



HAL
open science

Etude de la consommation énergétique de systèmes de communications numériques sans fil implantés sur cible FPGA

Jordane Lorandel

► **To cite this version:**

Jordane Lorandel. Etude de la consommation énergétique de systèmes de communications numériques sans fil implantés sur cible FPGA. Electronique. INSA de Rennes, 2015. Français. NNT : 2015ISAR0036 . tel-01411846

HAL Id: tel-01411846

<https://theses.hal.science/tel-01411846>

Submitted on 7 Dec 2016

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

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



THÈSE INSA Rennes
sous le sceau de l'Université Européenne de Bretagne
pour obtenir le grade de
DOCTEUR DE L'INSA DE RENNES
Spécialité : Électronique et Télécommunications

présentée par

Jordane Lorandel

ÉCOLE DOCTORALE : MATISSE
LABORATOIRE : IETR

Étude de la
consommation
énergétique de
systèmes de
communications
numériques sans fil
implantés sur cible
FPGA

Thèse soutenue le 8 décembre 2015

devant le jury composé de :

Olivier Berder

Professeur à l'IUT de Lannion / *Président*

Emmanuel Boutillon

Professeur à l'Université de Bretagne Sud (Lorient) / *Rapporteur*

Bertrand Granado

Professeur à l'Université Pierre et Marie Curie (Paris) / *Rapporteur*

Dominique Houzet

Professeur à l'Université de Grenoble / *Examineur*

Cécile Belleudy

Maître de conférences à l'Université de Nice / *Examinatrice*

Maryline Hélard

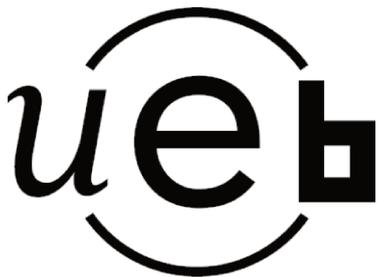
Professeur à l'INSA de Rennes / *Directrice de thèse*

Jean-Christophe Prévotet

Maître de conférences à l'INSA de Rennes / *Co-encadrant de thèse*

Etude de la consommation énergétique de systèmes de communications numériques sans fil implantés sur cible FPGA

Jordane LORANDEL



En partenariat avec

--	--	--	--	--

*Je ne cherche pas à connaître les réponses,
je cherche à comprendre les questions.*
Confucius

Remerciements

Dans un premier temps, je tiens à remercier chaleureusement Maryline H elard, Professeur   l'INSA de Rennes et Jean-Christophe Pr evotet, Ma tre de conf erences   l'INSA de Rennes, pour m'avoir donn e l'opportunit e de pouvoir effectuer cette th ese, dans de tr es bonnes conditions. Leurs disponibilit es, ainsi que leurs aides respectives, leurs conseils et leurs encouragements, m'ont  t e d'une aide pr ecieuse tout au long de ces 3 ann ees. Ce f ut une collaboration des plus enrichissantes pour moi, aussi bien professionnellement qu'humanement, et je leur en suis tr es reconnaissant.

Je remercie aussi l'ensemble des membres du jury qui m'ont fait l'honneur d' tre pr esents lors de ma soutenance. Ainsi, je remercie Olivier BERDER, Professeur   l'IUT de Lannion, qui me fait l'honneur de pr esider ce jury. Je remercie M. Granado Bertrand et M. Boutillon Emmanuel d'avoir bien voulu accepter la charge de rapporteur. Leurs remarques constructives m'ont permis d'enrichir et d'am liorer le contenu de ce manuscrit. Je remercie aussi Mme Belleudy C ecile et M. Houzet Dominique pour avoir jug e mes travaux.

Une partie des travaux a  t e men e   travers un contrat de recherche,  tabli conjointement entre l'INSA de Rennes et Orange Labs de Rennes. Je tiens   remercier l'ensemble des collaborateurs, avec qui j'ai eu l'occasion de travailler, pour leurs conseils et analyses qui ont permis d'orienter et de faire progresser ces travaux.

Je tiens aussi   remercier l'ensemble du personnel du b atiment 6 de l'INSA de Rennes que j'ai c otoy e durant mon stage de fin de Master puis durant ma th ese. Je fait un clin d'oeil particulier   l'ensemble des doctorants/post-docs/ATERs et stagiaires qui ont fait vivre ce fameux 'bocal'.

Je tiens aussi   remercier tout particuli erement Yvan Kokar, Ing enieur de recherche   l'INSA de Rennes, pour les nombreuses discussions passionnantes, son expertise ainsi que sa disponibilit e.

Pour finir, je tiens   avoir une pens ee pour ma famille et mes proches qui m'ont soutenu durant les diff erentes  tapes de mes  tudes et de mon doctorat.

Table des matières

Table des matières	1
Table des figures	4
Acronymes	7
Introduction	9
1 Les systèmes de communications numériques sans fil	14
1.1 Introduction	14
1.2 Caractéristiques d'un système de communications numériques	14
1.2.1 Le canal et ses effets	15
1.3 La modulation OFDM	18
1.3.1 Caractéristiques de l'OFDM	18
1.3.2 Réalisation matérielle	22
1.3.3 Avantages et limitations de l'OFDM	23
1.4 OFDM dans les réseaux mobiles	23
1.4.1 Le réseau mobile LTE et ses propriétés	24
1.4.2 Configuration de systèmes à multiples antennes	29
1.4.2.1 Notion de diversité et de capacité	29
1.4.2.2 Technique de retournement temporel	31
1.5 Performances, enjeux énergétiques et problématiques	33
1.5.1 Évaluation des performances	34
1.5.2 Estimation de la consommation d'énergie	34
1.5.3 Problématiques	35
1.6 Conclusion	37
2 La consommation d'énergie dans les circuits de type FPGA	38
2.1 Introduction	38
2.2 Les circuits FPGA	39
2.2.1 Qu'est-ce qu'un FPGA ?	39
2.2.2 Flexibilité versus performances	42
2.2.3 Flot de conception d'un FPGA	43
2.3 Origines de la consommation dans un FPGA	44
2.3.1 Puissance statique	44
2.3.2 Puissance dynamique	47
2.3.3 Quelques techniques de réduction de la consommation énergétique	50

2.4	Problématique de l'estimation de la consommation des circuits FPGA	51
2.5	Conclusion	52
3	Techniques d'estimation de la consommation de systèmes sur FPGA	54
3.1	Techniques d'estimation de la consommation	56
3.1.1	Niveau technologique	56
3.1.2	Niveau Porte logique	57
3.1.3	Niveau Register Transfer Level (RTL)	60
3.1.4	Mesures réelles et émulation	62
3.1.5	Niveau Système	63
3.2	Conclusion	68
4	Méthodologie d'estimation de la consommation de systèmes sur FPGA	70
4.1	Introduction	70
4.2	Le concept de scénario	71
4.3	Caractérisation des blocs IP matériels	74
4.4	Modélisation haut-niveau en SystemC des blocs IP matériels	78
4.5	Conclusion	82
5	Évaluations haut-niveau de systèmes implantés sur FPGA	84
5.1	Introduction	84
5.2	Descriptions et évaluations des chaînes de communications numériques	84
5.2.1	Chaîne d'émission-réception SISO-OFDM	84
5.2.1.1	Évaluation de la puissance dynamique	86
5.2.1.2	Évaluation du temps d'estimation	88
5.2.1.3	Évaluation de l'efficacité énergétique	90
5.2.2	Émetteur LTE SISO-OFDM en voie descendante du canal PDSCH	94
5.2.2.1	Évaluation de la puissance dynamique	95
5.2.2.2	Évaluation des temps d'estimation	96
5.2.2.3	Évaluation de l'efficacité énergétique	97
5.2.3	Émetteur LTE MISO-OFDM 2x1 en voie descendante avec encodage d'Alamouti	100
5.2.3.1	Évaluation de la puissance dynamique	100
5.2.3.2	Évaluation des temps d'estimation	101
5.2.4	Émetteur d'une chaîne MISO-OFDM à X antennes avec pré-codage de retournement temporel en fréquence	102
5.3	Analyse de l'efficacité énergétique de techniques coopératives dans le contexte LTE	108
5.3.1	Schémas de relayage et protocoles de retransmission en LTE	108
5.3.2	Modélisation énergétique	110
5.3.2.1	Analyse de l'efficacité énergétique	113
5.3.2.2	Analyse du compromis efficacité spectrale-efficacité éner- gétique	119
5.4	Conclusion	121

6	Optimisations de la méthodologie d'estimation	124
6.1	Limitations de l'approche proposée	124
6.2	Modélisation sous forme analytique de la consommation	124
6.2.1	Modélisation de blocs IP	125
6.2.2	Extension à des systèmes complets	127
6.2.3	Limitations	128
6.2.4	Généralisation à différentes familles de FPGA	129
6.3	Bilan de la modélisation de bloc IP sous forme analytique	131
6.4	Modélisation de blocs IP à l'aide de réseaux de neurones	131
6.4.1	Les réseaux de neurones	131
6.4.2	Modélisation	133
6.4.3	Entraînement, tests et validation	134
6.4.4	Exemples de modélisation	134
6.5	Conclusion	137
A	Annexes A	143
A.1	Chaîne d'émission-réception SISO-OFDM	143
A.2	Émetteur LTE SISO-OFDM en voie descendante du canal PDSCH	143
	Bibliographie	152

Table des figures

1.1	Schéma de base des traitements en bande de base d'une chaîne de télécommunication	15
1.2	Détermination de la réponse impulsionnelle d'un canal	16
1.3	Division d'une bande W en N sous-bandes	18
1.4	Spectre fréquentiel d'un signal OFDM	19
1.5	Comparaisons entre une modulation mono-porteuse et l'OFDM	20
1.6	Schématique d'un modulateur OFDM	20
1.7	Constellations en BPSK, QPSK et QAM16	21
1.8	Illustration du principe de l'intervalle de garde face aux interférences entre symboles OFDM	22
1.9	Schéma d'un modulateur OFDM à base d'IFFT	22
1.10	Schéma d'un démodulateur OFDM à base de FFT	23
1.11	Évolution des technologies des réseaux mobiles [YB12]	24
1.12	Schéma représentatif du réseau eUTRAN [YB12]	25
1.13	Structure de la trame LTE de Type 1 (FDD)	26
1.14	Structure de la trame LTE de Type 1 (FDD)	27
1.15	Allocation de ressource en LTE	28
1.16	Configurations de systèmes SISO, MISO, SIMO et MIMO	29
1.17	Principe du Retournement Temporel [HPCK15]	31
1.18	Pic d'auto-corrélation en réception	32
1.19	Intégration du Retournement Temporel dans un système	32
1.20	Illustration du principe de Retournement Temporel dans un système MISO	33
1.21	Répartition de la consommation des éléments d'une station de base suivant différents types de déploiement	36
2.1	Évolution de la puissance dissipée, comparaison de la technologie bipolaire et CMOS [CSC99]	39
2.2	Positionnement des FPGAs	40
2.3	Architecture simplifiée des FPGAs	41
2.4	Exemple d'architecture de FPGA	42
2.5	Flot de conception typique de FPGA	43
2.6	Courants de fuite d'un transistor NMOS	45
2.7	Evolution de la puissance statique en fonction de l'évolution technologique	47
2.8	Illustration d'une cellule CMOS	48
2.9	Illustration du phénomène de glitch	49
2.10	Illustration de la technique du <i>clock gating</i>	51

2.11	Évolution des puissances dynamique et statique en fonction de l'évolution technologique	52
3.1	Représentation des niveaux d'abstraction	55
3.2	Illustration du Transaction Level Modelling	56
3.3	Flot d'utilisation de l'outil XPower Analyzer	60
4.1	Introduction du concept de scénario	71
4.2	Schéma d'un émetteur et d'un récepteur d'une chaîne SISO-OFDM	73
4.3	Illustration de temps d'activité différents pour 2 applications	74
4.4	Caractérisation de chaque bloc IP matériel	75
4.5	Modèle d'implémentation de l'IP	79
4.6	Illustration de l'interface générique basée sur le standard AXI	80
5.1	Description de la chaîne SISO-OFDM	86
5.2	Schéma de comparaison de l'approche proposée et celle d'XPower	87
5.3	Répartition de la consommation par IP pour l'émetteur de la chaîne SISO-OFDM	89
5.4	Répartition de la consommation par IP pour le récepteur de la chaîne SISO-OFDM	90
5.5	Efficacité énergétique moyenne (bit/J) en fonction de la puissance totale consommée (dBm) et de la puissance de transmission	93
5.6	Émetteur LTE SISO-OFDM en voie descendante	95
5.7	Résultats d'estimation entre notre méthodologie et XPower pour un émetteur SISO-OFDM en LTE	97
5.8	Efficacité énergétique moyenne (bit/J) en fonction de la puissance totale consommée (dBm) et de la puissance de transmission (dBm)	99
5.9	Traitements en bande de base du canal PDSCH en LTE DL dans une configuration MISO 2x1 avec codage d'Alamouti	100
5.10	Répartition de la puissance dynamique consommée de l'émetteur MISO-OFDM 2x1 en fonction de la bande de fréquence	101
5.11	Schématique d'émetteurs SISO et MISO-OFDM avec Retournement Temporel appliqué dans le domaine fréquentiel	102
5.12	Illustration de la méthode d'évaluation de l'EE à puissance de transmission P_t constante	103
5.13	Efficacité énergétique vs puissance totale consommée pour plusieurs configurations avec P_{CSISO} et $P_{CRT} = 0W$	105
5.14	Efficacité énergétique vs puissance totale consommée pour plusieurs configurations avec P_{CSISO} et P_{CRT} considérées	105
5.15	Efficacité énergétique vs puissance de transmission P_t pour plusieurs configurations avec P_{CSISO} or P_{CRT} considérées	107
5.16	Protocoles de relayage avec mécanismes de retransmission HARQ	109
5.17	Protocole de retransmission HARQ-CC pour $N_{max} = 3$	110
5.18	Protocole de retransmission HARQ-IR pour $N_{max} = 3$	111
5.19	Schéma des traitements en bande de base considérés	111
5.20	Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-CC	114

5.21	Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-IR	114
5.22	Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-CC, toutes les sources de consommation sont considérées . . .	115
5.23	Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-IR, toutes les sources de consommation sont considérées . . .	116
5.24	Efficacité énergétique versus puissance de transmission (dBm) pour différentes considérations lors du calcul de l'efficacité énergétique	117
5.25	Maximum d'efficacité énergétique en fonction de la puissance de transmission	118
5.26	Efficacité énergétique versus Efficacité spectrale : les deux protocoles (HARQ-IR et HARQ-CC), les 3 modulations (QPSK, 16QAM 64QAM) et toutes les sources de consommation sont considérés	120
5.27	Efficacité énergétique pour les schémas AF et DF versus puissance de transmission (dBm) pour $N_{\max} = 2$	121
5.28	Efficacité énergétique pour les schémas AF et DF versus puissance de transmission (dBm) pour $N_{\max} = 6$	122
6.1	Comparaison de la précision du modèle analytique et des valeurs originales -FPGA Virtex 6 - 14bits	126
6.2	Puissance dynamique moyenne en fonction des différentes familles de FPGAs	129
6.3	Schéma d'un réseau de type perceptron multi-couches	132
6.4	Schéma du procédé de modélisation à partir des deux réseaux de neurones .	134
6.5	Schéma de l'émetteur basique OFDM modélisé	135

Acronymes

3GPP : 3rd Generation Partnership Project
AF : Amplify and Forward
ASIC : Application-Specific Integrated Circuit
AT : Approximately Timed
BPSK : Binary Phase Shift Keying
CA : Cycle Accurate
CC : Chase Combining
CDMA : Code Division Multiple Access
CMOS : Complementary Metal Oxide Semi-conductor
DF : Decode and Forward
DSP : Digital Signal Processor
eUTRAN : evolved Universal Terrestrial Radio Access Network
FPGA : Field Programmable Gate Array
GSM : Global System for Mobile communication
HARQ : Hybrid Automatic Repeat ReQuest
IEEE : Institute of Electrical and Electronics Engineers
IFFT : Inverse Fast Fourier Transform
IP : Intellectual Property
IR : Incremental Redondancy
LT : Loosely Timed
LTE : Long Term Evolution
MIMO : Mutliple-Inputs Mutliple-Outputs
MISO : Mutliple-Inputs Single-Output
MPSoC : Multiple Processors System on Chip
OFDM : Orthogonal Frequency Division Multiplexing
PDSCH : Physical Downlink Shared CHannel
QAM : Quadrature Amplitude Modulation
QPSK : Quadrature Phase Shift Keying
ROM : Read-Only Memory
RTL : Register Transfer Level
SISO : Single-Input Single-Output
SIMO : Single-Input Multiple-Outputs
SoC : System on Chip
SRAM : Static Random Access Memory
TR : Time Reversal
UMTS : Univeral Mobile Telecommunications System
VHDL : Very high speed integrated circuit Hardware Description Language

Introduction

Contexte

Les systèmes de télécommunication n'ont cessé d'évoluer ces dernières années, poussés par une forte demande du marché en systèmes toujours plus autonomes, performants et efficaces en énergie. Le secteur des Technologies de l'Information et de la Communication (TIC), qui regroupe de nombreux domaines dont les télécommunications et internet, connaît une croissance incroyable depuis quelques années, en majorité portée par la commercialisation de nombreux objets connectés comme les tablettes tactiles, les smartphones et les ordinateurs portables. Ainsi, ce type de périphériques utilise généralement des réseaux sans fil comme le *Wireless Fidelity* (Wi-Fi) ou bien des réseaux mobiles comme le *Long-Term Evolution* (LTE).

En France, cette croissance se traduit par une augmentation considérable du volume de données échangées par les clients des opérateurs mobiles. En effet, rien que le volume de données (sms et communications vocales) a atteint une croissance annuelle de 107% au troisième trimestre 2014 et est en perpétuelle augmentation [ARC15]. Ainsi, le volume est passé d'environ 40000 téraoctets¹ au 3ème trimestre 2013 à plus de 83000 téraoctets au dernier trimestre 2014.

A l'échelle mondiale, le nombre de données sur le réseau mobile a augmenté de 69% en 2014. Une augmentation de 57% par an doit se poursuivre entre 2014 et 2019 [CIS15].

La multiplication des équipements capables d'échanger et de transmettre des données via Internet étant considérable, on parle aujourd'hui de l'internet des objets, *Internet of Things*. Comme le montre l'étude dans [Str14], il y aurait aujourd'hui plus de 12 milliards d'objets connectés à internet à travers le monde. D'ici à 2020, ce chiffre pourrait atteindre les 33 milliards.

La plupart de ces appareils communicants sont généralement composés de microprocesseurs associés à de multiples périphériques matériels. Néanmoins, c'est l'ensemble des acteurs du secteur des TIC qui doit faire face à des demandes toujours plus importantes en termes de performances comme l'augmentation des débits, de la qualité de service, etc. En ce qui concerne le domaine des télécommunications, l'ensemble des acteurs est concerné par ce développement, que ce soit les acteurs agissant au niveau du réseau (déploiement, constitution des réseaux, opérateurs...) que les concepteurs des systèmes embarqués et des composants électroniques.

1. 1 téraoctet = 10^{12} octets

Parallèlement à cet essor, de nouveaux enjeux liés à l'environnement et à la consommation d'énergie sont apparus. En effet, les TIC seraient à l'origine d'environ 2% des émissions de gaz carbonique dans le monde et leur empreinte continuerait à croître au fil des années. Néanmoins, le rapport de [BBFH⁺08] souligne que si des innovations technologiques interviennent dans le domaine des TICS alors des réductions de l'ordre de 98% des émissions de gaz carbonique seraient envisageables dues notamment à l'utilisation massive de ces technologies dans de nombreux domaines tels que l'aéronautique, les transports, l'électronique grand public, etc.

Les systèmes de communications sans fil sont généralement très consommateurs d'énergie. Or, ces systèmes sont la plupart du temps embarqués et reposent souvent sur une source d'énergie limitée telle qu'une simple batterie. Atteindre et maintenir un haut niveau de performances implique traditionnellement une importante consommation d'énergie ainsi que de fortes dissipations de puissance, particulièrement au niveau des stations de bases, éléments de base du réseau de téléphonie mobile. Dans ce cas, des investissements coûteux en termes de refroidissement peuvent survenir. De plus, des problèmes de fiabilité des composants sont encourus sur le long terme.

A partir de ces constats, de nombreuses recherches, projets et consortiums [Gre14, OCF⁺13, EAR12] ont été créés dont les objectifs sont de réduire l'empreinte énergétique des systèmes de télécommunications tout en cherchant à améliorer les performances.

Au fil des avancées technologiques, les systèmes embarqués voient leurs dimensions se réduire quand parallèlement, la complexité des systèmes qu'ils doivent intégrer augmente. Ainsi, depuis quelques décennies, **les circuits logiques programmables de type FPGA**, *Field Programmable Gate Array*, sont devenus une technologie attractive dans le domaine des systèmes embarqués.

Historiquement, les circuits FPGA étaient utilisés comme accélérateurs matériels et pour effectuer du prototypage rapide d'ASIC, *Application Specific Integrated Circuit* à faible coût.

De nos jours, ils représentent une véritable alternative aux ASIC car ils permettent d'atteindre un bon niveau de performances dans un temps de développement très court. Ceci convient donc très bien au fait que le contexte actuel impose que le temps de mise sur le marché, *time-to-market*, se doit d'être toujours plus court.

Dans ce contexte à fortes contraintes, la conception de systèmes embarqués s'avère une tâche difficile pouvant être longue et fastidieuse et source de nombreuses erreurs tout au long du flot de conception. Afin de réduire de manière significative le temps de développement d'un circuit sur FPGA, il est possible de ré-utiliser des blocs matériels développés antérieurement, appelés *Intellectual Property* (IP). Un autre point fort des FPGAs est leur reconfiguration dynamique qui peut être totale ou bien partielle. Cela permet d'adapter l'utilisation des ressources pour répondre aux besoins de nombreuses applications en cours de fonctionnement. Malgré les nombreux bénéfices apportés, la spécificité de leur architecture ainsi que leur caractère reprogrammable font que les FPGA consomment beaucoup plus d'énergie que leurs principaux concurrents ASIC.

Début 2015, certains circuits FPGA possèdent plusieurs millions de cellules logiques, des blocs de calcul spécifiques, d'importantes capacités mémoires, etc. Ceci leur confère la capacité de mettre en œuvre toutes sortes d'applications, même complexes, telles que des circuits d'émission-réception de différentes technologies et standards comme le LTE ou le WiMAX (*Worldwide Interoperability for Microwave Access*).

Les FPGAs semblent donc une solution appropriée pour la conception rapide des futurs systèmes de télécommunications. Néanmoins, les aspects énergétiques sont devenus des enjeux majeurs et ces nouvelles contraintes de conception doivent être prises en compte au plus tôt dans le flot de conception afin d'éviter des coûts supplémentaires de re-développement. Afin d'envisager l'utilisation de tels circuits pour satisfaire la complexité croissante des systèmes de télécommunications, il semble primordial de permettre aux concepteurs de futurs systèmes, de pouvoir explorer l'espace de conception ainsi que d'estimer la consommation au plus tôt dans le flot de conception. De cette manière, il sera alors possible de dimensionner convenablement son système et ainsi d'effectuer des choix de conception déterminants.

Dans le cadre de la thèse, différents systèmes de communications numériques sans fil seront implantés sur FPGA et étudiés en termes de performances et consommation. Classiquement, les performances de systèmes de communications numériques sont évaluées à haut-niveau à l'aide d'un modèle fonctionnel du système, très flexible et permettant d'effectuer des simulations rapides. Évaluer la consommation énergétique d'un système à haut-niveau peut conduire à de grandes approximations et erreurs d'estimation à cause du manque d'informations sur la manière dont le système est implanté physiquement. Pour palier cela, le système peut être raffiné à travers différentes étapes du flot de conception. De cette manière, une estimation précise de la consommation peut être réalisée. Néanmoins, les temps de simulations et de conception deviennent très importants ainsi que les coûts de re-développement en cas de non-respect du budget énergétique ou des performances.

Organisation du manuscrit

Le manuscrit s'articule autour de deux thématiques majeures que sont les systèmes de communications numériques sans fil et leur implantation sur des circuits de type FPGA.

Dans un premier temps, le chapitre 1 illustre les principes généraux de tout système de communications numériques sans fil. Un focus sur les systèmes à base de modulation à porteuses multiples de type OFDM, *Orthogonal Frequency Division Multiplexing* sera effectué. Cette technique est très utilisée dans les systèmes de communications numériques actuels et permet d'atteindre d'important débit.

Une description de la technologie mobile de 4ème génération, le *Long-Term Evolution* sera effectuée. En effet, une étude de la couche physique sur la consommation relative aux différentes fonctions a été abordée dans la thèse.

Le chapitre 2 présente les caractéristiques inhérentes aux circuits logiques reconfigurables de type FPGA. Les origines de la consommation d'énergie dans les FPGAs seront présentées. D'autre part, les principales techniques de réduction qui sont classiquement mises en œuvre pour ce type de circuit seront décrites.

Le chapitre 3 trace l'état de l'art visant à répondre aux principales questions et problématiques soulevées lors des deux chapitres précédents. Premièrement, la consommation énergétique d'un système de communications numériques peut s'exprimer et évoluer en fonction de nombreux paramètres, technologiques ou systèmes. A partir de ce constat, comment évaluer rapidement un grand nombre de configurations possibles d'un système ? D'autre part, comment évaluer précisément et rapidement la consommation à haut-niveau ? Dans ce chapitre, les limites actuelles des méthodologies d'estimation de la consommation énergétique sur circuit FPGA seront présentées.

Dans le chapitre 4, une méthodologie d'estimation développée dans le cadre de la thèse est présentée. Celle-ci permet de répondre aux principales problématiques soulevées dans les chapitres précédents. Un point particulier de la méthodologie est de faciliter la comparaison entre différents systèmes, à haut-niveau, ainsi que l'exploration de l'espace de conception et des multiples configurations possibles. Un autre caractère innovant est la prise en compte des temps d'activité relatifs à chaque bloc-IP du système. De cette manière, les aspects dynamiques d'une application peuvent être considérés lors de l'estimation de la consommation.

Ensuite, le chapitre 5 propose l'évaluation de la méthodologie proposée à travers l'étude de plusieurs systèmes typiques de communications numériques. Tout d'abord, la méthodologie a été validée sur une chaîne à une antenne d'émission et de réception utilisant la modulation OFDM. Ensuite, des émetteurs LTE à une ou plusieurs antennes d'émission ont été évalués en termes d'efficacité énergétique et de consommation. Une chaîne intégrant le pré-codage de retournement temporel a aussi été évaluée. Pour finir, des travaux menés en commun au sein du laboratoire ont permis d'étudier l'efficacité énergétique de techniques coopératives dans le contexte LTE.

Le chapitre 6 propose des voies d'optimisation de la méthodologie proposée. Pour cela, des approches de modélisation de la puissance consommée par sous forme analytique en fonction des paramètres d'intérêt est présentée. De plus, une nouvelle approche de modélisation de bloc matériel à base de réseaux de neurones est proposée.

Pour finir, nous concluons sur l'ensemble des travaux effectués durant cette thèse et les perspectives de travaux futurs sont abordées.

Liste des contributions scientifiques

Conférences Internationales avec comité de lecture :

- J. Lorandel, J-C. Prévotet and M. Hélar, *Fast Power and Energy Efficiency Analysis of FPGA-based Wireless Base-band Processing*, in International Workshop on High Performance Energy Efficient Embedded Systems (HIP3ES 2016 4th Edition), Prague, Czech Republic, 18 Jan. 2016.

- J. Lorandel, J-C. Prévotet and M. H elard, *Dynamic Power Evaluation of LTE Wireless Baseband Processing on FPGA*, Conference on Design & Architectures for Signal & Image Processing (DASIP 2015), Cracow, Poland, 23-25 Sept. 2015.
- J. Lorandel, J-C. Prévotet and M. H elard, *Dynamic Power Estimation of FPGA-based Wireless Communication Systems*, The International Conference on Wireless Technologies, embedded and intelligent Systems WITS-2015, Fes, Marocco, 29-30 Apr. 2015.

Revue et lettres internationales :

- M. Maaz, J. Lorandel, P. Mary, J-C. Prévotet and M. H elard, *Energy Efficiency Analysis of Hybrid-ARQ Relay-Assisted Schemes in LTE Based Systems*, EURASIP Journal on Wireless Communications and Networking, Vol. 1, 2016 (soumise le 1 octobre 2015, accept ee le 7 janvier 2016, publi ee le 19 janvier 2016)
- J. Lorandel, J-C. Prévotet and M. H elard, *Fast Power and Performance Evaluation of FPGA-based Wireless Communication Systems*, IEEE Access (soumise 5 avril 2016, accept ee le 20 avril 2016, publi ee le 27 avril 2016))

Conf erences nationales, colloques nationaux,  coles th ematiques :

- Jordane Lorandel, *Dynamic Power Estimation Methodology for FPGA-based Wireless Communication Systems*, Pr esentation orale   la Journ ee th ematique de l'Eco-Radio du GDR ISIS,   T el ecom. ParisTech, le 6 Mai 2015.
- Jordane Lorandel, *Dynamic Power Estimation Methodology for FPGA-based Wireless Communication Systems*, Pr esentation poster au colloque GDR SOC-SIP 2014,   T el ecom. ParisTech, 11-13 juin 2014.
- Ecole th ematique COnception FAible Consommation pour les syst emes embarqu es et temps r eels, Ecofac 2014, 19-23 mai 2014   Lorient.

Chapitre 1

Les systèmes de communications numériques sans fil

1.1 Introduction

De nos jours, les équipements numériques sont devenus incontournables et omniprésents. Ces équipements tels que les téléphones mobiles, ordinateurs, sont capables de transmettre et d'échanger des données via les réseaux sans fil.

Préalablement à la description des systèmes actuels, il semble important de rappeler comment fonctionne un système de communications numériques sans fil et quelles sont leurs principales caractéristiques. De plus, des notions relatives au support physique de propagation de l'onde électromagnétique, appelé canal de transmission seront rappelées.

Par la suite, une présentation de la modulation numérique multi-porteuses OFDM, très utilisée dans les réseaux mobiles actuels tels que le LTE, sera effectuée. Finalement, les enjeux et les problématiques énergétiques dans les systèmes de communications numériques sans fil seront abordés.

1.2 Caractéristiques d'un système de communications numériques

Le schéma simplifié des traitements en bande de base d'une chaîne de télécommunications est rappelé en Figure 1.1. Il représente les différents éléments nécessaires pour effectuer une communication d'une source à un destinataire. La plupart des systèmes, aussi perfectionnés soient-ils, dérivent de ce schéma de base.

La source binaire délivre les bits d'information à transmettre. Ceux-ci sont tout d'abord encodés de manière à rendre l'information à transmettre plus robuste durant la communication. En effet, une information redondante est ajoutée afin de protéger les données face aux effets du canal, qui seront abordés dans les paragraphes suivants.

Ensuite, l'information encodée ne peut pas encore être transmise à travers le canal. Pour

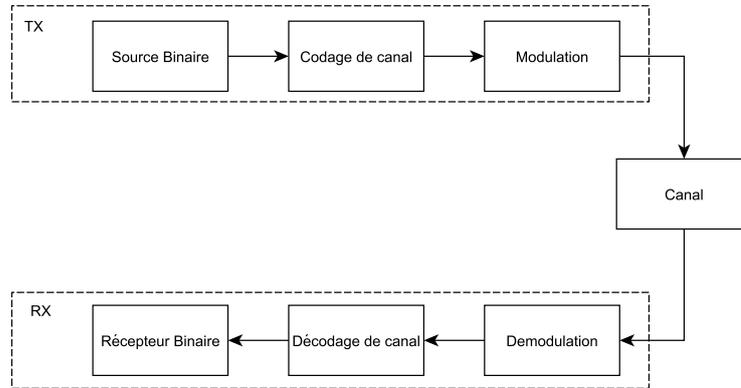


FIGURE 1.1 – Schéma de base des traitements en bande de base d'une chaîne de télécommunications

cela, l'information binaire est transformée en signal à travers une étape de modulation, qui correspond à un codage binaire/M-aire et à un codage/mapping M-aire à signal (aussi appelé mise en forme). Si l'on considère que pour chaque ensemble de m_k bits du message est associé un symbole a_k alors le signal continu issu du modulateur, noté $s(t)$, peut s'écrire sous la forme suivante :

$$s(t) = \sum_k a_k h(t - kTs) \quad (1.1)$$

avec a_k pouvant prendre $M = 2^m$ valeurs, $h(t)$ est la forme d'onde de durée Ts (durée d'un symbole). Au niveau du récepteur, le signal reçu est traité de manière duale à travers les étapes de démodulation et de décodage de canal.

Comme énoncé précédemment, tout système de télécommunication requiert l'introduction de technique de codage de canal afin de protéger l'information à transmettre. Dans les paragraphes suivants, nous expliquerons les effets introduits sur le signal suite à sa propagation à travers le canal.

1.2.1 Le canal et ses effets

Transmettre un signal à travers un canal engendre des modifications sur celui-ci. Pour mieux comprendre, on peut représenter un canal par sa fonction de transfert, qui affecte une ou plusieurs de ses entrées et les restitue à sa sortie.

$$Y = H.X \quad (1.2)$$

avec Y le signal reçu, X le signal émis et H la fonction de transfert du canal.

Pour caractériser un canal, un moyen généralement retenu est de déterminer sa réponse impulsionnelle (RI) en émettant une impulsion de l'émetteur vers le récepteur. Ainsi, on peut estimer les perturbations (atténuations, amplifications) que le signal va subir au cours de sa transmission. Le schéma de principe est illustré en Figure 1.2.

A l'issue de l'évaluation de la RI du canal, plusieurs paramètres statistiques sont couramment utilisés pour le caractériser :

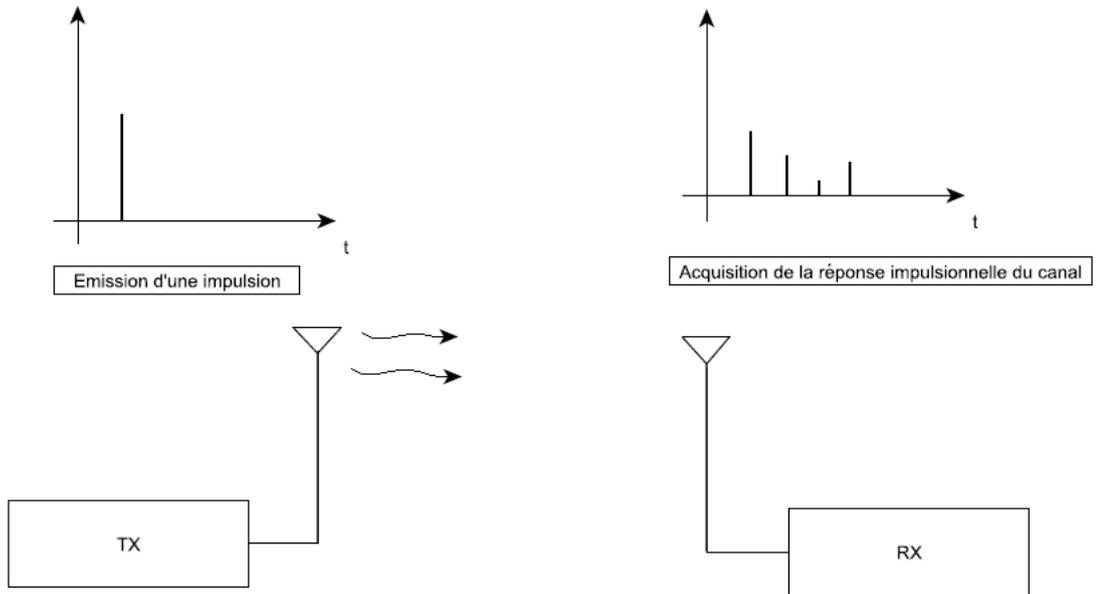


FIGURE 1.2 – Détermination de la réponse impulsionnelle d'un canal

- l'étalement maximal des retards τ_{max} ,
- la dispersion des retards σ_T^2 ,
- la bande de cohérence du canal B_c ,
- le temps de cohérence du canal T_c et,
- la fréquence Doppler maximale Fd_{max} .

L'étalement maximal des retards, τ_{max} , représente la durée entre le premier et le dernier trajet significatif de la réponse impulsionnelle mesurée en réception. Ainsi, lors de l'émission d'une impulsion, plusieurs trajets (d'amplitude et de phase différentes) peuvent arriver à différents instants au niveau du récepteur en fonction de l'environnement de propagation. Ces trajets peuvent se recombinaer au niveau du récepteur de manière constructive (amplification du signal) ou bien destructive. On parle alors d'évanouissements.

La dispersion des retards, σ_T^2 , est une valeur statistique correspondant à l'écart-type de la densité de probabilité moyenne des retards.

Concernant **la bande de cohérence du canal**, B_c , elle est définie comme étant la bande de fréquence pendant laquelle le canal est considéré comme non sélectif c-à-d que la fonction de transfert du canal est constante. Une définition de la bande de cohérence est donnée par l'équation 1.3 :

$$B_c \propto \frac{1}{\sigma_T} \quad (1.3)$$

Par analogie au domaine temporel, le temps de cohérence est également défini. **Le temps de cohérence du canal** correspond à la durée pendant laquelle le canal peut être considéré comme invariant. Ce paramètre est un indicateur dans le cas où il y a mobilité

dans le canal, lorsqu'un des points de communication est en mouvement.

La fréquence Doppler maximale, notée $F_{d_{max}}$, est engendrée par exemple lors de la transmission d'un signal vers un récepteur en mouvement. En effet, le récepteur parcourt une distance au cours du temps entre l'émission et la réception de l'onde. Ceci engendre des variations de phase du signal et donc un décalage de fréquence, appelée fréquence Doppler. Elle est calculée par l'équation suivante 1.4 :

$$F_{d_{max}} = \frac{v_r}{\lambda} \quad (1.4)$$

avec v_r , la vitesse du récepteur et λ la longueur d'onde de l'onde transmise.

La caractérisation d'un canal permet de définir sa sélectivité. Ainsi, un canal est dit sélectif en fréquence lorsque les perturbations apportées sur les fréquences du signal ne sont pas les mêmes selon les fréquences. Ainsi, un canal peut être propice aux trajets multiples et engendrer des évanouissements fréquentiels sur le signal. En effet, il est possible dans ce type d'environnement qu'un écho de puissance variable arrive au même instant que le signal attendu au niveau du récepteur. Réciproquement, ce type de canal est qualifié de dispersif en temps.

Un canal est dit non sélectif en fréquence si la bande de fréquence d'un signal B_s est très inférieure à la bande de cohérence B_c du canal. Par analogie dans le domaine temporel, si un signal possède une durée T_s très supérieure à l'étalement des retards, alors le canal est dit non sélectif. Finalement, on résume le cas d'un canal sélectif en fréquence ainsi :

$$\begin{aligned} T_s &\leq \sigma_T \\ \text{ou } B_s &\geq B_c \end{aligned} \quad (1.5)$$

Pour qu'un canal soit non-sélectif en fréquence :

$$\begin{aligned} T_s &\gg \sigma_T \\ \text{ou } B_s &\ll B_c \end{aligned} \quad (1.6)$$

Pour transmettre une information à travers le canal, une étape de modulation est nécessaire. Cette étape permet d'adapter l'information pour la transmission sur le canal. L'information à transmettre module généralement un signal sinusoïdal qualifié de porteuse. Le signal s'écrit alors de la manière suivante :

$$S(t) = A \cos(2\pi \cdot f_0 \cdot t + \phi_0) \quad (1.7)$$

avec A , l'amplitude, f_0 la fréquence et ϕ_0 la phase.

Ces différents paramètres peuvent alors être modifiés de manière à réaliser une modulation de fréquence, de phase et d'amplitude du signal sinusoïdal.

Il a pu être décrit précédemment que l'information binaire est transformée en signal à l'aide de l'étape de modulation, qui correspond à un codage binaire/M-aire et codage/mapping M-aire à signal. Le codage binaire/M-aire permet d'associer à un paquet de m bits, un état de modulation (aussi appelé symbole) parmi les $M = 2^m$ états possibles.

Plusieurs types de modulation peuvent être utilisés :

- Modulation de phase (*Phase Shift Keying (PSK)*),
- Modulation d'amplitude de porteuses en quadrature (*Quadrature Amplitude Modulation (QAM)*),
- Modulation de fréquence (*Frequency Shift Keying (FSK)*),
- ...

Au fil des avancées technologiques et des recherches, des procédés de codage des signaux numériques ont été développés de manière à pouvoir répondre aux contraintes de chaque application. Certaines, comme l'OFDM, permettent de répartir l'information à transmettre sur un grand nombre de sous-porteuses et sont présentes dans les systèmes de communications numériques actuels.

1.3 La modulation OFDM

L'OFDM, *Orthogonal Frequency Division Multiplexing*, est aujourd'hui utilisée dans différentes technologies radios et dans certains standards tels qu'en LTE et en WiMAX (IEEE 802.16e [IEE12]). Elle est aussi présente dans de nombreux domaines d'application tels que la Télévision Numérique Terrestre (TNT) et l'ADSL (*Asynchronous Digital Subscriber Line*).

1.3.1 Caractéristiques de l'OFDM

Afin d'illustrer les principales caractéristiques de la modulation OFDM, celle-ci sera comparée à une modulation classique mono-porteuse. L'OFDM a un fort intérêt en présence de canaux sélectifs en fréquence présentant des trajets multiples. Pour palier cette sélectivité, l'information à transmettre est répartie sur un grand nombre de sous-porteuses. Ainsi, pour une bande de fréquences notée W , N sous-bandes de largeur Δ_f Hz sont créées. Chaque sous-canal ainsi réalisé possède une bande de fréquence très étroite, de taille inférieure à la bande de cohérence, ayant une réponse fréquentielle pouvant être considérée comme non sélective en fréquence.

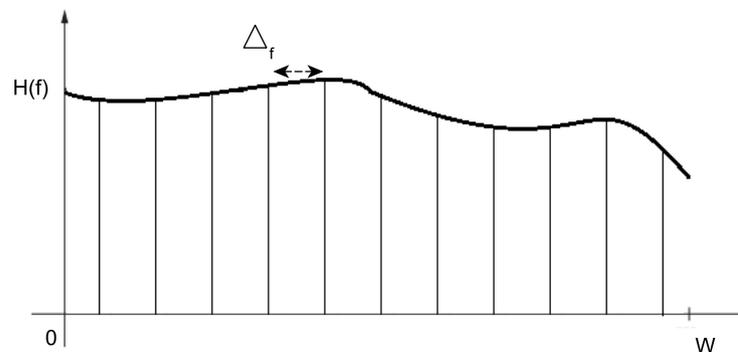


FIGURE 1.3 – Division d'une bande W en N sous-bandes

Une occupation maximale de la bande est possible grâce à une autre spécificité de l'OFDM qui est l'orthogonalité entre chaque sous-porteuses. Les sous-porteuses sont séparées par un espace de fréquence valant $\Delta_f = \frac{1}{T_s}$, correspondant à l'espace inter-porteuse, avec T_s représentant la durée du fenêtrage appliqué durant la modulation. Dans le domaine fréquentiel, ce type de fenêtrage engendre un spectre en sinus cardinal, représenté en Figure 1.4, s'annulant tous les $k\frac{1}{T_s}$. La transformée de Fourier d'un sinus cardinal est une fonction porte. Ainsi, aucune des sous-porteuses n'interfère avec une autre. On évite donc le phénomène d'interférence inter-porteuse (ICI *Inter-Carrier Interferences*) permettant une occupation maximale de la bande.

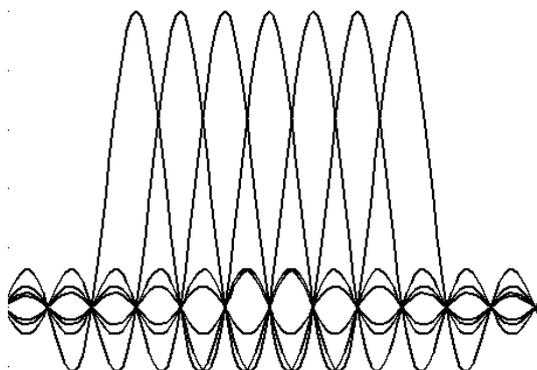


FIGURE 1.4 – Représentation fréquentielle d'un signal OFDM

La Figure 1.5 permet de mettre en évidence les caractéristiques à la fois dans les domaines temporel et fréquentiel de l'OFDM. On suppose que les modulations OFDM et mono-porteuse utilise une même bande de fréquence notée W . De plus, on considère un roll-off nul ainsi qu'un intervalle de garde nul. Si T_s est le temps symbole en mono-porteuse et que N sous-porteuses sont utilisées en OFDM alors le temps d'un symbole OFDM sera $T_{sOFDM} = T_s.N$.

Dans le cas d'une modulation mono-porteuse, si le débit symbole est noté D_s alors on a $D_s = W = \frac{1}{T_s}$. Dans le cas de la modulation OFDM, $D_{sOFDM} = N \cdot \frac{1}{T_{sOFDM}} = N \cdot \frac{1}{N.T_s}$. On remarque donc que pour une même bande donnée W , les deux modulations possèdent au final le même débit symbole. Néanmoins, sachant que N sous-porteuses sont utilisées en OFDM, le débit symbole de chaque sous-porteuse est de D_s/N . Cette réduction de débit sur chaque sous-porteuse rend cette modulation plus robuste aux interférences.

Un schéma d'un modulateur OFDM est présenté en Figure 1.6. En amont du modulateur OFDM, les données binaires à transmettre sont préalablement modulées en phase et/ou en amplitude par un modulateur QAM. Ainsi, ce dernier associe à chaque mot de m bit(s), un symbole complexe noté c_k . Ce symbole complexe correspond aux coordonnées d'un point de la constellation de modulation visible à la Figure 1.7. On désigne par m , le nombre de bits par symbole complexe modulé. A titre d'exemple, 1 bit est nécessaire pour coder 2 états en modulation BPSK, 2 bits sont requis pour coder 4 (2^2) états pour la modulation QPSK, 4 bits en 16QAM, 6 bits en 64QAM etc. A noter que l'ordre de

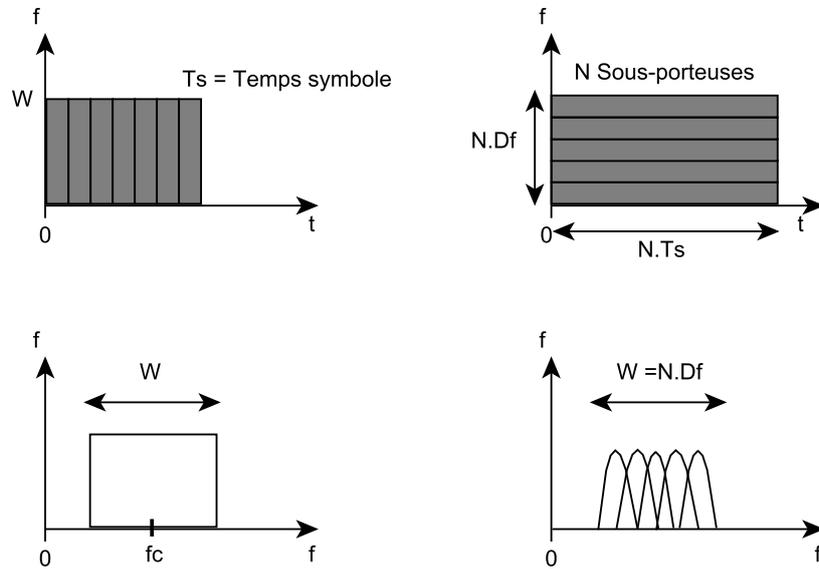


FIGURE 1.5 – Comparaisons entre une modulation mono-porteuse (gauche) et l’OFDM (droite)

modulation noté M est égale à 2^m . L’augmentation du nombre m de bits par symbole complexe modulé permet d’augmenter le débit binaire tout en conservant un même débit symbole. A noter que l’on parle aussi d’étape de ‘mapping’ des données binaires sur la constellation. Dans le cas d’un système à multiples antennes, une étape dite de pré-codage des symboles complexes issus du modulateur QAM est nécessaire avant l’envoi des données aux entrées des modulateurs OFDM (cf. section 1.4.2).

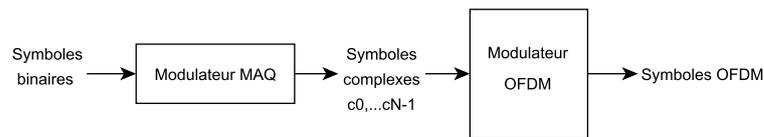


FIGURE 1.6 – Schématique d’un modulateur OFDM précédé d’un modulateur de phase et d’amplitude (MAQ)

On note c_k , les symboles complexes issus du codage binaire/ M -aire et du mapping de la constellation et T_{sOFDM} la durée utile d’un symbole OFDM.

Le modulateur OFDM va associer à chaque donnée c_k un signal de fréquence f_k . Il en résulte que le signal s’écrit de la manière suivante :

$$c_k e^{(2j\pi k t / T_{sOFDM})} \quad (1.8)$$

A l’issue de cette étape, les sous-porteuses sont groupées et sommées afin de former un symbole OFDM. En sortie du modulateur OFDM, le signal final $s(t)$ correspondant à un symbole OFDM s’écrit :

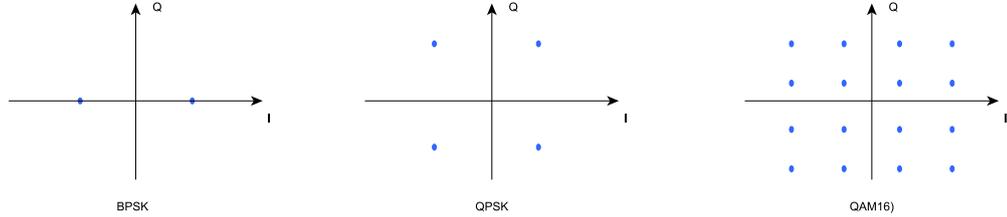


FIGURE 1.7 – Constellations en BPSK, QPSK et QAM16

$$s(t) = \sum_{k=0}^{N-1} c_k e^{(2j\pi kt/T_{sOFDM})} \quad (1.9)$$

avec k , l'indice des sous-porteuses, N la taille du symbole OFDM.

Dans des environnements propices aux trajets multiples, il est possible d'obtenir des perturbations entre deux symboles consécutifs, comme illustré à la Figure 1.8. Ces perturbations proviennent de répliques du signal retardé et engendrent un phénomène appelé interférences entre symboles (ISI, *Inter-Symbol Interferences*). Pour palier cela, une technique consiste à ajouter un intervalle de garde entre chaque symbole. Son dimensionnement est déterminant afin qu'il puisse permettre l'absorption de ces trajets avant le symbole utile. Ainsi, il faut que sa taille soit supérieure à l'étalement maximum des retards τ_{max} . L'intervalle de garde est appelé préfixe cyclique lorsque la fin d'un symbole OFDM est recopiée à son début. Ces deux techniques augmentent la robustesse aux trajets multiples et engendrent néanmoins une perte de débit utile.

En réception, le fait d'avoir ajouté le préfixe cyclique permet d'extraire le signal utile du signal reçu de manière relativement simple. En effet, une égalisation à un coefficient par sous-porteuses est alors possible, réduisant ainsi la complexité du récepteur dans le domaine fréquentiel.

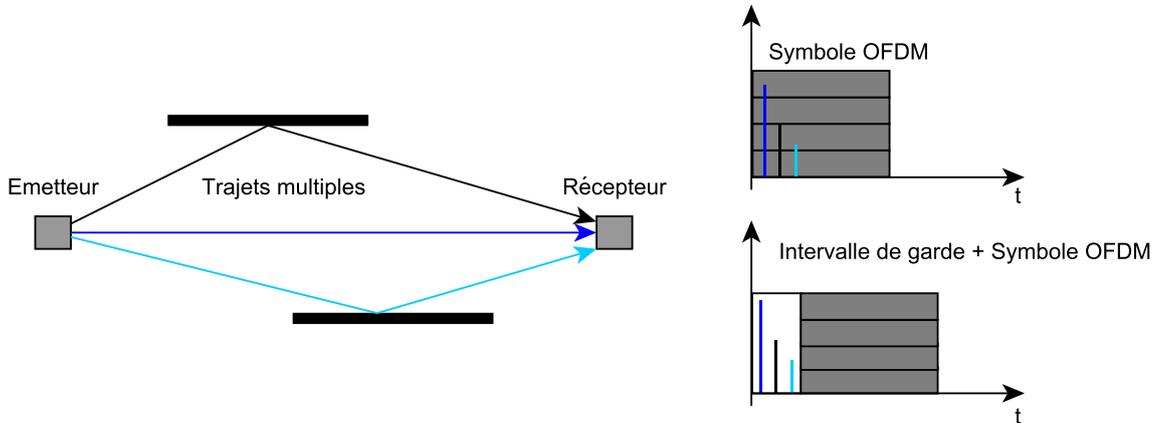


FIGURE 1.8 – Illustration du principe de l'intervalle de garde face aux interférences entre symboles OFDM

1.3.2 Réalisation matérielle

Plusieurs recherches ont montré que le modulateur OFDM pouvait être réalisé à partir d'une Transformée de Fourier Inverse Discrète (IDFT). D'un point de vue de la réalisation matérielle, c'est une IFFT, transformée de Fourier rapide inverse qui est privilégiée car relativement peu complexe. Ceci a eu pour conséquence de grandement favoriser l'utilisation de l'OFDM dans de nombreux systèmes de communications numériques. Sur la Figure 1.9 est présenté un schéma du modulateur OFDM à base d'IFFT, incluant l'étape d'insertion du préfixe cyclique. En réception, le démodulateur OFDM est réalisé à partir d'une FFT dont le schéma est visible en Figure 1.10.

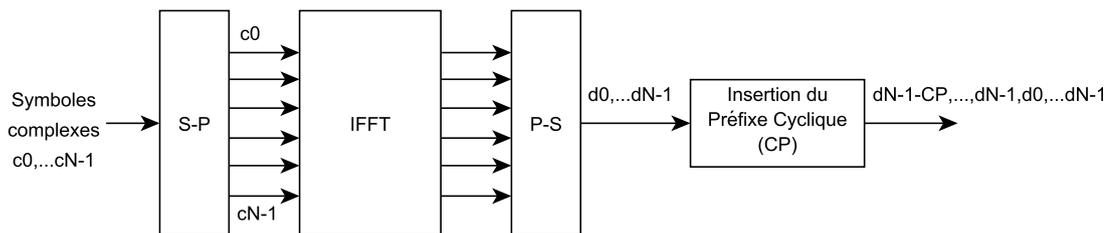


FIGURE 1.9 – Schéma d'un modulateur OFDM à base d'IFFT

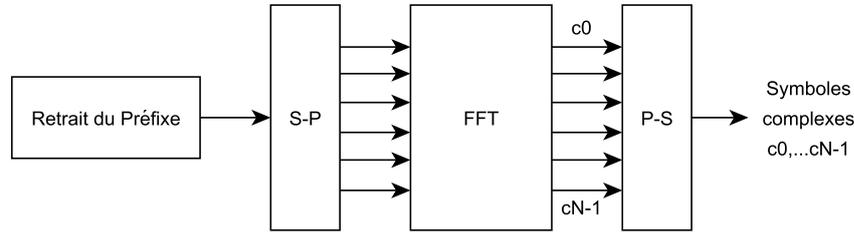


FIGURE 1.10 – Schéma d'un démodulateur OFDM à base de FFT

1.3.3 Avantages et limitations de l'OFDM

La modulation OFDM est très plébiscitée car elle offre un grand nombre d'avantages :

- une occupation optimale de la bande,
- la robustesse face aux interférences,
- de bonnes performances pour les canaux sélectifs en fréquence,
- une réalisation matérielle relativement peu complexe.

