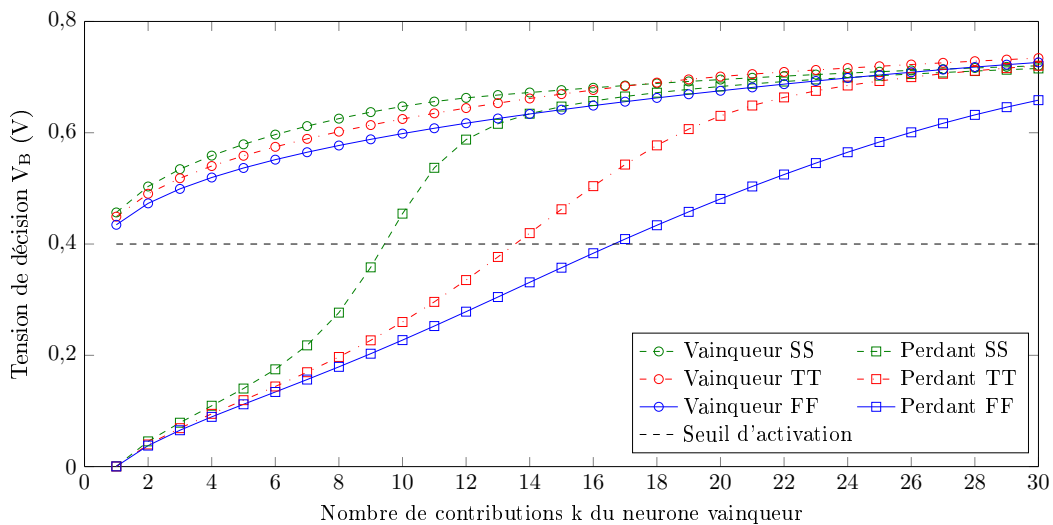


FIGURE 3.9 – Simulation du neurone de [Lar15] pour une différence entre vainqueur et perdant de I_{unit} .

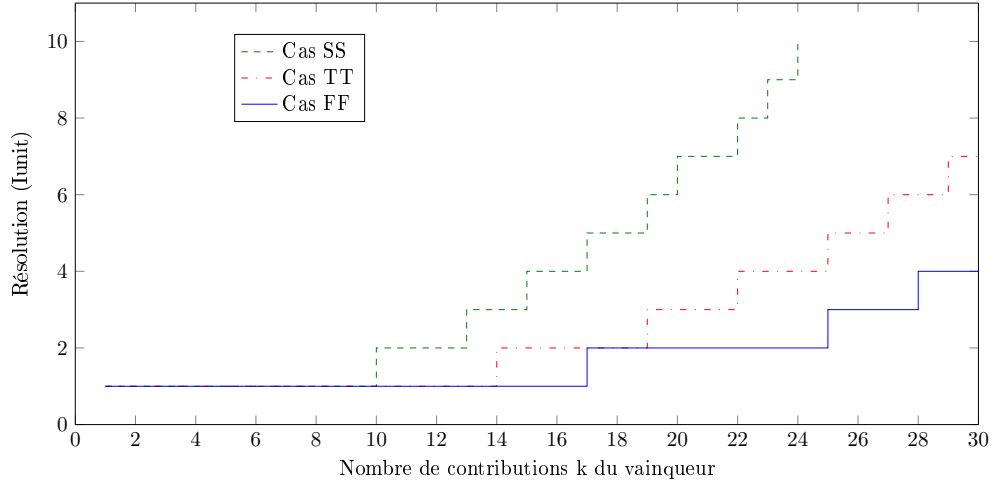


FIGURE 3.10 – Évolution de la résolution avec le nombre de contributions.

- *SS* : process *Slow* pour tous les transistors, température de 80°C et alimentation de 0,9 V.
- *FF* : process *Fast* pour les transistors, température de 0°C et alimentation de 1,1 V.

Le neurone gagnant possède k contributions actives et le perdant $k - 1$. On constate que dans le cas *SS*, le circuit n'est plus capable de détecter une différence de I_{unit} à partir de dix contributions actives. La limite dans le cas le plus favorable est de 16. Des simulations similaires ont été effectuées afin d'obtenir les courbes de résolution de la Figure 3.10. Si cette résolution est convenable pour le réseau proposé dans [Lar15], elle est trop faible pour l'*ASIC* désiré pour le prototype. L'*ASIC* à concevoir ayant 31 synapses par neurone, il est nécessaire d'avoir une faible résolution sur l'ensemble de la plage [1;31] contributions.

3.2.3.2 Modification du circuit du neurone

Afin d'améliorer la résolution du circuit il est nécessaire d'augmenter la différence de tension V_B entre deux neurones ayant une contribution de différence. Pour maximiser cette différence, il faut que la tension V_B du neurone vainqueur soit la plus proche possible de la tension limite de saturation du transistor M_1 . Du fait du V_{GS1} commun, un autre neurone parcouru par un courant plus faible sera obligé d'abaisser sa tension V_B . En se plaçant dans la branche du neurone gagnant, on a :

$$V_{DS1} = V_{GS1} + V_{DS2} \quad (3.17)$$

avec V_{DS2} la tension aux bornes de la diode M_2 . De plus,

$$V_{GS1} = V_{DS1,SAT} + V_{th} \quad (3.18)$$

avec $V_{DS1,SAT}$ la tension de saturation du transistor M_1 et V_{th} sa tension de seuil. Ainsi pour avoir une tension V_B proche de la tension de saturation il est préférable d'utiliser un transistor M_1 ayant

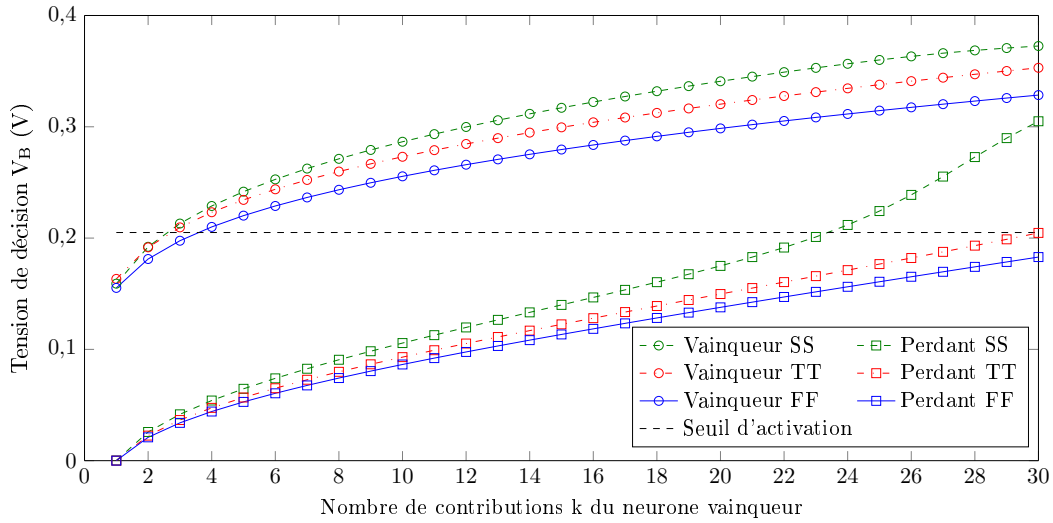


FIGURE 3.11 – Simulation du neurone modifié pour une différence entre vainqueur et perdant de I_{unit} .

un V_{th} faible. Avec la technologie utilisée, les V_{th} possibles sont de l'ordre de 150, 200 et 300 mV selon le type de transistor.

Le circuit a donc été modifié afin de changer le type de transistor pour M_1 . De plus, le nombre de contributions pouvant être plus grand, le miroir ($M_3 ; M_4$) est adapté afin d'être dimensionné pour un courant de $15 \cdot I_{unit}$. La Figure 3.11 montre les nouvelles tension V_B suite aux modifications apportées. La différence entre les deux neurones est toujours de I_{unit} et le nombre de neurones en parallèle reste de six. On peut constater que suite à l'abaissement du V_{th} , les tensions sont plus faibles. Il est donc nécessaire d'adapter le seuil de décision afin d'avoir le fonctionnement désiré. Une valeur possible pour le seuil est de 205 mV. Avec un tel seuil, le circuit est capable de correctement détecter le vainqueur jusqu'à 30 contributions dans les cas TT et FF . Dans le cas SS la limite devient 24, c'est une amélioration de 14 contributions par rapport au circuit original.

Il est aussi à noter que le nombre minimal de contributions pour fonctionner est de trois en SS et TT et quatre en FF . Afin de pallier ce problème, il est possible de modifier le miroir d'une des synapses en ajoutant un gain. Il sera alors possible de choisir d'utiliser cette dernière pour la réalisation de petits réseaux avec moins de quatre connexions par neurones. Lors de l'utilisation de cette synapse améliorée, la limite haute de fonctionnement est réduite.

L'évolution de la résolution de ce nouveau neurone est illustré en Figure 3.12. Ces courbes ne considèrent pas l'utilisation des synapses améliorées. Une valeur de la résolution de zéro correspond à un neurone gagnant qui ne s'active pas. Cela correspond à un fonctionnement incorrect du circuit. On constate que la modification apportée permet d'avoir une résolution de un sur l'ensemble de la plage [4;30] pour FF et TT . Pour 30 contributions, la résolution en SS est de trois. C'est une amélioration par rapport au neurone original qui avait une résolution supérieure à dix.

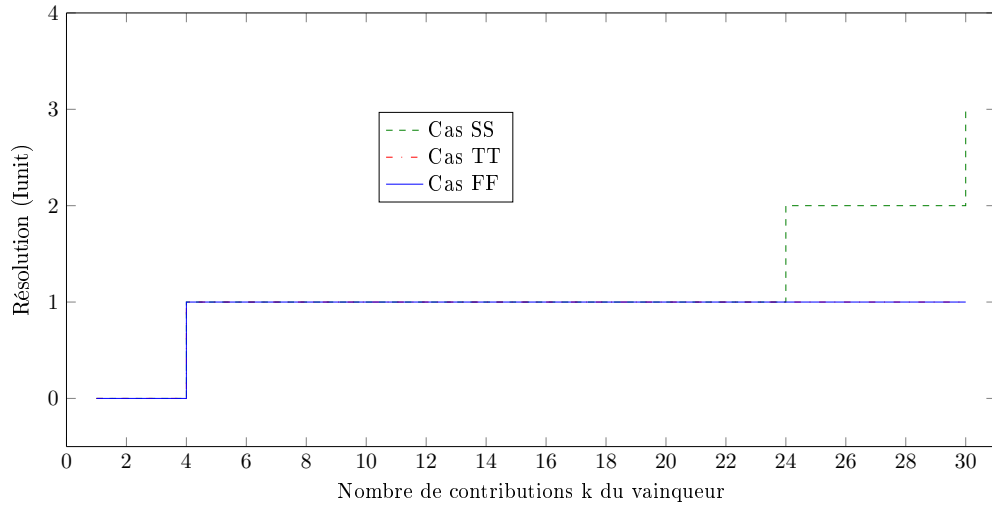


FIGURE 3.12 – Évolution de la résolution avec le nombre de contributions suite aux modifications.

3.2.3.3 Conception du réseau

Suite à la modification du circuit du neurone, l'étude de ce dernier pour le passage à l'échelle est réalisée. En effet, les simulations effectuées dans la section précédente sont réalisées avec seulement six neurones en parallèle. Avec l'augmentation du nombre de neurones, le nombre de diodes connectées en C augmente aussi. Ces dernières n'étant pas idéales, un courant de fuite proportionnel au nombre de neurones est prélevé à l'entrée du neurone vainqueur. Ce courant de fuite étant prélevé uniquement au neurone vainqueur, la différence de courant avec les autres neurones est réduite. Cela se traduit par une dégradation de la résolution. De plus, l'augmentation de ce courant de fuite conduit à une augmentation de la tension V_{DS2} aux bornes de la diode M_2 et donc de V_B d'après l'équation (3.17). La tension de saturation du transistor M_1 s'éloigne donc de la limite de saturation, conduisant à une dégradation de la résolution. On constate donc que l'ajout de neurones est doublement pénalisant. Les simulations de la mise en parallèle de 128 neurones de la Figure 3.13 confirme cette analyse. Seul le cas *SS* qui est le plus défavorable est illustré.

La dégradation de la résolution est accompagnée d'une forte augmentation du temps de réponse des neurones. En effet, pour chaque neurone en plus, le nombre de grilles connectées au point C augmente de deux. L'établissement de la tension V_{GS2} dans les neurones est donc ralenti. La Figure 3.14 montre le temps de réponse dans le pire cas pour des réseaux de six et 128 neurones. Les seuils de décisions sont adaptés afin de prendre en compte l'augmentation de la tension de décision avec l'ajout de neurones. Le temps de réponse est défini comme le temps nécessaire au réseau pour déterminer le vainqueur. On peut constater que pour un réseau de six neurones le temps de réponse est $T_{r,6} \approx 10$ ns et qu'il augmente à $T_{r,128} \approx 200$ ns pour 128 neurones.

Si le temps de réponse n'est pas critique pour l'application de classification de signaux ECG, il est toujours préférable d'avoir un système rapide. En effet, plus la réponse est rapide, plus l'alimentation peut être coupée tôt, et ainsi réduire la consommation du système. De même si

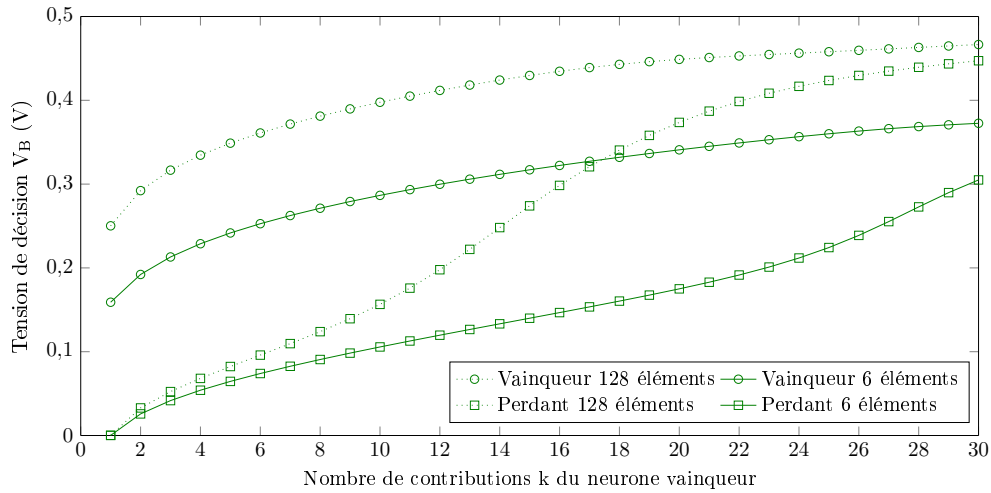


FIGURE 3.13 – Effet de l’augmentation du nombre de neurones en parallèles sur la tension V_B .

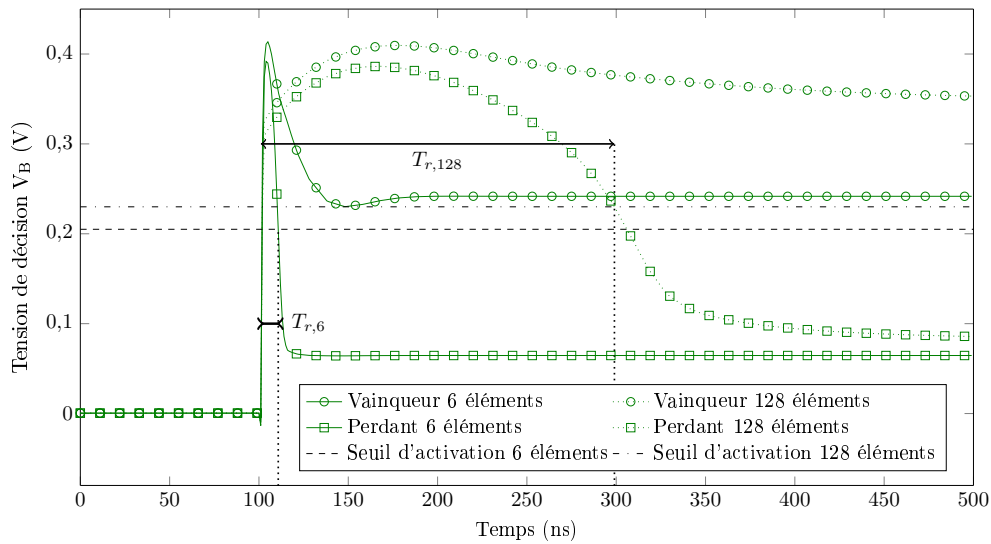


FIGURE 3.14 – Effet de l’augmentation du nombre de neurones sur le temps de réponse.

la dégradation de la résolution du réseau avec 128 neurones en parallèle ne rend pas impossible l’application souhaitée, elle limite fortement la flexibilité du système.

Pour contourner ces limitations, une architecture à base de bloc $WTA-k$ est mise en place. Un exemple de bloc WTA de k neurones est illustré sur la Figure 3.15. Ce bloc reçoit k courants d’entrée qui correspondent à la somme des courants issus des bancs de synapses. En plus de connecter ensemble les neurones, le bloc met en place le conditionnement de la sortie via le signal *Reset*. Ce signal met à zéro les sorties numériques de chaque neurone et connecte le point C à la masse via l’interrupteur M_9 . Cette mise à la masse permet de réinitialiser le WTA entre deux utilisations.

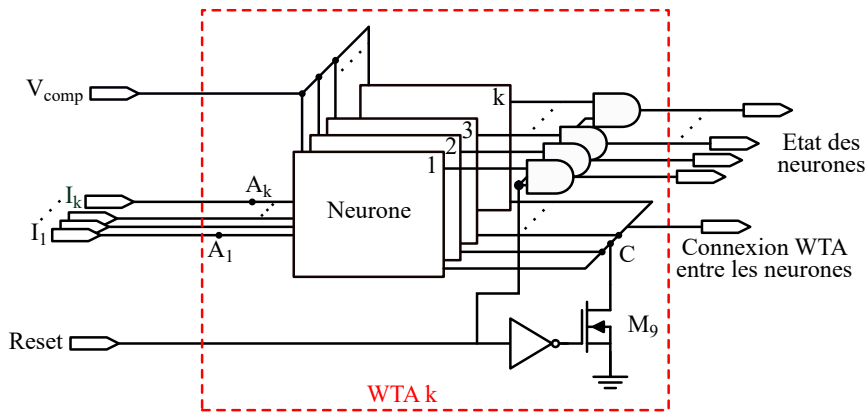


FIGURE 3.15 – Mise en parallèle de k neurones pour former un bloc $WTA-k$.

L'architecture proposée pour réaliser l'opération de WTA sur 128 éléments est présentée en Figure 3.16. Un premier étage de huit blocs $WTA-16$ cherche les maxima locaux. Le courant vainqueur de chacun de bloc est recopié en entrée du deuxième étage qui cherche le plus grand des maxima locaux. La recopie du courant vainqueur en sortie de chaque bloc $WTA-16$ est effectuée par le miroir ($M_1 ; M_{10}$). Les transistors M_{10} et M_1 étant identiques, le déblocage de la diode M_2

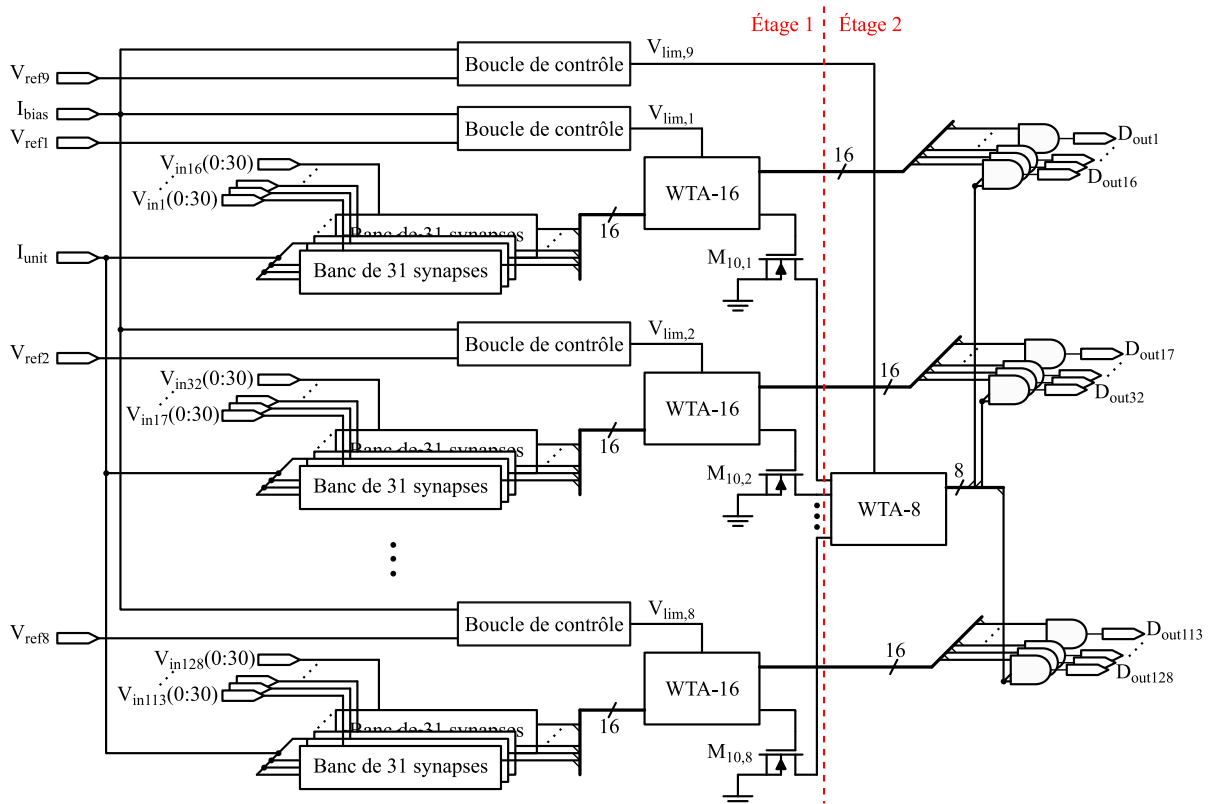


FIGURE 3.16 – Circuit réalisant le WTA sur 128 neurones. Deux étages sont utilisés afin de réduire le nombre d'éléments en parallèle.

dans la branche gagnante créé un miroir de courant. Pour être actif, un neurone doit être le plus fort localement et appartenir au bloc gagnant. La sortie de chaque neurone est donc obtenue via une opération de ET logique entre la sortie du bloc *WTA-16* et celle du *WTA-8* correspondant.