Néanmoins, l'OFDM présente quelques désavantages. En effet, cette modulation est très sensible aux erreurs de synchronisation. Un autre problème rencontré en OFDM, et qui est sujet à de nombreuses recherches est le *Peak to Average Power Ratio* (PAPR). Cette métrique traduit les variations d'amplitude-crête du signal au cours du temps. Ces fortes disparités d'amplitude du signal au cours du temps sont problématiques au niveau de l'amplificateur de puissance, élément non-linéaire, qui se trouve en aval du modulateur OFDM. Ceci impacte le rendement de l'amplificateur et donc sa consommation et provoque d'importantes distorsions.

1.4 OFDM dans les réseaux mobiles

Les réseaux mobiles n'ont cessé d'évoluer depuis les dernières décennies et cette évolution est généralement représentée par un changement de 'génération' qui s'est traduit par des choix technologiques importants comme par exemple l'intégration de l'OFDM pour le LTE, *Long Term Evolution*.

Dans le chapitre 1 de [YB12], dont les paragraphes suivants sont inspirés, l'évolution des réseaux mobiles et des technologies radios sont décrites de la première génération (1G) jusqu'aux dernières générations avec un focus sur les avancées qui ont permis le développement du LTE. En partant de leur analyse, le schéma en figure 1.11 permet de résumer de manière concise et schématique cette évolution.

Historiquement, la première génération de réseau mobile était majoritairement basée sur des systèmes à codages et modulations analogiques qui devaient permettre d'assurer un service en mobilité. Les ressources radios étaient divisées entre les utilisateurs à l'aide d'une technique d'accès multiple telle que le FDMA, *Frequency Division Multiple Access*. La deuxième génération (2G) a été principalement marquée par le développement des modulations numériques et du GSM, *Global System for Mobile communication*, dont les

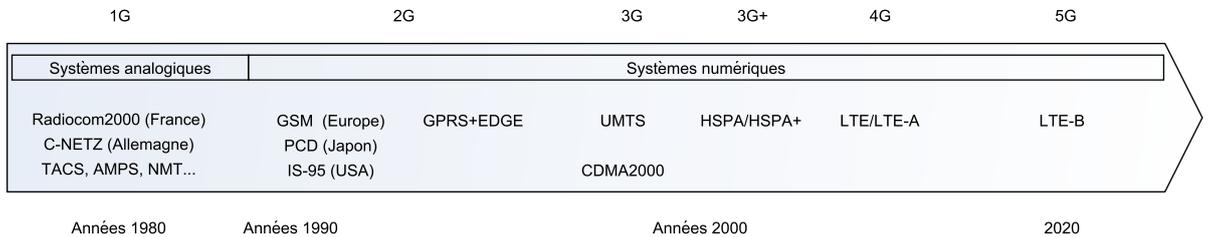


FIGURE 1.11 – Évolution des technologies des réseaux mobiles [YB12]

réseaux couvrent aujourd’hui plus de 90% de la population mondiale. La troisième génération (3G) a vu deux succès notables que sont l’UMTS et le CDMA2000. C’est à partir d’une série d’évolution de l’UMTS, HSPA et HSPA+, que sera envisagé le LTE en 2004 [YB12]. Cette technologie mobile sera standardisée par l’organisme 3GPP et aura comme principaux objectifs de répondre aux enjeux de l’amélioration des débits, de l’efficacité spectrale tout en réduisant les latences. De plus, elle doit aussi permettre un déploiement plus facile du réseau pour répondre rapidement à la croissance des besoins de transmission de données. A noter que les réseaux de cinquième génération (5G) devraient voir le jour à l’horizon 2020. Les principaux objectifs, d’après un des leaders actuels de la technologie 5G [HUA13] sont notamment :

- réduire la latence (1ms),
- réduire la consommation d’énergie,
- fournir des débits de pointe de plus de 10 Gbit/s,
- augmenter la capacité des réseaux.

Dans le cadre de la thèse, des études ont été menées sur la couche physique du LTE afin notamment d’en évaluer l’efficacité énergétique. Aussi, le réseau mobile LTE et ses propriétés sont présentés dans la section suivante.

1.4.1 Le réseau mobile LTE et ses propriétés

Le réseau universel évolué d’accès radio terrestre, eUTRAN, *Evolved Universal Terrestrial Radio Access Network*, dont un schéma de l’architecture est visible en Figure 1.12, est constitué d’une part de l’équipement utilisateur (UE) et d’autre part de l’infrastructure réseau contenant les équipements propres à l’opérateur. L’équipement utilisateur est relié à l’infrastructure réseau par le biais de son interface radio. L’équipement faisant office d’interlocuteur, côté réseau est appelé station de base, *Base Station* ou *eNodeB*.

Il existe deux voies de communication entre l’UE et la station de base. La voie dite ‘montante’, *UpLink*, entre l’UE et la station de base et la voie descendante, *DownLink*, entre la station de base et l’UE. Pour emprunter ces deux voies de communication, il existe plusieurs modes dont le FDD, *Frequency Division Duplex* et le TDD, *Time Division Duplex*, sont utilisés en LTE. Le mode FDD consiste à séparer les communications montantes et descendantes en fréquence. Le mode TDD, quant à lui, permet aux équipements d’utiliser les mêmes fréquences mais à des instants différents [YB12]. En LTE, c’est l’OFDMA, *Orthogonal Frequency Division Multiple Access* qui est utilisée en voie descendante. En

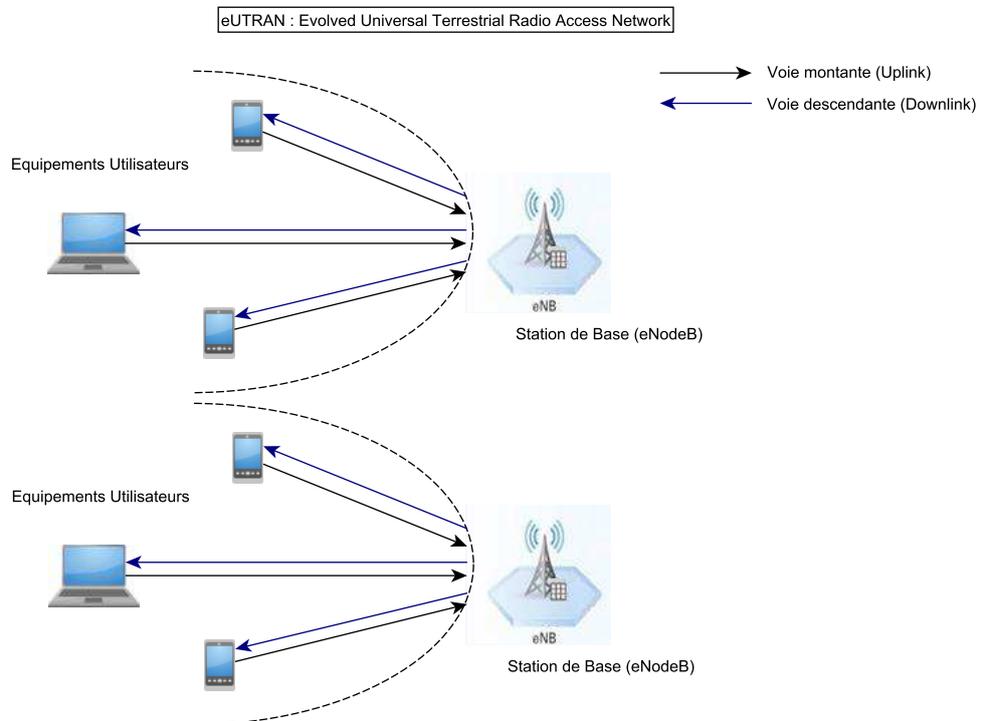


FIGURE 1.12 – Schéma représentatif du réseau eUTRAN [YB12]

voie montante, c'est le SC-FDMA, *Single Carrier Frequency Division Multiple Access*.

La norme LTE, standardisée par le 3GPP, introduit un certain nombre de spécificités :

- concept de canaux et de signaux physiques,
- deux types possibles de structure de trame : Type 1 ou Type 2,
- allocation de ressources aux utilisateurs sous forme de blocs de ressources.

Selon la norme [ETS09], la trame porte des données qui sont issues **de canaux physiques et de signaux physiques**. Le concept de canal permet notamment d'identifier les types de données transportées sur l'interface radio, provenant des couches supérieures, ainsi que les caractéristiques de qualité de service associées et les paramètres liés à la transmission [YB12]. Il existe trois types de canaux : logiques, de transport et physiques.

Au niveau de la couche physique, les canaux sont :

- PDSCH *Physical Downlink Shared CHannel*,
- PBCH *Physical Broadcast CHannel*,
- PCFICH *Physical Control Format Indicator CHannel*,
- PDCCH *Physical Downlink Control CHannel*
- PMCH *Physical Multicast CHannel*
- PHICH *Physical Hybrid ARQ Indicator CHannel*

Le canal PDSCH fait référence au canal dédié aux données des utilisateurs. Le canal PBCH porte des informations d'identification de cellule et de contrôle d'accès. Le canal

PCFICH contient des données représentatives du nombre de ressources allouées pour la transmission du PDCCH qui porte des informations sur l'allocation de ressources à l'équipement utilisateur (DCI, *Downlink Control Information*). Les canaux PMCH et PHICH portent respectivement des informations pour le mode de MBMS *Multimedia Broadcast and Multicast Services* et des informations d'acquittements et des identifiants nécessaires à la mise en œuvre du protocole de retransmission HARQ. Différents traitements sont effectués sur les données en fonction du type de canal, comme indiqué en Figure 1.13.

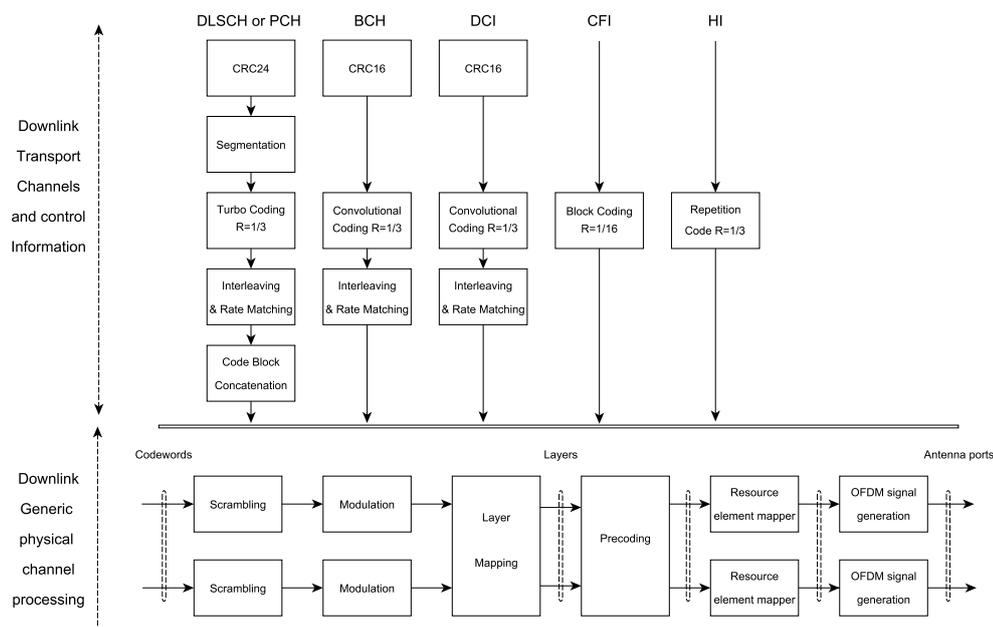


FIGURE 1.13 – Traitements à l'émission associés aux différents canaux en voie descendante

Les signaux physiques correspondent aussi à des éléments de ressources qui sont utilisés en support des canaux physiques mais qui prennent leur origine dans la couche physique. Il existe deux types de signaux physiques :

- les signaux de références *Reference Signals (RS)*,
- et les signaux de synchronisation primaires et secondaires *Synchronisation Signals (PSS, SSS)*.

Les signaux physiques de référence sont connus du récepteur et sont classés en trois types : les RS spécifiques à la cellule (CRS), ceux spécifiques à l'équipement utilisateur et les RS de type MBSFN utilisés dans le cas d'une transmission de type diffusion, *broadcast*. L'ensemble de ces signaux de référence permettent notamment la démodulation de certains canaux (PDSCH, PBCH, PDCCH, PMCH), d'estimer le canal, d'effectuer la formation de faisceaux, *beamforming*, etc. Ils sont conçus à partir de symboles QPSK générés grâce à une séquence pseudo-aléatoire (SPA) de Gold.

Comme indiqué par la norme [ETS09], les signaux de référence sont disposés précisément en temps et en fréquence en fonction du nombre de ports d'antenne utilisés. Les

signaux de synchronisation (PSS et SSS) sont utilisés dans le cadre de la recherche de cellule et de synchronisations temporelles et fréquentielles.

La structure de la trame LTE peut être décrite suivant deux types : type 1 ou 2. Le type utilisé va dépendre du mode de communication privilégié. Une structure de trame de type 1, illustrée en Figure 1.14, est adaptée au mode de communication FDD en full-duplex et half-duplex. Ainsi, 10 sous-trames de 1ms sont utilisées pour les transmissions en voies descendante et montante [ETS09]. Concernant la structure de type 2, elle est adaptée au TDD. Certaines sous-trames spécifiques sont alors réservées à la voie descendante et d'autres à la voie montante¹.

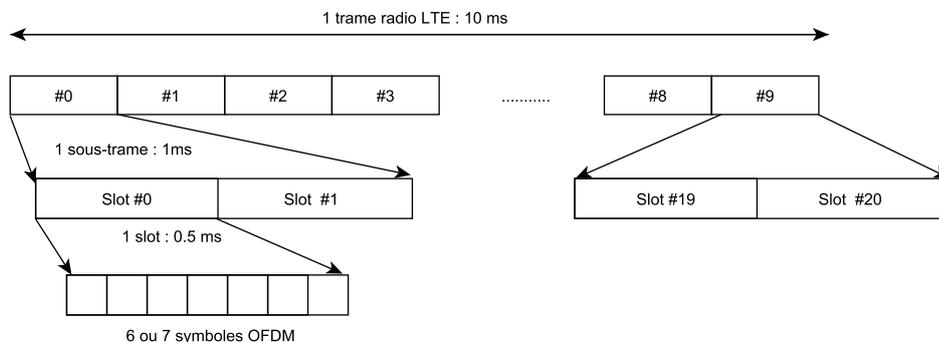


FIGURE 1.14 – Structure de la trame LTE de Type 1 (FDD) [ETS09]

D'après la Figure 1.14, la trame radio en LTE possède une durée totale de 10ms. Elle est divisée en 10 sous-trames de 1ms chacune qui sont elles-mêmes chacune constituées de deux *slots*. Chaque slot peut contenir 6 ou 7 symboles OFDM en fonction de la durée du préfixe cyclique (normale ou étendue).

L'allocation des ressources radios aux utilisateurs en LTE est généralement représentée sous la forme d'une grille temps-fréquence, comme indiquée en Figure 1.15. La plus petite unité est appelée élément de ressource, *resource element* (RE), et correspond typiquement à une sous-porteuse allouée pendant un symbole OFDM. La plus petite ressource radio allouable par un utilisateur est appelée un bloc de ressources, *Resource Block* (RB), correspondant à un bloc de 12 sous-porteuses contigües sur la durée d'un slot, c-à-d 6 ou 7 symboles OFDM.

1. La structure de type 1 FDD est majoritairement utilisée en Europe pour l'UMTS [YB12]; c'est pourquoi nous détaillons principalement la structure de la trame dans cette configuration.

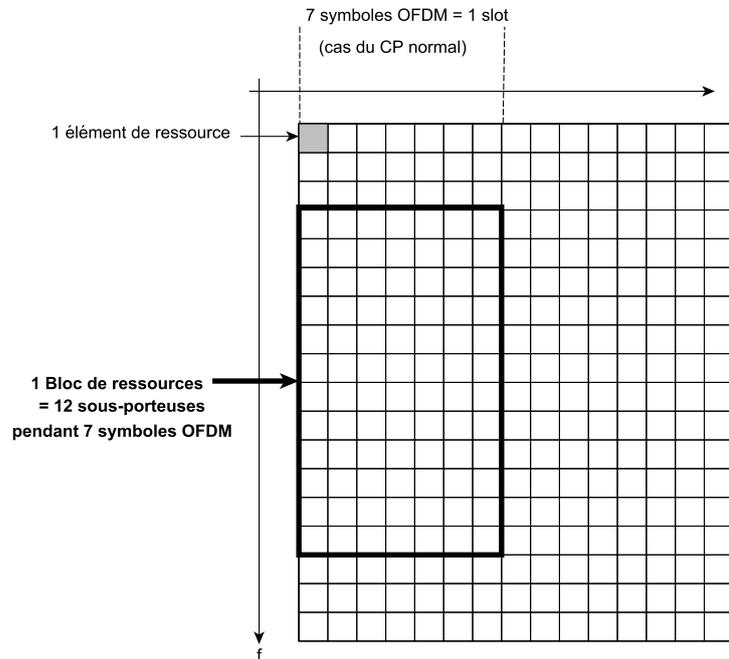


FIGURE 1.15 – Allocation de ressources en LTE (cas du CP normal)

Dans le cadre de la thèse, nous étudierons plus précisément les traitements en bande de base de la couche physique du LTE en liaison descendante (*Downlink*) portant les données utilisateurs (PDSCH). Les principaux paramètres OFDM sont récapitulés dans le Tableau 1.1.

TABLE 1.1 – Paramètres OFDM de la couche physique LTE en voie descendante

Bande spectrale occupée (MHz)	1.4	3	5	10	15	20
Durée de trame (ms)	10					
Nombre de symboles OFDM par slot	7/6 (CP normal/étendu)					
Espace entre sous-porteuses (kHz)	15					
Fréquence d'échantillonnage (MHz)	1.92	3.84	7.68	15.36	23.04	30.72
Taille de (I)FFT	128	256	512	1024	1536	2048
Nombre de sous-porteuses utilisées	72	180	300	600	900	1200
Nombre de blocs de ressources (RB)	6	12	25	50	75	100
Longueur du Préfixe Cyclique (CP court) (μs)	5.2 (1er Symbole) / 4.69 (symboles suivants)					
Longueur du Préfixe Cyclique (CP long) (μs)	16.67					

Comme décrit dans le tableau 1.1, on peut noter que le LTE supporte l'utilisation de bande de fréquence de différentes tailles allant de 1.4MHz à 20MHz. Chaque bande de fréquence est associée à une taille d'IFFT et donc une taille de symbole OFDM. De plus, on peut aussi constater que l'ensemble des sous-porteuses ne sont pas utilisées. Ainsi pour une bande de 1.4MHz, seules les 72 sous-porteuses centrales sont utilisées sur les 128 disponibles. A noter aussi que l'espace inter-porteuse est constant quelle que soit la taille de bande et fixé à 15kHz.

Dans le LTE comme dans d'autres standard, différentes configurations c-à-d un nombre d'antennes d'émission et de réception différent de 1, du système sont possibles.

1.4.2 Configuration de systèmes à multiples antennes

Augmenter le nombre d'antennes d'émission et/ou de réception d'un système permet d'obtenir certaines propriétés. On distingue d'abord les différentes configurations possibles comme indiqué par la Figure 1.16 :

- MISO : plusieurs antennes d'émission, 1 antenne de réception,
- SIMO : 1 antenne d'émission, plusieurs antennes de réception,
- MIMO : plusieurs antennes d'émission, plusieurs antennes de réception .

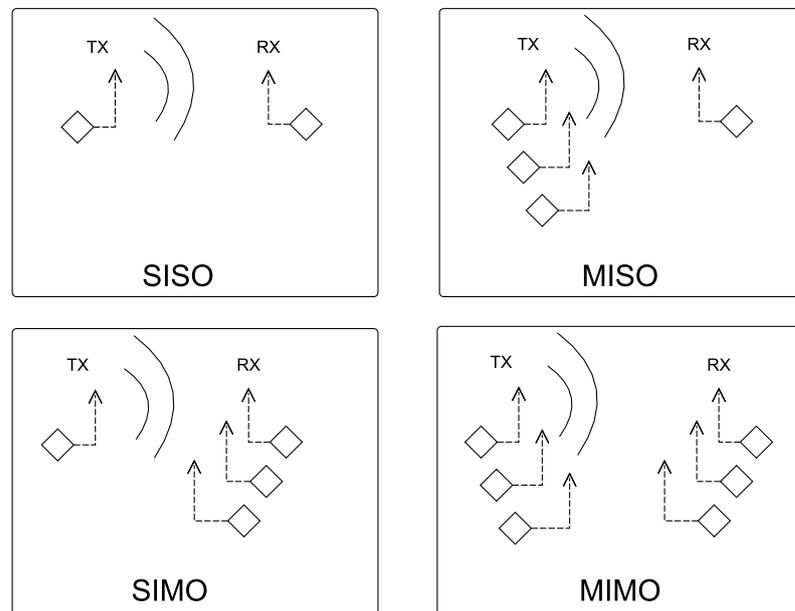


FIGURE 1.16 – Configurations de systèmes SISO, MISO, SIMO et MIMO

A travers ces différentes configurations, plusieurs techniques de transmission ont été développées afin de tirer bénéfices du nombre d'antennes.

1.4.2.1 Notion de diversité et de capacité

Un gain en diversité peut être observé lorsque plusieurs images du signal émis sont disponibles en réception. Cela se traduit par une augmentation de la probabilité de recevoir correctement le signal (de subir un évanouissement au même instant diminue). Le nombre de versions reçues est appelé ordre de diversité ou diversité. Dans le cas de configurations à multiples antennes, une diversité spatiale peut être exploitée. On notera qu'il est aussi possible d'obtenir de la diversité par d'autres moyens comme le relayage et l'utilisation de protocoles de retransmission etc.

Plusieurs types de diversité sont possibles :

1. diversité temporelle,

2. diversité fréquentielle,
3. diversité spatiale,
4. diversité de polarisation.

La diversité temporelle est obtenue en émettant une image d'un même signal à différents instants (temps de cohérence du canal) via plusieurs antennes. L'information est préalablement codée puis transmise de manière spécifique sur chaque antenne d'émission. Les codes espace-temps comme l'encodage d'Alamouti peuvent être mis en œuvre. Le schéma d'encodage d'Alamouti pour une configuration MISO 2x1 avec deux antennes d'émission et 1 antenne de réception est indiqué dans le tableau 1.2.

TABLE 1.2 – Alamouti Encoding

	Antenna 1	Antenna 2
Temps t	x_1	x_2
Temps $t+T$	$-x_2^*$	x_1^*

A l'instant t , les antennes 1 et 2 transmettent le couple de données (x_1, x_2) puis à l'instant $t + T$, le couple $(-x_2^*, x_1^*)$ où $*$ est le complexe conjugué.

La diversité fréquentielle consiste à transmettre sur des sous-bandes différentes et la diversité spatiale peut être obtenue à condition que les antennes soient suffisamment espacées spatialement. On peut alors considérer que les canaux vus par chaque antenne et leurs perturbations sont indépendants (non corrélés).

L'augmentation du nombre d'antennes permet d'obtenir des propriétés intéressantes comme l'augmentation de la capacité. Celle-ci est définie comme étant le nombre maximal de bits d'information pouvant être transmis sans erreur par unité de temps dans une bande de fréquence donnée. Ainsi, des techniques s'appliquant à des configurations MIMO comme le *beamforming* analogique c-à-d la formation de faisceau concentrant l'énergie transmise dans une direction ou encore des techniques de pré-codage numériques (*Zero Forcing* (ZF), *Minimum Mean Square Error* (MMSE), etc.), les codages espaces-temps, le multiplexage spatial pour un utilisateur (*Single User -MIMO*) ou plusieurs utilisateurs (*Multiple User*) peuvent être employées pour améliorer les performances. Dans le cadre de la thèse, une technique particulière de pré-codage appelé Retournement Temporel (RT, *Time Reversal*) a été étudiée et est décrite dans les paragraphes suivants. Les propriétés et performances de cette technique, appliquée à des systèmes multi-antennes OFDM, ont été étudiées dans [Dub13, Bou14].

La diversité de polarisation peut être obtenue à partir de réseau d'antennes à polarisation orthogonale entre elles. De cette manière, les évanouissements subis lors de la transmission par chaque composante sont généralement indépendants.

1.4.2.2 Technique de retournement temporel

La technique de Retournement temporel est une technique de pré-codage particulière qui permet de focaliser et transmettre un maximum d'énergie d'un émetteur vers un récepteur. Pour cela, il faut être en mesure de connaître la réponse fréquentielle ou temporelle du canal vue de l'émetteur (lien descendant), de manière à adapter les signaux transmis pour subir le moins de pertes possibles en énergie. Une hypothèse de réciprocité des canaux sur les deux liens est alors émise.

Le principe de retournement temporel peut s'appliquer dans le domaine temporel comme dans le domaine fréquentiel. Deux étapes sont nécessaires pour mettre en œuvre ce principe, qui est illustré pour un système SISO en Figure 1.17 :

1. le sondage de canal,
2. la focalisation.

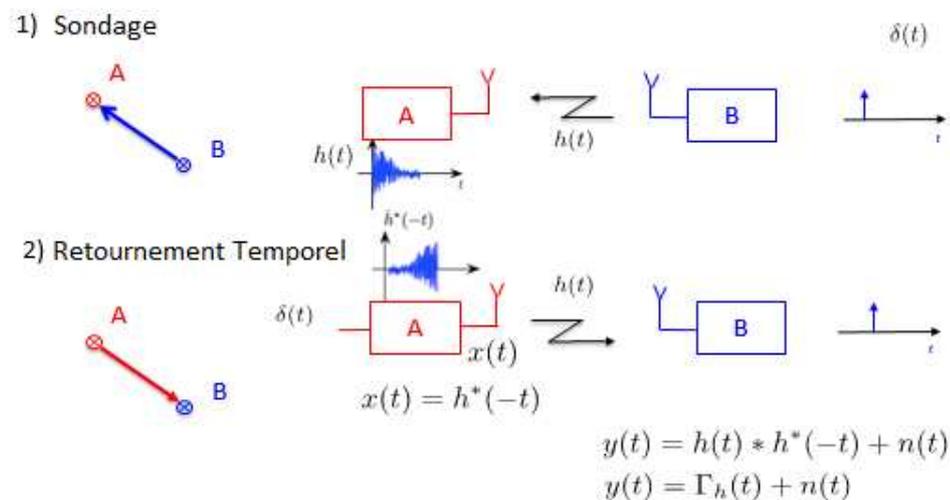


FIGURE 1.17 – Principe du Retournement Temporel [HPCK15]

Deux hypothèses sont fixées :

1. canal supposé invariant dans le temps,
2. réciprocité des canaux (liens montant et descendant).

La première étape consiste à sonder le canal du récepteur vers l'émetteur de manière à déterminer la réponse impulsionnelle du canal. Une fois ce sondage effectué, l'émetteur transmet à son tour une version retournée temporellement de la réponse impulsionnelle obtenue précédemment. Cela signifie que les derniers échos arrivés au niveau de l'émetteur lors de l'étape de sondage, sont transmis en premier. De cette manière, l'ensemble des trajets générés permet d'obtenir en réception un pic de signal d'une forte puissance comme

indiquée en Figure 1.18.

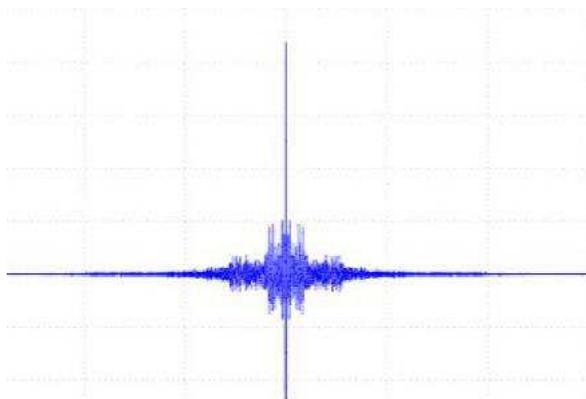


FIGURE 1.18 – Pic d’auto-corrélation en réception

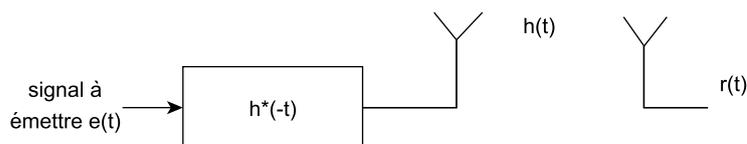


FIGURE 1.19 – Intégration du Retournement Temporel dans un système [Dub13]

D’un point de vue théorique, l’intégration du retournement temporel est très simple. En effet, elle consiste à appliquer un filtre sur le signal à transmettre. Comme indiqué sur la Figure 1.19, si l’on note $h(t)$ la réponse impulsionnelle du canal obtenue durant la phase de sondage, en réception, le signal reçu peut s’exprimer de la manière suivante :

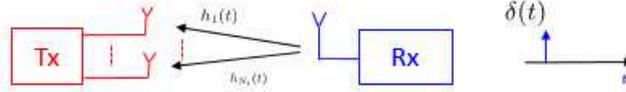
$$y(t) = x(t) \otimes h^*(-t) \otimes h(t) + n(t) \quad (1.10)$$

$$= x(t) \otimes \Gamma_h(t) + n(t) \quad (1.11)$$

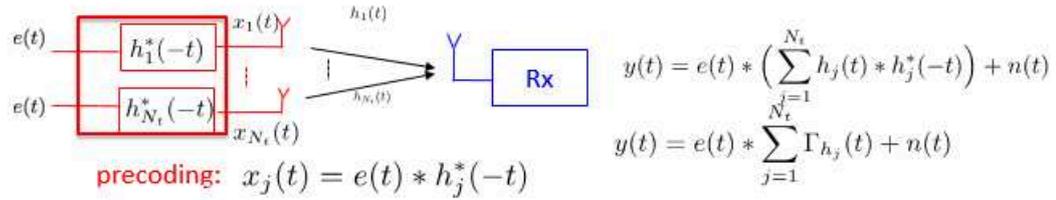
avec $x(t)$ le signal à émettre, $n(t)$ le bruit blanc additif gaussien, $\Gamma_h(t)$ l’auto-corrélation de la réponse impulsionnelle du canal et $*$ représente le conjugué. Le canal équivalent est l’auto-corrélation de la réponse impulsionnelle du canal. Dans le cas d’un système parfait, $\Gamma_h(t)$ tend vers l’impulsion de Dirac $\delta(t)$ [Dub13]. Dans ce cas, le récepteur est grandement simplifié puisque qu’il n’y a plus d’égalisation.

Sur la Figure 1.20, le schéma de principe du RT est présenté dans le cas d’un système MISO à N antennes d’émission [HPCK15]. Dans ce cas, le canal équivalent correspond à la somme des fonctions d’auto-corrélation. Plus le nombre d’antennes augmente, plus la focalisation spatiale et temporelle est grande.

1) Sondage



2) Retournement Temporel



Equivalent channel

$$\sum_{j=1}^{N_t} h_j(t) * h_j^*(-t) = \sum_{j=1}^{N_t} \Gamma_{h_j}(t)$$

FIGURE 1.20 – Illustration du principe de Retournement Temporel dans un système MISO [HPCK15]

D'un point de vue pratique, lorsque le nombre N d'antennes du système augmente, il est généralement préféré d'appliquer le principe de Retournement Temporel en fréquence afin de limiter la complexité. En effet, dans le domaine fréquentiel, une multiplication complexe à 1 coefficient sera nécessaire par antenne d'émission tandis que dans le domaine temporel, un filtre par antenne d'émission sera requis.

La focalisation spatiale permet notamment d'introduire de la sécurité en s'assurant d'émettre localement autour du récepteur, tout en réduisant les interférences vis-à-vis des récepteurs voisins. La focalisation temporelle permet de réduire les interférences entre symboles. Un des principaux inconvénients reste l'hypothèse de réciprocity du canal. En effet, à cause des différents éléments des étages RF, le canal du point de vue du récepteur n'est pas forcément identique à celui perçu par l'émetteur. Néanmoins, cette technique semble prometteuse dans le cadre de communications dites vertes, où d'une part le récepteur peut être grandement simplifié et d'autre part, la possibilité de réduire la puissance de transmission grâce à la focalisation spatiale et temporelle [WWH⁺11, CHY⁺14]

1.5 Performances, enjeux énergétiques et problématiques

Dans le contexte actuel, de nouveaux enjeux apparaissent pour continuer à augmenter les performances tout en satisfaisant des contraintes énergétiques. L'amélioration des performances est toujours un objectif primordial dans les systèmes de communications numériques sans fil. Néanmoins, la complexité croissante notamment des couches basses

(physique et MAC) due au grand nombre de ressources de calculs, conduit à des consommations énergétiques et des dissipations de puissance importantes.

Dans le domaine des communications numériques, lors de l'implantation matérielle de telles chaînes, un certain nombre de questions sont soulevées autour de la fiabilité, de l'autonomie, de la durée de vie ainsi que des coûts de refroidissement liés à l'utilisation de ces technologies. Cette contrainte est d'autant plus forte lorsque le système est embarqué.

1.5.1 Évaluation des performances

Évaluer les performances des systèmes de communications numériques sans fil s'effectue généralement à haut-niveau à l'aide d'outils permettant de modéliser la fonctionnalité du système comme par exemple Matlab. Ainsi, la fonctionnalité du système peut être testée et validée, sans aucune information précise sur le système réel. Ceci permet de valider rapidement le système en faisant abstraction du matériel.

Plusieurs métriques de performances peuvent être ainsi déterminées :

- le taux d'erreurs binaire (TEB ou *BER*, *Bit Error Rate*),
- le taux d'erreurs par paquet (TEP *PER*, *Packet Error Rate*),
- le débit,
- l'efficacité spectrale,
- etc.

Le taux d'erreurs binaire est évalué en déterminant le nombre de bits correctement reçu par rapport au nombre de bits émis. Le débit correspond au nombre de bits transmis par seconde. L'efficacité spectrale est définie par le rapport du débit sur la bande passante du canal de transmission. Elle traduit la capacité d'un système à transmettre une quantité de bits utiles par seconde et par Hz de la bande de transmission.

1.5.2 Estimation de la consommation d'énergie

Évaluer la consommation énergétique d'un système, que ce soit dans le domaine des communications numériques ou un tout autre domaine, est un enjeu déterminant afin de permettre aux concepteurs d'effectuer des choix algorithmiques, de configuration, etc.

Or, un moyen d'obtenir des résultats précis est de mesurer directement la consommation sur un prototype. Cela requiert donc que le système soit entièrement conçu ainsi qu'un banc de mesures approprié.

Lorsque le temps et les moyens nécessaires au développement d'un prototype ne sont pas envisageables, des estimations de consommation peuvent être effectuées à partir d'un modèle du système. Dans ce cas, sachant que le système est modélisé à l'aide d'outils ou langages, c'est le niveau de détails de description du système et le choix de la technique d'estimation qui permettent d'obtenir une estimation plus ou moins précise. Un autre facteur important à considérer est le temps de simulation. En effet, plus les niveaux de précision et de finesse de description du système seront élevés, plus les temps de simulation

seront grands.

Il y a donc un compromis à faire entre la précision de l'estimation et la rapidité d'estimation.

Dans l'optique de comparer plusieurs systèmes entre eux d'un point de vue énergétique, la métrique d'efficacité énergétique est généralement évaluée. Elle permet de quantifier la capacité d'un système à transmettre un maximum de bits d'information en utilisant un minimum d'énergie.

L'efficacité énergétique [Bou14, HIXR15, LXX⁺11] peut être définie par l'équation 1.12 :

$$EE(\text{bit}/J) = \frac{\text{Capacité } [\text{bit}/s]}{P_{\text{totale}} [W = J/s]} \quad (1.12)$$

avec P_{totale} , la puissance moyenne totale consommée qui est égale à la somme de la puissance allouée à la transmission des données à travers le canal (Puissance rayonnée) et de la puissance moyenne consommée par le circuit.

1.5.3 Problématiques

Dans le cadre de la thèse, il est envisagé de comparer les traitements effectués en bande de base de plusieurs systèmes de communications numériques, en termes de performances et de consommation énergétique.

Généralement, la fonctionnalité du système étudié est modélisée à haut-niveau, c-à-d en faisant abstraction de la manière dont le système sera conçu matériellement. Dans le domaine des communications numériques, des outils comme Matlab sont couramment utilisés pour évaluer rapidement les performances d'un système. Ce type d'outil ne permet pas d'évaluer la consommation énergétique du système modélisé car aucune cible matérielle n'est définie. Pour obtenir des estimations de consommation précises, le système doit être entièrement défini et implanté physiquement sur une plate-forme matérielle connue.

A partir de ce constat, il semble primordial de pouvoir remonter des informations fiables et précises provenant des niveaux bas (où le système est entièrement défini) jusqu'au niveau système où les performances sont étudiées. De cette manière, il sera envisageable de pouvoir estimer précisément et rapidement la consommation d'un système de communication numériques, tout en évaluant conjointement ses performances.

D'autre part, il peut être constaté qu'un grand nombre de paramètres comme par exemple le choix de taille d'IFFT (dans le cadre de la modulation OFDM), le nombre d'antennes d'émission et de réception, le choix des techniques de codage de canal et de décodage, peuvent avoir un impact important à la fois sur les performances mais aussi sur la consommation finale du système. Faciliter l'exploration de l'espace de conception est aussi un objectif important afin de permettre de tester un grand nombre de configurations tout en évaluant le compromis performances-consommation.

Nous avons choisi de focaliser notre étude sur la comparaison de différents traitements en bande de base dans le domaine des communications numériques sans fil. Il est géné-

ralement considéré que c'est l'amplificateur de puissance qui consomme la majorité de la puissance et ceci de manière très significative devant les autres éléments d'une chaîne.

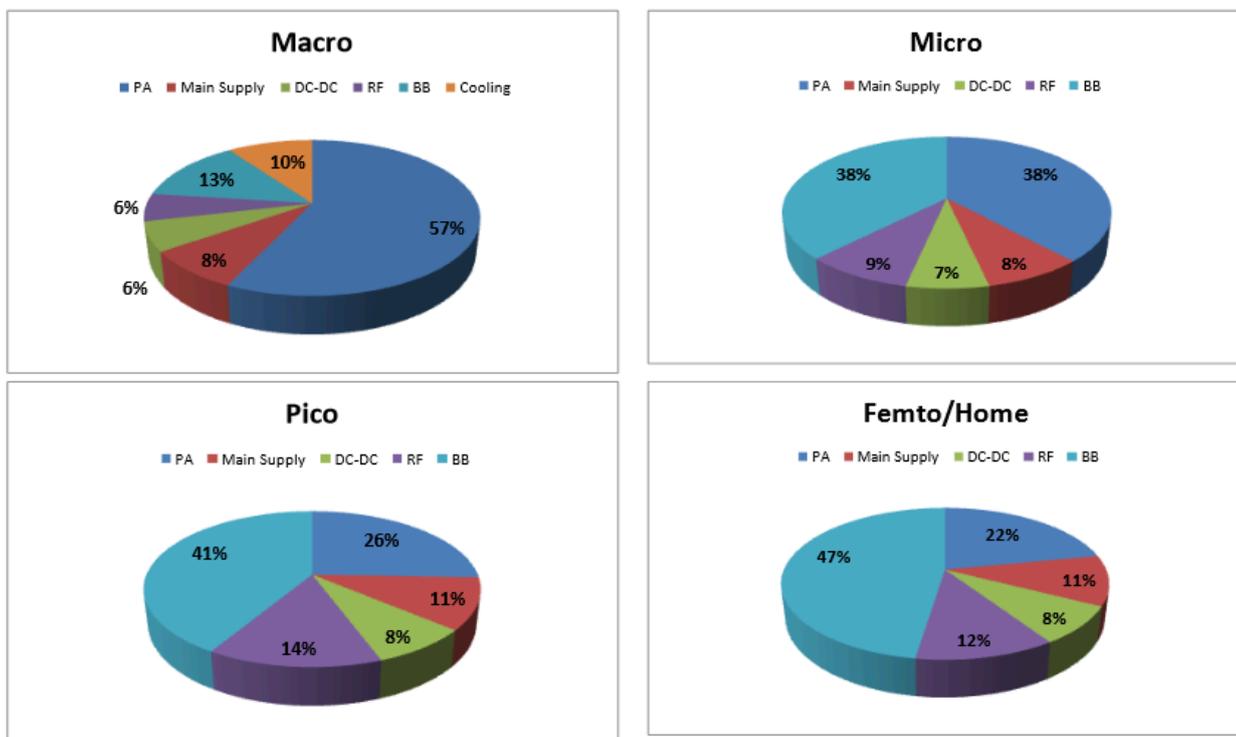


FIGURE 1.21 – Répartition de la consommation des éléments d'une station de base suivant différents types de déploiement [A⁺12]

D'après une étude du projet EARTH, la Figure 1.21 présente les répartitions de la consommation des différents éléments d'une station de base en voie descendante pour plusieurs scénarios de déploiement dans le contexte du LTE. On peut remarquer que, pour une station de base pouvant couvrir des centaines d'utilisateurs sur plusieurs centaines de mètres (cellule de réseau mobile dite Macro), l'amplificateur de puissance consomme plus de la moitié (57%) de la puissance totale consommée. Plus la taille de cellule diminue, plus l'impact de la puissance consommée par les traitements en bande de base sur la puissance totale augmente. Ainsi, pour une taille de cellule très faible (Femto/Home) on constate que l'amplificateur ne consomme plus que 22% de la puissance totale contre 47% pour les traitements en bande de base.

Dans le contexte de la multiplication des équipements des objets connectés, il semble raisonnable de considérer que l'impact énergétique des traitements en bande de base sera aussi important que celui de l'amplificateur de puissance. En effet, les puissances allouées à la transmission des données pourraient être relativement faibles, pour de faibles distances, devant la puissance consommée par les traitements en bande de base. De plus, la complexité des traitements en bande de base a tendance à augmenter.

A travers ces différentes constatations, le choix de l'étude de la consommation des traitements en bande de base de systèmes de communications numériques est ainsi justifié.

1.6 Conclusion

Dans ce chapitre, nous avons pu faire une description générale d'un système de communications numériques sans fil ainsi que des caractéristiques du canal de propagation. Ensuite, une présentation de la modulation OFDM ainsi que son utilisation dans la couche physique de la technologie du réseau mobile LTE ont été effectuées. Plusieurs problématiques autour des enjeux énergétiques et de l'évaluation des performances des systèmes de communications numériques sans fil ont été soulevées :

1. Comment remonter l'information de consommation à haut-niveau afin d'évaluer conjointement le compromis performances versus consommation énergétique ?
2. Comment estimer rapidement et facilement la consommation d'énergie des traitements en bande de base à haut-niveau ? Pour de multiples configurations ?

Répondre à ces problématiques devrait permettre de comparer différents systèmes entre eux de manière efficace en tenant compte à la fois des performances et des aspects énergétiques.

Pour évaluer précisément la consommation énergétique d'un système, il est nécessaire de réaliser des étapes de conception par rapport à une cible matérielle. Ces étapes peuvent être coûteuse en temps. Des circuits logiques programmables tels que les FPGAs sont généralement développés dans l'objectif d'effectuer du prototypage rapide et ainsi obtenir des valeurs de consommation et performances qui peuvent permettre d'obtenir des ordres de grandeurs avant l'implantation d'un système sur ASIC. De plus, les FPGAs permettent d'introduire de la flexibilité notamment grâce à la possibilité de les reconfigurer. A travers ces quelques exemples non-exhaustifs, les FPGAs semblent être une technologie attractive, qui permettrait de répondre et de s'adapter aux contraintes actuelles des systèmes de communications numériques.

Chapitre 2

La consommation d'énergie dans les circuits de type FPGA

2.1 Introduction

Le transistor est l'élément de base de l'électronique actuelle. Son évolution, ainsi que ses procédés de fabrication rythment encore aujourd'hui l'évolution des systèmes. Dans les années 60, Gordon Moore proposait une loi, faisant référence, prédisant que le nombre de transistors par unité de surface doublerait tous les deux ou trois ans. De nos jours, l'évolution des systèmes est confrontée à de nouvelles limites technologiques et physiques. Les densités de puissance par unité de surface des puces sont telles que leur fiabilité peut être compromise et les coûts liés au refroidissement de la puissance dissipée deviennent très importants.

Une avancée majeure dans la création des circuits intégrés (CI) a été le développement de la technologie CMOS (*Complementary Metal Oxide Semiconductor*). Comme illustré à la Figure 2.1, extraite des travaux de [CSC99], cette technologie a remplacé les transistors bipolaires qui étaient notamment beaucoup plus énergivores.

La technologie CMOS est aujourd'hui largement utilisée dans les circuits FPGA développés à base de cellules de type SRAM (*Static Random Access Memory*) dont chaque cellule est composée d'environ 4 à 6 transistors.

L'apparition du premier FPGA, fabriqué autour des cellules SRAM, daterait de 1967 [KTR08]. C'est par le biais de l'intégration de la technologie SRAM qu'a été introduite la re-programmabilité des éléments logiques ainsi que des interconnexions. Aujourd'hui, c'est une des familles de FPGA la plus répandue.

Actuellement, le marché des FPGAs est toujours en pleine croissance. Deux fournisseurs de FPGA se disputent principalement ce marché : Xilinx et Altera (environ 80% à eux deux). Néanmoins on peut aussi citer d'autres acteurs comme Atmel et Lattice. Pour souligner l'essor de cette technologie, on peut noter que la société Altera vient d'être acquise par Intel au cours de l'année 2015 dans l'objectif d'être compétitif face au partenariat entre Xilinx et ARM. Ceci souligne l'attrait de la technologie FPGA pour la conception

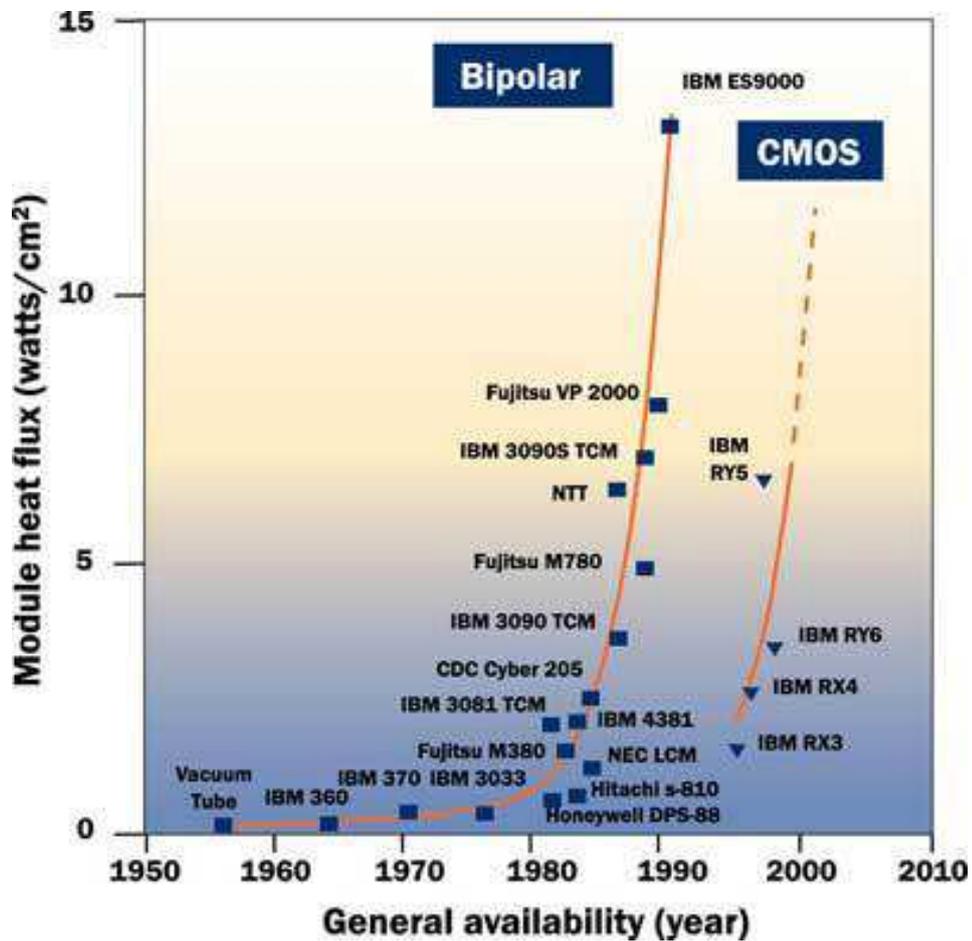


FIGURE 2.1 – Évolution de la puissance dissipée, comparaison de la technologie bipolaire et CMOS [CSC99]

des futurs systèmes.

2.2 Les circuits FPGA

2.2.1 Qu'est-ce qu'un FPGA ?

Cet acronyme de *Field Programmable Gate Array* désigne littéralement une 'Matrice/réseau de portes logiques programmables'. Ces circuits programmables électriquement peuvent être destinés à intégrer toutes sortes de systèmes numériques et peuvent être reprogrammés de nombreuses fois.

Comme indiqué sur la Figure 2.2, les FPGAs font partie de la famille des circuits logiques programmables (CPL). On y retrouve en plus des FPGAs, les CPLDs (*Complex*

Programmable Logic Device) ainsi que les PALs (*Programmable Array Logic*).

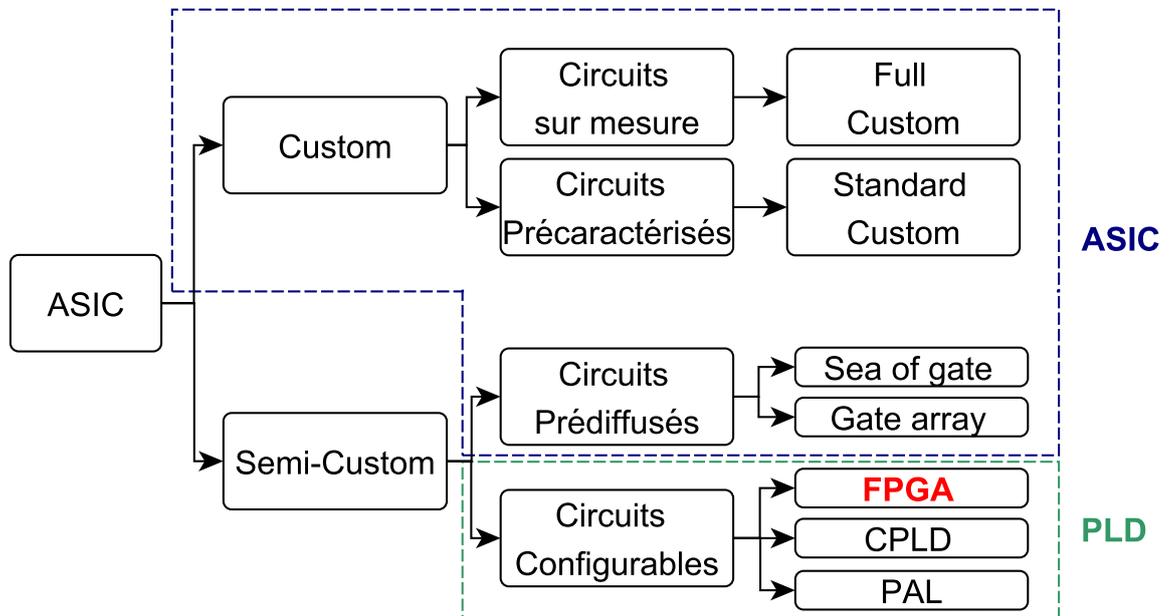


FIGURE 2.2 – Positionnement des FPGA

Les PALs sont des matrices de portes ET logiques programmables connectées aux sorties par des portes OU logiques. Les CPLDs sont des circuits qui regroupent plusieurs PALs à l'aide d'une matrice d'interconnexions supplémentaire. Quant aux FPGAs, ils possèdent généralement beaucoup plus de ressources que les CPLDs mais ont une structure relativement différente.

Les circuits ASICs dits "*full custom*", *entièrement personnalisés* sont généralement spécifiques à une application donnée. L'ensemble du circuit peut être défini par l'utilisateur, ce qui engendre des coûts relativement importants. Néanmoins, le circuit est optimisé pour une application et des critères donnés.

Les circuits dits *pré-caractérisés* sont généralement construits autour de cellules élémentaires contenues dans des bibliothèques. Ces cellules peuvent contenir de la logique, des micro-processeurs ou encore des mémoires.

Pour les circuits dits *pré-diffusés*, certains éléments sont déjà présents dans le circuit mais ne sont pas connectés entre eux. Les circuits FPGAs font partie de la famille des circuits dits *semi-personnalisés*.

Chaque famille de FPGA possède des caractéristiques physiques qui lui sont propres. Les procédés de fabrication, la technologie de conception peuvent être néanmoins différents. La famille la plus répandue est celle fabriquée autour de cellules SRAM, fabriquées à partir de la technologie CMOS. D'autres technologies existent comme la technologie à fusibles (programmable 1 fois) etc.

Chaque FPGA est généralement composé de plusieurs éléments de base que sont :

1. des blocs logiques programmables réalisant des fonctions logiques simples ou de la mémorisation,
2. des éléments d'interconnexion entre les différentes ressources,
3. des broches d'entrées/sorties (*I/O*),
4. des éléments mémoires embarqués,
5. des blocs de calcul spécifiques.

Tous les FPGAs sont constitués d'un nombre fixe de ressources, comportant des blocs logiques, des interconnexions programmables et des blocs d'entrées/sorties, permettant de concevoir une multitude de circuits numériques.

Sur la Figure 2.3-a, est schématisée une architecture classique de FPGA où l'on retrouve les principaux éléments énumérés précédemment. L'élément de base du FPGA est appelé un bloc logique configurable (CLB).

Sur la Figure 2.3-b, on peut remarquer que chaque bloc logique configurable est constitué de bascules (FF) et de tables de vérité (*Look-Up Table (LUT)*) ainsi que de multiplexeurs. Il est ainsi possible d'effectuer des traitements purement combinatoires en utilisant uniquement la LUT, ou des traitements séquentiels à l'aide de la bascule, ou encore les deux types de traitements à la fois. Ce sont les LUTs qui permettent l'implantation des opérations logiques de base ('ET', 'OU', etc).

Dans les FPGAs récents, d'autres éléments matériels ont été ajoutés tels que des blocs mémoires RAM, des blocs multiplieurs accumulateurs appelés blocs DSP ainsi que des modules de gestion d'horloge (DCM pour *Digital Clock Manager*) comme indiqué sur la Figure 2.4.

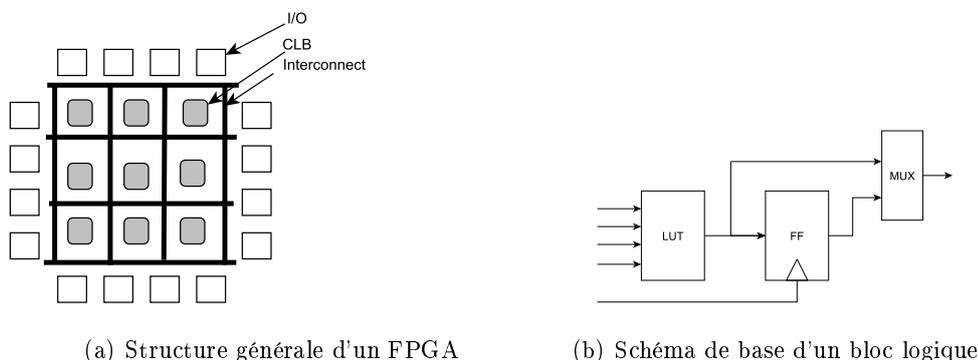


FIGURE 2.3 – Architecture simplifiée des FPGAs

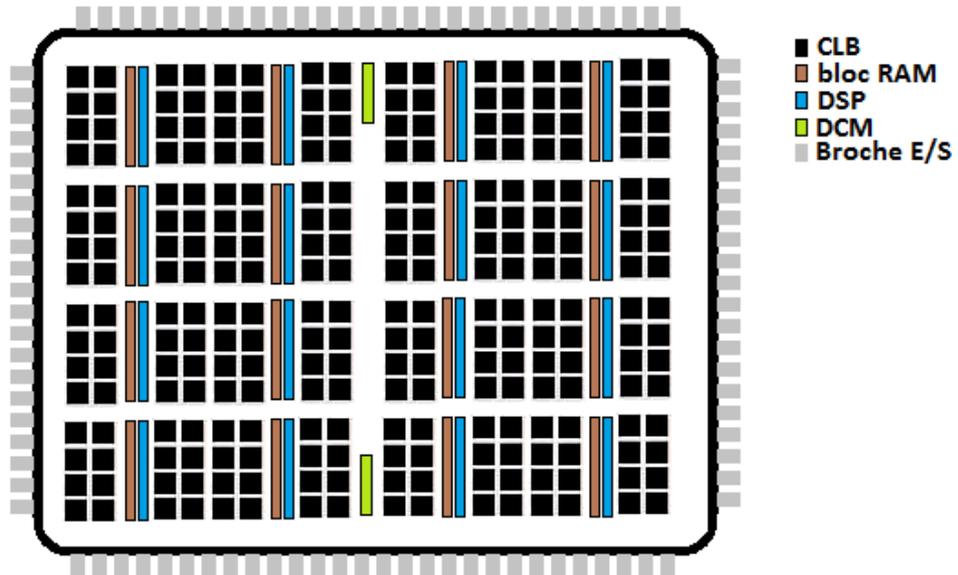


FIGURE 2.4 – Exemple d’architecture de FPGA

2.2.2 Flexibilité versus performances

Les FPGAs permettent d’introduire de la flexibilité dans un système, en comparaison à d’autres circuits comme les ASICs qui possèdent une architecture extrêmement optimisée pour une application donnée. La flexibilité des FPGAs reste leur plus grand avantage face aux ASICs au détriment d’une plus grande consommation énergétique et des performances dégradées [KR07]. Néanmoins, le temps de développement est nettement plus court dans le cas des FPGAs et les coûts associés sont relativement faibles. Le temps de développement d’un ASIC est environ plus long de 55%, que celui d’un FPGA [Par04b].

De part la flexibilité des FPGA, il est possible de reconfigurer dynamiquement une partie ou encore la totalité du FPGA. De cette manière, la fonctionnalité du circuit peut être modifiée en fonction des besoins de l’application. Une comparaison des principaux avantages et inconvénients entre ASIC et FPGA est présentée dans le tableau 2.1.

Critère	FPGA	ASIC
Temps de développement	faible	élevé
Coûts	faibles à moyens	élevés
Consommation d’énergie	élevée	faible
Performances	moyennes	élevées
Flexibilité	élevée	faible

TABLE 2.1 – ASIC versus FPGA

2.2.3 Flot de conception d'un FPGA

Plusieurs étapes sont nécessaires afin de réaliser la conception d'un circuit sur FPGA. Sur la Figure 2.5, la fonctionnalité d'un système est tout d'abord décrite à partir d'un ou plusieurs fichiers de langage de description matérielle (*Hardware Description Language, HDL*) de type VHDL ou Verilog. Une simulation dite comportementale, *behavioural*, permet de valider la fonctionnalité du système avant de passer à l'étape suivante. Ensuite, une étape de synthèse du design permet de traduire le langage de description matérielle sous la forme d'un ensemble d'éléments logiques (c-à-d macro) appelé *netlist*. Ces 'macros' peuvent être par exemple des additionneurs, des blocs mémoires, etc. Différents outils de synthèse existent comme Precision synthesis de MentorGraphics, XST de l'environnement ISE de Xilinx, Synplify^R synthesis de Synopsys ou encore Quartus II d'Altera (liste non-exhaustive).

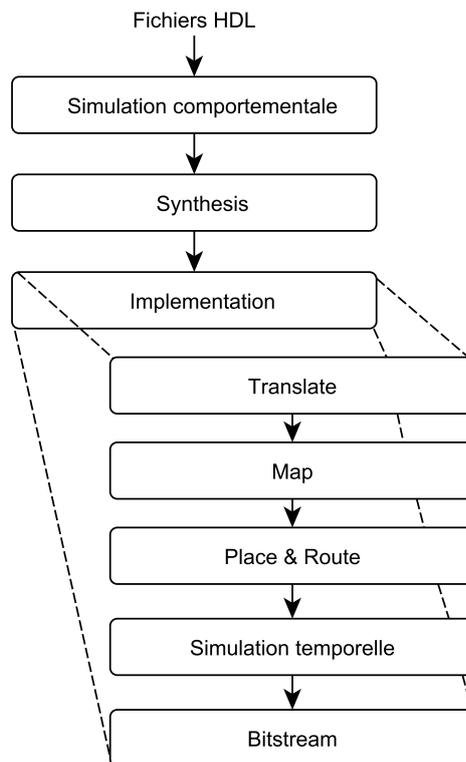


FIGURE 2.5 – Flot de conception typique de FPGA

Une fois l'étape de synthèse achevée, la phase d'implémentation du FPGA débute. Comme son nom l'indique, cette phase permet l'implantation physique du design pour une cible FPGA donnée. Pour cela, l'implémentation est réalisée en 3 étapes :

1. Translate,
2. Mapping,

3. et Placement et Routage.

Les étapes de *Translate* et de *Mapping* visent à associer et allouer les macros à des ressources physiques du FPGA. Ensuite, l'étape de placement/routage vise à placer ces ressources dans le FPGA et à les relier entre elles par le biais d'interconnexions. On parle alors de routage. Une autre simulation est alors effectuée. Cette fois-ci, une simulation dite 'temporelle', qui en comparaison à la simulation comportementale, va permettre de prendre en compte les contraintes de timings, c-à-d des temps de propagation de chaque élément du circuit réalisé. Une fois les contraintes temporelles respectées, une dernière étape permet de générer un fichier de configuration du FPGA appelé *bitstream*. Celui-ci sera ensuite téléchargé dans le FPGA afin d'effectuer sa programmation.

2.3 Origines de la consommation dans un FPGA

Dans le cas de FPGA de type SRAM, la puissance totale dissipée provient de plusieurs sources : la puissance dynamique et la puissance statique, comme indiquée par l'équation 2.1 :

$$P_{FPGA} = P_{Dynamique} + P_{Statique} \quad (2.1)$$

Chacune de ces puissances dissipées est décrite plus en détails dans les paragraphes suivants.

2.3.1 Puissance statique

Précédemment, il a pu être constaté que les FPGAs de type SRAM sont réalisés à partir de transistors de technologie CMOS. Lorsque ces transistors sont inactifs, une puissance dite statique est dissipée. Cette puissance est générée par la présence de nombreux courants de fuite au niveau de ces transistors dont les origines et les mécanismes sont relatés dans [RMMM03]. Les trois principaux courants sont indiqués sur la Figure 2.6 et listés ci-dessous :

1. courants sous la tension de seuil (*I1*),
2. courants tunnels au niveau de l'oxyde de grille (*I2*),
3. courants de fuite au niveau des jonction PN (*I3*).

Les courants de fuite sous la tension de seuil, *sub-threshold leakage current*, apparaissent lorsqu'une tension inférieure à la tension de seuil, V_{th} , est appliquée entre la grille et la source du transistor, V_{GS} .

D'après l'équation reprise de [LDSY07], le courant de fuite sous la tension de seuil peut s'exprimer de la manière suivante :

$$I_{subthreshold} \cong \frac{W}{L} \left(\frac{kT}{q}\right)^2 e^{\frac{q(V_{GS}-V_{Th})}{nkT}} \quad (2.2)$$

où W et L correspondent respectivement à la largeur et la longueur du canal, k la constante de Boltzmann, T la température (K), q la constante de charge d'un électron, V_{GS} la tension

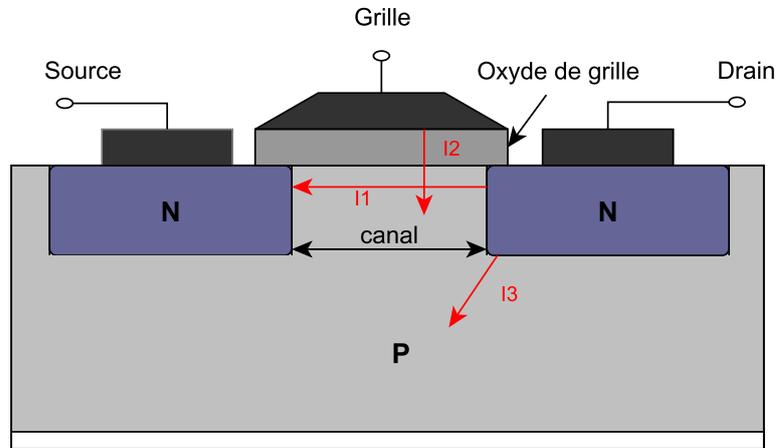


FIGURE 2.6 – Courants de fuite d'un transistor NMOS

entre la grille et la source, n un coefficient dépendant des caractéristiques technologiques du transistor et V_{Th} la tension de seuil. D'après l'équation 2.2, on peut remarquer que ce courant varie de manière exponentielle en fonction de la tension de seuil V_{Th} et de l'inverse de la température.

La puissance statique est aussi induite par des courants présents au niveau de l'oxyde de grille du transistor. Des courants dits 'tunnels' se forment au passage d'électrons entre le substrat et la grille. Les courants tunnels présents au niveau de l'oxyde de grille sont directement liés à la miniaturisation des transistors, impliquant la réduction des épaisseurs des oxydes. Plus les épaisseurs sont fines, plus ces courants de fuites sont importants.

On dénote aussi une autre source de courant de fuite qui se situe au niveau des jonctions P/N des transistors entre le drain et le substrat. En effet, ces jonctions forment des diodes P/N qui possèdent des courants de polarisation inverses. Ce courant de fuite est appelé courant de fuite inverse des jonctions P/N, *PN Junction Reverse-Bias Current*. Ce courant noté I_{PN} s'exprime de la manière suivante [KL03] :

$$I_{PN} = I_s \cdot e^{\left(\frac{qV_p}{kT} - 1\right)} \quad (2.3)$$

avec I_s le courant de saturation et V_p la tension de polarisation. On peut noter que ce courant dépend essentiellement des paramètres technologiques des transistors.

Finalement, la puissance statique dans un FPGA est généralement modélisée sous l'équation suivante 2.4 :

$$P_{Statique} = V_{cc} * \sum i_{fuite} \quad (2.4)$$

avec V_{cc} , la tension d'alimentation et $\sum i_{fuite}$ la somme des courants de fuites.

Les principaux facteurs ayant un impact sur la puissance statique sont :

- la tension d'alimentation,
- la température,

– et le procédé technologique.

Dans le cas d'un circuit FPGA, des courants de fuites sont aussi présents au niveau des broches d'entrées et de sorties du FPGA, où des courants continus sont présents. L'ensemble des blocs de génération d'horloges comme les boucles à verrouillage de phase (*Phase-locked loop*, PLLs) contribue donc à cette puissance statique. De plus, un certain nombre de transistors du plan de configuration du FPGA qui sert à maintenir les interconnexions lorsque le FPGA est configuré, dissipe de la puissance statique.

Longtemps négligée, la puissance statique est désormais relativement importante dans les circuits FPGAs récents. En effet, la miniaturisation des transistors est un des premiers facteurs de l'augmentation de la puissance statique. La diminution des tailles des transistors a un impact direct sur les épaisseurs des oxydes servant d'isolant électrique. Parallèlement, la miniaturisation permet l'accroissement de la densité d'intégration c-à-d du nombre de transistors par unité de surface. A titre d'exemple, début 2014, le pas technologique était de 28nm pour les Virtex-7 de Xilinx. Début 2015, des puces à base de transistors de 22nm sont réalisées notamment par Xilinx et Intel.

Sur la Figure 2.7 provenant de l'ITRS, on constate bien la contribution croissante des courants de fuites à la puissance dissipée en fonction de la réduction des dimensions des transistors (axe ordonnée gauche) au fil des années (axe des abscisses).

Cette puissance statique peut être réduite mais reste difficilement optimisable par l'utilisateur du FPGA. Une première stratégie de réduction de la puissance statique consiste à diminuer la tension d'alimentation des transistors. Or, diminuer la tension d'alimentation augmente les temps de commutation des transistors réduisant ainsi les performances.

Cette puissance dissipée est fortement dépendante de la technologie de conception du FPGA. Ce sont donc les constructeurs de FPGA qui sont les premiers acteurs pouvant réduire cette puissance statique par le biais de l'introduction d'innovations technologiques.

Une innovation importante a été l'introduction d'un oxyde de grille efficace. En effet, des diélectriques possédant une constante diélectrique élevée (*high-k*) ont peu à peu remplacé le silicium polycristallin conventionnel. En combinaison avec l'utilisation d'une grille en métal, les courants de fuite ont pu être réduits. La série 7 de FPGA de Xilinx est notamment fabriquée à partir de la technologie HKMG et du procédé technologique en 28nm. 65% de la puissance statique peut être ainsi préservée par rapport à son précédent pas technologique à 40nm [HKH15].

Néanmoins, d'autres approches peuvent être mises en œuvre. Par exemple, une technique couramment employée est l'utilisation de transistors à multiples tension de seuil, *MT-CMOS*. Ainsi, des transistors ayant une faible tension de seuil sont utilisés pour atteindre de hautes performances tandis que d'autres transistors, qui possèdent des tensions de seuils élevées, sont utilisés lorsque les performances requises sont plus faibles afin de dissiper moins de puissance statique. Cette technique est actuellement utilisée pour certains FPGAs de Xilinx et Altera par exemple. Dans [KA06], une réduction de 55% de

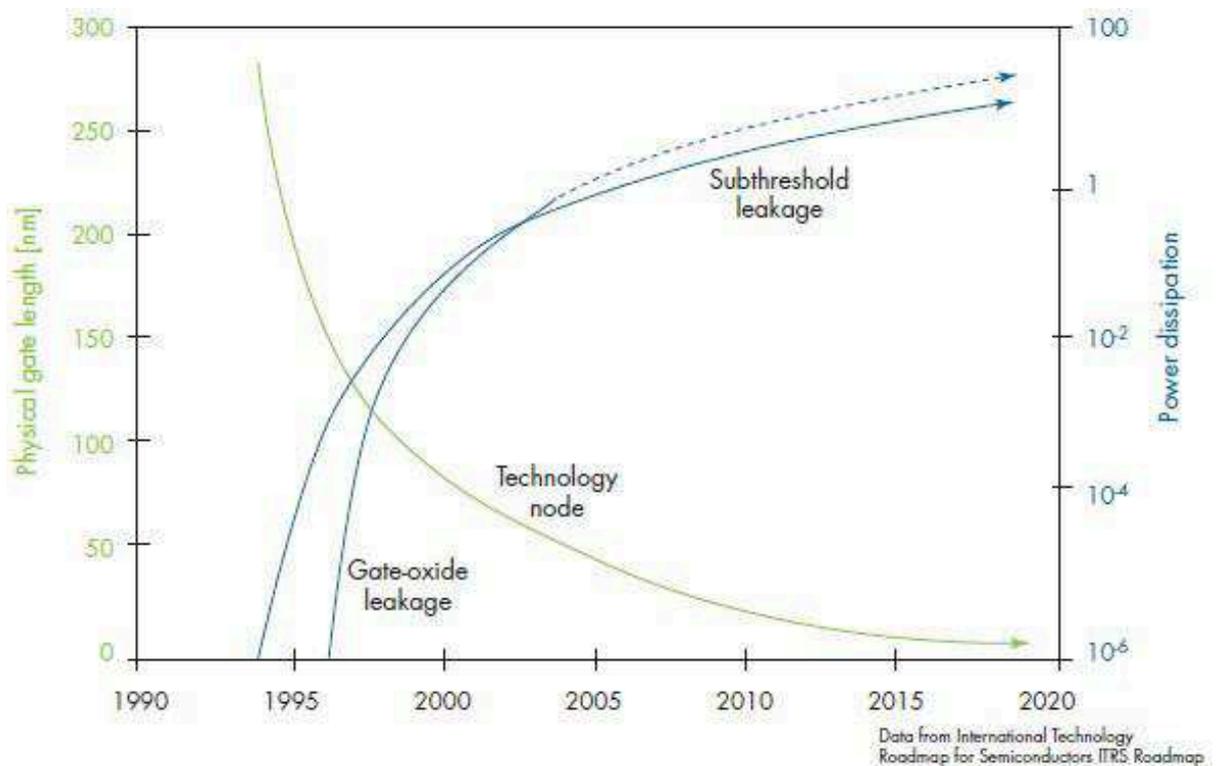


FIGURE 2.7 – Evolution de la puissance statique en fonction de l'évolution technologique [fS00]

la puissance statique est obtenue grâce à l'utilisation de transistors à fortes et à faibles tensions de seuil.

D'autres techniques de réduction des courants de fuite sous la tension de seuil sont revues dans [DN07] comme les transistors SCCMOS (*Super Cut-off CMOS*) ou encore des techniques efficaces lorsque les transistors sont en mode actif.

Le principal point faible des FPGAs basés sur la technologie SRAM réside dans le besoin d'une mémoire externe non-volatile (généralement de technologie FLASH) afin de préserver le *bitstream* et le contexte d'utilisation. En effet, pour réduire la puissance statique, des techniques de *Power Gating*, consistant à couper l'alimentation de régions du FPGA, peuvent être mises en place mais l'information contenue en mémoire est alors perdue. Des technologies innovantes comme la RRAM, *Resistive RAM*, peuvent contribuer à réduire cette consommation [TOR⁺14].

2.3.2 Puissance dynamique

La puissance dynamique correspond à la puissance dissipée lorsque le circuit est actif et donc la puissance engendrée par les commutations des transistors afin de réaliser la

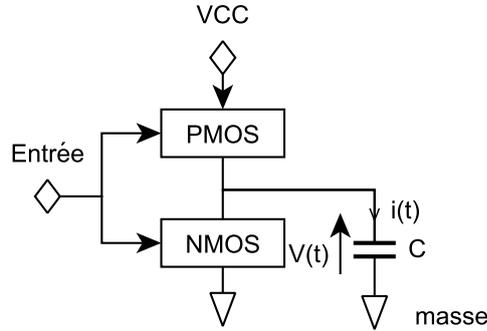


FIGURE 2.8 – Illustration d’une cellule CMOS

fonctionnalité du circuit. La puissance dynamique est composée de 2 sources :

$$P_{dynamique} = P_{sw} + P_{cc} \quad (2.5)$$

avec P_{sw} , la puissance dissipée par l’activité de commutation des transistors, P_{cc} la puissance dissipée lorsque les transistors de la cellule CMOS sont en court-circuit.

En effet, les commutations des transistors ne sont pas instantanées et pendant un court instant, les transistors N et P de la cellule CMOS sont dans un état de court-circuit. D’après [KL03], la valeur moyenne de P_{cc} s’écrit :

$$P_{cc} \cong K\tau f_{clk}(V_{cc} - 2V_{th})^3 \quad (2.6)$$

avec K une constante dépendant de paramètres électriques, τ le temps de montée/descente du signal d’entrée, f_{clk} la fréquence d’horloge, V_{cc} la tension d’alimentation, V_{th} la tension de seuil.

Toujours d’après [KL03], cette puissance peut être réduite de manière significative si l’on considère des temps de montée et descente du signal d’entrée très faibles.

Concernant P_{sw} , on peut distinguer deux types de transitions. Premièrement, certaines transitions dans le circuit vont être nécessaires à la réalisation de la fonctionnalité. Deuxièmement, d’autres commutations appelées *glitches* peuvent subvenir lorsque le circuit est actif. Ces transitions supplémentaires au niveau des transistors sont majoritairement présentes au niveau de la logique combinatoire. Ces transitions, appelées *glitches*, sont générées lorsque des signaux présents sur les entrées d’un circuit combinatoire possèdent des temps d’arrivée différents, induits notamment par les temps de propagation des portes situées en amont. Ce phénomène est mis en évidence sur la Figure 2.9. Dans le cas d’une porte ET à 3 entrées, des transitions supplémentaires peuvent être générées en sortie, avant que la sortie n’atteigne son état logique final.

Les *glitches* seraient à l’origine d’une part importante de la puissance dynamique. Si l’objectif est de concevoir un circuit à faible dissipation, différentes techniques comme l’insertion de délais programmables montrent qu’il est possible d’éliminer 87% des *glitches* et

ainsi réduire de 17% la puissance dynamique [LLW08]. L'utilisation de pipeline peut aussi permettre de réduire l'activité liée aux *glitches* [RW05].

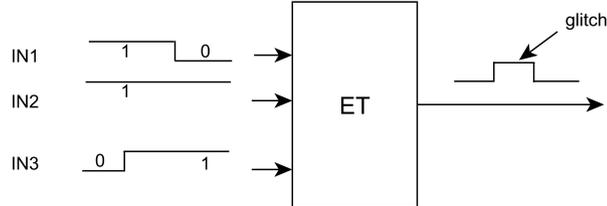


FIGURE 2.9 – Illustration du phénomène de glitch

L'architecture spécifique des FPGAs comprenant un grand nombre d'interconnexions, est particulièrement propice et sensible à ce type phénomène.

D'après la Figure 2.8, l'énergie de charge de la capacité C d'un transistor MOS notée E_c vaut :

$$E_c = \int V_i(t) dt \quad (2.7)$$

avec $i(t) = C \frac{dV}{dt}$.

De plus, si la tension d'alimentation est notée V_{cc} , alors on a :

$$E_c = \frac{1}{2} C V_{cc}^2 \quad (2.8)$$

Dans le cas de circuits utilisant un grand nombre de transistors, les commutations de ces derniers ne sont pas toutes produites en même temps. Le terme α permet de définir le nombre moyen de transitions par cycle d'horloge tandis que le produit αf définit l'activité de commutation.

Dans le cas de charge et décharge de la capacité à la fréquence f , on obtient bien comme puissance moyenne liée aux commutations [KL03] :

$$P_{sw} = E_c f = \alpha C V_{cc}^2 f \quad (2.9)$$

Dans la littérature, il est généralement considéré que la puissance totale dynamique est environ égale à la puissance P_{sw} car très supérieure à P_{cc} . De cette manière, la puissance dynamique totale est alors approximée par la puissance P_{sw} :

$$P_{dynamique} \approx \alpha C_{eff} V_{cc}^2 f \quad (2.10)$$

avec α le nombre moyen de transitions par cycle d'horloge, C_{eff} la capacité effective (F) du circuit, V_{cc} la tension d'alimentation (V) et f la fréquence d'horloge (MHz).

La puissance dynamique est prépondérante dans les circuits de technologie ancienne (> 100 nm), dans lesquels la puissance statique était aussi généralement négligée. Au fur et à mesure des avancées technologiques et de la diminution de la taille des transistors, la puissance statique a considérablement évoluée. On peut alors s'interroger sur les techniques pouvant être mises en œuvre pour réduire ces puissances statique et dynamique.

2.3.3 Quelques techniques de réduction de la consommation énergétique

On peut remarquer d'après l'équation 2.10 que la puissance dynamique évolue linéairement avec la fréquence et de manière quadratique avec la tension. Une première approche pour réduire la puissance dynamique consiste à optimiser la tension d'alimentation.

Le *dynamic voltage and frequency scaling* permet d'ajuster dynamiquement en fonction des besoins de l'application, le niveau de tension d'alimentation ainsi que la fréquence de fonctionnement [GS14]. Généralement, une tension d'alimentation faible permet d'augmenter la fréquence de fonctionnement lorsque l'application requiert de hautes performances. Dans [BZ00], la technique de gestion dynamique de fréquence est mise en œuvre afin de minimiser les pertes d'énergie liées aux horloges dans les FPGAs. Pour cela, un nouveau diviseur d'horloge plus robuste aux *glitches* est présenté. Le *Dynamic voltage scaling* (DVS) est utilisé dans les travaux de [CTL⁺05] dans lesquels la gestion dynamique de tension pour un FPGA Xilinx Virtex 300E est effectuée. Une réduction de consommation de puissance pouvant aller de 4% jusqu'à 54% est montrée.

Une autre technique appelée *Power gating* prévoit d'éteindre l'alimentation des régions du FPGA qui ne sont pas utilisées. Il est alors possible de réduire à la fois la puissance dynamique et statique [HNY14]. D'autres approches de *Power gating* prévoient d'appliquer cette technique à grain plus fin, c-à-d pour chaque LUT [IHK11]. Dans le cas de FPGA-SRAM, si l'alimentation est coupée, l'information est perdue ce qui peut s'avérer problématique et engendrer des contraintes de mémorisation et de contrôle supplémentaires, comme décrit précédemment.

Les FPGAs possèdent généralement un arbre d'horloge contribuant à la dissipation de la puissance dynamique. Le *clock gating* est une technique qui consiste à couper l'horloge d'un certain nombre de ressources logiques qui ne sont pas utilisées par l'intermédiaire d'une porte logique connectée avec l'entrée d'horloge. Le schéma de principe est indiqué sur la Figure 2.10. Sans *clock gating*, la bascule D commute tous les fronts montants d'horloge comme tous les éléments situés en aval de cette bascule, notamment la logique combinatoire. Par l'ajout d'une porte logique sur l'horloge et d'un signal d'activation (*Enable*), des réductions de la consommation de puissance dynamique peuvent être obtenues. Cette technique est désormais employée dans certains FPGAs de Xilinx et permet d'obtenir une réduction de 30% à 80% de la puissance dynamique liée à l'arbre d'horloge, conduisant à une réduction de 30% de la puissance dynamique totale [HKH15].

Une solution de réduction pour s'affranchir de la consommation liée à l'horloge serait l'utilisation de circuits Globalement Asynchrone Localement Synchrones (GALS). Un exemple autour d'un décodeur de Reed-Salomon montrant une réduction d'environ 19% de la puissance dynamique est présenté dans [NSN⁺05].

Un avantage majeur des FPGAs est de pouvoir être reconfigurés dynamiquement au cours du temps. Il a pu être montré dans [LPF10] que l'utilisation de la reconfiguration partielle de certaines zones inactives du design peut permettre de réduire la consommation totale d'un système (statique et dynamique). Néanmoins, il est nécessaire d'effectuer la reconfiguration très rapidement afin d'éviter des surcoûts préjudiciables.

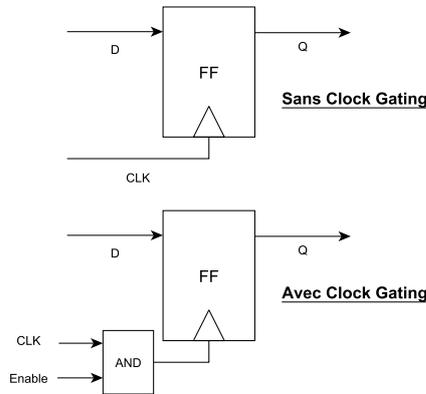


FIGURE 2.10 – Illustration de la technique du *clock gating*

Les principales techniques applicables à des circuits de type FPGA sont référencées dans le tableau 2.2. Ces techniques sont à mettre en œuvre si l'objectif principal est la conception de circuit dit *low power*, à faible consommation. Pour des informations supplémentaires, d'autres techniques de réduction sont référencées dans [LL08].

Techniques de réduction	Références	Puissance Statique	Puissance Dynamique
Circuits avec pipelines	[RW05]	✓	✗
Insertion de délais programmables	[LLW08]	✓	✗
DVS	[CTL ⁺ 05]	✓	✓
DVFS	[GS14]	✓	✓
Power gating	[HNY14, IHK11]	✓	✗
Clock gating	[HKH15]	✗	✓
Circuit asynchrone	[NSN ⁺ 05]	✗	✓
Reconfiguration dynamique	[LPF10]	✓	✓

TABLE 2.2 – Principales techniques de réduction de la consommation pour FPGA

2.4 Problématique de l'estimation de la consommation des circuits FPGA

Dans la Figure 2.11, issue des travaux de [Mor14], sont représentées les évolutions des technologies radio ainsi que des procédés technologiques et des puissances dissipées au cours du temps. On peut constater qu'au fur et à mesure des années, le débit atteignable en voie descendante augmenterait considérablement, passant de moins de 1 Mbps au début des années 1990 à plus de 1 Gbps en 2020. Ceci se traduit aussi par une augmentation de la complexité des systèmes qui possèdent une densité d'intégration croissante.

La puissance dynamique augmente de manière importante avec la miniaturisation. Celle-ci est passée d'environ 0.1W dans les années 90 à plus de 5W en 2015. La puissance statique qui était négligeable sur des technologies anciennes, contribue à la puissance totale dissipée de manière plus importante qu'il y a quelques années. D'autre part, des limites technologiques apparaissent et nécessitent de trouver des solutions pour permettre aux

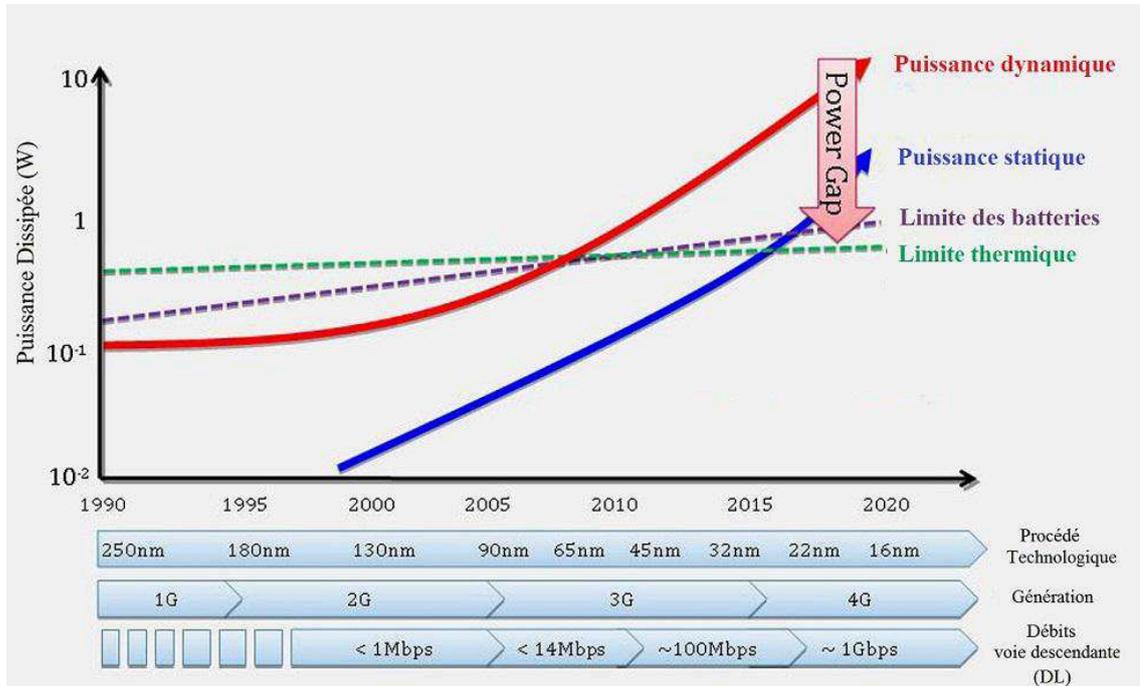


FIGURE 2.11 – Évolution des puissances dynamique et statique en fonction de l'évolution technologique [Mor14]

concepteurs des futurs systèmes de mieux faire face à cette augmentation de puissance dissipée. D'après l'ITRS [fS00], on peut constater que c'est à haut niveau d'abstraction, c'est à dire dès les premières étapes de conception, que les techniques de réduction de consommation devront être appliquées.

2.5 Conclusion

Dans ce chapitre, une description des circuits FPGA ainsi que les origines de la consommation dans ces circuits ont été présentées. Les principaux facteurs ayant un impact sur les puissances dissipées ont été décrits. Il a été montré que la puissance statique reste très fortement dépendante de la technologie choisie par le concepteur du système. Plusieurs stratégies d'optimisation ont été présentées, ces dernières pouvant être appliquées afin de réduire les courants de fuites. Néanmoins, ce sont les aspects dynamiques qui semblent être les plus importants dans la mesure où ils peuvent être directement optimisés au niveau architectural. Cette optimisation peut être obtenue en comparant efficacement plusieurs configurations de systèmes au sein d'un FPGA et en retenant la meilleure solution.

Néanmoins, évaluer la puissance dynamique sur de nombreux systèmes, implantés sur une cible FPGA donnée (où la puissance statique est relativement constante) pose les questions suivantes :

1. Comment estimer rapidement et précisément la consommation de systèmes implantés

sur FPGA ?

2. Comment estimer rapidement et de manière précise la consommation dynamique d'énergie des traitements en bande de base à haut-niveau tout en faisant face à la complexité croissante ?
3. Comment effectuer une exploration rapide de l'espace de conception, à partir des informations de consommation et de performances ?

Chapitre 3

Techniques d'estimation de la consommation de systèmes sur FPGA

Pour les concepteurs de futurs systèmes, être capables d'estimer la consommation d'un circuit au plus tôt durant son développement est crucial afin d'éviter des coûts importants liés aux re-développements notamment si le budget énergétique n'est pas respecté.

Dans les chapitres précédents, on a pu constater qu'un système peut être représenté en fonction d'un niveau d'abstraction. Dans la littérature, les principaux niveaux d'abstraction vont du niveau technologique, qui correspond au niveau le plus bas, jusqu'au niveau système qui correspond au niveau d'abstraction le plus haut, comme indiqué à la Figure 3.1. Ainsi, des techniques et outils d'estimation de la consommation ont été développés par des académiques et des industriels à ces différents niveaux.

De nos jours, les concepteurs de systèmes à base de FPGA suivent une approche partant des niveaux d'abstraction haut vers les niveaux bas. Au niveau système, la fonctionnalité du système est modélisée à l'aide de langages comme le C/C++ ou encore à l'aide d'outils tels que Matlab. La fonctionnalité du système peut être ainsi validée avec des outils de simulation haut-niveau comme Simulink par exemple.

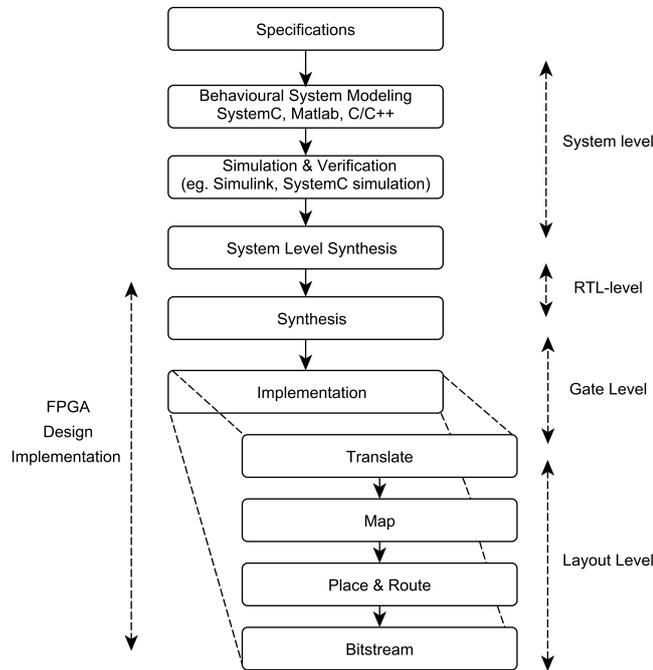


FIGURE 3.1 – Représentation des niveaux d’abstraction

De nombreuses approches et outils de modélisation existent et sont adoptés dans différentes industries et dans le monde académique. Pour palier cette diversité et donc améliorer l’interopérabilité des modèles entre eux, l’organisme *Open SystemC Initiative (OSCI)* a raffiné et standardisé le concept de modélisation au niveau transactionnel [Mon07]. L’objectif de ce standard est de permettre de résoudre de nombreux problèmes de la conception haut-niveau. D’après [Mon07], la nouvelle version du standard TLM 2.0 de 2007 doit :

1. faciliter le développement logiciel,
2. faciliter l’exploration d’architecture,
3. permettre la vérification et les tests fonctionnels du matériel,
4. faciliter l’intégration et le partitionnement logiciel/matériel,
5. uniformiser les modèles conçus (interopérabilité des modèles).

Dans ce standard, différents styles de codage des modèles ont été proposés, permettant un raffinement et une précision de la description :

- *Loosely-timed models (LT)*
- *Approximately-timed models (AT)*
- *Cycle-Accurate (CA)*

Ces différents styles de codage doivent orienter le concepteur sur la façon de décrire son modèle à travers la précision de la dépendance des données vis-à-vis du temps. Une modélisation avec un style de codage LT sera moins précise qu’un modèle décrit avec un style de codage AT et encore moins qu’un modèle décrit au cycle près. Dans le cas de modèles LT, ces modèles sont capables de fournir des données dès lors qu’une transaction

entre deux modèles est initiée. Ces modèles ne dépendent pas de l'avancée du temps de simulation pour produire une réponse. Dans le cas de modèles AT, ces modèles dépendent à la fois d'événements internes ou externes et de l'avancée du temps pour fournir une réponse. On peut noter que ce type de modélisation requiert plus de synchronisation, conduisant à des temps de simulations plus longs [Mon07].

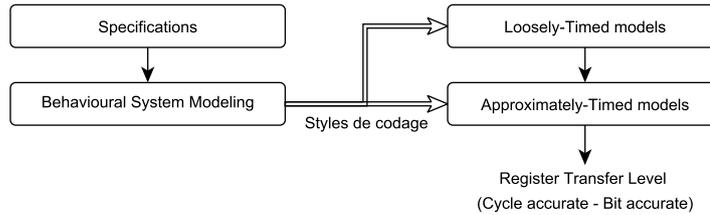


FIGURE 3.2 – Illustration du Transaction Level Modelling

A partir des différents styles de codage, le système modélisé peut ainsi être testé, validé et raffiné. Le passage du niveau système au niveau RTL (*Register Transfer Level* (*High-Level Synthesis* (HLS))). Ces outils permettent le passage automatique de langages de modélisation en une description de niveau RTL. En effet, il est très difficile et coûteux en ressources (temps, main d'œuvre) de passer directement d'un langage de modélisation à un code RTL synthétisable. Ces dernières années, de nombreuses recherches ont conduit au développement d'outils de synthèse de haut niveau pour tenter de palier ce problème [Ca15] dont certains sont énumérés dans [CLN⁺11, GS09]. Ces outils prennent généralement en entrée des langages de modélisation basés sur le C++ comme SystemC, HardwareC, SpecC ou encore Handel-C. Certains outils sont aussi disponibles en open-source comme LegUp [CCA⁺11].

Une fois le système décrit au niveau RTL, le flot de conception traditionnel débute avec l'étape de synthèse logique qui utilise un langage de description matérielle pour générer une netlist compatible avec les outils d'implémentation des constructeurs de FPGA. On peut noter que ces niveaux sont communs à toutes les implémentations matérielles et sont adaptés pour le développement de systèmes de communications numériques.

Les sections suivantes identifient les principales méthodologies, outils et techniques permettant d'estimer la consommation énergétique (puissances statique et/ou dynamique) avec un focus sur celles utilisées dans le domaine des communications numériques et ayant pour cible un FPGA. Ces techniques sont classées en partant des niveaux d'abstraction les plus bas jusqu'aux niveaux hauts.

3.1 Techniques d'estimation de la consommation

3.1.1 Niveau technologique

Le niveau technologique (*Layout* ou *niveau transistor*) est le niveau dans lequel le système est entièrement défini. Les concepteurs possèdent alors le plus d'informations possibles sur le système. Ainsi, chaque transistor est caractérisé par rapport à une cible

matérielle définie et une technologie donnée. Une estimation très précise de la puissance dissipée peut être obtenue, soit par l'intermédiaire de mesures effectuées directement sur le circuit, soit grâce à des simulations. En revanche, les temps de simulation sont très longs à cause de la quantité de détails nécessaire à la description du système.

A ce niveau d'abstraction, des outils comme SPICE [Nag75] développés par l'Université de Berkeley font généralement office de référence grâce à leur très bonne précision. Cet outil permet de simuler les performances électriques de circuits électroniques, notamment ceux basés sur la technologie CMOS à l'aide de modèles de composants, par exemple, résistances, capacités, etc. PowerMill [HZDS95] est un autre outil permettant d'obtenir des estimations de consommation au niveau transistor, qui requiert un temps de simulation 100 à 1000 fois plus court que SPICE [Inc00], à l'aide d'une simulation événementielle et grâce à l'utilisation de tables de correspondance. IRSIM [SH89] est un autre exemple d'outil de simulation au niveau transistor.

Néanmoins, estimer la consommation à partir de ces outils est très coûteux en temps de simulation, notamment à cause de la complexité croissante des systèmes et le volume de données à traiter. Les derniers FPGAs possèdent des milliards de transistors et il est donc nécessaire de réduire le niveau de détails en élevant le niveau d'abstraction au détriment d'une plus faible précision.

3.1.2 Niveau Porte logique

Au niveau Porte Logique, (*Gate-Level*), le système est le résultat de l'étape de synthèse. Il est décrit sous la forme d'un ensemble d'éléments logiques et d'interconnexions. Ainsi, le circuit est simplifié par rapport à une description au niveau transistor. Dans le cas où des mesures réelles ne sont pas facilement réalisables, il est nécessaire d'estimer la consommation du circuit. Deux catégories de techniques sont alors citées dans la littérature afin d'estimer la puissance dissipée par un circuit :

- les approches basées sur les simulations, aussi appelées méthodes statistiques,
- les approches basées sur les probabilités.

Les méthodes d'estimation basées sur les simulations consistent à simuler un circuit un grand nombre de fois afin d'obtenir des informations, soit sur l'activité du circuit, soit sur le courant utilisé. Pour cela, des stimuli aléatoires sont fournis par l'utilisateur aux entrées du circuit. La précision des résultats est grandement influencée par la représentativité des stimuli appliqués aux entrées [CKCW06, NX98]. En outre, évaluer la puissance dissipée d'un circuit pour l'ensemble des combinaisons possibles d'entrées peut s'avérer très compliqué à mettre en œuvre, voire impossible. Néanmoins, des simulations logiques de type événementiel peuvent être effectuées afin de réduire les temps de simulation.

Dans le cas de circuits combinatoires, une puissance moyenne peut être obtenue facilement sans tenir compte des délais de propagation. Ce type de circuit ne possède pas de boucle de rétroaction et est donc moins difficile à analyser que des circuits séquentiels. Dans ce cas, des simulations dites de "Monte Carlo" peuvent être utilisées afin d'éviter un nombre trop exhaustif de simulations. En effet, l'objectif recherché est de minimiser l'impact de la nature des stimuli d'entrée en effectuant un nombre important de simula-

tions. Des stimuli aléatoires sont appliqués aux entrées tandis que la puissance dynamique est mesurée par unité de temps jusqu'à ce qu'une erreur d'estimation, relative ou absolue, atteigne une précision donnée [BNYT93, KN97, TBAV05]. Dans [KN97], l'approche proposée permet de déterminer les bornes supérieures et inférieures de la consommation avec une précision renseignée par l'utilisateur.

L'autre type de méthode d'estimation est basée sur les probabilités de transition. L'avantage majeur de ce type d'approche est qu'une seule étape de calcul est requise. En effet, l'utilisateur fournit les probabilités de transition des entrées qui sont ensuite propagées jusqu'aux sorties du circuit de manière à déterminer l'activité interne de ce dernier. C'est une approche très efficace pour déterminer la puissance dissipée et qui ne nécessite pas une connaissance approfondie des stimuli d'entrées. Dans la plupart des approches, ce sont des probabilités moyennes de transitions qui sont appliquées aux entrées tandis qu'avec les approches basées sur les simulations, l'utilisateur doit fournir les caractéristiques les plus précises possibles sur les stimuli d'entrées. Fournir des probabilités moyennes de transition sur des signaux d'entrée est généralement suffisant pour modéliser le comportement des entrées [CKCW06]. Les techniques basées sur les probabilités de transition sont alors qualifiées de faiblement dépendantes des stimuli d'entrée (*weakly pattern-dependent*) contrairement aux approches basées sur les simulations [NaJ94]. Néanmoins, la précision des estimations est moins bonne qu'avec les simulations à cause de l'approximation en moyenne faite sur les probabilités de transition des entrées. De plus, ce type d'approche ne permet pas forcément de tenir compte des *glitches* et de leurs propagations dans le circuit.

L'ensemble des deux types d'approche a été largement étudié et exploité dans la littérature afin d'améliorer la précision, la rapidité des estimations, à partir des informations à disposition.

L'avantage des techniques probabilistes est d'effectuer une seule propagation des probabilités de transition afin de déterminer l'activité du circuit. Or, un certain nombre de paramètres est négligé durant l'estimation de l'activité d'un circuit [LW06a, AN04] conduisant à des estimations plus ou moins précises :

1. les temps de propagation des portes logiques ainsi que ceux des interconnexions ne sont pas considérés, *zero delay activity*,
2. les temps de propagation des portes logiques sont considérés mais pas ceux des interconnexions, *logic delay activity*,
3. les temps de propagation des portes logiques et des interconnexions sont pris en compte, *routed delay activity*.

Afin d'obtenir des estimations précises, l'ensemble des temps de propagation doivent être considérés de manière à tenir compte des *glitches*, qui contribuent de manière significative à la dissipation de puissance dynamique. De plus, il est d'autant plus important de les considérer au regard de l'architecture spécifique des FPGAs c-à-d avec un grand nombre d'interconnexions. Il est aussi nécessaire de prendre en compte les corrélations temporelles et spatiales entre les entrées [LW06a].

La corrélation spatiale [HAE09] signifie que la valeur logique d'un signal à un instant t , dépend d'un autre signal tandis que la corrélation temporelle est présente lorsque la valeur logique d'un signal à l'instant t dépend de sa valeur à l'instant $t - 1$.

A partir des observations précédentes, un outil appelé ACE-2.0 a été proposé [LW06b]. Son objectif est d'estimer automatiquement l'activité en considérant les délais de propagation des portes et des interconnexions ainsi que les corrélations spatiales et temporelles. Dans les différents travaux cités, ces considérations n'étaient pas forcément prises en compte, tout comme l'activité générée par les *glitches*. L'activité d'un circuit peut aussi être reliée à la puissance moyenne dissipée à travers l'utilisation de techniques comme les densités de transition, *transition density* [CKCW06] ou encore à travers l'utilisation de chaînes de Markov [MMP94].

Contrairement aux approches probabilistes, des simulations sont effectuées dans [SKB02]. Tout d'abord, les capacités de chaque ressource d'un FPGA Virtex-2 ont été identifiées à l'aide de simulations au niveau transistor avec l'outil SPICE. En partant du flot de conception de Xilinx, des programmes en Perl permettent de récupérer des informations sur l'utilisation des ressources à partir d'une netlist obtenue après placement routage (NCD). Ensuite, ce fichier est annoté, avec notamment les retards de propagation pour chaque ressource. Une simulation temporelle après placement et routage est ensuite effectuée sous Modelsim afin de déterminer l'activité de chaque ressource. Au final, la puissance dynamique est estimée par l'équation 3.1 :

$$P = 0.5 \times V^2 \times f \times \sum C_i U_i S_i \quad (3.1)$$

avec V la tension d'alimentation, f la fréquence de fonctionnement. $C_i U_i S_i$ sont la capacité effective, l'utilisation et l'activité de commutation de chaque ressource respectivement. Une méthodologie similaire a été développée dans [DT05].

Dans le monde industriel, des outils ont aussi été développés et permettent à l'utilisateur d'effectuer les deux types d'approches. A titre d'exemple, Xilinx [Xil12c] a développé XPower Analyzer (XPA). Cet outil permet d'estimer la consommation moyenne de puissance d'un circuit implanté sur FPGA. Sur la Figure 3.3, son schéma d'utilisation permettant d'obtenir des estimations précises de consommation est donné. Ainsi, XPA peut délivrer des estimations de consommation lors de deux cas :

1. lorsqu'un fichier d'activité lui est fourni,
2. à partir d'un algorithme estimation vectorielle.

Dans le premier cas, l'utilisateur peut fournir un fichier d'activité (.saif ou vcd) issu d'une simulation, contenant l'ensemble de l'activité interne. Pour cela, une simulation, généralement effectuée après placement et routage du design, est effectuée. Dans le second cas, aucun fichier d'activité n'est spécifié en entrée d'XPA. Dans ce cas, des probabilités de transition des signaux d'entrée, définis par défaut ou bien paramétrables par l'utilisateur, sont utilisés. Ensuite, un algorithme d'estimation vectorielle propage les probabilités de transition en partant des entrées vers les sorties. L'outil PowerPlay Analyzer [Alt14] développé par Altera, fonctionne sur le même principe qu'XPower Analyzer

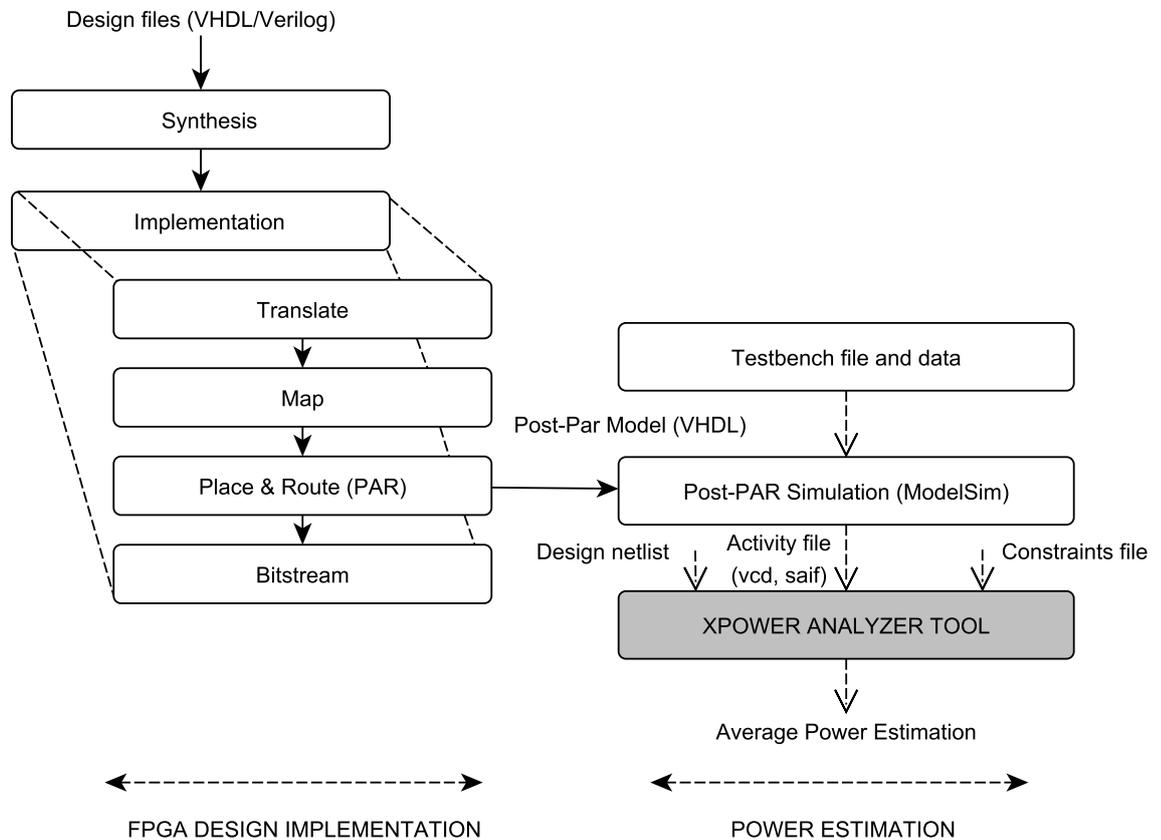


FIGURE 3.3 – Principe d'utilisation de l'outil XPower Analyzer

mais est dédié pour leur cible FPGA. Ainsi, l'utilisation de ce type d'outil nécessite la réalisation du flot de conception du FPGA ainsi que des simulations temporelles qui seront d'autant plus longues que le système est complexe. De plus, pour les deux types d'approche, le nombre d'information sur l'implantation physique du circuit réduit considérablement les temps d'estimation.

De manière à accélérer l'estimation de la consommation d'un circuit, d'autres solutions ont été introduites à un niveau d'abstraction supérieur, le niveau RTL.

3.1.3 Niveau Register Transfer Level (RTL)

Au niveau RTL, le système est généralement décrit à l'aide de langages de description matérielle comme le VHDL ou le Verilog. Le système est alors représenté par l'intermédiaire de signaux et d'échanges de données entre différents éléments. Ces éléments sont aussi appelés macros et peuvent faire références à des blocs de plus ou moins grande complexité tels que des blocs multiplieurs, additionneurs ou encore des blocs IPs (*Intellectual Property*).

A ce niveau d'abstraction, la puissance dissipée d'un système est généralement évaluée à l'aide de modèles de consommation développés pour chaque sous-élément du système, c-à-d des macros. Des techniques, initialement développées pour des ASICs, dites de *macro-modelling* sont utilisées de manière à exprimer la puissance dissipée en fonction de paramètres statistiques tels que la probabilité de transition etc. Les techniques de *macro-modelling* couramment utilisées dans la littérature, reposent généralement sur l'utilisation de tables (de valeurs ou d'équations) ou bien sur l'utilisation de techniques de régression linéaire. Des simulations peuvent être aussi requises lors d'une phase de caractérisation pour le développement des modèles.