Une boucle de contrôle est utilisée pour chaque bloc afin d'assurer une valeur correcte pour le seuil de décision des neurones. Cette boucle de contrôle est identique à celle présentée dans [Lar15] et permet de régler la tension V_{lim} des neurones. La tension V_{ref} à fournir correspond à la tension de seuil désirée dans le buffer de décision des neurones. Cette tension est de 220 mV pour les blocs *WTA-16* et de 190 mV pour le bloc *WTA-8*. Avec un tel choix de bloc *WTA*, le nombre minimal de contributions pour activer un neurone est de trois. Ainsi la synapse améliorée est réglée de manière à valoir trois contributions.

Le circuit contenant 128 neurones de 31 synapses chacun, le nombre total d'entrées du circuit est de 3968. Afin de réduire le nombre d'entrées/sorties, des registres *SIPO* sont utilisés pour chaque bloc de 31 synapses connecté à un *WTA-16*, conduisant à huit registres *SIPO* de taille 496. De même un registre *PISO* est utilisé afin de récupérer les 128 sorties numériques du réseau. Les registres *SIPO* de la Figure 3.17a sont contrôlés par les signaux LR_{in} et L_{in} . Le signal LR_{in} active le décalage des données dans le registre et L_{in} permet d'envoyer les données du registre vers le réseau. Lorsque L_{in} est à zéro, les entrées du réseau sont forcées à zéro. De même, le signal L_{out} de la Figure 3.17b permet de sélectionner le signal à l'entrée des bascules du registre *PISO*. Lorsqu'il est inactif, les tensions $D_{out,i}$ sont copiées dans le registre. Lorsqu'il est actif, le registre est isolé des sorties du réseau. Le signal LR_{out} commande le décalage du registre *PISO* afin de sortir les données sur *Serial_out*. Enfin le signal $Reset_{reg}$ force à zéro les registres *SIPO* et le registre *PISO*. Pour chaque *SIPO* les données sont présentées dans en commençant par les synapses du dernier neurone puis l'avant-dernier et ainsi de suite. De même sortie du registre *PISO* est dans l'ordre opposé à celui des neurones. Ainsi la première sortie série du registre est celle du neurone 128 et la dernière celle du premier neurone.

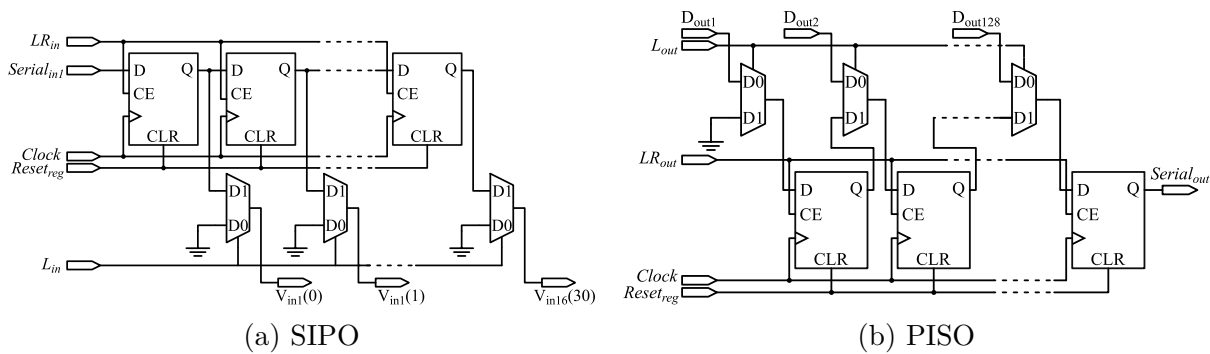


FIGURE 3.17 – Architecture des interfaces numériques.

3.2.3.4 Simulations du réseau

Il n'est pas possible de faire des simulations exhaustives du système complet de part sa complexité. Ainsi, seuls des cas limites sont testés. La Figure 3.18 montre l'évolution des signaux de contrôle pour une stimulation des neurones un et huit du premier bloc *WTA-16*. Le premier neurone est stimulé avec trois contributions et le huitième avec une seule. La période d'horloge est de 10 ns. La phase de chargement consiste à stocker les entrées dans les registres *SIPO*. Les données sont ensuite basculées vers le réseau via le signal L_{in} . Simultanément, le cœur analogique du réseau est activé via le signal *Reset*. Le réseau calcule alors les sorties durant la phase de convergence. Cette dernière a une durée T_{conv} de 500 ns. Suite à cela, les données sont récupérées dans le registre *PISO* en activant les signaux L_{out} et LR_{out} . Ce dernier est activé une période d'horloge plus tôt que L_{out} . Durant la période de récupération, la sortie du réseau peut être lue en série sur *Serialout*. Enfin le réseau et les registres sont remis à zéro en attendant la stimulation suivante.

Les Figures 3.19 et 3.20 présentent l'évolution des signaux dans les deux étages du circuit. Les courants I_{in} et I_{in2} correspondent respectivement aux courants totaux entrant dans les neurones des deux étages respectifs. De même V_B et V_{B2} sont les tensions aux points *B* des neurones du premier et deuxième étage respectivement. La Figure 3.19 montre les deux neurones stimulés afin de comparer le comportement de ces derniers. Le courant d'entrée est bien d'environ $3 \cdot I_{unit}$ pour le neurone gagnant et de I_{unit} pour le perdant. La tension de sortie, V_B , reste supérieure au seuil de décision pour le gagnant alors que celle du perdant retombe rapidement proche de zéro. Sur la Figure 3.20, le gagnant correspond au signal arrivant du premier bloc *WTA-16* et le perdant à tous les autres. Le courant gagnant du premier bloc est bien recopié en entrée du deuxième. Comme il n'y a aucune stimulation sur les neurones des autres blocs, le courant du perdant est réduit à un courant parasite. La tension de seuil du bloc *WTA-8* étant de 190 mV, le fonctionnement

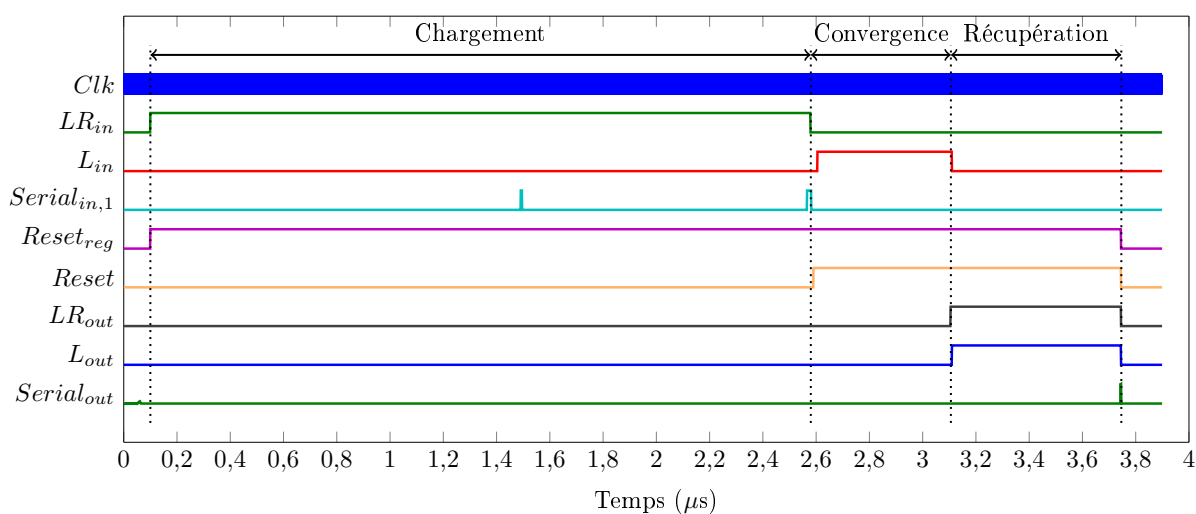
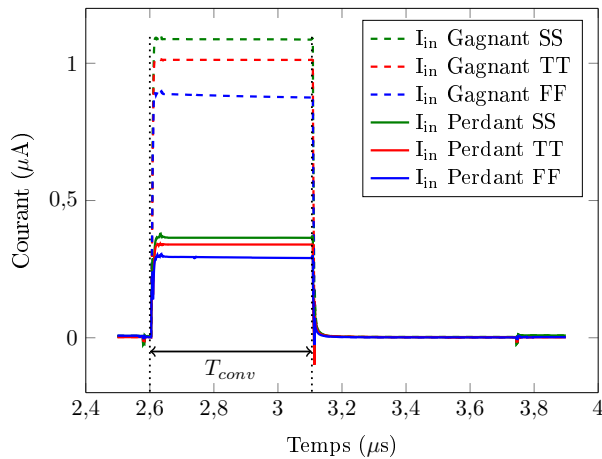
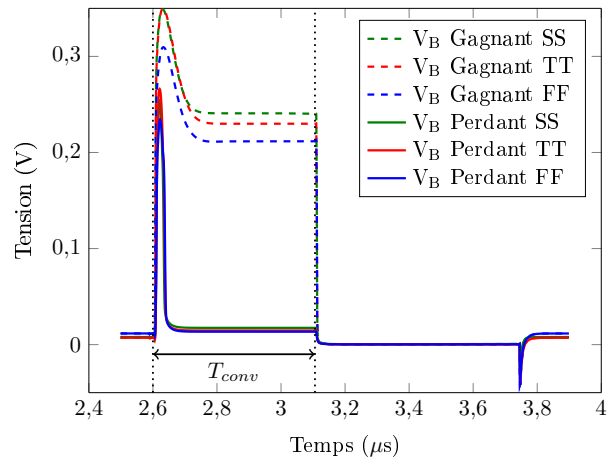


FIGURE 3.18 – Signaux de contrôle pour pour une stimulation des premier et huitième neurones avec respectivement trois et une contributions.

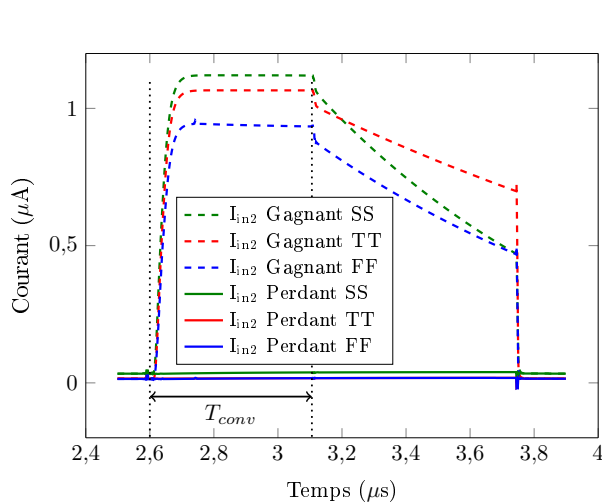


(a) Courant en entrée des neurones stimulés

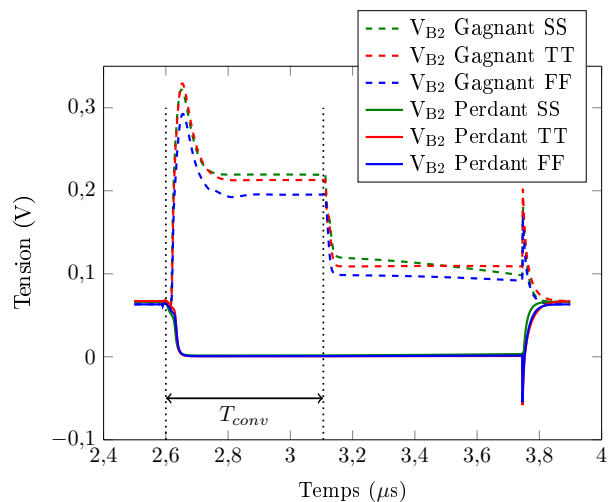


(b) Tension de sortie des neurones

FIGURE 3.19 – Évolution des signaux dans le premier bloc *WTA-16*.



(a) Courant en entrée du deuxième étage



(b) Tension de sortie correspondante

FIGURE 3.20 – Évolution des signaux dans le bloc *WTA-8*.

de ce dernier est validé en simulation. Le comportement observé à partir de la fin de la période de convergence correspond au comportement suite à la désactivation des entrées (passage à zéro de L_{in}). Les données étant alors déjà copiées dans le registre de sortie, ce comportement n'est pas souhaité puisqu'il induit une consommation inutile du circuit. Afin d'éviter cela, le signal d'activation du cœur analogique *Reset* peut être désactivé dès la fin de la période de convergence. C'est qui est fait lors des tests du circuit intégré.

La Figure 3.21 présente l'évolution des sorties numériques des deux neurones stimulés. Le

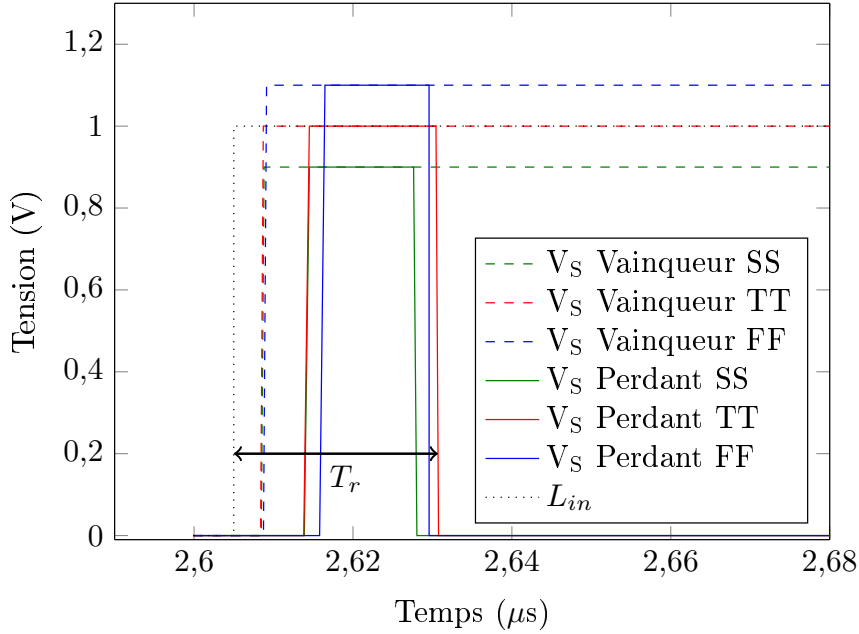


FIGURE 3.21 – Signaux de sortie numérique pour la stimulation considérée.

neurone gagnant est bien activé quelles que soient les conditions de simulation. Le perdant tente de s'activer mais quelles que soient les conditions, il finit par se désactiver. En mesurant le délai entre l'activation des entrées et la désactivation du neurone perdant, le temps de réponse T_r peut être trouvé. Il est évalué en simulation à 30 ns dans le pire des cas (TT). Une stimulation similaire est réalisée en utilisant des neurones de deux blocs *WTA-16* différents. Dans ce cas le temps de réponse est augmenté à 45 ns.

Les différentes simulations présentées valident le bon fonctionnement du circuit en simulation. L'étape suivante consiste donc à intégrer le circuit sur un *ASIC* afin de faire des mesures sur le circuit. L'intégration et les tests du circuit sont présentés dans la section suivante.

3.3 Intégration et test du réseau

3.3.1 Intégration du réseau

Le circuit précédemment présenté est intégré en utilisant la technologie *CMOS* 65 nm de *STMicroelectronics*. Les layouts d'un bloc *WTA-16* ainsi que deux neurones avec leur banc de synapses sont présentés sur la Figure 3.22. Les neurones sont agencés par groupe de quatre et les bancs de synapses sont disposés de part et d'autre. La surface d'un neurone avec ses synapses est de $192 \mu m^2$. La partie analogique du bloc *WTA-16* mesure $8\,884 \mu m^2$ tandis que les registres occupent une surface de $15\,533 \mu m^2$.

La structure du *WTA-16* est répétée afin d'obtenir les huit blocs. De même les huit éléments *WTA* du deuxième étage sont ajoutés. Le résultat final est visible sur la photo de la puce de la

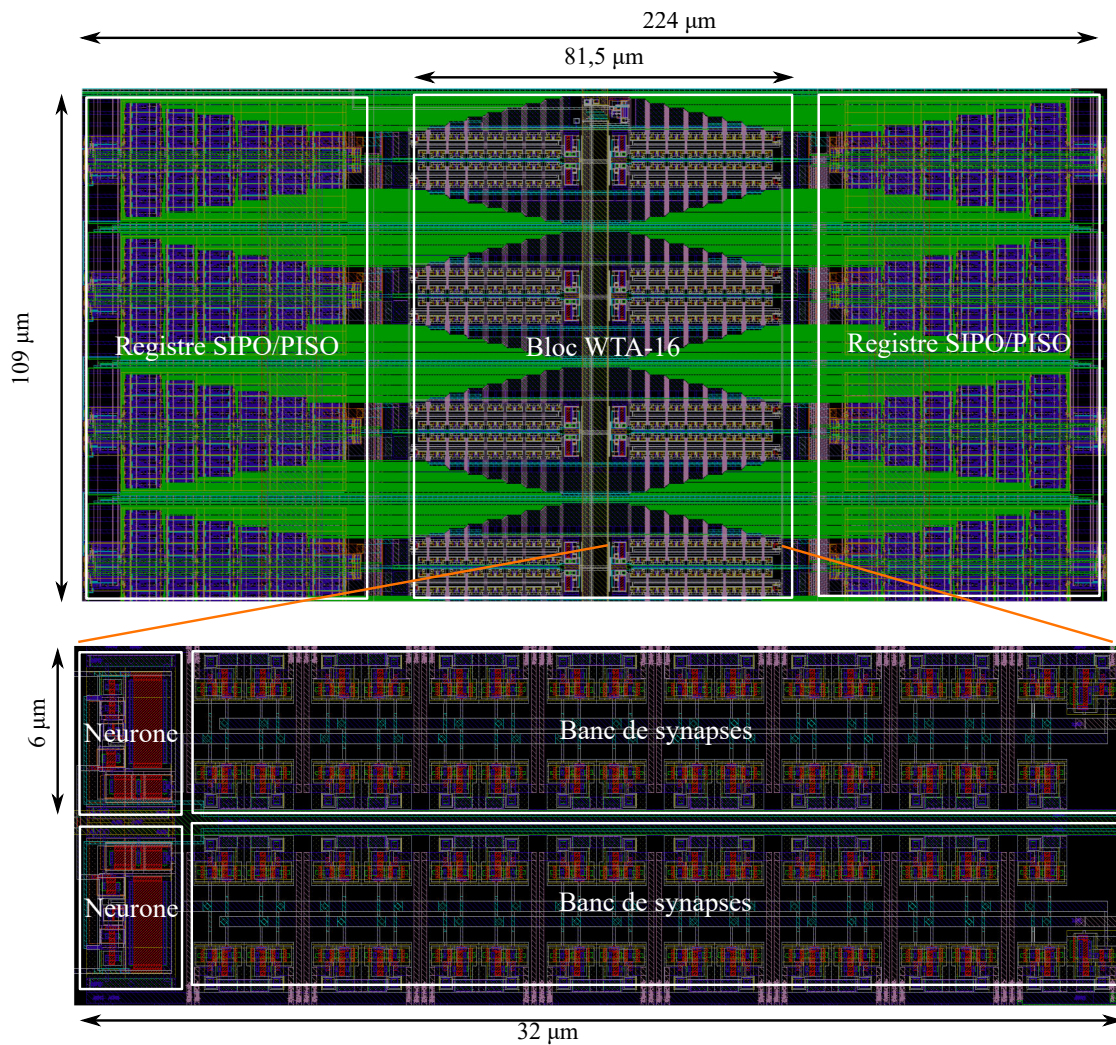


FIGURE 3.22 – Layout d'un bloc *WTA-16* et zoom sur des neurones.

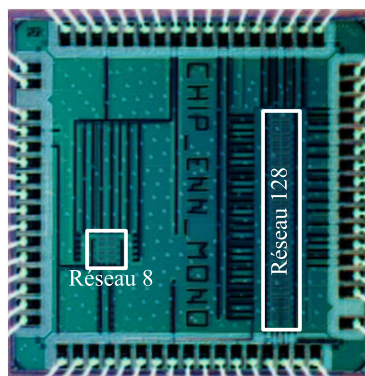


FIGURE 3.23 – Photo de la puce prise au microscope électronique.

Figure 3.23. Le réseau de 128 neurones a une surface totale de $0,21 \text{ mm}^2$ avec $79\,316 \mu\text{m}^2$ occupés par les neurones et synapses analogiques.

La puce contient aussi un réseau de huit neurones de 60 synapses chacun. L'objectif de ce réseau est d'étudier l'évolution de la résolution de l'élément *WTA* avec une réduction du courant unitaire I_{unit} . La puce est un carré de 1,5 mm de côté pour une surface totale de 2,25 mm². Parmi les 64 entrées/sorties du package, dix ne sont pas utilisées, les autres sont réparties de la manière suivante : 24 pour les signaux numériques, 15 pour les signaux analogiques et 15 pour les alimentations. L'anneau de plot est séparé en deux sous-anneaux, un pour les signaux analogiques et un pour les signaux numériques. Afin de limiter les interférences entre les parties numérique et analogique, ces dernières sont isolées via des domaines d'alimentation séparés. Ainsi, la puce possède trois domaines d'alimentation : un pour l'ensemble des traitements numériques, un pour les éléments analogiques du réseau 128 et le dernier pour le réseau de huit neurones.