Dans les travaux de [DR⁺06, DA14], des valeurs de puissances dissipées de blocs IP sont stockées dans des tables (LUT) en fonction de paramètres statistiques des entrées. Les paramètres considérés pour la réalisation des macro-modèles sont les corrélations temporelles et spatiales des entrées et des sorties ainsi que la densité de transition et les probabilités moyennes de transition des entrées et des sorties. Des simulations de "Monte-Carlo" sont utilisées, sans tenir compte des délais de propagation afin de construire les tables. Leur approche leur permet d'atteindre une précision entre 9% et 15% pour des systèmes basés sur plusieurs blocs IPs préalablement caractérisés. Un autre exemple d'illustration de macro-modelling est décrit dans [LP05]

Dans [RSN06], une approche d'estimation développée pour des ASICs est adaptée pour les FPGAs. Différentes métriques permettant l'analyse des transitions comme la distance de Hamming (pourcentage de bits qui changent d'état pendant une transition), la distance signal et la 0 distance (pourcentage de bits restant à l'état haut et à l'état bas respectivement par transition) et la quantification des données sont utilisées pour la création de modèles. Leur approche utilise plusieurs outils comme XPower, MacAnova pour les régressions ainsi que l'outil ORINOCO [SKN⁺01] qui permet d'effectuer une synthèse de haut niveau de descriptions en C/C++ ou SystemC.

On peut remarquer que les approches basées sur le développement de tables (LUT) requièrent un grand nombre de tests de manière à couvrir le plus largement possible les combinaisons d'entrées. De plus, la taille des tables peut devenir considérable tout comme le temps de simulation lorsque le nombre de paramètres augmente.

Pour palier cela, des modèles analytiques peuvent être développés afin d'exprimer facilement la puissance dissipée en fonction de paramètres d'intérêt [NMH12]. Dans [SJK⁺11], des modèles liant l'activité à la puissance dissipée et à d'autres paramètres tels que le nombre de ressources utilisées et la quantification des données, ont été développés à l'aide de techniques de régression. Des probabilités de transition sont ensuite générées et propagées à travers le circuit à l'aide d'un outil dédié. De cette manière, l'approche suivie par les auteurs délivre une précision moyenne d'environ 8% pour la puissance totale et 17% pour la puissance dynamique pour différentes familles de FPGAs en comparaison à une estimation au niveau porte. Le développement de modèles analytiques réduit la précision de l'estimation mais offre néanmoins un gain en flexibilité.

D'autres approches visent à développer des modèles analytiques autour des principaux

éléments matériels disponibles dans un FPGA comme la logique, les blocs multiplieurs DSPs, etc. afin de pouvoir estimer la consommation ainsi que la surface utilisée par les blocs IPs [DSZC11]. Dans l'approche proposée, XPower Analyzer est utilisé pour estimer la consommation à partir d'un circuit après placement et routage.

Dans le cas de création de modèles par régression, on peut donc constater qu'une importante phase de caractérisation pour élaborer les modèles est nécessaire. D'autre part, une des limitations majeures reste l'impossibilité à estimer une IP qui ne ferait pas partie de la librairie. Pour cela, des outils comme RTLEst [BCMP01], PrEsto [SWPC10] permettent d'automatiser l'estimation de la puissance dissipée de systèmes spécifiques. PrEsto permet la génération automatique de modèles analytiques afin d'estimer la consommation de systèmes d'exploitation sur des processeurs embarqués. Concernant RTLEst, l'avantage de cette approche est de pouvoir générer automatiquement des modèles analytiques obtenus par régressions linéaires en combinaison avec des tables de valeurs (LUTs).

Au niveau RTL, il a pu être noté que le développement de modèles repose généralement sur une phase importante de caractérisation qui peut être effectuée à des niveaux plus bas. D'autre part, le développement de modèles analytiques réduit la précision de l'estimation mais offre un gain en flexibilité.

3.1.4 Mesures réelles et émulation

Différents travaux de la littérature visent à mesurer directement sur plate-forme, le courant aux bornes du système pour en déduire une relation avec la puissance consommée [HASLVE08, MGP11]. Ainsi, leurs estimations sont précises mais cela nécessite un temps de développement conséquent. En effet, il est nécessaire de réaliser l'implantation réelle du système ainsi que le développement du procédé d'acquisition avant de pouvoir effectuer les mesures. Néanmoins, ces mesures peuvent être analysées dans l'objectif de concevoir un système à basse consommation ou encore pour servir à la création de modèles à plus haut niveau. Par contre, une fois le système développé, les gains en termes de précision sont alors très intéressants sur la consommation ainsi que les temps de mesure en comparaison des approches d'estimation utilisant de longues simulations temporelles (après placement et routage).

Les mesures réelles peuvent aussi permettre d'évaluer et quantifier l'impact des options d'optimisation qui peuvent être choisies durant le flot de conception. Dans [MGP11], leur approche consiste à évaluer l'impact des choix des paramètres de conception en phase de synthèse et d'implémentation, pour des applications de sécurité/cryptographie. Les mesures obtenues sont ensuite comparées aux prédictions faites par XPower Analyzer de Xilinx. Néanmoins, les tests effectués ne précisent pas si une simulation temporelle est effectuée pour analyser l'activité interne ou bien si les taux d'activité par défaut sont utilisés. Le cas échéant, il paraît logique d'obtenir pour certains circuits des différences relativement importantes au niveau des estimations.

Une autre méthode envisagée dans [AVB14] est d'évaluer la consommation de systèmes à base de processeurs embarqués sur FPGA de type MicroBlaze à l'aide de l'outil *ChipScope Analyzer* de Xilinx. La méthode consiste à relever la consommation à l'aide *ChipScope Analyzer*, en faisant varier le nombre d'unités de traitement ainsi que les applications logicielles développées sur le MicroBlaze afin d'estimer la consommation d'une seule unité

de traitement.

D'autres approches tirent bénéfice des mesures réelles pour développer des modèles précis. Ainsi dans [EJH⁺04], les consommations des puissances statique et dynamique des mémoires d'un FPGA de type Virtex-400E de Xilinx sont modélisées. Après avoir sélectionné les paramètres haut-niveau (algorithmiques et architecturaux), une série de mesures est effectuée et les modèles sont réalisés par régression linéaire. En comparaison à XPower, la précision du modèle est d'environ 8% avec des erreurs maximales de 22%.

Dans le cadre d'un projet, le développement d'une plate-forme libre et accessible à distance, appelée Open-People [SCZ⁺12], permet aux utilisateurs d'estimer la consommation de puissance de leurs systèmes embarqués (OS, FPGA, DSP, MPSoC...) et d'explorer de nouvelles architectures et techniques d'optimisation afin de réduire la consommation de puissance et d'énergie. Des modèles de consommation peuvent être créés suivant différents niveaux d'abstraction à partir d'un banc de mesure automatisé. Ce type de projet est adapté pour les concepteurs qui visent à optimiser la consommation de leurs systèmes hétérogènes en décidant d'effectuer une tâche particulière sur FPGA ou bien sur un processeur.

Dans les cas où des mesures réelles sont effectuées, le banc de mesure doit être robuste afin d'acquérir un nombre suffisant de mesures sans surcoût important de temps tout en gardant une bonne précision. Ce type d'approche peut néanmoins s'avérer relativement peu flexible dans le cas de changement de configuration ou de blocs IP dans un système complet.

Une autre technique utilisant les propriétés des FPGAs s'avère être de plus en plus utilisée. L'émulation permet d'accélérer l'estimation de la consommation d'un système en exécutant directement une application sur le FPGA [HBB14]. En effet, le principe réside dans l'implémentation matérielle du modèle analytique d'estimation. Ainsi, la puissance dissipée est estimée à partir de la vitesse de traitement du FPGA.

Dans le cas de circuits complexes, un grand nombre de modèles est nécessaire. Pour cela, des techniques de corrélation et de *clustering*, permettant entre autres le partage de ressources, peuvent être mises en œuvre pour réduire le nombre de ressources matérielles utilisées [CRR05]. La mise en œuvre de ces techniques peut conduire à des accélérations allant d'un facteur x10 à x500 en comparaison avec d'autres outils d'estimation. Néanmoins, malgré des facteurs d'accélération relativement importants, ce type d'approche requiert à la fois le développement des modèles analytiques en combinaison avec l'implémentation du circuit de test.

3.1.5 Niveau Système

Au niveau Système est généralement modélisé ce que le système doit faire, c-à-d sa fonctionnalité. Celle-ci peut être décrite sous forme de spécifications ou encore d'algorithmes, eux-mêmes représentés de façon spécifique comme, par exemple, sous forme de tâches (matérielles et/ou logicielles), sous forme graphique ou encore schématique. La fonctionnalité du circuit peut aussi être représentée sous la forme d'un certain nombre d'instructions exécutées séquentiellement.

Dans la littérature, le terme d'*Electronic System Level* (ESL) peut être employé de

différentes manières et prêter à confusion en qualifiant à tort un niveau d'abstraction. D'après [GBP10], ESL est un concept qui regroupe l'ensemble des méthodologies et des outils facilitant la conception mais aussi la vérification et le test de nombreux systèmes (hardware/software) sur différentes plate-formes (SoC, MPSoC, FPGA). Par analogie avec la Figure 3.1, ESL regrouperait les techniques et outils permettant la conception d'un système en partant des spécifications jusqu'à l'étape de synthèse, incluant les outils de synthèse de haut-niveau (*High-Level Synthesis*, HLS) ainsi que les techniques de modélisation TLM. ESL prendrait donc fin quand l'implémentation commence.

C'est à ce niveau d'abstraction que les choix effectués par les concepteurs des futurs systèmes ont un impact maximum. L'exploration de l'espace de conception est facilitée à ce niveau malgré une précision sur la consommation estimée, dépendante des modèles utilisés. De plus, ce type d'information n'est généralement pas disponible à ce niveau.

Dans le domaine des communications numériques sans fil, les performances théoriques d'un système sont généralement évaluées au niveau système. Les concepteurs commencent par décrire la fonctionnalité du système après avoir analysé les spécifications.

D'une manière générale, lorsque les concepteurs d'un système de communications numériques sans fil désirent estimer la consommation de leur circuit, des modèles analytiques très généraux peuvent être utilisés. Dans les travaux [DDG⁺12], [ARFB10] ou encore [DDVJM12], des modèles analytiques ou des valeurs de référence issues de la littérature ont été associés à chaque sous-élément d'un système de communications sans fil, incluant les traitements en bande de base, les étages RF, l'amplificateur de puissance, les micro-processeurs etc. La puissance consommée par le système entier est alors obtenue à l'aide d'outils de simulation (respectivement Matlab de MathWorks [Mat15] et un outil de déploiement de réseau).

Dans l'ensemble de ces travaux, la précision de l'estimation de la consommation n'est pas l'objectif premier. En utilisant de telles valeurs, l'objectif est de dégager des tendances de manière à prendre rapidement une décision. Néanmoins, comparer précisément différents systèmes entre eux en termes de consommation et de performances peut s'avérer difficile, sans aucune information sur leur implantation réelle.

Plusieurs outils et langages comme Matlab sont couramment utilisés dans la communauté. Ce langage permet de modéliser et tester la fonctionnalité d'un système décrit dans un langage commun (Matlab). Néanmoins, un ou plusieurs outils supplémentaires sont nécessaires afin de concevoir le circuit réel sur FPGA.

Des outils comme System Generator [Xil12a] et DSP builder [Alt15a], respectivement de Xilinx et Altera, sont des actuellement intégrés dans l'environnement Matlab. Ils permettent la génération de code HDL à partir d'une description graphique sous Simulink (utilitaire de Matlab). Le code HDL généré peut alors être traduit en une netlist puis en bitstream par l'intermédiaire de la réalisation du flot classique de conception. La consommation du système peut alors être soit mesurée à l'aide d'un prototype, soit estimée à l'aide d'outil comme XPower Analyzer pour des cibles FPGA de Xilinx. Néanmoins, l'exploration de l'espace de conception est limitée. En effet, l'ensemble du flot de conception

doit être effectué à chaque changement de configuration du système, ce qui peut s'avérer long et très coûteux en temps dans le cas de système complexe.

Pour palier le manque de détails sur l'implantation matérielle, un outil nommé PyGen, basé sur l'utilisation de Matlab/Simulink, est présenté dans [JP04]. L'outil propose de ramener à haut niveau sous Matlab, des informations de consommation et de performances (latence/ressources utilisées) de blocs IP qui ont été caractérisés à bas niveau. Chaque bloc IP du système est ainsi synthétisé à l'aide de System Generator pour l'ensemble des paramètres possibles. L'estimation de la puissance totale du système est de cette manière rendue possible sous Matlab/Simulink. Néanmoins, l'approche proposée ne permet pas de tenir compte des caractères dynamiques d'une application où chaque bloc IP peut être actif ou non. En fait, les auteurs supposent que tous les modules sont actifs en même temps, ce qui n'est pas forcément le cas dans un système réel et peut conduire à de grandes erreurs d'estimation. On peut aussi noter que d'autres outils, en plus de System Generator, peuvent être utilisés pour traduire des modèles au niveau système en code synthétisable e.g. SPW de Synopsys [Syn15] ou encore DSP designer [ACE⁺12].

Dans le tableau 3.1, quelques outils, environnements de conception, de simulation, de modélisation et de prototypage qui sont utilisés dans le domaine du traitement du signal et des communications numériques sont énumérés (liste non-exhaustive).

Bien que très performants et utiles, ces outils ne permettent pas, par défaut, aux concepteurs, d'obtenir une information rapide et précise sur la consommation de leurs algorithmes. En effet, aucune étape d'implémentation n'étant réalisée, il semble évident de ne pas pouvoir obtenir d'estimations précises.

Comme relaté dans le tableau 3.1, les outils permettent le plus souvent d'effectuer du prototypage et de l'exploration d'architectures à l'aide de modèles du système. Ainsi, ils permettent un raffinement du système des niveaux haut vers les niveaux les plus bas. Néanmoins, estimer de manière précise la consommation du système est seulement rendue possible si l'utilisateur possède déjà ses propres modèles de consommation.

Comme Matlab/Simulink ou encore Labview [Nat15], les langages de programmation tels que le C/C++ sont très utilisés pour modéliser et déterminer les performances théoriques d'un système. A travers leur utilisation, les concepteurs décrivent le système en connectant des blocs de base qui échangent des données de manière séquentielle.

Dans le cadre de la modélisation TLM, plusieurs langages peuvent donc être utilisés. Néanmoins, SystemC [Ini15] est un des langages les plus utilisés dans la littérature, notamment grâce de la standardisation du TLM 2.0 [Mon07]. SystemC n'est pas un langage de modélisation mais se réfère à une bibliothèque C++ possédant son propre noyau de simulation. La modélisation en SystemC a un grand avantage. Chaque objet communique avec d'autres à travers un canal. La fonctionnalité du système est alors séparée du procédé de communication. Cette classe permet aussi aux concepteurs de modéliser conjointement les parties logicielles et matérielles à travers un langage commun. Plusieurs degrés de raffinement sont possibles, permettant d'inclure ou non des détails spécifiques liés à l'implémentation. Les modèles développés peuvent être des modèles fonctionnels ou

TABLE 3.1 – Exemples d’outils utilisés au niveau système dans le domaine du traitement du signal

Nom	Académique ou Commercial (A/C)	Intègre l’estimation de puissance (O/N)	Description
Matlab/Simulink [Mat15]	C	N	Outil de modélisation et simulation graphique
System Development Suite [Cad11]	C	N	Permet le développement logiciel/matériel, le prototypage virtuel à partir de modèles TLM (<i>Transaction-Level Models</i>), vérification fonctionnelle et optimisation système avant les étapes d’implémentation
SystemVue [Key15]	C	N	Environnement de conception assistée qui permet la simulation et le test automatique de systèmes de communications sans fil (bande de base, RF)
LabVIEW FPGA [Nat15]	C	N	Basé sur la plate-forme de développement graphique LabVIEW, il permet la simulation, vérification et déploiement de système sur FPGA
Ptolemy [Dep15]	A	N	Framework logiciel, disponible en open-source, permettant la conception, modélisation et la simulation autour d’acteurs (composants logiciels qui exécutent simultanément et communiquent par messages envoyés via les ports interconnectés)
System Studio [Syn15]	C	N	Outil de conception et d’analyse d’algorithme de traitement du signal, basé sur les modèles flot de données en C.
Vista [Sha15]	C	N	Outil d’exploration d’architecture, de vérification et de prototypage virtuel de SoC basé sur le kit TLM 2.0.

bien des modèles permettant simplement l'évaluation de performances. Dans le dernier cas, une information de contrôle est généralement requise afin d'assurer de bons timings. Des simulations précises au cycle près peuvent aussi être effectuées sous SystemC.

Néanmoins, SystemC ne permet pas d'estimer la consommation du système modélisé. Pour cela, une information supplémentaire a besoin d'y être intégrée. Partant de ce principe, des techniques ont été proposées dans la littérature de manière à pouvoir permettre l'estimation du système modélisé.

Dans [BJ03], une technique de modélisation de la consommation de SoC permet d'associer aux états de blocs IP ou de processeurs, des valeurs de puissance dissipée par l'intermédiaire de *power state machines*. L'évolution des états est définie par l'intermédiaire de diagrammes de décision binaires. A travers une simulation symbolique, la consommation totale du système peut être évaluée ainsi que les bornes supérieures et inférieures de la consommation. La précision de l'approche dépend des paramètres saisis par l'utilisateur comme l'activité de transition, la tension d'alimentation etc. L'approche des auteurs n'est pas dédiée aux FPGAs et la précision de l'estimation est conditionnée par la connaissance du concepteur de chaque élément du système. De plus, les paramètres modifiables sont fortement technologiques (activité, tension d'alimentation). Pour notre étude, des paramètres liés au domaine d'application doivent être considérés.

Un autre outil utilisant SystemC et permettant l'estimation de la consommation de SoC, constitué de blocs matériels est présenté dans [HLF⁺11]. L'approche des auteurs est basée sur le développement de modèles analytiques construits à partir de régressions linéaires sur des valeurs obtenues après simulation au niveau transistor. Ils introduisent l'utilisation de signaux clés afin de permettre la reconnaissance des modes de chaque IP. Pour la modélisation des processeurs, la technique ILPA pour *Instruction-Level Power Analysis* a été utilisée. Cette technique consiste typiquement à caractériser chaque instruction exécutée sur le processeur par des mesures de consommation. Après les étapes de modélisation, l'outil appelé Power Depot, correspondant à une classe C++ ajoutée à SystemC, permet la transformation des modèles de consommation en un module décrit en SystemC. A l'issue d'une simulation sous SystemC, la puissance totale consommée par le système peut être évaluée. La précision de l'approche permet d'obtenir moins de 2% d'erreur, grâce à l'utilisation de simulations au niveau transistor.

Dans [SRH⁺11], SystemC est utilisé pour modéliser une application ciblant des systèmes sur puce hétérogènes à multiples processeurs (*Multi-Processor System-on-Chip, (MPSoC)*). La consommation est évaluée à l'aide de modèles développés à partir de la méthodologie *Functional-Level Power Analysis* (FLPA). Cette méthodologie a été précédemment utilisée pour modéliser la consommation de micro-processeurs [LJSM04] ainsi que d'IP [EJH06]. Les modèles sont construits après une analyse fonctionnelle permettant d'identifier les paramètres ayant une influence sur la consommation. Dans une approche hybride [RBA⁺11], les estimations peuvent délivrer une erreur inférieure à 1.2% en comparaison avec des mesures réelles (modèles aussi construits à partir de mesures réelles). Dans [AFJS07], une extension de la méthodologie FLPA est présentée. En effet, les modèles de consommation supportent plusieurs niveaux d'abstraction afin d'estimer la consomma-

tion du système tout au long du flot de conception FPGA. Les modèles sont ainsi raffinés après chaque étape.

Une autre approche de modélisation analytique, appelée *Functional Level Power Analysis and Modelling (FLPAM)*, a été développée de manière à modéliser la consommation ainsi que la surface utilisée par des IP matérielles pour FPGA [CA07]. La modélisation de la consommation dynamique est effectuée à partir de régressions linéaires en fonction de paramètres tels que la surface, la fréquence etc. L'outil XPower de Xilinx ainsi que des simulations sont utilisés pour obtenir des consommations moyennes. La puissance dynamique de l'IP est ainsi rendue générique et facilement exploitable à haut-niveau par l'utilisateur.

Une autre technique combinant cette fois-ci, l'utilisation de macro-modèles et de SystemC, est présentée dans [DS07]. La librairie SystemC est utilisée pour modéliser le système tandis que les macro-modèles expriment la consommation en fonction des paramètres statistiques des signaux d'entrées. Les macro-modèles développés correspondent à des éléments de relativement faible complexité tels que des additionneurs, des portes logiques, etc. Bien que l'approche semble relativement rapide pour de petits éléments grâce au *macro-modelling*, ce type d'approche ne semble pas approprié pour l'estimation de systèmes complexes.

Les fabricants de FPGAs ont, eux aussi, développé des outils permettant l'estimation de la consommation au niveau système, appelés *spreadsheets*, littéralement des feuilles de calculs. Les plus connus sont Xilinx Power Estimator (XPE) [Xil12c] de Xilinx et PowerPlay [Alt14] d'Altera. Ces outils permettent à l'utilisateur d'estimer la consommation de puissance moyenne d'un circuit sans effectuer aucune étape d'implémentation. Des formules analytiques permettent d'estimer une consommation moyenne après que l'utilisateur ait spécifié des informations comme le nombre et le type de ressources, le FPGA cible, la fréquence d'horloge, l'activité des signaux, la tension d'alimentation. Leur précision dépend de la connaissance de l'utilisateur sur le circuit. L'utilisateur peut tout de même raffiner ces estimations au fur et à mesure du développement du système, directement après l'étape de synthèse où plus d'informations peuvent être obtenues. Ce type d'approche peut s'avérer utile pour des modules de faible complexité comme des additionneurs ou des blocs multiplieurs mais s'avère peu efficace pour des systèmes complexes.

Pour conclure, le tableau 3.2 récapitule les principales techniques, outils et méthodologies d'estimation décrites précédemment.

3.2 Conclusion

A l'issue de cet état de l'art, plusieurs constats peuvent être dressés. Premièrement, c'est à haut-niveau que les techniques d'estimation de la consommation doivent être développées afin de permettre l'exploration de l'espace de conception dans des temps de simulations courts. C'est à ce niveau que les choix des concepteurs auront le maximum d'impact sur les performances et la consommation du système final.

TABLE 3.2 – Exemples de méthodologies, outils et techniques d’estimation et de modélisation de la consommation de puissance

Niveau d’abstraction	Nom	Références
Transistor	SPICE PowerMill IRSIM	[Nag75] [Inc00] [SH89]
Porte Logique	approches statistiques Approches probabilistes	[BNYT93, KN97, TBAV05, SKB02, DT05] [Naj94, LW06a, AN04, HAE09, CKCW06, MMP94]
RTL	Macro-modelling Régression, tables de valeurs/d’équations Outils	[DA14, SJK ⁺ 11, SJ01, LP05, DR ⁺ 06] [NMH12, DSZC11] [SWPC10, BCMP01, SKN ⁺ 01]
Système	Modèles flexibles haut-niveau Spreadsheets SystemC+ Power state machine SystemC+ macro-modèles	[DDG ⁺ 12, ARFB10, DDVJM12] [Xi12c, Alt14] [HLF ⁺ 11, SRH ⁺ 11] [DS07]

Les approches à bas niveaux permettent d’obtenir une bonne précision sur l’estimation de la consommation mais la quantité importante de détails liés à l’implantation physique augmente considérablement le temps et les ressources de simulation. Dans le cas de notre étude, de multiples configurations de système doivent être étudiés, ce qui ne semble pas envisageable à faibles niveaux d’abstraction.

Au niveau Système, les temps de simulations sont plus courts et des détails liés à l’implémentation sont abstraits. Une solution opportune serait de ramener l’information précise de consommation des niveaux bas au niveau système.

D’autre part, l’utilisation de la librairie spécifique SystemC semble être un choix judicieux pour la modélisation de système numériques à haut-niveau. Les temps de simulation peuvent être grandement réduits et la fonctionnalité ainsi que les performances du système peuvent être évaluées rapidement. De plus, le système modélisé est nettement plus flexible qu’à faible niveau d’abstraction. Dès lors, tester de multiples configurations du système dans des délais de simulation court semble réalisable.

De rares approches permettent de considérer les aspects dynamiques des applications au cours du temps. En effet, tous les éléments du système sont généralement supposés actifs, ce qui n’est pas forcément le cas dans un système réel. Tout ceci peut conduire à des sur-estimations. D’autre part, et à notre connaissance, il n’y a pas d’approche dans la littérature permettant l’estimation de la consommation de systèmes purement matériels, sur cibles FPGA, pour des chaînes entières du domaine des communications numériques sans fil. Dans ce sens, une méthodologie d’estimation de la consommation de système de communications numériques implantés sur FPGA a donc été développée et sera détaillée au chapitre suivant.

Chapitre 4

Méthodologie d'estimation de la consommation de systèmes sur FPGA

4.1 Introduction

La méthodologie développée dans le cadre de la thèse, vise à répondre aux problématiques soulevées dans les chapitres précédents. Cette méthodologie a été développée pour des systèmes entièrement matériels, sans éléments logiciels.

Elle repose sur l'introduction du concept de *scénario*. Ce terme a déjà été employé dans la littérature mais dans un sens différent [GPH⁺09, VSP10, ZPR⁺13]. Tout d'abord, on retrouve la notion de scénario au sens de cas d'étude, *use-case scenario*. Dans ce sens, ce terme regroupe les différentes réactions et comportements qu'un système peut adopter durant son utilisation. Il fait également référence au contexte dans lequel le système sera utilisé, c-à-d les entrées du système.

Il existe une autre notion appelée scénario système (*system scenario*). Ce terme se rapporte généralement à une méthodologie de conception qui permet de classer les comportements du système en se basant sur des compromis entre différents critères de coûts durant l'implémentation [GPH⁺09, ZPR⁺13].

Dans notre cas, le concept de scénario a été imaginé dans l'objectif de faciliter la comparaison de différentes applications dans le domaine des communications numériques sans fil. Dans les chapitres précédents, il a pu être constaté qu'un grand nombre de paramètres peut influencer à la fois les performances et la consommation d'un système. Le scénario doit faciliter et permettre une comparaison efficace de systèmes qui seront implantés sur FPGA en tenant compte de multiples critères que sont les performances intrinsèques du système (débit, latence, etc.) ainsi que de la consommation. La définition de ce scénario constitue une première étape pour les concepteurs du système. Ainsi, ils peuvent explorer l'espace de conception afin d'évaluer l'impact de leurs choix, sans avoir à dérouler le flot de conception du système entier, ce qui est souvent source d'erreurs et très long. Avec cette approche, les concepteurs des systèmes peuvent rapidement comparer différentes solutions matérielles entre elles et évaluer l'impact de paramètres d'intérêt (systèmes ou dépendants

de la technologie).

Un autre point innovant de la méthodologie est de prendre en compte l'activité temporelle des blocs IPs grâce à la modélisation des signaux de contrôle relatifs à ses composants IP matériels. Ceci est relativement important afin d'obtenir des résultats d'estimation précis. En effet, pour un circuit donné, si l'on applique deux testbenchs différents sur les entrées, le comportement ainsi que la consommation peuvent varier alors que le même circuit est utilisé.

La méthodologie se déroule en 3 étapes :

- la définition du scénario,
- une phase de caractérisation et
- une phase de simulation haut-niveau.

4.2 Le concept de scénario

Le concept de scénario est illustré à la Figure 4.1. Premièrement, les concepteurs système doivent définir le scénario qui est le point d'entrée de la méthodologie. Le terme de scénario fait référence à un ensemble de paramètres communs à plusieurs applications. Ce dernier permet de regrouper à la fois des paramètres de haut niveau et des paramètres dépendants de la technologie (c-à-d de la cible matérielle) qui ont un impact sur les performances et/ou sur la consommation. Aussi, chaque application correspond à un paramétrage complet d'un scénario et constitue donc une instance de scénario. Dans un autre sens, le scénario peut être vu comme le méta-modèle des applications.

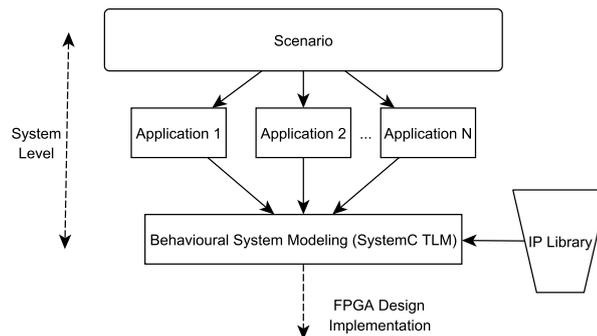


FIGURE 4.1 – Introduction du concept de scénario

Dans un second temps, chaque application est modélisée à haut niveau en SystemC à partir d'une librairie de modèles dédiés. Cette librairie est construite autour de modèles haut-niveau décrits en SystemC, relatifs à des blocs IP. Dans le domaine des communications numériques, tous ces modèles, aussi appelés IPs matérielles, sont des éléments qui constituent une chaîne de communication numérique comme les encodeurs, les modulateurs QAM, l'IFFT, les estimateurs de canal, les égaliseurs, etc.

Afin de construire cette librairie de modèles, un grand nombre de blocs IP a été entièrement caractérisé pour l'ensemble de leurs paramètres intrinsèques. En effet, chaque bloc IP a été caractérisé en termes de ressources, puissance moyenne consommée, vitesse (fréquence d'horloge maximale), etc. Chaque modèle haut-niveau, en SystemC, peut décrire le comportement d'un composant de type bloc IP auquel il correspond, en fonction des paramètres définis pour chaque application. Par exemple, dans le cas d'un bloc IP de type IFFT, les paramètres peuvent être la taille des données traitées c-à-d la quantification, la taille de la transformée, etc. Le procédé de caractérisation des blocs IP et la création des modèles SystemC seront décrits dans les sections suivantes.

L'étape finale de la méthodologie consiste à développer un modèle haut-niveau du système à partir des modèles décrits en SystemC. Ensuite, les concepteurs peuvent fournir un testbench c-à-d une séquence de données à appliquer en entrée de leur système. En fonction du scénario qui a été défini par les utilisateurs de la méthodologie, chaque application est simulée à haut-niveau à l'aide du noyau de simulation intégré à SystemC. Pour chaque application, les modèles SystemC sont paramétrés et les utilisateurs peuvent ainsi évaluer l'impact des choix de paramètres ainsi qu'explorer l'espace de conception à haut-niveau. A l'issue de l'étape de simulation du modèle du système, le comportement temporel des composants IP pour chaque application peut être déterminé à l'aide des modèles SystemC. En effet, on rappelle qu'il est important de déterminer le comportement temporel de chaque composant du système de manière à pouvoir obtenir de bonnes estimations de consommation. De plus, cette étape de simulation est nécessaire dans l'objectif de valider la fonctionnalité du système ainsi que d'évaluer ses performances à l'aide de certaines métriques comme le taux d'erreurs binaire (TEB), le débit ou encore la latence. Ces métriques sont déterminantes pour comparer plusieurs systèmes entre eux.

Dans notre approche, les aspects énergétiques ont un rôle significatif. La métrique de consommation d'énergie, de puissance dynamique moyenne du circuit, est évaluée en déterminant le comportement de chaque bloc IP du système pour un temps de simulation donné. D'ailleurs, il est aussi possible de fixer d'autres critères d'arrêt de la simulation comme :

- un nombre d'erreurs à atteindre en réception,
- un volume de données à émettre,
- etc.

Dans un dernier temps, la valeur de consommation du système global peut être estimée à partir des contributions de chaque élément du système et ceci, pour différentes applications. Il est aussi possible d'explorer les nombreuses possibilités de paramétrage des modèles (et donc des composants IP matériels) afin de trouver le meilleur compromis entre les performances et l'énergie consommée.

Afin d'illustrer l'utilisation de la méthodologie, la Figure 4.2 fournit un exemple d'une chaîne de communications numériques à base de modulation OFDM, en configuration SISO. Pour ce système, les utilisateurs peuvent proposer un scénario incluant des paramètres haut niveau et des paramètres orientés matériels tels que :

- la modulation utilisée,
- la taille de (I)FFT,
- la taille du préfixe cyclique,
- la quantification des données,
- la fréquence d'horloge et
- le type de FPGA...

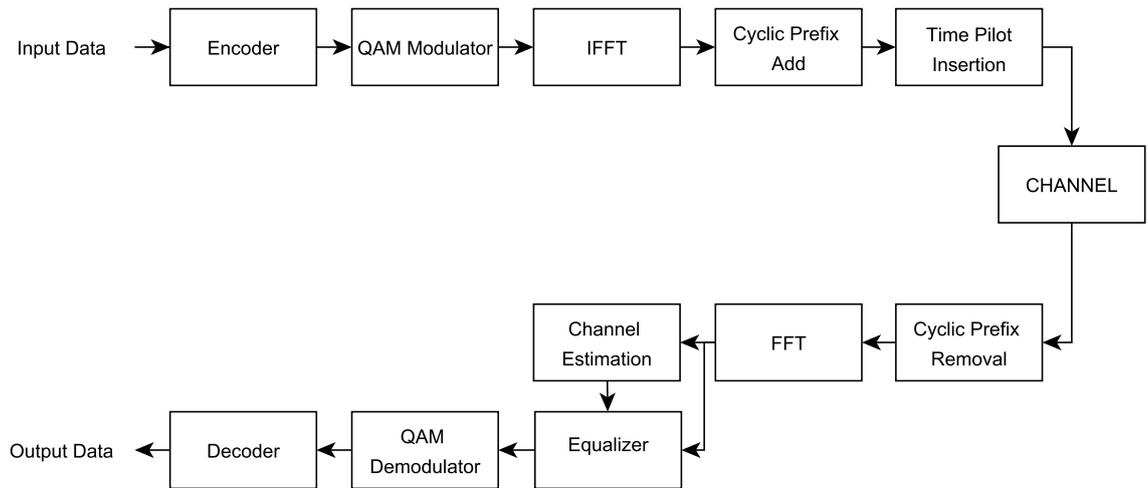


FIGURE 4.2 – Schéma d'un émetteur et d'un récepteur d'une chaîne SISO-OFDM

A partir de ce scénario, des applications, c-à-d des instances du scénario, sont déduites et pourront être testées. Un autre exemple de scénario ainsi que deux exemples d'application sont indiqués dans le tableau 4.1. Les paramètres d'intérêt ciblés par l'utilisateur, comme le rendement du codeur de canal, la taille de FFT ou encore le nombre de sous-porteuses utiles, constituent le scénario. De cette manière, les utilisateurs peuvent évaluer l'impact de chaque paramètre sur les métriques de performances et de consommation.

TABLE 4.1 – Exemples d'application pour la chaîne SISO-OFDM

Scénario	Application 1	Application 2
Nombre d'antennes d'émission	1	2
Rendement du codeur de canal	1/3	
Taille de FFT	512 pts	1024 pts
Modulation utilisée	QPSK	16QAM
Sous-porteuses utiles	100%	75%
Fréquence d'horloge	50 MHz	25 MHz
FPGA	Virtex 4	

Pour un scénario donné, différentes applications peuvent donc être réalisées, conduisant à des consommations énergétiques différentes. En effet, tous les blocs de traitement ne sont généralement pas tous actifs en même temps, comme indiqué sur la Figure 4.3.

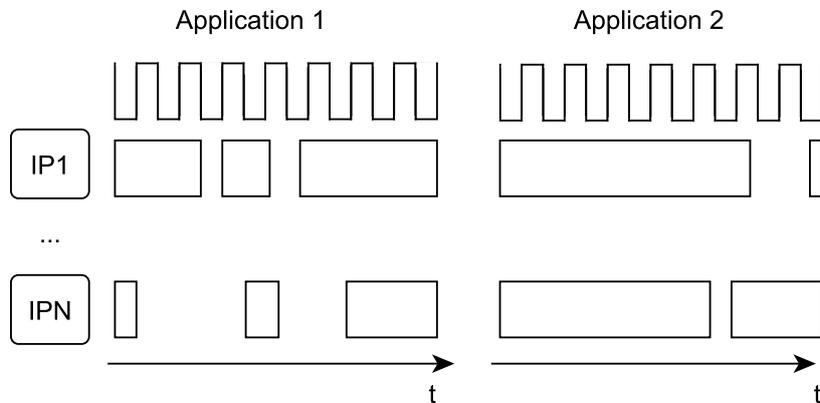


FIGURE 4.3 – Illustration de temps d’activité différents pour 2 applications

Ceci peut aussi être le cas, pour un circuit donné avec des testbenchs différents. Chaque application possède ses propres caractéristiques qui influencent le comportement temporel de chaque bloc. Cet aspect dynamique de l’application au cours du temps peut conduire à des estimations de consommation non réalistes. Il est donc nécessaire d’en tenir compte.

Dans les prochaines sections, les étapes de caractérisation et de modélisation/simulation haut niveau de la méthodologie sont décrites en détails.

4.3 Caractérisation des blocs IP matériels

Dans l’approche proposée, nous faisons l’hypothèse qu’un système peut être entièrement représenté par un ensemble de modèles haut-niveau connectés entre eux. Chaque modèle haut-niveau fait référence à un ensemble de configurations (C_i , avec i allant de 1 à N) correspondant à une combinaison de paramètres (quantification des données, fréquence d’horloge etc.). Pour chaque configuration C_i du modèle de haut-niveau, la description RTL du composant IP matériel existe dans une librairie dédiée. Ces descriptions RTL sont généralement décrites en utilisant des langages de description matériel (HDL) tels que le VHDL ou le VERILOG. Elles peuvent aussi être directement générées à l’aide d’outil d’un fabricant de FPGA (par exemple, l’interface graphique Core Generator de Xilinx). Dans ce cas, le composant IP a généralement besoin d’un contrôleur associé pour effectuer sa configuration et son contrôle.

Comme indiqué sur la Figure 4.4, chaque bloc IP, qui correspond à une configuration spécifique de son modèle haut-niveau, est entièrement caractérisé suivant les différentes étapes constituant le flot de conception d’un FPGA. Son implémentation est effectuée en suivant les étapes de synthèse, mapping et placement-routage. On peut remarquer que ces étapes sont spécifiques à un FPGA donné, conduisant à l’utilisation d’un nombre de ressources spécifiques ainsi que des temps de propagation définis.

Des options d'implémentation peuvent être choisies par les utilisateurs comme par exemple les choix d'optimisation lors de l'étape de synthèse (niveau de performances, surface utilisée, consommation), de l'étape de placement et de routage, etc. De plus, des contraintes peuvent être fournies aux outils de conception par l'intermédiaire d'un fichier dédié. Toutes ces considérations peuvent avoir un impact important sur les performances (surface utilisée, vitesse du circuit) finales du système ainsi que sur sa consommation d'énergie [MGP11]. Afin de s'affranchir des choix de ces options et ainsi limiter l'impact des optimisations logicielles sur le circuit final et sur chaque bloc IP, ces options ont, soit été réduites à une valeur standard, soit fixées à leurs valeurs par défaut. De cette manière, à l'issue des phases d'implémentation, un circuit avec le minimum d'optimisations est obtenu, conduisant à une valeur non-optimisée de consommation. On peut alors considérer que, dès lors que les utilisateurs appliqueront des contraintes d'optimisation sur la consommation, les valeurs de consommation du circuit seront alors inférieures à celles obtenues en phase de caractérisation. A noter que durant la thèse, l'environnement de développement ISE v14.4 de Xilinx a été utilisé [Xil15].

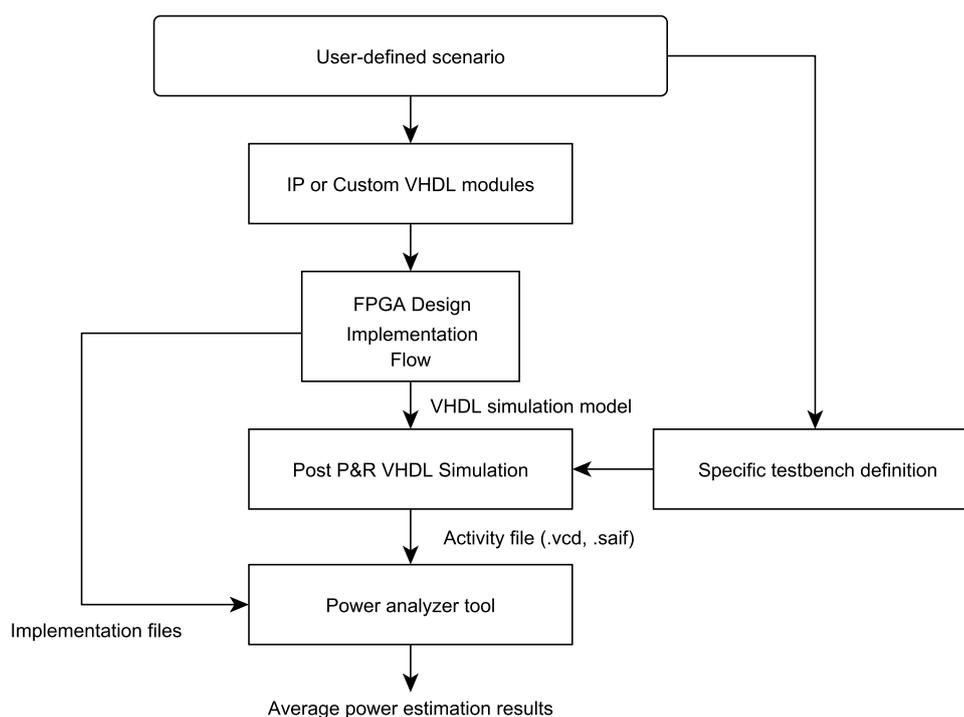


FIGURE 4.4 – Caractérisation de chaque bloc IP matériel

Après les étapes d'implémentation des blocs IP, un modèle de simulation obtenu après placement et routage et décrit en VHDL, est généré. Ce modèle fournit des informations précises concernant les retards et les temps de propagation dans la netlist du circuit final. Comme indiqué sur la Figure 4.4, le modèle de simulation de bas-niveau est ensuite simulé. Pour cela, l'outil Modelsim 10.1c de Mentor Graphics [Men15] a été utilisé comme simulateur. Un fichier de *testbench*, décrit en VHDL, a aussi été généré afin d'appliquer les signaux appropriés aux entrées du modèle bas niveau du composant IP à tester.

Durant cette simulation temporelle, le phénomène de *glitches* (voir section 2.3.2 du chapitre 2) peut être précisément évalué. De plus, l'activité interne du circuit est sauvegardée dans un fichier d'activité (.vcd ou .saif).

Deux simulations temporelles sont effectuées afin d'évaluer l'activité interne du circuit lorsque :

1. l'IP est active durant toute la durée de la simulation,
2. l'IP est inactive et qu'il n'y a pas d'activité.

A ce niveau de la phase de caractérisation, des outils d'analyse de la consommation comme XPower Analyzer [Xil12c] de Xilinx ou encore PowerPlay Analyzer [Alt14] d'Altera permettent, à partir des fichiers d'activité obtenus durant les simulations et de fichiers relatifs à l'implémentation (c-à-d la netlist du circuit et de fichier de contraintes), de fournir des estimations des puissances statique et dynamique du circuit.

La première simulation permet d'obtenir une estimation de la puissance dynamique moyenne lorsque le bloc IP est actif durant toute la simulation.

La seconde simulation permet de déterminer une valeur de puissance dynamique moyenne lorsque le bloc IP n'est pas actif, généralement lorsqu'une entrée de contrôle de type *clock enable* ou *chip enable* est égale à 0. De cette manière, des valeurs de consommation moyenne sont connues lorsque les composants IP sont dans ces deux états.

Remarques sur les simulations en phase de caractérisation :

Les points importants sont la durée de simulation et le choix des stimuli (ou *patterns*) d'entrée. En effet, l'objectif est d'éviter aux utilisateurs d'effectuer un grand nombre de simulations qui pourraient être très coûteuses en temps. Une étude menée dans le cadre de la thèse, dont une partie des résultats est indiqué dans le Tableau 4.2, a montrée que la durée de la simulation en phase de caractérisation dépend de l'architecture du composant IP.

Par exemple, si le composant possède des étages de pipeline, il est suffisant d'étudier l'activité de ce dernier uniquement lorsque le pipeline est plein (nous qualifions cet état de régime permanent). Si le composant ne possède pas d'étages de pipeline, la longueur de simulation doit être suffisamment longue afin d'obtenir une valeur représentative de la puissance dynamique moyenne.

Dans le tableau 4.2 est résumé les valeurs de consommations dynamiques moyennes obtenues pour le bloc IP IFFT (pour une quantification des données sur 10bits, une taille de transformée de 512pts et pour un FPGA Virtex-6 LX240T) en faisant varier la durée de simulation (exprimés en nombre de symboles OFDM générés) ainsi que le type de pattern appliqué (données binaires aléatoires ou données typiques telles que des symboles complexes modulés QPSK).

On peut constater que, pour des blocs IPs possédant un pipeline, les valeurs de consommation dynamique moyenne obtenues sont relativement constantes, indépendamment de la nature des stimuli d'entrée et de la durée de simulation. De plus, nous avons vérifié et

TABLE 4.2 – Tableau d'évaluation de la puissance dynamique moyenne consommée pour différentes durées de simulation et de patterns appliqués

Bloc IP IFFT : 512pts, 10bits 50MHz, Virtex-6 LX240T					
Durée d'évaluation : 1 symbole OFDM					
Patterns	1	2	3	4	5
	Puissance dynamique moyenne (mW)				
Typiques	58.74	59.36	59.2	59.04	58.59
Aléatoires	59.99	59.85	60.02	60.03	59.85
Durée d'évaluation : 10 symboles OFDM					
Patterns	1	2	3	4	5
	Puissance dynamique moyenne (mW)				
Typiques	58.74	59.36	59.2	59.04	58.95
Aléatoires	60.06	60.07	59.97	59.92	59.98
Durée d'évaluation : 90 symboles OFDM					
Patterns	1	2	3	4	5
	Puissance dynamique moyenne (mW)				
Typiques	58.88	59.02	59.06	59.04	59.01
Aléatoires	59.99	60.08	60.01	59.98	59.98

validé les hypothèses émises, pour d'autres configurations de ce bloc IP (tailles de transformée allant de 256 à 2048, ainsi que pour des valeurs de quantifications différentes allant de 10 à 18 bits). A noter que l'objectif n'est pas d'obtenir une précision extrêmement fine mais plutôt d'obtenir des valeurs qui reflètent les valeurs typiques de consommation du bloc IP dans chaque état.

Pour les blocs IP qui ne possèdent pas de pipeline, les stimuli appliqués aux entrées doivent être les plus réalistes possibles par rapport à ceux utilisés dans le système final. Pour palier cela, durant les simulations, nous avons considéré des données binaires typiques, par exemple des symboles complexes QPSK au module suivant le modulateur QPSK, ou des données aléatoires ayant le même type d'activité que dans le système final. Si ces conditions ne sont pas réalisées, les estimations délivrées par des outils comme XPower et PowerPlay peuvent varier d'environ +/- 20% [Alt15b].

A l'aide d'outils comme XPower et PowerPlay, il est possible d'évaluer chacune des composantes de la puissance dynamique. Ainsi, la puissance moyenne dynamique d'un système ou d'un bloc IP sur FPGA est répartie suivant plusieurs termes, comme indiqué par l'équation 4.1 :

$$P_{Dynamique_{IP}} = P_{Horloge_{IP}} + P_{Logique_{IP}} + P_{Signal_{IP}} + P_{E/Ss_{IP}} + P_{BRAM_{IP}} + P_{DSP_{IP}} \quad (4.1)$$

avec

- $P_{Horloge_{IP}}$ la puissance dynamique moyenne consommée par l'arbre d'horloge incluant les buffers et les éléments de routage,

- $P_{Logique_{IP}}$ la puissance dynamique moyenne consommée par l'ensemble des éléments logiques (*Configurable Logic Block* (CLB) ou *Logic Element* (LE)) incluant les LUTs et les bascules,
- $P_{Signal_{IP}}$ la puissance dynamique moyenne consommée par les éléments de routage et les données,
- $P_{E/Ss_{IP}}$ la puissance dynamique moyenne consommée par les broches d'entrées/sorties,
- $P_{BRAM_{IP}}$ la puissance dynamique moyenne consommée par les mémoires,
- $P_{DSP_{IP}}$ la puissance dynamique moyenne consommée par les blocs multiplieurs, *Digital Signal Processing* (DSP).

Idéalement, la phase de caractérisation doit être effectuée pour chaque configuration d'un composant IP et pour un FPGA donné. Actuellement, la librairie contient de multiples analyses de dizaines de composants différents. Dans le tableau 4.3, une liste non-exhaustive de résultats d'estimation de puissance dynamique obtenus pour différents composants est présentée. Les résultats ont été obtenus pour une fréquence d'horloge de 50MHz, avec une quantification des données de 14 bits et pour une cible FPGA Virtex6-LX240T.

Un point important de la méthodologie a été l'automatisation de la phase de caractérisation. Pour cela, des scripts TCL (*Tool Command Language*) ont été développés de manière à réduire l'effort de caractérisation. Les utilisateurs n'ont plus à entrer dans le flot de conception mais seulement à configurer ces scripts pour pouvoir caractériser leurs propres composants. Les résultats pourront aussi permettre de compléter la bibliothèque de composants déjà caractérisés.

TABLE 4.3 – Exemples de puissances dynamiques moyennes pour différents composants IP

Nom ¹	Configuration utilisée	Puissance dynamique moyenne (mW)
IFFT Xilinx - IP	Architecture pipelinée Scaled [256, 512, 1024, 2048]	[59.45, 71.02, 79.8, 93.8]
FFT Xilinx - IP	Architecture pipelinée Unscaled [256, 512, 1024, 2048]	[67.98, 80.31, 86.89, 103.13]
Modulateur QAM - C	[QPSK, 16QAM, 64QAM]	[2.19, 2.44, 2.88]
Démodulateur QAM - C	[QPSK, 16QAM, 64QAM]	[1.9, 2.88, 2.3.58]
LTE Turbo Encoder Xilinx - IP	R=1/3 (bloc de 1024 bits)	8.89
LTE Mixed Mode Turbo Decoder Xilinx - IP	R=1/3 (bloc de 1024 bits)	179.27
Channel Estimator Equalizer -C	Zero-Forcing	[73.64,20.12]

¹ IP ou VHDL 'Custom' (C)

4.4 Modélisation haut-niveau en SystemC des blocs IP matériels

Une fois que l'étape de caractérisation des blocs IP est effectuée, des modèles correspondants sont développés à haut niveau et sauvegardés dans une librairie. Dans un second

temps, le modèle SystemC du système final est construit à partir de ces modèles.

Chaque modèle haut-niveau a été construit suivant un modèle d'implémentation précis et commun à tous les modèles. Ce modèle d'implémentation a été choisi car il est particulièrement bien adapté à la modélisation des blocs IP utilisés dans le domaine des communications numériques. Ce modèle d'implémentation, décrit à la Figure 4.5, est composé d'un chemin de contrôle et d'un chemin de données (*Data Path*).

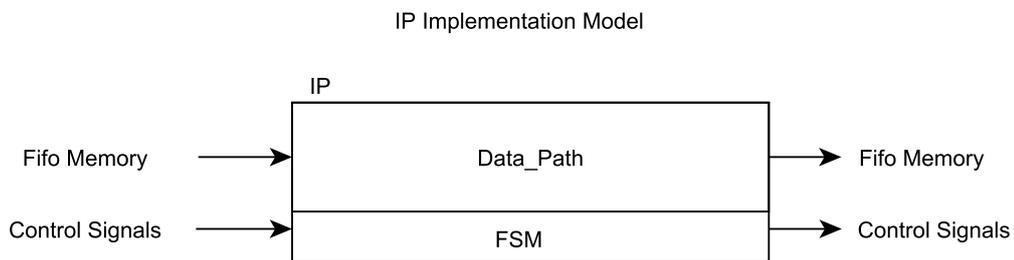


FIGURE 4.5 – Modèle d'implémentation de l'IP

Le chemin de données permet de réaliser la fonctionnalité du composant à partir des données d'entrée. Deux cas de figures peuvent alors se présenter en fonction de la connaissance des concepteurs sur le bloc IP.

Premièrement, les utilisateurs ne possèdent pas de détails liés à l'implémentation matérielle du bloc IP. Dans ce cas, la fonctionnalité peut être décrite de manière relativement abstraite. Dans le second cas, les concepteurs ont connaissance du bloc IP qu'ils souhaitent utiliser. Actuellement, les vendeurs de blocs IP comme Xilinx fournissent généralement des modèles fonctionnels de leurs blocs IPs, décrits en langage C/C++. Ces modèles permettent de délivrer à haut-niveau, les mêmes résultats que le bloc IP matériel, avec une précision au bit près. Ces modèles peuvent alors être facilement intégrés, à haut niveau, dans le modèle SystemC de manière à fournir des résultats précis sur la fonctionnalité du composant. De plus, SystemC supporte à la fois une précision en virgule flottante ou fixe, tout comme ces modèles. Il est alors possible d'évaluer l'impact de la quantification des données sur les performances du système de manière très précise. Cependant, ce type de modèle n'est pas précis au cycle près puisqu'aucune notion de temps n'est considérée.

Le modèle d'implémentation est aussi constitué d'un chemin de contrôle. Ce dernier est composé d'une machine à états (*Finite State Machine* (FSM)) qui évolue en fonction des valeurs des signaux de contrôle des entrées ainsi que des paramètres de configuration. Ces derniers sont relatifs à chaque bloc IP (latence de calcul, simulation de pipeline, etc...) et leurs valeurs sont définies pour chaque application.

Les signaux de contrôle de sortie d'un modèle sont reliés aux entrées du modèle suivant et ainsi de suite pour l'ensemble du système. On peut noter que peu d'informations de bas niveau (c-à-d liées à l'implémentation) sont nécessaires car le comportement d'un bloc IP est généralement contrôlé par quelques signaux clés (comme le *clock enable*, le *chip enable*,

une entrée validation d'écriture des données d'entrée, une sortie indiquant des données valides, etc.). De telles informations sont généralement accessibles très tôt à partir des documentations constructeurs (*datasheet*) des blocs IP par exemple.

Dans un souci de rendre les modèles haut-niveau compatibles avec les blocs IP matériels, les modèles en SystemC possèdent la même interface, visible sur la Figure 4.6. L'interface choisie est basée sur l'*Advanced eXtensible Interface* (AXI) [Xil11a] qui permet de partager des informations suivant un protocole spécifique. Ce type d'interface est générique et permet d'inter-changer les modèles facilement. De plus, l'interface est très répandue et des plus utilisées parmi les fabricants de FPGA et reste compatible avec la grande majorité des cibles.

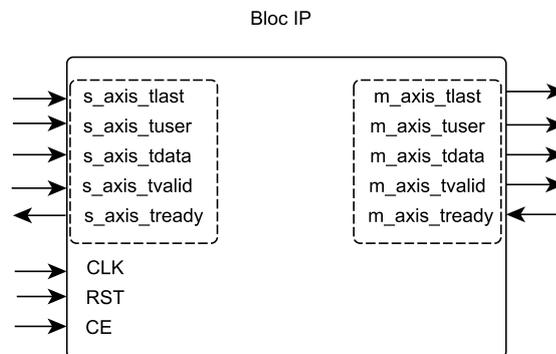


FIGURE 4.6 – Illustration de l'interface générique basée sur le standard AXI

Chaque modèle possède donc deux interfaces : une maître et une esclave. Ces deux interfaces contiennent différents signaux de contrôle dont un indicateur de fin d'échange de données (X_axis_tlast), un indicateur de données valides (X_axis_tvalid), un indicateur de bloc prêt (X_axis_tready). Les paramètres de configuration de chaque bloc IP peuvent être modélisés à l'aide de l'entrée X_axis_tuser . Pour finir, les données transitent via les ports X_axis_tdata et des entrées d'horloge (CLK), de remise à zéro (RST) et de *clock enable* (CE) sont présentes.

Dans le cadre de la méthodologie, un des caractères innovants réside dans la supervision de l'activité de chaque modèle. Pour cela, ce sont les signaux clés de contrôle des modèles SystemC qui sont supervisés. Lors des simulations, ces signaux vont évoluer et il est alors possible de déterminer facilement l'état du modèle (état actif ou inactif) au fur et à mesure de la simulation. L'activité temporelle de chaque modèle peut alors être scrutée de manière précise, au cycle près, dans le cas d'une description fine de la FSM du modèle haut-niveau. En outre, des valeurs de consommation peuvent ainsi corrélées en fonction de l'activité des signaux de contrôle pour chaque modèle. On peut noter que la spécification par l'utilisateur de ces signaux conditionne nettement la précision de l'estimation finale de la consommation.

La supervision des signaux de contrôle des différents modèles est réalisée par un modèle

spécifique, appelé *Power Monitor*. Ce modèle prend comme entrées, les différents signaux clés de contrôle définis par l'utilisateur. Son rôle est de délivrer en sortie des coefficients relatifs à l'activité temporelle de chaque modèle. Ces coefficients sont calculés de la manière suivante :

$$Act_{Coeef} = \frac{T_{IP_a}}{T_s} \quad (4.2)$$

avec T_{IP_a} le temps durant lequel le modèle du bloc IP est actif, T_s le temps durant lequel le système entier est actif.

Ces coefficients représentent donc la fraction de temps durant laquelle le modèle (ou le bloc IP) est actif par rapport au temps total au cours duquel le système est actif. Ces coefficients sont calculés directement à partir des signaux de contrôle de chaque modèle.

Finalement, une estimation de la consommation de puissance dynamique du système entier peut être obtenue en considérant l'ensemble des contributions des modèles haut-niveau des blocs IP du système et les coefficients calculés précédemment :

$$P_{Dynamique_{système}} = P_{Horloge_{système}} + P_{Logique_{système}} + P_{Signal_{système}} + P_{BRAM_{système}} + P_{DSP_{système}} \quad (4.3)$$

avec :

$$P_{Horloge_{système}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) \times ((P_{Horloge_{IP_{active}}}(i) - P_{Horloge_{IP_{inactive}}}(i))) + (P_{Horloge_{IP_{inactive}}}(i) - P_{buf_{Horloge}}(i)) \quad (4.4)$$

avec $P_{buf_{Horloge}}$, la puissance dynamique consommée par les buffers de l'arbre d'horloge (déterminée durant la phase de caractérisation). En effet durant l'implémentation de chaque bloc IP, des buffers spécifiques pour chaque signal d'horloge sont insérés. La puissance consommée par ces buffers ne doit pas être prise en compte pour l'estimation du système complet (tout comme la puissance dynamique des E/S). On a aussi :

$$P_{Logique_{système}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) \times P_{Logique_{IP_{active}}}(i) \quad (4.5)$$

$$P_{Signal_{système}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) \times (P_{Signal_{IP_{active}}}(i) - P_{buf_{Signal_{IP_{active}}}}(i)) \quad (4.6)$$

avec $P_{buf_{Signal_{IP_{active}}}}$, la puissance dynamique des buffers des signaux d'entrées et de sorties (déterminée durant la phase de caractérisation). De la même manière que pour les buffers au niveau de l'arbre d'horloge, on retire la puissance consommée par des buffers spécifiques sur certains signaux. On a aussi :

$$P_{BRAM_{\text{système}}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) \times P_{BRAM_{IP_{active}}}(i) \quad (4.7)$$

$$P_{DSP_{\text{système}}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) \times P_{DSP_{IP_{active}}}(i) \quad (4.8)$$

avec NB_{IP} le nombre de blocs IP utilisés dans le système entier.

A noter que l'on ne considère pas la consommation dynamique liée aux broches d'E/S dans le calcul de chaque bloc IP. En effet, celle-ci ne doit pas être prise en compte.

Dans le cas où l'on possède relativement peu d'information sur le bloc IP et son implémentation matérielle, on peut aussi exprimer la consommation d'un système entier à plusieurs blocs IP sous la forme suivante :

$$P_{\text{système}} = \sum_{i=1}^{NB_{IP}} Act_{Coeef}(i) (P_{Dynamique_{IP_{actif}}}(i) - P_{Dynamique_{IP_{inactif}}}(i)) + P_{Dynamique_{IP_{inactif}}}(i) \quad (4.9)$$

avec $P_{Dynamique_{IP_{actif}}}$ la puissance dynamique totale (sans tenir compte de la puissance consommée par les buffers et les broches E/S) lorsque le bloc IP est actif et $P_{Dynamique_{IP_{inactif}}}$ la puissance dynamique totale lorsque le bloc IP est inactif. Pour certains blocs IP qui ne possèdent pas d'entrée de *clock enable* (signal de contrôle permettant la désactivation de l'horloge), une puissance dynamique est générée via l'activité de l'arbre d'horloge. La puissance dynamique du bloc IP considéré n'est alors pas nulle et doit être prise en compte.

Dans le cas où l'on possède une information détaillée sur le bloc IP, il est possible, à l'aide de certaines documentations constructeurs, de développer plusieurs signaux de contrôle spécifiques aux BRAMs, aux DSPs, afin d'obtenir des coefficients d'activité pour chaque type de ressource. Ainsi, plusieurs coefficients d'activité peuvent être obtenus pour un même composant IP de manière à raffiner les estimations. A titre d'exemple, l'entrée *BUSY* du module de Transformée de Fourier v7.1 [Xil11c] permet d'indiquer lorsque le bloc IP est en train de calculer la transformée. Lorsque ce signal est à l'état bas, on peut en déduire que les blocs multiplieurs DSP sont aussi inactifs.

4.5 Conclusion

Dans ce chapitre, une méthodologie permettant l'estimation de la consommation d'un système implanté sur FPGA a été présentée. Cette méthodologie est réalisée en deux étapes. La première étape consiste en la caractérisation des composants IP utilisés dans le système, associée aux développements de modèles SystemC haut-niveau de ces composants. La seconde étape est la représentation du système global à partir des modèles haut

niveau représentatifs de chaque bloc IP.

L'introduction du concept de scénario est un point clé de la méthodologie. Le scénario permet d'étudier et d'évaluer l'impact des choix de paramètres fait par l'utilisateur sur les performances et la consommation. De plus, cet outil devrait permettre de comparer facilement et de manière formelle, plusieurs applications entre elles.

D'autre part, la méthodologie d'estimation proposée permet de considérer les comportements temporels des blocs IPs, qui peuvent varier de multiples manières (données en entrées, contrôle extérieur, choix de configuration, etc.). L'estimation de la consommation peut ainsi être ajustée en conséquence.

Une liste des avantages supplémentaires de la méthodologie est présentée ci-dessous :

- l'exploration de l'espace de conception est facilitée,
- les comportements temporels des composants IP dans chaque application sont considérés,
- le développement du système dans sa globalité est évité,
- l'estimation de la consommation est plus précise (car elle repose sur une caractérisation de bas niveau),
- les bibliothèques de blocs IP matériels ainsi que des modèles SystemC peuvent être ré-utilisées et étendues.

Dans le chapitre suivant, les apports de la méthodologie seront présentés à travers l'étude de différents systèmes de communications numériques sans fil. Ces systèmes seront évalués et comparés à la fois en termes de consommation et de performances.

Chapitre 5

Évaluations haut-niveau de systèmes implantés sur FPGA

5.1 Introduction

A l'aide de la méthodologie d'estimation proposée au chapitre 4, une comparaison des traitements en bande de base de plusieurs systèmes de communications numériques sans fil a été effectuée, à la fois en termes de performances et de consommation énergétique. Pour cela, différentes chaînes de traitement en bande de base ont été développées dans le cadre de la thèse :

1. une chaîne d'émission-réception SISO-OFDM,
2. un émetteur LTE SISO-OFDM en voie descendante,
3. un émetteur LTE MISO-OFDM 2x1 en voie descendante avec encodage d'Alamouti,
4. un émetteur d'une chaîne MISO-OFDM à X antennes intégrant le principe de retournement temporel appliqué dans le domaine fréquentiel.

Ces systèmes vont permettre de servir de références et de bases de validation pour l'approche proposée. En appliquant la méthodologie proposée, les différents blocs IPs de chaque chaîne vont tout d'abord être caractérisés. Ensuite, des modèles haut-niveau en SystemC de ces blocs IP vont être développés. Finalement, des représentations haut-niveau de tous les systèmes seront réalisées à partir de ces modèles et des simulations seront effectuées vis-à-vis des différentes applications définies. Ainsi, la précision et la rapidité de l'approche seront évaluées efficacement par rapport à un système réellement implémenté.

5.2 Descriptions et évaluations des chaînes de communications numériques

5.2.1 Chaîne d'émission-réception SISO-OFDM

Sur la figure 5.1 est décrite la chaîne SISO-OFDM qui a été réalisée. Cette chaîne est un exemple typique de système de communications numériques sans fil basé sur la modulation OFDM. La chaîne a été complètement décrite en VHDL et est entièrement paramétrable.

L'émetteur est composé d'une source permettant de délivrer les bits d'information à transmettre. Les données binaires sont ensuite transmises au codeur de canal puis elles sont envoyées au modulateur MAQ. Ce modulateur permet la réalisation des modulations QPSK, 16QAM et 64QAM. Il permet d'effectuer le mapping de mots binaires (groupes de bits) en symboles I/Q complexes représentatifs d'une constellation. Le bloc IP appelé *Carrier Mapper* permet d'affecter les symboles I/Q modulés aux sous-porteuses choisies par l'utilisateur. En effet, dans des systèmes réels, toutes les sous-porteuses ne sont généralement pas utilisées, comme les sous-porteuses proches des bords de la bande, afin de réduire les dégradations subies. Ensuite, la modulation OFDM est réalisée à l'aide d'un bloc IP IFFT ainsi que l'insertion du préfixe cyclique. Les données issues de l'IFFT sont ensuite mises à l'échelle à l'aide du module appelé *DAC Scaling*. Le dernier module d'émission permet l'insertion de symboles OFDM appelés 'pilotes' dans le domaine temporel. Ce type de symbole a été généré à partir d'une séquence de symboles BPSK qui sera connue du récepteur. De cette manière, le canal pourra être estimé en réception à partir de ces symboles qui sont placés au début de chaque trame. A noter que les modules de contrôle ne sont pas représentés sur la figure 5.1.

Le récepteur réalise les opérations duales. Tout d'abord, le préfixe cyclique est retiré. Une FFT est ensuite effectuée pour démoduler les symboles OFDM. Les sous-porteuses utiles, c-à-d portant de l'information de l'utilisateur, sont gardées à l'aide du bloc *Carrier Unmapper*. L'estimation de canal est ensuite réalisée afin de déterminer la réponse fréquentielle du canal sur chaque sous-porteuse. Les données sont ensuite égalisées et une démodulation des symboles QAM est effectuée afin de retrouver l'information binaire codée. Finalement, un décodeur de canal est utilisé afin de décoder et ainsi restituer les bits d'information.

La chaîne a été construite autour de blocs IP décrits directement en VHDL ainsi qu'un certain nombre de composants IP provenant de Xilinx. Une liste non-exhaustive des paramètres modifiables de la chaîne est fournie ci-dessous :

- le nombre de symboles OFDM par trame (c-à-d entre deux symboles OFDM pilotes),
- le nombre de bits de quantification des données complexes I/Q,
- le choix de modulation,
- le nombre et la sélection des sous-porteuses actives,
- le rendement du codeur de canal,
- la taille d'(I)FFT,
- la fréquence d'horloge,
- le composant FPGA,
- etc.

Afin d'appliquer la méthodologie, un scénario et 4 applications ont été définis. Ils sont listés dans le tableau 5.1

On peut constater que les 4 applications possèdent la même efficacité spectrale car elles utilisent la même bande de fréquence, la même modulation (c-à-d QPSK) et un même rendement de codage (c-à-d $R = 1/3$). Pour les applications 1 et 2, la quantification des symboles I/Q complexes est fixée à 10 bits tandis que pour l'application 3 et 4, cette dernière est de 14 bits. D'autre part, la taille de d'IFFT diffère selon les applications. L'IFFT

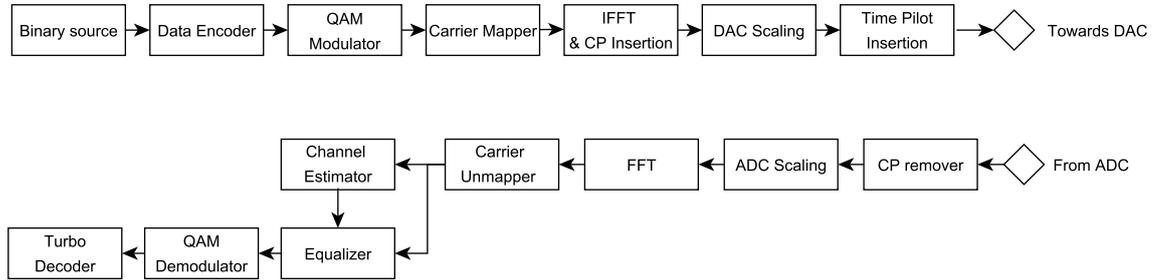


FIGURE 5.1 – Description de la chaîne SISO-OFDM

TABLE 5.1 – Scénario et applications testées pour la chaîne SISO-OFDM

Scénario	Appli. 1	Appli. 2	Appli. 3	Appli. 4
Codage de canal	R=1/3			
Décodage de canal	Turbo décodeur (1 itération)			
Modulation	QPSK			
Taille d'(I)FFT	256	2048	256	2048
Longueur du préfixe cyclique (en Symboles QAM)	32	256	32	256
Nombre de sous-porteuses utiles	256/256	2048/2048	256/256	2048/2048
Quantification	10 bits		14 bits	
Trame	1 symbole OFDM pilote tous les 10 symboles OFDM de données			
FPGA	Xilinx Virtex-6 LX240T			
Fréquence d'horloge	50 MHz			

sera de 256 points pour les applications 1 et 3 tandis qu'elle sera de 2048 points pour les applications 2 et 4. Par conséquent, l'espace inter-porteuse $\frac{W}{N} = \Delta_f$ est différent selon la taille de transformée de Fourier inverse effectuée. De plus, la fréquence d'horloge du système est identique pour les 4 applications. A noter aussi que le format de trame et la cible FPGA sont identiques pour les 4 applications.

On résume les hypothèses émises :

- efficacité spectrale identique pour les 4 applications,
- espace inter-porteuse différent

A travers ces hypothèses et la définition de ce scénario, on pourra alors évaluer l'impact de la quantification et du choix de taille d'IFFT sur la consommation totale du circuit ainsi que la métrique d'efficacité énergétique.

5.2.1.1 Évaluation de la puissance dynamique

Après avoir entièrement caractérisé tous les modules individuellement et développé les modèles haut-niveau correspondants, des simulations du système à haut-niveau ont été

effectuées pour les 4 applications considérées. Ainsi, des estimations de puissance dynamique moyenne ont été obtenues.

Ces estimations ont été comparées à celles obtenues en utilisant l’approche classique d’utilisation d’XPA pour le système entier. En effet, le flot de conception a été réalisé pour l’émetteur et le récepteur. Des simulations temporelles (après placement routage) ont permis d’obtenir les activités internes respectives de chaque circuit. Ces fichiers sont requis afin d’obtenir des estimations précises de consommation sous XPA.

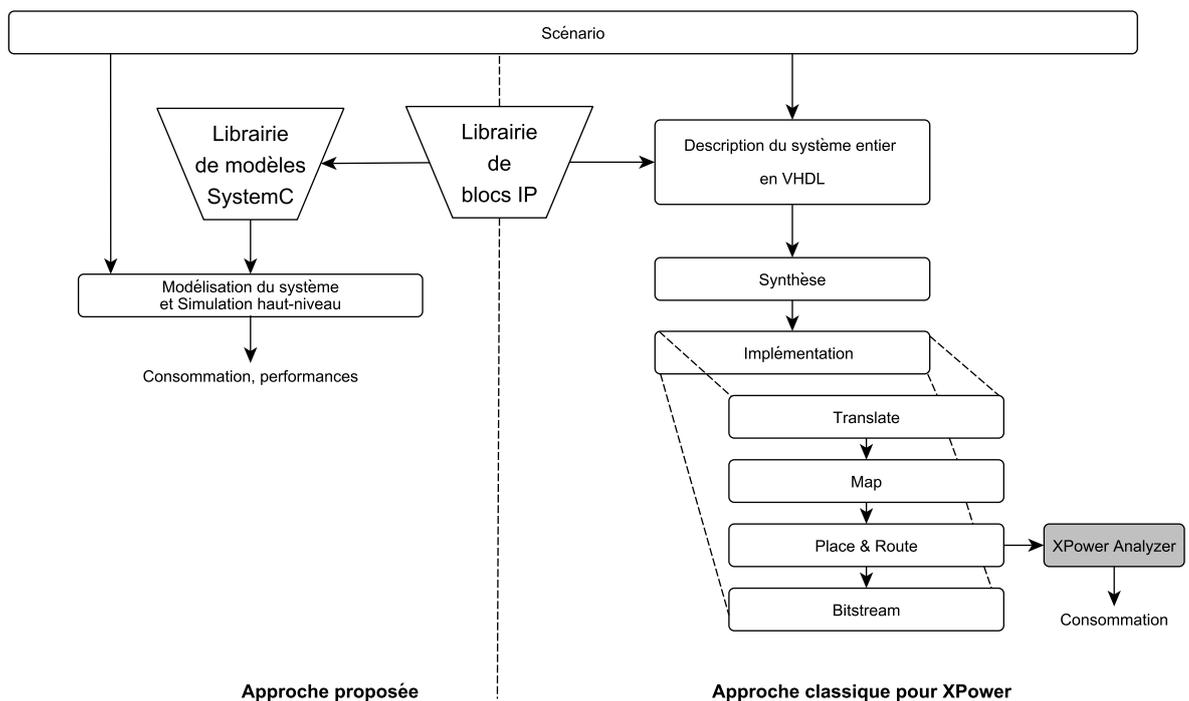


FIGURE 5.2 – Schéma de comparaison de l’approche proposée et celle d’XPower

L’ensemble des résultats est fourni dans le tableau 5.2. Le procédé de comparaison est résumé sur la figure 5.2. A noter que les résultats présentés concernant les puissances dynamiques ne tiennent pas compte des puissances dissipées au niveau des broches d’E/S. En effet, cette puissance dynamique ne doit pas être prise en compte lors de la phase de caractérisation des IPs car elle ne correspond pas à celle du système implanté. D’autre part, le temps de simulation du système a été fixé à 5 ms, afin d’obtenir une valeur moyenne de consommation dynamique réaliste, résultant de la génération d’un certain nombre de symboles OFDM.

On peut constater que les estimations obtenues à l’aide de la méthodologie proposée permettent d’obtenir des estimations précises en comparaison des résultats obtenus à l’aide de l’approche classique d’XPA. En effet, l’erreur maximale pour ces exemples est d’environ 5%. On peut aussi remarquer qu’entre l’application 3 et 4, la puissance dynamique moyenne augmente seulement de 33 mW tandis que la taille de l’IFFT et du préfixe cyclique est multipliée par un facteur 8. De plus, l’utilisation du codage de canal a un grand

TABLE 5.2 – Comparaison des estimations de puissance dynamique par la méthodologie proposée et l’approche classique d’XPower Analyzer

		Méthodologie proposée (mW)	XPA (mW)	Erreur Absolue (%)
Application 1	TX	58.19	55.36	5.07
Application 2	TX	80.7	84.58	4.81
	RX	245.7	244.26	0.59
Application 3	TX	68.76	65.48	5
	RX	266.84	263.52	1.26
Application 4	TX	100.04	98.73	1.32
	RX	308.11	308.38	0.1

impact au niveau du récepteur tandis que l’impact est relativement faible au niveau de l’émetteur.

On remarque aussi que la puissance dynamique évolue de manière significative avec l’augmentation de la quantification des données, notamment lorsque le système devient complexe.

A titre de comparaison avec une approche basée sur les feuilles de calcul, qui consisterait à effectuer la somme des puissances dynamiques moyennes de chaque bloc IP, la puissance dynamique correspondante à l’application 2 a été estimée. Dans ce cas, en faisant uniquement la somme des puissances dynamiques lorsque les blocs IPs sont actifs, c-à-d sans considérer les temps d’activités, la puissance dynamique pour l’émetteur de l’application 2 est de 120.93 mW (au lieu de 84.58 mW par XPA, cf. Table 5.2) et pour le récepteur de 358.55 mW (au lieu de 244.26 mW). Pour l’application 3, la puissance estimée à partir de l’approche cumulée est de 107.6 mW pour l’émetteur (au lieu de 65.48 mW) et de 396.54mW (au lieu de 263.52 mW) pour le récepteur.

A travers cet exemple, on constate que l’approche proposée permet de prendre en compte les spécificités de chaque application. D’autre part, une erreur importante peut être commise en utilisant une approche classique d’estimation en faisant la somme des puissances dynamiques moyennes sans considérer les temps d’activités.

5.2.1.2 Évaluation du temps d’estimation

Un autre point déterminant est le temps nécessaire à l’obtention de ces estimations. L’approche classique d’utilisation de XPA nécessite la réalisation du flot de conception du système entier, ce qui peut s’avérer très long, notamment si l’utilisateur fournit de fortes contraintes sur la réalisation. De plus, la simulation temporelle requiert encore un temps supplémentaire pouvant atteindre plusieurs heures pour simuler quelques millisecondes du fonctionnement d’une seule application.

Comme indiqué précédemment, le temps de simulation a été défini à 5 ms pour les deux approches. Dans le tableau 5.3, les différents temps d’estimation obtenus à l’aide de l’approche classique d’XPower ainsi que pour l’approche proposée sont comparés.

TABLE 5.3 – Comparaison des temps d’estimation de la puissance dynamique par la méthodologie proposée et l’approche classique d’XPower Analyzer pour les émetteurs

		Approche proposée	XPA
Application 1	TX	1.51s	12h09
Application 2	TX	1.67s	19h51
	RX		35h07
Application 3	TX	2.11s	15h46
	RX		30h15
Application 4	TX	1.72s	22h
	RX		50h56

Contrairement à l’approche classique, seulement quelques secondes sont nécessaires pour simuler une application à haut-niveau et effectuer le test d’une autre application. On peut aussi penser que le temps de caractérisation de tous les blocs IP est important et que le gain dans l’évaluation d’un système n’est pas avantageux. L’objectif étant la comparaison de plusieurs systèmes de communications numériques sans fil, un grand nombre de ces blocs est ré-utilisé et le surcoût de caractérisation peut s’avérer relativement faible.

Un autre besoin important pour les concepteurs est de pouvoir déterminer des stratégies d’optimisation ainsi que des choix pertinents concernant l’utilisation de blocs IPs. En reprenant l’exemple de la chaîne SISO-OFDM et de l’application 3, indiquée dans le tableau 5.1, la répartition de la consommation des puissances pour les différents modules a été obtenue et est présentée sur les Figures 5.3 et 5.4.

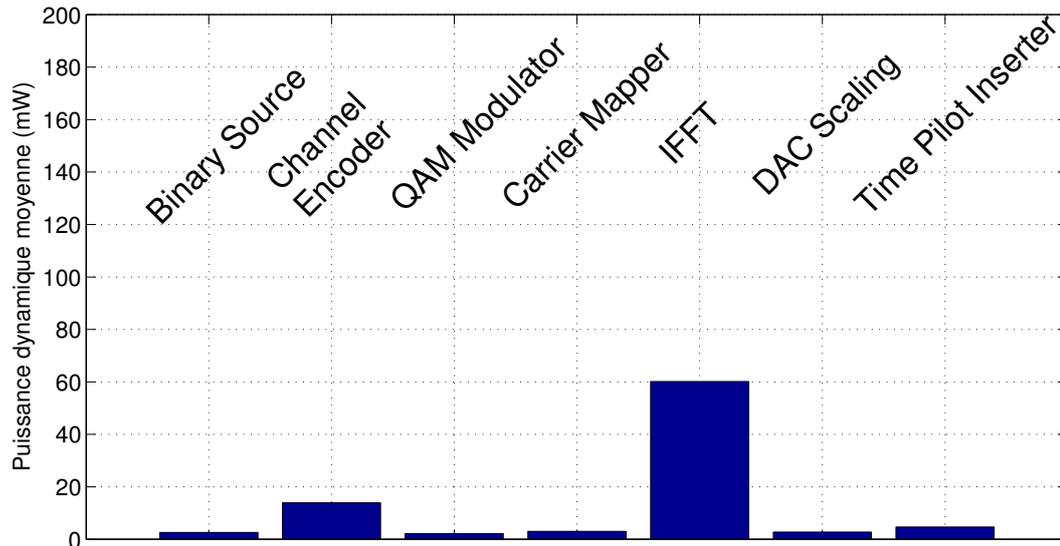


FIGURE 5.3 – Répartition de la consommation par IP pour l’émetteur de la chaîne SISO-OFDM

Les Figures 5.3 et 5.4 permettent respectivement de comparer les consommations des

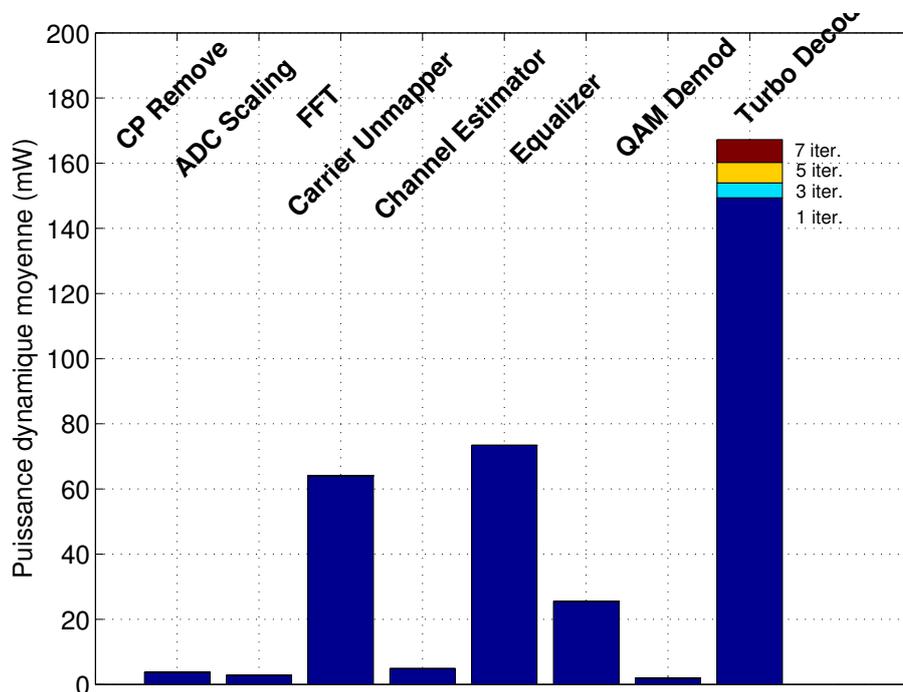


FIGURE 5.4 – Répartition de la consommation par IP pour le récepteur de la chaîne SISO-OFDM

différents modules à l'émission et à la réception.

Sur la Figure 5.3, on peut constater que c'est le bloc IP réalisant l'IFFT qui contribue significativement à la puissance totale dissipée à l'émission. Au niveau du récepteur, on peut constater que c'est le turbo-décodeur qui consomme une grande partie de la puissance totale, suivi par l'estimateur de canal, la FFT et l'égaliseur. On peut également noter que la consommation dynamique moyenne du turbo décodeur augmente avec le nombre d'itérations. A travers ces résultats, les concepteurs peuvent ainsi être guidés dans leurs choix ainsi que dans les stratégies d'optimisation en identifiant les sources majeures de consommation.

5.2.1.3 Évaluation de l'efficacité énergétique

Comme décrit à la sous-section 1.5.2 du chapitre 1, l'efficacité énergétique est une métrique très utilisée dans le domaine des systèmes de communications numériques sans fil. Elle permet d'exprimer à la fois la faculté d'un système à transmettre le maximum d'information tout en considérant la consommation énergétique du système. Cette métrique est généralement évaluée à haut-niveau. Dans la plupart des approches de la littérature, seule la puissance de transmission P_t allouée à la transmission des données est considérée.

Dans cette étude, nous allons à la fois considérer la puissance de transmission allouée

aux données mais aussi tenir compte de la puissance consommée par les traitements en bande de base sur FPGA, **effectués à l'émission**, notée P_{CSISO} . A noter que l'on pourrait considérer l'ensemble des puissances consommées, ou bien encore se focaliser sur les puissances consommées au niveau du récepteur.

On rappelle que l'efficacité énergétique est généralement définie pour un canal SISO de la manière suivante :

$$EE(bit/J) = \frac{C_{SISO}}{P_{TotaleSISO}} = \frac{W.E[\log_2(1 + \frac{|h|^2 P_t}{N_0 W P_L})]}{P_{TotaleSISO}} \quad (5.1)$$

avec C_{SISO} , la capacité moyenne en bit/s en configuration SISO, P_t la puissance moyenne allouée à la transmission des données ($W=J/s$), W la bande de fréquence (Hz), N_0 la densité spectrale mono-latérale de bruit, P_L les pertes de propagation, h les coefficients d'évanouissement (fading) du canal générés suivant une loi gaussienne et où :

$$P_{TotaleSISO} = P_t + P_{CSISO} \quad (5.2)$$

avec P_{CSISO} la puissance dynamique moyenne consommée par les traitements en bande de base à l'émission effectués sur cible FPGA en configuration SISO.

Si l'on souhaite évaluer l'EE à partir de la capacité, il est important de considérer qu'un biais est introduit. En effet, la capacité ne permet pas de tenir compte du schéma de modulation et du codage défini. En effet, la capacité croît avec l'augmentation de la puissance de transmission tandis que l'ES converge vers une valeur maximale imposée par le MCS. Par exemple, pour une modulation QPSK et un rendement de codage d'1/3, l'ES maximale est de 0.66bit/s/Hz. Une autre manière d'évaluer l'EE est alors la suivante :

$$EE(bit/J) = \frac{V}{P_{TotaleSISO} * T} \quad (5.3)$$

avec V , un volume de données à transmettre, nécessitant le temps T et $P_{TotaleSISO}$ la puissance totale consommée comme précédemment.

Evaluer l'EE à partir de l'équation 5.3 conduit à des résultats d'EE plus précis qu'à partir de la capacité. Ainsi, pour un MCS donné, l'EE est :

$$EE(bit/J) = \frac{W.min(E[\log_2(1 + \frac{|h|^2 P_t}{N_0 W P_L})], ES(MCS))}{P_{TotaleSISO}} \quad (5.4)$$

avec $ES(MCS)$ l'efficacité spectrale maximale obtenue pour un MCS donnée, par exemple $ES(QPSK,1/3)=0.66bit/s/Hz$.

- Hypothèses

Dans cette étude, nous avons considéré une fréquence d'horloge égale à 20MHz. De cette manière, la fréquence du système correspond à la bande de fréquence totale utilisée, soit

$W = 20MHz$. Les puissances consommées par le circuit pour les différentes applications à cette fréquence sont données en Annexe A.1. Celles-ci sont plus faibles car la puissance dynamique diminue linéairement avec la fréquence.

De plus, nous avons choisi un modèle de *path loss* suivant : $P_L(dB) = 16.9 * \log_{10}(d) + 32.8 + 20 * \log_{10}(fc)$

avec fc la fréquence porteuse (GHz) et d la distance émetteur récepteur (m).

Ce modèle de *path loss* permet de traduire les pertes de propagation dans un environnement de type *indoor* en femto-cell. De plus, ce modèle est valide pour une distance d entre un émetteur et un récepteur en ligne de vue (*Lign-of-Sight*, LOS). Dans notre étude, nous avons choisi une distance d de 50m, ce qui correspond à une valeur standard dans ce type d'environnement. Ce modèle est notamment utilisé dans le standard LTE (cf. 3GPP TR 36.914 v9.0.0 page 93 [3GP10]) et est valide pour des fréquences comprises entre 2 et 6GHz et des distances d , allant de 10 à 150m. Nous avons donc fixé la fréquence porteuse à 2.4GHz, valeur standard en LTE et en Wifi. La densité spectrale mono-latérale de bruit a été fixée à -174 dBm/Hz.

Il est important de considérer qu'il existe un SNR minimal requis en réception afin de décoder correctement l'information. Ce SNR impose une valeur minimale de puissance de transmission. D'après [B⁺10] à la page 30, les auteurs ont déterminé que le SNR minimal requis est de 10dB pour une modulation QPSK et un rendement de codage d'1/3. A partir de cette valeur, nous avons pu déterminer la puissance de transmission correspondante. Pour notre scénario, la puissance de transmission minimale est de -22dBm. Pour des puissances de transmission supérieures à -22dBm, le SNR de réception sera donc supérieur au SNR minimal requis. Nous avons donc défini une plage d'étude pour laquelle la puissance de transmission varie de -20dBm à 25dBm. De cette manière, les conditions seront requises pour permettre un décodage de l'information en réception. D'autre part, on rappelle que la puissance maximale de transmission allouée en LTE pour une station de base (femtocell) est de 20dBm.

Dans les paragraphes suivants sont présentés les résultats d'EE pour les différentes applications données. A noter que les résultats d'EE obtenus par à partir des équations 5.4 et 5.3 conduisent à des résultats similaires. Les résultats présentés dans les figures suivantes sont ceux calculés à partir de l'équation 5.3.

Sur la figure 5.5-a, on peut observer l'efficacité énergétique (bit/J) en fonction de la puissance totale consommée (dBm) lorsque la puissance consommée par le circuit n'est pas considérée, c-à-d $P_{CSISO} = 0$, pour les différentes applications considérées (cf. tableau 5.1). La puissance totale consommée est donc égale à la puissance de transmission. On peut vérifier que l'EE est la même pour les différentes applications. L'EE décroît avec l'augmentation de la puissance de transmission, P_t .

Sur la figure 5.5-b, on observe l'EE en fonction de la puissance totale consommée sauf que cette fois-ci la puissance consommée par le circuit est considérée durant l'évaluation de l'EE.

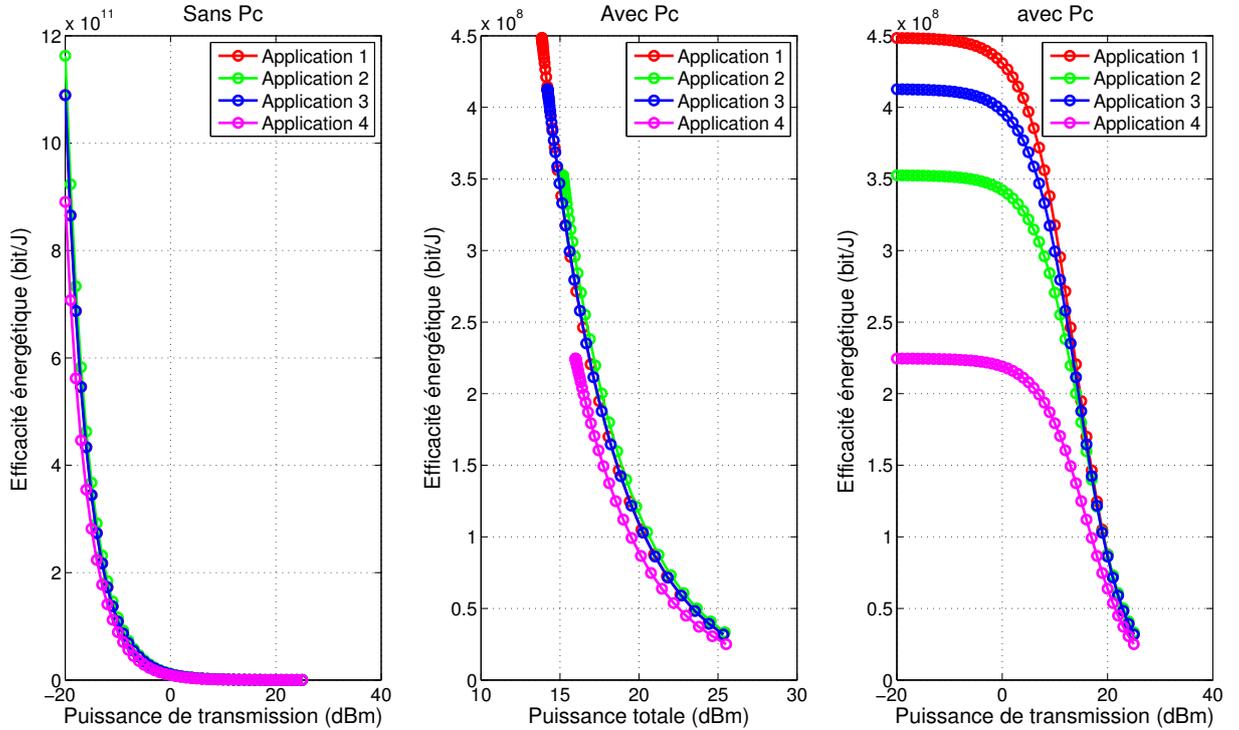


FIGURE 5.5 – Efficacité énergétique moyenne (bit/J) en fonction de la puissance totale consommée (dBm) et de la puissance de transmission

En comparant les figures 5.5 a et b, on peut constater que prendre en compte la puissance consommée par le circuit a un impact significatif sur l'EE. En effet, on observe les courbes se décalent suivant l'axe des abscisses de manière d'autant plus significative que la puissance consommée par le circuit est grande. En effet, l'application 1 possède la consommation la plus faible suivie par l'application 3 puis 2 et 4.

Sur la figure 5.5-c, l'EE est représentée en tenant compte de la puissance consommée par le circuit en fonction de la puissance de transmission uniquement. Elle permet de comparer l'EE des différentes applications pour une puissance de transmission donnée. On peut remarquer que :

1. pour des puissances de transmission faibles, $P_t < 0\text{dBm}$, l'EE des différentes applications tend vers une même valeur,
2. pour des puissances de transmissions supérieures, $P_t > 0\text{dBm}$, toutes les EE diminuent et convergent vers 0.

Pour des puissances de transmission faibles ($P_t < 0\text{dBm}$), l'EE varie peu. En effet, étant donné que la puissance de transmission à l'émission permet d'obtenir un SNR supérieur au SNR minimal requis, l'EE est donc égale à $\frac{V}{P_{TotaleSISO} * T}$. L'EE décroît ensuite à cause

de l'augmentation de la puissance de transmission. A noter qu'à faibles puissances de transmission, il n'est plus possible de discriminer les différentes applications par la prise en compte de la puissance consommée par le circuit.

Remarque : L'EE n'est jamais nulle car le SNR minimal requis en réception est assurée pour la gamme de puissance de transmission étudiée. On peut donc considérer que les résultats d'EE obtenus permettent de déterminer l'EE optimale pour les différentes applications. Ces valeurs ont été obtenues pour un MCS donné. En perspective, il serait intéressant d'évaluer les EE optimales pour différents MCS et ainsi tracer la courbe d'EE pour le système donné en fonction de nombreux MCS.

Au regard du choix de quantification, l'impact sur l'EE est observable entre les courbes rouges et bleues. L'application 3 possède une EE légèrement inférieure à l'application 1 à cause de l'augmentation de la quantification des données.

L'impact du choix de la taille d'IFFT peut être évalué en mesurant les différences d'EE entre les applications 1 et 2 et puis entre les applications 3 et 4, c-à-d pour des quantifications de données respectivement de 10 et 14 bits. On constate alors que pour une bande de fréquence donnée (ici 20MHz), l'impact de l'augmentation d'un facteur 8 de la taille d'IFFT a pour effet de diminuer l'EE d'un facteur d'environ 1.28 en moyenne, à faibles puissances de transmission.

D'autre part, si un concepteur de système possède une puissance totale à ne pas dépasser, au regard de la figure 5.5-b, on peut alors discriminer plusieurs choix de configurations afin de respecter le budget énergétique tout en satisfaisant le critère d'EE. Par exemple, si un concepteur a pour critère une puissance totale consommée notée P_{max} valant 16dBm (environ 40mW), alors l'application 4 pourrait être écartée et le choix se porterait alors sur une des trois autres applications.

A partir de cette étude, on peut conclure qu'il est nécessaire de prendre en compte la puissance dynamique consommée par le circuit, tout comme la bande utilisée et autre paramètre système. Cette prise en compte est d'autant plus importante que les puissances de transmissions sont faibles. Des résultats réalistes peuvent être ainsi obtenus et une comparaison efficace de plusieurs applications peut être effectuée. Finalement, les concepteurs du système peuvent choisir le meilleur compromis entre les performances et la puissance dynamique moyenne consommée.

5.2.2 Émetteur LTE SISO-OFDM en voie descendante du canal PDSCH

Dans le cadre d'une nouvelle étude, un émetteur simplifié LTE en configuration SISO, intégrant les traitements de la bande de base en voie descendante du canal des données utilisateurs (PDSCH *Physical Downlink Shared CHannel*), a été développé conjointement en VHDL et à l'aide de blocs IPs de Xilinx. Comme indiqué à la Figure 5.6, les différents modules correspondant aux traitements du canal PDSCH portant les données utilisateurs ont été développés.

L'émetteur est composé d'un codeur de canal, bloc IP de Xilinx appelé *DL Channel Encoder* [Xil11b]. Ce bloc IP permet l'ajout du contrôle de redondance cyclique (*Cyclic Redundancy Check* (CRC)) aux blocs de transport des données utilisateurs. Ensuite, il réalise la segmentation des blocs de transport, leur encodage ainsi que le *rate matching*.

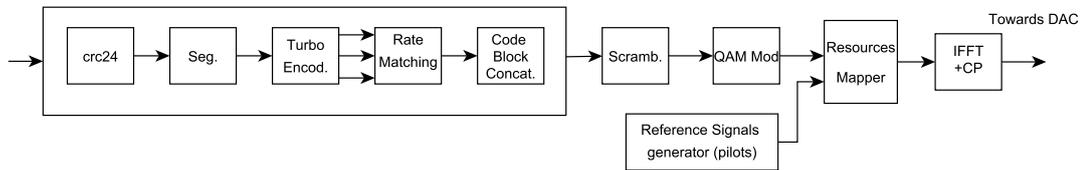


FIGURE 5.6 – Émetteur LTE SISO-OFDM en voie descendante

Ensuite, les données sont embrouillées à l’aide d’une séquence pseudo-aléatoire (*scrambling*). Les données sont ensuite modulées en symboles complexes I/Q. Parallèlement, les signaux de référence sont générés et le bloc nommé *Resource Mapper* permet l’allocation des ressources dans la trame LTE. Finalement, la modulation OFDM est réalisée et le préfixe cyclique est inséré.

Pour ce système, certaines hypothèses ont été posées :

1. le système supporte une structure de trame LTE de type 1 (cf. chapitre 1, section 1.4),
2. l’identifiant de la cellule a été fixé pour la génération des signaux de référence,
3. tous les blocs de ressources sont utilisés par des données ou des pilotes, les autres éléments de ressources sont délibérément forcés à 0.

En effet, la couche physique du LTE dépend d’un nombre important de paramètres, provenant des couches supérieures (MAC, etc.). Ainsi, des paramètres comme l’identifiant cellule ou encore l’identifiant RNTI (*Radio Network Temporary Identifier*) pour l’équipement utilisateur ont été délibérément fixés afin de limiter le cadre des développements de la couche physique.

5.2.2.1 Évaluation de la puissance dynamique

Un nouveau scénario a donc été défini ainsi que 4 nouvelles applications comme indiqué dans le tableau 5.4. Les différentes applications ont été définies en accord avec les paramètres de la couche physique définis dans la norme LTE (cf chapitre 1. Ainsi, toutes les sous-porteuses du spectre ne sont pas allouées et leur nombre varie en fonction de la bande utilisée, tout comme la taille des transformées de Fourier.

Finalement, les différents blocs IP du système ont été caractérisés et les modèles haut-niveau ont été décrits en SystemC. Le temps de simulation du système a été fixé de manière à correspondre à la génération de 70 symboles OFDM, c’est-à-dire 10 slots ou 5 sous-trames LTE. On notera aussi que la fréquence d’horloge pour tous les blocs IPs est de 50MHz.

A titre de comparaison avec les émetteurs décrits précédemment (cf section 5.2.1), le nombre de sous-porteuses utiles varie ainsi que la taille du préfixe cyclique.

Les estimations de puissance dynamique obtenues à l’aide de la méthodologie sont comparées à celles d’XPA comme pour l’exemple précédent. Les résultats sont reportés

TABLE 5.4 – Scénario et 4 applications de la chaîne SISO-OFDM en LTE DL

Scénario	Appli. 1	Appli. 2	Appli. 3	Appli. 4
Codage de canal	Rendement=1/3 Taille de bloc = 1024 bits			
Modulation QAM	QPSK			
Taille (I)FFT	256	512	1024	2048
Préfixe Cyclique	cas normal : $5.21\mu s$ 1er symbole OFDM, $4.67\mu s$ pour les 6 autres			
Sous-porteuses utiles	180	300	600	1200
Quantification des données	14 bits			
FPGA	Xilinx Virtex-6 LX240T			
Fréquence d'horloge	50 MHz			
Temps de simulation	Génération de 70 symboles OFDM (5 sous-trames)			

dans le tableau 5.5 et affichés à la figure 5.7.

TABLE 5.5 – Estimation de la puissance dynamique de la chaîne SISO-OFDM en LTE DL

	Méthodologie Proposée (mW)	XPA (mW)	Erreur Abs. (%)
Application 1	83.63	77.14	10.83
Application 2	100.38	101.54	1.14
Application 3	127.41	127.15	0.21
Application 4	143.27	139.89	2.41

A l'aide de notre méthodologie, l'erreur maximale absolue est d'environ 10%, ce qui reste relativement faible, de part les faibles valeurs de consommation (erreur maximale d'environ 6 mW). Il peut aussi être remarqué que l'application 4 consomme seulement 60mW (70% d'augmentation) de plus que l'application 1 tandis que la bande de fréquence de l'application 4 est 8 fois plus grande que pour l'application 1.

5.2.2.2 Évaluation des temps d'estimation

Le temps d'estimation a été évalué pour la méthodologie proposée ainsi que pour l'approche classique d'XPA. D'après le tableau 5.6, on peut noter que le facteur d'accélération pour estimer la consommation du système entier est très important, d'ordre 3 et 4.

Ce gain sera d'autant plus important si l'utilisateur veut explorer d'autres jeux de paramètres.

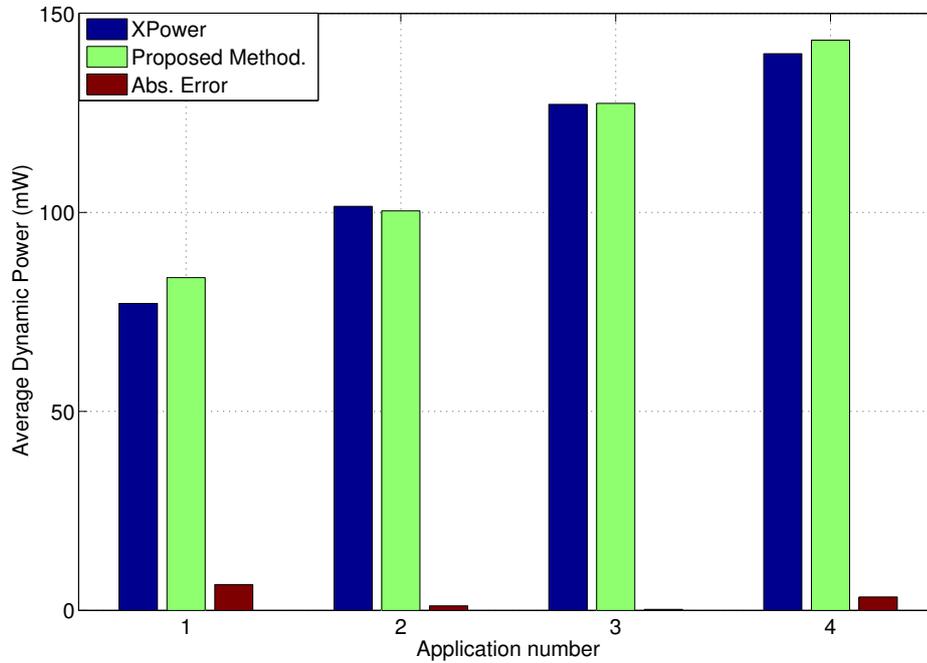


FIGURE 5.7 – Résultats d’estimation entre notre méthodologie et XPower pour un émetteur SISO-OFDM dans le contexte LTE

TABLE 5.6 – Comparaison des temps approximatifs d’estimation

	Méthodologie Proposée	XPA	Facteur d’accélération
Application 1	1.18s	1h12	x3661
Application 2	1.42s	2h45	x6971
Application 3	2.45s	8h16	x12146
Application 4	3.91s	15h30	x14271
Total	8.96s	99780s	x11136

5.2.2.3 Évaluation de l’efficacité énergétique

Au regard des résultats précédents, on peut, de prime abord, penser que c’est l’application consommant le moins de puissance dynamique qui possédera la meilleure efficacité énergétique, d’autant plus que l’application 1 est moins complexe et utilise moins de ressources que les autres applications. On rappelle que le choix de taille d’IFFT a un impact significatif sur la consommation du système.

– Hypothèses

Durant l'évaluation de l'efficacité énergétique, nous avons considéré le même modèle de *path loss* que pour l'émetteur SISO-OFDM précédent ($P_L(dB) = 16.9 * \log_{10}(d) + 32.8 + 20 * \log_{10}(fc)$) et les mêmes paramètres N_0 et fc .

D'autre part, le standard LTE prévoyant des fréquences d'échantillonnage multiples de 3.84MHz, les fréquences d'horloge des différentes applications seront donc de 3.84MHz, 7.68MHz, 15.26MHz et 30.52MHz, respectivement. Les résultats des consommations dynamiques des différentes applications pour ces fréquences sont donnés en Annexe A.2.

Sur la figure 5.8, l'efficacité énergétique est affichée en fonction de la puissance moyenne totale consommée et de la puissance de transmission en tenant compte ou non de la puissance consommée par le circuit. Afin de prendre en compte la limitation de l'ES par le MCS utilisé, l'EE a été évaluée à partir de l'équation 5.4. On rappelle la formule de l'efficacité énergétique donnée à l'équation 5.4 :

$$EE(\text{bit}/J) = \frac{\min(C_{SISO}, ES)}{P_{Totale_{SISO}}} = \frac{W \cdot \min(E[\log_2(1 + \frac{|h|^2 P_t}{N_0 W P_L})], ES(MCS))}{P_t + P_{CSISO}} \quad (5.5)$$

avec C_{SISO} , la capacité moyenne en bit/s en configuration SISO, ES l'efficacité spectrale en bit/s/Hz dépendante du MCS choisi, P_t la puissance moyenne allouée à la transmission des données (W), W la bande de fréquence (Hz), N_0 la densité spectrale de puissance mono-latérale du bruit (W/Hz), h les coefficients d'évanouissement du canal et P_L les pertes de propagation du canal.

Sur la Figure 5.8-a, l'EE est calculée lorsque P_{CSISO} n'est pas considérée. On peut constater que l'EE de l'application 4 est supérieure à celle des autres applications quelle que soit la puissance allouée à la transmission des données P_t . En effet, la capacité augmentant avec la bande de fréquence W , c'est l'application 4 qui possède la plus meilleure efficacité énergétique. L'EE décroît ensuite avec l'augmentation de la puissance de transmission.

Sur la Figure 5.8-b, l'EE est tracée en fonction de la puissance totale consommée lorsque la puissance dynamique consommée par le circuit, P_{CSISO} , est considérée. Si un concepteur connaît la puissance totale maximale que son application ne doit pas dépasser, alors des choix d'application peuvent être facilement effectués.

Sur la Figure 5.8-c, l'EE est affichée en fonction de la puissance de transmission, P_t . On peut constater qu'à faible puissance de transmission, c-à-d $P_t < 9dBm$, l'application 1 est plus efficace énergétiquement. Ceci s'explique notamment par une puissance consommée par le circuit qui est moins importante que pour les autres applications. Cette observation est à corréliser avec la taille d'IFFT, qui a un fort impact sur la consommation du circuit. A partir d'une puissance de transmission supérieure à 9dBm, on peut constater que c'est l'application 4 qui possède la meilleure efficacité énergétique. En effet, au fur et à mesure de l'augmentation de la puissance allouée à la transmission des données, les capacités des

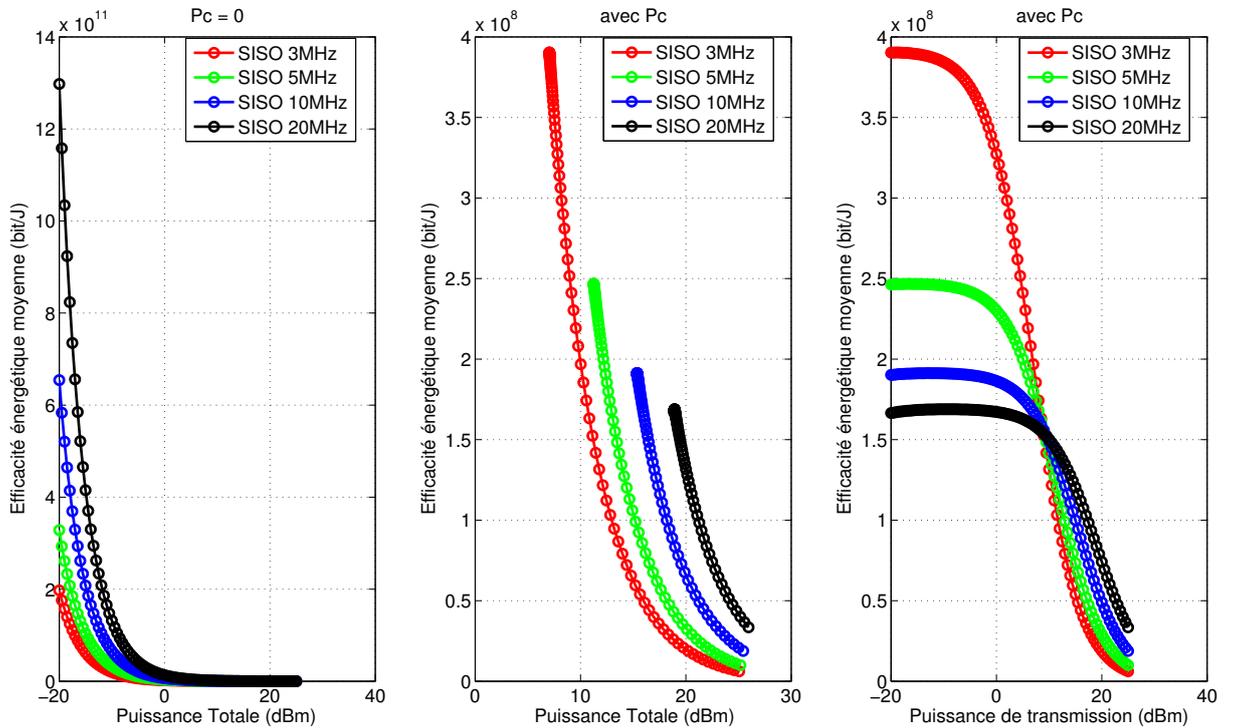


FIGURE 5.8 – Efficacité énergétique moyenne (bit/J) en fonction de la puissance totale consommée (dBm) et de la puissance de transmission (dBm)

applications augmentent. A partir d'une puissance de transmission de 9dBm, l'augmentation de la capacité vient contre-balancer la puissance consommée par le circuit. Dès lors, l'EE devient meilleure pour l'application 4 que pour les autres applications car sa capacité est plus grande (bande de fréquence plus large). Finalement, si la puissance de transmission continue d'augmenter alors les EE de toutes les applications décroissent.