3.3.2 Présentation du banc de test

Un environnement de test est mis en place afin de fournir les données et signaux de contrôle à la puce. L'interconnexion avec la puce se fait au moyen de la carte de test de la Figure 3.24. Elle permet d'apporter les tensions d'alimentation des anneaux (AVDDE et DVDDE) ainsi que les

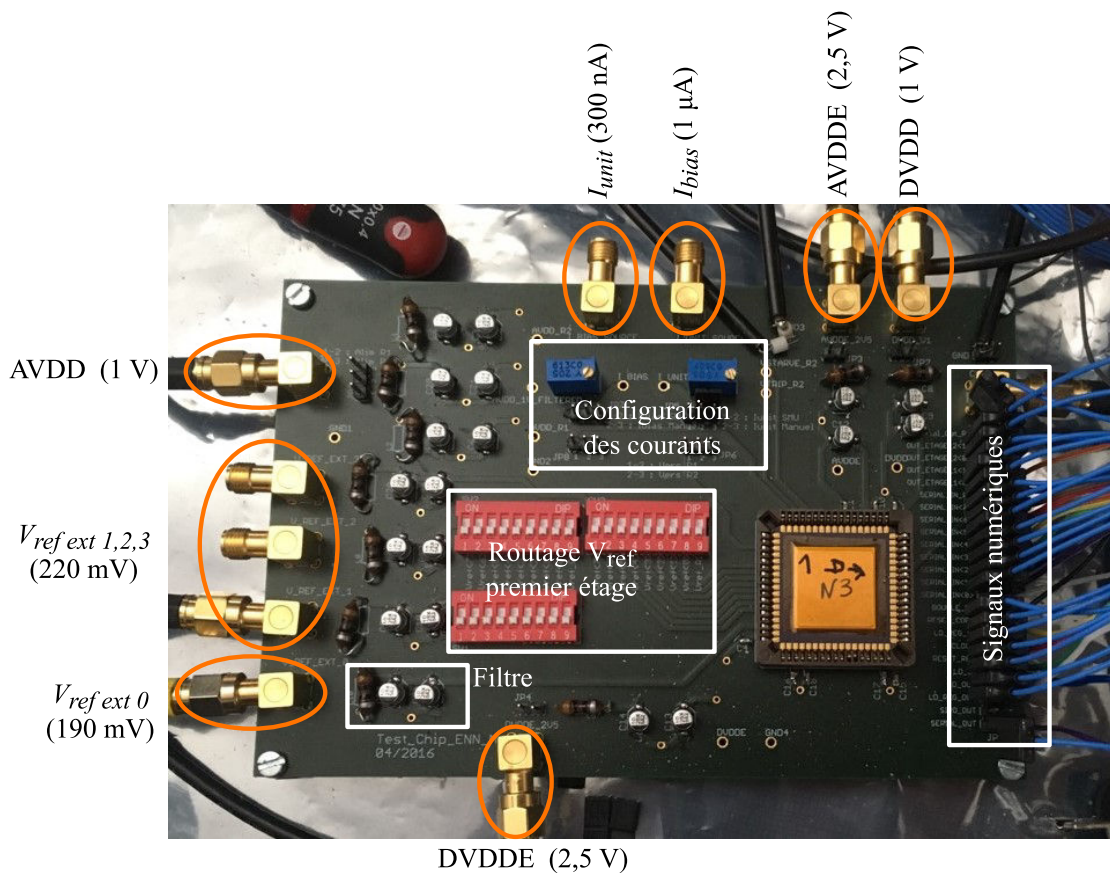


FIGURE 3.24 – Photo de la carte de test conçue.

tensions d'alimentation pour le cœur de la puce (AVDD et DVDD). Toutes les sources de tension arrivant sur la carte sont d'abord filtrées afin d'avoir un signal le plus propre possible. Un cavalier permet de router la tension AVDD afin de choisir le réseau à alimenter. Pour assurer une mesure précise du courant consommé, il n'est pas possible d'alimenter les deux clusters en même temps. Cela garantit que le courant mesuré sur AVDD correspond uniquement à celui consommé par le réseau testé. Un connecteur directement relié aux entrées/sorties numériques de la puce permet de fournir les signaux de contrôle et de données à cette dernière.

Les courants de polarisation peuvent être directement fournis à la puce par une source de courant externe via les entrées I_{unit} et I_{bias} . Il est aussi possible de les générer à partir de l'alimentation AVDD via des résistances variables. Le mode de fonctionnement est réglé au niveau la zone configuration des courants de la carte de test.

La tension de référence V_{ref9} de la boucle de contrôle du deuxième étage est directement connectée à l'entrée $V_{refext0}$ de la carte de test. Les huit autres tensions de référence étant identiques, il n'est pas nécessaire d'avoir une entrée dédiée pour chacune. Il a été choisi d'avoir trois entrées différentes sur la carte de test. Les interrupteurs de routage de la carte permettent de choisir, pour chaque tension V_{refi} la tension extérieure à laquelle elle est connectée. L'utilisation de plusieurs tensions de référence extérieures permet de faire des tests de comportement avec des seuils différents pour les blocs *WTA-16*. Il est ainsi possible d'émuler une chute de la tension de décision d'un bloc et d'analyser le comportement de la puce.

Le banc de test complet est présenté en Figure 3.25. Les tensions d'alimentation et de référence sont générées par des générateurs *Metrix AX503* et *Rohde & Schwarz HMP4040*. La tension

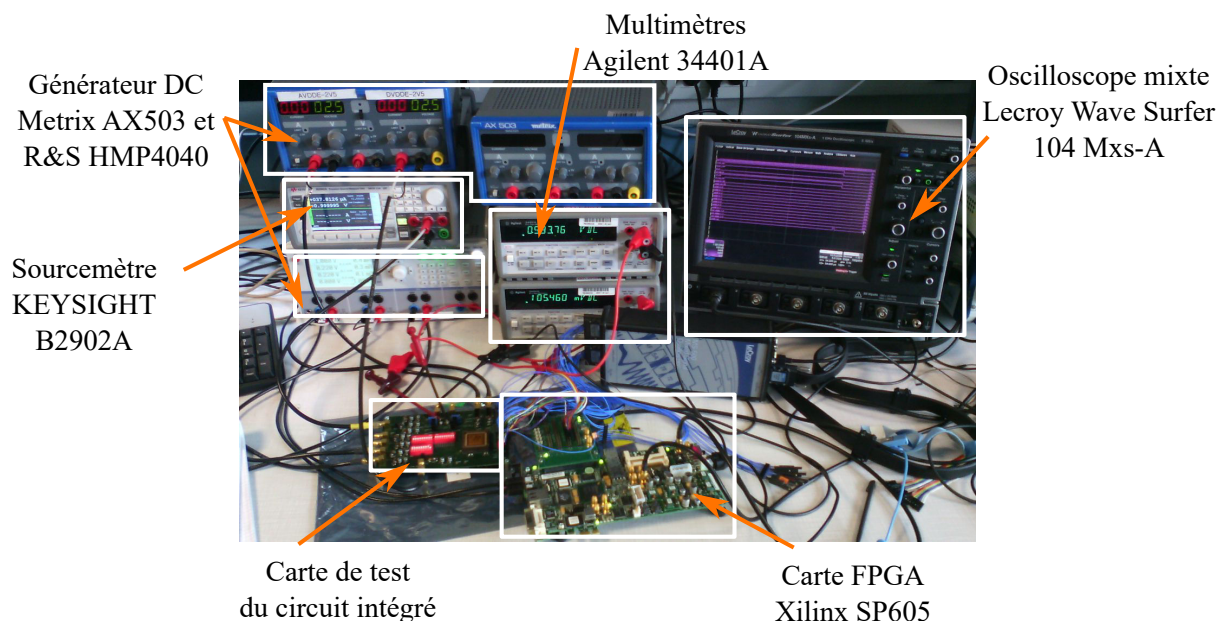


FIGURE 3.25 – Photo du banc de test.

d'alimentation analogique AVDD est fournie par un source-mètre (*Source/Measure Unit*) *Keysight B2902A* qui permet une précision de mesure du courant délivré de 10 pA . La fréquence d'acquisition maximale pour la mesure du courant est de 50 kHz. Deux multimètres *Agilent 34401A* sont utilisés afin de mesurer la tension à différents point de la carte. Il sont notamment utilisée lors du réglage des courants de polarisation via les résistances variables. Enfin une carte d'évaluation *Xilinx SP605* embarquant un *FPGA Spartan-6* est utilisée afin de réaliser la partie contrôle du classifieur.

Le *FPGA* est en communication avec un ordinateur. Cette connexion permet à la fois d'envoyer les données de stimulation et de récupérer la réponse finale du classifieur pour l'analyser. Dans un premier temps, l'ordinateur envoie la stimulation au *FPGA*. Ce dernier se charge ensuite de réaliser la partie contrôle du classifieur et envoie les signaux au circuit intégré. Après le décodage, le *FPGA* renvoie la réponse du classifieur à l'ordinateur. Le protocole de communication mis en place entre l'ordinateur et le *FPGA* permet de créer facilement des scripts pour réaliser différents tests sur le classifieur et sur la puce.

3.3.3 Test et mesures sur le prototype

Le banc de test présenté est utilisé afin d'évaluer les performances du classifieur proposé. Avant de valider le classifieur, les neurones du circuit intégré sont testés un à un afin de vérifier leur bon fonctionnement.

3.3.3.1 Tests unitaires des neurones

Le test unitaire des neurones est effectué dans les conditions nominales suivantes :

- Alimentation $AVDD = 1 V$
- $I_{unit} = 300 \text{ nA}$
- $I_{bias} = 1 \mu\text{A}$
- $V_{ref1,2,\dots,8} = 220 \text{ mV}$
- $V_{ref9} = 190 \text{ mV}$.

Le premier test effectué consiste à s'assurer que les neurones s'activent correctement lorsqu'ils sont stimulés. Cela permet de valider à la fois le bon fonctionnement des registres *SIPO* et *PISO* et celui des neurones. Pour le test, les neurones sont stimulés un par un avec dix contributions actives. Chaque test est répété 100 fois par neurones afin de vérifier la répétabilité des résultats. Le temps de convergence est fixé à 500 ns soit une fréquence de 2 MHz. Les résultats de ce test sont présentés sur la Figure 3.26. La courbe étoilée correspond, pour chaque neurone, au nombre d'activations lorsque que ce dernier est stimulé. Les test étant répétés 100 fois, on constate que la majorité des neurones répond correctement de manière consistante. On constate aussi que les deux premiers neurones ne sont pas fonctionnels puisque le premier ne s'est activé que quatre fois et le deuxième 57. Ces deux neurones ne seront donc pas utilisés pour la suite des tests.

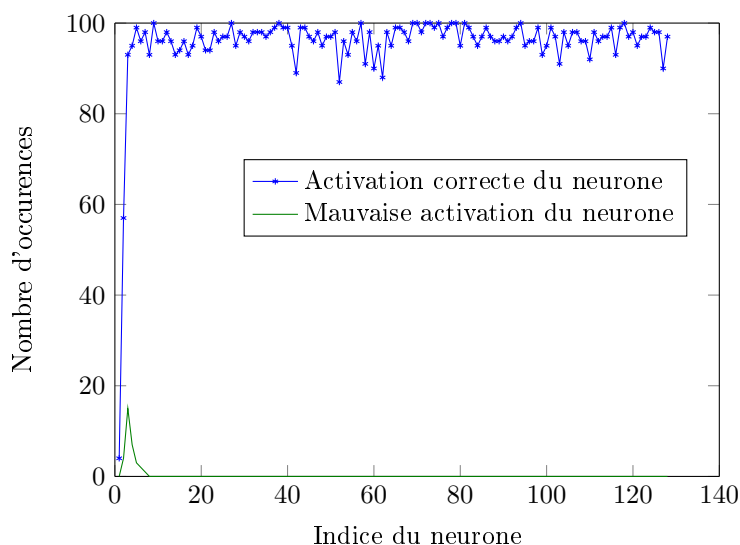


FIGURE 3.26 – Résultats du test des stimulations individuelles à dix contributions des neurones.

La courbe pleine représente, pour chaque neurone, le nombre d’activations alors qu’il n’était pas stimulé. La majorité des neurones restent bien inactifs quand ils ne sont pas stimulés. Cependant, le troisième neurone s’est activé 15 fois alors qu’il n’était pas stimulé. De même, les neurones voisins se sont activés alors qu’ils ne devaient pas. L’analyse détaillée des résultats montre que tous les cas de mauvaises activations apparaissent lors de la stimulation des deux premiers neurones. Cela confirme le dysfonctionnement des deux premiers neurones.

Le second test réalisé est celui de la résolution des éléments *WTA*. Il consiste à réaliser tous les cas de stimulations possible entre deux neurones. Lors de ce test, la synapse valant trois contributions n’est pas considérée. Cela revient à avoir 30 synapse par neurones et donc 900 cas de stimulations possibles. Il n’est pas possible d’effectuer ce test avec tous les couples de neurones du circuit. Ainsi, seuls deux neurones par bloc *WTA-16* sont sélectionnés pour un total de 16 neurones et 120 couples. Les résultats majoritaires sont présentés sur la figure 3.27a. Pour chaque combinaison d’entrées, le résultat reporté est le résultat majoritaire sur les 120 couples testés.

On peut constater que lorsque le nombre de contributions est inférieur à neuf, les cas d’égalité ne sont pas détectés et l’un des deux neurones prend l’avantage. Cela vient du fait que la copie de courant n’est pas parfaite dans les miroirs, causant ainsi une différence de courant dans les éléments *WTA*. À partir de neuf contributions, l’égalité est bien détectée et une zone d’incertitude apparaît. Cette zone s’élargie jusqu’à 18 contributions. Cela signifie qu’à partir de 18 contributions pour les deux neurones, le circuit n’arrive plus à déterminer le gagnant. Cette limite est beaucoup plus basse que celle obtenue en simulation et restreint les cas d’application du circuit. En effet la résolution obtenue pour trente contributions est de trois en simulation. En mesure, elle est de $31 - 18 = 13$. Cependant cela ne pose pas de problème pour l’application de classification des signaux ECG qui présente au maximum huit contributions actives sur chaque neurone. Cette limite très inférieure à celle obtenue en simulation s’explique par le fait d’utiliser deux étages de comparaison. En effet,

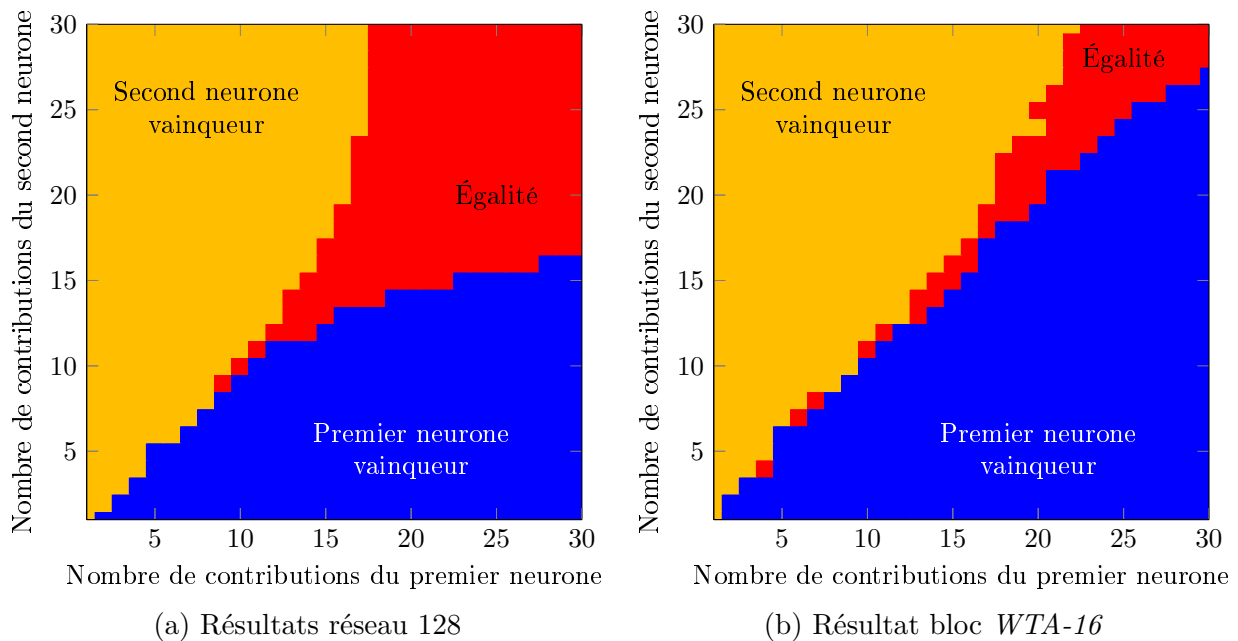


FIGURE 3.27 – Réponse du réseau et des blocs *WTA-16* lors des duels.

le deuxième étage ajoute deux miroirs de courant, augmentant ainsi les chances d'erreurs et de mauvais comportements. De plus, les simulations présentées en Section 3.2.3.2 ne concernent qu'un seul étage de comparaison avec six neurones.

Afin de vérifier le résultat des simulations de manière plus précise, les cas de duels avec des neurones issus du même bloc *WTA-16* sont isolés. Le fait de prendre des neurones issus du même bloc permet de regarder uniquement le comportement du premier étage puisque le deuxième n'aura qu'une entrée non nulle. Les résultats sont présentés sur la Figure 3.27b. On constate que la zone d'égalité est beaucoup plus petite et que le point limite de détection passe de 18 à 23.

La conclusion des tests unitaires est que deux neurones parmi les 128 sont non fonctionnels. Tous les autres neurones répondent d'une manière similaire. Cela démontre une robustesse aux variations de processus de fabrication. De plus, l'analyse de la résolution montre qu'il est préférable de se limiter à des réseaux avec un nombre de connexions par neurone inférieur à 18 s'il y a plus de 16 neurones. Il est possible d'augmenter le nombre de connexions à 23 mais pour ce faire, il faudra se limiter à un maximum de 16 neurones afin de n'utiliser qu'un bloc *WTA-16*. Malgré la résolution restreinte, le circuit est parfaitement adapté pour la réalisation du classifieur pour les ECG. La section suivante présente les résultats obtenus lors de la mise en place du classifieur.

3.3.3.2 Performances de classification

Le test pour vérifier les performances de classification se fait dans les mêmes conditions nominales que les tests unitaires. Un ensemble de 27 neurones est sélectionné sur la puce et les données des mémoires de connexions sont envoyées au FPGA. Les 1572 signaux de dix secondes de la base de données sont pré-traités via *Matlab*[®] afin d'obtenir le message de stimulation sur 24 bits. Chaque message est ensuite transféré au classifieur et le décodage est réalisé. La sortie est ensuite renvoyée à l'ordinateur qui la compare à l'entrée pour produire la matrice de confusion. Le temps de convergence est d'un cycle d'horloge à la fréquence de 2 MHz. La matrice de confusion obtenue est présentée dans le Tableau 3.2. Elle est très similaire à celle du Tableau 2.4 qui présente les résultats simulés. Ces résultats se traduisent par les performances moyennes du Tableau 3.3. On constate que les performances mesurées sont très proches des valeurs simulées. Ces résultats ont fait l'objet d'une présentation [CLAS17] à la conférence EMBC 2017. Le circuit permet donc bien de réaliser la classification des signaux ECG.

La fonctionnalité du réseau proposé étant validée, l'analyse de la consommation du circuit est réalisée. Pour ce faire, le courant consommé durant le décodage d'un message est analysé. La consommation est mesurée dans le pire des cas qui correspond à un message d'entrée étant proche de deux cliques. L'appareil de mesure ne permettant pas une fréquence d'acquisition supérieure à 50 kHz, elle est effectuée avec une fréquence de fonctionnement de 10 Hz. Ainsi, le temps de convergence de 100 ms pour une stimulation assure de pouvoir analyser la courbe de courant lors du décodage. Cette dernière est présentée sur la Figure 3.28. Les quatre traits pleins représentent le début des itérations du décodage *WsTA* ainsi que le décodage *WTA* de sortie. Les itérations de *WsTA* étant elles-même itératives, le départ de chaque nouvelle stimulation du réseau est marquée par un trait pointillé. Le profil obtenu est cohérent avec ce qui est attendu. En effet, entre chaque itération de décodage *WsTA*, le nombre de synapses activées augmente. Cela s'observe bien avec

TABLEAU 3.2 – Matrice de confusion obtenue avec le prototype.

		Classe prédite		
		<i>LBBB</i>	<i>RBBB</i>	<i>PB</i>
Classe d'entrée	<i>LBBB</i>	435	51	5
	<i>RBBB</i>	14	491	24
	<i>PB</i>	0	25	542

TABLEAU 3.3 – Performances de classification simulées et mesurées.

	Simulation	Mesure
Sensibilité	92,6 %	92,2 %
Spécificité	96,2 %	95,9 %
Exactitude	95 %	94,7 %

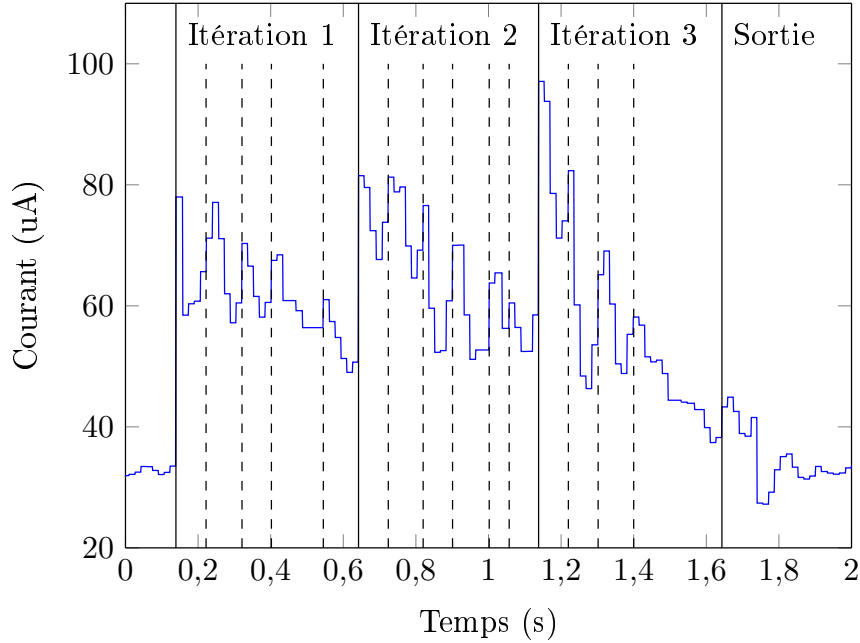


FIGURE 3.28 – Profil de la consommation de courant durant le décodage du pire cas.

une augmentation du courant maximum lors des différentes itérations. Lors d’une itération de *WsTA*, chaque sous-itération correspond à la désactivation d’un ou plusieurs neurones suite à son/leur activation. Ainsi, au sein même d’une itération, la consommation diminue au fur et à mesure que des neurones s’activent. Le décodage de sortie ne concernant que trois neurones, on s’attend bien à ce que la consommation associée soit inférieure à celle des itérations précédentes. La consommation statique représente environ 66 % de la consommation moyenne durant le décodage. Elle est constituée de la consommation des boucles de contrôle ainsi que des courants de fuite des nombreux neurones et synapses.