A travers ces résultats, on souligne le bénéfice de considérer à la fois la consommation du circuit et d'autres métriques de performances comme la capacité, l'efficacité spectrale. Si l'on ne considère pas la puissance consommée par le circuit, l'application 4 aurait été choisie car elle possède la meilleure EE. Néanmoins, lorsque celle-ci est considérée dans le calcul de l'EE, on constate qu'à faible puissance de transmission, il serait plus efficace énergétiquement de choisir l'application 1 tandis qu'à plus fortes puissances de transmission, l'application 4 serait à privilégier.

En conclusion, la prise en compte de la puissance dynamique consommée par les traitements en bande de base est d'autant plus importante que la comparaison des systèmes s'effectue à faible puissance de transmission.

5.2.3 Émetteur LTE MISO-OFDM 2x1 en voie descendante avec encodage d'Alamouti

Dans cette nouvelle étude, un émetteur MISO-OFDM 2x1 dans le contexte LTE est évalué. Comme présenté à la Figure 5.9, l'émetteur est l'extension de la configuration SISO mais avec 2 antennes d'émission (cf. chapitre 1, section 1.4). De plus, un code espace-temps de type Alamouti (*Space Frequency Block Coding* SFBC) est appliqué dans le domaine fréquentiel avant l'étape de modulation OFDM. On peut aussi constater que les signaux de référence ne sont pas utilisés dans ce système. Les éléments de ressources qui ne sont pas alloués pour les données des utilisateurs (PDSCH) sont délibérément forcés à 0. Néanmoins, tous les blocs sont configurés de manière à respecter le standard LTE.

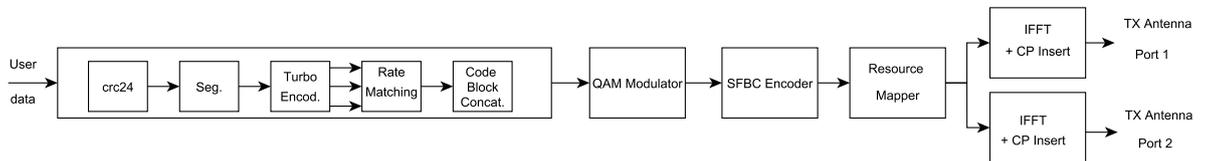


FIGURE 5.9 – Traitements en bande de base du canal PDSCH en LTE DL dans une configuration MISO 2x1 avec codage d'Alamouti

5.2.3.1 Évaluation de la puissance dynamique

Un scénario et 4 applications ont été définis et sont identiques à ceux utilisés dans la configuration SISO (cf. table 5.4).

A travers le tableau 5.7, on peut constater que l'erreur maximale absolue reste faible, inférieure à 4% pour les 4 applications considérées, bien que la complexité du système soit accrue par rapport à la configuration SISO.

TABLE 5.7 – Résultats d'estimation pour les 4 applications

	Méthodologie proposée (mW)	XPA (mW)	Erreur Abs. (%)
Appli. 1	122.72	118.64	3.44
Appli. 2	163.30	159.01	2.7
Appli. 3	196.22	195.07	0.59
Appli. 4	222.11	227.01	2.16

La répartition de la consommation dynamique moyenne par module est visible sur la Figure 5.10. Étant donné les spécificités du LTE, cette répartition de puissance a été tracé

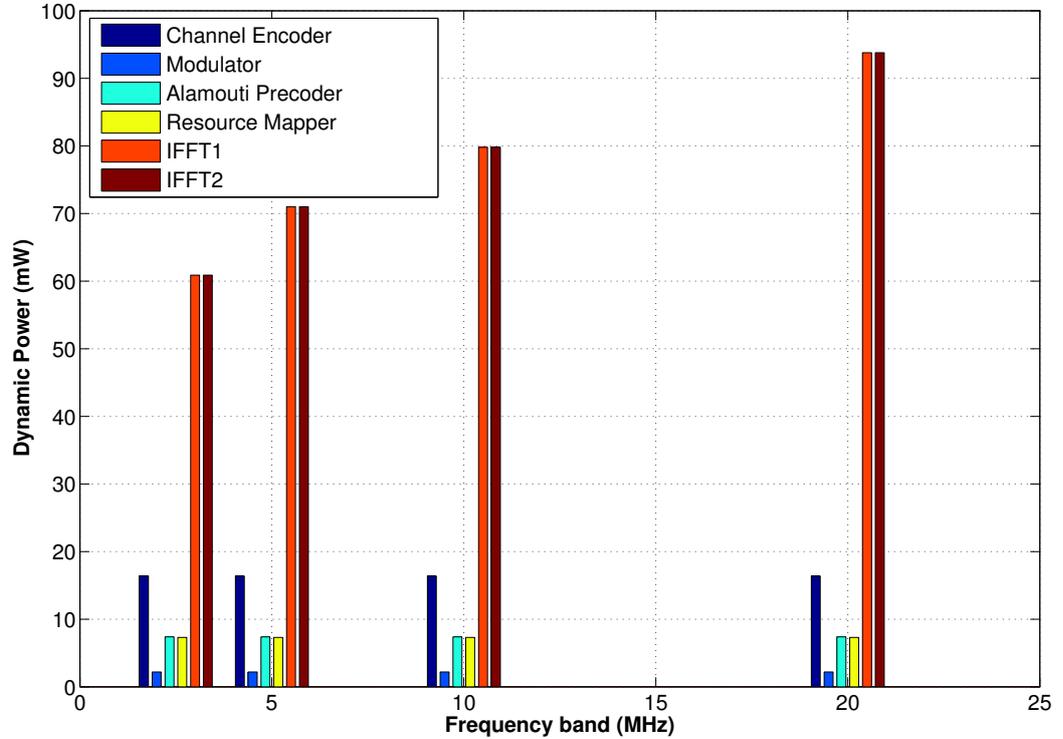


FIGURE 5.10 – Répartition de la puissance dynamique consommée de l'émetteur MISO-OFDM 2x1 en fonction de la bande de fréquence

en fonction de la bande de fréquence utilisée.

On peut remarquer que pour chaque système, ce sont les deux IFFTs qui contribuent de manière significative à la consommation totale par rapport aux autres modules de l'émetteur. D'ailleurs, il est intéressant de remarquer que les consommations de puissance moyenne des IFFTs évoluent de façon logarithmique en fonction de la bande de fréquence. Ce comportement est directement corrélé à l'évolution de la complexité calculatoire de la transformée de Fourier. On peut aussi noter que l'impact des autres modules d'émission est relativement faible sur la consommation dynamique totale.

A la suite de ces remarques, il peut alors être envisagé de remplacer le bloc IP IFFT ou bien de le modifier/configurer autrement de manière à diminuer son impact sur la consommation dynamique totale du système.

5.2.3.2 Évaluation des temps d'estimation

Dans le tableau 5.8 sont relevés les temps d'estimation pour ce nouveau système et les 4 applications. L'estimation à haut-niveau est grandement accélérée. En effet, les facteurs d'accélération sont d'ordre 3 ou 4. Plusieurs heures sont requises pour effectuer les simu-

lations temporelles ainsi que les estimations à l'aide d'XPA tandis que quelques secondes suffisent à l'aide de l'approche proposée. D'autre part, la complexité croissante augmente légèrement le temps de simulation en comparaison avec le système à une antenne d'émission.

TABLE 5.8 – Comparaison des temps d'estimation

	Méthodologie proposée	XPA	Facteur d'accélération
Appli. 1	1.25s	2h25	x6960
Appli. 2	2s	6h53	x12390
Appli. 3	3.42s	14h	x14736
Appli. 4	6.65s	27h	x14616

5.2.4 Émetteur d'une chaîne MISO-OFDM à X antennes avec pré-codage de retournement temporel en fréquence

Dans le cadre d'un nouvel exemple de validation de la méthodologie, nous avons évalué l'efficacité énergétique d'un émetteur OFDM intégrant le principe du retournement temporel en fréquence. Plusieurs configurations ont été évaluées, dans lesquelles le nombre d'antennes a varié comme indiqué dans la figure 5.11.

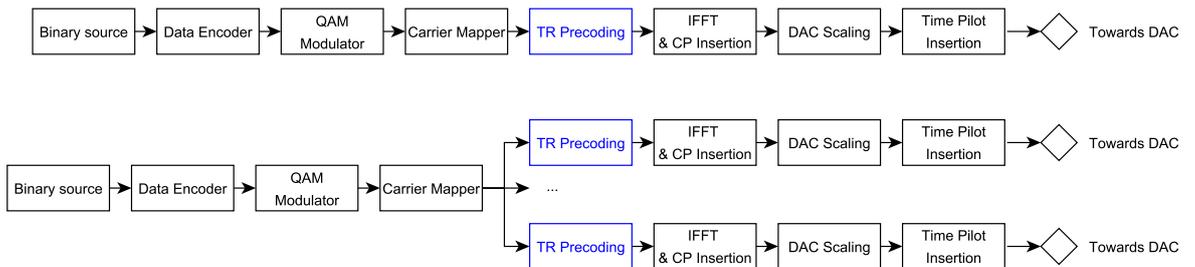


FIGURE 5.11 – Schématisation d'émetteurs SISO et MISO-OFDM avec Retournement Temporel appliqué dans le domaine fréquentiel

De manière à comparer efficacement les différents systèmes selon la méthodologie, un scénario a été défini et est résumé dans le tableau 5.9. De plus, nous avons choisi de fixer la puissance moyenne allouée à la transmission des données pour l'ensemble des configurations, comme indiqué sur la figure 5.12. De cette manière, si une puissance moyenne allouée à la transmission des données notée P_t est égale à P en configuration SISO, chaque antenne aura une puissance moyenne allouée à la transmission des données de $\frac{P}{Nt}$ en configuration MISO à Nt antennes d'émission.

TABLE 5.9 – Scénario et application

Scénario	Application 1
Rendement de codage	R=1/3
Taille des blocs	1024
Modulation QAM	QPSK
(I)FFT	256
Taille du Préfixe Cyclique	32
Sous-porteuses utiles	256/256
Quantification des données	14 bits
FPGA	Xilinx Virtex-6 LX240T
Fréquence d'horloge	50 MHz

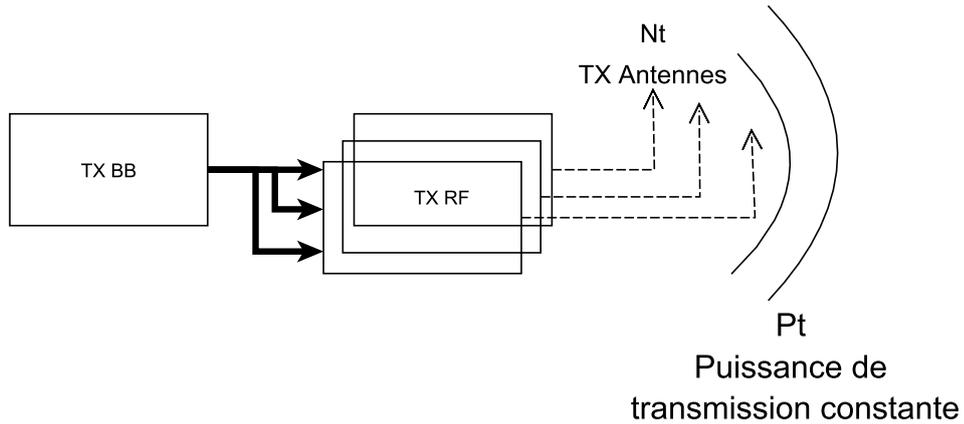


FIGURE 5.12 – Illustration de la méthode d'évaluation de l'EE à P_t constante

En rappel de l'équation de l'efficacité énergétique en configuration SISO 5.1, l'EE en configuration MISO est définie de la manière suivante :

$$EE(\text{bit}/J) = \frac{C_{TR}[\text{bit}/s]}{P_{Totale_{TR}}[W = J/s]} = \frac{W.E[\log_2(1 + \frac{(\sum_{n=0}^{N_T} |h|^2)^2 P_t}{N_0 W N_t P_L})]}{P_{Totale_{TR}}} \quad (5.6)$$

avec C_{TR} , la capacité en bit/s en RT, P_t la puissance moyenne allouée à la transmission des données (W), W la bande de fréquence (Hz), N_t le nombre d'antennes de transmission, W la bande de fréquence (Hz), N_0 la densité spectrale mono-latérale du bruit (W/Hz), P_L les pertes de propagation dans le canal, h les coefficients d'évanouissement du canal (*fading*) et où :

$$P_{Totale_{TR}} = P_t + P_{cRT} \quad (5.7)$$

avec P_{cRT} la puissance dynamique moyenne consommée par les traitements en bande de base effectués sur cible FPGA dans le cas MISO à base de RT.

Les estimations de la consommation dynamique des systèmes considérés ont été obtenues à l'aide de la méthodologie proposée.

– **Hypothèses**

Pour cette étude, des hypothèses ont été fixées. Tout d'abord, nous avons choisi de considérer $\frac{1}{(N_0.W.P_L)} = 1$. Cette hypothèse a été vérifiée pour $W = 50\text{MHz}$ et $N_0 = -174\text{dBm/Hz}$ ainsi qu'avec plusieurs modèles de *Path loss* utilisés notamment dans standard LTE [ETS10]. Cette hypothèse est compatible avec des distances émetteurs-récepteurs inférieures au kilomètre. Ainsi, l'équation précédente devient :

$$EE(\text{bit}/J) = \frac{C_{TR}[\text{bit}/s]}{P_{Totale_{TR}}[W = J/s]} = \frac{W.E[\log_2(1 + \frac{(\sum_{n=0}^{NT} |h|^2)^2 P_t}{Nt})]}{P_{Totale_{TR}}} \quad (5.8)$$

A noter aussi que la gamme d'étude de la puissance de transmission est très importante afin d'identifier le comportement générale de l'EE mais qu'elle pourrait se limiter à des puissances de transmission allant de -20 dBm à 20 dBm, comme pour l'étude de la chaîne SISO-OFDM étudiée en section 5.2.1.3.

Sur la Figure 5.13, l'EE est représentée en fonction de la puissance totale consommée lorsque la puissance dynamique moyenne consommée par les traitements en bande de base est négligée c-à-d $P_{CSISO} = P_{CRT} = 0$. Plusieurs configurations ont été évaluées :

- émetteur SISO à 1 antenne,
- émetteur SISO à 1 antenne avec filtre de RT,
- émetteur MISO à 2 antennes avec filtre de RT,
- émetteur MISO à 4 antennes avec filtre de RT,
- émetteur MISO à 8 antennes avec filtre de RT et
- émetteur MISO à 16 antennes avec filtre de RT.

Sur la Figure 5.14, l'EE est représentée en fonction de la puissance de transmission lorsque cette fois-ci, P_{CSISO} et P_{CRT} sont considérées.

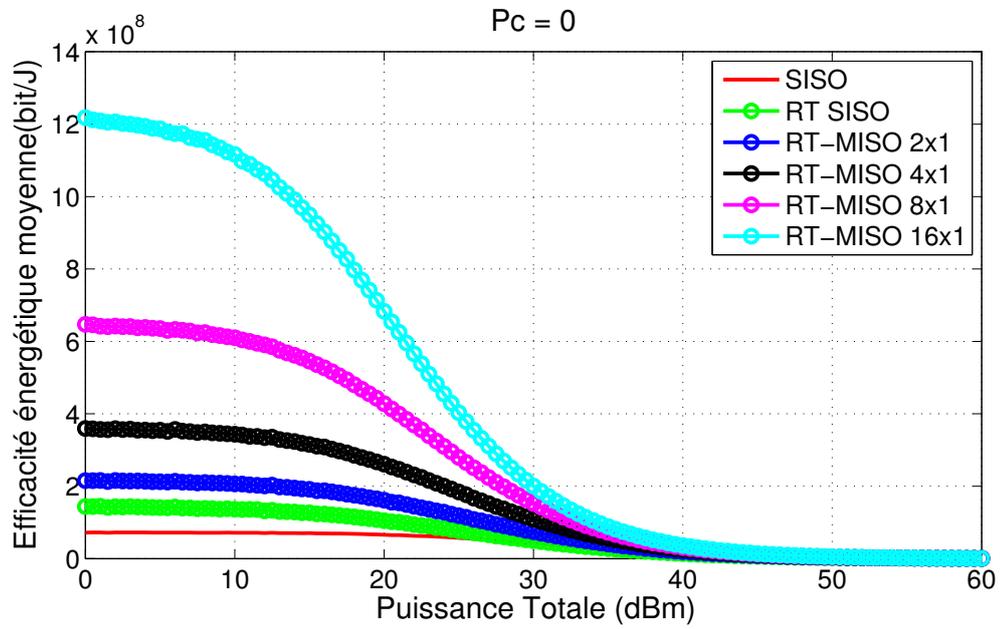


FIGURE 5.13 – Efficacité énergétique vs puissance totale consommée pour plusieurs configurations avec P_{cSISO} et $P_{cRT} = 0W$

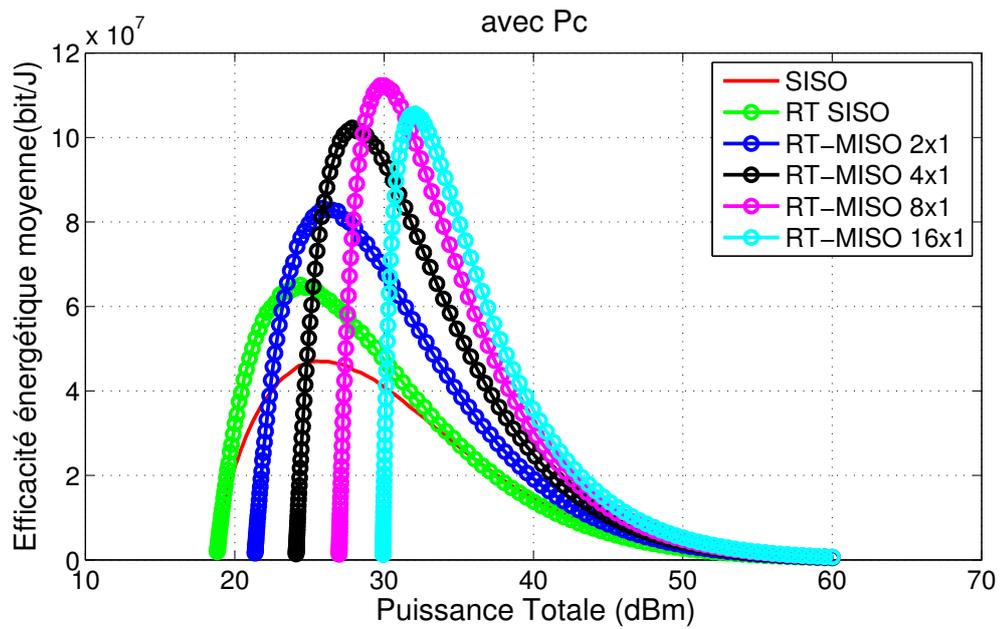


FIGURE 5.14 – Efficacité énergétique vs puissance totale consommée pour plusieurs configurations avec P_{cSISO} et P_{cRT} considérées

A partir des Figures 5.13 et 5.14, on peut noter que la puissance dynamique moyenne consommée par les traitements en bande de base effectués sur la cible FPGA, a un impact important sur l'efficacité énergétique en fonction de la puissance totale consommée.

Tout d'abord, sur la Figure 5.13 lorsque la puissance consommée par le circuit n'est pas considérée, c-à-d $P_c = 0$, l'EE du système à 16 antennes d'émission est la meilleure, quelle que soit la puissance de transmission. En effet, la capacité est la plus importante dans cette configuration.

En Figure 5.14, on peut alors constater que lorsque la puissance consommée par les traitements en bande de base est considérée, différents maximums d'EE sont obtenus. Tout d'abord, on note que l'intégration de la technique de RT apporte un gain en EE en comparaison avec la configuration SISO sans RT. Bien qu'un filtre de RT soit ajouté dans le système RT-SISO, conduisant à une puissance dynamique légèrement plus élevée, l'EE d'un système à base de RT reste meilleure grâce à une plus grande capacité (cf. Figure 5.13). On peut aussi constater sur la Figure 5.14 que les courbes d'EE sont décalées sur l'axe des abscisses à cause de la prise en compte de la puissance consommée par le circuit. D'autre part, on note que lorsque les puissances de transmission sont faibles, l'augmentation de capacité fait croître l'EE des différents systèmes jusqu'à une valeur maximale. Pour une puissance de transmission supérieure, cette dernière devient trop importante par rapport à l'augmentation de la capacité et fait donc décroître l'EE. A travers ces résultats pour les différentes configurations, on met en avant l'importance de tenir compte de la puissance consommée par le circuit durant l'évaluation de l'EE. Le concepteur d'un système peut alors effectuer des choix sous contraintes de puissance totale consommée.

En considérant une puissance totale consommée inférieure à 23 dBm, un système à 1 antenne d'émission possède la meilleure EE. Entre 23 dBm et 25dBm, on peut constater qu'un système à base de RT et à 2 antennes d'émission possède une meilleure EE que des systèmes à 1 antenne d'émission. Pour une puissance totale consommée comprise entre 25 dBm et 28 dBm, l'EE de l'émetteur MISO-OFDM à 4 antennes est la meilleure. Entre 28 dBm et 32 dBm, c'est l'émetteur à 8 antennes d'émission et pour des puissances totales consommées supérieures, c'est l'émetteur avec 16 antennes d'émission qui atteint la meilleure EE.

On peut aussi noter sur cette figure, que le maximum d'EE est obtenue pour le système à 8 antennes d'émission. Cette valeur traduit un compromis entre la puissance de transmission, la capacité et la puissance consommée par le circuit. La courbe représentative de l'EE en fonction de la puissance totale consommée pour un émetteur RT à 16 antennes, possède un maximum d'EE plus faible qu'un émetteur à 8 antennes. Dans ce cas, l'augmentation de capacité ne compense plus l'augmentation de la puissance consommée par le circuit.

Ce constat est aussi visible en Figure 5.15 où l'EE est affichée cette fois-ci en fonction de la puissance allouée à la transmission des données, tout en considérant la puissance consommée par le circuit. Cette figure permet d'observer l'évolution de l'EE des différents systèmes pour une puissance de transmission donnée. On constate que le RT apporte un

gain d'EE en comparaison à des systèmes sans pré-codage, notamment à faible puissance de transmission. Par ailleurs, on observe qu'un compromis entre l'augmentation de la puissance consommée par le circuit et le nombre d'antenne devra être trouvé, tout en tenant compte des performances.

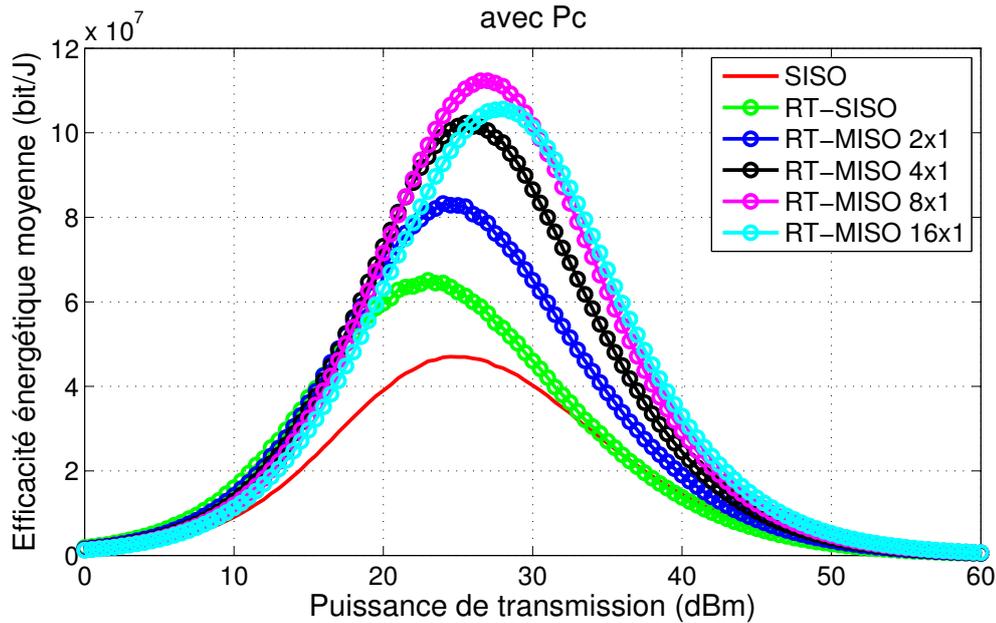


FIGURE 5.15 – Efficacité énergétique vs puissance de transmission P_t pour plusieurs configurations avec P_{cSISO} or P_{cRT} considérées

– Limitations de l'évaluation et pistes d'amélioration

Durant cette étude, il est important de noter que les performances des systèmes sont différentes pour une puissance de transmission donnée. Or, il a été démontré que les performances de ces systèmes s'améliorent avec l'augmentation du nombre d'antennes [Dub13]. Afin de prendre en compte ces améliorations, il pourrait être envisagé dans la suite des travaux :

1. de déterminer la puissance de transmission nécessaire de chaque système permettant d'atteindre un taux d'erreur binaire donné en réception (ex : 10^{-3}),
2. d'évaluer l'efficacité énergétique en fonction du SNR pour les puissances de transmissions correspondantes.

En conclusion, les nombreux résultats présentés permettent de mettre en évidence l'importance de tenir compte de la consommation induite par le circuit dans l'objectif d'obtenir des résultats réalistes d'efficacité énergétique. Dès lors, le concepteur du système peut tester un grand nombre de configurations et de paramètres possibles à l'aide de la

méthodologie de manière à optimiser l'EE ou bien le compromis EE-ES ou encore d'autres métriques d'intérêt. Le retournement temporel apporte un gain d'efficacité énergétique pour une puissance de transmission donnée en comparaison avec des systèmes classiques sans pré-codage. L'augmentation du nombre d'antenne d'émission permet aussi d'améliorer l'EE.

5.3 Analyse de l'efficacité énergétique de techniques coopératives dans le contexte LTE

D'importantes consommations énergétiques sont présentes lorsqu'il est nécessaire d'émettre, de générer et de traiter les données dans les réseaux mobiles.¹ Ces pertes doivent être prises en compte notamment à cause de la diminution des distances entre les stations de base et les équipements mobiles. Les puissances allouées à la transmission des données sont alors relativement faibles en comparaison à la puissance consommée par le circuit [I⁺12]. Par conséquent, optimiser la consommation sur circuits semble donc nécessaire dans l'optique de la conception des futurs systèmes de communications numériques sans fil.

Parallèlement, il a aussi été démontré que les techniques à relais permettent d'exploiter, dans certains cas, un gain dit de *path-loss*. Ce dernier réfère à un gain qui est obtenu lorsqu'il n'y a pas de visibilité directe sur le lien entre l'antenne d'émission et le récepteur, à cause d'obstacles naturels ou non et qui dégradent la qualité du signal en réception. L'utilisation de relais peut permettre d'améliorer les chances de bien recevoir les informations transmises au niveau du récepteur.

5.3.1 Schémas de relayage et protocoles de retransmission en LTE

Dans le cadre du LTE, différents schémas à relais ont été proposés [3GP14, 3GP13], *Amplify-and-Forward* (AF) et *Decode-and-Forward* (DF). Ces deux schémas de relayage sont détaillés en Figure 5.16.

En DF, il est important de noter que le relais ne retransmettra une information que si elle est correctement décodée. Dans ce schéma de relayage, si une information transmise par une station de base est correctement reçue et décodée par l'équipement utilisateur, alors un message d'ACK est transmis au relais et à la station de base. En cas d'échec de décodage, la station de base retransmet le même paquet d'information au relais et à l'équipement utilisateur. Cette fois-ci, si le relais décode correctement l'information (et l'équipement utilisateur non), le relais transmet un ACK à la station de base. Dans ce cas, c'est uniquement le relais qui retransmet l'information jusqu'à ce que l'équipement utilisateur décode correctement le paquet ou bien que le nombre maximal de transmission N_{\max} soit atteint.

En AF, le signal est directement retransmis au niveau du relais tandis qu'en DF, les données sont traitées (démodulation du signal, décodage des données, etc.) au niveau du relais avant d'être retransmises vers le récepteur. On peut alors présumer que, dans le

1. Les travaux présentés dans cette section ont fait l'objet d'une publication [MLM⁺16] et sont le fruit de recherches communes, menées au laboratoire.

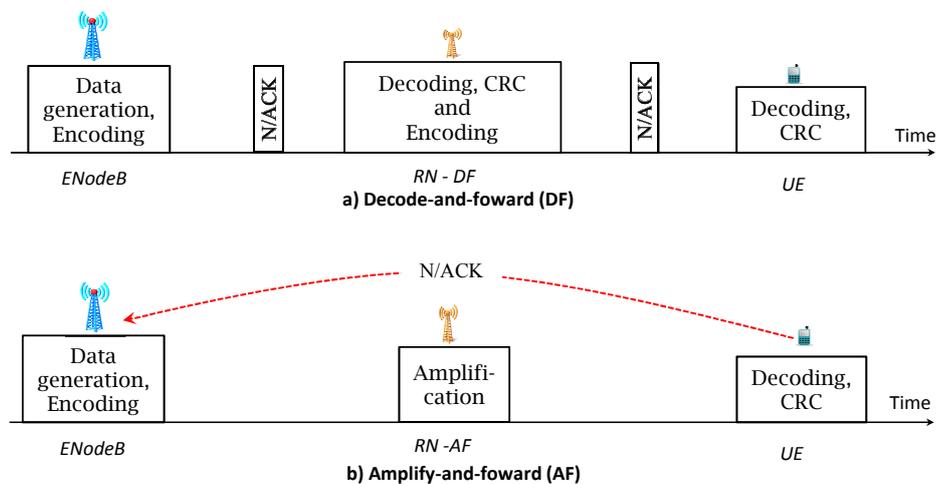


FIGURE 5.16 – Protocoles de relayage avec mécanismes de retransmission HARQ

schéma de relayage DF, la consommation d'énergie sera supérieure à celle du schéma de relayage AF à cause des traitements supplémentaires réalisés au niveau du relais. Néanmoins, la communication en DF peut s'avérer plus robuste au bruit contrairement au schéma AF. D'autre part, dans le cas du schéma AF, une latence négligeable est induite au niveau du relais tandis que dans le cas du schéma DF,

En plus de ces schémas de relayage, différents protocoles de retransmission, appelés protocoles *Hybrid-Automatic Repeat reQuest* (HARQ) ont été adoptés dans le standard LTE. Néanmoins, les travaux présents dans la littérature se focalisent généralement sur l'étude d'un protocole de retransmission HARQ avec un schéma de relayage. Dans le cas de l'étude, les protocoles HARQ-IR (*Incremental Redundancy*) et HARQ-CC (*Chase Combining*) ont été comparés en termes de performances et de consommation.

Dans le cas du protocole HARQ-CC, illustré en figure 5.17, lors de la première transmission, si un paquet n'est pas correctement décodé au niveau du récepteur, un message de NACK est renvoyé à l'émetteur. Dès lors, le message entier est retransmis jusqu'à ce que le paquet soit correctement décodé ou bien que le nombre maximum de retransmission ait été atteint. Au niveau du récepteur, celui-ci sauvegarde l'ensemble des informations reçues et les combine de manière à améliorer le décodage. La probabilité de bien décodé un paquet augmente notamment grâce à la diversité temporelle pour chaque paquet retransmis.

Pour le protocole HARQ-IR, ce n'est pas le même message d'information qui est retransmis en cas de non-décodage en réception. Pour ce protocole, ce sont des bits de redondance, différents selon les retransmissions, qui sont émis. En effet, comme illustré en Figure 5.18 et pour un nombre de retransmission maximum N_{max} égale à 3, l'information est divisée en N_{max} sous-codes, correspondant à différentes versions de bits redondance. A la première émission, les bits d'information et la première version des bits de redondance sont transmis. En réception, en cas d'échec de décodage, un NACK est renvoyé à l'émetteur. Dès lors, une nouvelle version des bits de redondance est retransmise. Ceci a pour conséquence de diminuer le rendement du codage en réception avec le nombre de

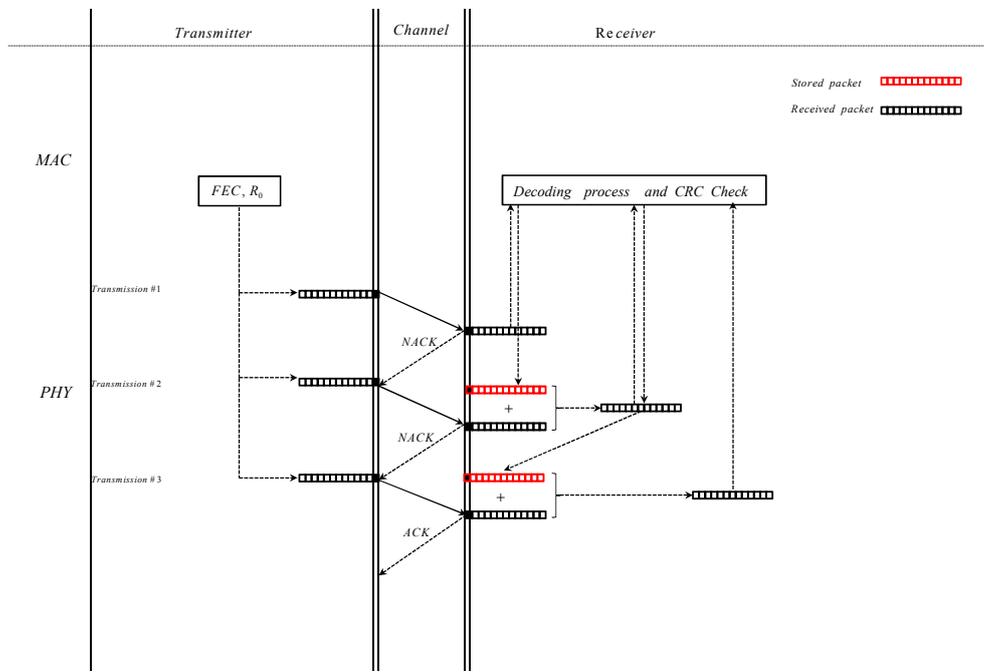


FIGURE 5.17 – Protocole de retransmission HARQ-CC pour $N_{\max} = 3$.

retransmission, ce qui augmente les probabilités de décoder correctement l'information en réception. Finalement, le récepteur essaie alors de décoder l'information préalablement stockée à l'aide de l'information de redondance supplémentaire reçue. En cas d'échec de décodage et que le nombre maximum de retransmission est atteint, alors le paquet est jeté.

A titre de comparaison entre les deux protocoles, le protocole HARQ-IR s'avère plus efficace au niveau du débit. En effet, seul les bits de redondance sont retransmis tandis que pour le protocole HARQ-CC, l'ensemble du paquet est ré-émis. Néanmoins, le protocole HARQ-IR s'avère plus compliqué à mettre en œuvre à cause de l'augmentation et le changement de taille du buffer du décodeur en réception.

5.3.2 Modélisation énergétique

L'autre point innovant de l'étude est la prise en compte de la consommation relative au circuit incluant à la fois les puissances dynamique et statique des traitements en bande de base implémentés sur une cible FPGA, ainsi que l'introduction de valeurs de puissances consommées par les étages RF. Pour cela, la consommation énergétique des principaux blocs IP du canal PDSCH, définis dans le standard LTE, a été évaluée. Ces blocs IP sont décrits sur la figure 5.19. On peut y observer les blocs IP de codage et décodage de canal,

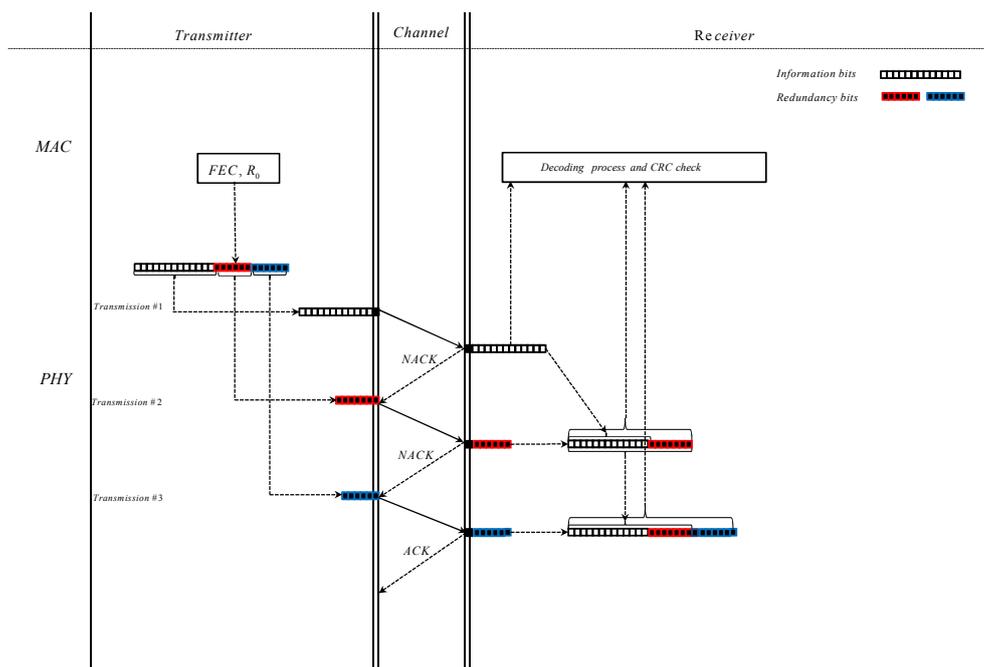


FIGURE 5.18 – Protocole de retransmission HARQ-IR pour $N_{\max} = 3$.

de modulation et démodulation QAM ainsi que les modulateur et démodulateur OFDM. Les puissances consommées relatives à ces différents éléments ont été estimées à l'aide de l'outil XPA.

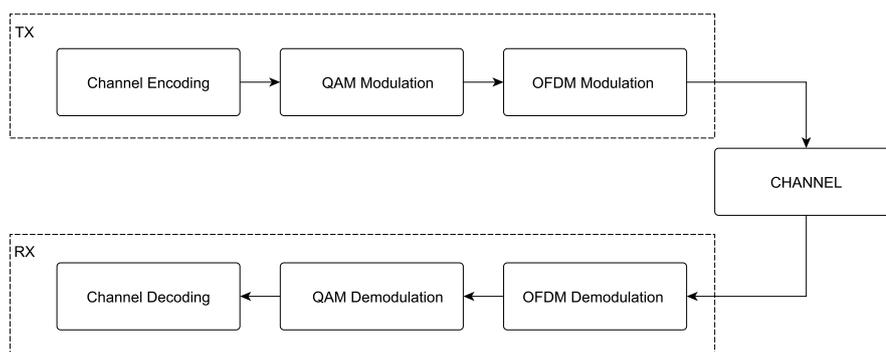


FIGURE 5.19 – Schéma des traitements en bande de base considérés

Afin d'analyser judicieusement l'efficacité énergétique (EE) pour des différents protocoles et schémas de relayage, un modèle énergétique a été construit. En effet, il est important de considérer le délais moyen de retransmission ainsi que la latence induite par les traitements en bande de base, lors de l'évaluation de cette métrique. De plus, on a pu constater dans les sections précédents, qu'il était important de prendre en compte la puissance consommée par les traitements en bande de base, tout comme la puissance consommée relative aux étages RF.

Dans le cadre de ces travaux, l'EE a été évaluée à l'aide de l'équation suivante :

$$\eta_E = \frac{L(1 - \text{PER}(\bar{\gamma}))}{\bar{E}_{\text{totale}}} \quad \text{bits/joule} \quad (5.9)$$

avec $\text{PER}(\bar{\gamma})$ le taux d'erreurs paquet atteint pour un rapport signal-à-bruit moyen (SNR) $\bar{\gamma}$ ppu transmettre L bits.

L'énergie totale consommée par la station de base, le relais et l'équipement utilisateur, notée \bar{E}_{totale} , est exprimée par :

$$\bar{E}_{\text{totale}} = \bar{E}_{\text{BS-UE}} \cdot \mathcal{H}_{\text{direct}} + (\bar{E}_{\text{BS-RN}} + \bar{E}_{\text{RN-UE}}) \mathcal{H}_{\text{coop}}, \quad (5.10)$$

avec $\mathcal{H}_{\text{direct}(\text{coop})}$ une fonction échelon qui vaut 1 lorsque le relais est utilisé, 0 sinon (lien direct). Les différents termes énergétiques peuvent s'exprimer de la manière suivante :

$$\bar{E}_x = (E_{\text{ST},x} + E_{\text{DYN},x} + E_{\text{T},x}) \bar{N}_x, \quad (5.11)$$

avec $E_{\text{ST},x}$, $E_{\text{DYN},x}$ et $E_{\text{T},x}$, les énergies statique, dynamique et de transmission du lien x respectivement, où x peut être le lien station de base (BS) - équipement utilisateur (UE), BS- Relais (RN) ou RN-UE.

De plus, $E_{\text{ST},x} = E_{\text{ST,BB},x} + E_{\text{ST,RF},x}$, avec $E_{\text{ST,BB},x}$ et $E_{\text{ST,RF},x}$ les énergies statiques relatives respectivement aux traitements en bande de base (BB) et à la RF pour le lien x . L'énergie dynamique, notée $E_{\text{DYN},x}$, correspond à l'énergie consommée par les traitement en bande de base sur le FPGA, lorsque celui-ci traite des données. L'énergie consommée liée à la transmission des données, $E_{\text{T},x}$, s'exprime sous la forme suivante :

$$E_{\text{T},x} = N_{\text{OFDM}}(T_{\text{OFDM}} + T_{\text{CP}})P_{\text{T},x} \quad (5.12)$$

avec N_{OFDM} , T_{OFDM} , T_{CP} et $P_{\text{T},x}$ correspondant respectivement au nombre de symboles OFDM, à la durée d'un symbole OFDM, à la durée du préfixe cyclique et à la puissance allouée à la transmission des données.

Au final, \bar{N}_x est le nombre moyen de retransmission par paquet sur le lien x . Ce nombre est mesuré en *Time Slots (TS)* et pour chaque paquet, un nombre maximal de retransmissions noté N_{max} est alloué. Dans ce sens, le délai moyen entre deux nœuds de communication \bar{N}_x peut s'exprimer de la manière suivante :

$$\bar{N}_x = \lim_{\tau \rightarrow \infty} \left(\frac{1}{\tau} \right) \sum_{k=1}^{\tau} N(k) \quad \text{TS} \quad (5.13)$$

avec $N(k) \leq N_{\text{max}} \forall k$ le délai instantané induit lors de la k ième transmission d'un paquet.

L'énergie dynamique consommée au sein du FPGA pour traiter un paquet peut alors s'écrire de la manière suivante :

$$E_{\text{DYN},x} = E_{\text{BB,t},x} + E_{\text{BB,r},x} \quad (5.14)$$

avec $E_{\text{BB,tx,x}}$ et $E_{\text{BB,rx,x}}$ les énergies consommées en bande de base par les chaînes d'émission et de réception respectivement.

Au niveau d'émetteur, l'énergie dynamique consommée par les traitements en bande de base du système décrit en Figure 5.19 vaut :

$$E_{\text{BB,tx,x}} = P_{\text{ENC}}T_{\text{ENC}} + \frac{N_b}{\log 2(M)}P_{\text{MOD}}T_{\text{MOD}} + N_{\text{OFDM}}P_{\text{IFFT}}T_{\text{IFFT}} \quad (5.15)$$

avec P_{ENC} , P_{MOD} and P_{IFFT} les puissances dynamiques moyennes consommées par le codeur de canal, le modulateur QAM et l'IFFT respectivement, N_{OFDM} est le nombre de symboles OFDM, N_b le nombre de bits dans un paquet et M est l'ordre de modulation. De plus, T_{ENC} , T_{MOD} et T_{IFFT} représentent les latences induites par chaque bloc IP.

Au niveau du récepteur, l'énergie dynamique relative aux traitements en bande de base peut s'exprimer de la manière suivante :

$$E_{\text{BB,rx,x}} = \frac{N_b}{\log 2(M)}P_{\text{DEMOD}}T_{\text{DEMOD}} + N_{\text{OFDM}}P_{\text{FFT}}T_{\text{FFT}} + N_{\text{ITER}}P_{\text{DEC}}T_{\text{DEC}} \quad (5.16)$$

avec N_{ITER} , le nombre d'itération du turbo-décodeur. P_{DEMOD} , P_{FFT} et P_{DEC} correspondent aux puissances dynamiques moyennes consommées par le démodulateur QAM, la FFT et le turbo-décodeur respectivement. De la même manière, T_{DEMOD} , T_{FFT} et T_{DEC} correspondent aux latences de chaque bloc IP.

Les valeurs des consommations pour chaque bloc de traitement en bande de base ont été déterminées pour une quantification de 14 bits, une cible FPGA donnée c-à-d un Virtex-6 LX240T et une fréquence d'horloge de 50 MHz. La puissance statique, $P_{\text{ST,BB,x}}$, a été estimée pour ce type de FPGA à 2.950W. D'autre part, la puissance statique de la RF a été délibérément définie à une valeur de 26 dBm qui peut correspondre à une puissance consommée par un utilisateur dans une cellule de petite taille (valeur cohérente avec les valeurs proposées dans le projet EARTH [A⁺12]).

5.3.2.1 Analyse de l'efficacité énergétique

Protocole DF :

Dans un premier temps, l'EE du protocole DF pour le système considéré a été évaluée en fonction de la puissance de transmission uniquement lorsque la puissance dynamique, notée $P_{\text{DYN,BB}}$, relative aux traitements en bande de base est considérée. Sur les Figures 5.20 et 5.21, l'EE a été évaluée respectivement pour les deux protocoles HARQ-CC et HARQ-IR, avec ou sans coopération à l'aide du relais et pour plusieurs modulations adoptées telles que les modulations QPSK, 16QAM et 64QAM.

Sur les Figures 5.20 et 5.21, on peut observer que pour des faibles puissances de transmission, l'EE est meilleure en mode coopératif qu'avec le lien direct, pour les deux protocoles. En effet, sur la Figure 5.20, on peut observer avec protocole HARQ-CC, un plateau autour des 3Mbits/Joule pour les différentes modulations 4-QAM, 16QAM et 64QAM et

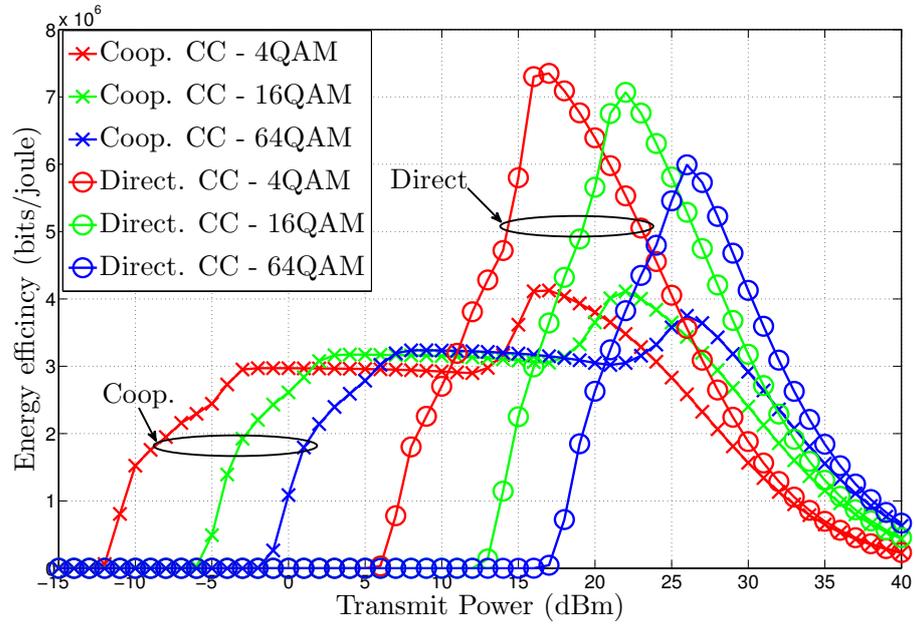


FIGURE 5.20 – Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-CC

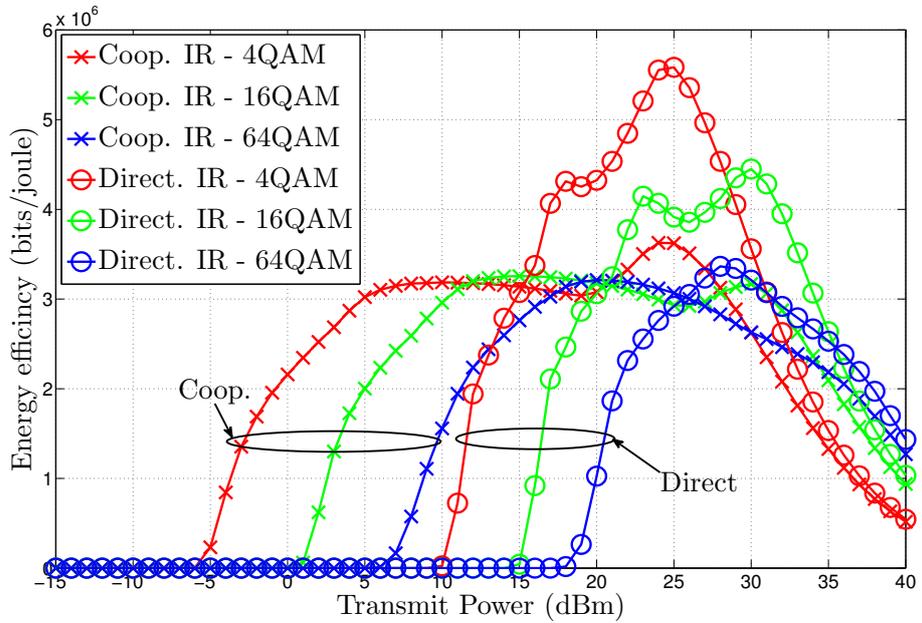


FIGURE 5.21 – Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-IR

pour des puissances de transmission allant de -4dBm à 22dBm . Au fur et à mesure de l'augmentation de la puissance de transmission, l'EE augmente grâce à l'augmentation de la probabilité de décodage avec succès en réception, conduisant à une réduction des délais. Un maximum d'EE est atteint pour le lien avec coopération et le lien direct. On peut noter que, l'EE maximale est 1.85, 1.7 et 1.6 fois plus importante pour le lien directe que pour le mode coopératif, pour respectivement les modulations QPSK, 16QAM et 64QAM. En effet, la puissance de transmission devient très importante et l'utilisation du relais engendre une consommation supplémentaire inutile. De plus à forte puissance de transmission, la probabilité de décoder correctement l'information devient très importante. Le délais atteint sa valeur minimale et aucun gain supplémentaire ne peut être atteint, l'EE décroît donc pour tous les modes de transmission. Le même type d'analyse peut être conduit pour le protocole HARQ-IR en Figure 5.21.

Sur les Figures 5.22 et 5.23, l'EE a cette fois-ci été calculée en tenant compte de l'ensemble des sources de consommation, c-à-d $P_{\text{DYN, BB}}$, $P_{\text{ST, BB}}$ et $P_{\text{ST, RF}}$.

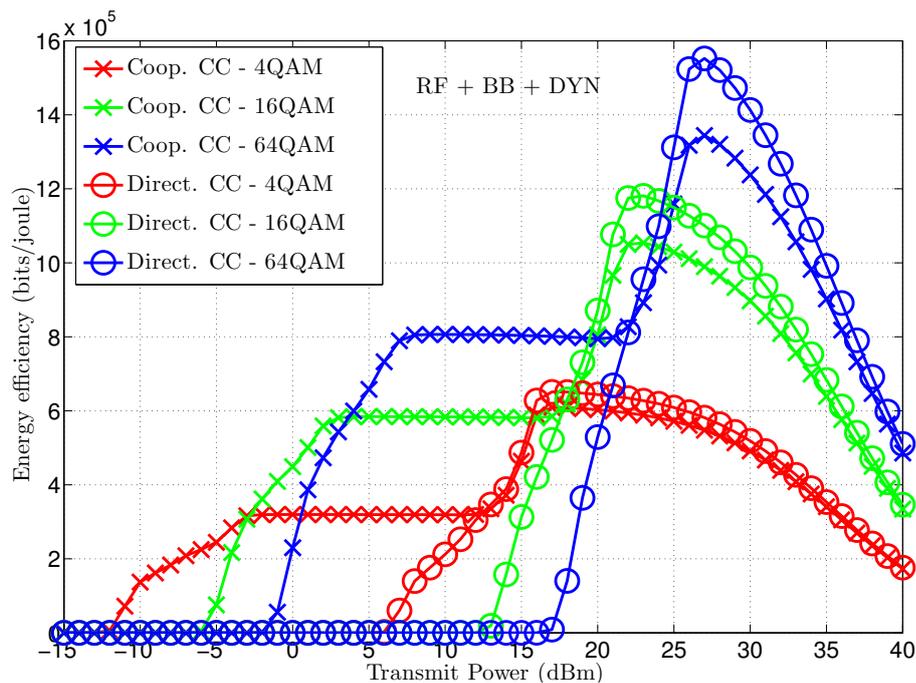


FIGURE 5.22 – Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-CC, toutes les sources de consommation sont considérées

Dans un premier temps, il peut être souligné que l'EE est globalement inférieure lorsque toutes les sources de consommation sont considérées, d'un facteur d'ordre 1 environ. Ensuite, le plateau décrit pour les figures 5.20 et 5.21 est grandement impacté par la modulation utilisée. En effet, le délai moyen dépend de la modulation choisie, l'EE dépend aussi de la modulation. Il peut aussi être constaté que la différence entre les maximums

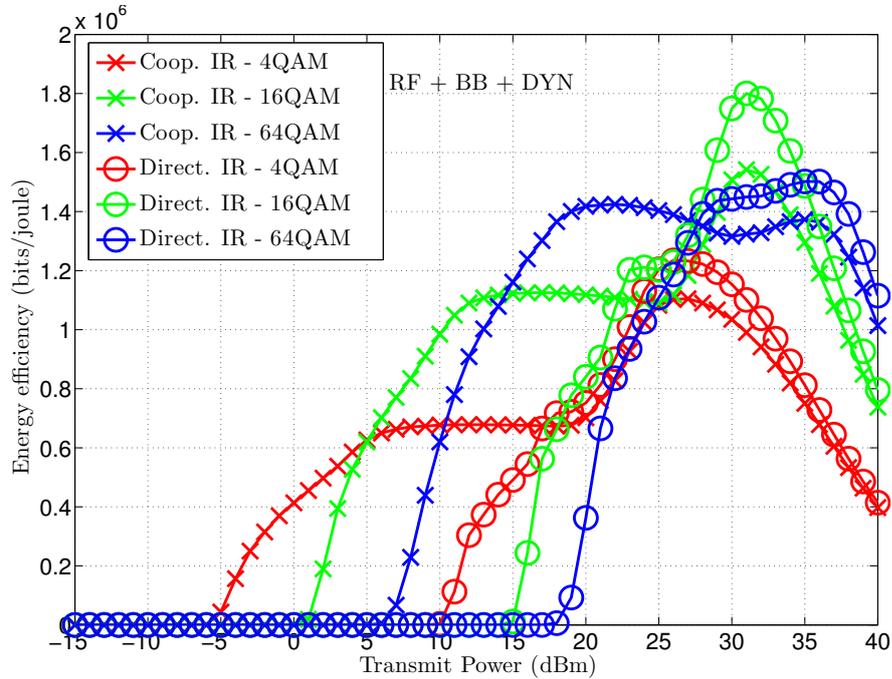


FIGURE 5.23 – Efficacité énergétique versus puissance de transmission (dBm) pour le protocole HARQ-IR, toutes les sources de consommation sont considérées

d'EE atteints en lien direct et pour une transmission coopérative est relativement faible. A titre d'exemple, pour le protocole HARQ-CC et une modulation de 64QAM, un maximum d'EE de 1.5 Mbits/Joule est atteint sur le lien direct tandis qu'un maximum d'EE de 0.8 Mbits/Joule est atteint en mode coopératif. Les différences sont encore plus faibles pour les modulations d'ordre plus faible. En modulation QPSK, on peut même constater qu'il n'y a presque aucun bénéfice sans l'utilisation du relais. Pour le protocole HARQ-IR, les comportements sont semblables. Néanmoins, pour ce type de protocole, un léger gain est obtenu lors de l'utilisation de la communication directe au lieu de la coopération. Cependant, on peut aussi remarquer que le maximum d'EE pour la communication directe est atteint pour une modulation 16QAM. En effet, pour une puissance de transmission $P_t = 31\text{dBm}$, le maximum d'EE sur le lien direct est de 1.8 Mbits/Joule tandis que les EE pour les autres modulations sont plus faibles.