Le nombre de stimulations pour effectuer chaque itération de *WsTA* n’est pas constant et peut varier en fonction des conditions. En effet, le fonctionnement du contrôleur présenté en Section 3.2.1 stimule le réseau jusqu’à avoir huit neurones actifs. La mesure du courant est donc effectuée dix fois afin d’avoir une valeur moyenne du courant par décodage. Ce courant I_{moy} est de $49,7 \mu\text{A}$. Pour obtenir l’énergie par classification avec un fonctionnement à 2 MHz la formule suivante est utilisée :

$$E = AVDD \cdot T_{conv} \cdot N_{stim} \cdot I_{moy}, \quad (3.19)$$

où T_{conv} est le temps de convergence, soit 500 ns, et N_{stim} le nombre de fois que le réseau est stimulé. La valeur N_{stim} est obtenue en moyennant le nombre de stimulations par décodage sur une dizaine de décodages du même message. Dans les conditions nominales à 2 MHz, il est de 22. L’énergie par classification est donc de 657 pJ. Cette énergie correspond à l’énergie totale utilisée par les 128 neurones du réseau. Il est possible d’estimer la consommation uniquement des 27 neurones utilisés. En rapportant la consommation statique à chaque neurone, il est possible de

TABLEAU 3.4 – Comparaison entre le système proposé et un système purement transmetteur.

	Système transmetteur	Système proposé
CAN	4,8 nJ	4,8 nJ
Ondelettes	X	876 nJ
Extraction	X	472 nJ
Transformation	X	37 nJ
Classification	X	243 pJ
Transmission	147 μ J	11,1 nJ
Total	147 μ J	1,401 μ J

retirer la consommation statique des 101 neurones non utilisés. L'énergie consommée par les 27 neurones durant le décodage est alors de 243 pJ.

Le circuit conçu permet donc de bien classifier les signaux ECG tout en ne consommant que 243 pJ par classification. Bien que cette consommation ne concerne que la partie analogique, la simplicité de la partie contrôle nécessitant seulement 135 octets de mémoire et une machine d'état simple permet d'avoir un système de classification des ECG très basse consommation. Il est donc tout à fait possible d'imaginer concevoir un circuit dédié à l'application afin de réaliser une classification embarquée dans le capteur. De plus, le classifieur intégré nécessite 2 millions de fois moins d'énergie que [BMD⁺14] et [LV13] et 300 fois moins que [CMM⁺13] tout en améliorant les performances de 2 à 3 %.

Le Tableau 3.4 fait une synthèse des estimations de la consommation des différents éléments du système d'analyse des ECG proposé. Ce dernier permet de réduire la consommation d'un facteur 105 par rapport à un système ne faisant que transmettre. Cela permet d'augmenter l'autonomie du capteur tout en réduisant l'utilisation de l'étage RF. L'objectif de traiter et classifier les signaux ECG afin de détecter le type d'arythmie est donc atteint. Le système est capable de correctement classer le signal avec une exactitude moyenne de 95%, tout en consommant 105 fois moins d'énergie qu'un système transmettant toute l'information. En plus de permettre d'atteindre l'objectif, cette étude permet aussi de montrer qu'il y a un intérêt à déplacer une partie des traitements sur les capteurs plutôt que de tout centraliser sur une machine très puissante.

3.3.3.3 Réduction de l'énergie du classifieur.

Dans cette section, deux pistes sont ouvertes pour tenter de réduire l'énergie nécessaire pour la classification. La première piste consiste à étudier le comportement du circuit en modifiant le courant I_{unit} . En effet, la majorité de la consommation dynamique étant issue de l'activité des synapses, il semble intéressant de voir son évolution avec une modification du courant.

Les courants considérés varient de 450 à 150 nA par pas de 50 nA. Les performances de classification ont été mesurées pour chacun de ces courants. Elles sont identiques quelque soit le courant. Le courant moyen durant le décodage est ensuite mesuré comme précédemment. De même, le nombre de stimulations nécessaire pour effectuer le décodage à 2 MHz est relevé afin de calculer

TABLEAU 3.5 – Évolution de la consommation avec le courant unitaire.

Courant unitaire (nA)	450	400	350	300	250	200	150
Courant moyen consommé (μA)	111	108	62	50	46	43	39
Nombre de stimulations	18	17	19	22	30	> 64	> 64
Énergie (pJ/Classif)	1101	1003	682	657	840	> 1685	> 1554

l'énergie consommée. Le Tableau 3.5 présente les résultats obtenus pour l'ensemble du réseau. On constate que le point de fonctionnement à $I_{unit} = 300 \text{ nA}$ est optimal parmi les valeurs testées. Si l'augmentation du courant a pour effet de réduire le nombre de stimulations nécessaires, le surplus de consommation est trop important pour réduire la consommation. De même, la réduction du courant ne permet pas une diminution du courant suffisante pour compenser l'augmentation du nombre de stimulations nécessaires. Pour les courants unitaires de 200 et 150 nA, le nombre de stimulations exact n'a pas pu être mesuré du fait d'une taille de compteur trop petite. Cela ne pose pas de problèmes puisque, à ces courants, le nombre de stimulations du réseau est trop important pour permettre une faible consommation.

Un autre axe pour réduire la consommation est l'augmentation de la fréquence de fonctionnement. En effet, une grande partie de la consommation étant issue du régime statique, la réduction du temps de décodage peut permettre des gains énergétiques. Cette étude a été menée tardivement du fait de la non disponibilité de l'équipement afin de générer une horloge externe suffisamment grande. Avec la possibilité de modifier facilement la fréquence de fonctionnement, la vitesse limite de fonctionnement des neurones a été évaluée. Pour ce faire la fréquence est augmentée progressivement jusqu'à ne plus avoir de réponse lors d'une stimulation unitaire de 10 contributions des neurones. La fréquence limite est de 12 MHz. En considérant qu'en fonctionnant à cette fréquence le nombre de stimulations reste identique que durant les opérations à 2 MHz, l'énergie par classification passerait de 657 pJ à 110 pJ par classification. Des tests sont donc envisagés pour trouver la fréquence de fonctionnement pour avoir l'énergie minimale par classification.

Conclusion

Dans ce chapitre un prototype est mis en place pour réaliser les opérations du système d'analyse du chapitre 2. Il s'appuie sur une architecture numérique pour réaliser le pré-traitement et sur un système mixte pour le classifieur. La conception mixte proposée est implantée dans un circuit intégré. Le circuit réalisé améliore le circuit de neurone existant [Lar15] en offrant une meilleure résolution, tout en effectuant un passage à l'échelle de six à 128 neurones en parallèle. À cela s'ajoute l'apport de la flexibilité des connexions afin de pouvoir utiliser le circuit pour d'autres applications. Le circuit intégré est interconnecté avec un FPGA pour réaliser le contrôle du classifieur et apporter la flexibilité. L'exactitude de classification moyenne de 95 % est en accord avec celle obtenue en

simulation et les mesures permettent de confronter la réalité à la simulation.

L'estimation de la consommation du pré-traitement montre qu'il faut $1,39 \mu\text{J}$ pour obtenir le message de 24 bits servant d'entrée au classifieur. La consommation du classifieur intégré est de 243 pJ par classification, ce qui est six ordres de grandeur plus faible que d'autres classifieurs embarqués [BMD⁺14] [CMM⁺13] [LV13]. Bien que le prototype proposé ne soit que partiellement intégré sur ASIC, il montre que l'intégration du classifieur est une solution intéressante pour l'amélioration de l'autonomie des capteurs dans les *BASN*. L'architecture proposée permet en effet de diviser par 105 la consommation du capteur. Une évolution de ces travaux consiste à réaliser une intégration complète de la chaîne de traitement.

La réduction drastique de la consommation du classifieur ainsi que la transmission d'une faible quantité de données conduisent à un système dont 98,8 % de l'énergie est utilisée pour l'extraction de paramètres. Afin de réduire encore plus la consommation, il semble donc intéressant de se focaliser sur la manière d'acquérir ces paramètres. Le chapitre suivant présente une technique récente d'acquisition des signaux laissant espérer une réduction d'énergie pour l'extraction des paramètres tout en conservant des performances de classification acceptables. Ainsi la combinaison de cette technique avec un classifieur embarqué basse consommation pourrait donner naissance à des capteurs ne consommant que quelques nJ par classification.

Chapitre 4

Un nouveau système d'acquisition parcimonieux des ECG

Introduction

Le système d'analyse des chapitres 2 et 3 réduit la consommation du capteur via un traitement embarqué des données. La consommation est réduite pour deux raisons : l'utilisation d'un classifieur en mixte analogique/numérique et l'utilisation parcimonieuse du transmetteur. Cela permet de diviser le budget énergétique du capteur par 105 par rapport à un capteur ne faisant que transmettre l'information. Les réductions de consommation apportées permettent d'avoir un capteur pour lequel acquisition, classification et transmission ne représentent que 1,2 % de la consommation totale. Les 98,8 % restant sont utilisés pour l'extraction de paramètres. Ainsi, réduire encore plus la consommation, il faut s'intéresser à la manière d'extraire les paramètres.

Dans un premier temps, une nouvelle méthode d'acquisition des signaux [VB15] est présentée. Elle combine les étapes d'acquisition et d'extraction de paramètres afin de réduire la consommation. Pour cela, il s'appuie sur le principe des CAN extracteurs de paramètres aussi appelés convertisseurs analogique vers paramètres (CAP). La combinaison d'un CAP et d'un classifieur permet de bénéficier des avantages du traitement analogique tout en réduisant ces limitations. En effet, certains défauts des circuits analogiques comme les non-linéarités sont pris en compte durant l'apprentissage et n'ont donc pas besoin d'être compensés par des mécanismes complexes.

Ce chapitre propose ensuite un système de traitement des ECG via la mise en place d'un CAP. Le CAP proposé s'appuie sur l'échantillonnage non uniforme par ondelettes initialement conçu dans le cadre de l'échantillonnage compressif. L'utilisation d'un tel système nécessite une optimisation conjointe du CAP et du classifieur. Un algorithme de sélection adaptatif est proposé afin de maximiser les performances tout en minimisant le nombre de paramètres. Le système d'acquisition est simulé sous *Matlab*[®] afin d'évaluer ses performances dans la détection d'anomalies sur les signaux ECG. Enfin une discussion est faite sur les points importants du système afin de permettre l'implantation de ce dernier.

4.1 Le concept de convertisseur analogique vers paramètres

4.1.1 Principe

Le conversion analogique vers paramètres est une méthode d'acquisition récente qui apparaît avec le développement des réseaux de capteurs et de l'internet des objets. Les capteurs ont pour objectif d'acquérir une grandeur physique afin de détecter un événement particulier. La connaissance du signal complet n'est pas nécessaire puisque seules certaines informations spécifiques sont extraites. Ces informations sont ensuite utilisées pour faire une classification.

Le Tableau 4.1 présente la répartition de l'énergie pour différents capteurs effectuant de la classification sur les ECG. [CLAS17] correspond au système présenté au Chapitre 3. Dans [BMD⁺14] et [CMM⁺13], les systèmes présentés ne comprennent que l'extraction des paramètres et la classification. On constate que dans tous les cas, la majorité de l'énergie est consommée pour l'extraction des paramètres.

Pour réduire la consommation de cette dernière fonction, il faut analyser la manière dont les paramètres sont extraits. Pour cela on considère le signal ayant le spectre de la Figure 4.1. L'information de ce signal est contenue dans les trois bandes B_1 , B_2 et B_3 . Pour une application spécifique, considérons que l'information utile est contenue dans B_2 . D'une manière générale l'acquisition des paramètres est effectuée en faisant d'abord une acquisition complète du signal à la

TABLEAU 4.1 – Répartition de l'énergie pour des circuits réalisant la classification d'ECG.

Ref	[CLAS17]	[BMD ⁺ 14]	[CMM ⁺ 13]
Acquisition	0,3 %	X	X
Extraction des paramètres	98,8 %	59,2 %	90,7 %
Classification	0,02 %	40,8 %	9,3 %
Transmission	0,88 %	X	X

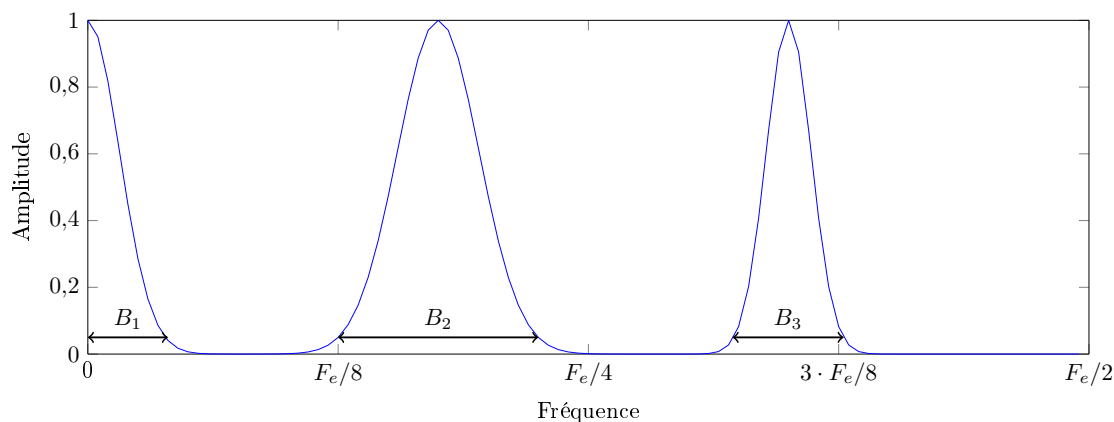


FIGURE 4.1 – Exemple de spectre d'un signal à acquérir. La bande utile pour l'application considérée est B_2 .

fréquence d'échantillonnage F_e . Ensuite, le signal numérique est filtré pour ne conserver que la bande d'intérêt et extraire les paramètres. Cette méthode n'est pas efficace d'un point de vue énergétique. En effet, le système produit un grand nombre d'échantillons afin de capter la totalité de l'information. Ces derniers sont ensuite filtrés pour n'en garder qu'un nombre réduit. Une méthode plus efficace consisterait à ne faire l'acquisition que de la partie utile du signal pour l'application cible.

C'est l'objectif de la conversion analogique vers paramètres qui vise à extraire du signal analogique uniquement l'information nécessaire. La notion d'information utile étant très dépendante de l'application, ce type d'acquisition l'est aussi. Elle se limite en effet aux applications de type classification et faisant appel à une extraction de paramètres. Avec l'explosion des applications autour des réseaux de capteurs, l'intérêt pour ce type d'acquisition grandit. Si elle peut permettre des gains énergétiques, la mise en place concrète de convertisseurs analogique vers paramètres présente de nombreux défis. La section suivante présente architecture générale d'un CAP ainsi que ses avantages et inconvénients.

4.1.2 Architecture générale

La conversion analogique vers paramètres vise à combiner les étapes d'acquisition et d'extraction de paramètres. Ainsi, il n'est pas nécessaire de faire une acquisition complète du signal suivi d'un traitement numérique. L'architecture générale d'un CAP, proposée par [VB15], est illustrée en Figure 4.2. Elle est composée d'un ensemble d'extracteurs analogiques de paramètres travaillant directement sur le signal. Chaque extracteur peut extraire un ou plusieurs paramètres. Les paramètres analogiques ainsi extraits sont connectés à un multiplexeur dont la sortie est reliée à un CAN. En parcourant l'ensemble des extracteurs via le multiplexeur, le CAN produit les différents paramètres numérisés utilisés par un classifieur. Le nombre de paramètres extraits étant faible, la fréquence d'échantillonnage du CAN sera inférieure à la fréquence de Nyquist. Une architecture

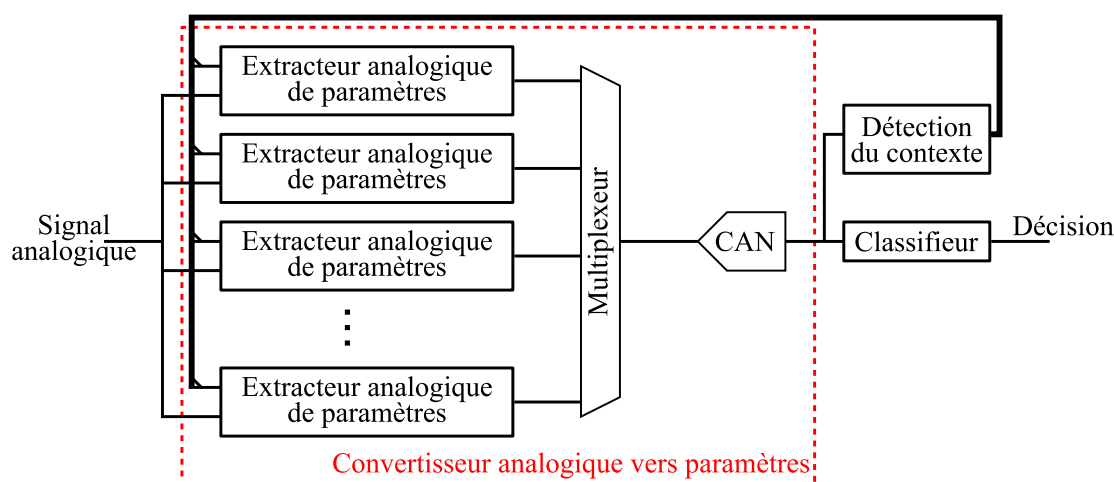


FIGURE 4.2 – Architecture d'un CAP telle que présentée dans [VB15].

alternative consiste à ne pas utiliser de multiplexeur et à placer un CAN à la sortie de chaque extracteur. Cela permet de réduire encore plus la fréquence d'acquisition des CAN au prix d'une surface plus grande.

Pour des systèmes au fonctionnement continu, il est possible d'avoir besoin de plusieurs jeux de paramètres utilisant des extracteurs différents selon le contexte. Un détecteur de contexte est donc utilisé afin de n'activer que les extracteurs utiles. Cela permet de réduire la consommation en ne faisant que l'acquisition des données nécessaires.

L'utilisation de traitements analogiques permet d'avoir une meilleure efficacité énergétique. Les auteurs de [LRYR16] et [TSH⁺17] utilisent un traitement analogique afin de détecter les instants d'échantillonnage d'un capteur d'impédance de la peau. Les échantillons acquis sont ensuite directement utilisés pour calculer la pression sanguine. Grâce au traitement analogique, le nombre d'échantillons ainsi que la consommation sont réduits. En plus d'être efficace énergétiquement, l'utilisation de méthodes de conception mixte permet l'ajout de flexibilité dans les circuits analogiques. Il est, par exemple, possible de choisir les caractéristiques d'un filtre [GCDB07] [LLG⁺13] sans avoir à dupliquer le matériel.

L'intérêt principal du CAP vient de l'apprentissage supervisé effectué sur les paramètres extraits. L'utilisation du classifieur peut permettre de lever de nombreuses contraintes sur la conception des extracteurs analogiques. Le classifieur étant capable de classifications non-linéaires, il pourra s'adapter aux imperfections de calcul lors de la phase d'apprentissage. Il n'est donc pas forcément nécessaire de dépenser du temps de conception, de la surface et de l'énergie pour corriger les non-linéarités ou les problèmes d'appariement des circuits analogiques. Cela offre un fort potentiel de simplification des fonctions analogiques du CAP, laissant espérer un temps de conception plus court et une consommation réduite. De plus, cela permet d'avoir une reproductibilité des performances ainsi qu'une meilleure fiabilité. Il est possible d'obtenir une robustesse face au vieillissement via la mise à jour du classifieur à intervalles réguliers.

L'architecture présentée en Figure 4.2 est mise en place dans [BLMV15] pour faire de la détection de parole. L'intégration en *CMOS* 90 nm nécessite 6 μ W pour extraire et classifier le signal. La consommation est réduite d'un facteur dix par rapport à des solutions classiques tout en ayant des performances similaires de l'ordre de 89 % de bonne détection.

La difficulté principale pour la mise en place d'un CAP est la sélection des paramètres. Il faut en effet choisir des paramètres permettant à la fois une extraction basse consommation et offrant des performances de classifications satisfaisantes. Cela revient à utiliser des paramètres avec une faible complexité d'extraction. De plus, les performances de classification étant naturellement dépendantes du classifieur, ce dernier doit être pris en compte lors de l'étude du système. Ainsi, la mise en place d'un CAP nécessite des compétences en conception de circuit, traitement du signal et apprentissage machine. La section suivante présente les travaux réalisés pour la mise en place d'un CAP pour la détection d'anomalie pour les signaux ECG.

4.2 Acquisition parcimonieuse de signaux ECG : conversion analogique vers paramètres

4.2.1 Objectifs

L'utilisation d'un CAP étant très spécifique à une application, il est important de bien définir cette dernière. D'une manière similaire au détecteur de parole de [BLMV15], l'objectif ici est la conception d'une acquisition de paramètres permettant de détecter une anomalie sur le signal ECG. À la différence du système proposé dans le Chapitre 2, la classification est strictement binaire entre signal normal et anomalie présente. L'architecture doit permettre une détection à chaque battement, et une réduction de la consommation par rapport au système du Chapitre 3. D'après le Tableau 3.4, ce dernier consomme $1,39 \mu\text{J}$ pour l'obtention des paramètres de dix secondes de signal.

Les signaux utilisés pour évaluer les performances de classification du système sont ceux issus de la base de données MIT-BIH [MIT]. Pour effectuer une classification battement par battement, un bloc de données de 256 échantillons avec le complexe QRS centré sur le 100^{ième} échantillon est considéré. Ainsi, pour chaque battement annoté dans la base de données, un vecteur est formé en prenant les 100 échantillons précédant et les 156 suivant le complexe. Dix signaux normaux et dix anormaux sont présentés en Figure 4.3. Le signal étant échantillonné à 360 Hz, chaque vecteur représente 0,71 s de signal avec le complexe localisé à 0,28 s. Chaque vecteur correspond au signal qui serait obtenu suite à une acquisition uniforme classique. Cette segmentation de la base conduit à la création de 107 045 vecteurs dont 75 019 correspondent à un signal normal et le reste à des signaux anormaux.

La section suivante présente l'architecture proposée pour faire une acquisition parcimonieuse du signal ECG et le classifier.