De manière à bien souligner l'influence des différentes sources de consommation, l'EE a été calculée de plusieurs manières, en considérant progressivement les différentes sources de consommation et pour une modulation donnée, c-à-d 16QAM. Les résultats sont présentés sur la Figure 5.24 pour les deux protocoles HARQ-CC et HARQ-IR. Pour chaque protocole, deux cas ont été considérés : lorsque les puissances statiques sont considérées et lorsque les puissances statiques et dynamiques sont considérées. Par exemple, les courbes bleue et rouge (HARQ-IR) représentent respectivement l'EE lorsque la puissance dyna-

mique $P_{\text{DYN, BB}}$ relative aux traitements en bande de base est considérée ou non. A noter que l'EE diminue si le nombre de sources de consommation considérés augmente.

Il a pu être démontré que les résultats de la littérature conduisent à des résultats optimistes de l'EE, à cause du manque de prise en compte des différentes sources de consommation. D'autre part, l'impact de la consommation relative aux traitements en bande de base, est d'autant plus important lorsque la puissance statique est considérée. On peut aussi remarquer HARQ-IR est plus efficace énergétiquement que le protocole HARQ-CC à moyennes et hautes puissances de transmission mais que le protocole HARQ-CC est meilleur à faibles puissances de transmission, c-à-d entre -5 et 2 dBm.

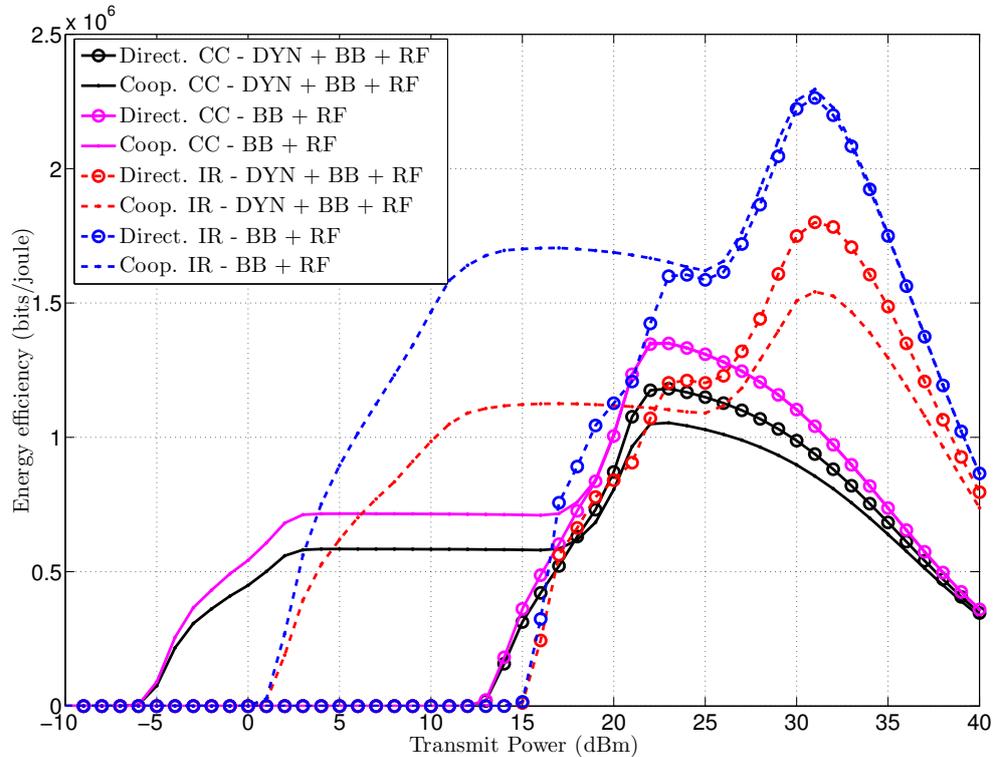


FIGURE 5.24 – Efficacité énergétique versus puissance de transmission (dBm) pour différentes considérations lors du calcul de l'efficacité énergétique

Dans le cas précédent, la modulation choisie était fixe (16QAM) quelle que soit la puissance de transmission. Dans un système réel, le choix de la puissance de transmission s'effectue généralement pour une modulation et un codage donné (*Modulation and Coding Scheme*, (*MCS*)). En Figure 5.25, le maximum d'efficacité énergétique a été représenté pour différents jeux de paramètres donnés dont les protocoles et les différentes modulations mises en œuvre (QPSK, 16QAM, 64QAM).

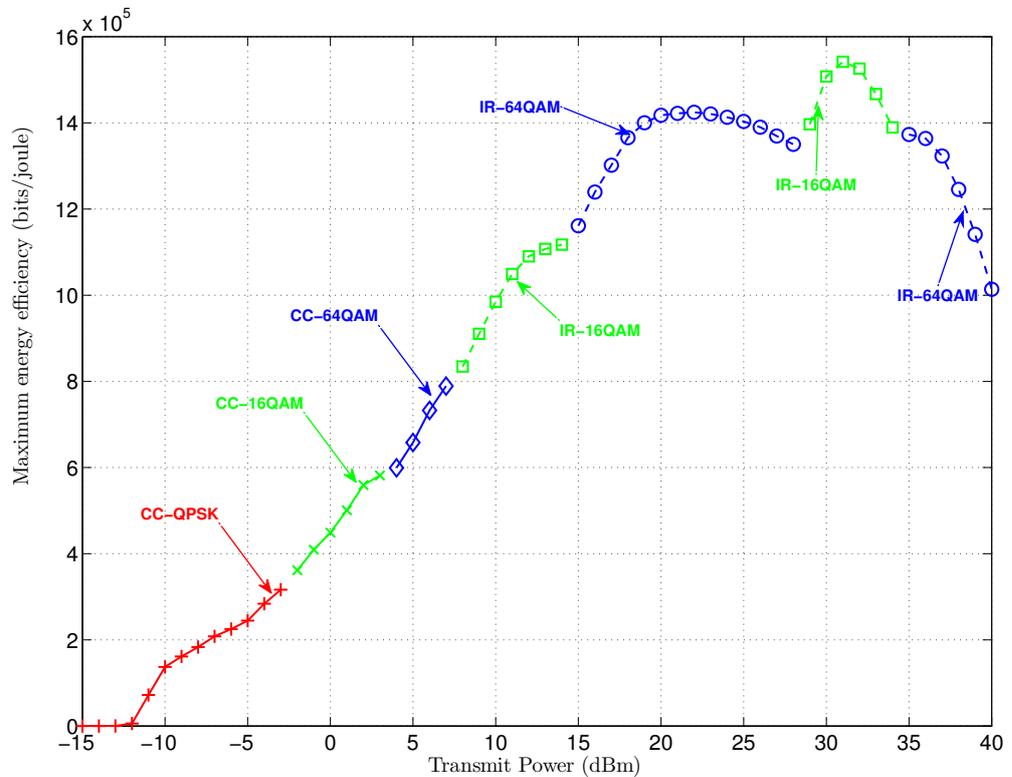


FIGURE 5.25 – Maximum d’efficacité énergétique en fonction de la puissance de transmission
 Transmissions coopératives, protocoles HARQ-IR et HARQ-CC ainsi que 3 modulations (QPSK, 16QAM 64QAM) sont considérés.

A partir de ces résultats, on peut donc identifier la modulation et le protocole qui maximisent l’efficacité énergétique pour une puissance de transmission donnée. A noter que pour une transmission de puissance d’environ 30dBm, le maximum d’EE est atteint pour une modulation 16QAM au lieu d’une modulation 64QAM. Ce type de conclusion ne peut seulement être obtenu que lorsque la puissance dynamique liée aux traitements en bande de base est considérée. Ceci s’explique dans les calculs des énergies consommées à l’émission et en réception lors du changement de modulation de 16QAM à 64QAM. En effet, le nombre de symboles OFDM nécessaire pour coder un nombre de bits donné diminue tandis que le taux d’erreurs par paquet (*Paquet Error Rate (PER)*) augmente considérablement. L’augmentation du PER ne compense pas la diminution de la consommation liée à la diminution du nombre de paquets transmis.

Au regard de cette étude, une comparaison efficace, axée sur l’évaluation de l’efficacité énergétique, a été présentée. Cette étude souligne l’importance de tenir compte de toutes les sources de consommation afin d’obtenir des résultats d’études réalistes. De plus,

dans des systèmes où les puissances de transmission sont faibles, l'impact de la puissance consommée par les traitements en bande de base ainsi que les différentes puissances statiques n'est pas négligeable.

5.3.2.2 Analyse du compromis efficacité spectrale-efficacité énergétique

L'EE n'est pas le seul critère afin d'optimiser un système de communications numériques. Pour cela, l'efficacité spectrale est évaluée. Ces deux métriques sont généralement opposées c-à-d que maximiser l'ES n'est possible que pour de faibles valeurs d'EE et inversement. Sur la Figure 5.26, le compromis EE-ES est évalué lorsque toutes les sources de consommation sont considérées, pour les deux protocoles HARQ-IR et HARQ-CC ainsi que pour les 3 modulations. Une manière d'estimer l'ES est de diviser le débit, noté η_T , par la bande de fréquence utilisée, notée W :

$$\eta_{SE} = \frac{\eta_T}{W} \quad \text{bits/sec/Hz} \quad \text{where} \quad \eta_T = \frac{L(1 - \text{PER}(\bar{\gamma}))}{\bar{\mathcal{D}}} \quad \text{bits/sec} \quad (5.17)$$

$$(5.18)$$

avec $\bar{\mathcal{D}}$, le délai moyen en seconds pour transmettre L bits d'information. L'ES est directement lié au choix du schéma de modulation et de codage (MCS).

Pour de faibles valeurs d'ES, l'EE évolue presque linéairement puis décroît au fur et à mesure de l'augmentation de l'ES. Ce comportement s'explique par la présence de consommations statiques qui sont indépendantes de la puissance de transmission. En effet, à faible puissance de transmission, c-à-d à faibles ES, les consommations statiques sont prédominantes. Cependant, lorsque la puissance de transmission augmente, la puissance dynamique devient importante jusqu'à devenir majoritaire, conduisant à des pertes d'énergie et donc à des diminutions de l'EE.

Le protocole HARQ-IR permet d'atteindre un meilleur compromis EE-ES en comparaison au protocole HARQ-CC pour une modulation donnée. Pour la modulation 16QAM, le maximum d'EE atteint pour le protocole HARQ-IR est d'environ 1.5 Mbits/joule pour une ES de 2.1 bits/s/Hz, tandis que pour le protocole HARQ-CC, l'EE est de 1.08 Mbits/Joule pour une ES de 0.9 bits/s/Hz. Ceci est principalement lié à une utilisation judicieuse de la retransmission en ne transmettant uniquement que des sous-blocs de redondance, ce qui s'avère plus efficace. Le taux de codage de base (avant retransmission des bits de redondance supplémentaire) en HARQ-IR est haut, c-à-d $4/5$, tandis qu'avec le protocole HARQ-CC, l'ensemble du paquet doit être retransmis avec un taux de codage fixe d' $1/3$.

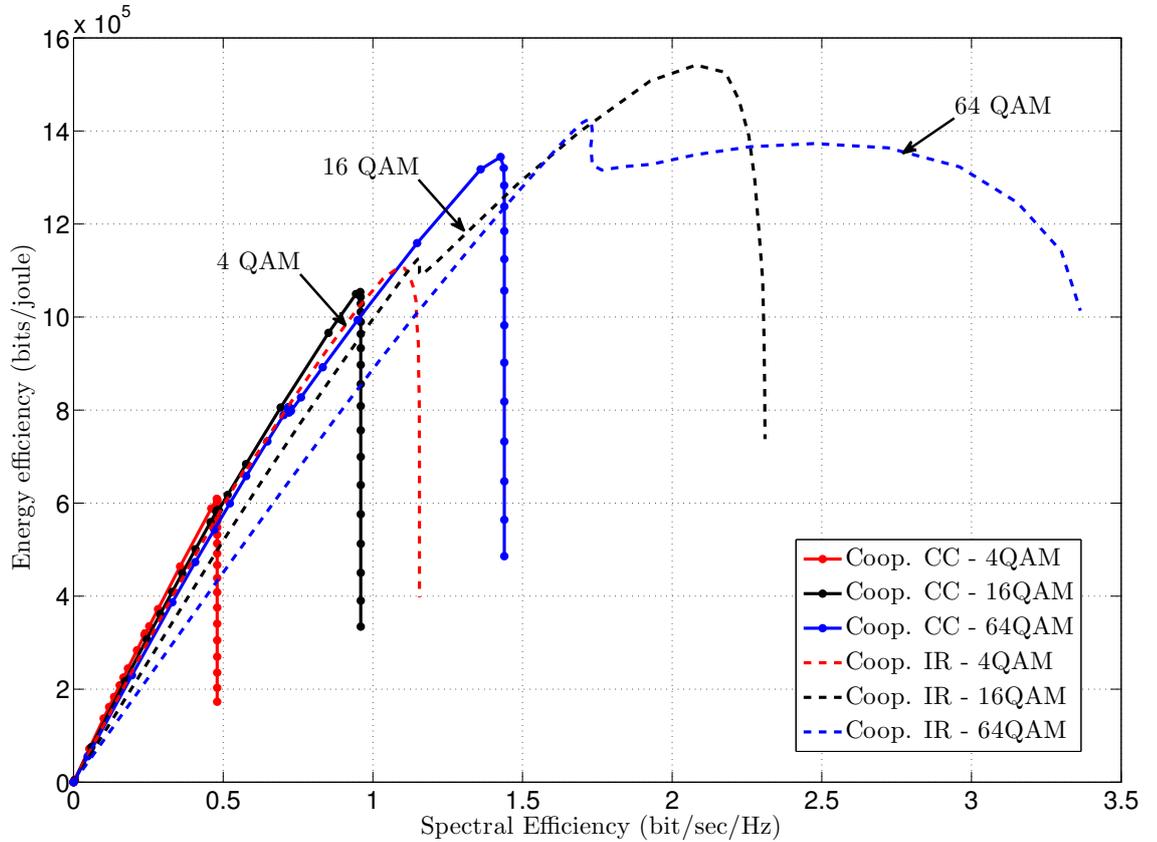


FIGURE 5.26 – Efficacité énergétique versus Efficacité spectrale : les deux protocoles (HARQ-IR et HARQ-CC), les 3 modulations (QPSK, 16QAM 64QAM) et toutes les sources de consommation sont considérés

Comparaison des schémas de relayage AF et DF

Dans cette partie, EE des schémas de relayage AF et DF sont évaluées en fonction de la puissance de transmission et pour une modulation donnée, c-à-d QPSK, permettant de fixer le débit. D'autre part, le protocole HARQ-CC a été utilisé et l'ensemble des sources de consommation a été considéré pour le calcul de l'EE, c-à-d $P_{\text{DYN,BB}} + P_{\text{ST,BB}} + P_{\text{ST,RF}}$.

Sur les Figures 5.27 et 5.28, l'EE est évaluée pour un nombre maximum de retransmission N_{max} valant 2 et 6 respectivement. Trois zones peuvent être distinguées, deux zones non efficaces énergétiquement à faibles et hautes puissances de transmission et une zone efficace pour des puissances de transmission moyennes. A faible puissance de transmission, la probabilité de décoder correctement le paquet est très faible tandis qu'à forte puissance de transmission, l'énergie dépensée est trop importante, ce qui conduit dans les deux cas à de faibles EE. Pour des puissances de transmission moyennes, EE pour le schéma AF est meilleure que pour DF. Pour $N_{\text{max}} = 2$ AF possède une meilleure efficacité énergétique à partir d'une puissance de transmission de -4dBm . Ensuite, EE atteint un plateau à environ 0.59 Mbits/Joule pour le schéma AF contre 0.32 Mbits/Joule pour DF jusqu'à des

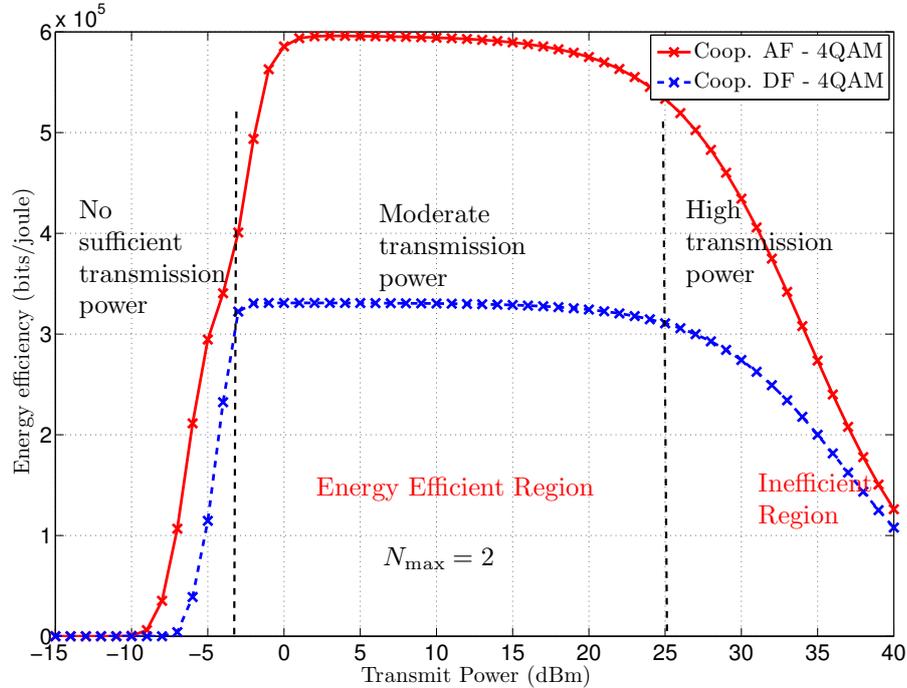


FIGURE 5.27 – Efficacité énergétique pour les schémas AF et DF versus puissance de transmission (dBm) pour $N_{\max} = 2$

puissances de transmission d'environ 20 – 25dBm.

Pour un nombre de retransmission $N_{\max} = 6$ en Figure 5.28, le même comportement est décelé avec approximativement les mêmes valeurs que pour $N_{\max} = 2$. Une remarque intéressante se situe pour des faibles puissances de transmission, c-à-d de -10 à -6 dBm, où DF possède une meilleure EE par rapport au schéma AF. Ceci s'explique par la très faible probabilité de décoder correctement un paquet dans le schéma AF par rapport au schéma DF. En effet, le bruit étant très important à ce niveau de puissance de transmission, DF possède une plus grande probabilité de pouvoir décoder correctement le paquet au niveau du relais. Au final, l'EE pour le schéma AF est bien meilleure pour des puissances de transmissions moyennes à cause de l'importance des consommations statiques et dynamiques au niveau du relais dans le schéma DF et des choix pour le modèle réalisé (cf. section 5.3.2).

5.4 Conclusion

Dans ce chapitre, la méthodologie présentée au chapitre précédent a été utilisée de manière à évaluer la consommation de différents systèmes : des chaînes SISO et MISO, avec ou sans RT et des chaînes avec techniques à relais. Ces chaînes ont été entièrement développées en VHDL afin de pouvoir évaluer et comparer efficacement l'approche proposée.

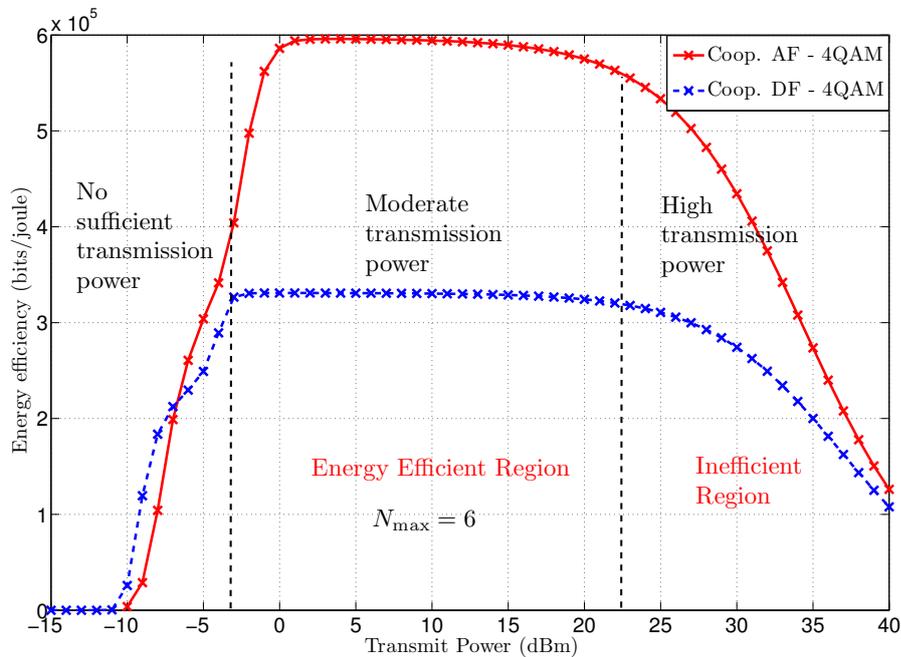


FIGURE 5.28 – Efficacité énergétique pour les schémas AF et DF versus puissance de transmission (dBm) pour $N_{\max} = 6$

Dans un premier temps, la méthodologie a été validée sur une chaîne d'émission réception SISO-OFDM. Ensuite, un émetteur LTE SISO-OFDM relatif aux traitements du canal PDSCH ainsi qu'un émetteur LTE MISO-OFDM 2x1 avec encodage d'Alamouti ont été développés et évalués en termes de consommation et d'EE. Pour finir, une chaîne MISO-OFDM à X antennes d'émission intégrant le principe du retournement temporel en fréquence a été réalisée.

La méthodologie a permis d'évaluer très rapidement et précisément la consommation de ces systèmes pour différentes configurations et paramètres. L'introduction du concept de scénario rend la comparaison des systèmes plus facile et la formalise. D'autre part, l'exploration de l'espace de conception est ainsi facilitée de manière à pouvoir dimensionner correctement son système (nombre d'antennes d'émission/réception, type de modulation, type de codage, nombre de sous-porteuses utiles, bande de fréquence utilisée, etc.) de manière à effectuer des choix judicieux de configurations, de blocs IP ou encore de technologies. Ainsi, on peut imaginer, à partir de cette méthodologie, comparer efficacement des systèmes à base de LTE (ou encore de WiMAX) à la fois en termes de consommation et de performances. Bien d'autres études peuvent être envisagées.

A titre d'exemple, des mesures de consommations ont été intégrées dans une étude de l'efficacité énergétique des techniques à relais dans le contexte LTE. Les EE des schémas AF et DF ont été évaluées ainsi que les protocoles HARQ-CC et HARQ-IR. Un modèle

énergétique a été construit à l'aide des estimations de consommation obtenues à l'aide d'XPower. De plus, différentes sources de consommation, statique et dynamique, des différents éléments d'une station de base ont été considérés. Le compromis EE-ES a aussi été évalué. Les principaux résultats de l'étude sont rappelés ci-dessous :

1. le protocole HARQ-IR combiné au schéma de relayage DF offre un meilleur compromis EE-ES en comparaison du protocole HARQ-CC grâce à l'utilisation efficace de la retransmission,
2. le schéma de relayage AF délivre une meilleure EE que DF à cause de la présence des consommations statiques et dynamiques au niveau du relais dans le cas DF et du temps de retransmission,
3. l'utilisation de techniques coopératives permettent d'obtenir une meilleure EE que les transmissions directes à faibles puissances de transmission,
4. prendre en considération l'ensemble des sources de consommation conduit à des résultats réalistes et permet de quantifier leurs impacts à différentes puissances de transmission.

Finalement, l'utilisation des techniques coopératives semble être une bonne opportunité dans l'objectif de concevoir des techniques vertes de communications sans fil.

Au regard de l'ensemble de la méthodologie proposée, plusieurs limitations peuvent être néanmoins identifiées :

- la phase de caractérisation peut être importante,
- il manque une généralisation des résultats à d'autres cibles FPGAs,
- un certain manque de flexibilité sur les résultats d'estimation peut être déploré.

En effet, la phase de caractérisation, malgré l'automatisation à l'aide de scripts, peut s'avérer longue. Ce temps est d'autant plus grand que le bloc IP possède de nombreux paramètres. L'autre limitation de la méthodologie réside dans le manque de généralisation pour d'autres familles de FPGA. En effet, il serait nécessaire d'effectuer autant de fois l'étape de caractérisation qu'il y a de cibles FPGA envisagées par le concepteur.

Dans le chapitre 6, ces limitations sont adressées notamment à l'aide de modélisation haut-niveau des blocs IP.

Chapitre 6

Optimisations de la méthodologie d'estimation

6.1 Limitations de l'approche proposée

Au regard de la méthodologie d'estimation proposée dans le cadre de la thèse, plusieurs limitations ont pu être soulevées. Premièrement, si le bloc IP n'est pas présent dans la bibliothèque et donc n'a pas été caractérisé, il est nécessaire d'effectuer la phase de caractérisation ainsi que la création du modèle SystemC correspondant. Bien que l'ensemble des procédures ait été automatisé, cette phase reste inévitable et doit être effectuée pour chaque cible FPGA.

Il existe donc un certain manque de flexibilité et de généralisation vis-à-vis des résultats obtenus dans la bibliothèque. Pour palier cela, il a été envisagé de développer des modèles analytiques qui permettent d'une part, d'exprimer la consommation en fonction de paramètres d'intérêt et d'autre part d'obtenir des modèles plus généralistes afin d'effectuer des extrapolations à haut-niveau. L'utilisation de modèles analytiques permet de réduire le volume d'informations nécessaire à sauvegarder dans la bibliothèque. Néanmoins, gagner en généralité provoque inévitablement des approximations et donc une perte de précision.

Les sections suivantes décrivent les procédés que nous avons employés pour palier les limitations de la méthodologie.

6.2 Modélisation sous forme analytique de la consommation

Dans le chapitre 3, plusieurs techniques de modélisation ont été décrites. Notre étude s'inspire de la méthodologie FLPA qui a notamment été appliquée à la modélisation de processeurs [LJSM04] et de blocs IP [EJH06]. Cette technique a pour objectif d'identifier les paramètres fonctionnels de blocs IP ayant un impact sur la consommation. Ensuite, un modèle analytique exprimant une relation entre la consommation et ces paramètres est réalisé par régression.

Dans le cadre de la thèse, cette technique peut être appliquée à des blocs IPs du domaine des communications numériques sans fil.

6.2.1 Modélisation de blocs IP

Un focus a été réalisé sur le développement d'un modèle analytique autour du bloc IP IFFT de Xilinx. En effet, ce dernier est l'un des blocs matériels qui consomme le plus d'énergie dans les chaînes de communications numériques qui ont pu être étudiées. La consommation de ce bloc IP a été modélisée pour un FPGA de type Virtex-6 LX240T, pour une taille de transformée variant de 128 à 2048 et des fréquences allant de 10 à 100MHz.

Le tableau 6.1 résume l'ensemble des paramètres utilisés pour la génération de l'IP IFFT de Xilinx (référence : Fast Fourier Transform v7.1) à l'aide de l'outil Core Generator.

TABLE 6.1 – Paramètres de génération de l'IP IFFT

Paramètres	Valeur
Channel	1
Transform Length	[128 :2048]
Implementation Options :	Architecture Pipelined-Streaming
Target Clock Frequency	100 MHz
Data Format	Fixed Point
Precision Options	Input Data Width = [10 14 or 16] Phase Factor Width = [10 14 or 16]
Scaling Option	Scaled
Rounding Mode	Truncation
Optional Pins	CE, SCLR, OVFL0
Output Ordering	Natural Order and Cyclic Prefix insertion
Input data timing	no offset
Number of Stage using Block Ram	Default value
Reorder Buffer	Block Ram - Default value
Optimize Options	Use 3-multiplier structure
Butterfly Arithmetic	CLB

Pour modéliser la consommation de ce bloc IP, l'approche classique d'utilisation d'XPA a été utilisée (Implémentation + simulation après placement routage). Par la suite, l'ensemble des mesures a été analysé sous Matlab afin de construire le modèle.

A l'issue de la modélisation, le modèle analytique obtenu a été décrit par l'équation 6.1 (NB : la puissance dynamique relative aux E/S n'est pas considérée dans cette équation) :

$$P_{\text{dynamique}(mW)} = p00 + p10 \ln(N_s) + p01 f_{\text{clock}} + p20 \ln(N_s)^2 + p11 \ln(N_s) f_{\text{clock}} \quad (6.1)$$

avec les coefficients $p00$, $p10$, $p01$, $p20$ et $p11$, donnés dans le tableau 6.2.

Ces modèles ont été étendus pour plusieurs valeurs de quantification des données c-à-d le nombre de bits servant à représenter les données. Il est important de noter que le bloc IP

ne tolère que des valeurs comprises entre -1 et 1 en entrée. Le format de quantification est donc, dans ce cas, d'un bit de signe, les bits restants représentant la partie fractionnaire.

TABLE 6.2 – Coefficients du modèle analytique d'estimation de la consommation du module IFFT - Virtex6

Paramètres	p_{00}	p_{10}	p_{01}	p_{20}	p_{11}	<i>Erreur Quadratique Moyenne (linéaire)</i>
10 bits	8.053	-2.626	-0.3537	0.2099	0.2432	2.321
14 bits	-4.613	1.531	-0.3551	-0.1225	0.2863	1.808
16 bits	10.26	-3.352	-0.2922	0.2688	0.2918	0.5225

Sur la Figure 6.1, la puissance dynamique estimée par le modèle est comparée aux valeurs utilisées pour construire le modèle (numéro d'échantillon sur l'axe des abscisses), pour une quantification des données sur 14 bits. La précision du modèle réalisé est alors relativement bonne. En effet, le modèle a été validé sur la même série de valeurs que la base de construction du modèle. Afin de valider complètement le modèle, une autre série de valeurs indépendantes de la base de construction du modèle devra être envisagée.

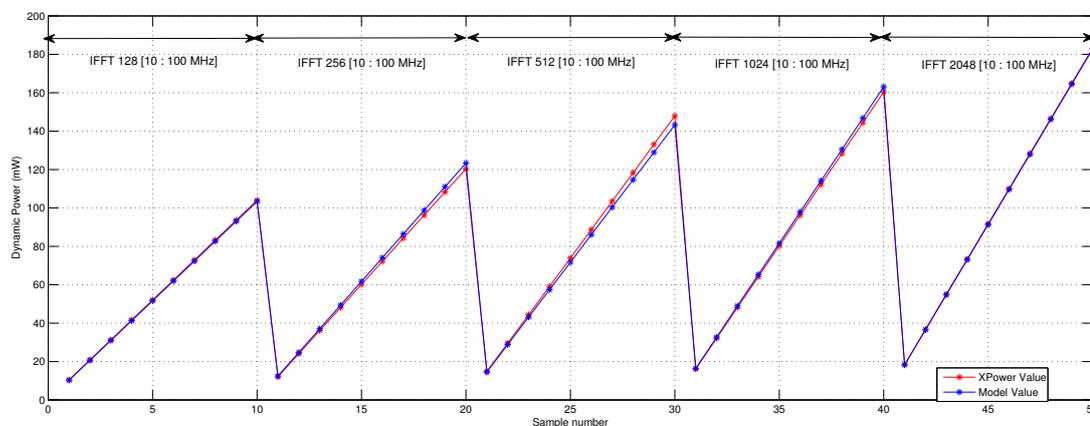


FIGURE 6.1 – Comparaison de la précision du modèle analytique et des valeurs originales -FPGA Virtex 6 - 14bits

A travers cet exemple, un modèle analytique peut être facilement obtenu en fonction des paramètres d'intérêt, à partir des résultats issus de la phase de caractérisation. Cette approche peut être étendue à d'autres éléments que l'IFFT tels que l'encodeur/décodeur, le (dé)modulateur QAM, afin finalement de modéliser l'ensemble des éléments d'un système de communications numériques sans fil. De cette manière, la puissance totale du système peut être approximée par les contributions de chaque élément. Dans un second temps, ces modèles peuvent être directement intégrés dans les modèles haut-niveau SystemC.

6.2.2 Extension à des systèmes complets

Dans de nombreux travaux de modélisation tirés de la littérature, les modélisations portent généralement sur des blocs IP spécifiques mais ne considèrent pas un système complet.

Dans cet exemple, nous avons repris la chaîne SISO-OFDM (cf. sous-section 5.2.1 du chapitre 5) et effectué la modélisation des blocs IPs de l'émetteur de ce système. Les modèles réalisés sont visibles dans le tableau 6.3.

Certaines hypothèses ont été émises lors de la création des modèles analytiques :

1. tailles d'IFFT de 256 à 2048,
2. quantification des données : $q1 = 1$ si $q=10$ bits sinon 0,
 $q2 = 1$ si $q=14$ bits sinon 0,
3. taille supportée de blocs de codage notée $CBS = [40, 1024, 2048, 3200, 4096, 6144]$,
4. fréquence d'horloge notée f : 10 à 100MHz,
5. modulations supportées : QPSK ($mod1= 1, mod2=0, mod3= 0$),
16QAM ($mod1= 0, mod2=1, mod3= 0$),
64QAM ($mod1= 0, mod2=0, mod3= 1$),
6. 1 symbole pilote est inséré au début de chaque trame de données, constituée de 10 symboles OFDM,
7. FPGA considéré : Virtex-6 LX240T.

TABLE 6.3 – Modèles analytiques des modules de l'émetteur SISO-OFDM

IP/module VHDL	Modèles Analytiques
Codeur de canal	$P = 0.0521 + 0.1785 * f + 1.835.10^{-5} * CBS$
Modulateur	$P = q1(mod1 * (0.0424 * f - 0.00533) + mod2 * (0.04425 * f - 0.004) + mod3 * (0.0514 * f - 0.00133)) + q2(mod1 * (0.04515 * f - 0.002) + mod2 * (0.0479 * f + 0.00067) + mod3 * (0.0591 * f - 0.0013))$
Carrier Mapper	$P = q1(0.0853 * f - 0.002667) + q2(0.106 * f - 0.002667)$
IFFT	$P = q1(8.053 - 2.626 * \log(FFT_{SIZE}) - 0.3537 * f + 0.2099 * \log(FFT_{SIZE})^2 + 0.2432 * \log(FFT_{SIZE}) * f) + q2(-4.613 + 1.531 * \log(FFT_{SIZE}) - 0.3551 * f - 0.1225 * \log(FFT_{SIZE})^2 + 0.2863 * \log(FFT_{SIZE}) * f)$
DAC Scaling	$P = q1(0.0503 * f - 0.003333) + q2(0.06015 * f - 0.004)$
Time-Pilot Inserter	$P = q1(0.09736 * f - 0.004667) + q2(0.1112 * f - 0.0006667)$

A partir des modèles analytiques développés, on peut estimer rapidement la consommation moyenne de l'émetteur pour une application donnée peut s'effectuer rapidement.

On considère un émetteur avec une IFFT de taille 256, une quantification des données sur 10 bits ($q_1=1, q_2=0$), une modulation QPSK ($\text{mod}_1=1, \text{mod}_2=\text{mod}_3=0$), une taille de bloc de codage de 1024 bits et une fréquence d'horloge de 50 MHz. Trois mémoires de type FIFO sont nécessaires à la synchronisation des données et ont été caractérisées afin d'être prises en compte dans le calcul. D'après les modèles analytiques et les valeurs de puissances consommées par les mémoires, la puissance totale de l'émetteur pour cette application est estimée à 72.44 mW. Le système total a été implémenté sur FPGA et la consommation totale a été estimée à l'aide de XPower et est de 67.22 mW.

Comme second exemple, on considère un émetteur avec une IFFT de taille 2048, une quantification des données sur 14bits ($q_1=0, q_2=1$), une modulation QPSK ($\text{mod}_1=1, \text{mod}_2=\text{mod}_3=0$), une taille de bloc de codage de 1024 bits et une fréquence d'horloge de 50Mhz. De la même façon, les mémoires de synchronisation ont été caractérisées. Pour cette seconde application, la puissance estimée via les modèles analytiques est de 116.44 mW. A l'aide de XPA, la consommation estimée pour ce système entier est de 100.14 mW.

6.2.3 Limitations

On peut noter que l'erreur d'estimation reste relativement faible. Celle-ci est liée en partie à des optimisations faites par l'outil de conception lors du déroulement du flot de conception. D'autre part, il est nécessaire de considérer les aspects dynamiques des applications (modules actifs/inactifs) qui peuvent avoir un impact significatif sur la consommation moyenne totale du système.

Ainsi, en combinant les modèles de consommation et une simulation haut-niveau à partir des modèles SystemC, des coefficients d'activité peuvent être obtenus pour les 2 applications, de manière à affiner les résultats, et tenir compte des temps d'activités des blocs IP.

Dans le tableau 6.4, on peut constater que l'approche combinant l'utilisation des modèles analytiques et de la simulation haut-niveau permettant de déterminer les coefficients d'activité, délivre des résultats plus précis que l'utilisation seule de modèles analytiques.

Le pourcentage d'erreur avec les modèles analytiques, pour les deux applications est de 7% et 16% respectivement, tandis qu'avec l'approche combinant les modèles analytiques et la simulation haut-niveau délivre une erreur de 1.44% et 6.64%.

TABLE 6.4 – Modèles analytiques des modules de l'émetteur SISO-OFDM

	Estimation (mW)		
	XPower (référence)	Modèles Analytiques	Modèles Analytiques et coefficients d'activité
Appli. 1	67.22	72.44	68.19
Appli. 2	100.14	116.44	106.79

6.2.4 Généralisation à différentes familles de FPGA

Il a pu être constaté qu'une des limitations majeures était la généralisation à d'autres familles de cibles FPGAs, différentes de celle définie en phase de caractérisation. En effet, un grand nombre de familles de FPGA existe, chacune possédant des caractéristiques technologiques qui lui sont propres. Estimer facilement la consommation d'un système implanté dans un FPGA d'une autre famille n'est pas trivial.

Établir une relation entre la consommation d'un système implanté sur différentes familles de FPGA est une problématique qui a été abordée lors de nos travaux. Pour cela, des scripts ont permis d'effectuer automatiquement les étapes d'implémentation et de caractérisation de plusieurs blocs IP pour différents FPGAs.

Sur la Figure 6.2, des valeurs de puissance dynamique moyenne ont été obtenues pour un bloc IP de type Filtre FIR. La taille du filtre varie de 2 à 112 taps (de 10 en 10), permettant ainsi de faire évoluer la complexité et le nombre de ressources utilisées. Le Filtre FIR a été implémenté pour 9 FPGAs de la famille des Virtex-4 (résultats indiqués en rouge), 7 FPGAs de la famille Virtex-5 (bleu) et 6 composants FPGAs de la famille Virtex-6 (vert).

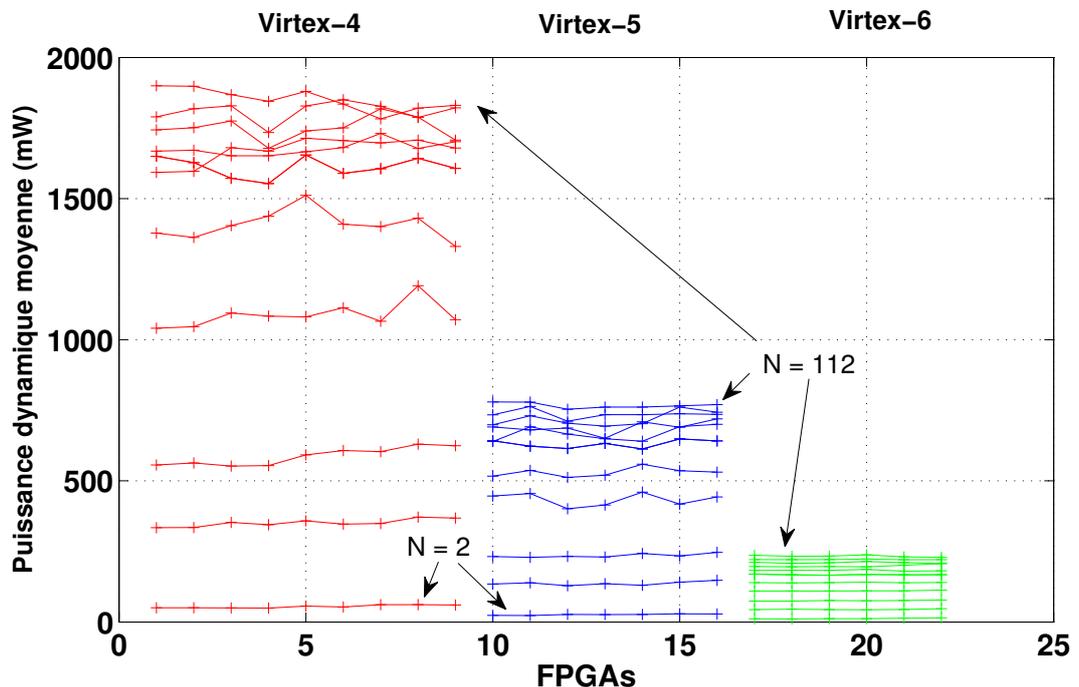


FIGURE 6.2 – Puissance dynamique moyenne (mW) d'un filtre FIR en fonction des différentes familles de FPGAs

Comme indiqué sur la Figure 6.2, on peut constater que pour une taille N de filtre donné (chaque ligne représente la puissance dynamique estimée pour une taille de filtre donnée pour différents FPGA d'une même famille), la puissance dynamique moyenne pour une famille ne varie pas de manière significative. Ceci est d'autant plus vrai en fonction

TABLE 6.5 – Facteurs de mise à l'échelle de la puissance dynamique moyenne de blocs IP pour différentes familles de FPGA

Bloc IP vs FPGA		Facteurs			Puissance Dynamique (mW)		
		Virtex 4 (ref)	Virtex 5	Virtex 6	Virtex 4	Virtex 5	Virtex 6
Filtre FIR	N=2 Taps	1	0.479	0.225	54.67	26.19	12.285
	32 Taps	1	0.399	0.1015	1087	433.79	110.41
	62 Taps	1	0.391	0.104	1611	630.72	167.58
	112 Taps	1	0.414	0.1261	1850	767.14	233.42
Mod. QAM	QPSK	1	0.415	0.263	8.33	3.47	2.19
	16QAM	1	0.306	0.304	8.02	2.46	2.44
	64QAM	1	0.317	0.23	12.44	3.94	2.88
Démod. QAM	QPSK	1	0.368	0.266	7.15	2.63	1.9
	16QAM	1	0.259	0.212	13.56	3.51	2.88
	64QAM	1	0.395	0.21	17.09	6.75	3.58
IFFT	128pts	1	0.465	0.332	697	324	231
	256pts	1	0.438	0.303	846.7	370.6	256.52
	512pts	1	0.393	0.277	1159	456	321
	1024pts	1	0.363	0.258	1407	511	364
	2048pts¹	1	0.478	0.293	1143	546.5	335.6

¹ Puissance dynamique plus faible liée à l'utilisation forcée de mémoires par l'outil de conception

des avancées technologiques.

Au regard de ces résultats, des facteurs de mise à l'échelle peuvent être *a priori* déterminés afin de pouvoir estimer la puissance dynamique moyenne de ce circuit pour une autre famille de FPGA. Dans le tableau 6.5, plusieurs facteurs de mise à l'échelle ont été déterminés pour différents blocs IPs. A noter que les estimations de puissance dynamique obtenues pour les FPGAs Virtex-4 servent de référence dans le calcul des facteurs.

A travers la mise en œuvre de ces facteurs de mise à l'échelle pour différents blocs IP d'une chaîne de communications numériques, il est alors possible d'évaluer rapidement l'impact du choix de cible FPGA sur la consommation dynamique moyenne du système. De plus, l'introduction de ces facteurs permet de généraliser les résultats obtenus à l'aide de notre méthodologie afin de limiter l'impact du temps de caractérisation, au détriment d'une perte en précision toutefois.

6.3 Bilan de la modélisation de bloc IP sous forme analytique

Le développement de modèles analytiques permet d'obtenir un modèle flexible et d'exprimer la consommation en fonction de paramètres d'intérêt. Il a aussi été montré qu'une extrapolation des estimations de puissance dynamique pour différentes familles de FPGA pouvait être réalisée à l'aide de facteurs de mise à l'échelle. Néanmoins, ces facteurs doivent aussi être déterminés pour chaque bloc IP et pour chaque famille de FPGA.

Les modèles décrits dans les sections précédentes ne permettent pas d'exprimer la puissance dynamique en fonction de l'activité du circuit. En effet, les modèles développés permettent d'estimer la valeur de la consommation dynamique moyenne d'un bloc IP ou d'un système en fonction de paramètre d'intérêt et ces valeurs de consommation, ont été obtenues à l'aide de XPA, pour un jeu de stimuli typique et réaliste pour chaque bloc IP.

Une limitation de l'approche proposée est donc mise en évidence. Celle-ci ne permet pas d'estimer la consommation pour des stimuli totalement différents qui peuvent conduire à un comportement différent et donc une consommation dynamique différente du bloc IP. Pour cela, une nouvelle approche de modélisation est proposée de manière à tenir compte de la dépendance de la consommation dynamique d'un bloc IP vis-à-vis des stimuli qui lui sont appliqués.

6.4 Modélisation de blocs IP à l'aide de réseaux de neurones

6.4.1 Les réseaux de neurones

Les réseaux de neurones sont des modèles mathématiques de calcul qui utilisent le principe de l'apprentissage par expérience. Ce modèle est issu en partie d'une analogie avec les réseaux de neurones biologiques que les humains et animaux possèdent. Ils sont très souvent employés pour résoudre des problèmes d'optimisation, de prédiction, de modélisation, de classification, de reconnaissance, etc. On les retrouve dans de nombreux domaines comme le traitement du signal et de l'image, la médecine, la robotique, la finance, les communications numériques, etc.

Les réseaux de neurones permettent de résoudre très rapidement des problèmes non-linéaires complexes à l'aide d'une structure relativement simple, en comparaison à d'autres algorithmes. Plusieurs types de structure de réseau existent, dont le choix du type dépend principalement de la nature du problème à résoudre.

Durant ces travaux, la structure du réseau utilisé est de type *Feed-Forward*, c-à-d un réseau non bouclé dans lequel l'information est propagée couche après couche sans retour en arrière. Néanmoins, d'autres structures existent. On peut citer les plus utilisées [Gra07, DB02] :

- Perceptron monocouche,
- Perceptron multicouches (PMC),
- Adaline (*Adaptive Linear Neuron*) ...

Plusieurs structures de PMC ont été étudiées mais un réseau possédant une couche cachée a été retenue pour la modélisation car ce type de réseau s'est avéré suffisamment performant. La structure d'un tel réseau de neurones est représentée sur la Figure 6.3.

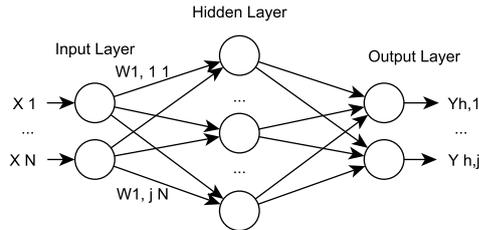


FIGURE 6.3 – Schéma d'un réseau de type perceptron multi-couches

Le modèle mathématique d'un neurone s'écrit sous la forme [Par04a] :

$$n = \sum_{j=1}^N X_{1,j} w_j - \theta \quad (6.2)$$

avec n la sortie du neurone correspondant à la somme pondérée des N poids w et des N entrées X moins le biais θ du neurone (aussi appelé seuil d'activation). La sortie n correspond au niveau d'activation atteint.

Une hypothèse régulièrement émise dans les réseaux artificiels est que tous les neurones sont supposés synchrones et délivrent un résultat à chaque instant t . Ainsi, tous les neurones fournissent une sortie $\alpha(t) = f(n(t))$ [Par04a].

Finalement, la matrice de poids notée W_h est multipliée par le vecteur d'entrée $X = (x_1, x_2, x_3, \dots, x_n)^T$ de façon à calculer la sortie de la couche cachée de la manière suivante :

$$y_{h,j} = f\left(\sum_{i=1}^{N_i} w_{h,ji} x_i - \theta\right) \quad (6.3)$$

avec $w_{h,ji}$ le poids connectant l'entrée i à l'unité neuronale j de la couche cachée. θ est un terme de biais.

Plusieurs fonctions d'activation existent dans la littérature comme notamment la fonction seuil, la fonction linéaire, etc. Dans notre cas, la fonction f est une fonction d'activation non-linéaire de type sigmoïde :

$$f(\alpha) = \frac{1}{1 + e^{-\alpha}}$$

Ce type de fonction d'activation représente un bon compromis, notamment en évitant les discontinuités apportés par une fonction seuil par exemple et permet d'obtenir plus de modularité qu'une fonction linéaire ayant une seule pente [Par04a].

Processus d'apprentissage

Les réseaux de neurones sont capables d'apprendre à partir d'une série de vecteurs de tests (stimuli) présentée aux entrées du réseau. D'autre part, plusieurs modes d'apprentissage sont possibles. Le mode d'apprentissage dit *supervisé*, qualifie un réseau connaissant la réponse à certaines entrées, c-à-d la sortie qu'il doit prédire pour un jeu de stimuli donné. De cette manière, on force le réseau à converger vers une valeur précise. Dans le cas du mode non-supervisé, le réseau ne possède pas d'information vers laquelle converger, le réseau est libre de converger vers telle ou telle valeur en sortie.

Dans la littérature, il existe une multitude d'algorithmes d'apprentissage. Dans le cadre de notre étude, un algorithme simple de rétro-propagation du gradient de l'erreur a été utilisé afin d'entraîner les réseaux. Plusieurs versions de cet algorithme existent [DB02] :

- l'algorithme de rétro-propagation de Levenberg-Marquardt,
- l'algorithme de rétro-propagation du gradient conjugué,
- ...

Ce type d'algorithme cherche à minimiser un indice de performance, basé sur l'erreur commise observée suivant la méthode de descente du gradient. Dans le cas de notre étude, l'erreur quadratique moyenne (*Mean Square Error (MSE)*) a été définie comme indice de performance.

Le processus d'apprentissage se déroule de manière itérative. Ainsi, au fur et à mesure de l'apprentissage, la valeur de poids synaptiques de chaque neurone est modifiée. L'erreur commise par rapport à la sortie attendue est ensuite propagée. Les poids synaptiques où l'erreur est importante se verront modifiés de manière plus significative que dans le cas d'une erreur faible. Le processus étant itératif, il se termine lorsqu'une erreur minimale est obtenue sur tous les exemples. Autrement dit, l'objectif des méthodes d'apprentissage consiste à modifier les poids synaptiques de manière à minimiser l'indice de performance.

Une fois l'apprentissage effectué, 2 autres bases sont requises : les bases de test et de validation. La base de test contient des stimuli, différents de la base d'apprentissage, qui sont appliqués au réseau afin de tester l'apprentissage effectué par le réseau. Une base supplémentaire peut être aussi définie et est dite de validation. Celle-ci permet, à l'aide de stimuli différents de ceux utilisés pour les 2 autres bases, de vérifier si le réseau est capable de généraliser, c-à-d si le réseau est capable de fournir une réponse réaliste pour n'importe quel jeu de stimuli, présenté à ses entrées.

6.4.2 Modélisation

Dans le cadre de notre étude, chaque bloc IP pourra être représenté sous la forme d'une boîte noire, possédant des entrées et des sorties. Il a pu être constaté, dans la formule de l'équation de la puissance dynamique, que la puissance dynamique d'un circuit est fortement dépendante de son activité. Ainsi, l'activité d'un circuit dépend de l'activité de ses signaux d'entrée et donc des stimuli appliqués.

Comme indiqué sur la Figure 6.4, chaque bloc IP est ainsi modélisé à l'aide de deux réseaux de neurones. Le premier réseau permet d'estimer l'activité des signaux de sortie par

rapport à l'activité des signaux d'entrée. Le second réseau permet d'estimer la puissance dynamique moyenne du bloc IP, toujours par rapport à l'activité des signaux des entrées.

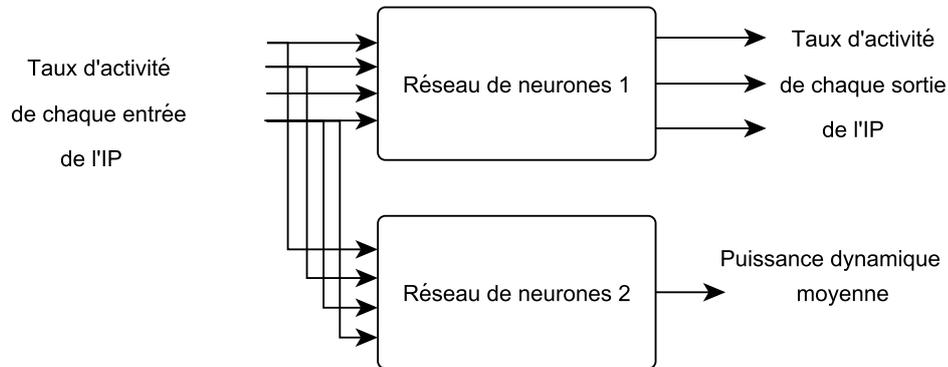


FIGURE 6.4 – Schéma du procédé de modélisation à partir des deux réseaux de neurones

Sous XPower, l'activité des signaux des entrées et des sorties est représentée à l'aide de deux valeurs :

1. le nombre de millions de transitions par seconde, aussi appelé *Signal Rate*,
2. le pourcentage du temps durant lequel le signal est à l'état logique haut.