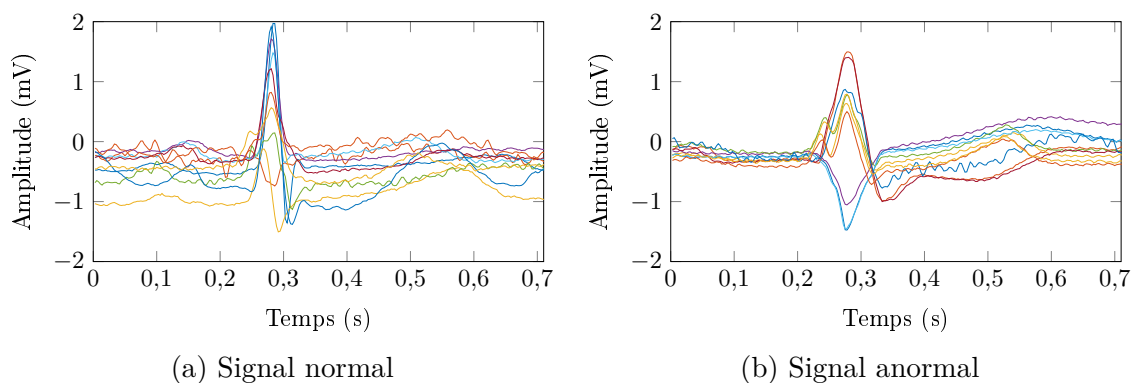


FIGURE 4.3 – Exemple de signaux alignés sur le complexe QRS issus de la base de données MIT-BIH [MIT] annotée.

4.2.2 Échantillonnage non uniforme par ondelettes

Comme il a été montré dans le Chapitre 2, les paramètres les plus souvent utilisés sont soit des échantillons spécifiques du signal temporel, soit des informations issues d'une transformation en ondelettes discrète du signal. De plus l'élément le plus énergivore du système proposé au Chapitre 3 est la transformation en ondelettes. Il semble donc pertinent de se focaliser sur ce type de paramètres pour réduire la consommation.

L'architecture proposée pour acquérir les signaux ECG s'appuie sur l'échantillonnage non uniforme par ondelettes initialement proposé dans [PS17] pour l'échantillonnage compressif de signaux RF. L'architecture illustrée en Figure 4.4 présente le CAP proposé. Le signal analogique issu des électrodes est d'abord amplifié par le frontal analogique avant d'être mélangé avec une ondelette analogique. Le résultat du mixage est ensuite intégré avant d'être numérisé par un CAN. L'échantillon en sortie du CAN est un paramètre qui sera utilisé par le classifieur. Ainsi, pour chaque ondelette générée, un paramètre est produit. L'ondelette est générée par un générateur flexible permettant de choisir les paramètres de cette dernière. Pour une forme d'ondelette donnée, il existe trois degrés de liberté sur lesquels il est possible d'agir : la fréquence d'oscillation f_c , la position temporelle t_c et la largeur du support l . Le support l d'une fonction est défini comme étant l'intervalle de temps sur lequel cette dernière est non nulle. L'expression générale temporelle

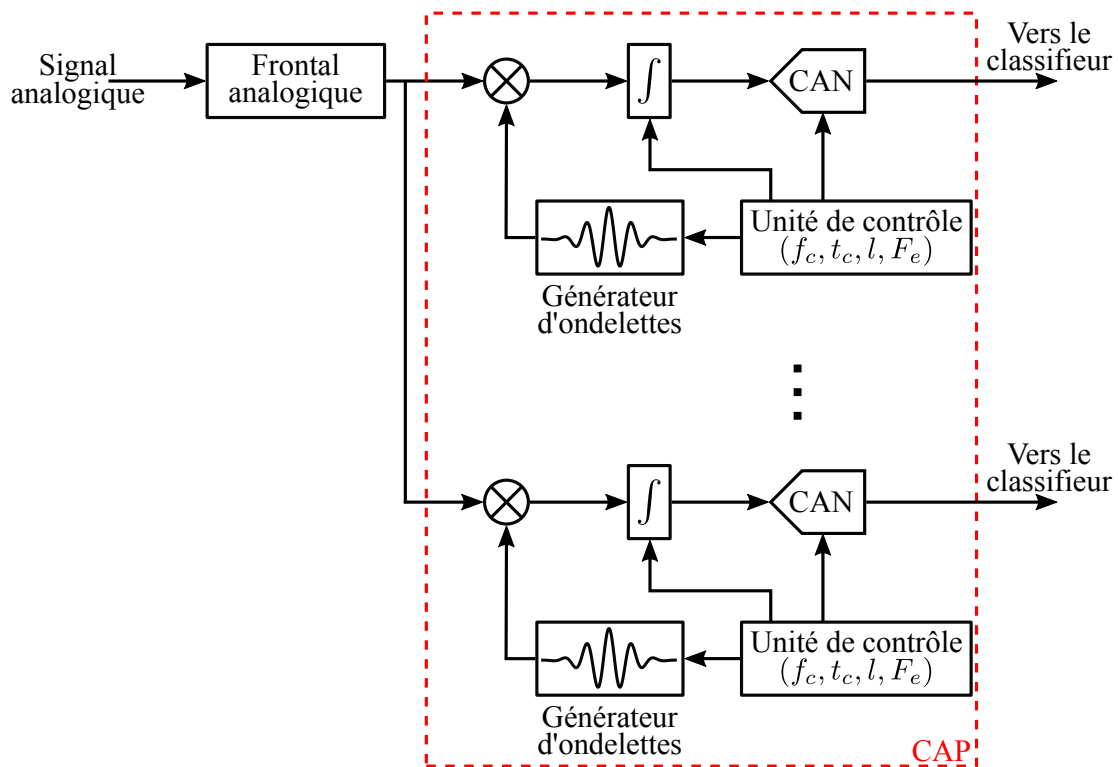


FIGURE 4.4 – Architecture du CAP proposé pour l'acquisition des signaux ECG.

d'une ondelette s'écrit :

$$\psi_{f_c, t_c, l}(t) = p_l(t - t_c) \cdot g_{f_c}(t - t_c), \quad (4.1)$$

où $p_l(t)$ est une fonction fenêtre centrée avec un support de taille l et $g_f(t)$ une fonction périodique de période $1/f$. Une ondelette est donc une fonction paramétrée de trois paramètres. Le type de l'ondelette est déterminé par les fonctions $p(t)$ et $g(t)$.

Les contrôleurs du CAP permettent de choisir les paramètres des ondelettes à générer, la remise à zéro de l'intégrateur et les instants d'échantillonnage du CAN. L'architecture de la Figure 4.4 présente un CAP avec plusieurs branches similaires. Chacune de ces branches est capable d'extraire un ou plusieurs paramètres. Le nombre de paramètres par branche dépend des ondelettes associées aux paramètres sélectionnés. Une branche est capable d'extraire des paramètres tant que les supports des ondelettes utilisées sont distincts. Ainsi, pour limiter la complexité et la consommation il est important de minimiser le nombre paramètres pour réduire le nombre de branches.

Le mélange avec une ondelette, suivi d'une intégration, est équivalent à obtenir directement un coefficient d'une transformation en ondelettes. Ainsi, il est possible d'extraire des informations à la fois temporelle et fréquentielle sur le signal. La position de l'information temporelle est centrée sur t_c et celle de l'information fréquentielle sur f_c . Le paramètre l permet de choisir la bande passante de l'ondelette. Plus le support temporel sera grand plus l'information sera localisée en fréquence. À l'inverse plus le support sera faible plus l'information sera localisée en temps.

Le système proposé devant être intégré, il est nécessaire d'utiliser des ondelettes pouvant être générées de manière analogique. Dans [PS17] les ondelettes de Gabor sont considérées. L'ondelette de Gabor est le résultat du produit d'une exponentielle complexe avec une fenêtre gaussienne. Cela équivaut à :

$$p_l^{Gabor}(t) = \exp\left(\frac{-t^2}{\alpha \cdot l^2}\right), \quad (4.2)$$

$$g_{f_c}^{Gabor}(t) = \exp(2 \cdot \pi \cdot j \cdot f_c \cdot t), \quad (4.3)$$

où α est un paramètre pour choisir la largeur de la gaussienne. L'intérêt de l'ondelette de Gabor est que la fenêtre utilisée est une gaussienne dans le domaine temporel et dans le domaine fréquentiel. La génération d'une fonction gaussienne paramétrable pouvant se révéler complexe, une ondelette similaire utilisant une fenêtre carrée est aussi considérée. Une troisième ondelette très simple est aussi évaluée. Il s'agit d'un signal carré entre -1 et 1 à la fréquence f_c avec une fenêtre carrée. Cette ondelette offre l'avantage de pouvoir être générée très simplement de manière numérique. De plus, elle ne prend que trois valeurs, ce qui simplifie la conception du multiplieur. La Figure 4.5 présente un exemple de chacune de ces ondelettes avec $f_c = 22,5$ Hz et $l = 180$ ms.

Avec l'architecture proposée, le nombre de paramètres extractibles est potentiellement infini. En effet, un paramètre peut être extrait pour chaque triplet (f_c, t_c, l) avec chaque élément pouvant prendre une infinité de valeurs. Faire une sélection d'un nombre fini de paramètres dans cet ensemble infini n'est pas réalisable. Une première étape consiste donc à effectuer une discrétisation

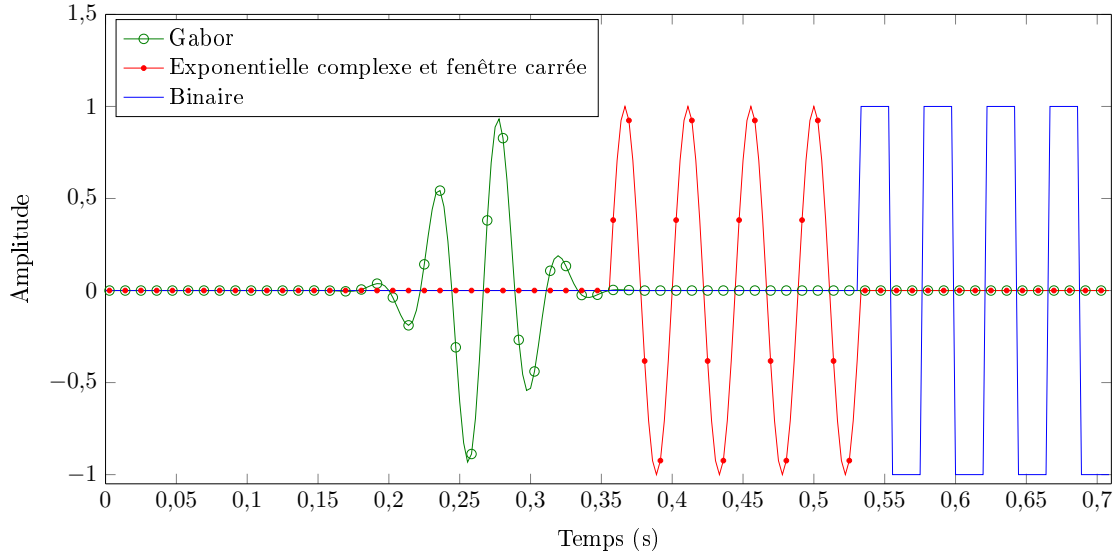


FIGURE 4.5 – Exemple des trois types d'ondelettes considérés avec un support de 180 ms et une fréquence de 22,5 Hz.

de cet ensemble de paramètres afin de pouvoir réaliser la sélection des paramètres sur un ensemble réduit. Pour se faire, une décomposition dyadique en temps et en fréquence est réalisée. Le signal d'entrée considéré étant un signal de 256 échantillons pris à 360 Hz, l'intervalle temporel maximal pour obtenir un paramètre est de 0,71 s. Il est noté T dans la suite. Les valeurs possibles pour le support l des ondelettes sont :

$$l(k) = \frac{T}{2^k}, \quad \forall k = 0, 1, \dots, n, \quad (4.4)$$

avec $n = 7$. Cette valeur de n conduit à un support minimal de 5,6 ms, soit deux échantillons. Cela permet de considérer toutes les tailles possibles de support à partir des données disponibles. De même la position temporelle de l'ondelette est limitée aux valeurs :

$$t_c(k, m) = (2 \cdot m + 1) \cdot \frac{T}{2^{k+1}}, \quad \forall k = 0, 1, \dots, n, \quad \forall m = 0, 1, \dots, k. \quad (4.5)$$

Enfin la fréquence des ondelettes est aussi discrétisée de la manière suivante :

$$f_c(k) = \frac{F_s}{2^{k+1}}, \quad \forall k = 0, 1, \dots, n, \quad (4.6)$$

avec F_s la fréquence d'échantillonnage uniforme de 360 Hz du signal considéré. Ces choix conduisent à huit valeurs possibles pour l et f_c et 255 pour pour t_c . Le nombre total de paramètres possibles est donc de 16 320.

Une telle discrétisation des paramètres des ondelettes ne tient pas compte des relations entre les paramètres. Une contrainte supplémentaire est ajoutée afin d'assurer que, pour chaque ondelette considérée, au moins une période soit contenue dans le support. Cela se traduit par la suppression

de tous les cas pour lesquels $l < 1/f_c$. Une seconde contrainte s'assure que la totalité de l'ondelette soit comprise dans l'intervalle de temps $[0; T]$. Elle peut s'écrire sous la forme des deux conditions suivantes :

$$t_c - l/2 \geq 0 \quad \text{et} \quad t_c + l/2 \leq T. \quad (4.7)$$

L'utilisation de ces deux contraintes permet d'assurer de n'utiliser que des ondelettes complètes pour extraire les paramètres. Après l'application des contraintes, il reste 502 ondelettes possibles.

La Figure 4.6 montre les 26 premières ondelettes pour chaque type considéré. Pour chaque ligne l'évolution de la partie réelle dans le temps est illustrée via une échelle de couleurs. Les ondelettes sont classées par fréquences croissantes. Pour chaque fréquence, elle sont ensuite classées par support décroissant. Pour chacune des tailles de supports, les positions temporelles sont ajoutées de manière incrémentale. Ainsi la première ligne contient l'ondelette à la fréquence la plus faible de 1,4 Hz. Il y qu'une seule ligne contenant cette fréquence car il n'est pas possible de réduire le support tout en conservant une période complète. Avec cette représentation on constate bien que ces ondelettes permettent de sélectionner des informations très diverses.

L'étape suivante pour la mise en place du CAP consiste à optimiser la sélection des paramètres et du classifieur. Cette optimisation doit offrir le maximum de performances possible avec un minimum de paramètres.

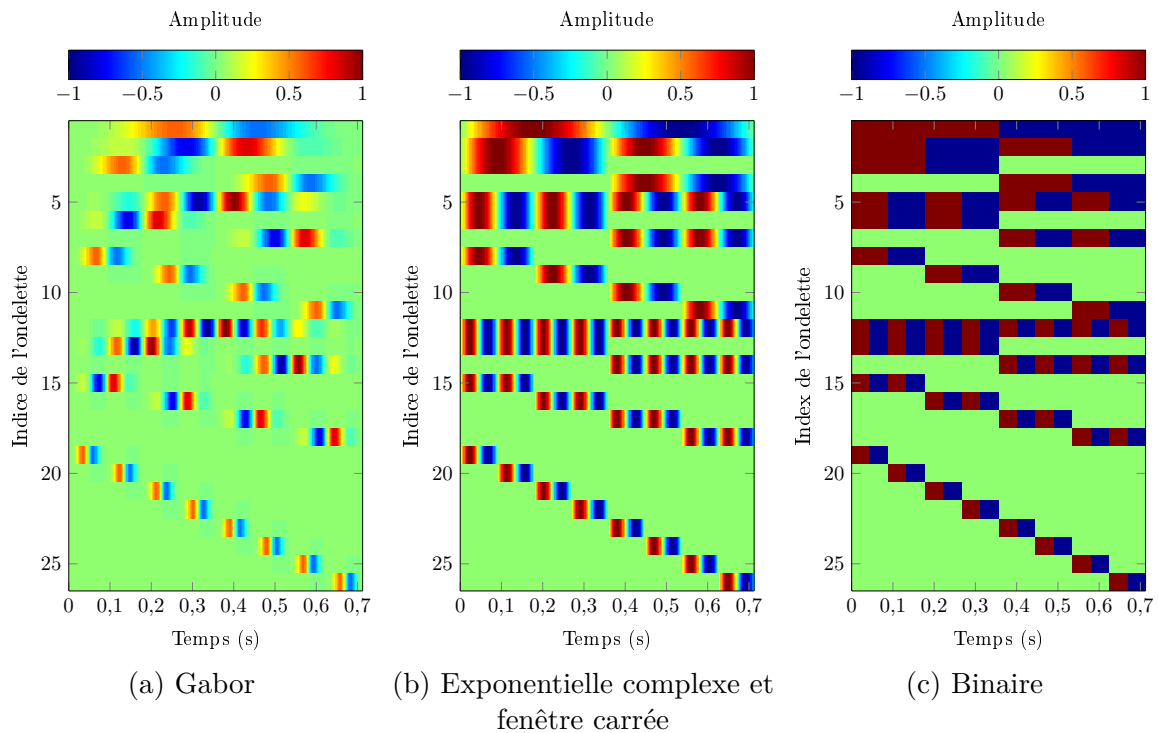


FIGURE 4.6 – Évolution temporelle des 26 premières ondelettes suite à la discrétisation.

4.2.3 Optimisation conjointe de la sélection des paramètres et du classifieur

L'objectif de cette section est de choisir les paramètres et le classifieur pour la mise en place du CAP. Lors de l'étude, les trois types d'ondelettes sont considérés de manière indépendante. Il n'est donc pas possible de prendre des paramètres issus d'ondelettes de différents types.

Les classifieurs considérés sont le perceptron et la forêt aléatoire. Ils sont choisis car ils permettent tout deux d'effectuer des classifications non linéaires tout en étant intégrable pour une consommation limitée [ZAAB16] [HHL⁺16]. Le perceptron utilisé possède une couche cachée de dix neurones et deux neurones de sortie. La forêt est une forêt contenant 20 arbres. L'apprentissage et l'évaluation sont effectués via l'outil d'apprentissage machine de *Matlab*[®]. Dans les cas testés, l'apprentissage est réalisé sur 70 % des données et le test sur les 30 % restantes. La séparation des données est effectuée de manière aléatoire.

Cette étude nécessitant d'effectuer un très grand nombre d'apprentissages sur les classifieurs, les réseaux à cliques des Chapitre 2 et 3 ne sont pas considérés. En effet, la méthode pour créer la connaissance *a priori* étant manuelle, il n'est pas possible mettre en place un réseau pour chaque test dans un temps raisonnable. L'apprentissage d'un perceptron ou d'une forêt est de l'ordre de la seconde ou de la minute tandis que la mise en place d'un réseau à cliques nécessite plusieurs jours voir semaines.

La suite de cette section présente les méthodes testées pour choisir les paramètres et le classifieur. Dans un premier temps deux méthodes classiques de sélection de paramètres sont mises en place afin de sélectionner le meilleur type d'ondelette et de valider l'intérêt de l'architecture. Ensuite un algorithme de sélection séquentielle de paramètres est créé afin de prendre en compte les spécificités de la méthode d'acquisition.

4.2.3.1 Sélection des paramètres indépendamment de l'architecture

La sélection des paramètres est une étape importante afin de maximiser les performances de classification. Pour se faire il est nécessaire de déterminer des critères de sélection permettant d'évaluer la pertinence des différents paramètres.

La première méthode utilisée repose sur les scores de Fisher [TAL14]. Cette méthode attribue un score à chaque paramètre possible afin d'estimer son importance pour la classification désirée. C'est une méthode supervisée, c'est-à-dire qu'il est nécessaire de connaître la classe des vecteurs d'observations utilisés. Le calcul du score d'un paramètre repose sur l'idée qu'un bon paramètre prendra des valeurs proches pour des observations d'une même classe et différentes pour des classes concurrentes. À partir de cette idée le score S_i du paramètre i est définie comme :

$$S_i = \frac{\sum_{j=1}^C n_j \cdot (\mu_{ij} - \mu_i)^2}{\sum_{j=1}^C n_j \cdot \rho_{ij}^2}, \quad (4.8)$$

avec C le nombre de classes, n_j le nombre d'observations appartenant à la classe j , μ_i et μ_{ij} respectivement la moyenne globale du paramètre et la moyenne du paramètres dans la classe j . Enfin ρ_{ij} représente la variance du paramètre pour la classe j . Les paramètres ayant les scores les plus élevés seront sélectionnés en priorité.

Cette méthode ne requiert pas de calcul complexe et permet d'obtenir rapidement un classement des paramètres par importance. De plus, elle est indépendante du classifieur. L'avantage est que la sélection ne nécessite pas d'apprentissage de classifieur pour choisir les paramètres, la rendant polyvalente. Cette polyvalence est aussi un défaut car la méthode ne prend pas en considération les spécificités du classifieur choisi. Cela peut conduire à une sélection sous-optimale comparée à une méthode qui utilise l'association avec un classifieur pour choisir les paramètres.

L'inconvénient majeur des scores de Fisher est que ces derniers sont calculés de manière indépendante pour chaque paramètre. Ainsi les corrélations pouvant exister entre les paramètres ne sont pas prises en compte, conduisant à une sélection sous-optimale. En effet, si un paramètre est présent deux fois dans l'ensemble de paramètres disponible, il aura deux fois le même score. Si ce score est grand, alors les deux paramètres identiques seront conservés alors qu'il est évident qu'une fois le premier utilisé, le second n'apporte aucune nouvelle information. Cette méthode est donc plus adaptée pour une sélection globale de paramètres que pour une sélection précise.

Une deuxième méthode de sélection liée à l'apprentissage d'une forêt aléatoire est aussi utilisée. Cette méthode vise à classer les paramètres par ordre d'importance. Elle repose sur l'idée que si un paramètre est important alors un changement de sa valeur devrait se traduire par une perte des performances de classification. L'apprentissage d'une forêt aléatoire utilise l'ensachage décrit en Section 1.4.2.2. Ainsi pour chaque arbre de la forêt, 37 % des données d'entraînement ne sont pas sélectionnées pour créer l'arbre. Ces données sont appelées observations *out-of-bag* (*OOB*). Elles sont utilisées pour avoir une estimation des performances de classification lors de l'apprentissage et pour évaluer l'importance d'un paramètre.

La mesure de l'importance d'un paramètre se fait arbre par arbre. Pour chaque arbre, chaque paramètre est évalué en permutant sa valeur entre les observations OOB de l'arbre. L'augmentation de l'erreur de classification suite aux permutations est mesurée pour chaque paramètre et chaque arbre. La mesure finale de l'importance d'un paramètre est obtenue en moyennant la mesure sur l'ensemble des arbres puis en divisant par l'écart-type. Enfin les paramètres avec les plus grandes métriques sont sélectionnés. Cette méthode est plus efficace avec un nombre réduit de paramètres au départ. En effet, l'utilisation d'un paramètre dans le modèle de la forêt est plus rare lorsque le nombre de paramètres augmente. Cela se traduit par une mesure de l'importance des paramètres moins précise.

Cette méthode utilise l'entraînement d'une forêt aléatoire pour évaluer l'importance d'un paramètre. À l'inverse des scores de Fisher, le choix des paramètres est effectué en utilisant un critère s'appuyant sur un résultat de classification. Par contre, le calcul de l'importance des paramètres est encore effectué de manière indépendante pour chaque paramètre. La différence est que lors

du calcul, les autres paramètres interviennent quand même dans l'arbre. Il est donc possible de considérer que cette méthode permet une sélection des paramètres en prenant plus en compte les corrélations que les scores de Fisher.

Une troisième méthode hybride est proposée. Elle consiste à faire une pré-sélection des 128 meilleurs paramètres via les scores de Fisher (SF) puis d'appliquer la sélection par importance de la forêt aléatoire (IF). Ainsi, les scores de Fisher sont utilisés de manière globale puis la sélection est affinée avec la méthode IF. Cela permet de tirer parti du meilleur de chacune des deux méthodes.

4.2.3.2 Choix de l'ondelette et validation de l'architecture

Dans un premier temps, les trois types d'ondelette considérés sont comparés. Pour cela, l'ensemble des 107 045 signaux d'entrée est mélangé avec les trois jeux d'ondelettes. Les données sont ensuite séparées entre apprentissage et test. Les performances de classification obtenues en fonction du nombre de paramètres sont présentées en Figure 4.7. La sélection est effectuée avec les scores de Fisher et le classifieur utilisé est le perceptron. Ce dernier prenant des entrées réelles, les paramètres complexes sont séparés en deux nouveaux paramètres, l'un pour la partie réelle et l'autre pour la partie imaginaire.

La limite d'exactitude à 70 % vient du fait que, dans la base de données, 70 % des signaux correspondent à un cas normal. Cette limite correspond aux performances d'un classifieur qui attribuerait la classe normale quelle que soit l'entrée. On constate que, d'une manière générale,

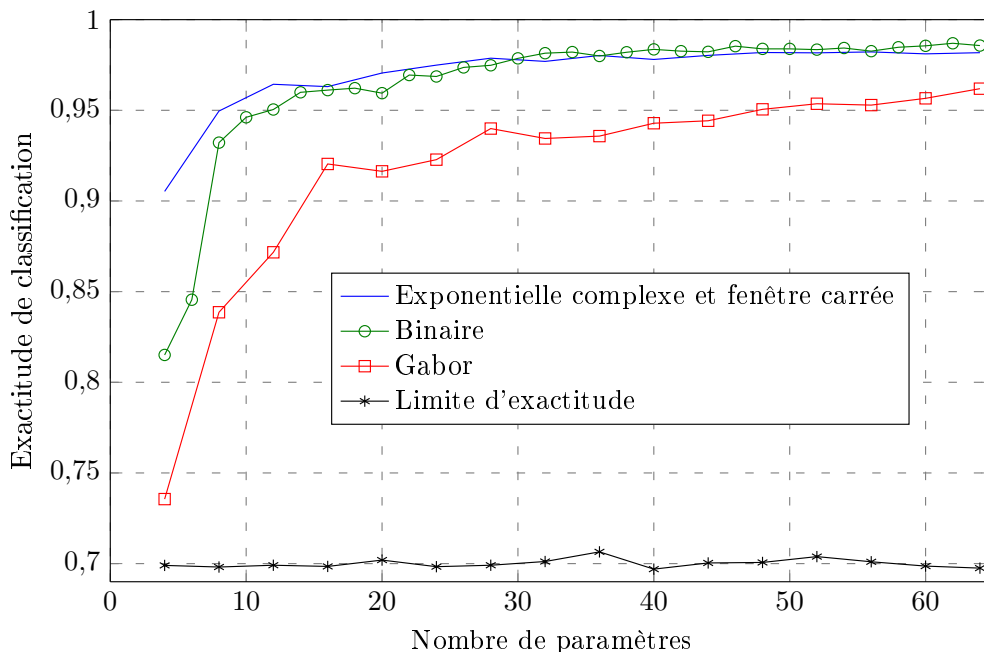


FIGURE 4.7 – Exactitude de classification pour les différents types d'ondelettes considérés.

l'exactitude de classification augmente avec le nombre de paramètres. Cette augmentation est très forte jusqu'à environ dix paramètres. Au delà, l'ajout de nouveaux paramètres ne permet qu'une faible amélioration des performances. En prenant 64 paramètres, la pire exactitude de classification obtenue est de 96,2 %. L'utilisation de 64 paramètres équivaut à une réduction du nombre d'échantillons acquis d'un facteur quatre par rapport à une acquisition uniforme. Cela permet de valider l'intérêt d'une telle architecture pour acquérir les signaux ECG.

On constate que pour moins de 30 paramètres, l'exponentielle complexe avec une fenêtre carrée offre les meilleures performances. Au delà, c'est l'ondelette binaire qui devient meilleure pour atteindre une exactitude de 98,6 % avec 64 paramètres. Pour la suite, seule l'ondelette binaire est sélectionnée car elle offre le meilleur compromis entre performances et simplicité de l'architecture.

Les différentes combinaisons de méthode de sélection et classifieur sont ensuite évaluées. Les résultats obtenus sont présentés sur la Figure 4.8. On constate que le perceptron se comporte moins bien que la forêt aléatoire. L'augmentation du nombre de neurones cachés n'a pas permis d'améliorer les performances. On constate aussi que les méthodes SF et IF présentent des performances assez similaires pour un classifieur donné. Par contre, une nette amélioration apparaît avec l'utilisation de méthode hybride. L'association de la méthode hybride avec la forêt aléatoire offre une exactitude de classification de 96,7 % avec seulement quatre paramètres. Une telle architecture permet donc d'acquérir 64 fois moins d'échantillons tout en offrant une bonne exactitude de classification.

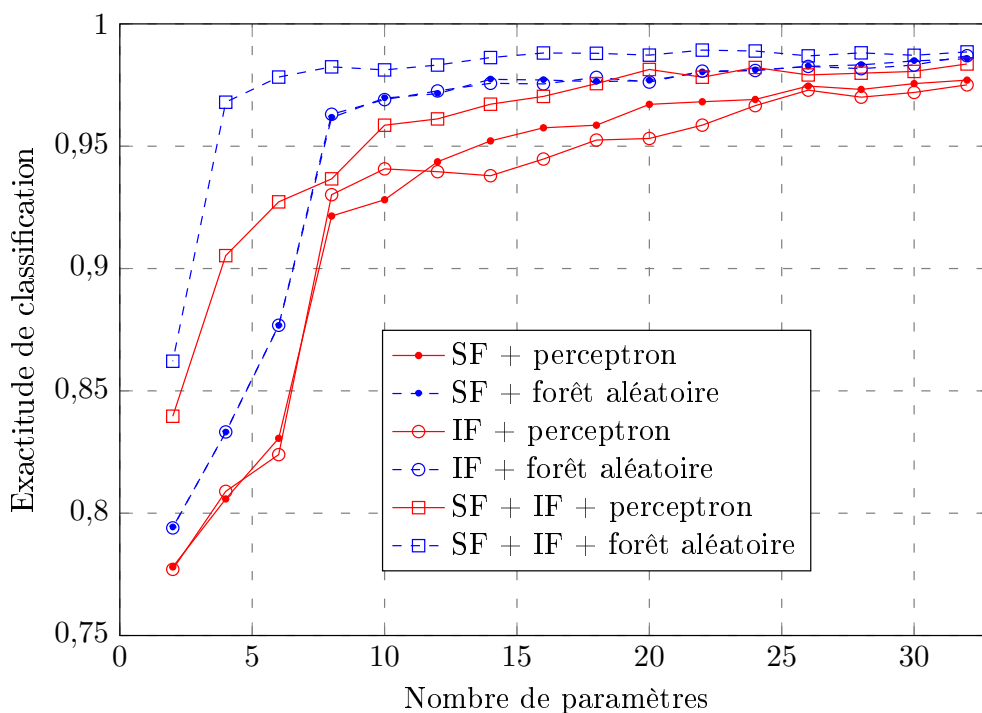


FIGURE 4.8 – Exactitude de classification pour les différentes méthodes de sélection et classifieurs.

Suite à ces tests, la méthode hybride combinée à la forêt aléatoire semble être le meilleur choix. Si les méthodes de sélection testées permettent de bonnes performances, elles sont complètement indépendantes de l'architecture d'acquisition. La section suivante présente un algorithme de sélection permettant de choisir des paramètres en prenant en compte l'architecture.

4.2.3.3 Algorithme de sélection adapté à l'architecture

Afin de réduire au maximum la consommation de l'architecture proposée en Figure 4.4, il est nécessaire de minimiser le nombre de branches de cette dernière. Pour cela il faut tenir compte du support des différentes ondelettes lors de la sélection des paramètres. Cela n'est pas possible avec les méthodes de sélection précédentes qui évaluent indépendamment chaque paramètre. Un algorithme de sélection itératif permettant de choisir le nombre de branches est donc proposé.

L'algorithme proposé pour faire la sélection de paramètres est un algorithme de sélection séquentielle. Il prend comme entrées l'ensemble des 502 ondelettes et le nombre de branches désiré N_b . Trois ensembles sont considérés : l'ensemble des ondelettes sélectionnées, l'ensemble des ondelettes candidates et l'ensemble des intervalles interdits. L'ensemble des ondelettes candidates regroupe toutes les ondelettes pouvant être sélectionnées. L'ensemble des intervalles interdits correspond à l'union des intervalles inclus N_b fois dans les supports des ondelette sélectionnées. Il représente l'ensemble des intervalles de temps qu'il n'est plus possible d'utiliser car déjà occupés sur toutes les branches.

Au départ, l'ensemble des ondelettes sélectionnées et celui des intervalles interdits sont vides. L'ensemble des ondelettes candidates contient la totalité des 502 ondelettes possibles. L'algorithme itère tant que l'ensemble des ondelettes candidates n'est pas vide. À chaque itération, pour chaque candidate restante, un classifieur est entraîné en utilisant l'ensemble des ondelettes déjà sélectionnées et la candidate. Les performances de chacun des classifieurs sont évaluées. La candidate issue du classifieur avec les meilleures performances est ajoutée à l'ensemble des ondelettes sélectionnées et est retirée de celui des candidates. Suite à la sélection d'une nouvelle ondelette, l'ensemble des intervalles interdits est mis à jour. Toutes les ondelettes candidates dont le support inclus une partie de l'ensemble des intervalles interdits sont retirées de l'ensemble des candidates. Si ce dernier n'est pas vide, l'itération suivante commence. Dans le cas contraire, l'algorithme s'achève.

Cet algorithme est utilisé avec une forêt aléatoire comme classifieur. Un grand nombre d'apprentissages étant nécessaire, l'ensemble des candidates de départ est réduit aux 128 ondelettes ayant les plus forts scores de Fisher. Cette réduction a pour but de réduire le temps d'exécution de l'algorithme. Deux cas extrêmes sont considérés : un nombre illimités de branches et une branche unique. Les performances de classification obtenues pour ces deux scénarios sont présentées en Figure 4.9. On constate que pour un nombre illimité de branches les performances sont très similaires à celles obtenues avec la méthode hybride. En se limitant à une seule branche, l'algorithme n'a pu sélectionner que six paramètres. Les performances sont réduites de 1,5 % mais la complexité de l'architecture est simplifiée à une branche unique. En supposant une acquisition des paramètres sur 11 bits le taux de compression atteint est de 97,7 % avec six paramètres. C'est 14,4 points de

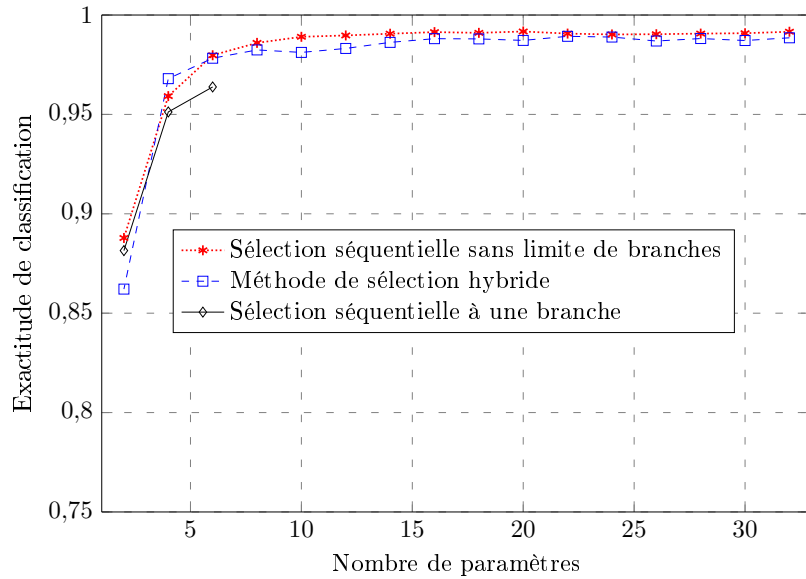


FIGURE 4.9 – Exactitude de classification pour l’algorithme proposée pour une branche et sans limite de branches.

plus que le convertisseur A2I de [GAD⁺14] qui utilise 22 branches.

Cet algorithme offre de nombreux avantages. Dans un premier temps il permet de s’assurer que les paramètres sélectionnés pourront tous être extraits avec le nombre de branches choisi. Il peut aussi être utilisé avec n’importe quel type de classifieur et n’est pas limité à un seul type. Enfin, l’étape de calcul des performances permet de choisir le critère à maximiser. Pour la Figure 4.9 le critère utilisé est l’exactitude de classification. Il peut être intéressant d’utiliser un critère faisant intervenir les performances par joule. Cela offrirait la possibilité de réduire encore plus la consommation tout en conservant de bonnes performances. Pour faire cela, il est nécessaire d’avoir un modèle de la consommation du CAP.

La section suivante présente les réflexions menées pour effectuer la conception matérielle du CAP. Elle ouvre des pistes pour l’obtention d’un modèle de la consommation du CAP

4.3 Discussion sur l’implantation matérielle

Dans cette section les gains de consommation potentiels de l’architecture de CAP proposée sont évalués via les résultats de circuits existant réalisant des opérations similaires. De plus, deux éléments indispensables pour la conception matérielles sont discutés : la synchronisation des données et la quantification.

4.3.1 Discussion sur la consommation

L’architecture proposée en Figure 4.4 est une architecture très simple. Les ondelettes utilisées étant binaires, leur génération peut être très facilement effectuée à partir d’une horloge numérique à

la fréquence maximale considérée de 180 Hz. De plus, le mélangeur peut être très simple puisque les ondelettes ne peuvent prendre que trois valeurs $\{-1; 0; 1\}$. La consommation espérée est calculée à partir du convertisseur $A2I$ de [GAD⁺14] de la Figure 4.10. L'architecture du convertisseur est très similaire à celle du CAP proposé dans ce chapitre. Il y a trois principales différences : la seconde entrée du mélangeur est codée sur six bits, un générateur aléatoire est nécessaire et chaque branche ne fait l'acquisition que d'un échantillon.

Les architectures étant très similaires il paraît intéressant de s'appuyer sur les résultats de consommation de ce circuit afin d'évaluer le potentiel du CAP proposé. Le circuit est fabriqué en technologie $CMOS$ $0,13 \mu m$ et utilise des transistors fonctionnant sous le seuil. La consommation d'une branche pour une fréquence d'acquisition de 2 kHz est de 28 nW. Cela équivaut à une énergie de 14 pJ par échantillon. En utilisant l'architecture de [GAD⁺14], l'énergie nécessaire au CAP pour acquérir six paramètres est de 84 pJ. Les six paramètres étant extraits sur 0,71 s de signal, l'acquisition de dix secondes de signal nécessite 1,18 nJ. La Figure 4.11 montre la consommation énergétique du traitement de dix secondes de signal des différents systèmes considérés dans ce document. Le système transmetteur fait une acquisition de Nyquist suivi de la transmission directe des données. Le traitement embarqué correspond au système proposé aux Chapitres 2 et 3 avec traitement et classification sur le capteur. Enfin la combinaison du CAP avec le traitement embarqué combine l'acquisition directe des paramètres avec la classification embarquée. Le CAP produisant des paramètres à chaque battement, l'énergie pour la classification de dix secondes de signal est décuplée. De même, la quantité de données transmises est augmentée.

L'utilisation d'un CAP permet de réduire la consommation d'un facteur 26 par rapport à un

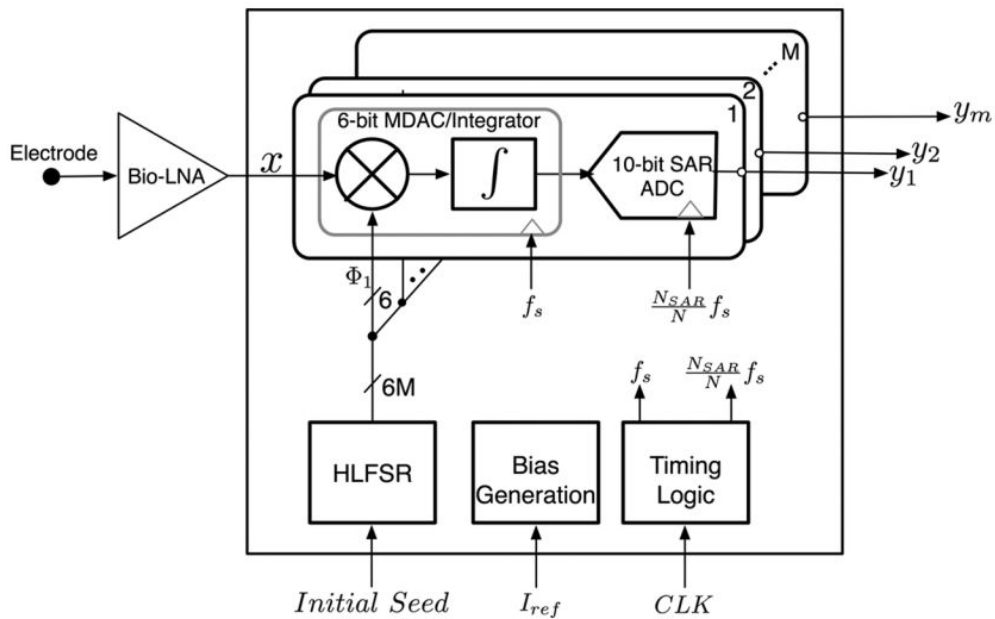


FIGURE 4.10 – Architecture du convertisseur $A2I$ tirée de [GAD⁺14].

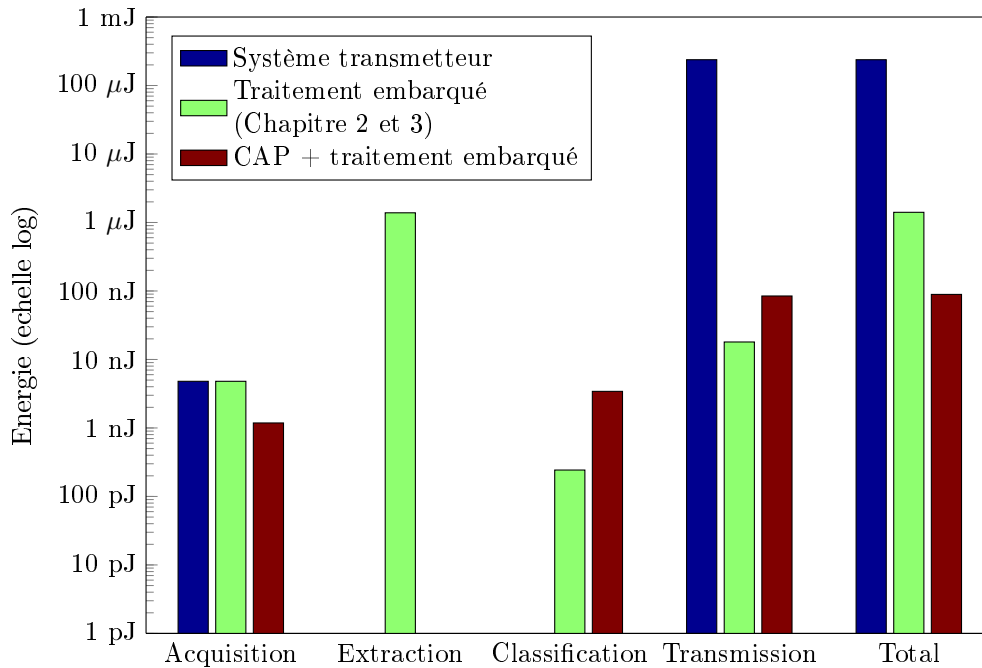


FIGURE 4.11 – Consommation des différents éléments des trois architectures considérées : système purement transmetteur, traitement embarqué et combinaison CAP et traitement embarqué. La consommation correspond au traitement de dix secondes de signal.

système faisant une acquisition classique. Comparé à un système n’effectuant aucun traitement, le gain est de 2 592. Ainsi, l’utilisation des CAP pour l’acquisition et la détection d’anomalie sur les ECG est vraiment intéressante. La combinaison du CAP et d’un classifieur embarqué basse consommation peuvent laisser entrevoir la possibilité d’un système capable de détecter des anomalies pour une énergie de l’ordre du nanojoule ou de la dizaine de nanojoules.

4.3.2 Synchronisation des données

Toutes les simulations réalisées jusqu’à présent supposent que la position du complexe QRS est parfaitement connue. Du fait de la forte localisation des ondelettes dans le temps, un décalage du signal d’entrée peut avoir des conséquences dramatiques sur les performances du système. L’effet d’un décalage temporel du signal est étudié dans cette section. La méthode de sélection des paramètres est limitée aux scores de Fisher et le classifieur utilisé est un perceptron.

Pour évaluer la robustesse du système, des blocs de données similaires à ceux présentés en Section 4.2.1 sont utilisés. La différence est que la position du complexe n’est plus exactement au 100^{ième} échantillon mais suit une loi normale de moyenne 100 et de variance σ^2 . Les performances sont évaluées pour différentes valeurs de σ_t^2 pour le signal d’entrée. Des classifieurs sont entraînés sur ces signaux mal alignés afin de voir l’intérêt de l’apprentissage sur des données plus réalistes. Pour chaque variance d’erreur σ_a^2 considérée lors de l’apprentissage, les paramètres sont sélectionnés via les scores de Fisher. Le nombre de paramètres choisis est de 20. Les résultats sont présentés

sur la Figure 4.12.

On constate qu'avec un apprentissage sur des signaux parfaitement alignés ($\sigma_a^2 = 0$), l'effet d'un mauvais alignement conduit à de fortes pertes de performances. Les performances chutent même sous la limite d'exactitude pour une variance de l'erreur supérieure à 23. L'utilisation de données présentant une erreur d'alignement permet de significativement réduire la chute des performances. Plus la variance utilisée pour l'apprentissage est grande, plus la perte est réduite et plus l'exactitude pour un signal parfaitement aligné est faible. Il faut donc trouver un compromis entre robustesse et performances. Ce compromis dépend de la précision temporelle de la détection du complexe QRS.