Le *signal rate* est calculé de la manière suivante :

$$SR(MTr/s) = F_{horloge}(Mhz) * \beta$$

avec β représentant le nombre moyen de commutations par cycle d'horloge [Xil12c] et $F_{horloge}$, la fréquence d'horloge en MHz.

Dans la suite du rapport, nous utiliserons la notion de taux d'activité qui fera référence au *Signal Rate* et au pourcentage du temps à l'état haut de chaque signal.

6.4.3 Entraînement, tests et validation

Dans l'objectif de construire les différentes bases nécessaires à l'entraînement, le test et la validation des réseaux, 10000 *patterns* aléatoires ont été générés. Chaque *pattern* fournit les taux d'activité de chaque signal représentant le *Signal Rate* et le pourcentage du temps dans lequel le signal se trouve à l'état haut, pour le bloc IP considéré. Une fois les *patterns* générés, des fichiers de configuration pour XPower sont annotés avec ces informations. A l'aide de scripts TCL, il est possible d'automatiser l'estimation de la consommation du bloc IP considéré pour chacun des *patterns*. Une fois les valeurs obtenues, la création, l'entraînement, le test et la validation des réseaux de neurones peuvent être réalisés. Typiquement, 70% des 10000 *patterns* servent à l'entraînement des réseaux, 15% des exemples constituent la base de test et les 15% restants constituent la base de validation.

6.4.4 Exemples de modélisation

Afin de valider les différents modèles réalisés, l'exemple d'un émetteur du domaine des communications numériques sans fil est présenté en Figure 6.5. L'émetteur simplifié est

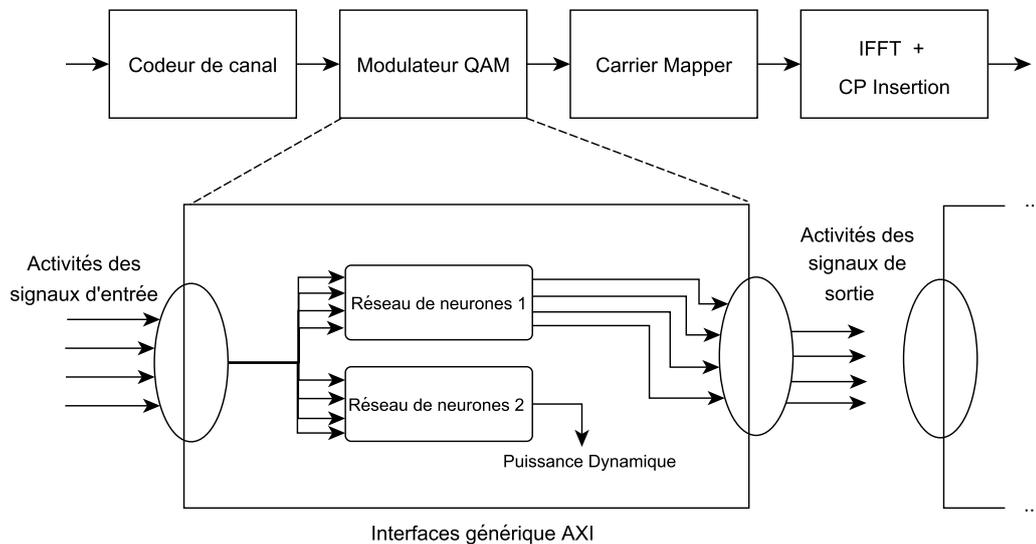


FIGURE 6.5 – Schéma de l'émetteur basique OFDM modélisé

constitué d'un codeur de canal, d'un modulateur QAM, d'un bloc permettant la sélection des sous-porteuses actives et d'un modulateur OFDM à base d'IFFT. Ces différents blocs IP ont donc été implémentés pour un FPGA donné (Virtex-6 LX240T) et testés de manière à pouvoir valider l'approche.

Sur la Figure 6.5, on peut aussi remarquer que chaque bloc IP est modélisé à l'aide de deux réseaux. Chaque modèle communique avec les autres à l'aide d'une interface générique, basée sur l'interface AXI. Ainsi, les activités relatives aux signaux peuvent être propagées de modèle en modèle sans soucis d'interopérabilité.

Dans le tableau 6.6, la puissance dynamique estimée à l'aide des réseaux de neurones est présentée pour les blocs IP considérés. De même, la précision des estimations des taux d'activité des signaux de sortie en fonction des taux d'activité des signaux d'entrée pour chaque bloc IP a été évaluée.

On peut constater à partir du tableau 6.6 que la précision obtenue à l'aide des réseaux de neurones est très bonne, c-à-d avec une erreur quadratique moyenne inférieure à 1% en comparaison aux estimations fournies par l'outil XPower Analyzer (XPA).

Bien que le modèle réalisé à l'aide des réseaux de neurones soit très précis (limité à la précision de l'outil XPower), il est important que ce modèle soit rapide. Pour cela, les temps d'estimations à l'aide d'XPower et des réseaux de neurones ont été évalués et sont présentés dans le tableau 6.7. On constate qu'à partir des modèles basés sur les réseaux de neurones, un facteur d'accélération d'ordre 3 peut être obtenu en comparaison avec le temps d'estimation de l'outil XPower.

D'autre part, l'utilisation de tels modèles augmente significativement la flexibilité de l'estimation. Si l'on considère un système constitué de N blocs IP, il est alors nécessaire

TABLE 6.6 – Précision des modèles par rapport à XPower Analyzer (Virtex-6)

Bloc IP		Erreur quadratique moyenne	Gamme des valeurs
IFFT	Puissance Dynamique*	0.087	[40-195] mW
	Activité	0.16	[0-100]
Carrier Mapper	Puissance Dynamique	4.22e-04	[6-8] mW
	Activité	0.353	[0-100]
Modulateur QAM	Puissance Dynamique	1.179e-4	[3-5] mW
	Activité	0.1498	[0-100]
LTE Downlink Channel Encoder	Puissance Dynamique	1e-3	[31-40] mW
	Activité	3.6e-3	[0-100]

* Puissance dynamique relative aux E/S n'est pas considérée

d'utiliser N fois l'outil XPA dans l'objectif de déterminer, pour chaque bloc, les taux d'activités des sorties et de les propager au bloc suivant. D'autre part, si le concepteur du système décide de modifier le système, par exemple à travers un changement de bloc IP, l'ensemble des caractérisations ainsi que la propagation des taux d'activité doivent être effectués de nouveau à partir du bloc IP modifié.

A l'aide des modèles basés sur les réseaux de neurones, tout est effectué automatiquement. Les taux d'activités sont propagés d'un réseau d'un bloc IP à un autre. Si un changement de bloc IP intervient, une seule simulation est requise pour obtenir les taux d'activité ainsi que la puissance consommée dans cette nouvelle configuration.

TABLE 6.7 – Comparaison des temps d'estimation

Bloc IP		XPower Résultats	Modèle proposé	Facteur d'accélération		
IFFT	Puissance dynamique	29s	20ms	1450		
	Activité					
Carrier Mapper	Puissance dynamique	12s		20ms	600	
	Activité					
Modulateur QAM	Puissance dynamique	12s			20ms	600
	Activité					
LTE Downlink Channel Encoder	Puissance dynamique	20s	20ms			1000
	Activité					
Système entier (fig.6.5)	Puissance dynamique	73s		80ms		912
	Activité					

Finalement, cette approche permet à ses utilisateurs de tester un grand nombre de configurations possibles pour leurs systèmes à base de blocs IP grâce à une grande flexibilité et une rapidité importante par rapport aux outils classiques.

6.5 Conclusion

A travers ce chapitre, les différentes limitations de la méthodologie proposée dans le cadre de cette thèse ont pu être adressées. Dans un premier temps, le développement de modèles analytiques, permettant de lier la puissance dynamique à des paramètres d'intérêt, a été réalisé. La flexibilité apportée par ce type d'approche est très importante et les temps d'estimation à l'aide de ce type d'approche est très court. Néanmoins, les aspects dynamiques d'une application ne sont pas forcément pris en considération. Un raffinement des estimations peut alors être effectué en combinaison avec une simulation fonctionnelle du système et de la détermination des temps d'activité.

Dans un second temps, une innovante modélisation de la consommation dynamique de blocs IP sur FPGA a été proposée à l'aide de réseau de neurones. A notre connaissance, ce type d'approche de modélisation n'a pas encore été présenté dans la littérature, pour modéliser la consommation de blocs IP pour FPGA. Les performances en termes de rapidité d'estimation et de précision sont très bonnes en comparaison à des outils classiques. Un facteur d'ordre 3 est obtenu en termes de rapidité tandis qu'une erreur quadratique moyenne inférieure à 1% a été observée pour plusieurs blocs IP. Finalement, l'utilisation de ce type d'approche permet d'offrir la possibilité aux concepteurs de système sur FPGA, de tester très rapidement un grand nombre de configurations de leur système constitué de blocs IP, tout en gardant une bonne précision sur les valeurs estimées de consommation.

Conclusion et Perspectives

Conclusion

Ces travaux de thèse ont été consacrés au développement d'une méthodologie d'estimation de la consommation de systèmes de communications numériques sans fil sur FPGA. Cette méthodologie s'inscrit dans le contexte actuel des systèmes embarqués et vise à répondre aux problématiques de rapidité et de précision de l'estimation tout en essayant de s'affranchir de la complexité croissante des systèmes. De plus, la méthodologie essaie de poser un cadre formel afin de faciliter la comparaison de systèmes entre eux, étant donné la multitude de paramètres possibles pouvant influencer la consommation et les performances ainsi que la complexité croissante des systèmes visés. Pour cela, des approches combinant une modélisation au niveau système avec des estimations bas-niveau ont démontré que la méthodologie était adéquate pour répondre à ces contraintes.

La méthodologie proposée dans ce manuscrit, repose sur l'hypothèse qu'un système de communications numériques peut être représenté sous la forme d'un ensemble de modèles haut-niveau communicants entre eux, et représentatifs de blocs IP matériels. Durant la première étape de la méthodologie, une phase de caractérisation de chaque bloc IP est effectuée. Chaque bloc IP a été entièrement caractérisé en termes de ressources utilisées, de consommation, etc, pour l'ensemble de leurs paramètres intrinsèques. A l'issue de cet étape, un modèle haut-niveau de chaque bloc IP est décrit en SystemC, suivant un modèle d'implémentation précis. Dans la seconde phase de l'approche proposée, une représentation haut-niveau du système est développée à l'aide des modèles SystemC de chaque bloc IP. Ces modèles sont connectés entre eux permettant ainsi de valider la fonctionnalité du système modélisé. Le concept de scénario a aussi été introduit et illustré à travers de nombreux exemples. Ce concept permet de comparer simplement et efficacement plusieurs applications entre elles, dans un cadre formel, tout en permettant d'évaluer l'impact des choix de paramètres fait par l'utilisateur c-à-d sur les performances et sur la consommation finale du système. Une fois le scénario défini par l'utilisateur, des simulations du système modélisé peuvent être effectuées pour chaque application (instance de scénario représentant un paramétrage complet du scénario). De plus, les temps d'activité des différents modèles des blocs IP du système sont mesurés afin de pouvoir raffiner les estimations de consommation du système total. De cette manière, les aspects dynamiques relatifs à chaque application peuvent être ainsi considérés. Ceci différencie grandement notre approche de celles illustrées dans la littérature.

L'efficacité de l'approche a été démontrée à travers de nombreux exemples, présentés

dans le chapitre 5, du domaine des communications numériques sans fil :

1. une chaîne d'émission réception SISO-OFDM,
2. un émetteur LTE-SISO-OFDM,
3. un émetteur MISO-OFDM à plusieurs antennes d'émission intégrant ou non le principe de RT en fréquence,
4. des chaînes avec techniques de transmission à relais.

Ces systèmes ont entièrement été développés en VHDL afin de valider efficacement l'approche proposée. La puissance dynamique moyenne ainsi que les temps nécessaires aux estimations ont été évalués. La méthodologie a permis d'évaluer très rapidement et précisément la consommation de ces systèmes, pour différentes configurations et paramètres, en comparaison à un outil classique d'estimation, c-à-d XPower Analyzer de Xilinx. Les facteurs d'accélération évalués sont de l'ordre de 3 à 4.

A partir des estimations de la consommation dynamique des traitements en bande de base de ces systèmes, la métrique d'efficacité énergétique a été évaluée. A travers ces études, il a pu être démontré qu'il était important de prendre en compte la consommation de ces traitements dans le calcul de cette métrique. Au regard du contexte actuel, cette considération est d'autant plus importante que les puissances allouées à la transmission des données deviennent faibles (taille des cellules du réseau diminue).

Des techniques de focalisation comme le retournement temporel ou encore l'utilisation de techniques à relais peuvent alors être facilement évaluées en y intégrant des modèles et des estimations de consommation réalistes. Dans le cadre de l'étude sur les systèmes à base de retournement temporel, il a pu être constaté que cette technique permettait d'atteindre une meilleure efficacité énergétique en comparaison à des systèmes sans pré-codage, pour une puissance de transmission donnée.

Une étude a aussi été effectuée à partir des estimations faites durant ces travaux de thèse. Menée en commun dans le laboratoire, elle portait sur l'évaluation des performances des techniques coopératives dans le cadre du LTE. Ainsi, des mesures de consommations y ont été intégrées afin de construire un modèle dont l'objectif est d'évaluer l'efficacité énergétique des schémas de relayage AF et DF et des protocoles HARQ-IR et HARQ-CC, définis dans standard LTE. A travers le paramétrage de ce modèle, différentes sources de consommation, statique et dynamique, des différents éléments d'une station de base ont été considérées. De cette manière, l'impact de sources de consommation a pu être évalué, notamment sur le compromis EE-ES. Les principaux résultats de l'étude sont rappelés ci-dessous :

1. le protocole HARQ-IR combiné au schéma de relayage DF achève un meilleur compromis EE-ES en comparaison avec le protocole HARQ-CC grâce à l'utilisation efficace de la retransmission,
2. le schéma de relayage AF délivre une meilleure EE que DF à cause de la présence des consommations statiques et dynamiques au niveau du relais dans le cas DF,

3. l'utilisation de techniques coopératives permettent d'obtenir une meilleure EE que les transmissions directes à faibles puissances de transmission,
4. prendre en considération l'ensemble des sources de consommation conduit à des résultats réalistes et permet de quantifier leurs impacts à différentes puissances de transmission.

Au final, l'utilisation des techniques coopératives semble être une bonne opportunité dans l'objectif de concevoir des techniques vertes de communications sans fil. L'utilisation de technique de focalisation spatiale et temporelle comme le retournement temporel semble prometteuse afin d'augmenter l'EE de système à faible puissance de transmission, tout en n'augmentant pas la complexité du récepteur.

Dans le chapitre 6, les principales limitations de la méthodologie ont été identifiées et des solutions d'optimisations ont été proposées puis mises en place. L'utilisation de modèles analytiques permet en effet de réduire la taille de la librairie requise, d'estimer très rapidement la consommation tout en apportant un réel gain en flexibilité (généralisation). Néanmoins, obtenir un modèle général se traduit dans la plupart des cas par une perte de précision. Plusieurs modèles analytiques ont été développés, pour différents éléments d'une chaîne d'émission-réception SISO-OFDM. Il a pu être constaté qu'en combinant l'utilisation de modèles analytiques avec une simulation haut-niveau du système modélisé (à l'aide de l'approche proposée), l'erreur d'estimation a pu être réduite pour une application de 7% à 1.44% et pour une seconde application de 16% à 6.64%.

Une technique de modélisation basée sur les réseaux de neurones, largement répandue dans la littérature, a été utilisée de manière innovante pour modéliser, à la fois la puissance dynamique d'un bloc IP matériel et les taux d'activité des signaux sorties. Pour cela, deux réseaux de neurones sont construits, prenant en entrées des taux d'activités des signaux d'entrées, et délivrant respectivement la puissance dynamique moyenne et les taux d'activité des signaux de sortie. Des résultats d'estimation pour plusieurs blocs IP d'un émetteur SISO-OFDM ont été obtenus à partir de cette technique et ont permis de valider cette approche. A notre connaissance, aucune étude de la littérature n'a encore utilisé les réseaux de neurones pour modéliser la consommation dynamique ainsi que les activités des entrées/sorties de blocs IP matériels. Un des principaux avantages de ce type d'approche est la rapidité d'estimation en comparaison à des approches classiques d'estimation. En outre, l'activité d'un circuit, qui est une des métriques les plus difficiles à évaluer sur FPGA, peut alors être modélisée à l'aide de ce type d'approche. L'objectif final étant de disposer pour chaque bloc IP des modèles correspondants pour ensuite les intégrer à haut-niveau. Au final, les activités de chaque bloc IP d'un système sont ainsi propagées de modèle en modèle. Il est alors possible d'estimer relativement précisément la consommation d'un système pour n'importe quel type de stimuli d'entrée.

Perspectives

Ces travaux de thèse ouvrent la voie à de nombreuses perspectives. Tout d'abord, il pourrait être intéressant de profiter de la méthodologie proposée pour comparer d'autres systèmes. Par exemple, toujours dans le domaine des communications numériques, les

consommations relatives aux traitements en bande de base du WiMax et ceux du LTE pourraient être comparés. Ensuite, les systèmes étant amenés à intégrer un grand nombre d’antennes, il serait intéressant d’approfondir les études menées, en tenant compte de cette perspective. De plus, les traitements en bande de base pour des nouvelles techniques de modulation prometteuse comme par exemple l’OFDM/OQAM (*Offset Quadrature Amplitude Modulation*) pourraient être ainsi efficacement comparés aux analyses réalisées pour les systèmes OFDM au cours de cette thèse.

Dans les différents cas d’étude décrits au chapitre 5, l’EE a été évaluée pour un MCS donné. En effet, nous avons choisi une modulation QPSK et un rendement de codage de 1/3. Or, il serait intéressant d’évaluer l’EE pour beaucoup d’autres MCS à l’aide de l’outil. Cette étude permettrait de déterminer les MCS permettant d’optimiser l’EE en fonction de la puissance de transmission. Une courbe pourrait être tracée pour le système étudié et qui représenterait les maximums d’EE pour les différents MCS.

D’autre part, il pourrait être envisagé d’appliquer la méthodologie à des systèmes développés dans d’autres domaines comme le traitement d’image par exemple. Il peut aussi être imaginé d’estimer la consommation de systèmes sur ASIC. Dans ce cas, des outils spécifiques aux ASICs pourraient être intégrés durant la phase de caractérisation.

L’acquisition d’une plate-forme matérielle, en combinaison avec du matériel de mesure devraient permettre d’effectuer des mesures réelles de la consommation. De cette façon, la précision de notre approche pourrait être évaluée de manière plus fine qu’à partir de l’outil XPower, servant dans la thèse de référence. En outre, l’acquisition de mesures réelles à l’aide d’un banc de test permettrait de quantifier l’erreur produite par XPower et ainsi de borner les estimations faites avec les modèles courants. Les mesures réelles pourraient alors servir de base d’apprentissage des réseaux de neurones, de fournir des estimations précises de consommation durant la création des modèles analytiques.

Durant cette thèse, certains éléments indispensables d’une chaîne de communications numériques comme les convertisseurs A/N et N/A, les étages RF (filtres, mélangeurs, etc) et les amplificateurs de puissance n’ont pas été modélisés à haut-niveau en SystemC. A l’aide d’une librairie comme SystemC-AMS (*Analog Mixed Signal*, qui s’interfacerait très bien avec les modèles SystemC déjà réalisés, des modèles de ces composants pourraient être développés puis intégrés dans les systèmes existants. Dès lors, tous les éléments d’une chaîne de communications numériques seraient modélisés, permettant ainsi des études complètes pouvant être très représentatives d’un système réel. Les études menées dans le cadre de cette thèse, portant notamment sur l’évaluation de l’efficacité énergétique et de l’efficacité spectrale, pourraient dès lors être étendues.

Les travaux, présentés dans le chapitre 4, ont été publiés dans une revue internationale [LPH16] :

- J. Lorandel, J-C. Prévotet and M. Héland, *Fast Power and Performance Evaluation of FPGA-based Wireless Communication Systems*, IEEE Access (soumise 5 avril 2016, acceptée le 20 avril 2016, publiée le 27 avril 2016))

De plus, des travaux présentés dans le chapitre 5 ont fait l'objet de présentations au cours de conférences internationales :

- J. Lorandel, J-C. Prévotet and M. Héland, *Fast Power and Energy Efficiency Analysis of FPGA-based Wireless Base-band Processing*, in International Workshop on High Performance Energy Efficient Embedded Systems (HIP3ES 2016 4th Edition), Prague, Czech Republic, 18 Jan. 2016.
- J. Lorandel, J-C. Prévotet and M. Héland, *Dynamic Power Evaluation of LTE Wireless Baseband Processing on FPGA*, Conference on Design & Architectures for Signal & Image Processing (DASIP 2015), Cracow, Poland, 23-25 Sept. 2015.
- J. Lorandel, J-C. Prévotet and M. Héland, *Dynamic Power Estimation of FPGA-based Wireless Communication Systems*, The International Conference on Wireless Technologies, embedded and intelligent Systems WITS-2015, Fes, Marocco, 29-30 Apr. 2015.
Publication dans 'Revue Méditerranéenne de Télécommunications', vol. 5, n. 2. 2015

Les travaux portant sur l'analyse des techniques à relais dans le cadre du LTE, présentés au chapitre 5, ont fait l'objet d'une publication dans une revue internationale :

- M. Maaz, J. Lorandel, P. Mary, J-C. Prévotet and M. Héland, *Energy Efficiency Analysis of Hybrid-ARQ Relay-Assisted Schemes in LTE Based Systems*, EURASIP Journal on Wireless Communications and Networking, Vol. 1, 2016 (soumise le 1 octobre 2015, acceptée le 7 janvier 2016, publiée le 19 janvier 2016)

Les travaux présentés au chapitre 6, en section 6.4, portant sur la modélisation à l'aide des réseaux de neurones ont été soumis :

- J. Lorandel, J-C. Prévotet and M. Héland, *FPGA Power Estimation using Neural Networks*, IEEE Embedded System Letter (soumis le 14 octobre 2015)

Annexe A

Annexes A

A.1 Chaîne d'émission-réception SISO-OFDM

Dans cet annexe, les estimations de la consommation dynamique de la chaîne SISO-OFDM (cf. section 5.2.1 du Chapitre 5) sont indiquées dans le Tableau A.1 pour une fréquence d'horloge de 20MHz :

TABLE A.1 – Puissance dynamique des émetteurs de la chaîne SISO-OFDM pour une fréquence d'horloge de 20MHz

	Puissance dynamique (mW)
Application 1	24.28
Application 2	32.97
Application 3	26.39
Application 4	39.66

A.2 Émetteur LTE SISO-OFDM en voie descendante du canal PDSCH

Dans cet annexe, les estimations de consommation dynamique des émetteurs de la chaîne SISO-OFDM dans le contexte LTE (cf. section 5.2.2 du Chapitre 5) sont données dans le Tableau A.2. Ces puissances dynamiques ont été obtenues pour des fréquences d'horloges multiples de 3.84MHz.

TABLE A.2 – Puissance dynamique des émetteurs LTE SISO-OFDM en voie descendante

	Fréquence d'horloge (MHz)	Puissance dynamique (mW)
Application 1	3.84	5.05
Application 2	3.84 * 2	13.33
Application 3	3.84 * 4	34.4
Application 4	3.84 * 8	77.93

Bibliographie

- [3GP10] 3GPP E-UTRA. Further advancements for E-UTRA physical layer aspects (Release 9). Technical report, 2010. 3GPP TR 36.814 v9.0.0.
- [3GP13] 3GPP E-UTRA. Relay architectures for E-UTRA (Release 9). Technical report, 2013. 3GPP TR 36.806 v9.0.0.
- [3GP14] 3GPP E-UTRA. Physical layer for relaying operation (Release 12). Technical report, 2014. 3GPP TS 36.216 v12.0.0.
- [A⁺12] G. Auer et al. Energy efficiency analysis of the reference systems, areas of improvements and target breakdown. EARTH Deliverable D2.3, 2012.
- [ACE⁺12] H. Andrade, J. Correll, A. Ekbal, A. Ghosal, D. Kim, J. Kornerup, R. Limaye, A. Prasad, K. Ravindran, T. N. Tran, et al. From streaming models to FPGA implementations. In *Proceedings of the conference for engineering of reconfigurable systems and algorithms (ERSA-IS), Las Vegas, USA*, July 2012.
- [AFJS07] N. Abdelli, A.-M. Fouilliant, N. Julien, and E. Senn. High-Level Power Estimation of FPGA. In *Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on*, pages 925–930, June 2007.
- [Alt14] Altera Inc. Powerplay early power estimator user guide, July 2014. User Guide UG-01070.
- [Alt15a] Altera Inc. Dsp builder introduction, May. 4 2015. Handbook.
- [Alt15b] Altera Inc. PowerPlay Early Power Estimator User Guide. Technical report, Jan. 2015. UG-01070.
- [AN04] J.H. Anderson and F.N. Najm. Power estimation techniques for FPGAs. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 12(10) :1015–1027, Oct 2004.
- [ARC15] ARCEP - Autorité de régulation des communications électroniques et des postes. *Observatoire des marchés des communications électroniques en France Observatoire des investissements et de de l'emploi - 3e trimestre 2014 - résultats définitifs*, 2015.
- [ARFB10] O. Arnold, F. Richter, G. Fettweis, and O. Blume. Power consumption modeling of different base station types in heterogeneous cellular networks. In *Future Network and Mobile Summit, 2010*, pages 1–8, June 2010.
- [AVB14] S.M. Afifi, F. Verdier, and C. Belleudy. Power Estimation Method Based on Real Measurements for Processor-Based Designs on FPGA. In *Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*, volume 2, pages 260–263, March 2014.

- [B⁺10] O. Blume et al. Most Promising Tracks of Green Network Technologies. Technical report, 2010. EARTH Project Deliverable D3.1.
- [BBFH⁺08] H. Breuil, D. Burette, B. Flüry-Hérard, J. Cueugnet, D. Vignolles, and H. Boisson. TIC et Développement durable, Déc. 2008. Rapport du Conseil Général de l’Environnement et du Développement Durable et le Conseil Général des Technologies de l’Information.
- [BCMP01] A. Bogliolo, I. Colonescu, E. Macii, and M. Poncino. An RTL Power Estimation Tool with On-Line Model Building Capabilities. In *Proceedings of the International Workshop on Power and Amplifier*, pages 391–396, Sept 2001.
- [Bhu15] Bhutada, S. A scalable approach for TLM across SystemC and SystemVerilog, 2015. Mentor Graphics, White Paper.
- [BJ03] R. Bergamaschi and Y.W. Jiang. State-based power analysis for systems-on-chip. In *Design Automation Conference, 2003. Proceedings*, pages 638–641, June 2003.
- [BNYT93] R. Burch, F. N. Najm, P. Yang, and T. N. Trick. A Monte Carlo approach for power estimation. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 1(1) :63–71, 1993.
- [Bou14] M.A. Bouzigues. *Green Techniques of optimizations and adaptations of the wireless radio link*. PhD thesis, INSA, RENNES, France, 2014. Order Number : 14ISAR 29 /D14-29.
- [BZ00] Ian Brynjolfson and Zeljko Zilic. Dynamic clock management for low power applications in fpgas. In *Custom Integrated Circuits Conference, 2000. CICC. Proceedings of the IEEE 2000*, pages 139–142. IEEE, 2000.
- [CA07] S. Chandrasekaran and A. Amira. A New Behavioural Power Modelling Approach for FPGA based Custom Cores. In *Adaptive Hardware and Systems, 2007. AHS 2007. Second NASA/ESA Conference on*, pages 350–357, Aug 2007.
- [Cad11] Cadence. Cadence System Development Suite, 2011.
- [Cal15] Calypto Design Systems Inc. Catapult : Product family overview, 2015. Website.
- [CCA⁺11] A. Canis, J. Choi, M. Aldham, V. Zhang, A. Kammoona, J. H. Anderson, S. Brown, and T. Czajkowski. LegUp : high-level synthesis for FPGA-based processor/accelerator systems. In *Proceedings of the 19th ACM/SIGDA international symposium on Field programmable gate arrays*, pages 33–36. ACM, 2011.
- [CHY⁺14] Y. Chen, F. Han, Y. Yang, H. Ma, Y. Han, C. Jiang, H. Lai, D. Claffey, Z. Safar, and K.J.R. Liu. Time-reversal wireless paradigm for green internet of things : An overview. *Internet of Things Journal, IEEE*, 1(1) :81–98, 2014.
- [CIS15] CISCO. Cisco Visual Networking Index : Global Mobile Data Traffic Forecast Update, 2014 - 2019, Feb. 2015. White Paper.

- [CKCW06] N. Chan King Choy and S.J.E. Wilton. Activity-based power estimation and characterization of DSP and multiplier blocks in FPGAs. In *Field Programmable Technology, 2006. FPT 2006. IEEE International Conference on*, pages 253–256, Dec 2006.
- [CLN⁺11] J. Cong, B. Liu, S. Neuendorffer, J. Noguera, K. Vissers, and Z. Zhang. High-level synthesis for FPGAs : From prototyping to deployment. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 30(4) :473–491, 2011.
- [CRR05] J. Coburn, S. Ravi, and A. Raghunathan. Power emulation : a new paradigm for power estimation. In *Proceedings of the 42nd annual Design Automation Conference*, pages 700–705. ACM, 2005.
- [CSC99] R.C. Chu, R.E. Simons, and G.M. Chrysler. Experimental investigation of an enhanced thermosyphon heat loop for cooling of a high performance electronics module. In *Semiconductor Thermal Measurement and Management Symposium, 1999. Fifteenth Annual IEEE*, pages 1–9, March 1999.
- [CTL⁺05] C.T. Chow, L.S.M. Tsui, P.H.W. Leong, W. Luk, and S.J.E. Wilton. Dynamic voltage scaling for commercial fpgas. In *Field-Programmable Technology, 2005. Proceedings. 2005 IEEE International Conference on*, pages 173–180, Dec 2005.
- [DA14] Y. Durrani and T. Alcaide. High-Level Power Analysis for Intellectual Property-Based Digital Systems. *Circuits, Systems, and Signal Processing*, 33(4) :1035–1051, 2014.
- [DB02] H. Demuth and M. Beale. Neural Network Toolbox for use with Matlab, July 2002. User’s Guide version 4.
- [DDG⁺12] C. Desset, B. Debaillie, V. Giannini, A. Fehske, G. Auer, H. Holtkamp, W. Wajda, D. Sabella, F. Richter, M.J. Gonzalez, H. Klessig, I. Godor, M. Olsson, M.A. Imran, A. Ambrosy, and O. Blume. Flexible power modeling of LTE base stations. In *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pages 2858–2862, April 2012.
- [DDVJM12] M. Deruyck, D. De Vulder, W. Joseph, and L. Martens. Modelling the power consumption in femtocell networks. In *Wireless Communications and Networking Conference Workshops (WCNCW), 2012 IEEE*, pages 30–35, April 2012.
- [Dep15] UC Berkeley EECS Department. The Ptolemy Project, 2015. Website.
- [DN07] B.S. Deepaksubramanyan and A. Nunez. Analysis of subthreshold leakage reduction in CMOS digital circuits. In *Circuits and Systems, 2007. MWS-CAS 2007. 50th Midwest Symposium on*, pages 1400–1404, Aug 2007.
- [DR⁺06] Y. Durrani, T. Riesgo, et al. Power macromodeling for IP modules. In *Electronics, Circuits and Systems, 2006. ICECS’06. 13th IEEE International Conference on*, pages 1172–1175. IEEE, 2006.
- [DS07] R. Damasevicius and V. Stukys. Estimation of Power Consumption at Behavioral Modeling Level Using SystemC. In *EURASIP Journal on Embedded Systems, 2007*. Hindawi Publishing Corporation, May 2007.

- [DSZC11] L. Deng, K. Sobti, Y. Zhang, and C. Chakrabarti. Accurate area, time and power models for FPGA-based implementations. *Journal of Signal Processing Systems*, 63(1) :39–50, 2011.
- [DT05] V. Degalahal and T. Tuan. Methodology for high level estimation of FPGA power consumption. In *Design Automation Conference, 2005. Proceedings of the ASP-DAC 2005. Asia and South Pacific*, volume 1, pages 657–660 Vol. 1, Jan 2005.
- [Dub13] T. Dubois. *Application du Retourneement Temporel aux systèmes multipor-teuses : Propriétés et performances*. PhD thesis, INSA de RENNES-IETR, 2013.
- [EAR12] EARTH. Energy Aware Radio and neTwork tecHnologies project, 2012. Website.
- [EJH⁺04] D. Elleouet, N. Julien, D. Houzet, J.-G. Cousin, and E. Martin. Power consumption characterization and modeling of embedded memories in Xi-linx. In *Digital System Design, 2004. DSD 2004. Euromicro Symposium on*, pages 394–401, Aug 2004.
- [EJH06] D. Elleouet, N. Julien, and D. Houzet. A high level soc power estimation based on ip modeling. In *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, pages 4 pp.–, April 2006.
- [ETS09] ETSI. Evolved Universal Terrestrial Radio Access (E-UTRA)-Physical chan-nels and modulation, Jun. 2009. 3GPP TS 36.211 v8.7.0 Release 8.
- [ETS10] ETSI. Technical Specification Group Radio Access Network ; Evolved Uni-versal Terrestrial Radio Access (E-UTRA) ; Radio Frequency (RF) require-ments for LTE Pico Node B . Technical report, 2010. 3GPP TR 36.931 version 9.0.0 (Release 9).
- [fS00] International Technology Roadmap for Semiconductors. Design edition re-port 2011, 2000.
- [GBP10] M. Grant, B. Bailey, and A. Piziali. *ESL design and verification : a pres-cription for electronic system level methodology*. Morgan Kaufmann, 2010.
- [GPH⁺09] S. V. Gheorghita, M. Palkovic, J. Hamers, A. Vandecappelle, S. Mamagka-kis, T. Basten, L. Eeckhout, H. Corporaal, F. Catthoor, F. Vandeputte, et al. System-scenario-based design of dynamic embedded systems. *ACM Tran-sactions on Design Automation of Electronic Systems (TODAES)*, 14(1) :3, 2009.
- [Gra07] D. Graupe. *Principles of artificial neural networks*, volume 6. World Scien-tific, 2007.
- [Gre14] GreenTouch, 2014. GreenTouch Website.
- [GS09] M. Grant and G. Smith. High-level synthesis : Past, present, and future. *IEEE Design & Test of Computers*, (4) :18–25, 2009.
- [GS14] D. Garg and R. Sharma. Low Power Multiplier using Dynamic Voltage and Frequency Scaling (DVFS). In *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*, pages 560–564, Sept 2014.

- [HAE09] H.A. Hassan, M. Anis, and M. Elmasry. Total Power Modeling in FPGAs Under Spatial Correlation. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 17(4) :578–582, April 2009.
- [HASLVE08] I. Herrera-Alzu, M.A. Sanchez, M. Lopez-Vallejo, and P. Echeverria. Experimental methodology for power characterization of FPGAs. In *Electronics, Circuits and Systems, 2008. ICECS 2008. 15th IEEE International Conference on*, pages 582–585, Aug 2008.
- [HBB14] S. Hesselbarth, T. Baumgart, and H. Blume. Hardware-assisted power estimation for design-stage processors using FPGA emulation. In *Power and Timing Modeling, Optimization and Simulation (PATMOS), 2014 24th International Workshop on*, pages 1–8, Sept 2014.
- [HKH15] J. Hussein, M. Klain, and M. Hart. Lowering Power at 28 nm with Xilinx 7 Series Devices. *Xilinx White Paper WP389 (v1.3)*, Jan. 2015.
- [HLF⁺11] C. Hsu, J. Liao, S. Fang, C. Weng, S. Huang, Hsieh W., and J. Yeh. PowerDepot : Integrating IP-based power modeling with ESL power analysis for multi-core SoC designs. In *Design Automation Conference (DAC), 2011 48th ACM/EDAC/IEEE*, pages 47–52, June 2011.
- [HIXR15] S. Han, C. I. I, Z. Xu, and C. Rowell. Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Communications Magazine*, 53(1) :186–194, January 2015.
- [HNY14] M. Hosseinabady and J.L. Nunez-Yanez. Run-time power gating in hybrid ARM-FPGA devices. In *Field Programmable Logic and Applications (FPL), 2014 24th International Conference on*, pages 1–6. IEEE, 2014.
- [HPCK15] M. H elard, J.C. Pr evotet, M. Crussi ere, and Y. Kokar. Application of Time Reversal Principles to Wireless Communications : From Single-User SISO to Multiple User Large Scale MIMO Scenarios. Tutorial given at the Softcom International Conference on Software, Telecommunications and Computer Networks, INSA of Rennes and IRT B-Com, 2015. Bol, Croatia.
- [HUA13] HUAWEI. 5G : A Technology Vision, White Paper, 2013.
- [HZDS95] C.X. Huang, B. Zhang, A. Deng, and B. Swirski. The design and implementation of PowerMill. In *Proceedings of the 1995 international symposium on Low power design*, pages 105–110. ACM, 1995.
- [I⁺12] M. A. Imran et al. Most suitable efficiency metrics and utility functions. Technical report, 2012. EARTH Project Deliverable D2.4.
- [IEE12] IEEE Standards Association. IEEE Standard for Air Interface for Broadband Wireless Access Systems, Aug. 2012. IEEE Std 802.16-2012.
- [IHK11] S. Ishihara, M. Hariyama, and M. Kameyama. A low-power FPGA based on autonomous fine-grain power gating. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 19(8) :1394–1406, 2011.
- [Inc00] Synopsys Inc. PowerMill User Guide. 2000. Release 5.4.
- [Ini15] Accellera Systems Initiative. Systemc, 2015. Accellera Systems Initiative Website.

- [JP04] O. Jingzhao and V.K. Prasanna. PyGen : a MATLAB/Simulink based tool for synthesizing parameterized and energy efficient designs using FPGAs. In *Field-Programmable Custom Computing Machines, 2004. FCCM 2004. 12th Annual IEEE Symposium on*, pages 47–56, April 2004.
- [KA06] A. Kumar and M. Anis. Dual-Vt design of FPGAs for subthreshold leakage tolerance. In *Quality Electronic Design, 2006. ISQED '06. 7th International Symposium on*, pages 6 pp.–740, March 2006.
- [Key15] Keysight-Technologies. SystemVue Electronic System-Level (ESL) Design Software, 2015.
- [KL03] S. Kang and Y. Leblebici. *CMOS Digital Integrated Circuits Analysis & Design*. McGraw-Hill, Inc., New York, NY, USA, 3 edition, 2003.
- [KN97] J.N. Kozhaya and F.N. Najm. Accurate power estimation for large sequential circuits. In *Proceedings of the 1997 IEEE/ACM international conference on Computer-aided design*, pages 488–493. IEEE Computer Society, 1997.
- [KR07] I. Kuon and J. Rose. Measuring the Gap Between FPGAs and ASICs. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 26(2) :203–215, Feb 2007.
- [KTR08] I. Kuon, R. Tessier, and J. Rose. Fpga architecture : Survey and challenges. *Foundations and Trends in Electronic Design Automation*, 2(2) :135–253, 2008.
- [LDSY07] Y. Liu, R.P. Dick, L. Shang, and H. Yang. Accurate Temperature-Dependent Integrated Circuit Leakage Power Estimation is Easy. In *Design, Automation Test in Europe Conference Exhibition, 2007. DATE '07*, pages 1–6, April 2007.
- [LJSM04] J. Laurent, N. Julien, E. Senn, and E. Martin. Functional level power analysis : an efficient approach for modeling the power consumption of complex processors. In *Design, Automation and Test in Europe Conference and Exhibition, 2004. Proceedings*, volume 1, pages 666–667 Vol.1, Feb 2004.
- [LL08] J. Lamoureux and W. Luk. An overview of low-power techniques for field-programmable gate arrays. In *Adaptive Hardware and Systems, 2008. AHS'08. NASA/ESA Conference on*, pages 338–345. IEEE, 2008.
- [LLW08] J. Lamoureux, G.G.F. Lemieux, and S.J.E. Wilton. GlitchLess : Dynamic Power Minimization in FPGAs Through Edge Alignment and Glitch Filtering. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 16(11), Nov. 2008.
- [LP05] X. Liu and M.C. Papaefthymiou. HyPE : Hybrid power estimation for IP-based systems-on-chip. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 24(7) :1089–1103, 2005.
- [LPF10] S. Liu, R. N. Pittman, and A. Forin. Energy reduction with run-time partial reconfiguration. In *FPGA*, page 292, 2010.
- [LPH16] J. Lorandel, J. C. Prevotet, and M. Helard. Fast power and performance evaluation of fpga-based wireless communication systems. *IEEE Access*, PP(99) :1–1, 2016.

- [LW06a] J. Lamoureux and S.J.E. Wilton. Activity estimation for field-programmable gate arrays. In *Field Programmable Logic and Applications, 2006. FPL'06. International Conference on*, pages 1–8. IEEE, 2006.
- [LW06b] J. Lamoureux and S.J.E. Wilton. Activity Estimation for Field-Programmable Gate Arrays. In *Field Programmable Logic and Applications, 2006. FPL '06. International Conference on*, pages 1–8, Aug 2006.
- [LXX⁺11] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu. Energy-efficient wireless communications : tutorial, survey, and open issues. *IEEE Wireless Communications*, 18(6) :28–35, December 2011.
- [Mat15] MathWorks, 2015. Website.
- [Men15] Mentor Graphics. ModelSim, 2015.
- [MGP11] D. Meidanis, K. Georgopoulos, and I. Papaefstathiou. FPGA power consumption measurements and estimations under different implementation parameters. In *Field-Programmable Technology (FPT), 2011 International Conference on*, pages 1–6, Dec 2011.
- [MLM⁺16] M. Maaz, J. Lorandel, P. Mary, J.-C. Prévotet, and M. Hêlard. Energy efficiency analysis of hybrid-ARQ relay-assisted schemes in LTE-based systems. *EURASIP Journal on Wireless Communications and Networking*, 2016(1) :1–13, 2016.
- [MMP94] R. Marculescu, D. Marculescu, and M. Pedram. Logic level power estimation considering spatiotemporal correlations. In *In Proceedings of the IEEE International Conference on Computer Aided Design*, pages 294–299, 1994.
- [Mon07] M. Montoreano. Transaction level modeling using osci tlm 2.0. *Open SystemC Initiative (OSCI)*, 2007.
- [Mor14] G. Moretti. System Level Power Budgeting. *Chip Design Mag.*, mar. 2014.
- [Nag75] L. W. Nagel. *SPICE2 : A Computer Program to Simulate Semiconductor Circuits*. PhD thesis, EECS Department, University of California, Berkeley, 1975.
- [Naj94] F.N. Najm. A survey of power estimation techniques in VLSI circuits. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 2(4) :446–455, Dec 1994.
- [Nat15] National Instruments. LabVIEW System Design Software, 2015.
- [NMH12] C. Najoua, B. Mohamed, and B.M. Hedi. Analytical dynamic power model for LUT based components. In *Design & Technology of Integrated Systems in Nanoscale Era (DTIS), 2012 7th International Conference on*, pages 1–6. IEEE, 2012.
- [NSN⁺05] M. Najibi, K. Saleh, M. Naderi, H. Pedram, and M. Sedighi. Prototyping globally asynchronous locally synchronous circuits on commercial synchronous FPGAs. In *Rapid System Prototyping, 2005. (RSP 2005). The 16th IEEE International Workshop on*, pages 63–69, June 2005.
- [NX98] F. N. Najm and M. G. Xakellis. Statistical Estimation of the, Switching Activity in VLSI Circuits. *VLSI Design*, 7(3) :243–254, 1998.

- [OCF⁺13] M. Olsson, C. Cavdar, P. Frenger, S. Tombaz, D. Sabella, and R. Jantti. 5GrEEEn : Towards Green 5G mobile networks. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2013 IEEE 9th International Conference on*, pages 212–216, Oct 2013.
- [Par04] M. Parizeau. Réseaux de neurones, 2004.
- [PB04] K. Parnell and R. Bryner. Comparing and contrasting FPGA and micro-processor system design and development. *White Paper, Xilinx Corporation*, 2004.
- [RBAN⁺11] S.K. Rethinagiri, R. Ben Atitallah, S. Niar, E. Senn, and J. Dekeyser. Hybrid system level power consumption estimation for FPGA-based MPSoC. In *Computer Design (ICCD), 2011 IEEE 29th International Conference on*, pages 239–246, Oct 2011.
- [RMMM03] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits. *Proceedings of the IEEE*, 91(2) :305–327, Feb 2003.
- [RSN06] A. Reimer, A. Schulz, and W. Nebel. Modelling macromodules for high-level dynamic power estimation of FPGA-based digital designs. In *Proceedings of the 2006 international symposium on Low power electronics and design*, pages 151–154. ACM, 2006.
- [RW05] N. Rollins and M. J. Wirthlin. Reducing energy in FPGA multipliers through glitch reduction. In *International Conference on Military and Aerospace Programmable Logic Devices*, Sept. 2005.
- [SCZ⁺12] E. Senn, D. Chillet, O. Zendra, C. Belleudy, S.B. Bilavarn, R.B. Atitallah, C. Samoyeau, and A. Fritsch. Open-People : Open Power and Energy Optimization PLatform and Estimator. In *Digital System Design (DSD), 2012 15th Euromicro Conference on*, pages 668–675, Sept 2012.
- [SH89] A. Salz and M. Horowitz. IRSIM : An Incremental MOS Switch-Level Simulator. In *Design Automation, 1989. 26th Conference on*, pages 173–178, June 1989.
- [SJ01] L. Shand and N. K Jha. High-level power modeling of CPLDs and FPGAs. In *Computer Design, 2001. ICCD 2001. Proceedings. 2001 International Conference on*, pages 46–51. IEEE, 2001.
- [SJK⁺11] P. Schumacher, P. Jha, S. Kuntur, T. Burke, and A. Frost. Fast RTL Power Estimation for FPGA Designs. In *Field Programmable Logic and Applications (FPL), 2011 International Conference on*, pages 343–348, Sept 2011.
- [SKB02] Li S., A. S. Kaviani, and K. Bathala. Dynamic power consumption in Virtex-II FPGA family. In *Proceedings of the 2002 ACM/SIGDA tenth international symposium on Field-programmable gate arrays*, pages 157–164. ACM, 2002.
- [SKN⁺01] A. Stammermann, L. Kruse, W. Nebel, A. Pratsch, E. Schmidt, M. Schulte, and A. Schulz. System level optimization and design space exploration for low power. In *System Synthesis, 2001. Proceedings. The 14th International Symposium on*, pages 142–146, 2001.

- [SRH⁺11] M. Streubuhr, R. Rosales, R. Hasholzner, C. Haubelt, and J. Teich. ESL power and performance estimation for heterogeneous MPSoCs using SystemC. In *Specification and Design Languages (FDL), 2011 Forum on*, pages 1–8, Sept 2011.
- [Str14] Strategy-Analytics. 33 Billion Internet Devices By 2020 : Four Connected Devices For Every Person In World, Oct 2014. Press Release.
- [SWPC10] D. Sunwoo, G.Y. Wu, N. A. Patil, and D. Chiou. PrEsto : An FPGA-accelerated power estimation methodology for complex systems. In *Field Programmable Logic and Applications (FPL), 2010 International Conference on*, pages 310–317. IEEE, 2010.
- [Syn15] Synopsys. System studio, 2015. Website.
- [TBAV05] E. Todorovich, E. Boemo, F. Angarita, and J. Vails. Statistical power estimation for FPGAs. In *Field Programmable Logic and Applications, 2005. International Conference on*, pages 515–518. IEEE, Aug 2005.
- [TOR⁺14] O. Turkyilmaz, S. Onkaraiah, M. Reyboz, F. Clermidy, C. Anghel, J.M. Portal, M. Bocquet, et al. RRAM-based FPGA for "Normally Off, Instantly On" applications. *Journal of Parallel and Distributed Computing*, 74(6) :2441–2451, 2014.
- [VSP10] P. Van Stralen and A. Pimentel. Scenario-based design space exploration of MPSoCs. In *Computer Design (ICCD), 2010 IEEE International Conference on*, pages 305–312. IEEE, 2010.
- [WWH⁺11] B. Wang, Y. Wu, F. Han, Y.H. Yang, and K.J. Liu. Green wireless communications : A time-reversal paradigm. *Selected Areas in Communications, IEEE Journal on*, 29(8) :1698–1710, 2011.
- [Xil11a] Xilinx Inc. AXI Reference Guide. Technical report, March 2011. UG761 (v13.1).
- [Xil11b] Xilinx Inc. LogiCORE IP 3GPP LTE Turbo Encoder v3.1. Technical report, 2011. DS701 Product Specification.
- [Xil11c] Xilinx Inc. LogiCORE IP Fast Fourier Transform v7.1. Technical report, March 2011. DS260 Product Specification.
- [Xil12a] Xilinx Inc. System generator for dsp, Oct. 16 2012. User Guide UG640 (v 14.3).
- [Xil12b] Xilinx Inc. Xpower estimator user guide, Jan 2012. User Guide UG440 (v13.4).
- [Xil15] Xilinx Inc. ISE Design Suite, 2015. Xilinx Website.
- [YB12] F.X. Wolff Y. Bouguen, E. Hardouin. *LTE et les réseaux 4G*. Ed. Eyrolles, 2012.
- [ZPR⁺13] N. Zompakis, A. Papanikolaou, P. Raghavan, D. Soudris, and F. Catthoor. Enabling efficient system configurations for dynamic wireless applications using system scenarios. *International journal of wireless information networks*, 20(2) :140–156, 2013.

AVIS DU JURY SUR LA REPRODUCTION DE LA THESE SOUTENUE

Titre de la thèse:

Etude de la consommation énergétique de systèmes de communications numériques sans fils implantés sur cible FPGA

Nom Prénom de l'auteur : LORANDEL JORDANE

Membres du jury :

- Monsieur BERDER Olivier
- Madame BELLEUDY Cécile
- Madame HELARD Maryline
- Monsieur PREVOTET Jean-Christophe
- Monsieur HOUZET Dominique
- Monsieur GRANADO Bertrand
- Monsieur BOUTILLON Emmanuel

Président du jury : *BERDER Olivier*

Date de la soutenance : 08 Décembre 2015

Reproduction de la these soutenue

Thèse pouvant être reproduite en l'état

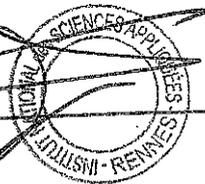
~~Thèse pouvant être reproduite après corrections suggérées.~~

Fait à Rennes, le 08 Décembre 2015

Signature du président de jury

Le Directeur,

M'hamed DRISSI



A handwritten signature in black ink, consisting of several loops and a final flourish.

Résumé

Les systèmes de communications sans fil n'ont cessé d'évoluer ces dernières années, poussés par de fortes demandes du marché en systèmes toujours plus autonomes et performants. Ainsi, de nouvelles contraintes de conception sont apparues de manière à mieux prendre en compte les aspects énergétiques et ainsi améliorer la durée de vie des batteries et des circuits. Actuellement, les systèmes de communications numériques sans fil consomment d'importantes quantités d'énergie. D'autre part, la complexité des systèmes croît de génération en génération afin de satisfaire toujours plus d'utilisateurs avec un haut niveau de performances.

Dans ce contexte à fortes contraintes, les circuits de type FPGA apparaissent comme une technologie attractive, pouvant supporter des circuits numériques complexes grâce à leur grand nombre de ressources. Pour pouvoir concevoir les futurs systèmes de communications numériques sans fil sur ce type de circuit, les concepteurs de tels systèmes doivent pouvoir estimer la consommation et les performances au plus tôt dans la phase de conception. De cette façon, ils pourront explorer l'espace de conception et effectuer des choix d'implémentation afin d'optimiser leurs systèmes.

Durant cette thèse, une méthodologie a été proposée dont les objectifs sont d'estimer rapidement et à haut-niveau la consommation de leurs circuits implantés sur FPGA ainsi que leurs performances, d'explorer l'espace de conception, de comparer efficacement plusieurs systèmes entre eux, tout en assurant une bonne précision de l'estimation.

La méthodologie repose sur une phase de caractérisation de composants IP matériels ainsi que de leur modélisation en SystemC. Dans un second temps, une représentation haut-niveau du système entier est réalisée à partir de la librairie des modèles SystemC de chaque IP. A travers des simulations haut-niveau, les utilisateurs peuvent tester rapidement de multiples configurations de leur système. Un des caractères innovants de l'approche repose sur l'utilisation de signaux clés qui permettent de tenir compte des comportements dynamiques des composants IP, c-à-d. leur temps d'activité (actif/inactif), au sein du système et ainsi obtenir des estimations précises.

Les nombreux gains de la méthodologie ont été démontrés à travers plusieurs exemples de systèmes de communications numériques sans fil comme une chaîne de traitement en bande de base de type SISO-OFDM générique, des émetteurs LTE etc.

Pour conclure, les limitations ont été adressées et des solutions d'optimisation ont pu être envisagées puis mises en place.

Abstract

Wireless communication systems are still evolving since the last decades, driven by the growing demand of the electronic market for energy efficient and high performance devices. Thereby, new design constraints have appeared that aim at taking into account power consumption in order to improve battery-life of circuits.

Current wireless communication systems commonly dissipate a lot of power. On the other hand, the complexity of such systems keeps on increasing through the generations to always satisfy more users at a high degree of performance.

In this highly constrained context, FPGA devices seem to be an attractive technology, able to support complex systems thanks to their important number of resources.

According to the FPGA nature, designers need to estimate the power consumption and the performance of their wireless communication systems as soon as possible in the design flow. In this way, they will be able to perform efficient design space exploration and make decisive implementation and optimization choices.

Throughout this thesis, a power estimation methodology for hardware-focused FPGA device is described and aims at making design space exploration a lot easier, providing early and fast power and performance estimation at high-level. It also proposes an efficient way to efficiently compare several systems.

The methodology is effective through an IP characterisation step and the development of their SystemC models. Then, a high-level description of the entire system is realized from the SystemC models that have been previously developed. High-level simulations enable to check the functionality and evaluate the power and performance of the system.

One of the contributions consists in monitoring the IP time-activities (active/idle) during the simulation. This enables to take into account dynamic effects related to each application. We show that this has an important impact on both power and performance.

The effectiveness of the methodology has been demonstrated throughout several baseband processing chains of the wireless communication domain such as a SISO-OFDM generic chain, LTE transmitters etc.

To conclude, the main limitations of the proposed methodology have been investigated and addressed.

Utiliser la police Arial Taille 9 en bleu dans les champs texte « résumé » et « abstract » - Texte justifié -
Ne pas dépasser le nombre de caractères des cadres de texte ci-dessus.
Ne pas modifier la taille des cadres de texte