Bien que les résultats présentés ici n'utilisent pas les méthodes offrant les meilleurs résultats, ils mettent en évidence l'importance d'avoir une détection précise de la position du complexe QRS. Afin de correctement extraire les paramètres, la position du complexe QRS doit être connue. Cela signifie qu'il faut soit utiliser un système de prédiction de la position du complexe, soit appliquer un délai sur le signal analogique afin de calculer en amont la position du complexe. Bien que l'architecture proposée soit très simple, elle nécessite la mise en place d'un calcul de position de complexe QRS. Si ces éléments n'ont pas été considérés dans cette première étude, il sera indispensable de les traiter pour avoir une évaluation globale de la solution.

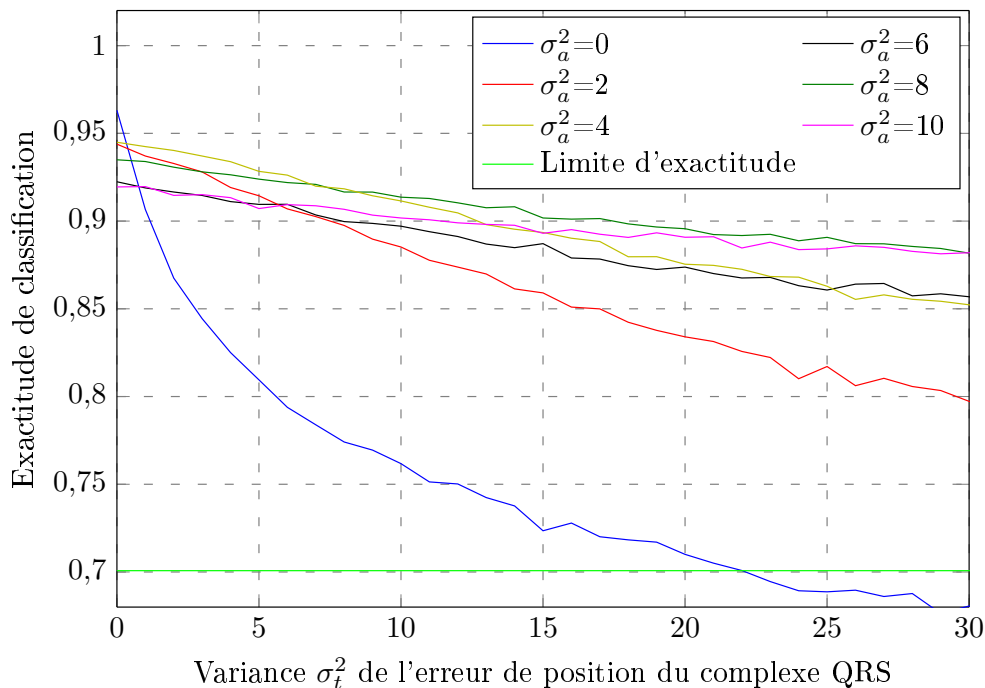


FIGURE 4.12 – Effet d'une mauvaise synchronisation sur les performances de classification. Les classifieurs sont entraînés avec une variance σ_a^2 et testé avec la variance σ_t^2 . Le nombre de paramètres utilisés est de 20.

4.3.3 Quantification des données

L'architecture du CAP proposée utilise un traitement analogique pour extraire des paramètres. Ces derniers sont alors échantillonnés et quantifiés par un CAN. La fréquence d'échantillonnage n'est pas fixe et dépend du support des ondelettes sélectionnées. En considérant les ondelettes sélectionnée par l'algorithme de la Section 4.2.3.3, la fréquence moyenne d'acquisition est divisée par 42,7 par rapport à une acquisition uniforme à 360 Hz. Si la fréquence d'échantillonnage est fortement réduite, le nombre de bits pour effectuer la quantification n'est pas pris en compte dans les simulations précédentes. En effet, ces dernières sont effectuées sous *Matlab*[®] et travaillent en virgule flottante double-précision, soit 64 bits.

Afin d'évaluer le nombre de bits nécessaires une quantification est appliquée suite au calcul des paramètres. Deux forêts aléatoires sont mises en place, l'une ayant été entraînée sur les données non quantifiées, l'autre sur les données quantifiées. Les résultats de classification obtenus sont présentés en Figure 4.13. Dans les deux cas, les paramètres sélectionnés sont ceux issus de l'algorithme de sélection appliqué à des données non quantifiées. Cela explique l'allure des courbes pour une quantification sur deux bits. En effet, le choix des paramètres n'est pas optimisé pour la quantification considérée. D'une manière générale, les performances sont très similaires quelles que soient les données utilisées pour l'apprentissage du classifieur.

On constate que pour un nombre de bits inférieur à quatre, les performances de classification chutent beaucoup. En effet, avec deux bits, dont un pour le signe, les paramètres sont quantifiés sur seulement deux niveaux. Quel que soit le mode d'apprentissage choisi, les quantifications sur six et huit bits présentent des performances quasiment identiques. Cela signifie que l'ajout de précision au delà de six bits n'est pas utile pour l'application considérée. L'utilisation de quatre bits au lieu de six se traduit par une perte de performances de 0,44 point à partir de six paramètres.

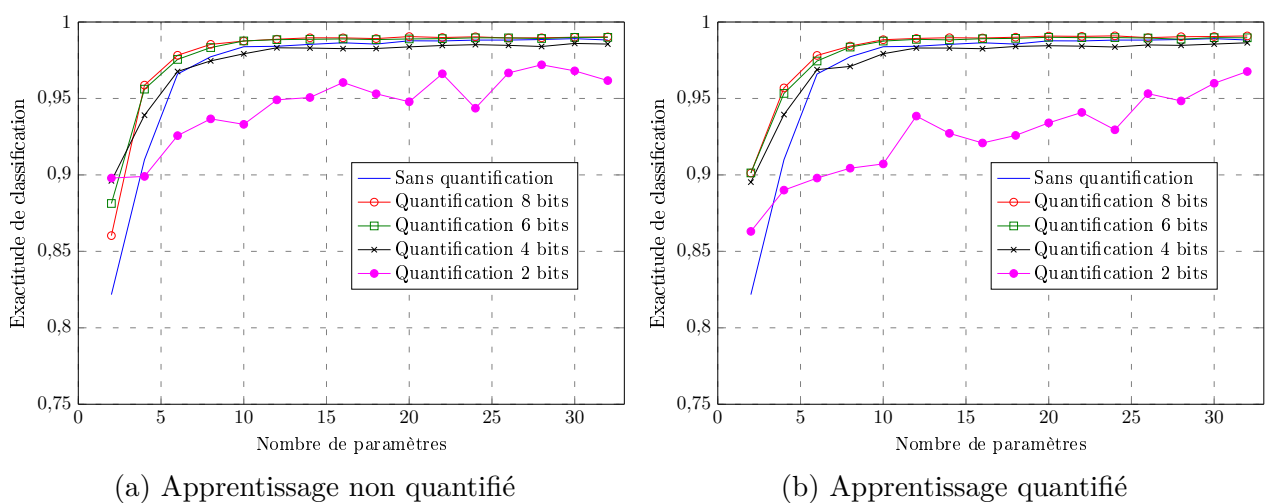


FIGURE 4.13 – Exactitude de classification pour différentes quantifications.

Un élément surprenant est que l'exactitude de classification pour six et huit bits est légèrement supérieure à celle sans quantification. Cela peut être un signe de sur-apprentissage du classifieur. En cas de sur-apprentissage, le classifieur n'apprend plus uniquement les spécificités des signaux de chaque classe mais aussi le bruit. Le bruit étant similaire pour tous les signaux, l'apprendre peut nuire aux performances de classification. Avec la quantification, la perte de précision agit en partie comme un filtrage du bruit de faible amplitude et permet d'améliorer les performances. Cela est particulièrement visible lorsque deux paramètres sont utilisés. Du fait du petit nombre de paramètres, le bruit sur ces derniers peut avoir un gros impact car il n'y a pas d'autres paramètres pour guider la décision du classifieur. L'intérêt de la quantification est visible par le gain de performances obtenu avec l'utilisation des données quantifiées pour l'apprentissage. Afin de réduire cet effet de sur-apprentissage il est possible de réduire la quantité de données utilisée pour l'apprentissage.

Ces tests ne réalisant pas la sélection de paramètres sur les données quantifiées, d'autres tests doivent être menés pour explorer cette piste. Les résultats attendus sont que les performances de classification seront au moins similaires à celles déjà obtenues si ce n'est meilleures. En effet, il semble très peu probable que la sélection de paramètres optimisés dégrade les performances. Si le choix actuel du nombre de bits est de six, les futurs tests permettront peut être de pouvoir descendre à quatre voire moins.

Jusqu'à présent, les tests menés sur la quantification montrent qu'à partir de quatre bits, l'augmentation de la précision ne conduit qu'à une faible augmentation de performances. Pour le CAP, cela signifie qu'il est possible de choisir le nombre de bits pour le CAN en fonction des performances désirées. Si dans un contexte d'application médicale, il n'y a pas de compromis possible sur l'exactitude de classification, cela peut être intéressant pour l'utilisation d'une telle architecture à d'autres applications comme la détection de parole par exemple. En considérant l'acquisition de six paramètres de quatre bits, le classifieur atteint une exactitude de classification de 96,8 %. Avec 12 paramètres, l'exactitude atteint son maximum de 98,3 % tout en ayant acquis seulement 48 bits. L'acquisition classique du signal ECG à 360 Hz sur 11 bits conduit à l'acquisition de 2 816 bits. Ainsi le facteur de compression est de 98,3 %. Bien que la quantité de données acquises soit extrêmement réduite, il est important de garder à l'esprit que ces dernières ne permettent pas de traitements poussés sur le signal. Une utilisation envisagée de cette architecture est en tant que système de réveil. Lorsqu'une anomalie est détectée, le système active un système d'acquisition classique ou d'acquisition compressée puis transmet les données à l'agrégateur pendant une durée déterminée. Ainsi, l'étage RF est utilisé uniquement quand un problème est détecté et l'autonomie du capteur peut être fortement prolongée.

Conclusion

Ce chapitre présente une nouvelle manière d'acquérir des signaux. Elle consiste à effectuer une extraction de paramètres directement à partir du signal analogique afin de réduire fortement la

quantité de données acquises. C'est un modèle très récent et les applications existantes restent très spécifiques à une application. Il permet de fortement réduire la consommation des systèmes de détection devant être constamment actif. Le chapitre propose une architecture de type échantillonnage non uniforme par ondelettes afin d'acquérir des signaux ECG et de détecter la présence d'une anomalie. Cette architecture permet d'extraire un très grand nombre de paramètres différents. La sélection des paramètres les plus pertinents est un défi majeur. Un algorithme est mis en place afin de sélectionner les paramètres offrant les meilleures performances tout en limitant la consommation. Avec un tel algorithme une exactitude de classification de 98% est atteinte pour l'acquisition de seulement six paramètres. De plus, les gains énergétiques, évalués à partir de la consommation d'un circuit ayant une architecture proche de celle du CAP proposé, sont de l'ordre de 26 par rapport au système de traitement numérique proposé aux Chapitres 2 et 3. Ces résultats permettent de montrer l'intérêt d'une telle solution pour réduire la consommation des systèmes d'acquisition et de traitement pour les *BASN*.

L'étude menée dans ce chapitre donne naissance à de nombreuses pistes pour la suite des travaux. La première consiste à perfectionner l'algorithme de sélection pour prendre en compte le coût énergétique ainsi que la robustesse des paramètres sélectionnés. En effet, il peut être préférable de sélectionner des paramètres offrant des performances légèrement réduites dans le cas idéal si ces dernières restent constantes pour des signaux d'entrée mal synchronisés. On voit donc que la sélection optimale des paramètres pour assurer à la fois performances, robustesse et faible énergie est un problème complexe qui doit être étudié. Afin de traiter cette problématique, il est nécessaire d'avoir un modèle de la consommation du système. Ainsi, une seconde perspective est la conception et la réalisation matérielle d'un système mixte réalisant l'architecture proposée.

Conclusion et perspectives

L'évolution des technologies semi-conductrices permet d'effectuer de plus en plus de calculs pour une énergie restreinte. Cette efficacité énergétique ouvre la voie pour l'internet des objets et les réseaux de capteurs, augmentant le nombre d'objets sans fil et connectés. Ces derniers doivent relever de nombreux défis comme la sécurité et la fiabilité des données transmises, l'interopérabilité des différents moyens de communication et l'autonomie énergétique. Une classe d'objets connectés où ces défis sont exacerbés est celle des capteurs corporels en réseau. Ce document apporte des éléments de réponses afin de réduire la consommation de ces derniers.

Dans un premier temps, la chaîne d'acquisition et de traitement classique a été analysée. Cela a permis d'identifier les différentes sources de consommation d'énergie. La transmission sans fil des données vers l'agrégateur a été identifiée comme ayant la moins bonne efficacité énergétique. Pour augmenter l'autonomie des capteurs il a été choisi d'ajouter de l'intelligence dans ces derniers afin de réduire l'utilisation du frontal RF. Suite à l'étude globale de la chaîne, deux axes de recherche ont été privilégiés : la mise en place d'un système de traitement basse consommation des signaux ECG et leur acquisition parcimonieuse.

Le premier axe a été traité via une étude des systèmes de détection d'arythmies sur les signaux ECG. Un système d'analyse des signaux ECG a été proposé dans ce manuscrit. Le système extrait 18 paramètres via une transformation en ondelettes discrètes de dix secondes de signal ECG. Les paramètres sont alors utilisés pour classer le signal en trois types d'arythmies. La nouveauté du système proposé repose dans l'utilisation des réseaux de neurones à cliques pour effectuer la classification. Ils ont été choisis car ils offrent des perspectives de faible consommation [LLAS13]. Les simulations comportementales ont mis en évidence une exactitude de classification de 95 % [CCL⁺16]. Ces résultats sont similaires à ceux obtenus par des circuits basse consommation dédiés au traitement des ECG.

L'architecture générale du système de traitement a ensuite été conçue afin d'estimer sa consommation. Elle est composée d'un système de traitement numérique effectuant l'extraction des paramètres et d'un *ASIC* mixte analogique/numérique pour effectuer la classification. La consommation du système numérique a été estimée à partir de la consommation donnée par *Synopsys Design Compiler*[®] pour des éléments de base de l'architecture. Suite aux tests du circuit de neurone existant [LCL⁺17][CLL⁺17], un *ASIC* flexible a été conçu. Un circuit de 128 neurones a été

fabriqué en utilisant la technologie ST CMOS 65 *nm*. Les performances de classification mesurées sont identiques à celles obtenues avec les simulations comportementales. La consommation de la partie analogique du circuit est de 243 pJ par classification [CLAS17], offrant une réduction de six ordres de grandeur par rapport à un circuit ayant des fonctionnalités similaires [LV13]. La consommation totale du système proposé est de 1,41 μJ pour l'acquisition, le traitement et la transmission des données pour dix secondes de signal. Comparé à un capteur transmetteur, l'intégration d'intelligence a permis une division du budget énergétique par 105.

Afin de réduire encore plus cette consommation, l'acquisition parcimonieuse des signaux ECG a été étudiée. Elle repose sur le concept de conversion analogique vers paramètres qui consiste à effectuer une extraction de paramètres directement sur le signal analogique. Une telle méthode améliore l'efficacité énergétique grâce à l'acquisition d'une quantité très réduite de données. Une architecture mixte analogique/numérique de convertisseur a été proposée afin de détecter la présence d'anomalies dans les ECG. Les simulations comportementales ont montré une exactitude de classification de 98,3 % tout en extrayant seulement six paramètres. Les gains énergétiques espérés sont de l'ordre de 26 par rapport au système précédemment proposé.

Suite à ces travaux, de nombreuses perspectives sont identifiées. Concernant le système d'analyse proposé, une première perspective consiste à automatiser le processus de conception des réseaux à cliques pour la classification. Cette étape est actuellement manuelle et nécessite plusieurs jours voire semaines de travail. Cela rend trop complexe l'ajout de classes dans le réseau ou l'extension à d'autres applications. Une deuxième perspective est de réaliser l'intégration complète du capteur intelligent. Cela permettrait d'obtenir des mesures de la consommation du système numérique au lieu d'estimations.

L'étude sur la conversion analogique vers paramètres ouvre des pistes pour la fabrication d'un premier circuit. Si la discussion sur la quantification montre que six bits semblent suffisants, cette dernière est effectuée avec un ensemble de paramètres non optimisés. Ainsi, une étude plus approfondie permettra de valider ce nombre de bits voire de le réduire. De plus, l'architecture proposée repose sur la supposition d'une connaissance préalable parfaite de la position du complexe QRS. Il est donc indispensable d'étudier plus en détails les effets d'une mauvaise synchronisation pour adapter l'architecture.

La conversion analogique vers paramètre est une opportunité pour augmenter l'attrait envers les méthodes de conception analogique et mixte. L'utilisation d'un classifieur en sortie du convertisseur permet de relâcher de nombreuses contraintes pour la conception de ce dernier. En effet, les imperfections, telles que les non-linéarités ou les problèmes d'appariement, seront prise en compte lors de l'apprentissage. Il est ainsi possible de tirer parti des avantages de l'analogique comme l'efficacité énergétique et la faible surface. En outre, le classifieur peut aisément être configurable et donc offrir une fiabilité accrue à la puce. En effet, cela permet d'effectuer une calibration puce par puce afin d'adapter le classifieur à la variabilité des procédés de fabrication. Il est aussi possible de s'adapter aux variations dues à une utilisation prolongée grâce à une mise à jour régulière du

modèle du classifieur.

Avec la poursuite des travaux sur les deux axes de recherche étudiés, il sera possible de les combiner. L'acquisition parcimonieuse de paramètres alliée à la très faible consommation d'un classifieur embarqué laisse espérer la possibilité de détecter des anomalies sur un signal ECG de 0,71 s avec une énergie de l'ordre de la dizaine de nJ. Une telle énergie permettrait de créer des capteurs fonctionnant à vie avec une simple pile de type PR41 ayant une capacité de 180 mAh et délivrant 1,45 V. Cela ouvre la voie à la démocratisation des réseaux de capteurs corporels améliorant la qualité et le coût des soins.

Bibliographie

- [AY16] M. A. Bin Altaf and J. Yoo. A 1.83 J/Classification, 8-Channel, Patient-Specific Epileptic Seizure Classification SoC Using a Non-Linear Support Vector Machine. *IEEE Transactions on Biomedical Circuits and Systems*, 10(1) :49–60, February 2016.
- [BD10] T. Blumensath and M. E. Davies. Normalized Iterative Hard Thresholding : Guaranteed Stability and Performance. *IEEE Journal of Selected Topics in Signal Processing*, 4(2) :298–309, April 2010.
- [BGM⁺10] A. Burns, B. R. Greene, M. J. McGrath, T. J. O’Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca. SHIMMERTM : A Wireless Sensor Platform for Non-invasive Biomedical Research. *IEEE Sensors Journal*, 10(9) :1527–1534, September 2010.
- [BLE] Adopted Specifications | Bluetooth Technology Website. <https://www.bluetooth.com/specifications/adopted-specifications>. En ligne, accès le 04/06/2017.
- [BLMV15] K. Badami, S. Lauwereins, W. Meert, and M. Verhelst. 24.2 Context-aware hierarchical information-sensing in a 6 μ W 90nm CMOS voice activity detector. In *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*, pages 1–3, February 2015.
- [BMD⁺14] R. Braojos, H. Mamaghanian, A. Dias, G. Ansaloni, D. Atienza, F. J. Rincón, and S. Murali. Ultra-low power design of wearable cardiac monitoring systems. In *2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, pages 1–6, June 2014.
- [BSL⁺12] J. Bae, K. Song, H. Lee, H. Cho, and H. J. Yoo. A 0.24-nJ/b Wireless Body-Area-Network Transceiver With Scalable Double-FSK Modulation. *IEEE Journal of Solid-State Circuits*, 47(1) :310–322, January 2012.
- [BTS⁺16] N. Bayasi, T. Tekeste, H. Saleh, B. Mohammad, A. Khandoker, and M. Ismail. Low-Power ECG-Based Processor for Predicting Ventricular Arrhythmia. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 24(5) :1962–1974, May 2016.
- [BVK⁺15] A. Ba, M. Vidojkovic, K. Kanda, N. F. Kiyani, M. Lont, X. Huang, X. Wang, C. Zhou, Y. H. Liu, M. Ding, B. Búsze, S. Masui, M. Hamaminato, H. Sato, K. Philips, and

- H. de Groot. A 0.33 nJ/bit IEEE802.15.6/Proprietary MICS/ISM Wireless Transceiver With Scalable Data Rate for Medical Implantable Applications. *IEEE Journal of Biomedical and Health Informatics*, 19(3) :920–929, May 2015.
- [CC10] Y. M. Chi and G. Cauwenberghs. Micropower integrated bioamplifier and auto-ranging ADC for wireless and implantable medical instrumentation. In *2010 Proceedings of ESSCIRC*, pages 334–337, September 2010.
- [CCL⁺16] P. Chollet, K. Colombier, C. Lahuec, M. Arzel, and F. Seguin. Toward sub-pJ per classification in Body Area Sensor Networks. In *2016 14th IEEE International New Circuits and Systems Conference (NEWCAS)*, pages 1–4, June 2016.
- [CCS12] F. Chen, A. P. Chandrakasan, and V. M. Stojanovic. Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors. *IEEE Journal of Solid-State Circuits*, 47(3) :744–756, March 2012.
- [Cda04] sur Wikipedia Cdang. Français : Schéma du cœur humain. https://commons.wikimedia.org/wiki/File:Coeur_schema.png?uselang=fr, 2004.
- [CDS01] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Rev.*, 43(1) :129–159, January 2001.
- [CLAS17] P. Chollet, C. Lahuec, M. Arzel, and F. Seguin. A Sub-nJ CMOS ECG Classifier for Wireless Smart Sensor. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017.
- [CLL⁺17] P. Chollet, B. Larras, C. Lahuec, F. Seguin, and M. Arzel. An ultra-low power iterative clique-based neural network integrated in 65-nm CMOS. In *2017 15th IEEE International New Circuits and Systems Conference (NEWCAS)*, pages 5–8, June 2017.
- [CMM⁺13] T. Chen, E. B. Mazomenos, K. Maharatna, S. Dasmahapatra, and M. Niranjan. Design of a Low-Power On-Body ECG Classifier for Remote Cardiovascular Monitoring Systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 3(1) :75–85, March 2013.
- [Coo] Daniel Cooper. 2015 is the year that wearables begin to grow up. <https://www.engadget.com/2015/01/10/2015-ces-wearables-wrap-up/>,. En ligne, accès le 12/04/2017.
- [COR04] Philip de Chazal, M. O’Dwyer, and R. B. Reilly. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, 51(7) :1196–1206, July 2004.
- [CRT06] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles : exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2) :489–509, February 2006.

- [CT06] E. J. Candes and T. Tao. Near-Optimal Signal Recovery From Random Projections : Universal Encoding Strategies? *IEEE Transactions on Information Theory*, 52(12) :5406–5425, December 2006.
- [CTH16] Y. C. Cheng, P. Y. Tsai, and M. H. Huang. Matrix-Inversion-Free Compressed Sensing With Variable Orthogonal Multi-Matching Pursuit Based on Prior Information for ECG Signals. *IEEE Transactions on Biomedical Circuits and Systems*, 10(4) :864–873, August 2016.
- [CW08] E. J. Candes and M. B. Wakin. An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2) :21–30, March 2008.
- [DAGA12] A. M. R. Dixon, E. G. Allstot, D. Gangopadhyay, and D. J. Allstot. Compressed Sensing System Considerations for ECG and EMG Wireless Biosensors. *IEEE Transactions on Biomedical Circuits and Systems*, 6(2) :156–166, April 2012.
- [Don06] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4) :1289–1306, April 2006.
- [ESH⁺16] Fatin A. Elhaj, Naomie Salim, Arief R. Harris, Tan Tian Swee, and Taqwa Ahmed. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Computer Methods and Programs in Biomedicine*, 127 :52–63, April 2016.
- [GAD⁺14] D. Gangopadhyay, E. G. Allstot, A. M. R. Dixon, K. Natarajan, S. Gupta, and D. J. Allstot. Compressed Sensing Analog Front-End for Bio-Sensor Applications. *IEEE Journal of Solid-State Circuits*, 49(2) :426–438, February 2014.
- [Gb05] İnan Güler and Elif Derya Übeyli. ECG beat classifier designed by combined neural network model. *Pattern Recognition*, 38(2) :199–208, February 2005.
- [GB11] V. Gripon and C. Berrou. Sparse Neural Networks With Large Learning Diversity. *IEEE Transactions on Neural Networks*, 22(7) :1087–1096, July 2011.
- [GCDB07] V. Giannini, J. Craninckx, S. D’Amico, and A. Baschiroto. Flexible Baseband Analog Circuits for Software-Defined Radio Front-Ends. *IEEE Journal of Solid-State Circuits*, 42(7) :1501–1512, July 2007.
- [GRR13] J. Garcia, S. Rodriguez, and A. Rusu. A Low-Power CT Incremental 3rd Order /spl Sigma//spl Delta/ ADC for Biosensor Applications. *IEEE Transactions on Circuits and Systems I : Regular Papers*, 60(1) :25–36, January 2013.
- [HHC⁺14] S. Y. Hsu, Y. Ho, P. Y. Chang, C. Su, and C. Y. Lee. A 48.6-to-105.2 μ W Machine Learning Assisted Cardiac Sensor SoC for Mobile Healthcare Applications. *IEEE Journal of Solid-State Circuits*, 49(4) :801–811, April 2014.
- [HHL⁺16] S. Huang, J. Han, X. Li, Z. Yang, and X. Zeng. A low-cost and energy-efficient EEG processor for continuous seizure detection using wavelet transform and AdaBoost. In

- 2016 *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 344–347, October 2016.
- [IEE12] IEEE Standard for Local and metropolitan area networks - Part 15.6 : Wireless Body Area Networks. *IEEE Std 802.15.6-2012*, pages 1–271, February 2012.
- [IEE16] IEEE Standard for Low-Rate Wireless Networks. *IEEE Std 802.15.4-2015 (Revision of IEEE Std 802.15.4-2011)*, pages 1–709, April 2016.
- [IJRV17] S. A. Imtiaz, Z. Jiang, and E. Rodriguez-Villegas. An Ultralow Power System on Chip for Automatic Sleep Staging. *IEEE Journal of Solid-State Circuits*, 52(3) :822–833, March 2017.
- [JA14] B. Johny and A. Anpalagan. Body Area Sensor Networks : Requirements, Operations, and Challenges. *IEEE Potentials*, 33(2) :21–25, March 2014.
- [JH16] J. W. Jhang and Y. H. Huang. A High-SNR Projection-Based Atom Selection OMP Processor for Compressive Sensing. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 24(12) :3477–3488, December 2016.
- [KFC⁺17] F. W. Kuo, S. Binsfeld Ferreira, H. N. R. Chen, L. C. Cho, C. P. Jou, F. L. Hsueh, I. Madadi, M. Tohidian, M. Shahmohammadi, M. Babaie, and R. B. Staszewski. A Bluetooth Low-Energy Transceiver With 3.7-mW All-Digital Transmitter, 2.75-mW High-IF Discrete-Time Receiver, and TX/RX Switchable On-Chip Matching Network. *IEEE Journal of Solid-State Circuits*, 52(4) :1144–1162, April 2017.
- [KHO02] B. U. Kohler, C. Hennig, and R. Orglmeister. The principles of software QRS detection. *IEEE Engineering in Medicine and Biology Magazine*, 21(1) :42–57, January 2002.
- [KK13] V. S. R. Kumari and P. R. Kumar. Cardiac arrhythmia prediction using improved multilayer perceptron neural network. *International Journal of Electronics, Communication, Instrumentation Engineering Research and Development*, 3(4) :73–80, 2013.
- [KKH⁺14] H. Kim, S. Kim, N. Van Helleputte, A. Artes, M. Konijnenburg, J. Huisken, C. Van Hoof, and R. F. Yazicioglu. A Configurable and Low-Power Mixed Signal SoC for Portable ECG Monitoring Applications. *IEEE Transactions on Biomedical Circuits and Systems*, 8(2) :257–267, April 2014.
- [KSC⁺08] Younghun Kim, Thomas Schmid, Zainul M. Charbiwala, Jonathan Friedman, and Mani B. Srivastava. NAWMS : Nonintrusive Autonomous Water Monitoring System. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, SenSys '08, pages 309–322, New York, NY, USA, 2008. ACM.
- [Lar15] Benoît Larras. *CMOS analog implementation of clique-based neural networks*. Theses, Télécom Bretagne ; Université de Bretagne Occidentale, December 2015.
- [LCL⁺17] B. Larras, P. Chollet, C. Lahuec, M. Arzel, and F. Seguin. A 65-nm CMOS 7fJ per synaptic event clique-based neural network in scalable architecture. In *2017 IEEE International Symposium on Circuits and Systems*, 2017.

- [LIY16] B. P. L. Lo, H. Ip, and G. Z. Yang. Transforming Health Care : Body Sensor Networks, Wearables, and the Internet of Things. *IEEE Pulse*, 7(1) :4–8, January 2016.
- [LKD⁺07] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud. Theory and Implementation of an Analog-to-Information Converter using Random Demodulation. In *2007 IEEE International Symposium on Circuits and Systems*, pages 1959–1962, May 2007.
- [LLAS13] B. Larras, C. Lahuec, M. Arzel, and F. Seguin. Analog implementation of encoded neural networks. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pages 1612–1615, May 2013.
- [LLG⁺13] S. T. Li, J. C. Li, X. C. Gu, H. Y. Wang, M. H. Tang, and Z. W. Zhuang. A Continuously and Widely Tunable 5 dB-NF 89.5 dB-Gain 85.5 dB-DR CMOS TV Receiver With Digitally-Assisted Calibration for Multi-Standard DBS Applications. *IEEE Journal of Solid-State Circuits*, 48(11) :2762–2774, November 2013.
- [LLSA16] B. Larras, C. Lahuec, F. Seguin, and M. Arzel. Ultra-Low-Energy Mixed-Signal IC Implementing Encoded Neural Networks. *IEEE Transactions on Circuits and Systems I : Regular Papers*, 63(11) :1974–1985, November 2016.
- [LLW⁺13] Xiaochen Lai, Quanli Liu, Xin Wei, Wei Wang, Guoqiao Zhou, and Guangyi Han. A Survey of Body Sensor Networks. *Sensors (Basel, Switzerland)*, 13(5) :5406–5447, April 2013.
- [LRYR16] C. Laurenson, F. Rivet, M. R. Yuce, and J. M. Redouté. A 180 nm CMOS analog adaptive sampler for blood pressure feature extraction. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4841–4844, August 2016.
- [LSCCM16] Eduardo José da S. Luz, William Robson Schwartz, Guillermo Cámara-Chávez, and David Menotti. ECG-based heartbeat classification for arrhythmia detection : A survey. *Computer Methods and Programs in Biomedicine*, 127 :144–164, April 2016.
- [LV13] K. H. Lee and N. Verma. A Low-Power Processor With Configurable Embedded Machine-Learning Accelerators for High-Order and Adaptive Analysis of Medical-Sensor Signals. *IEEE Journal of Solid-State Circuits*, 48(7) :1625–1637, July 2013.
- [MAM13] Roshan Joy Martis, U. Rajendra Acharya, and Lim Choo Min. ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, 8(5) :437–448, September 2013.
- [MCXF16] Y. Mohan, S. S. Chee, D. K. P. Xin, and L. P. Foong. Artificial neural network for classification of depressive and normal in EEG. In *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pages 286–290, December 2016.
- [MIT] MIT-BIH Arrhythmia Database. <http://physionet.org/physiobank/database/mitdb/>. En ligne, accès le 18/01/2017.

- [MKAV11] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vanderghenst. Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes. *IEEE Transactions on Biomedical Engineering*, 58(9) :2456–2466, September 2011.
- [MM92] George B Moody and Roger G Mark. MIT-BIH Atrial Fibrillation Database, 1992. DOI : 10.13026/C2MW2D.
- [MMCT09] M. Mollazadeh, K. Murari, G. Cauwenberghs, and N. Thakor. Micropower CMOS Integrated Low-Noise Amplification, Filtering, and Digitization of Multimodal Neuron Potentials. *IEEE Transactions on Biomedical Circuits and Systems*, 3(1) :1–10, February 2009.
- [MZM⁺11] T. Mar, S. Zaunseder, J. P. Martínez, M. Llamedo, and R. Poll. Optimization of ECG Classification by Means of Feature Selection. *IEEE Transactions on Biomedical Engineering*, 58(8) :2168–2177, August 2011.
- [NT09] D. Needell and J. A. Tropp. CoSaMP : Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3) :301–321, May 2009.
- [OAT14] Rami J. Oweis and Basim O. Al-Tabbaa. Qrs detection and heart rate variability analysis : a survey. *Biomed Sci Eng*, 2(1) :13–34, 2014.
- [OMS] OMS | Les 10 principales causes de mortalité. <http://www.who.int/mediacentre/factsheets/fs310/fr/>. En ligne, accès le 29/03/2017.
- [OST⁺16] H. Okuni, A. Sai, T. T. Ta, S. Kondo, T. Tokairin, M. Furuta, and T. Itakura. 26.1 A 5.5mw ADPLL-based receiver with hybrid-loop interference rejection for BLE application in 65nm CMOS. In *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 436–437, January 2016.
- [Pal13] Günther Palm. Neural associative memories and sparse coding. *Neural Networks*, 37 :165–171, January 2013.
- [PPG16] A. M. da Silva Pinho, N. Pombo, and N. M. Garcia. Sleep apnea detection using a feed-forward neural network on ECG signal. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–6, September 2016.
- [Pri] PricewaterhouseCoopers. Consumer Intelligence Series : The wearable life 2.0. <https://www.pwc.com/us/en/industry/entertainment-media/publications/consumer-intelligence-series/wearables.html>. En ligne, accès le 12/04/2017.
- [PS84] O. Pahlm and L. Sörnmo. Software QRS detection in ambulatory monitoring — a review. *Medical and Biological Engineering and Computing*, 22(4) :289–297, July 1984.
- [PS17] Michaël Pelissier and Christoph Studer. Non-Uniform Wavelet Sampling for RF Analog-to-Information Conversion. *IEEE Transactions on Circuits and Systems I : Regular Papers*, (To appear), July 2017.

- [PSS⁺15] A. Page, C. Sagedy, E. Smith, N. Attaran, T. Oates, and T. Mohsenin. A Flexible Multichannel EEG Feature Extractor and Classifier for Seizure Detection. *IEEE Transactions on Circuits and Systems II : Express Briefs*, 62(2) :109–113, February 2015.
- [PT85] J. Pan and W. J. Tompkins. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3) :230–236, March 1985.
- [RLN⁺08] T. Ragheb, J. N. Laska, H. Nejati, S. Kirolos, R. G. Baraniuk, and Y. Massoud. A prototype hardware for random demodulation based compressive analog-to-digital conversion. In *2008 51st Midwest Symposium on Circuits and Systems*, pages 37–40, August 2008.
- [RM16] F. Ren and D. Marković. A Configurable 12-237 kS/s 12.8 mW Sparse-Approximation Engine for Mobile Data Aggregation of Compressively Sampled Physiological Signals. *IEEE Journal of Solid-State Circuits*, 51(1) :68–78, January 2016.
- [RRDLL14] N. Ravanshad, H. Rezaee-Dehsorkh, R. Lotfi, and Y. Lian. A Level-Crossing Based QRS-Detection Algorithm for Wearable ECG Sensors. *IEEE Journal of Biomedical and Health Informatics*, 18(1) :183–192, January 2014.
- [RSCB14] Priyanka Rawat, Kamal Deep Singh, Hakima Chaouchi, and Jean Marie Bonnin. Wireless sensor networks : a survey on recent developments and potential synergies. *The Journal of Supercomputing*, 68(1) :1–48, April 2014.
- [SML⁺04] Gyula Simon, Miklós Maróti, Ákos Lédeczi, György Balogh, Branislav Kusy, András Nádas, Gábor Pap, János Sallai, and Ken Frampton. Sensor Network-based Countersniper System. In *Proceedings of the 2Nd International Conference on Embedded Networked Sensor Systems*, SenSys '04, pages 1–12, New York, NY, USA, 2004. ACM.
- [ST06] Brij N. Singh and Arvind K. Tiwari. Optimal selection of wavelet basis function applied to ECG signal denoising. *Digital Signal Processing*, 16(3) :275–287, May 2006.
- [TAL14] Jiliang Tang, Salem Alelyani, and Huan Liu. Feature selection for classification : A review. *Data Classification : Algorithms and Applications*, page 37, 2014.
- [TG07] J. A. Tropp and A. C. Gilbert. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory*, 53(12) :4655–4666, December 2007.
- [TSH⁺17] H. Töreyn, S. Shah, S. Hersek, O. T. Inan, and J. Hasler. Proof-of-concept energy-efficient and real-time hemodynamic feature extraction from bioimpedance signals using a mixed-signal field programmable analog array. In *2017 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 233–236, February 2017.
- [VB15] M. Verhelst and A. Bahai. Where Analog Meets Digital : Analog-to-Information Conversion and Beyond. *IEEE Solid-State Circuits Magazine*, 7(3) :67–80, 2015.

- [VC07] N. Verma and A. P. Chandrakasan. An Ultra Low Energy 12-bit Rate-Resolution Scalable SAR ADC for Wireless Sensor Nodes. *IEEE Journal of Solid-State Circuits*, 42(6) :1196–1205, June 2007.
- [VSB⁺10] N. Verma, A. Shoeb, J. Bohorquez, J. Dawson, J. Guttag, and A. P. Chandrakasan. A Micro-Power EEG Acquisition SoC With Integrated Feature Extraction Processor for a Chronic Seizure Detection System. *IEEE Journal of Solid-State Circuits*, 45(4) :804–816, April 2010.
- [WALR⁺06] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh. Deploying a wireless sensor network on an active volcano. *IEEE Internet Computing*, 10(2) :18–25, March 2006.
- [WBLH69] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222(5197) :960–962, June 1969.
- [WCHY13] Jeen-Shing Wang, Wei-Chun Chiang, Yu-Liang Hsu, and Ya-Ting C. Yang. ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing*, 116 :38–45, September 2013.
- [WDD⁺13] A. C. W. Wong, M. Dawkins, G. Devita, N. Kasparidis, A. Katsiamis, O. King, F. Lauria, J. Schiff, and A. J. Burdett. A 1 V 5 mA Multimode IEEE 802.15.6/Bluetooth Low-Energy WBAN Transceiver for Biotelemetry Applications. *IEEE Journal of Solid-State Circuits*, 48(1) :186–198, January 2013.
- [WFKA16] J. Wang, T. Fujiwara, T. Kato, and D. Anzai. Wearable ECG Based on Impulse-Radio-Type Human Body Communication. *IEEE Transactions on Biomedical Engineering*, 63(9) :1887–1894, September 2016.
- [WZL⁺15] C. Wang, J. Zhou, L. Liao, J. Lan, J. Luo, X. Liu, and M. Je. Near-Threshold Energy- and Area-Efficient Reconfigurable DWPT/DWT Processor for Healthcare-Monitoring Applications. *IEEE Transactions on Circuits and Systems II : Express Briefs*, 62(1) :70–74, January 2015.
- [YMPH08] R. F. Yazicioglu, P. Merken, R. Puers, and C. Van Hoof. A 200 W Eight-Channel EEG Acquisition ASIC for Ambulatory EEG Systems. *IEEE Journal of Solid-State Circuits*, 43(12) :3025–3038, December 2008.
- [YYED⁺13] J. Yoo, L. Yan, D. El-Damak, M. A. B. Altaf, A. H. Shoeb, and A. P. Chandrakasan. An 8-Channel Scalable EEG Acquisition SoC With Patient-Specific Seizure Classification and Recording Processor. *IEEE Journal of Solid-State Circuits*, 48(1) :214–228, January 2013.
- [YYKY10] L. Yan, J. Yoo, B. Kim, and H. J. Yoo. A 0.5- V 12- W Wirelessly Powered Patch-Type Healthcare Sensor for Wearable Body Sensor Network. *IEEE Journal of Solid-State Circuits*, 45(11) :2356–2365, November 2010.
- [ZAAB16] X. Zhai, A. A. S. Ali, A. Amira, and F. Bensaali. MLP Neural Network Based Gas Classification System on Zynq SoC. *IEEE Access*, 4 :8138–8146, 2016.

- [ZAY16] C. Zhang, M. A. Bin Altaf, and J. Yoo. Design and Implementation of an On-Chip Patient-Specific Closed-Loop Seizure Onset and Termination Detection System. *IEEE Journal of Biomedical and Health Informatics*, 20(4) :996–1007, July 2016.
- [ZDL⁺14] Zhancheng Zhang, Jun Dong, Xiaoqing Luo, Kup-Sze Choi, and Xiaojun Wu. Heart-beat classification using disease-specific feature selection. *Computers in Biology and Medicine*, 46 :79–89, March 2014.
- [Zig] Zigbee alliance. <http://www.zigbee.org/>. En ligne, accès le 26/04/2017.
- [ZJW⁺13] L. Zhang, H. Jiang, J. Wei, J. Dong, F. Li, W. Li, J. Gao, J. Cui, B. Chi, C. Zhang, and Z. Wang. A Reconfigurable Sliding-IF Transceiver for 400 MHz/2.4 GHz IEEE 802.15.6/ZigBee WBAN Hubs With Only 21% Tuning Range VCO. *IEEE Journal of Solid-State Circuits*, 48(11) :2705–2716, November 2013.
- [ZL14] X. Zhang and Y. Lian. A 300-mV 220-nW Event-Driven ADC With Real-Time QRS Detection for Wearable ECG Sensors. *IEEE Transactions on Biomedical Circuits and Systems*, 8(6) :834–843, December 2014.
- [ZWF14] C. Zhang, H. Wang, and R. Fu. Automated Detection of Driver Fatigue Based on Entropy and Complexity Measures. *IEEE Transactions on Intelligent Transportation Systems*, 15(1) :168–177, February 2014.

Les réseaux de capteurs corporels représentent un enjeu sociétal important en permettant des soins de meilleure qualité avec un coût réduit. Ces derniers sont utilisés pour détecter des anomalies dès leur apparition et ainsi intervenir au plus vite. Les capteurs sont soumis à de nombreuses contraintes de fiabilité, robustesse, taille et consommation.

Dans cette thèse, les différentes opérations réalisées par les réseaux de capteurs corporels sont analysées. La consommation de chacune d'elles est évaluée afin de guider les axes de recherche pour améliorer l'autonomie énergétique des capteurs. Un capteur pour la détection d'arythmie sur des signaux cardiaques est proposé. Il intègre un traitement du signal via l'utilisation d'un réseau de neurone à cliques. Le système proposé est simulé et offre une exactitude de classification de 95 % pour la détection de trois types d'arythmie. Le prototypage du système via la fabrication d'un circuit mixte analogique/numérique en CMOS 65 nm montre une consommation du capteur de l'ordre de 1,4 μ J.

Pour réduire encore plus l'énergie, une nouvelle méthode d'acquisition est utilisée. Une architecture de convertisseur est proposée pour l'acquisition et le traitement de signaux cardiaques. Cette dernière laisse espérer une consommation de l'ordre de 1,18 nJ pour acquérir les paramètres tout en offrant une exactitude de classification proche de 98 %. Cette étude permet d'ouvrir la voie vers la mise en place de capteurs très basse consommation pouvant durer toute une vie avec une simple pile.

Mots-clés : Réseaux de capteurs corporels, Intégration mixte analogique/numérique, Réseaux de neurones à cliques, Acquisition parcimonieuse

Body area sensor networks gained great focused through the promise of better quality and cheaper medical care system. They are used to detect anomalies and treat them as soon as they arise. Sensors are under heavy constraints such as reliability, sturdiness, size and power consumption.

This thesis analyzes the operations perform by a body area sensor network. The different energy requirements are evaluated in order to choose the focus of the research to improve the battery life of the sensors. A sensor for arrhythmia detection is proposed. It includes some signal processing through a clique-based neural network. The system simulations allow a classification between three types of arrhythmia with 95 % accuracy. The prototype, based on a 65 nm CMOS mixed signal circuit, requires only 1.4 μ J.

To further reduce energy consumption, a new sensing method is used. A converter architecture is proposed for heart beat acquisition. Simulations and estimation show a 1.18 nJ energy requirement for parameter acquisition while offering 98 % classification accuracy. This work leads the way to the development of low energy sensor with a lifetime battery life.

Keywords: Body area sensor networks, Mixed signal integration, Clique-based neural networks, Sparse sensing

